Principles and Recommendations for Population and Housing Censuses

Revision 4

United Nations

PART ONE. ESSENTIAL FEATURES AND CENSUS METHODOLOGY

I. Essential roles of the census

1.1. Evidence-based decision-making is a universally recognized paradigm of efficient management of economic and social affairs and of overall effective governing of societies today. Making use of relevant, accurate and timely statistics is essential to this model; producing detailed statistics for small areas and small population groups is its foundation. The role of the population and housing census is to produce and disseminate such detailed small-area statistics on population, its composition, characteristics, spatial distribution and organization into families and households. Censuses have been promoted internationally since the end of the nineteenth century, when the International Statistical Congress recommended that all countries in the world conduct them.[1] Since 1958, the United Nations has also been actively promoting the population and housing census by compiling the principles and recommendations for population and housing censuses and launching regular decennial worldwide programmes on population and housing censuses. In fact, the majority of the countries in the world conduct a population and housing census periodically. In order to increase the relevance of the data produced by censuses, several countries are now moving towards using alternative census methodologies to produce census-like statistics with a more frequent periodicity than the traditional decennial or quinquennial one.

1.2. While the roles of the population and housing census are many and will be elaborated in detail throughout the present revision of *Principles and Recommendations for Population and Housing Censuses*, several of the essential roles are listed below:

(a) **The population and housing census plays an essential role in public administration**. The results of a census are used as a critical reference to ensure equity in distribution of wealth, government services and representation nationwide by informing the distribution and allocation of government funds among various regions and districts for education, health services, delineating electoral districts at the national and local levels, and measuring the impact of industrial development, to name a few. Establishing a public consensus on priorities would be almost impossible to achieve if it were not built on census counts. A wide range of other users, including the corporate sector, academia, civil society and individuals, make use of census outputs.

(b) **The census also plays an essential role in all elements of the national statistical system**, including the economic and social components. Census statistics are used as benchmarks for statistical compilation or as a sampling frame for sample surveys. Today, the national statistical system of almost every country relies on sample surveys for efficient and reliable data collection, notwithstanding the increasing availability of non-traditional data sources. Without the sampling frame and population benchmarks derived from the population and housing census, the national statistical system would face difficulties in providing reliable official statistics for use by the government and the general

[1] *Report of the Proceedings of the Fourth Session of the International Statistics Congress, held in London July 16th, 1860, and the Five following Days*, printed by George Edward Eyre and William Spottiswoode, London, 1861.

public. For many countries with less developed statistical systems, a census is also an essential source providing data to integrate with survey data for producing model-based small-area estimations.

(c) **A basic function of a census is to generate statistics on small areas and small population groups**. While statistics on small areas are useful per se, they are important because they can be used to produce statistics on any geographic unit with arbitrary boundaries. For example, in planning the location of a school, it is necessary to have the data on the distribution of school-age children by school area, which may not necessarily correspond to the administrative area units. Similarly, small-area data from the census can be combined to approximate natural regions (for example, water catchments or vegetation zones) that do not follow administrative boundaries. Since census data can be tabulated for any geographic unit, it is possible to provide the required statistics in a remarkably flexible manner. This versatile feature of the census is also invaluable for use in the private sector for applications such as business planning and market analyses.

(d) **Census data are used as a benchmark for research and analysis**. They are used to study population dynamics. Population projections are one of the most important analytical outputs based on census data; future population projections are crucial for all segments of the public and private sectors.

1.3. It is critically important to produce detailed statistics for small areas and small population groups **as a building block for efficient governance at all levels**. For many countries the method for assembling this building block will be by conducting a population and housing census through universal and simultaneous individual enumeration of each set of living quarters and of each individual within the country's boundaries. Some countries will adopt alternative approaches; yet, all of these methods should produce detailed, timely statistics for small areas and small population groups at a well-defined point or period in time.

II. Definitions and key features

A. Definitions

1. Population census

1.4. A population census is the total process of planning, collecting, compiling, evaluating, disseminating and analysing demographic, economic and social data at the smallest geographic level pertaining, at a specified time, to all persons in a country or in a well-delimited part of a country.

1.5. Population is basic to the production and distribution of material wealth. In order to plan for, and implement, economic and social development, administrative activity or scientific research, it is necessary to have reliable and detailed data on the size, distribution and composition of population. The population census is a primary source of these basic benchmark statistics, covering not only the settled population but also homeless persons and nomadic groups. Data from population censuses should allow presentation and analysis of statistics on persons and households and for a wide variety of geographic units, ranging from the country as a whole to individual small localities or city blocks.

2. Housing census

1.6. A housing census is the total process of planning, collecting, compiling, evaluating, disseminating and analysing statistical data relating to the number and condition of housing units and facilities as available to the households pertaining, at a specified time, to all living quarters[2] and occupants thereof in a country or in a well-delimited part of a country.

1.7. The census must provide information on the supply of housing units together with information on the structural characteristics and facilities that have a bearing upon the maintenance of privacy and health and living conditions. Sufficient demographic, social and economic data concerning the occupants must be collected to furnish a description of housing conditions and also to provide basic data for analysing the causes of housing deficiencies and for studying possibilities for remedial action. In this connection, data obtained as part of the population census, including data on homeless persons,[3] are often used in the presentation and analysis of the results of the housing census, if both operations are conducted together or there is a link between them.

B. Key features

1.8. The key features of population and housing censuses are individual enumeration, universality within a defined territory, simultaneity, defined periodicity and small-area statistics.

1. Individual enumeration

1.9. The term "census" implies that each individual and each set of living quarters is enumerated separately and that the characteristics thereof are separately recorded. Only by this procedure can the data on the various characteristics be cross-classified. The requirement of individual enumeration can be met by the collection of information in the field, by the use of information contained in an appropriate administrative register or set of registers, or by a combination of these methods.

2. Universality within a defined territory

1.10. The census should cover a precisely defined territory (for example, the entire country or a well-delimited part of it). The population census should include every person present and/or residing within its scope, depending upon the type of population count required, and should include them only once. The housing census should include every set of living quarters irrespective of type, and should include each one only once. This does not preclude the use of sampling techniques for obtaining data on specified characteristics, provided that the sample design is consistent with the size of the areas for which the data are to be tabulated and the degree of detail in the cross-tabulations to be made.

3. Simultaneity

[2] For the definition of "living quarters", see paragraph 4.421.

[3] For the definition of “homeless persons”, see paragraph 2.37.

1.11. Each person and each set of living quarters should be enumerated as of the same well-defined point in time and the data collected should refer to a well-defined reference period. The time reference period need not, however, be identical for all of the data collected. For most of the data, it will be the day of the census; in some instances, it may be a period prior to the census.[4] Not all census methodologies satisfy this criterion. [5]

Comments FWN: This part needs to start with a sentence explaining why we care about either simultaneity or at least approximately comparable reference periods. That way we can account for the variation among census methodologies, saying ok they might not all refer to exactly the same point but the need for them to refer to a reasonably similar point in time remains irrespective of methodology.

4. Defined periodicity

1.12. Censuses should, ideally, be taken at regular intervals so that comparable information is made available in a fixed sequence. A series of censuses makes it possible to appraise the past, accurately describe the present and estimate the future. It is recommended that a national census be taken at least every 10 years. Some countries may find it necessary to carry out censuses more frequently because of the rapidity of major changes in their population and/or its housing circumstances.

1.13. The census data of any country are of greater value nationally, regionally and internationally if they can be compared with the results of censuses of other countries that were taken at approximately the same time. Therefore, countries should make all efforts to undertake a census in years ending in "0" or at a time as near to those years as possible[6]. It is obvious, however, that legal, administrative, financial and other considerations often make it inadvisable for a country to adhere to a standard international pattern in the timing of its censuses. Some regional groupings of countries, such as the European Union, have their own requirements regulating the dates of censuses among their members. In fixing a census date, therefore, such national and regional factors could be given greater weight than the desirability of international simultaneity.

5. Capacity to produce small-area statistics

1.14. The census should produce data on the number and characteristics of the population and housing units down to the lowest appropriate geographic level, compatible with national circumstances, and for small population groups, awhile protecting confidentiality of personal information on each individual. Not all census methodologies satisfy this criterion.[7]

[4] For example, collecting information on the core topic of household deaths in the past 12 months (see paragraph 4.250).

[5] Simultaneity is an essential feature of censuses that rely on full field enumeration (traditional approach) (see paragraphs 1.69–...) as well as of some types of register-based censuses (see paragraph 1.91). However, for censuses that employ other methodologies such as ..., this feature may not apply (see paragraphs ...–…).

[6] For EU countries it is mandatory to have censuses in years ending in "1".

[7] The capacity to produce data for small areas or for small population groups is an essential feature of censuses that rely on full field enumeration (traditional approach) or that are fully register-based. For other types of census methodologies, the capacity to produce small-area statistics data may depend on the size of the sample survey (see paragraphs 1.127, 1.130, ..), the ability to integrate sources (see paragraphs 1.29) or the extent to which samples from different surveys can be combined (see paragraphs 1.79, 1.108, ...).

III. Uses of population and housing censuses

1.15. Population and housing censuses are a principal means of collecting basic population and housing statistics. They form the core of an integrated programme of data collection and compilation aimed at providing a comprehensive source of statistical information for economic and social development planning, administration, assessing conditions in human settlements, research and commercial and other uses. Population and housing censuses are often a basis for sample surveys and provide insights into buildings, household and population characteristics as well as increasingly being used as a source of environmental statistics.

1.16. The value of population and housing censuses is increased if the results can be employed together with other data sources, as in the use of the census data as a basis or benchmark for current statistics, and if it can furnish the information needed for conducting other statistical data collection. It can, for example, provide a statistical frame for other censuses or sample surveys. The population census is also important in developing the population estimates needed to calculate vital rates in combination with civil registration data (see paragraphs 1.57–1.59). In addition, these censuses are a major source of data used in official compilations of social indicators, particularly on topics that usually change slowly over time. It is important, therefore, to take into account the requirements of a continuous, coordinated programme of data collection and compilation when planning a census. When the relationships between the population census, the housing census and other statistical investigations are considered from the outset, their shared purposes can best be served and their benefits best harnessed. The use of consistent concepts and definitions throughout an integrated programme of data collection and compilation is essential if the advantages of these relationships are to be fully realized. Of course, census-type information can also be derived from population registers and can also be estimated from sample surveys without undertaking a complete enumeration. These alternative data sources are presented under "Census methodology" in paragraphs 1.63–1.119.

1.17. Indicators are required by countries to track the progress towards the Sustainable Developmental Goals and as such efforts must be made by census offices to produce relevant indicators to meet this need.

1.18. Increasingly, Censuses of Population and Housing are a key source of data in the compilation of environment statistics. They provide environmental agencies with a range of information to analyze energy consumption, identify conservation opportunities, and forecast energy needs. Census data help national and local government and relief agencies in planning relief operations to assist populations and areas affected by natural disasters such as floods, hurricanes, tornadoes, and earthquakes. It also provides a rich source of data for public health authorities in identifying prevalence levels of disease, responses to pandemics and measurement of outcomes of public health initiatives.

1.19. A population and housing census also serves as the logical starting point for work on the organization and construction of computerized statistical products to serve continuing national and local needs for data in the intercensal period.[8]

[8] See Chapter X, Part Three of this publication.

1.20. In addition to the statistical value obtained directly from the census results themselves, there are further, indirect benefits from taking a census, particularly to the organization responsible for the census, or the national statistical office. These benefits include:

(i) **Improved skills and experience:** varied sets of skills are often required for administering a census that are not necessarily prominent in other parts of the organization, such as project management, procurement and commercial, communication, human resources, and information technology (IT) skills.

(ii) **Technological advancement:** often a census requires new technology to support complex data collection and processing requirements. These developments may be reused for other exercises within the national statistical office or lead to new technological developments.

(iii) **New methods:** the development of methods for enumerating the whole population, or statistical methods (such as edit and imputation) developed for processing census results, can often be reused for other statistical exercises within the national statistical office such as stratification of the population, estimation and modelling population characteristics, or the development of indicators.

(iv) **Halo effect:** the extensive promotion of the census may also have a positive effect, the "halo effect", on other surveys, resulting in increases in response rates.[9]

A. Uses of population censuses

1. Uses for policymaking, planning and administrative purposes

1.21. The fundamental purpose of the population census is to provide the facts essential to national policymaking, planning and administration. Information on the size, distribution and characteristics of a country's population is essential for describing and assessing its economic, social and demographic circumstances and for developing sound policies and programmes aimed at fostering the welfare of a country and its population. The population census, by providing comparable basic statistics for a country as a whole and for each administrative unit, locality and small area therein, can make an important contribution to the overall planning process and the management of national affairs. Counts of the population overall, or of subgroups within the population, by geographic region are often used for the distribution of government funding, development programmes and services. Population censuses in many countries represent the very foundation of their national statistical systems, with census data providing important baseline data for policy development and planning, for managing and evaluating programme activities across a broad range of sectoral applications, and for monitoring overall development progress. The population census is used by international and regional organizations. An emerging use for census data is the assessment of good governance by civil society groups and measuring the achievement of sustainable Development Goals. The performance of a democratically elected government in improving the welfare of its citizens can be monitored from one census to the other by ordinary citizens through the widespread and timely dissemination of census results.

[9] The halo effect is a cognitive bias in which an observer's overall impression of a person, company, brand or product influences the observer's feelings and thoughts about that entity's character or properties. It was named by psychologist Edward Thorndike in reference to a person being perceived as having a halo or aureole. Subsequent researchers have studied it in relation to attractiveness and its bearing on the judicial and educational systems. The halo effect is a specific type of confirmation bias, wherein positive feelings in one area cause ambiguous or neutral traits to be viewed positively. Edward Thorndike originally coined the term referring only to people; however, its use has been greatly expanded, especially in the area of brand marketing.

1.22. Population censuses serve many programme needs by providing statistical information on demographic, human settlement, social and economic issues for local, national, regional and international purposes. For example, population censuses provide basic information for the preparation of population estimates or projections and detailed demographic and socioeconomic analysis of the population. The census can be used to update intercensal population estimates and population projections as well as provide data for the calculation of demographic and social indicators, particularly those that may be observed infrequently because they measure phenomena that change slowly over time, and those that are needed for small areas or small population groups.

1.23. In humanitarian settings, comprehensive and sound census outputs can be used to support the identification of crisis affected populations and specifically vulnerable groups.

2. Uses for research purposes

1.24. In addition to serving specific governmental policy purposes, the population census provides indispensable data for the scientific analysis and appraisal of the composition, distribution and past and prospective growth of the population. The changing patterns of urban–rural concentration, the development of urbanized areas, the geographic distribution of the population according to such variables as occupation and education, the changes in the sex and age structure of the population, and the mortality , fertility and net emigration differentials for various population groups, as well as the economic and social characteristics of the population and the labour force, are questions of scientific interest that are of importance both to research and for solving practical problems of industrial and commercial growth and population dynamics management.

3. Uses for business, industry and labour

1.25. In addition to those uses given above, the census has many important uses for individuals and institutions in business, industry and labour. Reliable estimates of consumer demand for an ever-expanding variety of goods and services depend on accurate information on the size of the population in subnational areas and its distribution at least by sex and age, since these characteristics heavily influence the demand for housing, transport, furnishings, food, clothing, recreational facilities, medical supplies and so forth. Furthermore, the census can be used to generate statistics on the size and characteristics of the supply and demand of labour needed for the production and distribution of such commodities and services in conformity with International Labour Organization statistical standards.[10] Such statistics on the local availability of labour may be important in determining the location and organization of enterprises.

4. Uses for boundary delimitation

1.26. One of the basic administrative uses of census data is to support political and administrative mapping. Detailed information on the geographic distribution of the population is indispensable for this purpose. Certain aspects of the legal or administrative status of territorial divisions may also depend on the size and characteristics of their populations, for example, whether a previously rural area is now to be declared as urban.

[10] See Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (October 2013).

1.27. A compelling use of census data is in the redrawing of electoral constituency boundaries in most countries. This is often enshrined in the country's constitution and provides a legal basis for census taking. The current distribution of a country's population is thereby used to assign the number of elected officials who will represent people in the country's legislature.

5. Use as a sampling frame for surveys

1.28. For countries without comprehensive population registers, population censuses serves as a foundation for construction of a sampling frame for surveys during the intercensal years on many topics, such as the labour force, poverty, fertility, mortality, disability and international migration.

1.29. Censuses provide an area frame that contains hierarchical geographical areas from the largest area (at the national level) to the smallest geographic division, usually called enumeration areas (EAs) and a list frame that contains the list of households located within each EA. For countries with resources address-based sampling frames have been gaining popularity given their efficiency and quality. The addresses are updated regularly and important auxiliary variables are available on the frame to help improve sampling efficiency. For countries that do not have the resources or capacity to maintain a comprehensive list of addresses, a master sample frame is often used. With a master sample frame, the address list is updated only for selected enumeration areas.

1.30. It is important to give careful consideration to the construction of a census for subsequent use as a survey sample frame when the census is in the planning stage. The census frame should be complete in coverage. This means, for example, that care must be taken to ensure that the entire country is divided into enumeration areas without any omissions or overlaps, i.e. all land area belongs to one and only one enumeration area. The coverage of population within the enumeration areas should be complete, regardless whether they reside in private households or not.

1.31. Population and household counts for the enumeration areas, taken from the census, are also a highly useful ingredient for post-census sample survey design planning. This information is often used to establish measures of size for the selection of first- or second-stage sampling units, or to help in various stratification schemes. Characteristics of living quarters, the households and individuals collected in the census are also useful as stratification variables to support efficient sampling design for surveys. For example, these may include the degree of urbanization, age, sex, education attainment, labour force status, type of living quarter, as well as other characteristics that may be used to support design of specialized surveys targeting rare and/or difficult-to-reach populations such as international migrants, persons with disabilities, indigenous population, etc.

1.32. Efforts should be made to ensure census records are geospatially enabled, i.e., geocoded to a specific location – this can facilitate selecting samples for household surveys as well as data integration.

1.33. Many countries use telephones to follow up with respondents to address missing values and nonresponse.[11] For countries without a good telephone frame for phone surveys, phone numbers collected during censuses can be used for subsequent surveys. However, this requires the following of strict protocols and consent from respondents.

[11] United Nations Statistical Commission. 51st session, item 3(j), background document. Report on the Results of the UNSD Survey on 2020 round population and housing censuses. 2020. Available from: https://unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3j-Survey-E.pdf

6. Use in humanitarian settings

1.34. Census data can be used in humanitarian settings in a number of ways, including:

- To assess the needs of the population: Census data can be used to identify the number and location of people who have been affected by a humanitarian crisis. This information can be used to plan and deliver relief and recovery efforts. For example, census data can be used to identify the number of people who need food, water, shelter, and medical assistance.
- To target assistance to the most vulnerable: Census data can be used to identify the most vulnerable populations in a humanitarian setting. This information can be used to target assistance to those who need it most. For example, census data can be used to identify women, children, and the elderly.
- To monitor the impact of humanitarian interventions: Census data can be used to monitor the impact of humanitarian interventions. This information can be used to improve the effectiveness of interventions and to ensure that they are reaching those who need them most. For example, census data can be used to track the number of people who have received food assistance, the number of children who have been vaccinated, and the number of people who have been able to return to their homes.

1.35. In addition to these general uses, census data can also be used to address specific humanitarian challenges. For example, census data can be used to:

- Track the spread of disease: Census data can be used to track the spread of disease in a humanitarian setting. This information can be used to prevent the spread of disease and to target public health interventions.
- Identify and protect vulnerable groups: Census data can be used to identify and protect vulnerable groups in a humanitarian setting. This information can be used to prevent violence and exploitation. For example, census data can be used to identify refugee camps and other areas where vulnerable people are concentrated.
- Plan for long-term recovery: Census data can be used to plan for long-term recovery from a humanitarian crisis. This information can be used to rebuild infrastructure and to support economic development. For example, census data can be used to identify areas where new schools and hospitals are needed.

1.36. Census data is an essential tool for humanitarian organizations. By providing accurate and timely information on the population, census data can help humanitarian organizations to save lives and to build a better future for people affected by crisis.

7. Use for monitoring internationally agreed development goals

1.37. Census data is essential for monitoring international development goals. The Sustainable Development Goals (SDGs), adopted by all United Nations Member States in 2015, provide a blueprint for a better and more sustainable future for all. They address the global challenges we face, including poverty, inequality, climate change, environmental degradation, peace and justice.
Census data can be used to track progress towards the SDGs in a number of ways, including:

- Measuring progress towards SDG indicators: Census data can be used to measure progress towards many of the SDG indicators. For example, census data can be used to measure SDG 1 on poverty, SDG 3 on health, and SDG 4 on education.

- Identifying disparities: Census data can be used to identify disparities in progress towards the SDGs. For example, census data can be used to identify disparities in access to education and healthcare between different groups of people, such as women and men, rural and urban populations, and different ethnic groups.
- Monitoring the impact of interventions: Census data can be used to monitor the impact of interventions designed to achieve the SDGs. For example, census data can be used to monitor the impact of a program to reduce poverty or improve educational attainment.

1.38. Census data is a valuable tool for monitoring progress towards the SDGs. By providing accurate and timely information on the population, census data can help to ensure that no one is left behind.

8. Use for population estimates and projections

1.39. Census data are essential for making population estimates and projections. Census data are used to make population estimates by adjusting the most recent census population counts for births, deaths, and migration. Census data are also used to make population projections by using a variety of demographic models that take into account factors such as fertility, mortality, and migration.

1.40. It is important to note that population estimates and projections are not perfect. They are based on assumptions about future trends in fertility, mortality, and migration. These assumptions can change, which can lead to inaccurate estimates and projections. However, population estimates and projections are still the best tool that we have for understanding the future of our population. and regular re-basing of them using a population census helps to limit inaccuracies.

B. Uses of housing censuses

1. Uses for development of benchmark housing statistics

1.41. The housing census produces benchmark statistics on the current housing situation and is vital for developing national housing and human settlements programmes. The housing census is also valuable for providing the sampling frame for special housing and related surveys during the intercensal years.

1.42. Housing benchmark statistics are also critical for emergency planning for response to natural hazards (such as destructive storms, earthquakes, tsunami and fires), or post-conflict situations. Following such situations, these statistics can be used to estimate the numbers of people and structures affected, the need for emergency response and reconstruction requirements.

1.43. National statistical authorities would need to develop, from housing censuses, the sort of benchmark statistics in housing that could be supplemented by current building and construction statistics and would provide continuous up-to-date information of the housing position needed for the consideration of housing programmes. Since not all the basic information required to assess housing needs or to formulate housing programmes can be obtained through a housing census, additional data must be obtained through the population census, special housing surveys and environmental surveys, and from vital statistics, economic statistics and so forth; but data obtained from the housing census will constitute the basic framework within which the estimates are made, indices computed and further statistical enquiries planned.

1.44. When population and housing censuses are carried out as a single operation or independently but in a well-coordinated fashion, the combined information provided is of much higher value than when conducted in isolation, since the essential features of both censuses are interrelated. The information on housing may be analyzed in association with the demographic and socioeconomic conditions of the occupants and vice versa.

2. Uses for the formulation of housing policy and programmes

1.45. The formulation of housing policies and programmes is one of the principal uses of housing census data. Housing policy is often influenced by social and economic as well as political considerations, and available factual data concerning the housing situation provide an objective basis on which policymakers can take their decisions.

1.46. In most countries, housing programmes encompass both governmental and private activity. The data derived from a housing census are used by governmental authorities for making an analysis or diagnosis of the housing situation. Housing conditions are analysed in quantitative and qualitative terms and data from previous censuses are used to indicate the changes in the housing situation that have occurred during the intercensal periods; the housing stock and future housing requirements are estimated and compared with the rates of dwelling production being attained; and the characteristics of the households in need of housing are considered in relation to the availability and cost of housing. As part of overall development plans, such an analysis is necessary for the formulation of national housing programmes and for their execution.

1.47. Commercial users also study housing census data. Those engaged in the construction industry, financing institutions, and manufacturers of housing fixtures and equipment and household appliances use housing census data, in combination with population census data, to assess the possible demand for these goods.

3. Assessment of the quality of housing

1.48. The materials used for the construction of housing units (roof, walls, floors) are a significant indication of the quality of life experienced by the occupants and, viewed in aggregate across multiple dwellings in an area and alongside other information such as weather conditions and date of construction, can provide an indicator of disaster risk. Trends indicated by census data in the type of housing materials used can show improvements in the welfare of the citizenry, such as where the percentage of poor-quality or slum-like housing facilities has decreased. Additionally, measures of water, sanitation and hygiene (WASH), as well as energy usage and telecommunication data are valuable for the planning and development of future infrastructure needs.

C. Relationship between the population census and the housing census

1.49. There is an especially close association between population censuses and housing censuses. The two censuses may constitute one statistical operation or they may be two separate but well-coordinated activities, but in either case they should never be considered completely independently of each other because essential elements of each census are common to both. For example, an essential feature of a population census is the identification of each occupied set of living quarters and of the persons living

therein, and an essential feature of a housing census is the collection of information on the characteristics of each set of living quarters in association with the number and characteristics of its occupants.

1.50. In many countries, the population and housing censuses are taken concurrently, often with the use of a single schedule. In this way, the information on population and living quarters can be more readily matched, processing is facilitated and extensive analysis can be carried out. This also makes it possible to relate the information on demographic and economic characteristics of each household member that is routinely collected in the population census to the housing census data.

1.51. In censuses that are conducted with direct enumeration, the advantages of simultaneous investigation may be offset to some extent by the additional burden on the respondent and the enumerator resulting from the increased amount of information that must be collected at one time. In countries where this is likely to be a serious problem, consideration might be given to collecting data for a limited number of topics on the basis of a complete enumeration in the population and housing census, with more complex data in both fields being collected on a sample basis only, either concurrently with or immediately following the full enumeration. Alternatively, consideration might be given to carrying out the housing census as part of the advance household listing operations of the population census.

1.52. The relationship between the population census and the housing census will affect the means by which data on homeless persons are obtained. In the case of simultaneous censuses of population and housing, data on homeless persons will be obtained as part of the population census. Where the housing census is carried out independently of the population census, it may be necessary to assign resources to enumerate homeless persons in the housing census. Information collected from enumerating homeless persons may reflect, among other things, the magnitude of the housing problem in a given locality.

D. Relationship of population and housing censuses to intercensal sample surveys

1.53. The rapidity of current changes in the size and other characteristics of populations, and the demand for additional detailed data on social and economic characteristics of population and housing characteristics that are not appropriate for collection in a full-scale census, have brought about the need for continuing programmes of intercensal household sampling surveys to collect current and detailed information on many topics.[12]

1.54. The population and housing census can provide the frame for scientific sample design in connection with such surveys (see paragraphs 1.25–1.29); at the same time, it provides benchmark data for evaluating the reasonableness of the overall survey results as well as a base against which changes in the characteristics investigated in both enquiries can be measured. To allow for the comparison of census and survey results, the definitions and classifications employed should be either identical or harmonized, while remaining consistent with the aims of each investigation. Because of the relative permanence of living quarters, the lists available from the housing census (with suitable updating) may also provide a convenient frame for carrying out enquiries dealing with topics other than population and housing.

[12] *Designing Household Survey Samples: Practical Guidelines No. 98* (United Nations publication, Sales No. E.06.XVII.13).

E. Relationship of population and/or housing censuses to other types of censuses and other statistical investigations

1. Census of agriculture

1.55. While the population and housing censuses have a close relationship, their relationship with the agricultural census is less well defined. However, as the result of increasing integration within programmes of data collection, the relationship between the population and housing census and the agricultural census is now far closer than in the past and countries are increasingly looking at new ways to strengthen this relationship.

1.56. One conceptual issue in relating the two censuses is that they use different units of enumeration. The unit of enumeration in the agricultural census is the agricultural holding,[13] which is the economic unit of agricultural production, while the units of enumeration in the population census are the household and the individual within the household. In many developing countries, however, there is usually a one-to-one relationship between households with own-account agricultural production[14] and agricultural holdings. In these cases the same unit is enumerated in both types of censuses. For countries where most agricultural production activities are carried out by households (that is, in the household sector), establishing links between the two censuses is particularly relevant.

1.57. The agricultural census collects various household or individual data for members of the agricultural holder's household. The *World Programme for the Census of Agriculture 2010*[15] recommends the collection of data on household size and limited data on demographic characteristics and economic activity of members of the agricultural holder's household, as well as some optional data on farm labour, such as persons working as employees on the holding. Users may find some agricultural activity data from the agricultural census more comprehensive than from the population census because the latter normally investigates the main work activity of each person during a short time reference period and this may not identify persons connected with agricultural activity on a seasonal basis or as a secondary activity. On the other hand, the population census provides data on all persons working in agriculture, including as paid employees. Such information is not available from the agricultural census, which only covers households with own-account agricultural production. To get a complete picture, agricultural data users will need both agricultural census data and population census data.

1.58. In planning the population and housing census, every opportunity for developing the relationship between this census and the agricultural census should be explored. This can take several forms. Definitions used in the population and housing censuses should be compatible with those used in the agricultural census so that meaningful comparisons can be made between the two data sets. The population and housing census can also be of use in the preparation of the agricultural census, such as in

[13] An agricultural holding is an economic unit of agricultural production under single management comprising all livestock kept and all land used wholly or partly for agricultural production purposes, without regard to title, legal form or size.

[14] That is, households with members engaged in agricultural activities as self-employed workers or as own-use producers.

[15] *A System of Integrated Agricultural Censuses and Surveys, Volume 1: World Programme for the Census of Agriculture 2010*, Food and Agriculture Organization of the United Nations, Statistical Development Series No. 11, Rome, 2005. [To be updated with the latest recommendations from the FAO.]

the demarcation of enumeration areas, the preparation of the frame for the agricultural census or, if applicable, the sample design.

1.59. In planning the national census programme, serious consideration should be given to the possibility of collecting key agricultural information as part of the population and housing census exercise that would facilitate the preparation of the frame of agricultural holdings in the household sector for a subsequent agricultural census. This could be done as part of the pre-census cartographic work and/or listing exercise or by adding a few questions to the census questionnaire (as elaborated in paragraphs 4.387–4.396). In the latter case, additional items at the household level could be included to identify whether any member of the household is engaged in own-account agricultural production activities. It may also be useful to collect additional data at the individual person level to identify persons involved in agricultural production activities during a longer period, such as a year. Information on occupation or industry and status in employment, and participation in own-use production of agricultural goods where relevant, could together facilitate identification of households with own-account agricultural production. Where countries choose to adopt this approach of using the population and housing census to establish a frame for the agricultural census, the agricultural census should be synchronized with the population and housing census and conducted as soon as possible after the population and housing census, while the frame is still up to date. This approach is detailed in the *Guidelines for Linking Population and Housing Censuses with Agricultural Censuses*.[16]

1.60. Linking population and agricultural census data can bring many benefits. This could add considerable analytical value to data sets from both censuses and save on data collection costs. Many of the demographic and activity status data collected in the population census are also collected in the agricultural census. If data from the two censuses could be linked, it would no longer be necessary to collect these data again in the agricultural census, while still allowing for comprehensive cross-tabulations.

1.61. A few countries conduct the data collection for the population and agricultural censuses as a joint field operation. Normally, each census retains its separate identity and uses its own questionnaire, but field operations are synchronized so that the two data collections can be done at the same time by the same enumerators. Occasionally, the two censuses are merged into one. This may have a number of advantages; however, as this is an increasingly complex operation, its impact on field operations and data quality needs to be carefully considered.

2. Census of establishments

1.62. Although the collection of information on industrial and commercial establishments does not constitute a part of the population census, the information that is collected from employers and own-account workers regarding the economic units they operate, such as the kind of economic activity and the size of the unit, can be used for preparing listings of the proprietors of such establishments. The information can also be used to prepare listings of the establishments themselves, if information is requested on their location or when the establishments are located within the living quarters (or dwellings). Experience shows that these listings can be used in a subsequent census of establishments or for supplementing the registers of establishments maintained by most countries and utilized as a list-based or area-based sampling frame for their establishment surveys.

[16] *Guidelines for Linking Population and Housing Censuses with Agricultural Censuses with Selected Country Practices*, Food and Agriculture Organization of the United Nations and the United Nations Population Fund, Rome, 2012.

1.63. Many business registers cover only establishments with fixed visible premises in which more than some minimum of persons (usually 5 or 10) are employed. In these cases, the population census can be used to collect basic information (such as kind of activity and size) on those establishments with employment below this minimum number of persons, by identifying the self-employed persons that operate them. However, special care should be taken in the choice of the unit of enumeration to ensure that there is no double counting of establishments.

1.64. When the information from the population census is to be used to construct a list-based sample frame, it is essential that the information from the population census be available and used shortly after the enumeration is carried out because this information can quickly become outdated. This requirement is less imperative when the information is to be used to construct an area-based sample frame.

1.65. The population census information needed for these purposes is the status in employment, in order to identify employers and own-account workers. For this subset of workers, information needs to be collected on the number of establishments operated, and for each of these, the kind of economic activity, the name and address of the establishment (if any), the number of workers engaged (including contributing family workers and employees) and whether the establishment is operated in partnership with other persons. If all of this information appears in the census questionnaire, the number of small establishments can be extracted from the schedule or from the processing documents after the enumeration.

3. Census of buildings

1.66. In certain circumstances, it may be necessary, as part of the housing census operations, to enquire whether buildings (both residential and non-residential) are occupied. Thus, it may be convenient to record basic information for all buildings at the time of the housing census, even though detailed data may be collected only for those in which housing units or other sets of living quarters are located. The comprehensive list thus obtained sometimes provides the basis for a census of buildings, carried out concurrently with, or subsequent to, the housing census, or it may provide for the identification of special types of buildings significant for other enquiries, such as the census of establishments or the census of schools. If a listing of households is to be carried out before the actual enumeration, this would be most ideal for carrying out such an exercise.

4. System of current housing statistics

1.67. Current housing statistics refer to latest data and trends in the housing market. They reflect the number of dwellings constructed and certain related information such as value, number of rooms, floor space, and so forth, as well as number of dwellings destroyed or demolished. These data are usually obtained from a system of data collection based on the administrative procedures required in connection with activities in the housing market. For example, construction statistics may be derived from permits issued for the construction of dwellings, from records of dwelling starts or completions, or from certificates of occupancy. Statistics on dwellings destroyed may be obtained from the records maintained for the levying of rates and the collection of taxes. Compiled monthly or quarterly, current housing statistics reflect changes in the housing inventory and, although they may serve other purposes, they are also used to update the benchmark data obtained from housing censuses.

F. Relationship between the population census and/or housing census and administrative registers

1. Civil registration and vital statistics

1.68. Population censuses are a fundamental baseline for the Civil Registration and Vital Statistics (CRVS) systems. By comparing census records with CRVS records, gaps in registration can be identified. In addition, Census data can:

a. help identify vulnerable populations, such as refugees, internally displaced persons, or marginalized communities, who may face challenges in accessing CRVS services;
b. be used to estimate mortality rates (e.g., age-specific death rates) and fertility rates, which are essential components of vital statistics.
c. be used as a benchmark for evaluating the performance of CRVS systems.

1.69. The United Nations Statistics Division (UNSD) in Handbook on Civil Registration and Vital statistics Systems define civil registration as "the continuous, permanent, compulsory, and universal recording of the occurrence and characteristics of vital events pertaining to the population, as provided through decree or regulation in accordance with the legal requirements in each country". A well-functioning civil registration and vital statistics (CRVS) system helps ensure that every person has a legal identity, facilitating access to the public benefits and social protections and human rights. It also produces vital statistics on population health and serves as an optimal source of data on the prevalence, distribution, and cause-specific mortality for evidence-based decision policy, planning and programme implementation.

1.70. The 2030 sustainable development goal (SDG) agenda highlights the importance of CRVS system with many targets and indicators dependent on data from the CRVS system. Additionally, data from complete, reliable, and timely CRVS systems can provide cause-specific mortality to monitor many targets and indicators listed in the SDGs. Complete CRVS system can provide reliable population denominators which are often needed for measuring progress on any indicators that mention a population ratio, such as "per 1000 population", "live births" or "per capita".

1.71. Population census data can serve as denominators for the computation of vital rates, especially rates specific for characteristics normally investigated only at the time of the census. Conversely, census results, time-adjusted by vital and migration statistics, can provide estimates of the future size, distribution and other characteristics of the population of the total country and subnational areas. Furthermore, census data on fertility can provide a benchmark check on the reliability of current birth statistics, and vice versa. It is consequently desirable that procedures for the collection of population census data, vital statistics and migration statistics be closely coordinated with regard to coverage, concepts, definitions, classifications and tabulations.

1.72. Some countries have linked individual census returns for infants less than 1 year of age with birth registration reports for the year preceding the census date as a means of checking on the completeness of one or the other type of investigation. Linkage of death reports with census returns has been used to

compare the information on characteristics of the deceased as reported in the two sources.[17] While the many problems posed in the past by the one-to-one matching of two types of records have not been entirely solved, their severity has been mitigated by developments in computer technology. Before undertaking either of the procedures, however, countries should consider carefully the possible advantages of using household sample survey returns rather than census returns in the operation. Moreover, such operations have to be carried out in complete accord with national laws and policies governing the confidentiality of information obtained in the census if public confidence in the census is to be maintained.

1.73. In the establishment of a civil registration system, census results on the geographic distribution of the population can be useful for determining appropriate locations for registration offices.

2. Population register

1.74. The term "population register" was defined in 1969, in the publication entitled *Methodology and Evaluation of Population Registers and Similar Systems* (United Nations, 1969), as "an individualized data system, that is, a mechanism of continuous recording, and/or of coordinated linkage, of selected information pertaining to each member of the resident population of a country in such a way to provide the possibility of determining up-to-date information concerning the size and characteristics of that population at selected time intervals" (chap. I.A). Thus, the population register is the product of a continuous process, in which notifications of certain events, which may have been recorded originally in different administrative systems, are automatically linked to it on a current basis. The method and sources of updating should cover all changes so that the characteristics of individuals in the register remain current. Because of the nature of a population register, its organization, as well as its operation, should have a legal basis.

1.75. Basic characteristics that may be included in a population register are date and place of birth, sex, date and place of death, date of arrival/departure, citizenship(s) and marital status. Depending on the possibility of proper linking with other registers, much additional information may be added to the single record, such as language(s), ethnicity, educational attainment, parity, activity status and occupation. In order to be useful, any additional information must be kept up to date. If complete, population registers can produce data on both internal and international migration through the recording of changes of residence as well as the recording of international arrivals and departures.

1.76. There is no universally adopted definition for a population register. However, a population register may be taken to mean an administrative register primarily used in an administrative information system by public agencies for their own specific purposes such as establishing personal identification, voting, education and military service, social insurance and welfare, and for police and court reference.

1.77. A key requirement is that each unit in the Population register should be always uniquely identifiable. This is best achieved by using a system of identification codes (keys). Identification is however also possible without such a code being used, if sufficient information is available on the units (for persons: name, Address, date of birth, etc.).

[17] An elaboration of comparison between census and vital statistics is provided in *Principles and Recommendations for a Vital Statistics System, Revision 3*, United Nations publication, Sales Number E.13.XVII.10, United Nations, New York, 2013 (see paragraphs 595–597).

3. Building/dwelling or address register

1.78. An address register is an up-to-date comprehensive list of all known physical addresses within a country. In order for it to be a key piece of statistical infrastructure that can be used for a census it should contain key pieces of information such as address text, address use, a latitude and longitude, dwelling structure for residential addresses and non-private dwelling information (i.e. hotels and caravan parks).

1.79. An address register can enable the running of a digital Census by providing a mail out frame which is a dataset of residential and non-private dwellings (such as hotels and caravan parks) and is used to support the mail out of census materials. The latitude and longitude and dwelling structure is particularly useful for helping to locate dwellings in remote areas.

1.80. An address register can be also be used to provide an administrative census by using common variables (i.e. address text) to link to other government data sources in order to add further value to the current information, for example, employment and income data. Utilities data could also be linked to an address register to help determine if a residential address is being actively lived in or whether it is only used at certain times across the year.

1.81. As there are certain fields (particularly in an administrative census) that exist as inputs from an administrative register and as outputs of the census it is important to ensure that the descriptions and concepts of these fields are harmonised to make it fit for purpose.
For instance, non-private dwellings and dwelling structure both appear as an input from an address register but also as an output in a census, in order to ensure the outputs are accurate the definitions of the different non-private dwellings and dwelling structures need to be consistent across the address register and the census.

4. Business register

1.82. Statistical Business Registers (SBRs) provide the infrastructure to support the collection of high quality, reliable and consistent economic data and the production of economic statistics. SBRs are maintained by National Statistical Organisations and used for statistical purposes.

1.83. To ensure effective and complete coverage, SBRs must be regularly maintained and updated. Administrative data sources such as legal or taxation business registers including unique identifiers are used to update less significant units with simpler structures. Large, complex and significant businesses are updated through profiling activity to transform these businesses from administrative data to statistical units. Other sources of updates may include economic censuses, survey feedback, big data, web scraping or data mining.

1.84. SBRs provide a coherent set of statistical units and classifications to collect and assemble data. SBRs include units represented in the economic units model, which describes the characteristics of businesses and the structural relationships between related businesses. The types of units that may be part of SBRs include:

- enterprise group - a set of enterprises grouped by financial or legal arrangements and controlled by a group head
- enterprise - units grouped by institutional sector

- legal entity - can differ widely between and within countries, these are units that own goods or assets, incur liabilities and can enter into contracts, with single or multiple legal entities forming an enterprise
- kind-of-activity unit - units grouped by a single principle productive activity
- establishment - single locations with one principal productive activity.

1.85. SBRs are used to create regular point-in-time frames (frozen or common frames) from which survey or sampling frames (subsets) for most economic surveys are drawn, ultimately feeding into the System of National Accounts. SBRs are also used to analyse and publish business demography statistics. Business demography includes events such as creation and cessation of units, survival information, and changes in business characteristics over time including type of activity, size measures (employment, turnover), location, institutional sector, employing versus non-employing units.

5. Other administrative data sets

1.86. Other admin data include such as records of foreigners and foreign workers, people living in institutional places, enrolment and graduation records, social insurance, employment records, etc. The availability and use of these sources is very country specific, available in some, not in others.

1.87. There is an increasing availability of a range of government and private data sets containing information on the persons or the households within a country. The utility of these data sets for statistical analysis may be limited, in some countries, by their lack of population coverage, data accuracy or range of characteristics, especially in the case of newly-developed sources and those designed for purposes in which completeness and accuracy are not key quality criteria. The linkage of these data sets with the census file, with its complete coverage of the persons and households, can provide the ability to create new insights and new statistical products to leverage more value from the census.

1.88. Administrative data may replace data obtained by direct enumeration – for example, in some countries income data from the taxation or revenue department can replace the need to directly collect those data in the census. Administrative data can also extend census data – for example, census data can be linked with visa information or health information to extend the census data set into areas that may be too sensitive to collect on the census form, or with past education data to analyse longitudinally the impact of education on labour force outcomes. Administrative data can also replace missing data – for example, in one country health records have been used to impute the count and characteristics of usual residents who were non-responding during the census enumeration period.

1.89. As described above in paragraph 1.58, linkage operations should be undertaken with caution, ensuring not only that all national laws are met but also that the trust of the public in the census and the statistical systems is maintained.

G. Integration of census data with other data sources

1.90. Data integration is the process of combining data from two or more sources at the unit level (for example, persons, dwellings, businesses) or at the aggregated level. Data integration has the potential to improve cost-effectiveness and relevance of any single data collection activity. It also has the potential to enhance the coverage, timeliness and the richness of available data, allowing more complex questions to be analysed and more insights to be drawn than would be possible from a single source. Many integrated

data assets are also longitudinal in nature, meaning they allow changes and patterns in population, economy and environment to be better understood over time.

1.91. Integrating data can also serve to reduce respondent burden by reducing the number of questions that need to be asked in censuses or surveys as shown in some countries when data from administrative sources are used with proper consent from respondents. Censuses are often used as a spine[18] for linking population data from multiple sources.

1.92. Data integration at unit level is made possible with the use of key linking variables. In the case of people, this may be unique identifiers available from both data sources; or names, sex, addresses, date of birth etc., if unique identifiers are not available for the linkage.

1.93. Data integration at the aggregated level required is carried out through modelling, for two different sources borrowing strength from each other. One example is the small area estimation that has been traditionally used for poverty mapping for lower geographical areas but has now been extended to cover a wide variety of outcome indicators including employment and social welfare. The method is also being used by countries to produce estimates for hard-to-reach population groups such as persons with disability.

1.94. For countries that do not have access to good administrative data, population censuses have been an important source to be integrated with household surveys, for two reasons. First, population censuses tend to be more comprehensive in coverage compared to household surveys. Second, population censuses collect a number of variables that can be used as auxiliary variables for the data integration. On the other hand, household surveys collected a lot of more in-depth information with limitations on coverage. Combining the two sources can produce data that are more granular and timelier.

1.95. Successful integration of census with other data sources requires careful planning and designing of the census data collection, so it is "interoperable" by design. This might include, for example, using the same concepts and definitions for key socio-demographic variables across data sources. Georeferencing censuses is another important element to ensure that census data can be integrated with other sources through geospatial information.

1.96. Geospatial information (e.g. address, x- and y-coordinates, or a geographical name) can play an important role in bringing together information from various domains by enabling integration of datasets from different sources using the location information as a matching key variable (e.g. integrating administrative data with survey data using address or postal code that exists in both datasets). To ensure the quality of the integration, standardising the geospatial information in the different datasets is critical. This standardisation would normally take place before the integration of datasets and can be done through, for example, matching location information in the datasets with a centralised standard system (e.g. address matching, geocoding) which should be part of the national spatial data infrastructure. This linkage, ideally done through consistent, unambiguous and persistent identifiers (PIDs), can also allow the dataset to use various additional geospatial information within the address registry or geocode database.

[18] A **population spine** is a concept related to a reference population that is used to link together and manage different data sets about individuals. It is often a term used in relation to register-based censuses, which makes it easier to track changes in the population over time.

In the absence of such a system, organisations may rely on other ways to reference location (e.g. GPS coordinates), alternative sources (e.g. address registry from utility provider) or higher-level geography (e.g. large geographical area).

IV. Census methodology

1.1. There are different methods of data collection, including traditional collection directly from the public in a field census, or the use of data sources based on administrative records or methods mixing administrative records with data collected from the field.

1.2. Summarizing the experiences of the previous population and housing census round,[19] it became evident that a number of countries were exploring the use of alternative methodologies with respect to the conventional, field-enumeration based census for producing census statistics. The use of registers – primarily population registers – in combination with other sources is being considered in a number of countries for the purpose of producing detailed small-area statistics on population and housing, as well as the application of continuous survey methodology for the same purpose. Furthermore, these alternatives to the traditional method of conducting population and housing censuses are becoming more diverse in terms of developing combinations of various data collection methods (see paragraph 1.95), and it is thus a challenge to summarize and categorize them using generally accepted data source methodologies.

1.3. It should be noted that most countries are expected to continue using the traditional census approach – soliciting information from each household in a country – in the 2020 round of censuses,[20] while at the same time it is anticipated that increasing numbers of countries will intend to use alternative methodologies. There are quite a few reasons for exploring alternative approaches, and the following presents a sample: (a) the need to produce more frequent and timely statistics; (b) budgetary limitations for census taking; (c) reluctance of the population to participate in the census; (d) increased technical capacities to manipulate data sources; and, e) utilizing data available on more variables from multiple sources.

1.4. This section aims to briefly elaborate on possible methodologies for conducting censuses based on the recent experiences of countries. The section also describes the necessary conditions for using a specific methodology, its advantages and disadvantages, and its implications for the content and administration of the census. It should be kept in mind that countries using a specific census methodology might have significant differences in implementation of the methodology, arising from differing country conditions and expectations. Regardless of the approach, the crucial principle of providing detailed statistics at the lowest geographic level remains of paramount importance.

1.5. The various census methodologies are represented in a matrix in Table 1, where the rows describe data collection through field enumeration and the columns represent use of administrative or population registers as census data sources. The matrix presents only those options that either have been used or

[19] The 2010 round of population and housing censuses was inaugurated by the Economic and Social Council of the United Nations and covered the period from 2005 until 2014. [To be updated with experience from 2020 round.]
[20] The 2020 round of censuses covers the decade 2015–2024.

are likely to be used by countries and does not present all possible combinations, including theoretical ones that have yet to be tested by any country.

1.6. The different approaches are explained in Table 2.[21] First, the full field enumeration and the register-based census are presented; then the combined methodologies are described. Alternative approaches have been adopted in different ways by different countries, depending on national preferences and practices and the availability of appropriate data sources.

1.7. The columns in the matrix present different types of registers: administrative registers, statistical registers and base registers. Administrative registers are registers that are created and used mainly for administrative purposes outside the national statistical authorities. An administrative register will be edited, corrected and perhaps imputed into a statistical register inside a statistical institute and can then be used for statistical purposes. A statistical register can also be established inside a statistical institute for statistical purposes; one example could be a register of occupations that in many cases does not have any administrative purposes. Base registers are registers, such as the population register, dwelling register or enterprise register, that create a population base for individuals, dwellings and enterprises. Other registers will then be matched with a base register. The household register will be created by combining the population and dwelling registers.

[21] For more details see *Guidelines on the use of registers and administrative data for population and housing census,* UNECE, 2018; *Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices*, United Nations Economic Commission for Europe, ECE/CES/13, Geneva, 2011; and *Register-based statistics in the Nordic countries: Review of best practices with focus on population and social statistics*, United Nations publication, Sales Number E.07.II.E.11, United Nations, Geneva, 2007.

Table 1. Overview of census-taking methodological approaches

Type of data collection	Use of registers as census data sources		
	No registers (fully field enumeration based or sample surveys)[22]	**Base registers (individuals, households, dwellings)**	**Integrated administrative sources**[23]
Full field enumeration	Full field enumeration only (traditional census)	Base registers and full field enumeration	Integrated administrative sources and full field enumeration
Rolling surveys (continuous surveys)	Full field enumeration and rolling surveys Rolling census	n.a.	n.a.
Ad hoc sample surveys	n.a.	Base registers and ad hoc sample surveys	Integrated administrative sources and ad hoc sample surveys
Existing sample surveys	n.a.	Base registers and existing sample surveys	Integrated administrative sources and existing sample surveys
No field enumeration (fully register based)	n.a.	n.a.	Fully register based

[22] In a fully field enumeration-based census, data from registers are not used as a census data source, even though registers may be used as a frame and to support field operations.
[23] Integrated administrative sources with information on business, tax, education, employment and other relevant registers.

Table 2. Descriptions of approaches

Approach	Description
Full field enumeration (traditional census) or full field enumeration and rolling surveys (see E.1)	Information on census topics concerning individuals and households is collected by census questionnaire directly from respondents using enumerators or other modes of data collection (e.g. telephone interview, mail out/mail back, Internet), or by applying a combination of different modes of data collection. In a traditional census, full field enumeration may include an in-built ad hoc survey by use of the long form or can be combined with a rolling survey. It is common that a short form is used together with a long form. Both forms are utilized during the same time frame of the census. While the long form estimates are not based on full coverage, they are regarded as census output. Through modelling, the full field enumeration with the continuous cumulative survey is used to generate yearly (or other interval) estimates of detailed characteristics for different geographic levels. With this approach, the sample can be cumulated over time to produce statistics at the lowest levels of geographic detail to provide more frequent and relevant data.
Rolling census (see E.3)	Information on individuals and households is collected through a continuous cumulative survey covering the whole country over a period of time (generally years), rather than a particular day or short period of enumeration. The two main parameters of a rolling census are the length of the period of enumeration (which is linked to the frequency of updates required) and the sampling rate (which depends on the geographic levels required for dissemination purposes).
Combined methodology (see E.3)	Information on individuals and households is collected by combining data collected from one or more surveys or full field enumeration with administrative or statistical registers. Data from registers are employed not only as a frame or to support field operations, but directly as a data source for some census information. In some cases, register data are used to prefill the questionnaires to be verified or corrected during data collection. Ad hoc sample surveys are used to provide information on census topics not available from administrative sources or to adjust data that are of poor quality in registers.
Base register and existing sample surveys or integrated administrative sources and existing sample surveys (see E.2)	Information on individuals and households is collected from existing administrative sources, namely different kinds of registers, of which the following are of primary importance: individuals, households and dwellings. These are linked at the individual level with information from existing sample surveys. No field data collection will take place. Existing sample surveys include intercensal sample surveys on different topics, such as the labour force survey and the living standards survey.
Fully register based (see E.2)	Information on individuals and households is collected from existing administrative sources, namely different types of registers, of which the following are of primary importance: individuals, households and dwellings. These are linked at the individual level with information taken from other administrative or statistical sources, such as business, tax, education, employment and other relevant registers.

A. Full field enumeration (traditional census)

1.8. The full field enumeration or traditional approach comprises a complex operation of actively collecting information from individuals and households on a range of topics at a specified time, accompanied by the compilation, evaluation, analysis and dissemination of demographic, economic and social data pertaining to a country or a well-delimited part of the country. Members of the public respond to a census questionnaire, or interviewers are deployed to collect information from respondents. For interviewer-based censuses, enumerators assigned to different enumeration areas cover all households and persons in the enumeration area during a specified and usually short period of time in order to meet the requirements of universality and simultaneity.

1.9. Various methods can be used for collecting the data, including a mailed or dropped-off questionnaire, the telephone, the Internet, personal visit follow-up, or a combination of such methods, countries employing the traditional design may utilize very different collection approaches in doing so.

1.10. The traditional census has merit in providing a snapshot of the entire population at a specified period and data for small geographic domains. In that sense, the traditional census is perhaps unique in nature. This approach is particularly suitable for countries requiring population numbers by various social and economic characteristics simultaneously for all geographic levels to meet the needs of planning and the allocation of funds. The delimitation of electoral boundaries requires simultaneity, and for that reason also the traditional approach may be more appropriate. But at the same time, traditional censuses have been singled out as the most elaborate, complex and costly data collection activity that national statistical authorities undertake. In addition to costs, this complex task requires full awareness and agreement of the public to participate in it.

Necessary conditions

1.11. It is essential to have national legislation for conducting the population and housing census to ensure confidentiality, transparency and the cooperation of the population. A permanent central census organization, which may or may not be part of the statistical office, needs to exist in the country, which can be expanded during the time of the census. Since a traditional census requires substantial resources, sufficient funding for a field operation covering the entire country and subsequent data processing needs to be ensured. Other conditions necessary for this approach are the support of and general acceptance by the public to participate in the enumeration, and trust towards the statistical office or census agency.

Some considerations for census taking and content

1.12. Very precise planning is required for every stage of census taking in the traditional approach, due to the sheer volume of work and overlapping time frames. The recruitment and training of a large number of census takers adds to cost and complexity. Involvement of administrative machinery at the central, provincial and local levels is essential for successful field operations.

1.13. Since the data obtained in the traditional method are respondent and enumerator based, there is scope for error in canvassing the questions and in the quality of response. This, however, can be minimized through proper design of the questionnaire, effective training and wide publicity.

1.14. Data can be provided for every administrative level subject to privacy and confidentiality considerations, which may not always be possible with other methods if some parts of the data collection are based on sample surveys.

1.15. Essential features of a population and housing census are fully satisfied with the traditional census method.

Advantages and challenges (of traditional censuses)

1.16. The two biggest advantages of a traditional census are comprehensiveness of coverage and simultaneity. Another major advantage is the flexibility in deciding the topics to be covered and design of the questionnaire. There is lesser need for complex data adjustment since processing of raw data provides all inputs. The census frame becomes the base for all subsequent sampling frames. Finally, the focused and time-bound nature of the field operation implies that the data collection is finished in a short period and does not require long-term or constant monitoring.

1.17. One of the biggest disadvantages of a traditional census is its cost and administrative complexity. Another disadvantage is that it has a very long processing time. Also, since it can be conducted only after 5- or 10-year intervals, data tend to become outdated. The burden on respondents can be considered a disadvantage, especially in countries where participation in the enumeration is declining. Finally, many countries experience increasing difficulties in enumerating specific population groups, such as persons with high mobility or with multiple residences, or who are difficult to reach for other reasons. Unexpected emergencies such as natural disasters and public health incidents could also significantly disrupt the traditional census.

1. Types of censuses conducted with full-field enumeration

1.18. There are two main types of the traditional census: the census with a single questionnaire and the census with both short and long questionnaires. In the latter case, the short form contains only questions intended for universal coverage, while the long form is used to collect information from only a sample of households and population. In addition to these two main types, there are also a number of hybrid census approaches that combine elements of both types of censuses.

(i) Traditional census with a single questionnaire

1.19. A traditional census with a single questionnaire is a census in which all households receive the same questionnaire. This type of census is conducted using traditional methods, such as door-to-door enumeration or self-enumeration by mail or online.

1.20. There are several advantages to using a traditional census with a single questionnaire. First, it is a proven method that has been used successfully for many years. Second, it is relatively simple to administer and process. Third, it can help to ensure that all households are counted equally. However, there are also some disadvantages to using a traditional census with a single questionnaire. One is that it can be expensive and time-consuming. Another is that it can be difficult to reach certain populations, such as people living in remote areas or people who are homeless.

1.21. Overall, the decision of whether to use a traditional census with a single questionnaire or a different method depends on factors such as the budget, the resources available, and the specific needs of the country or region.

1.22. Traditional censuses with single questionnaires are still the most common method of conducting a census today, but countries are increasingly looking for ways to improve the efficiency and effectiveness of these censuses. They are also exploring alternative sources of data, such as administrative records.

(ii) Traditional census with both short and long questionnaires

1.23. A traditional census with both short and long questionnaires is a census in which all households receive a short questionnaire, and a sample of households receive a long questionnaire with additional questions. The short questionnaire, meant for universal coverage, collects basic demographic information, such as name, age, sex, marital status and ethnicity. The long questionnaire, on the other hand, collects more detailed information on a variety of socioeconomic and housing topics, such as education, employment, income, and housing. Both forms are utilized during the same time frame of the census. While the long form estimates are not based on full coverage, they are regarded as census output.

1.24. The use of two questionnaires allows the census to collect a broad range of data without burdening all households with the long questionnaire. However, this type of census can be more complex and expensive to administer than a census with a single questionnaire. Here are some of the advantages and disadvantages of using a traditional census with both short and long questionnaires:

- Advantages: Collects a broad range of data; Reduces the burden on households; Can be used to estimate population characteristics for small geographic areas.
- Disadvantages: More complex and expensive to administer than a census with a single questionnaire; May be subject to sampling error; May be difficult to reach certain populations.

1.25. The use of a single form can reduce the cost and complexity of the census, and ensure that all households are counted equally. However, the single-form census may not collect enough data to meet the needs of policymakers and researchers. The use of two-form census, or using alternative methods, such as administrative records and sample surveys, to collect more detailed data. The decision of whether to use a traditional census with both short and long questionnaires, a single-form census, or an alternative method depends on a number of factors, such as the budget, the resources available, and the specific needs of the country or region.

(iii) Traditional census with annual continuous surveys

1.26. A traditional census with annual continuous survey is a hybrid approach to population enumeration that combines a traditional census with a continuous survey. In a traditional census, typically conducted once every 10 years, all households are counted at a specific point in time. This means that the data collected from the traditional census can become outdated over time. In a continuous survey, a household sample is interviewed regularly, such as quarterly or annually. This means that the data collected from the annual continuous survey is more up-to-date than the data collected from the traditional census. However, the annual continuous survey only collects data from a sample of the population, so it is not as accurate as the traditional census. By combining the two methods, a traditional census with annual continuous survey can provide both timely and accurate data on the population. This

approach generates yearly (or at other intervals) estimates of detailed characteristics of population and housing. This can be useful for planning government programs and services, allocating resources, and conducting research.

1.27. The primary advantage of this approach is to provide more frequent and relevant data on population and housing than would be available when a census is conducted only once a decade. However, such a programme might be costly and technically difficult, as it requires a multi-year round of comprehensive planning, development and testing.

1.28. Here are some of the advantages and disadvantages of using a traditional census with continuous survey:

- Advantages: Produces more frequent and up-to-date population estimates; Can be used to estimate population characteristics for small geographic areas; Reduces the burden on households.
- Disadvantages: More complex and expensive to administer than a traditional census; May be subject to sampling error; May be difficult to reach certain populations.

1.29. A traditional census with continuous survey can produce more frequent and up-to-date population estimates while reducing the burden on households. However, this approach is also more complex and expensive to administer than a traditional census. The decision of whether to use a traditional census with continuous survey is complex and depends on factors such as the budget, resources available, and specific needs of the country or region.

(iv) Rolling census

1.30. The rolling census is a variation of the traditional census whereby the whole population of a country is covered over a period of time rather than on one particular day as the other types of censuses. The rolling census involves conducting a sample survey of households each year, and then combining the data from the sample survey with data from previous years to produce annual population estimates.

1.31. In a rolling census, information is collected on individuals, households and dwellings by a continuous cumulative survey covering the whole country over a period of time (generally years), rather than a particular day or short period of enumeration.[24] The two main parameters of a rolling census are the length of the period of enumeration and the sampling rate (which depends on the geographic levels required for dissemination purposes). For example, it is possible to build a sample framework in order to produce national results with one annual survey, regional results by cumulating three annual surveys, and small-area results by cumulating data over five years. Annual surveys may be conducted over the full course of the year or in a particular month or other shorter time frame.

1.32. Implementation of such an approach requires highly complex sampling and modelling techniques; a high-quality sampling frame in order to allow sampling at very low levels of geography; and successful consultation to gain acceptance of the approach with major stakeholders, including national and local governments and the user community.

Necessary conditions

[24] In the 2010 round of censuses France was the only country applying this concept.

1.33. The necessary conditions partly depend on the complexity of the sample framework. If the sampling units are addresses, a master address file is to be built first. But if the sampling units are larger, for example municipalities, it is only necessary to have enough information to spread the municipalities over the different years. It will be necessary to explain to statistics users the impact of the rolling sample on the use and interpretation of data, as many users are more used to snapshot data rather than period data.

Advantages and disadvantages

1.34. The main advantage of the rolling census approach is the higher frequency for updating data: a traditional census provides benchmarks every five or, more commonly, ten years. In contrast, the rolling census provides annual updates. Another advantage is the reduction in the burden on the public. The high peak costs and labour requirements of a traditional census are instead spread over a longer period. Furthermore, it is possible to improve the census process over time, and to test methodological refinements and new technologies as they emerge.

1.35. The disadvantage is that the rolling census approach no longer provides a snapshot of the whole population, complicating comparisons between areas due to different enumeration times. In addition, as the rolling census covers the whole country over a period of time, some respondents will move. Thus some people may be surveyed several times and others may not be surveyed at all. As a result, universality might not be ensured unless careful methodological adjustments are made. Another disadvantage is that the reference date is not the same for the entire population.

Some considerations for census taking and content

1.36. It is better to begin a rolling census just after a full traditional census, in order to exploit the recent census information to build the sample framework. As the operation is annual, the process must be very carefully prepared, since any delay can be problematic for the following stages.

1.37. A rolling census is able to include all usual census topics. There is also the possibility of changing the questions more regularly than in a decennial cycle. This enables the census to be more reactive to changes in the needs of users, even if comparability over time should in principle be preserved. However, only if the questions are stable over a number of years can a rolling census produce statistics at the same level of detail. Depending on the census organization and procedures, it may be possible to add some thematic surveys if required

2. Innovations in conducting traditional censuses

1.38. There are a number of innovations that NSOs can pursue to improve the efficiency, accuracy, and accessibility of traditional censuses. Some of these innovations include:

a) **Technology**. Technology is playing an increasingly important role in conducting traditional censuses. For example, many countries are now using mobile devices to collect census data in the field. This can help to improve the efficiency and accuracy of data collection. Some countries are experimenting with using machine learning to develop new methods for estimating the population of small geographic areas, and to improve the accuracy of annual population estimates.

b) **Administrative data**. Administrative data is providing useful information to improve the preparations and collection of a census. Some examples of this include:
   - Creation of address register
   - Inform on planning and collection operations
   - Dwelling occupancy verification
   - Informing on areas likely to respond using paper forms
   - Informing on likely areas of hard to reach populations
   - Geocoding special locations (eg places of education)
   - Determining non-response follow up contacts
   - Quality assurance of online responses

   Administrative data can also be used to enhance the processing and analysis of census data. For example:
   - Administrative enumerations to address Census non-response (whole records).
   - Providing missing characteristics (non-response to specific items)
   - Improving whole dwelling imputation
   - Improving person imputation (e.g. age imputation)
   - Quality assurance/validation
   - Resolving multiple responses
   - Bias adjustments (for population estimates)

c) **Data integration.** Countries are also increasingly integrating data from different sources to produce more complete and accurate census data. For example, some countries are using administrative records, such as birth and death records, to update their population data, while others are using satellite imagery to estimate the population of remote areas. . Others are using a combination of administrative records and sample surveys and geospatial data to produce more frequent population statistics.

d) **Using alternative population bases**: Generating population statistics using alternative population bases.

e) **Public engagement**: Countries are also working to improve public engagement with the census process. This is important because it helps to ensure that everyone is counted and that the census data is accurate.

1.39. These are just a few examples of the many innovations that are being used to improve traditional censuses. By using these innovations, countries can produce more accurate, timely, and accessible census data. In addition to the above innovations, countries are also exploring the use of new technologies, such as artificial intelligence and machine learning, to produce population statistics. These technologies have the potential to revolutionize the way that population statistics are collected and produced.

B. Register-based censuses

1.40.The concept of producing census-like results based on registers was developed in the 2000 round of censuses, although it has been debated and tested to various degrees since the 1970s, and several countries succeeded in using this approach to generate census data in the 1990 round of censuses. The philosophy underlying this concept is to take advantage of existing administrative sources, namely different kinds of registers on individuals, households and dwellings. These registers are linked at the individual record level with information held on business, tax, education, employment and other relevant registers. While it is theoretically possible to link records on the basis of the name and other unique details

of the individuals, the existence of a unique identification number for each individual, household and dwelling allows a much more effective and reliable linkage of records from different registers.
1.41.A precondition that the registers used in the census must meet is that the definition of the variables existing in them fits well with the definition of the census variable on which information must be provided.

1.42. Administrative registers can, depending on its content and quality, be used in all phases of census taking. In principle, where greater amounts of information can be obtained from administrative sources, the production of census-type statistics will be faster, cheaper and more complete. The most complete use of registers will be where all core, and in some countries non-core, census topics can be based on register information. It is possible to improve the quality of data collected through the administrative register by conducting a sample survey or surveys. The sample survey(s) may either use the register as a sampling frame, or else be completely independent of the register.

1.43. To provide specific demographic data and socio-economic characteristics, the Register-based Census relies on the collection of available administrative data and information from government and private entities, as well as other service providers. . The adoption of the new record-based approach will enable the provision of systematic, wide-ranging and timely statistical products that support informed decision-making and anticipate countries' future.

1.44. The use of register data may reduce the flexibility of the census exercise in terms of the variables that are available and their definitions. It may be difficult to change the variables as these are defined in line with administrative priorities. A significant potential risk for the success of the census exercise is that the administrative source will often be outside the control of the statistical authority. The influence of the statistical authority over the administrative source can be very limited. The content and availability of the administrative source may change at relatively short notice and without reference to statistical needs. For example, a change in taxation legislation may mean that a key administrative register may no longer collect information needed for the census. This risk can be minimized by establishing close and regular communication between the statistical authority and the owners of the administrative sources.

1.45. In some situations, only persons legally present in the various registers would get covered through this approach. For example, unregistered births, deaths or marriages, illegal immigrants, homeless persons, nomadic or floating populations, persons involved in illegal activities, etc., are not likely to be recorded on certain occasions in any such administrative register. On the other hand, registers may include persons who are actually not living (any longer) in the country, for example persons who emigrated but were not cancelled from the registers.

1.46. Summing up, subject to the caveats mentioned in the paragraphs above, the feature of individual enumeration is satisfied in this approach as separate information is collected regarding the characteristics of each individual. As regards “universality within a defined territory”, this criterion is satisfied as the enumeration is taken from a population register in which the fields for attributes are populated from subsidiary registers relating to specific topics. With regard to “simultaneity”, the timing of the census extraction may require careful thought where register update cycles vary. With respect to “periodicity”, this approach allows extraction at desired frequencies, including “at least once in 10 years”, noting again the need to manage the updating cycles for the registers. Finally, in most cases, the requirement in terms of producing small-area statistics is largely met, as the information in the registers allows for such aggregates to be generated.

1.47. It is possible the effect that countries with larger populations may have greater difficulties when making the transition to a register-based census, since the volume of information that must be analyzed and integrated is higher. On the other hand, it is also possible that larger countries, instead of having a single register at national level, have several registers at regional level, which would increase the number of data sources to be processed and complicate the transition to a register-based census.

1. Necessary pre-conditions for the use of administrative registers for censuses

1.48. Before administrative data can be effectively utilized for censuses, several key conditions must be met. These prerequisites are crucial to ensure the successful integration of administrative data into censuses, ultimately leading to more accurate, efficient, and cost-effective population enumeration processes. A well-established central population register, or systems of registers, must exist with the infrastructure to facilitate access to the data it houses. The data within the registers must be of high quality, ensuring its accuracy and reliability for statistical analysis, as well as have comprehensive geographic coverage, encompassing the entire population of the country. The registers must be updated continuously to ensure they reflect the most recent demographic changes. Furthermore, a consistent set of concepts and definitions must be utilized across all administrative registers to facilitate data integration and analysis .

1.49. Some of the essential preconditions that a country should have to conduct a register-based census are:

a) Legal framework

1.50. If administrative data are used for census purposes, statistical authorities should have a clear legal mandate to collect administrative data for statistical purposes. Individual historical, cultural and political factors of each country lead to highly diverse legal frameworks.

Data access
A legal basis should enable the statistical authority to collect administrative data. The required data sources should be described clearly. Data supply by governmental or private organizations should be specified as compulsory. Limitations to the data access (for example duration of access, confidentiality) should be described.

Privacy, integrity and security
To secure the handling of data and strengthen the trust from the general public, some legal acts should be in place. Examples of these are statistical act, privacy act and data act. These regulations should regulate how data can be transferred, handled and delivered inside the statistical institute and between the institute and other departments, organizations and users.

Data use
All variables of census relevance, with metadata, including identifiers of administrative data sources, should be listed completely and described clearly. Limitations to the data use (for example duration of use, deletion of microdata) should be described. Furthermore, it should be clearly defined that data compiled for statistical purposes will not be retransmitted to the data-providing organization or other governmental authorities.

b) Cooperation

1.51. A joint effort towards register-based statistics production requires firm and explicit commitment at the highest possible political level, as well as close collaboration among relevant authorities. Cooperation between statistical and administrative authorities generates a mutual and deeper understanding of the primary purpose of the registers and the needs of the statistical authority.

c) Confidentiality and public approval

1.52. In the context of a census, the most important principle for the population is the confidential use of individual information, as stated in the sixth Fundamental Principle of Official Statistics, which requires that the use by statistical authorities of individual data, whether they refer to natural or legal persons, be strictly confidential and used exclusively for statistical purposes. [add reference]

1.53. The political decision concerning the use of administrative data in a census can be highly influenced by public approval or refusal. In the run-up to implementing a new or modified census methodology it is helpful to inform the public about the project. It can be expected that people will become increasingly sensitive towards the collection and analysis of personal data by governmental authorities. In addition to outlining the general benefits and risks of the use of administrative data, information to the public should focus on the confidentiality of personal microdata. Clear limits and rules regarding the use of administrative data provide a common understanding that individual data collected for statistical purposes will not be passed on to other governmental authorities.

d) Administrative routines

1.54. The decision about the use of register data for statistical purposes largely depends on the nature of the register itself, including qualities such as integrity, reliability and lifespan. For this reason, there should be confidence in the administrative authority's capacity to be a reliable partner and data supplier. This implies the presence of administrative routines and safeguards. Does the administrative authority have extensive experience with the collection of the data that may be used for statistical purposes? Is the administrative authority well organized and is it anticipated that the necessary data collection will continue into the future? Are there existing quality guidelines for the administrative authority that guarantee long-term data quality? These are some of the questions that need to be elaborated upon in terms of assessing the feasibility of exploiting administrative registers for statistical purposes in general, and for the purposes of generating census statistics in particular.

e) Identifiers

1.55. Regardless of the census methodology adopted, it is extremely important that a unique primary key variable is used in all the data sources. The use of a unique identifier is essential in order to link information successfully. This primary key may already exist in the country – for example, a national personal identification number. Where it does not exist, or exists but with poor quality (for example, too many duplicates), it can be artificially created for statistical purposes. A statistical linkage key can be built from unchanging variables for persons, such as "family name at birth", "first name", "date of birth", "sex" and "place of birth". Care needs to be taken with alternative spellings, for example, incorrect or incomplete registry entries, transcription errors, and the varying transcription of foreign languages, names or place names. In any case, linkage between different registers can be made both by deterministic criteria (complete equality of one or more variables) or by probabilistic criteria (high similarity of one or more variables).

1.56. Unique identifiers assist in the detection (and correction as necessary) of identical statistical units (duplicates). Duplicate records most often arise when collecting data from more than one decentralized register, but are also possible within one centralized register. The problem with duplicate data entries is the risk of multiple counting of identical statistical units.

1.57. In the case of fully register-based censuses, information from different registers should be matched using good-quality identifiers. It is important to define how often information from different sources is updated and the reference date of the information stored in the different registers. When two or more data deliveries with the same content from the same administrative authority are planned, a linkage key will enable validation of data quality with regard to the statistical reference period.

1.58. Successful data linkage may be compromised by poor quality of the source data. Information stored or provided by data owners may have errors resulting in non-linkage or multi-linkage of records. In these situations, probabilistic approaches that choose the closest candidate, or the use of geographic information (starting from the lowest detail level and gradually increasing), can help to improve the linkage process.

2. Types of registers

1.59.Administrative registers are produced on the basis of administrative processes to collect information on units (persons and housing units) and variables that are defined by administrative rules and demands in a country. Although the content and process of registers would differ from one country to another, the types of registers are usually very similar. The following provides definitions of the main concepts used in the system of administrative registers.[25]

a) A *register* is defined as systematic collection of unit-level data organized in such a way that *updating* is possible. Updating is the processing of identifiable information with the purpose of establishing, updating, correcting or extending the register.

b) *Administrative registers* are registers primarily used in an administrative information system. This means that the registers are used in the production of goods and services in public or private institutions or companies, or that the information is a result of such production. Administrative registers used for statistical purposes are normally operated by the State or jointly by local authorities, but registers operated by private organizations are also used.

c) *Base registers* are kept as a basic resource for public administration. The function is to keep stock of the population and to maintain identification information. *Statistical base registers* are based on the corresponding administrative registers. Their principal tasks are to define important populations and contain links to other base registers.

d) *Specialized registers* are registers that, unlike base registers, serve one specific purpose or a clearly defined group of purposes. Specialized registers often receive information on the population and some basic data from a base register, but supply other data themselves (such as the register of vehicles).

[25] *Register-based statistics in the Nordic countries: Review of best practices with focus on population and social statistics*, United Nations publication, Sales Number E.07.II.E.11, United Nations, Geneva, 2007.

e) *Statistical registers* are created by processing data from administrative registers for statistical purposes. A statistical register could be based on one or several administrative registers. Statistical registers are also referred to as *secondary registers*.

1.60. The process involves collecting information on the characteristics of individuals, which has been provided to an administrative register for non-statistical purposes. To be effective, access to administrative data for statistical purposes must be given by law or by agreement, providing the capability to (a) transfer the data as individual records to the statistical database; or (b) temporarily link the registers to form a proxy register for statistical purposes.

1.61. Administrative registers are maintained primarily for administrative purposes. Units and variables of administrative data are described according to administrative rules and demands. Before a register is used for census purposes, the suitability of its data in terms of definitions, concepts, content, reference date, accuracy and other criteria should be statistically tested by comparing them with previous census and survey results, and conducting quality and compatibility surveys. A pilot census may be used for this purpose.

1.62. All persons within the defined territory who meet the register's rules are enumerated. The timing of the census extraction may require careful thought where register update cycles vary. Registration delays and administrative delays in updating between regional and national databases can otherwise have a serious impact on the quality of the output.

(i) Base registers

1.63. Base registers are those registers that hold the basic information relating to the stock of the entire set of 'population' units that are being counted in the census – typically persons, dwellings and buildings. Base registers are created and then linked at the unit record level to the information taken from specialised registers from which one or more of the characteristics that may be required for the census can be derived in order to create the requisite census attribute variables of the population and dwellings from which the census statistics are then produced.

1.64. The base registers that are most commonly used for the purpose of the census are:

(i) **Population register** – records all residents in the country, together with their basic demographic information, and updates for births, deaths, changes of marital status, international migration (both into and out of the country), and internal migration (by recording changes of address within the country).

(ii) **Building and/or dwelling register** - usually held by land and property valuation agencies or by local or municipal authorities responsible for the development of housing policies and urban planning. They can provide information on the location of all places where people might live (by distinguishing between wholly residential, partially residential and non-residential buildings) as well as some of the core census housing topics such as ownership, type of housing (for example whether a detached house, terrace house, or apartment), whether or not the housing is rented, living area, number of rooms, condition concerning availability of electricity, bathroom, kitchen, and toilet, and year of construction. With respect to a register-based population census such registers are of fundamental importance in the harmonisation

with the population register, in that every person that is allocated a dwelling number can be associated with a particular housing unit, and (in some cases) a distinct household.

(iii) **Establishment (or business) register** - usually contains information about business establishments (including their industry, institutional sector, size of workforce, and location), used in census tables.

(iv) **Address register** – records all addresses, by location, and identifies whether the address is used for accommodation, commercial (business) or office purposes. The address register may also include links to the building (such as an apartment building) and/or the dwelling.

1.65. Without these base registers, it is very difficult to conduct a fully register-based census. However, even without a population register, it may still be possible to conduct a combined census, using data from other registers and information collected in a field enumeration.

(ii) Specialized registers

1.66. In addition to the base registers described above, there will be other specialised (or supplementary/subject-matter) administrative registers available to the NSO that contain records relating to those characteristics for which census information is required and from which the requisite variables for core and non-core topics can be derived using the PIN for record linkage. Such registers may include the following:

a) Social security or pension registers
b) Tax registers
c) Employment, unemployment and jobseeker registers
d) Education and student registers
e) Health registers
f) Border control
g) Other supplementary registers

1.67. NSOs may also be able to access data from a range of other administrative sources either to provide topic-related census variables or to identify particular population sub-groups that are traditionally hard to count accurately (and for whom full information may thus not be recorded in the population register), such as the institutional population. For example (and this is by no means a complete list): electoral registers; registers of motor vehicles; registers of foreign nationals or of residence permits; lists of military service personnel; registers of persons living in institutional places; registers of building permits; farm registers; and, registers held by public facility service providers (such as for piped water supply, electricity and/or piped gas, and sewage and waste disposal).

1.68. Specialized registers may help to identify issues of coverage by, in particular, identifying 'signs of life' activities. The basis of this methodology is that over time, a person living in a country inevitably leaves certain traces or markers of administrative activity in the form of records in different registers. This means it is possible to verify the person's residence in the country, as well as connections between persons and their locations, on an annual basis. Such verification is based on signs of life, signs of partnership that are recorded in registers every year. For example: persons in the age 0-3 years old can be captured in health registers. School age population 4-17 years old can be captured in education enrolment registers. The remaining population may be captured in labour force and other registers.

3. Establishing a statistical population register

1.69. The statistical population register is a systematized and indexed collection of individual records for every resident (including nationals and foreign citizens) of the country. Where developed and functioning, it represents the backbone of the compilation of official statistics and represents a master population frame as well for designing and running statistical surveys. Connected on a regular basis and under the provisions of the law with administrative registers, starting with the administrative population register, it should be regularly updated with pertinent information. The legal framework for establishing, maintaining and exploiting the national statistical population register must ensure that it is used solely for compiling aggregate statistics and that it cannot be accessed for any other purposes either within or outside of the national statistical authorities except for approved research.

1.70. Creating a Statistical Population Register (SPR) is a crucial step towards achieving the use of administrative data for census purposes. It should be noted that it is important to distinguish between the registration of an individual unit (person) in the population register and the inclusion of that record in the computation of internationally comparable statistics on the stock of population and its basic characteristics, such as age and sex. While the listing of those persons in the register will be in accordance with national legislation for the requisite administrative purpose, for statistical purposes only those persons meeting the requirements of being residents of the country should be included. For instance, persons who have left the country but are still holding a permanent residence permit should be still listed in the population register, but they should be excluded from the census population of usual residents. Particular care should then be applied to properly classifying registered persons who have emigrated, especially in those countries where there is no incentive to declare the emigration to the authorities.

1.71. An SPR is usually generated from the existing administrative population register by the NSO. In countries that have a reliable population register (for both nationals and foreigners), a statistical population register is easily generated and updated from the existing population register. For example, in Austria, the local population registers were replaced by the Central Register of Residents (ZMR) maintained by the Ministry of Interior. On the basis of the ZMR, Statistics Austria set up a separate database, the Population Register (POPREG) which has been the main data source for producing statistics on the resident population and for producing migration statistics since then. Other registers established or in the process of being established include the register of enrolled pupils and students, the register of educational attainment, the housing register of dwelling and buildings, and the business register and their local units of employment34

1.72. On the other hand, if there is no centralized population register, a statistical population register can be constructed by integrating local population registers and linking them with other existing registers.

1.73. Thus, establishing a national statistical population register may involve some investment in collecting records from different existing administrative registers and/or administrative data sources, harmonizing them, linking information from different registers and running complex editing procedures to ensure the consistency and quality of individual records. Once established, and under the protocols spelled out in legal provisions, the national statistical population register is updated from administrative registers on a regular basis.

4. Constructing integrated statistical registers

1.74. A statistical register contains the data integrated from different sources including surveys and administrative data. The statistical registers are usually housed in the NSOs. The statistical agency usually can transform the data from administrative sources into the statistical data by editing, correcting the incorrect entries, remove inconsistencies, removing duplicates and allocating new statistical identifications. Among the statistical registers, the population register is the backbone for the register based statistical system.

1.75. A statistical register typically plays the role of a data coordination tool, integrating data from several sources, both statistical and administrative. This may be done by linking records using common identifiers or by using matching techniques (as discussed in section 4.5 below). It may sometimes be easier to use data from a single source, but in such cases it is often difficult to check the accuracy of that source. When several sources are used and integrated within a statistical register it is possible to have a much better view of the accuracy of the data. Unfortunately, the negative side of this is that it becomes necessary to have a strategy for dealing with conflicting data from different sources. However, if variables in statistical registers are stored with source codes and dates, automated algorithms can be used to prioritise sources and resolve most data conflicts.

1.76. As well as integrating data from different sources, a statistical register may also provide the possibility of deriving new (census-related) variables not otherwise available from any one single register (such as industry of occupation, level of overcrowding or a housing quality index.)

1.77. Traditionally, statistical registers have been used as sampling frames for surveys, but as has been already noted, a statistical population register is at the very heart of a register-based census. A population register may be established inside a statistical institute for statistical purposes. However, it is more common (and far simpler) for the statistical agency to transform the administrative register into a statistical register owned and managed by the NSO, if permitted by law and the NSO has the resources to do it. This transformation process is usually a relatively simple task and may include editing, correcting, removal of duplicates, and translating the administrative ID into a statistical ID. 4-15 The NSO may also then integrate data that the administrative register does not have with authenticated evidence of data from other sources (such as, for example, country of birth for people immigrating before the administrative register was established). The statistical population register should be updated regularly. In Norway, for example, it is updated five days a week with new transactions from the administrative register.

1.78. In addition to identifying the data sources to be used for any particular statistical operation, some of the key processes involved in the construction of a statistical register are:

- data quality management;
- linking the unit records from these sources;
- dealing with duplications;
- resolution of conflicting information relating to the same data items in different linked sources;
- updating; and
- editing and imputation.

5. Transition from traditional censuses to register-based approaches

1.79. Moving forward to register-based census from the traditional census has certain advantages like reduced cost in the long run, improvement in the quality of data and ability to produce more frequent information at lower geographic level.

1.80. Country practices indicate that changing census methodology from a traditional census to a wholly register-based census usually takes several decades. There is no ideal method or duration for this transition and therefore each country should plan this process based on the initial assessment of the availability of administrative registers as well as assessment of preconditions that have to be met before changing census methodology. It should be noted that direct transition from the traditional census to wholly register-based census is not realistic, and that this transition should be planned gradually, introducing more variables each time, providing that the registers have been proven to be of good quality.

1.81. If a country decides that it wants to move to a combined or register-based census, there are certain conditions that must be met before information from administrative registers (and other sources) can successfully be integrated to create the underlying statistical register.

1.82. Such a move is usually done in several stages, often by first adopting one form of a combined census approach, such as registers with full-field enumeration or registers with ad-hoc surveys, before moving to a register-based census. In this case, dependency on administrative data sources can be gradually increased and the cost of the change can then be spread over two or three census cycles. At the beginning of this process, key registers such as a population register might be incomplete or of insufficient quality to be used as source for the census. In such cases, it should be made clear from the beginning that a country should not attempt a move immediately and should continue conducting some form of traditional census. However, even when countries continue to do so, innovations making greater use of administrative data could help the NSO work more efficiently. It is beneficial if the relevant public authorities make administrative data sources available to the statistical institutes to produce proto-type register-based statistics. The government itself can help a great deal both by removing any legal barriers to data sharing and by subsidising the transition.

1.83. It should be noted that there is no unique way of moving to register-based approaches; actually national practices show that there are many different ways of designing this transition, especially for countries that are planning to use administrative data sources for producing census information for the first time. Therefore, this handbook provides only a general framework for moving to a combined or register-based censuses. Four key phases of the transition include:

(i) **Establishing/maintaining necessary conditions**: For a successful transition from a traditional to a combined or register-based census, it is necessary to establish legal framework for being able to use individual data for statistical purposes. Institutional capacity and establishing effective and continuous communication with all stakeholders are also very critical.

(ii) **Identifying administrative data sources and other sources**: This phase needs to be carefully planned and managed in collaboration with register-holders and users. NSO should build up a wide-ranging knowledge of the existing administrative data sources before making a decision on changing the census methodology.

(iii) **Transforming administrative data into statistical registers**. This phase includes a series of activities that are implemented for transforming administrative data into statistical data from which the census outputs can be produced. The design of this phase will differ from one country

to another depending on the quality of existing population register and other administrative registers. Countries with good quality of population register can more easily construct statistical population register. However, countries that do not have reliable population register, may have to conduct a full-field enumeration or ad-hoc surveys in order to be able to construct the statistical population register.

(iv) **Quality management and assessment**: Quality management and assessment is an overarching process implemented by NSOs throughout whole phases of this transition. For a successful design of this phase, it is important to develop a system for continuously monitoring the quality of activities and assessing the quality of each phase so as to be able to understand the effects of the quality on the next phase and, finally, on the census outputs.

6. Types of register-based censuses

(i) Fully register-based census

1.84. A census is said to be fully register-based when information on all variables is obtained from different registers. This option allows the countries that have it implemented, to obtain information from the different census variables as frequently as they deem appropriate. Due to the advantages that this alternative provides, many countries have the medium or long-term objective of carrying out a census using this methodology.

1.85. For a successful registry-based census, it must be verified that each and every one of the census variables can be obtained from registers with enough quality. Therefore, a detailed study must be made of the information provided by these sources and contrasted with other results to verify its reliability. On the other hand, its level of coverage must also be checked, which covers the entire population. Additionally, the updating of information must be correctly reflected in the registers.

(ii) Register with existing sample surveys

1.86. A special case of register-based census is when population and/or integrated administrative registers are combined with existing sample survey(s). Different data sources are integrated as part of a "virtual census"[26] process. The data for the census exercise are derived from many types of registers and surveys, covering different population groups throughout the country and its subregions. Compared to conventional census methods, this process is lower in cost and staff requirements, and uses more frequently updated data sources. More significantly, by combining data that are already available from other sources, the virtual census makes no additional respondent burden, increasing its public acceptability.

1.87. A weakness of the use of sample data within the census exercise is that for the variables derived from the sample surveys the sample size may not be sufficient to offer the geographically detailed outputs, or thematically disaggregated information with multiple variables, that are an essential feature of the census. Importantly though, the use of data from existing regular surveys will often mean that time series of data are available. By combining samples for several survey exercises, it may be possible to produce reliable estimates for small geographic areas (although there would be the problem that the reference date would not be homogeneous for all the individuals).

[26] The term "virtual census" originated in the Netherlands, where it was first developed, and here it refers to that specific methodology.

1.88. However, the use of existing surveys does present a number of problems compared to the use of ad hoc surveys. The timing, content, statistical definitions and sampling approaches used in an existing survey may not be appropriate to allow the data to be readily combined with data from the administrative source(s). For example, many major household surveys are not designed to cover persons living in institutional households (such as student accommodation, hospitals, prisons and military establishments), meaning that an additional source of information is needed for these persons.

7. Advantages and challenges

1.89. The primary advantages of a register-based approach are reduced costs and greater frequency of data. and eliminating respondent burden. The reduced cost of register-based census refers to the long-term costs involved in future censuses. There is always a considerable cost involved in setting up systems for moving from traditional census to the registered-based census. The availability of up to date and detailed regional statistics every year instead of every 5/10 years is another advantage of using this approach.

1.90. Another significant advantage may be the fact that by having information based on registers for many variables, the information contained in the surveys can be of a more specific nature.

1.91. Certain potential drawbacks with the use of administrative data sources also need to be taken into account. One limitation is that the scope of statistical topics, key definitions and, indeed, the population base of the exercise depend on the information that can be compiled from the available registers. These, in turn, will be based on the underlying administrative purpose and procedures of the registers. In addition, it is common for national legislation to restrict or prevent the use of administrative registers for other purposes, sometimes including the production of statistics. This may impose restrictions with respect to the characteristics that are available for description, and may also undermine international comparability. When a registered data item is changed, new or updated information is not always registered immediately. In certain cases, new or updated information may not be registered at all. Where this occurs, the register information does not accurately reflect real circumstances. Also, the choice of new and emerging topics is limited in registered-based approach as compared to traditional census questionnaire which has lot more flexibility. This also applies to several non-core topics that can be difficult to derive using administrative sources only.

1.92. Furthermore, when using different data sources based on registers, not all of them have the same reference date: there are those that are updated continuously, weekly, monthly, etc. This can be a problem, because in theory, all the Census variables should refer to the same reference date.

1.93. Depending on the amount of information available from different data sources and its degree of integration, data processing can be more complex with these census methods than with a traditional census, although good-quality results can be obtained.

1.94. The decision to use administrative sources in the statistical production process requires close collaboration between the administrative authorities and the national statistical offices. During preparations for data delivery, all parties concerned must agree on date(s) of delivery and the content of the data. This implies a bilateral agreement at a high hierarchical level on a detailed data set description, scheduled delivery dates and the statistical reference period. Test data deliveries help to solve or minimize problems with the subsequent data processing by the national statistical authority. Validation techniques

appropriate to administrative data should be applied, including checks on the plausibility, completeness and reference periods.

1.95. Register information may contain errors (for example, records showing people as being implausibly old, invalid occupations, information about migration that is not consistent with other data). Edit rules may be defined to highlight inconsistent or implausible information. If a data source based in registers has errors in a significant percentage of cases, then that data source should be discarded. Correction or imputation of records with errors can be attempted in different ways: first, if possible, using another data source (register) that also has information about that specific record and topic; or second, carrying out probabilistic imputation based on available information that is considered to be reliable.

1.96. Missing or implausible data can create serious problems for data analysis. Cases that have missing or implausible values may be deleted but this can result in a loss of representativeness and completeness, and the introduction of bias. Various imputation methods can be considered, with a general distinction between single and multiple imputation techniques.

1.97. Sometimes information about topics can be obtained through different data sources (registers and survey). In this situation, it is very common that calibration techniques are used in order to reduce inconsistencies between data from different sources. However, calibration only guarantees coherence to a certain geographic level, generally modifying or adjusting the sampling factors. It may be necessary to explain to users the reasons for any (remaining) differences. Furthermore, for the calibration results to be satisfactory, the survey must be sufficiently representative.

C. Combined census methodologies

1.98. The combined census is a methodological approach whereby some information is obtained from administrative sources such as population registers, while other information on selected census topics are collected through field operations, either by using census forms that are completed by all units, in a way similar to the traditional approach (full field enumeration), or by means of ad hoc sample surveys.

1.99. The combined census is a frequently used as a steppingstone strategy of NSOs aiming to move from the traditional census to a fully register-based census. Compared to a register-based census, the combined census method is more expensive, and imposes more response burden on the public. On the other hand, it does provide better results and it may help to increase the coverage and quality of administrative registers.

1.100. In recent years, it has been observed in a number of countries that:

(a) The quality of the administrative registers is relatively good (at least for certain key census variables);

(b) Information for some census topics is not available in the administrative registers or the quality is not sufficiently high;

(c) The population generally, and certain population groups (in particular people difficult to enumerate), are becoming more sensitive to the handling of personal information, and possibly more reluctant to cooperate with the statistical office or more difficult to enumerate due to their high mobility or other reasons.

(d)The administrative registers are available but the information on some topics cannot be shared with NSOs due to lack of coordination and/or legal constraints.

1.101. In these cases, a combined census that uses register(s) and questionnaire(s) could be an option. Essentially, the combined methodology makes use of registers relevant to a census, complemented by surveys or complete enumeration. The use of survey and enumeration data is intended to:

(a) Improve the accuracy of the population counts;

(b) Provide information for census variables that cannot be reliably based on administrative data;

(c) Check, update and improve the quality of census data derived from administrative sources;

(d) Add additional variables to the census;

(e) Be a linking frame in order to bring together different sources.

(f) To pave a way forward towards a complete register-based census by analysing the accuracy of available multiple data sources. The combined census methodology can be considered as a pilot test for the complete register-based census. It provides the opportunity to utilize/test the accuracy, completeness of the available administrative data sources and hence plan for a complete register-based approach in the upcoming censuses.

1.102. In short, when a country is in a state of transition towards a register-based census due to having a limited set of records with partial information, the combined census may be a good option.

1.103. Information on individuals, households and dwellings is collected by combining data from registers with data collected from one or more surveys. Data collection may be based on full field enumeration, an ad hoc sample and rolling survey methods. Data from registers are employed not only as a sampling frame or to support field operations, but also directly as the data source for some census information. In a case where registers are used along with total enumeration, data from registers may be prefilled in questionnaires, and respondents may be asked to check, update and confirm their details. Other questions relating to fields not available in the registers may also be canvassed during this exercise. In the case of new individuals, households or dwellings that do not feature in the register, all fields of information that are required for the register and the census are to be canvassed afresh. When registers are used along with sample surveys (ad hoc sample or rolling surveys), some census tables may be produced entirely from the information available on the register(s), while for other census tables, information from the survey(s), duly weighted to the population totals, could be used.

1.104. It is worth clarifying that the results obtained from the registers will not always be consistent with the results obtained from the sample surveys, so it would be desirable for both data sources to be calibrated so that the information produced is as coherent as possible. The surveys would also serve to evaluate the accuracy of the register counts.

1.105. The option of a register-based census with sample surveys can be adopted only if all necessary census information is available from the various administrative or survey sources, and it is possible to link the information from the different sources at the record level. During the process of integrating individual records, care should be taken to check the accuracy of the data and remove inconsistencies prior to the production of statistical outputs.

1.106. The data sources would include verified and accurate personal information (name, ID number, date of birth, sex, marital status**,** family structure, etc.) and a dwelling register. In an ideal situation, a "base" register can be envisaged, to include unified identity codes for both people and address components in order to link more efficiently the related register and survey data. The link between

persons and their dwellings is equally important, giving the household unit. Other administrative data sources include tax files, social security files, public records of unemployed and registers of educational qualifications. It is preferable to have a centralized base register. If this is not available, regional registers will need to be consolidated. The data sharing protocols/procedures be in place among the holders of administrative registers and the NSOs. These procedures should be placed under legal framework and ensure the confidentiality of individual records. The census reference period should be aligned with the regular updating cycle of the registers that has to be used for producing census outputs.

1.107. The different data sources of a combined census may have overlaps in some variables and the information contained may be different in certain cases. If this happens, it should be prioritized which data source is used in that situation.

1. Types of combined census methodologies

1.108. There are many different approaches to the implementation of combined census methodologies. These differences can be categorized according to whether a non-continuous or continuous approach is adopted, and the ways that and extent to which the different data sources are used. Methods used by countries using either a fully register-based census approach or one based partly on administrative sources are examined in detail in a research report.[27] This report presents information for seven fully register-based and three partly register-based countries, giving information on the methods used to combine register and survey data, and to compensate for missing information. The use of repeated weighting, register estimation methods, calibration of data, and microdata or macrodata integration are discussed.

(i) Base register[28] and/or integrated administrative sources[29] with full field enumeration

1.109. One approach can be to combine the full enumeration with a base register or registers.[30] The questionnaire used in the total enumeration then contains fewer questions compared to a traditional census questionnaire, but still covers the whole population of individuals, households and dwellings. Over time, countries may decide to adopt this model, increasing in successive censuses the use of integrated administrative registers and reducing the number of questions in the questionnaire. In any case, to carry out a Census that is as efficient as possible, it must be taken into account that the variables that can be obtained with sufficient quality from the administrative registers should not be part of the traditional census questionnaire. The registers can be used to prefill such information as name, address, family composition, education, occupation and dwelling characteristics on the questionnaire. These prefills can then be used to ask the respondents if the information is still valid or needs to be changed. That can greatly reduce the work involved in coding of the census questionnaire.

---

[27] Schulte Nordholt, Eric. *Efficiency in population censuses: the situation of the European register-based 2011 Censuses*, available at: http://www.statistics.gov.hk/wsc/STS067-P3-S.pdf.

[28] Examples of base registers could be population register and/or dwelling register; and address register and/or business register.

[29] Examples of integrated administrative sources could be administrative or statistical registers of education and/or occupation.

[30] In Europe in the 2010 round of censuses, four countries applied this approach: Estonia, Latvia, Liechtenstein and Lithuania. For more details, see *Measuring population and housing: practices of UNECE countries in the 2010 round of censuses,* United Nations Economic Commission for Europe, United Nations, New York and Geneva, 2014, page 12.

1.110. The main advantage with using this model is that it will reduce the response burden for respondents and reduce the cost of the census (if we compare it with a 100% traditional census). The model also allows the preparation of small-area statistics as all variables are collected as total counts. However, this model will still involve a large data collection exercise with the use of enumerators. Mail out/mail back data collection may sometimes be used, but a significant proportion of respondents may require enumerator follow-up and assistance in completing the questionnaire. This model will continue to require great planning in terms of data collection as it requires the response of the entire population.

(ii) Base register and/or integrated administrative sources with ad hoc sample survey(s)

Another model involves the use of an ad hoc sample survey instead of a full enumeration.[31] The backbone register could then be used as a sample frame and also to prefill some information such as name and address on the questionnaire. The ad hoc sample survey questionnaire can be specifically designed to complete and statistically correct the data coming from registers, covering those variables not available from the register. The sample can be sized and stratified in such a way that data are available for small groups and geographic areas.

1.111. This model requires far fewer enumerators than a full enumeration, so a more specific training operation with skilled and prepared professional interviewers can be carried out. The follow-up of the operation is also simpler. There is no need to obtain information from each member of the population, giving a clear reduction in response burden. Non-response can be corrected in the sample by the use of statistical techniques to ensure information is still representative of the population. In any case, it is important to pay attention to ensuring that all difficult-to-count groups from the Census point of view are also correctly represented in the survey responses.

1.112. By using a sample and a much smaller number of enumerators, a significant reduction in the cost of the census is possible, in particular when compared with the traditional approach.

1.113. The larger the size of the survey, the greater the geographical and thematic detail of the information that can be produced, but the higher the costs of the operation will also be. The extent to which this model can produce detailed statistics (in particular for municipalities or smaller geographic areas) will depend on the size of the sample survey. A larger sample should allow more detailed statistics to be produced but this will correspondingly increase the financial cost of the exercise. Even with a large sample, the availability of detailed information is likely to be lower than if a comprehensive approach were taken – such as with a traditional census or fully register-based exercise.

2. Advantages and challenges

1.114. The register-based census combined with ad-hoc surveys(s) offers several advantages:

[31] In the 2010 round of censuses, countries that applied this approach included Germany, Israel, Poland, Spain, Switzerland and Turkey. For more details, see *Measuring population and housing: practices of UNECE countries in the 2010 round of censuses,* United Nations Economic Commission for Europe, United Nations, New York and Geneva, 2014, page 12.

(a) It can be much cheaper than a traditional census with a full enumeration collecting all census items from the whole population;
(b) It will reduce the burden of enumerators and respondents;
(c) It will reduce non-response in case information is obtained from registers;
(d) It should be possible to correct the survey data for differing levels of non-response in different
(e) The combined census approach by having field enumeration using an ad-hoc sample survey provide greater flexibility on introducing questions on new/emerging topics (like COVID for example).
population groups.
(e) The planning of this operation will be simpler than in a Census that involves the entire population

1.115. Micro-integrated data might be expected to provide very reliable results, because they are based on a maximum amount of information. The coverage of subpopulations may be more reliable because when data are missing in one source, another source may be used. Another advantage of micro-integration is that there will be less reason for confusion among statistics users. For example, there will be one figure on each socioeconomic phenomenon, instead of several different figures depending on which sources have been used.

1.116. A disadvantage is that it involves more work to produce the tables from the sample survey microdata, as weighting problems may arise. As the combined census may lack the high public profile and publicity of a traditional census, there may be less interest in and use of the census results, as there is no longer a single census event to attract public attention. Other potential disadvantages may be a lack of transparency (no one external to the process may be able to reproduce the information) and data quality. Finally, this type of Census cannot have the same level of thematic and geographic detail as a traditional Census and this may require giving explanations to the different census users. Another disadvantage arises when the register data is combined with the pre-existing sample surveys. The timing of the existing survey can be different than the census reference period, also the objectives, target population and sampling approaches, which makes it difficult to combine the data with administrative registers in an efficient way.
1.117.

1.118. Data validation, processing and dissemination may be more complicated, as this approach involves both total counts based on the register and sample data from surveys. In addition, as some variables are based only on sample data, it may be impossible to meet the level of statistical and geographic detail required in some tables. It is advisable to be very didactic with census users: explain to them in a transparent way how the census results are obtained and the limitations they may have. On the other hand, the possibilities of reducing cost and response burden provide a very strong reason to adopt this approach.

1.119. Some of the required variables will need to be constructed from different sources. The census results obtained may differ to some extent from those that would be obtained from a full enumeration covering all census topics. This may have a negative impact on the comparability of results between countries and over time. An advantage of registers is that, in effect, they offer complete coverage subject to the quality of the data contained (see also paragraph 1.93). It is preferable that statistical authorities make full use of the register data that are available.

1.120. For the combined census method, a number of different options can be used to collect information, including paper, Internet, handheld devices and telephone interviews and emails. Electronic devices have important advantages that influence the quality of the information obtained: validation controls can be included in the different questions, time to answer the questions is reduced and the

analysis and dissemination of information is faster. Where data are collected via different routes (such as Internet collection in parallel with face-to-face interviews), controls are needed to avoid duplication of information.

1.121. For combined census methodologies, it is important to store control information and indicators at the lowest geographic level available in the central database, covering issues related to the census operation, such as progress with the fieldwork, response rates and comparisons with information in registers. This control information – normally based on web reports and analysed daily by project managers or regional offices – can be used to detect problems that appear during the fieldwork and to plan necessary actions to overcome these problems. If data are analysed on a daily basis, close monitoring of the continuing field operation and data entry is possible.

V. Methods for estimating population for areas not enumerated in the census

1.122. While conducting a census, some countries face security or access challenges in certain parts of their territory, with no possibility to send enumerators there. In that case, it is possible to estimate the population numbers and densities of inaccessible areas, using various approaches.

1.123. It is important to note that all estimation methods have some degree of error. Therefore, it is important to carefully consider the accuracy requirements when choosing an estimation method.

1.124. In addition to the methods described above, there are a number of new and innovative methods that are being developed to estimate the population of areas where enumeration is not possible. For example, researchers are developing methods to use mobile positioning data to estimate the population. As technology continues to develop, we can expect to see even more accurate and efficient methods for estimating the population of areas where enumeration is not possible.

A. Requirements for estimating population

1.125. Population estimates generally start with a base population, frequently obtained from a population census. A post-enumeration survey provides a means of measuring under- and over-coverage of the census. This approach utilizes a dual-capture framework, where the census provides one capture of the population, which is adjusted using the results of a second, independent capture of the population in the post-enumeration survey. This system works well but relies on high response rates to both the census and the post-enumeration survey to ensure that correlation bias is minimised.

1.126. From this population base, births are added, deaths are subtracted, and adjustments are made for internal and overseas migration to create the population estimate for the next time point. Vital registers provide information on births and deaths, or where registration processes are not well established, fertility and mortality rates can be used. Migration data may be more difficult to obtain, particularly internal migration, but administrative data sources can provide measures of internal migration.

1.127. As population estimates move further away from the base population time point, errors from each of the input data sources accumulate and the population estimates become less accurate.

B. Methodological approaches

1. Use of administrative data sources

1.128. Administrative data can be used to estimate the population not enumerated in a census in a number of ways. Administrative records can be used to estimate the population of areas where enumeration is not possible by using data such as birth records, death records, and school enrollment records.

1.129. It is important to note that administrative data of not sufficient coverage may not be a perfect substitute for a census. Administrative data sets may not be complete or accurate, and they may not cover all members of the population. However, administrative data can be used to improve the accuracy of population estimates, and to identify populations that are at risk of being undercounted in the census. Despite these challenges, administrative data can be a valuable tool for estimating the population not enumerated in a census. By carefully considering the strengths and weaknesses of administrative data, researchers can develop accurate and reliable estimates of the total population.

2. Use of sample surveys

[More content to be added]

3. Other methods

1.130. Within the 2030 Agenda, the growing requirement for spatially disaggregated population data has triggered the exploration of new data sources at different geographical scales, especially in highly stressed countries and countries without a recent population census. Advances in the availability of detailed satellite imagery, geo-positioning tools for field surveys, statistical methods and computational power are enabling the development and application of modelized approaches that can estimate population distributions at fine spatial scales across entire countries, in cases where population and housing censuses cannot be conducted.

1.131. In contexts where a population and housing census cannot be fully executed in all locations of a given territory, for instance because of security concerns, a hybrid approach can be adopted to produce population estimates for small areas or for uniform and detailed grids. Such an approach relies on complete population counts within small, defined areas, through a partial census, a census-like population survey or specifically designed “micro-census surveys” undertaken where it is possible to conduct field work. Statistical models can then be used to link these micro-census data to spatial data with full coverage over the regions of interest to predict population numbers in locations where no survey was conducted. By aggregating these high-resolution predictions, population totals can also be produced for administrative units or for the national level if required.

1.132. The success of this approach relies on the identification of relevant geospatial covariates and the definition of functional relationships between geospatial datasets and population density to predict population numbers in un-surveyed locations. It also requires robust geo-referencing of the geographical areas where population data are collected, and a design that captures as best as possible the range of population densities, demographics and environmental dimensions that exist across the regions of interest. The covariates need to be strongly correlated to population density and available consistently

across all areas where population estimation is required. Although access to high-quality and spatially comprehensive datasets has traditionally been difficult in resource-poor settings, the increasing availability of very high-resolution satellite imagery makes the production of high-quality covariates increasingly feasible for many settings.

1.133. Although the production of such modelled population estimates can never replace the rich production of data on the individual, family, household or community generated by a full-scale population and housing census, their high level of geographical disaggregation makes them useful in contexts where no other population counts are available.

VI. Strategies for producing more frequent population and housing statistics

1.134. There are a number of alternative approaches that could be used for producing more frequent population statistics such as intercensal or annual surveys for countries conducting a census with the traditional method. Some of these approaches include: Sample surveys; Administrative records; Geospatial data; and Combined approaches. A combination of administrative records and sample surveys can be used to produce more frequent population statistics. The best approach for producing more frequent population statistics will vary depending on the specific needs of the country or region. However, all of the approaches discussed above have the potential to improve the timeliness and accuracy of population statistics. In addition to the above approaches, countries are also increasingly exploring the use of new technologies, and alternative data sources, to produce population statistics. These sources can revolutionize the way population statistics are collected and produced, and make population statistics more frequent, accurate, and accessible to everyone.

A. The need for more frequent data on population and housing characteristics

1.135. There are a number of reasons why there is a need for more frequent data on population and housing characteristics. Some of these reasons include:

a) **Planning government programs and services**: Governments need timely and accurate data on population and housing characteristics in order to plan and deliver effective programs and services. For example, data on population growth and distribution can be used to plan for new schools and hospitals and the training of their staff to meet future needs, while data on housing occupancy and affordability can be used to plan for new housing developments and rental assistance programs.

b) **Allocating resources**: Governments also need data on population and housing characteristics in order to allocate resources fairly and efficiently. For example, data on poverty rates can be used to target resources to low-income communities, while data on educational attainment can be used to allocate resources to schools in areas with high concentrations of students from low-income families.

c) **Conducting research**: Researchers need data on population and housing characteristics in order to study a wide range of topics, such as economic trends, social inequality, and public health. For example, data on income and employment can be used to study the impact of economic policies, while data on race and ethnicity can be used to study the impact of racial discrimination.

1.136. In addition to these general reasons, there are also a number of specific reasons why more frequent data on population and housing characteristics is needed. For example, the COVID-19 pandemic has highlighted the need for more timely data on population health and housing conditions. The pandemic has also shown how quickly population characteristics can change, making it more important than ever to collect data on a regular basis. More frequent data can be used to respond to unexpected events, such as natural disasters or economic crises. For example, data on population distribution can be used to target resources to areas that have been affected by a natural disaster. More frequent data can be used to identify emerging trends in population and housing characteristics. For example, data on housing affordability can be used to identify areas where housing prices are rising rapidly and people are at risk of displacement.

1.137. Overall, there is a strong need for more frequent data on population and housing characteristics, which are essential tool for making informed decisions about the future. This data is essential for planning and delivering government programs and services, allocating resources fairly and efficiently, and conducting research on a wide range of topics. By investing in more frequent data collection, governments and researchers can ensure that they have the information they need to make informed decisions that benefit everyone.

B. Use of administrative registers for producing more frequent data

[More content to be added]

C. Intercensal surveys

1.138. Intercensal surveys can be used to produce more frequent population and housing statistics by collecting data from a sample of the population between censuses.

1.139. Intercensal surveys have a number of advantages over traditional censuses. First, they are less expensive and time-consuming to conduct. Second, they can be used to collect data on a wider range of topics than traditional censuses. Third, they can be used to produce more timely data on the population.

1.140. However, intercensal surveys also have some disadvantages. First, they are subject to sampling error, which means that the estimates produced from the survey may not be as accurate as the estimates produced from a census. Second, intercensal surveys may not be able to reach all segments of the population, which can lead to bias in the results.

1.141. Despite these disadvantages, intercensal surveys can be a valuable tool for producing more frequent population and housing statistics. By using intercensal surveys, countries can produce more timely and accurate data on the population, which can be used to inform decision-making on a range of issues.

[More content to be added]

D. Annual continuous surveys

1.142. Annual continuous surveys can be used to produce more frequent population and housing statistics by collecting data from a sample of the population on a continuous basis.

[More content to be added]

PART TWO. PLANNING, ORGANIZATION AND MANAGEMENT

I.Introduction

2.1. Part Two of the *Principles and Recommendations* elaborates on planning, organization and management of the population and housing census. Regardless of whether taking a traditional or non-traditional approach, this is a peculiar and most complex statistical exercise requiring a multidimensional and simultaneous approach to preparation and management.

II.Overall census planning

A. Overall census planning

2.2. A population and housing census (or a population census by itself) consists of a complex series of interrelated activities, and constitutes perhaps the single most extensive, complicated, and expensive operation that a country undertakes. Some of these activities, for example the printing of the census questionnaires, may be massive in scale; other activities, for example the training of the supervisory staff, must be carried out in a uniform manner in all parts of the country; and still others, for example the actual enumeration, must incorporate both features. For administrative or combined censuses that incorporate administrative data and traditional enumeration, there is also extensive work involved with the intake and integration of administrative data sources. Also, since censuses take place after five to ten years, the planning and preparation for each new census round must account for changes in field conditions, census methodology, technological innovations, stakeholders' requirements, census questions, personnel and societal conditions as well as changing administrative availability.

2.3. To ensure that the diverse operations occur in their proper sequence and in a timely manner, the entire census and its various component steps must be planned for carefully in advance. An apparently minor oversight in planning may lead to serious defects in the census results and to costly inefficiencies in the census operations. Careful planning is therefore critically important to a successful census, not only in countries with comparatively little statistical experience but also in those with a well-developed system of statistics. Coupled with the need for careful planning is the need for appropriate organizational and administrative arrangements and procedures. Such arrangements and procedures are necessary to ensure both that the extensive human and material resources mobilized for the census are effectively and efficiently used and that its very tight time schedules and massive logistic requirements are met.

2.4. It must be stressed, however, that at each stage of census planning and implementation, the various administrative arrangements developed will need to be guided by sound technical considerations. The quality and timeliness of the census data will almost certainly suffer unless sufficient and appropriate weight is given throughout the census to a wide range of subject matter and statistical

requirements. This is especially valid in the case of cross-cutting issues, such as information technology, present throughout many essential phases of the census. It is for this reason that the management of a large statistical operation, and especially a population and housing census, cannot be considered a routine administrative assignment.[39]

2.5. Not all censuses follow a uniform pattern but there are certain major elements that must be taken into account in every one of them. In general, census operations can be divided into seven phases: (a) preparatory work and testing, (b) enumeration, (c) data processing, (d) building of databases, (e) evaluation of the results, (f) dissemination of the results, and (g) analysis of the results. In addition, distinct sets of operations related to the systematic recording of census experience and the quality assurance and improvement programme must accompany and support the main census operations. It will be readily apparent that these phases are not entirely separate chronologically or mutually exclusive. For example, some census results might be released before all data-processing activities are completed; the analysis and the dissemination of census results overlap quite extensively; and the systematic recording of census experience should start at the beginning of the preparatory work and continue through all subsequent phases. Furthermore, certain elements that are discussed below, such as the budget and staff, may have to be amended according to the circumstances arising at a later stage of operations. The elements of each of these phases are discussed below in terms of their implications for sound census management.

B. Planning for both a population and housing census

2.6. Housing or population censuses may be conducted separately, or as one combined programme. A combined population and housing census will be more costly and complex than each census considered by itself, but less expensive than the total operation of carrying out both censuses independently, even if those independent censuses are managed to minimize extra expense. If conducted separately, there may be added expense and challenge for training, resource mobilization, logistics, procurement, communication and data processing, among other operations. In addition to the financial and logistical advantages, combining both a population and housing census will provide a greater wealth of data than both censuses carried out independently. In assessing whether to run population and housing censuses as one integrated program, or as separate censuses, each country will have to decide on the trade-offs involved in the light of its own needs and circumstances (see also paragraphs 1.38–1.41).

C. Special considerations when planning a combined or register-based census

2.7. When Census is conducted using administrative registers, a comprehensive study of the benefits and issues, and whether a move away from traditional enumeration is feasible, should be conducted. The base register from which the basic data is obtained, it is advisable that the register has been prepared under a law. Countries that are planning for a register-based census should assess their legal mandates to ensure they can use government and private data sets, some containing considerable demographic, social, cultural or statistical information, to enumerate the population and provide characteristic data.

2.8. Before a register is used for census purposes, the suitability of data in terms of definitions, concepts, content, coverage, reference date and other criteria should be ensured. Birth and death data are among the most important sources for creation and maintenance of registered based census and so these and all other data used for updating base data should be available in digital format. For linkage of records from different registers, the existence or creation of a unique identification number for each individual, household and/or dwelling will be required. In addition, modelling techniques may be required in order to integrate and select the best available information from the various administrative data sources. A well-defined methodology to link different registers and its maintenance must be prescribed, which should include a plan for the periodicity of updates.

2.9. The analysis of a country's readiness for a move to a combined or register-based census should include a thorough assessment of the public acceptability within the population of the large-scale use of their administrative data. Some portion of the population may be concerned with the privacy implications of using administrative data or registers to enumerate and characterize the population, so the national statistical office should ensure that it has the legal and technical capacity to use large scale administrative data in a way that guarantees the safety and security of the data it may have in its holdings.

III.Strategic objectives and management

A. Strategic management

2.10. The primary value of strategic management is to assist census organizations to operate successfully in a dynamic, complex environment. The strategy drives key strategic decisions and choices over the term of the census undertaking in response to external and internal forces. The crafting of a strategy for a project or programme such as a census is critical for successful execution, and is dependent on a variety of knowledge and skills in different areas. While the crafting of the strategy is an art, it is important that it is governed by a systematic process to ensure careful examination and consideration of all issues that might have an influence on the future state of the programme. Census organizations may consider adopting the following strategic management process that will guide census operations in all its phases.

2.11. The strategic management process consists of four phases, namely:

a) Strategy analysis;
b) Strategy formulation;
c) Strategy implementation;
d) Strategy monitoring and review.

2.12. Strategy analysis is about ascertaining the issues that need to be addressed to take corrective action or to chart a new direction. It seeks to change the organizational set-up to one that is systemic, holistic, comprehensive and coordinated. Strategy analysis consists of:

a) Setting the direction for the census operation, with the objective of reaffirming its purpose and the way it should conduct its business.

b) Setting strategic goals for the census operation by way of defining what the system aims to achieve in terms of defining its highest goals and strategic outcomes. Strategic outcomes and goals must be aligned to what the user can expect and must therefore address user needs and requirements. A strategy driven by outcomes means "planning backwards" from the desired outcome through how best to achieve it.

c) Strategic analysis of the census operations by defining its current and previous status or situation and identifying the key issues that need to be addressed. Analysing the internal and external environment provides the evidence base to inform the development of the strategy. This analysis of the situation forms the basis of the strategy and its objectives to be defined. Census organizations may consider using a SWOT (strengths, weaknesses, opportunities, threats) analysis to do a strategic analysis of the internal and external environment.

2.13. Strategy formulation is about defining how and where the census organization must respond. The key steps involved in this phase include (a) develop and review the value chain; (b) develop strategic objectives and subobjectives; (c) compile a strategy map; (d): define the strategic intent; (e) identify critical success factors; (f) identify strategic risks; and (f) identify or develop strategic interventions.

2.14. Strategic objectives are essential for a successful census. They provide direction on what should be done to achieve the goals and outcomes. All activities in the census should be linked to a strategic objective, whether it is a new activity, an improvement initiative or maintaining the current status. Strategic objectives provide specific direction to the activities of the census operation and form the foundation upon which decisions on actions are made. They also provide the direction for everyone in the organization and motivate people to achieve them, especially if they are rewarded. Strategic objectives affect other aspects of management, such as planning, organizing and leading, and provide a benchmark for performance measurement as well as a mechanism of control through provision of corrective measures. Finally, strategic objectives form the basis for delegation of authority. Good objectives are helpful in effective delegation of authority.

2.15. The strategic objectives, as formulated, must be specific, measurable, achievable, relevant and time bound (SMART).

2.16. Strategy implementation is the process that turns strategies and plans into actions in order to accomplish strategic goals, outcomes and objectives. Implementing the strategic plan is as important as, or even more important than, developing the strategy. The key steps involved in this phase include:

(a) Compiling a census strategic plan, work programme and operational plan with the following elements:

I. *Work planning*: Compilation of the operational plan by defining the inputs, processes and outputs;

II. *Resource planning*: Identifying what human and other resources are required and how these resources should be deployed and developed to create the competences need to deliver the strategy;

III. *Financial planning*: Compilation of the budget outlining the funding required to implement the strategy;

IV. *Risk planning*: Identifying the operational risks and developing control and mitigating actions in response to those risks.

(b) Translating the strategy into action through:

I. *Defining the body of work (identifying key performance indicators, outputs to be delivered, targets and milestones that will deliver the strategy);*

II. *Defining the method of work (developing a value chain at conceptual and operational levels, including the quality management process that defines how customer value will be delivered);*

III. *Defining the organization of work (designing an organizational structure that implement the strategy).*

(c) Managing the implementation and strategic change.

2.17. Strategy *monitoring and review* is about monitoring and reporting on the progress, achievements and challenges in the programme; taking corrective action where required; and evaluating the impact of the changes and improvements. Implementation must be monitored to be successful. Due to constantly changing external and internal conditions, census managers must continuously review both environments as new strengths, weaknesses, opportunities and threats may arise. The key element in strategy monitoring is to get the relevant and timely information to take corrective actions where required.

B. Strategic objectives

2.18. The development of plans for a census should include the early preparation of a set of strategic aims and objectives that may be used to guide the implementation of the plans, set standards and form a set of benchmarks against which outcomes can be assessed to help determine the success of the census. Ideally, the starting point for developing these objectives would lie in combining information derived from evaluating previous census experience, from understanding stakeholders' requirements for information from the census and from assessing changes in both society and technology. In practice, some of this information is difficult to obtain and often provides conflicting guidance. Nevertheless, such objectives can be used to assist in planning major elements of the process. Although the strategic objectives of the census will be specific to individual countries and will differ according to local circumstances and priorities, they can be described under the following headings: census content; impact on the public and on census staff; production of census results; cost-effectiveness. The cost, benefit and the ability to harness technology must be in line with the general situation of each country.

2.19. *Census content.* The aim is to ensure a relevant census, such that the topics are appropriate to meet the clear requirements of *stakeholders*, while taking into account the considerations of cost-effectiveness, human resources, time availability and respondent burden. Census content should be the result of: a) appropriate consultation with current and potential *stakeholders* at all stages; b) the establishment of measurable standards of reliability incorporating *stakeholders* views on priorities; and (c) adequate testing of new topics to ensure successful collection and production of reliable census products ; (d) ensuring that outputs are comparable internationally, nationally and regionally; (e) the suitability of the census as the source of the data. Furthermore, should any administrative-sourced data be conceptually different than what is desired by the stakeholder, these conceptual limitations in the administrative data should be considered when designing any household surveys that are associated with the census or within the national statistical organization.

2.20. *Impact on the public and on census staff.* The aim is to ensure that all the aspects of collection operations, administrative data intake and use, as well as the dissemination of results are acceptable to

the public and fully comply with legal and ethical standards for protecting the confidentiality of individual information. The public should be fully informed about census objectives, content and methods, as well as about their rights and obligations with respect to the census. Similarly, all census staff must be fully aware of their responsibilities. Subsidiary objectives include such issues as (a) keeping completed forms, accessed administrative records and other records containing personal information secure and confidential; (b) ensuring that public support for all aspects of the census is as strong as possible, whether traditional or administrative; and (c) producing requested customized output in a manner consistent with preventing disclosure of personal information, adhering to established reliability standards for the release of data, and implementing policies designed to safeguard the access to all census products by the various stakeholders

2.21. *Production of census results.* The aim is to deliver census products and services and to meet legal obligations and stakeholder needs within stated quality standards and a predetermined and acceptable timetable. Subsidiary objectives include (a) producing timely outputs with a minimum of error suitable for the purposes for which the data are to be used; (b) providing standard outputs for the main results and services for customized output; (c) providing access to output; (d) using geographic bases appropriate for collecting and referencing data for output; (e) improving methods of traditional or administrative enumeration, particularly in difficult areas, so as to reduce levels of net undercoverage and response error; (f) improving methods of evaluation and the means to convey findings to users; and (g) developing a measure of coverage, quality and targets.

2.22. *Cost-effectiveness*. The aim is to plan and carry out a census as inexpensively as possible without compromising other strategic objectives. Subsidiary objectives relate to minimizing costs by (a) adopting more efficient data collection, data capture and data-processing approaches and related technological innovations; (b) contracting out appropriate parts of the operation; (c) exploring possible sources of alternative funding and, if appropriate, developing proposals for cost recovery and income generation; (d) international and in-country collaboration and reuse of systems and other resources; (e) encouraging the public to self-complete forms online or on paper where required ; and (g) replacing direct collection of data with use of administrative data; (h) integrating and maintaining linkages of different administrative data sources; and (i) exploring shorter periodic production of census results (for administrative-based censuses).

2.23. *Cost–benefit.* The aim is to increase the value or benefit generated from the census while also managing the overall cost. A cost-benefit analysis may also outline the costs and consequences of not having the necessary information. This requires clearly articulated benefits that follow logically from Census activities and outputs, along with plans to realise these benefits. Some of these benefits can be quantified, for example benefits generated through the use of census data. Other benefits can be more difficult to measure, but are nonetheless important and should still be recorded. Subsidiary objectives include illustrating the value of the census as an educational tool and framework, for comparative purposes at national and international level and as a cornerstone of the national statistical system.

2.24. These objectives can be used as benchmarks to assess user requirements and may also be built into appraisal systems that, with suitable weighting, can be used to compare and review options. In general, strategic objectives of the population and housing census need to be clearly emphasized throughout the process of preparing, conducting and exploiting census data.

Comments by GB and CC: Add content on role of census in the national statistical system, given that in the majority of countries the census ends up becoming the backbone of the country's social statistics

system, where the conceptual definitions used in the census operate as the organizing framework for the country's social statistics over the next 10 years.

C. Avoiding gender biases and biases affecting indigenous peoples and minorities

2.25. Gender-based stereotypes can introduce serious biases in census data and the conclusions drawn from these data. These biases are discussed in more detail in Part Four (see for example paragraphs 4.129–4.135 and 4.289–4.351 relating to household relationships and economic characteristics, respectively). There is much that can be done in the preparatory stages of the census to help minimize gender-based biases. These preparatory activities are of two broad types: those related to census content and those related to census operations.

2.26. Issues of census content, including what information is sought and how, the definitions and classifications used, and the manner in which databases and tabulations are specified, are important in generating data needed to examine questions of gender equity. In addressing these content issues, census planners and users will need to be alert to prevailing stereotypes so as to develop a census that both minimizes the influence of the stereotypes that respondents and enumerators may hold and avoids further perpetuation of these stereotypes.

2.27. With regards to census operations, particular attention will need to be given to the selection, training and supervision of the field staff. This involves ensuring that both men and women are recruited to the field staff (both as interviewers and supervisors) and that manuals and training materials cover gender bias issues just as they do other important sources of error. Consultations with women's groups and others concerned with gender equity can help in addressing both content and operational issues.

2.28. Gender-related stereotypes and biases are concerns that have relevance for all countries. Census authorities in a number of countries must also be alert to the possibility of stereotypes and biases affecting data on certain disadvantaged population groups. Such groups may include ethnic, linguistic, national, sexual orientation, racial, religious, indigenous, and nomadic populations. Persons with disabilities may often be subject to similar bias. As with gender issues, the problem will need to be addressed in terms of both census content and census operations. Representatives of these disadvantaged groups can often provide census planners with important information and insights relevant to both census content and operations (for detailed information about difficult-to-enumerate groups, see paragraphs 3.125–3.134 and 4.48). Thus, special efforts should be made to consult when planning the census. In the case of indigenous and populations living in isolated settlements or enclaves, such consultations are often critical for minimizing underenumeration among these populations.

IV. Units, place and time of enumeration

A. Units of enumeration

2.29. Since individual enumeration is an essential feature of a population and housing census, clarity about the unit of enumeration is an essential element of census planning. In the case of the population census, the primary unit of enumeration is the person. There are two general frameworks within which individuals are identified: (a) households, and (b) institutions, as a subset of collective living quarters. The household is a general framework within which most individuals are identified, since the majority of the population live in households, and the household is also a unit of enumeration in its own right. Because the household is also a unit of enumeration for traditional censuses, careful identification as a preliminary step in the enumeration can facilitate the efficient collection of the data and the control of its completeness.. For administrative censuses, the source data are often at the individual level, but data requirements from stakeholders may necessitate the construction of households from the individual-level source data.

2.30. As mentioned in the previous paragraph, the second framework within which individuals are identified comprises "institutions", as a subset of collective living quarters. In addition to persons identified within households, there are persons living in institutions who are not members of a household. This group constitutes the "institutional population", which is also in scope for population censuses.

2.31. For the housing census, the household is one of the three units of enumeration; the other two units are living quarters (in other words, housing units and collective living quarters) and buildings. It is important to bear in mind that, in conceptual terms, these three units are clearly distinguishable. There is not necessarily an identity or exact correspondence among these concepts nor are the terms themselves interchangeable. Several households may live together in one set of living quarters and one household may occupy more than one set of living quarters. Similarly, several sets of living quarters may together occupy one building and one set of living quarters may occupy more than one building.

2.32. It is recognized that there may be difficulty in some countries in maintaining independent concepts of "household" and of "housing unit".[40] However, the advantages in terms of the usefulness of the data that result from preserving separate concepts usually outweigh the additional effort required in maintaining them.

2.33. In carrying out a census, it is essential that the units of enumeration be clearly defined and that the definitions be included in manuals of instruction for the enumeration and, to provide appropriate guidance for users of the resulting statistical information, in census reports. In order to reduce the possibility of difficulties in applying the definitions recommended below, countries may find it necessary to expand the definitions and to illustrate them in terms of national conditions and circumstances. Post-enumeration field checks can provide a useful means of determining to what extent the national definitions of the units of enumeration have been applied in the field and the consequent effects on census results.

1. Person

2.34. For census purposes, the term "person" denotes each individual falling within the scope of the census. As emphasized above (paragraph 2.27), a person can be identified as belonging to the household population (that is to say, the population living in households) or to the institutional population (that is to say, the population living in institutions, as a subset of collective living quarters), as defined in paragraph 2.39 below. Although each person must be included in the count of the population, there will be some variation in regard to the persons for whom information is collected on different topics. The

variations usually depend on the person's age (for example, questions relating to economic activity, in which case the age boundary may be driven by national legislation), sex (for example, questions relating to children born), or relationship to the head or other reference member of the household. It may be recommended that information on a particular topic should be investigated for less than the total population, and the group of persons for which a given topic should be investigated is indicated below under the definitions and specifications of such topics presented in Part Four, Chapter I, section IV. In addition, the recommended tabulations for population censuses on the website of the United Nations Statistics Division are accompanied by a description of the population to be included in each tabulation. Similarly, the scope of the census should clearly indicate the persons to be covered and those to be left out.

2. Household

2.35. The concept of "household" is based on the arrangements made by persons, individually or in groups, for providing themselves with food and other essentials for living. A household may be either (a) a one-person household, that is to say, a person who makes provision for his or her own food and other essentials for living without combining with any other person to form a multi-person household; or (b) a multi-person household, that is to say, a group of two or more persons living together who make common provision for food and other essentials for living. The persons in the group may pool their resources and may have a common budget; and they may be related or unrelated persons, or constitute a combination of persons both related and unrelated.

2.36. The concept of "household" provided in paragraph 2.33 is known as the "housekeeping concept". It does not assume that the number of households and housing units are or should be equal. A housing unit, as defined in paragraph 4.427, is a separate and independent place of abode that is intended for habitation by one household, but that may be occupied by more than one household or by a part of a household (for example, two nuclear households that share one housing unit for economic reasons or one household in a polygamous society routinely occupying two or more housing units).

2.37. Some countries use a concept different from the housekeeping concept described in the previous paragraph, namely, the "household dwelling" concept, which regards all persons living in a housing unit as belonging to the same household. According to this concept, there is one household per occupied housing unit. Therefore, the number of occupied housing units and the number of households occupying them are equal and the locations of the housing units and households are identical. However, this concept can obscure information on living arrangements, such as doubling up, that is relevant for evaluating housing needs.

2.38. Households usually occupy the whole or a part of, or more than, one housing unit, but they may also be found in camps, boarding houses or hotels or as administrative personnel in institutions, or they may be homeless. Households consisting of extended families that make common provision for food, or of potentially separate households with a common head resulting from polygamous unions, or households with vacation or other second homes, may occupy more than one housing unit. For more discussion of household occupancy, see paragraphs 4.471–4.475.

2.39. A household may also consist of one or more homeless people. The definition of the homeless can vary from country to country because homelessness is essentially a cultural definition based on concepts such as "adequate housing", "minimum community housing standard", or "security of

tenure",[41] which can be perceived in different ways by different communities. The following two categories or degrees of homelessness are recommended:

a. Primary homelessness (or rooflessness): This category includes persons living in streets or without a shelter that would fall within the scope of living quarters.
b. Secondary homelessness: This category may include the following groups:
   i. Persons with no place of usual residence who move frequently between various types of accommodation (including dwellings, shelters or other living quarters);
   ii. Persons usually resident in long-term (also called "transitional") shelters or similar arrangements for the homeless.

These definitions should be supported by a data collection strategy that ensures, for example, that dwellings are properly identified as shelters and not households.

2.40. For some topics investigated in housing censuses, the household may serve more efficiently than living quarters as the unit of enumeration. For example, tenure, if investigated in the census, should be collected with reference to households rather than living quarters. Information about household possessions that are normally included as part of the equipment of living quarters (radio and television receivers, for example) should be collected with reference to households. Information on rent, an item of significance in relation to both living quarters and households, would of necessity be collected in relation to the household.

3. Population in collective living quarters

2.41. As emphasized in paragraph 2.27, institutions represent the second general framework within which persons, as major units of enumeration, are identified. The institutional population comprises persons who are not members of households. These include persons living in military installations, correctional and penal institutions, dormitories of schools and universities, religious institutions, hospitals and so forth.[42] Personnel responsible for the running of an institution and not living in dormitories or similar accommodations should be excluded from the institutional population. Although the members residing in the institutional households constitute a single unit, keeping in mind that in future more and more people would submit the responses in census through self-enumeration and past data may also be used for maintenance in future censuses, each member may be enumerated as single member household.

2.42. Persons living in hotels or boarding houses are not part of the institutional population and should be distinguished as members of one-person or multi-person households, on the basis of the arrangements that they make for providing themselves with the essentials for living.

4. Building

2.43. The building is regarded as an indirect but important unit of enumeration for housing censuses since the information concerning the building (building type, material of construction and certain other characteristics) is required for proper description of the living quarters located within the building and for the formulation of housing programmes. In a housing census, the questions on building characteristics are normally framed in terms of the building in which the living quarters enumerated are located, and the information is recorded for each of the housing units or other living quarters located within it.

2.44. A building is any independent free-standing structure comprising one or more rooms[43] or other spaces, covered by a roof and usually enclosed within external walls or dividing walls[44] that extend from the foundations to the roof. However, in tropical areas, a building may consist of a roof with supports only, that is to say, one without constructed walls; in some cases, a roofless structure consisting of a space enclosed by walls may be considered a building.[45]

2.45. In some countries, it may be appropriate to use the "compound" as a unit of enumeration, either in addition to the building or as a substitute for it. In some areas of the world, living quarters are traditionally located within compounds and the grouping of living quarters in this way may have certain economic and social implications that it would be useful to study. In such cases it may be appropriate, during the census, to identify compounds and to record information suitable for linking them to the living quarters located within them.

5. Living quarters

2.46. The principal units of enumeration in a census of housing are living quarters. Only by precise recognition of these identities can data be obtained that will provide a meaningful description of the housing situation and a suitable basis for the formulation of housing programmes and policies.

2.47. Living quarters are structurally separate and independent places of abode. They may (a) have been constructed, built, converted or arranged for human habitation, provided that they are not at the time of the census used wholly for other purposes and that, in the case of improvised housing units and collective living quarters, they are occupied; or (b) although not intended for habitation, actually be in use for such a purpose at the time of the census.[46]

B. Place of enumeration

1. Concepts relating to the place of enumeration

2.48. In the context of the population census, for censuses with field data collection, a country may wish to enumerate all persons present in the territory and/or potentially belonging to the population of interest. *Population to be enumerated* is the group of persons who the country decides should be covered by the census, regardless of their later inclusion in a population count. In other words, it is the group of people that form the target population of the data collection activities (be that direct enumeration or compilation of data from registers or administrative sources). A country might collect data on everyone present at a point in time, but then only produce official statistics referring to those who are usually resident, for example.

2.49. The place of enumeration is either the place where the person is found or the place of usual residence of the person at the census reference moment. It should be ensured that each person has only one place of enumeration, to avoid double-counting. Countries should document the definition of place of enumeration adopted for their census. In those countries which use direct enumeration, they should also provide explicit instructions to enumerators on how this definition should be applied during an interview, or to respondents when filling in self-administered questionnaires.

2.50. In general, “usual residence” is defined for census purposes as the place at which the person lives at the time of the census, and has been there for some time or intends to stay there for some time

i.e. the geographic place where the person usually spends their daily rest, assessed over a defined period of time including the census reference time.

46. Most individuals enumerated have not changed residence for some time and thus defining their place of usual residence is unambiguous. For others, however, the application of the definition can lead to many interpretations, particularly if the person moves or has moved residence often.

2.51. It is recommended that countries apply a threshold of 12 months when considering place of usual residence according to one of the following two criteria:

a. The place at which the person has lived continuously for most of the 12 months before census reference day (that is, for at least six months and one day, not including temporary absences for holidays or work assignments) or intends to live for at least six months;
b. The place at which the person has lived continuously for at least 12 months before census reference day (not including temporary absences for holidays or work assignments) or intends to live for at least 12 months.[47] [32]

2.52. For register-based censuses, the place of usual residence could be taken to mean the place of legal or registered residence, defined with reference to a qualification period of 12 months or assessed with reference to the 12-month criterion.

2.53. Persons who move frequently and do not have a place of usual residence should be enumerated at the place where they are found at the time of the census.

2.54. Regardless of the criteria used to define the 12-month period, countries should ensure that each person has one and only one place of enumeration (i.e., place of usual residence or registered place of residence for census purposes).

2.55. There are various population groups for whom some uncertainty may arise about their inclusion in the usual resident population. The following persons should generally be considered in the usually resident population:

a. Persons found at the moment of enumeration who cannot identify a place of usual residence, such as those who change residence often;
b. National military, naval and diplomatic personnel and their families, located outside the country, irrespective of their duration of stay abroad;
c. Foreign persons working within the country for international businesses or organizations (but not including foreign diplomats or military forces) and their families, provided that they meet the criteria for the usual residence in the country;
d. Merchant seafarers and fishers usually resident in the country but at sea at the census reference time (including those who have no place of residence other than their quarters aboard ship);

[32] This approach is consistent with the Conference of European Statisticians Recommendations for the 2020 round of censuses.

It is also consistent with what is recommended in the Recommendations on Statistics of International Migration, Revision 1,

Statistical Papers No. 58, Rev. 1, (United Nations publication, Sales No. E.98.XVII.14).

e. Persons who are irregular or undocumented migrants, as well as asylum seekers and persons who have applied for, or been granted, refugee status or similar types of international protections, provided that they meet the criteria for usual residence in the country (the intention is not to distinguish these persons separately, but rather to ensure that they are not missed from the enumeration);

f. Persons who cross a frontier daily or weekly to work or study in another country, provided that they meet the criteria for usual residence in the country;

g. Children born in the 12 months before the census reference time and are usually resident in the country at the census reference time;

h. Persons of minor age studying abroad for one year or more to attain primary- or secondary-level education, regardless of the frequency of their return to the family home located within the country. If the person is also working abroad, the same rules for cross-border workers apply;

i. Persons who regularly live in more than one country during a year, if they reside in the country conducting the census most of the time, regardless of whether they are physically present in that country at the census reference time.

On the other hand, the following group of persons should generally be excluded from the usual resident population:

a. Foreign military, naval and diplomatic personnel and their families, located in the country, regardless of their duration of stay and/or place of usual residence;

b. Persons of minor age attending the primary or secondary level of education whose family home is located abroad, regardless of the duration of their stay. However, if these persons are also working in the country, then the identification of the place of usual residence follows the same rules as for cross-border workers;

c. Third-level students who are or intend to be absent from the country for one year or more;

d. Persons who regularly live in more than one country during a year, if they reside in the country conducting the census the least amount of time, regardless of whether they are physically present in that country at the census reference time.

Comment FWN: We should harmonize this section 100% with CES to avoid confusion, unless there is some strong reason for which they need to differ. ESTAT has an Excel sheet of current CES TF suggestions for revised wording.

2.56. The concept of usual residence is often referred to as though it is synonymous with the concept of de jure residence. In certain circumstances, however, the term “de jure” may carry with it a requirement that the person’s residence at that place has a basis in the legal system applicable to that specific place. In turn this implies that people without such a legal basis should not be enumerated in that area. It is not recommended that censuses of population and housing limit their enumeration only to those people with a legal right to be in a place but rather, as described in section 2 below, they should include either all those present at the census reference time or all those whose usual residence at the census reference time was at the place of enumeration.

Comment FWN: Maybe we need a brief bit on WHY it's not recommended to exclude those who don't have a legal right to be there. Example: one of the key reasons is to avoid them refusing to take part in the census.

2. Operational issues relating to the place of residence and the place of enumeration

2.57. In a field enumeration-based population census, information about each person can be collected and entered in the census questionnaire either where he or she is (or was) present on the day of the census or at his or her usual residence.

2.58. In compiling the census results by geographic areas, however, each person who is part of a household can be included in either (a) the household (and hence the geographic area) where the person was present on the day of the census; or (b) the household (and the geographic area) where he or she usually resides. The same also applies for the institutional population. This allocation is not necessarily dependent upon the place at which information was collected for the individual but it can be simplified by the proper choice of a place of enumeration.

2.59. If a "present-in-area" population distribution is wanted, it is logical to enumerate each person at the place where he or she is (or was) present at the time of the census. If a distribution by usual residence only is required, it is more satisfactory to collect the information about each person at the person's place of usual residence. It should be noted, however, that it is not always possible to collect information about each individual at his or her usual residence, as, for example, when an entire household is away from its usual residence at the time of the census. Some provision must therefore be made for collecting information about such persons at the place where they are found at the time of the census.

2.60. If the objective is to obtain information on *both* the usually-resident population and the present-in-area population, then either each person present in each household or institution on the census day, or each person present and each usual resident temporarily absent can be enumerated at the appropriate household or institution. A clear distinction must then be made in the questionnaire, as applicable, between (a) persons usually resident and present on the day of the census; (b) persons usually resident but temporarily absent on the day of the census; and (c) persons not usually resident but temporarily present on the day of the census.

2.61. Depending on the categories of persons enumerated at any given place, information may then be collected both on the usual residence (address) of those only temporarily present, and on the place (address) at which each temporarily absent person can be found. This information can be used for the purpose of allocating persons to the household (or institution) and geographic location at which they are to be counted and for checking that no person is counted twice (namely, at both the usual residence and the place where present). The procedures to be followed during the enumeration and through the subsequent allocation of persons must, however, be very carefully planned and strictly adhered to if the allocation is to be accurate.

2.62. With the exception of mobile housing units (see discussion in paragraph 2.63), living quarters and buildings have a fixed location. The place where they are to be enumerated does not have, therefore, to be considered in taking a housing census. Information on households, however, and on the persons in households, can be collected and entered in the housing census questionnaire (for census with direct enumeration) either where they are (or were) present on the day of the census or at the usual residence. The procedure followed in the housing census should be governed by that adopted in carrying out the population census if the two censuses are carried out simultaneously. If the housing

census is an independent operation, however, the procedure to be followed should be carefully considered since it may have a significant effect on the validity of the results of the housing census.

2.63. Where persons and households are allocated to the place of usual residence, they should also be allocated to the living quarters that they usually occupy. The living quarters that they are actually occupying at the time of the census should be counted as vacant if they are conventional dwellings or should be excluded from the census if they are non-conventional dwellings.[48] [33]

2.64. Mobile housing units represent a special case as far as the place of enumeration is concerned. They should be enumerated where they are found on the day of the census. However, in accordance with the procedure adopted for the allocation of the population, mobile housing units may also be allocated to the area where the occupants usually reside, provided that they are the usual living quarters of the occupants in the area of usual residence. Where they are not the usual living quarters of the occupants in the area of usual residence, the occupants will be allocated to their usual living quarters and the mobile housing unit will be excluded from the census.

C. Enumeration point of time

2.65. One of the essential features of population and housing censuses is that each person and each set of living quarters must be enumerated at the same, well-defined point of time. This is usually accomplished by fixing a census "moment" at midnight at the beginning of the date selected as the "census day". This moment is the "census reference moment".

2.66. For the population census, each living person at the census moment is included in a census schedule and counted in the total population, even if the process of completing the schedule or inputting the administrative data takes place before or after the census day.. Infants born or persons who have deceased after the census moment should not impact what is entered in a schedule or included via administrative data in the total population on census day, even if those events have transpired between the census day and when the enumeration actually took place.

2.67. For the housing census, each set of living quarters that has reached an established stage of completion and is not scheduled for, or in the process of, demolition should be included in a census schedule and counted as a part of the housing inventory even though the process of completing the schedule does not take place until after the census moment or even after the census day, and the living quarters may have been scheduled for demolition in the interim. Living quarters that have attained the prescribed state of completion after the census moment are not to be entered in a schedule (unless special instructions are issued for recording living quarters under construction), nor should they be included in the total number of sets of living quarters.

2.68. Where the amount of time allotted for enumeration in the census is considered to be so long that the population is not likely to be able to supply information as of a single moment in the past, it may be necessary to employ different points of time in the enumeration, even to the extent of using the night before the visit by the enumerator. If such a procedure is followed, it should be clearly explained in

[33] To be considered as living quarters, non-conventional housing units and collective living quarters are required to be occupied in order to be included in the census.

the census report and the total duration of the enumeration should be stated. For ease of reference and for the computation of intercensal indices, it is useful to designate a single date in the enumeration period as the official "census date". This date could be, for example, the day by which half of the population was enumerated. This date is the "census reference (average) day" or, if reference is made to a period of time, the "census reference period".

D. Time reference period for data on the characteristics of the population and of living quarters

2.69. The data collected about the characteristics of the population and of living quarters should refer to a well-defined reference period. The time reference period need not, however, be the same for all of the data collected. For most of the data, it will be the census moment or the census day; in some instances (as is the case for current economic characteristics and rental arrangements), however, it may be a brief period just prior to the census or (as is the case for fertility questions, usual economic activity and information on the period of construction of the building in which living quarters are located) a longer period of time.

V.Legal basis

2.70. Legal authority for the census undertaking is required for regulating primary administrative responsibility, for obtaining the necessary resources, for determining the general scope and timing of the census, and for placing a legal obligation upon the public to cooperate and provide truthful answers, provide a mandate to collect and use administrative data for an enumeration, provide a legal obligation upon the enumerator and census officials to record the responses faithfully, and specific responsibilities upon other census field personnel at various supervisory levels. In addition, the confidentiality of the individual information should be strongly and clearly established in the census legislation and guaranteed by adequate sanctions so as to provide a basis for the confident cooperation of the public. In countries that lack permanent legal authority for the taking of periodic censuses, it is important to act early to establish ad hoc legal authority or, preferably, legislation calling for a system of periodic censuses. In case of an ad hoc legal authority, attempts should be made to ensure the independency of the implementing agency is safeguarded and powers to make technical and administrative actions are delegated to the said agency.

2.71. The principle of conceptual and organizational flexibility should be observed in drafting the census legislation. The legislative provisions should ensure data security and confidentiality. However, the inclusion of provisions that may be too rigid regarding the type of data to be collected or the structure and relationships of the various parts of the census organization is undesirable. Rather, necessary details should be contained in the census regulations promulgated by the census authorities. Moreover, provision may have to be made, in either the legislation or the regulations, for sanctioning the use of simplified administrative procedures, including the appropriate delegations of authority for the procurement of equipment and supplies and the recruitment of personnel during the operational phase of the census.

2.72. While the content of the census legislation will inevitably depend on national legal practices and procedures, as well as on the organization of the national civil service, the following components are usually represented: the purpose of the law; the coverage of the census; assigning the mandate for conducting a census to a specific institution; purpose of the census; obligation and rights of the citizens; modes of financing the census; organization of the census; administering the census; technology to be used, rights and obligations of enumerators and supervisors; census data dissemination and exploitation; treatment of individual data; confidentiality and privacy of respondents and their data; disposal of the census equipment and materials, access and use of administrative data, and archiving.

2.73. A comprehensive and well-timed legislative framework is of utmost importance for ensuring the legality and authority for conducting the census itself. In the case of an ad hoc approach to the census legislation, that is, in the case where it is done before each census, it often also contains the census topics clearly spelled out, thus providing additional legal weight to the composition of the questionnaire and the content of the census.

Comments GB: Add content related to guidance on enhancing census legislation based on lessons learned from previous censuses, including to enhance privacy provisions.

VI. Financial management

A. Financial basis for censuses

2.74. A census is the primary source of data about the size and characteristics of the population disaggregated into lower geographic levels; it provides a demographic profile of a country and is the basis for developing area sampling frames for use in surveys. A census, however, is usually one of the largest and costliest statistical activities that governments and their national statistical offices undertake. As a result, countries have been forced to delay or even cancel a census owing to funding constraints. Countries that have been able to secure partial funds or secure funds but at a late stage of their census preparation have been forced to compromise their data collection, data processing and dissemination of census results. It is therefore recommended that all census operations, including planning, cartography, enumeration, processing, analysis and dissemination, be budgeted from the beginning and efforts made to mobilize the required resources, for the entirety of the census cycle (i.e. for all fiscal years). Inflation should be taken into account, keeping in mind that duration has an impact on cost.

2.75. Hence, there is growing pressure to look into the solutions to census funding, taking into account the role of key stakeholders, namely governments and their statistical agencies, and the greater involvement of national and international development partners, stakeholders that includes the private sector. Concurrently, cost-effective strategies need to be put in place that would reduce census costs without compromising the quality of census data.

2.76. It should be emphasized, however, that censuses cannot be carried out merely by national statistical and census offices alone. Rather, conducting a census should be seen as a national task involving all stakeholders. Thus, government departments, non-governmental organizations and private

sector end users should be consulted at all stages to ensure the legitimacy of and need for conducting the census and, at the same time, to improve advocacy for sufficient funding. Although conducting a census is principally financed by the government, the census must be designed in partnership with all political actors so as to obtain their support and involvement in the census process. A high-level committee consisting of the government, the private sector and civil society, including non-governmental organizations, communities and donors, could be formed to discuss issues related to the funding and resourcing for the census.

2.77. National statistical and census offices need to advocate the importance of investing in censuses within their own governments. It is also important for the national statistical and census authorities to ensure continuous feedback and promote the use of statistical data from previous censuses, in order for users to recognize and appreciate the importance of the population census as a source of vital statistical data and give their support. The possibility of reducing the financial costs through collaborations with other government departments with existing devolved infrastructure, such as education and health ministries, should be further explored. These institutions could be supportive in providing administrative data and/or logistics arrangements for the census, such as the use of existing infrastructure, transportation, and communications facilities, and sharing of employees of other government departments.

2.78. Excellence in planning is an essential prerequisite not only for achieving a cost-effective census (see paragraph 2.11) but also for securing comprehensive financial support for its funding. Technologies and methods that will be used in mapping, data collection, administrative data intake and management, processing, questionnaire design, and other activities must be decided upon in advance, as these have an influence on costs. Census planning must bring out the links between the various components, which will include types of resources (such as personnel, cost of equipment, stationery or printing) as well as tasks (including data collection and capture, data processing, and data management, dissemination and archiving). Cost tags must be attached to each of these components together with a justification. Experience and lessons from past censuses or similar activities must be considered when estimating costs for the next census. Where multiple modes of data collection and new technologies are being used for the first time, these must be tested for data quality and cost implications.

2.79. For each stage of the census the costs must be optimized. A careful choice of the appropriate technology will greatly assist in this. Recent advances in technologies throughout the traditional census process, such as digital mapping, computer-assisted or Internet data collection, scanning, data processing (including machine learning technologies for coding) and data management and archiving, and census data analysis and dissemination, may be of assistance in achieving significant efficiencies, allow more cost reductions or increased productivity. In addition, the proper selection and use of such technologies will speed up the computation of results and enhance their preservation. However, the choice of technology should be made only after carefully evaluating the costs and benefits of possible options. Some potential risks to canvass include the following: some approaches only become cost-effective for large operations; some are dependent on expensive and scarce inputs (for example very high-quality satellite images or paper for scanning); some are dependent on services that may not be available throughout the country (for example Internet access); and others require significant investments in high-quality computers and upfront investment in human resources. The options examined in the cost–benefit analysis could incorporate consideration of leasing (rather than purchasing) equipment or sharing/borrowing it between countries that are undertaking censuses at convenient times.

2.80. Outsourcing or collaborating with the private sector could be considered as another potential cost-saving option, particularly in the context of publicity and advocacy or for systems development for data collection, transmission, processing, dissemination and archiving. Outsourcing can contribute technical expertise or resources not readily available within the national statistical office or census implementing agency.

2.81. It is anticipated that international donors will continue to play a pivotal role in supporting funding census costs in many countries. Technical cooperation and assistance from international agencies have also contributed greatly to the success of censuses in many countries. It is worth noting that a population and housing census has some intangible positive values. It is an opportunity for mobilizing the whole country and reaching even the most remote corners of it. In the life of many citizens, a regular census is often the only time that the State reaches out to them and asks them some questions. Successfully conducting a census is a matter of pride in many countries and a welcome opportunity to recruit a massive labour force and generate jobs and train people in valuable tasks (such as data entry) or in other ways to add to the national infrastructure.

2.82. There is opportunity for cost savings, through the use of administrative data to replace or support traditional census field operations, usually the most expensive phase of census-taking. As countries adopt increased administrative data into their censuses, this can be an important driver of change. However, considerations should be made for the added expense associated with the acquisition, intake, processing and other management (e.g. modelling) of administrative data used for census-taking. Also, if administrative data cannot entirely replace what is collected traditionally, good financial planning will be required to supplement any continued surveying that must continue, post-transition to administrative data.

2.83. In general, population and housing censuses are exclusively the responsibility of national governments and structures; this is particularly true for funding the census. Thus, all activities related to funding need to be elaborated, documented, justified and presented to all stakeholders in a transparent and comprehensive manner.

B. Budget and cost control

2.84. While no universal system of census budgeting and cost control can be suggested since financial practices vary greatly among countries, a few generally accepted principles can be noted. First and foremost, effective planning and control of the various census operations are not possible without a very careful financial estimate of the cost of each census operation, including all of its components, no matter how small. It is recommended to draft a detailed list of activities related to censuses with timelines and, as much as possible, to draft the budget in such a way that it corresponds to this list of activities. Second, it is critical for this census plan and budget to be presented by national statistical and census agencies to their respective governments with adequate lead time, to facilitate the appropriation of sufficient resources from national budgets in time or, where required, from the international development community. Moreover, funding of the census must be accompanied and developed on a sound and adequate legal basis if effective national census operations are to be enabled.

2.85. Information on expenditures from the previous census classified by census phases, starting with the expenditure for different elements of the preparatory work and ending with expenditure for the

dissemination of the census results, provides an important basis for estimating the budget of the census. Figures from the previous census will of course have to be reviewed and modified in order to take into account quantitative and qualitative change in hardware and software, changes in wage rates and the costs of equipment, supplies and so on, planned changes in census content, methods and procedures, and anticipated changes in the population itself (for example, total size, percentage urban, and average household size), all of which may affect the cost structure of the census. In most countries, several cost elements tend to increase (for example, wage rates and the size of population) so that there is considerable pressure to achieve economies in other items of the census budget. In addition, time must be spent to identify new features of the next census, for which a past expenditure cannot provide the basis of a cost estimate. For example, future censuses might need to consider a new budget for cybersecurity, or administrative data acquisition and use.

2.86. The census offices need to implement transparent accounting procedures and financial management systems to ensure speedy disbursement of funds, proper receipting of expenditures and an efficient audit. This would enable prompt release of periodic allocations of census funds by national governments. A clean outcome from a financial audit adds credibility to the census process so that the government and civil society are more likely to accept the final results.

2.87. In the case of external or donor funds, the required conditions should be established well in advance by discussion between the donor and the national statistical or census office. This will avoid delay in the release of such funds for census operations.

2.88. Control measures and monitoring systems must be developed for cost-effectiveness. Activities to be outsourced must be clearly defined and contracts for outsourcing should be well prepared with clear deliverables and timelines.

2.89. For planning the costs of a census, detailed and precise data will be required on the following: (a) number and cost of census staff classified by function and manner of payment; (b) type of equipment and material used for the census, manner of acquisition (purchase or rental) and cost; (c) office space (surface measurement), classified by use and type of cost (that is, for construction or for rent); and (d) type of services used for census operations. The usefulness of the above information would be enhanced if the information could be recorded by source of funding, in other words, in terms of whether the expenditure has come from (a) the official census budget; (b) other funds of the census office (for example, a regular annual budget not specifically intended for census purposes, or general funds of the governmental agency or department of which the census office is a part); (c) other parts of the government; (d) non-governmental organizations; or (e) international donors. This information is needed not only for fiscal planning and control but also in order to examine the trade-offs in terms of costs and benefits among alternative ways of carrying out various census operations. Although cost experience from a previous census in a country may provide useful experience for planning the next census, much more caution should be exercised in using the cost parameters from other countries. Differences in census content, organization and operations, as well as in cost accounting, can introduce serious incompatibilities into such country-to-country cost comparisons.[49]

2.90. It is important that the persons at the administrative and supervisory levels who will be responsible for the execution of each operation participate in estimating the budget items. Such an organization of the work presupposes detailed advance planning and "cost consciousness" on the part of those responsible for a census.

2.91. The census plan as executed will certainly change in a number of respects after the making of the original calculations. Consequently, a perfect correspondence between the estimates and the final costs is not to be expected. Changes in the prices of major components of census costs should be monitored on a regular basis with either the census budget adjusted accordingly or the census plans modified. Indeed, the development of the census budget is usually an incremental process in which rough initial estimates are replaced by more detailed and precise statements of resource requirements. Throughout the period of census taking and compilation of census results, the budget will have to be re-examined and performance compared with plans. With detailed information on expenditure, the governmental and census authorities will be better able to exercise control over keeping the development of census operations within the census budget and to assess and control the effectiveness and efficiency of these operations. This information is also very useful for studying possible improvements in census techniques and census methodology.

2.92. As with any project, particularly ones as large and complex as the census, it will not go according to plan and there will be difficulties. Accepting this at the outset and making sure there are arrangements in place to deal with delays, changes or other unforeseen issues is essential. In particular, there must be resources set aside to enable such issues to be dealt with quickly. Therefore, some contingency funding should be included within the overall costs of the census, and some controls put in place to monitor and allocate the contingency pot. Different methods exist for estimating the cost of the contingency budget, such as risk modelling, but a good starting point might be to allocate a percentage of the annual budget (say 15 per cent) each year for such contingency.

VII. Resource mobilization

2.93. Census is a massive undertaking that requires huge resources to implement. The resources range from financial resources to fund the costs of various activities and equipment, huge number of personnel with various skills and qualifications to implement various activities, and materials and equipment. There is need therefore for the implementing agencies to develop a comprehensive resource mobilization strategy for a census prior to commencement of the exercise. The strategy should detail the required resources, timelines, proposed mechanisms or strategy for mobilizing the resources, the possible stakeholders/areas to get the resources from ……..and modalities of acquiring the required resources. The modalities may include borrowing, collaborations, ……

2.94. A resource mobilization committee for the census should be established well in time to develop, implement and monitor implementation of the resource mobilization strategy for the census. The strategy should have a detailed workplan and implementation plan to be monitored.

[More content to be added]

VIII.Project management

A. Development of workplans

2.95. The Project Management Institute defines projects as "temporary efforts to create value through unique products, services, and processes". Censuses are major projects which require very careful, advanced planning and governance. Their complexity necessitates a large number of people working together, with various skill types and levels. At the centre, orchestrating it all, should be a census manager and governance that pulls the necessary team members together to make planning and execution decisions.

2.96. During the planning phases of a census, this team managing a census will need to detail their workplans. Using previous censuses as a benchmark, the team should consider what is required for the next cycle or iteration of the census, and articulate a vision. This high-level plan for the next census will require adaptations or new development of multiple workplans. These workplans should describe what is planned for the next census, the resources (human, financial and other) required to achieve this outcome, and the timelines for planning, developing and implementing it.

2.97. Multiple workplans will be required. Each should be prepared by an assigned team member with the greatest expertise and approved by the census manager and their governance teams, after they are fully articulated with the appropriate inputs from as many elements of the census program as necessary. The assigned team member who prepares the workplan should also be the same person who presents updates on how plans are developing. In other words, they are the person accountable for planning and executing on the workplans.

2.98. To ensure clarity, census project teams developing their workplans should have "charters" that describe their function within the census program. These charters should describe roles and responsibilities that make it sufficiently clear as who should be the responsible team member taking the lead on the development of each workplan. The charter should also integrate with the overall census project charter, a document that provides the basic parameters for the next census, developed from previous cycles and adapted to incorporate the vision for the next.

2.99. A key product of the workplans should be a "critical path" of milestone dates for the completion of work to prepare for the next census. This critical path represents key dates when if one activity is not completed, it would have downstream impacts on other parts of the census plan. Other dates should be tracked as well, but the critical path are the most important dates that the census team should be reviewing regularly and with greater frequency at key points during a census cycle (e.g. in the months leading up to field operations of a traditional census).

2.100. It is recommended that all censuses have on staff at least one certified project management expert, who can develop project plans and other elements crucial to the successful completion of a census.

Comments GB: Add more elaboration on project management basics?

B. Emergency preparedness, risk management and contingency planning

2.101. The proper conduct of a population or housing census requires thorough planning that often spans years of work by dedicated and knowledgeable staff. However, those plans can be interrupted by unexpected and very serious events, such as natural disasters, pandemics or war. These serious events can lead to the failure, cancellation or delay of a census if there is no time or ability within a country to respond properly. But it is not just these serious situations that can cause census-takers great grief. Given the enormous complexity of censuses, less catastrophic incidents such as labour issues or technical challenges can also cause considerable problems for census managers.

2.102. Unfortunately, the context for census-taking is getting increasingly complex and risk-prone. Censuses are more complicated and expensive than ever, with new technologies and techniques being adopted to properly enumerate populations. In addition, it seems that these more complex censuses are being run in a world with increasing risks. For example, in the 2020 round of censuses, countries were faced with a world-wide pandemic, numerous natural disasters (some associated with climate change), new cybersecurity threats and tight labour markets. Some countries had time to manage and adapt to this rapidly changing context, and others adapted "on the fly", with varying levels of capability to do so.

2.103. In response, this chapter is intended to provide advice for countries wishing to manage issues and risks related to censuses. It is designed to support census-takers in all phases of a census but especially during high-risk periods such as during the field operations of a traditional census. This chapter should also support all census managers, regardless of the type of census that they are undertaking in their country, or the nature of the risks with which they are faced.

C. A framework for emergency preparedness and contingency planning

2.104. Emergency preparedness and contingency planning are broad topics which can mean different things to different people. The following guidance demonstrates how to avoid issues in the first place by effectively using the planning phase for risk mitigation, but also to support census takers when the inevitable happens during the execution phase. Also, the following is designed to provide support and ideas for when, despite good planning and emergency management, a census has a major break-down of its intended plan.

1. Effective documentation of inherent risks to the conduct of a census, and their possible mitigations (risk registry)

2.105. To avoid or manage risk, one must first identify it. Census takers should be aware of their most important risks and set up processes to document them, with a resulting "risk registry". As defined by the International Standards Organization (ISO) in their risk management standard ISO 31000, a risk is "…anything that generates uncertainty or creates a deviation from the expected"[i].

2.106. In a census context, these are issues that may threaten the proper execution of a census, the quality of the data, the security of the information collected, or the availability of human or financial resources. As one is planning to conduct a census – either traditional, administrative-based or combined – time should be dedicated to on-going discussions with census management teams, in order to identify and manage risks. Ask yourselves, "what are the threats to this census"? Build upon previous experiences when identifying risks – what happened the last time, and can it happen again?

2.107. Risks can be further assessed using categories such as "low", "moderate", "high" and "extreme" to characterize their inherent likelihood of occurrence and the level of impact if the risks were to materialize. With this approach, one should be better aware of the vulnerability of the census to various threats.

2.108. Next, standard risk management suggests that census managers and experts discuss mitigations, or ways to lower the likelihood of occurrence and/or the impact of any threat to a census. Management must often prioritize this work, focussing on mitigations for the most significant risks, and/or those risks which are most likely to affect the census in the near term.

2.109. The best way to deal with a risk is to avoid it altogether through mitigating steps, but not all risks are fully within the control of the statistical agency. It may be important to discuss risk mitigations with external partners, such as IT contractors, postal agencies, printing firms, labour providers, administrative data providers, or any other key stakeholder in the proper conduct of a census.

2. Proper emergency management, governance and decision-making (Emergency management plan)

2.110. A Census is one of the most complex projects that governments undertake. With complexity comes a large number of "moving parts" subject to events not controlled or even anticipated by the census taker. Natural disasters are a good example. Although census managers understand that natural disasters can affect their operations, there is no predicting when, where or to what extent. Even with the best risk management plan and mitigations in place, a census is too complex to anticipate every potential issue. As a result, the statistical agency must be prepared to respond to the unexpected. With proper emergency management, governance and decision-making, censuses can weather difficult storms effectively.

2.111. Every census should have a process in place to identify, classify and respond to an inevitable emergency. Key members of the census team should know how to respond to emergencies during crucial phases of their census. They need to be able to determine if an emergency is underway, how to classify the severity of the emergency, who to contact when it occurs and how contact can be made, so that solutions can be put into place.

2.112. Census team members must be equipped with the information necessary to detect new issues and to respond effectively using communication lines to other team members, senior officials and perhaps the general public. For example, in today's census environment, traditional censuses are making greater use of internet questionnaires and real time metrics which can be monitored to ensure the proper functionality of that service to citizens. As well, for both traditional and register-based census, massive databases require constant monitoring to detect anomalies and cyberthreats. Should there be an issue detected during such monitoring, the person viewing that information must be

prepared to interpret what is happening, classify the degree of the emergency and communicate to others to establish a resolution plan.

2.113. When the emergency is detected, is should be assessed for its severity. Census team members responsible for the component of the census impacted by the event should be trained on various severity definitions during the planning phase of the census (i.e. before the census is executed). In the preceding example, when an event impacting the internet response service is detected, the information should be shared with the census manager responsible for that service, who is trained to determine the emergency's severity level. Is what is being detected a slowdown of the service that might be of minor annoyance to the user, or is this a broad interruption of internet response, such as a cyberattack?

2.114. The severity level of the emergency as determined by the responsible manager should trigger a pre-determined response. For example, a high severity event could require a meeting of the full census management team within a short period (e.g. no later than one hour) after detection, and require faster resolution, relative to a low severity event. At that meeting, the situation should be described by the responsible manager, who should then present information on the cause of the emergency and any identified remedies. If possible, there should be an estimate of when impacted census operations will be fully operational again, which will be particularly valuable for communications and any down-steam activity that may be subject to a subsequent impact.

3. A new threat to censuses: dis/misinformation

2.115. Censuses face a new threat of dis/misinformation, which is unlikely to go away anytime soon. Exploiting weakness in internet literacy and mistrust in government, sophisticated actors can launch campaigns to divert attention away from a census or embarrass governments by disrupting operations. Censuses now need to invest in tools to monitor social media, the most common way that dis/misinformation campaigns are conducted.

2.116. Specialized "social listening" tools can comb through large databases of publicly available social media posts to find incorrect information being shared about censuses. When detected, a statistical agency can decide to intervene as appropriate with responsive communication tactics, to correct the record and re-establish the facts regarding the census. Not all situations will be the same, and some responses will require close working relationships with major social media companies. If a major dis/misinformation campaign is detected, the statistical office should be able to quickly contact the platform(s) on which the incorrect information is circulating, and request that posts be removed before more damage is inflicted on the census. Other interventions may be necessary, such as countering the offending posts with information in the social media accounts of the statistical agency.

2.117. While powerful and effective, social listening tools must be used carefully, however. While reviewing only publicly-available information, there is a risk that government census takers could be perceived as invading privacy. As a result, before these tools are used, privacy experts should be consulted and offered an opportunity to approve or reject the proposed social listening.

2.118. To detect and respond to an emergency, a census should have an Operations Centre where employees have access to tools that track relevant operational data. When a census is underway, a number of operational metrics are available to staff, who should be presented with those data in real-time and with an historical archive to reference as well. Prior to the execution of a census, those data can be specified in advance, using expertise from previous cycles. The resulting "Management

Information System" (MIS) will gather information from various systems and present them in a way that allows managers to detect and respond to incidents.

2.119. In the event of an emergency, census managers should be prepared to communicate within their organizations. As mentioned earlier, priority should be taken to communicate first with the census team, to ensure a timely response. However, internal communication to senior leaders in the organization and political staff may be required, depending upon the severity of the incident. Also, should they be implicated in the response, or impacted by the emergency, external contractors may need to be notified.

2.120. In addition to providing tools for employees tracking MIS systems for signals of trouble, the physical Operations Centre can be a place to meet to discuss the issues and respond. The Operations Centre should be a place where communication crises are managed and should have access to facilities and equipment to record video and develop other content to be broadcast on the internet and other media. To monitor any ongoing situations and detect any new ones, the facility should have access to television, radio and internet broadcasts.

2.121. It is recommended that the Operations Centre be both a physical and virtual space. Should a physical space not be available (which was the case for many countries during the COVID-19 emergency), staff should have all the same tools that exist in the Operations Centre, available to them while working remotely. From a location away from the census office, the fundamental requirements for staff should be to have access to MIS and other reports and have an ability to communicate via phone and video calls. Any capacity to respond to an emergency should not be impeded by the closure of a census management office location.

4. Use of practice sessions such as "tabletop exercises" to condition staff on how to react in the event of a crisis to a census (Team conditioning).

2.122. While it is impossible to anticipate all possible scenarios, key census staff should be prepared for all significant risks After establishing a risk registry and emergency management protocols, prior to the execution of a census, census teams should practice what they might encounter. These practice sessions can take various forms and can be adapted for different types or levels of census staff. For example, one might be targeted toward IT staff, while another might be for the management team.

2.123. These exercises are an effective method to condition census staff on emergency management. The exercise would involve staff members who should be made aware of all (generic) emergency management protocols already developed. With that knowledge in hand, participants to the exercise will be introduced to various scenarios that should be based on potential issues identified in the risk registry. When the scenarios are presented, staff should reflect on the best next steps that should be taken. Everyone at the sessions should be expected to contribute and provide feedback on the reflections of other team members to the scenarios.

2.124. During the conduct of practice exercises, team members may find themselves uncomfortable. This is part of the nature of the exercises – the whole point should be to get used to very difficult situations. As a result, participants will find themselves unsure of what to suggest as solutions to the situations they are practicing. Furthermore, these practices sessions should include scenarios which are

progressively more difficult to deal with. It will be impossible to imagine and practice all scenarios that census managers might find themselves in. However, they should be realistic and designed by knowledgeable individuals who use the previously developed risk registry. The sessions should not seek to develop an ability within the team to know all steps to follow in all emergencies, but instead the sessions should develop a more general ability within the census team to respond quickly and effectively during the execution of a census.

5. Ensuring fast workflow during an emergency

2.125. Census teams understand the importance of their work and therefore respond well during emergencies. Without structure and good management, some census teams will work to do "whatever it takes" to ensure success despite setbacks. However, techniques other than hard work and long hours can be applied in an emergency situation, to ensure the completion of the work without unnecessary negative effects on the people working during the emergency.

2.126. One likely requirement of a census emergency will be fast workflow and strong communication to ensure the impacted components of the census are aware of what is changing, so they can adapt as well. During a census emergency, solutions must be found quickly and acted upon appropriately. As described above, once an issue is detected a quick response will be required.

2.127. The census manager in the statistical organization has an obligation to orchestrate the census team such that the resolution to the problem is efficient and effective. Lean management techniques can be applied, such as Kanban boards and quick, frequent "stand up" or "sprint" meetings. Lean management techniques such as these should be applied in order to make sure the full census team is working on the correct priorities and that these new priorities are communicated to the affected team members.

2.128. Census takers run a sort of data factory. There is an assembly line of inputs that are being transformed into data outputs and the factory line must continue running, even when it is impacted by some unforeseen event. This data factory, unlike the manufacturing factory, has elements that remain hidden from view by the manager and team members. Depending upon the severity of the emergency, the census manager and the broader team may feel the need to find ways to see what changes are required to remedy a crisis, to ensure priorities are being actioned and at the speed necessary.

2.129. One way that census managers can "see" the factory line, is to create virtual and real Kanban boards during an emergency. The Kanban technique will have the team develop and act upon problem statements describing the response to an emergency. For example, if there has been an attempted cyber-attack a problem statement might declare an issue to be resolved, such as "Re-open internet portal" and a team is assigned to resolve the issue. There might be related issues created by the cyberattack, for the communication team or the downstream operations. After the creation of the various issue statements, they can be placed into a "to-do" list or Kanban board, which provides a visual representation on the work created by the emergency. The census manager can meet with the team regularly, perhaps in the form of quick sprint meetings, to review the Kanban board and make sure the work to re-establish regular operations is moving at a fast pace. As well, the Kanban board is an effective way for other members of the team to see if there are any implications of the work of another team, on their area of responsibility.

6. Developing a contingency in the event of main plan failure

2.130. Despite our best efforts, censuses can fail to reach their full objectives. Disruptions to the execution of a census are inevitable and may severely impact the expected data quality. When risks are looming, a census taker can respond in a number of ways. The best scenario is that the risk is fully mitigated such that it does not develop into an issue. Another situation might require a census taker to cancel or pause a census, as was the case for many countries in the 2020 round, impacted by the COVID-19 pandemic. However, if a census risk cannot be fully mitigated or avoided, the issues that are created may not always be manageable. So much can be beyond the control of the census taker, and with censuses becoming increasingly complex (such as with combined censuses or those using large amounts of administrative data from multiple sources), there is always the risk that the main plan fails. What does a statistical agency do in that case?

2.131. Unfortunately, the options may be limited. For countries conducting traditional censuses, there may be an option to accelerate plans to use administrative data. While the available administrative data may not be suitable to replace all planned data variables and on the original timelines, national statistical agencies may be able to perform the basic function of a census with an administrative population count. If the risk of main plan failure is high enough, a national statistical agency might consider a parallel administrative count, to ensure the timely delivery of population counts in any scenario.

2.132. While it may not be considered as such, a robust administrative data plan for traditional census takers can be an effective "insurance plan" and therefore an incentive for investment in administrative data, technology to manage it, and expertise within the census team on how to use administrative data properly for census taking.

2.133. For register-based censuses, a disruption of the census would likely involve a communication plan and an adjustment of the timeline of the census, to allow a re-start of the process when the disrupting event has passed. If this is the case, there would an impact on other programs dependent upon the census, such as an extension or temporary introduction of intercensal population estimates.

[ii] International Standards Organization. (2021, December 10). *Effective risk management. New international guidance on how to get it right.* https://www.iso.org/news/ref2773.html

D. Change management

2.134. In each census round, when planning for a new population and housing census begins, various innovations derived from technological developments, lessons learned from the previous census, and the ongoing need to improve the balance between budget and data quality are evaluated. When a new census involves significant methodological (from jure to facto) and/or technological changes, such as transitioning from paper questionnaires to electronic forms, or to a combination of collection methods (paper, electronic), or from traditional to an administrative data based census, change management becomes essential to ensure that these innovations are implemented effectively, and the desired results are achieved.

2.135. Change management is grounded in a set of principles, approaches, and strategies designed to facilitate the successful transition of an organization or project from its current state to a desired one. To achieve this, it is crucial to comprehend the reasons for the change and assess the prevailing conditions, challenges, and opportunities. For instance, when transitioning from a paper-based census

to one that is more electronic, it is imperative that staff comprehends the comprehensive benefits and impacts on the census process. Therefore, it's essential that senior management actively supports and effectively communicates the change vision and is actively engaged in the implementation process. Employee insights and concerns should be taken into consideration, and they should be provided with training in the new skills and knowledge necessary to adapt to the change.

2.136. Detailed planning is essential, encompassing clear objectives, timelines, responsibilities, allocated resources, and risk mitigation strategies. This plan should also maintain flexibility to address any deviations that may occur. Lastly, there must be a specific strategy in place to address any resistance to change from employees or the community.

2.137. In a census project, it is vital to establish a change management plan to ensure the effective implementation of new methods, technologies, tools, or processes. This plan should not only improve efficiency and reduce costs but also enhance data confidentiality and security, adapt to the demands of modern society, foster community acceptance and collaboration, and ultimately enhance the quality of census data.

IX.Administrative organization

A. Overall overview

2.138. In planning the organization and administration of a census, it is important to consider the role and relationship of the various executive and advisory organs. National, subnational and local commissions and committees are frequently useful in the planning and preparations of a census. Such bodies may be composed of representatives of governmental agencies, community leaders with due representation to all sections of society, and non-governmental users of the census data, particularly those involved in policy-oriented analysis of census results and analytical studies of the social, economic and demographic situation of the country. This ensures broad-based and complete participation of the people to enable proper canvassing of sensitive issues such as ethnicity, gender, disability, migration, and marginalized groups. It is important, however, that their advisory and promotional functions be clearly defined and that the final responsibility for planning and execution rest with the executive agency.

2.139. There are well documented and proven advantages in having an office continuously responsible for census work established as an integral part of the statistical system of a country. Such an office assures continuity in census work and is the principal centre for the formulation of the programme and the initiation of preparatory work for the next census. Its permanence permits the development of specialized and experienced personnel and the maintenance of statistical and cartographic information, including cross-cutting issues such as information technology, essential for planning the next census.

2.140. At the pre-enumeration stage, the census office will need to be expanded to form the nucleus of the full census organization, which must be capable of directing the field organization during the preparatory work as well as during the enumeration and processing. In order to provide immediate supervision in each area, field offices at various levels are needed for the later part of the preparatory

work, including staff recruitment and training, as well as for the enumeration period. Supervisory personnel in such offices should be persons who, being familiar with the particular area and the local language, are able to deal with local problems. This does not mean, however, that all supervisory positions need necessarily be filled by persons from the area. Personnel may be transferred from the central office or from other areas as the need arises. Prior training of all such available personnel is necessary, so that they have a working knowledge of all aspects of the current census programme. An essential part of the preparatory work is local administrative planning, which would set out the likely problems and challenges specific to the local area and how they are to be dealt with. Adequate coordination with local public authorities is always important so that the enumeration work is not interrupted due to other activities.

2.141. Subsequent to the enumeration, the census organization is usually readjusted to meet the needs involved in compiling, evaluating, analysing and publishing the results and to provide the continuity desirable for promoting the continued use of census materials. Census organizations need to pay special attention to continuity of knowledge and skills from one census to the next, since the intervening gap, which is usually a whole decade, is likely to cause loss of institutional memory and attrition of qualified personnel. Comprehensive documentation of census activities while they are being carried out is essential, as is training of younger personnel to create a pool of knowledgeable and experienced persons by the time the next census comes.

B. Roles and responsibilities of key stakeholders in the census process

2.142. Population and housing censuses are complex programs that require the involvement of several stakeholders. The roles and responsibilities of key stakeholders in the various phases of the census should be clarified and adapted to the local situation. For example, while roles and responsibilities of the government are at the central level, these can also be defined for sub-national governmental authorities.

Government

- Issue a proclamation or decree for the census with an adequate lead time for preparatory activities.
- Ensure allocation of adequate resources (financial, human, physical, etc) for the census implementation.
- Assure a conducive environment for the census undertaking.
- Reinforce the statistical purpose of the census and assure against politicization or other influences that can compromise the census.
- Contribute to the identification and prioritization of data gaps to be addressed through the census.
- Ensure census data serve a public good for improving population wellbeing.
- Provide administrative data for use in producing or supporting the census

National statistical organizations

- Ensure effective intragovernmental coordination for the successful implementation of the census.
- Ensure consultations and effective engagement with various census stakeholders throughout the census process.
- Lead the development of technical resources for strengthening capacity of the census personnel in all phases of the census.
- Ensure high quality, international standards for the census, including the selection of the most appropriate enumeration method(s) for the country context.
- Promote effective dissemination and use of the census data, including public data access.
- Ensure the safety and security of census data, as well as for employees.

The United Nations and other international organizations

- Coordinate the development and promote adoption of international principles and standards for the census.
- Provide technical assistance for the census through National Statistical Offices.
- Promote the census as a public good for improving population wellbeing.
- Contribute to the identification and prioritization data gaps to be addressed through the census.

Donors

- Provide financial and technical assistance for the census within the scope agreed with the government.
- Promote learning and capacity strengthening for census undertaking, including through South-to-South-and Triangular Cooperation.
- Contribute to the identification and prioritization of data gaps to be addressed through the census.

The private sector

- Provide financial and technical assistance for the census within the scope agreed with the government.
- Promote the adoption of technology for the census process in accordance with industry best standards.
- Contribute to the identification and prioritization of data gaps to be addressed through the census.
- Provide support services and goods for the conduct of a census
- Provide administrative data for use in producing or supporting the census

Academic and research institutions

- Contribute to the identification and prioritization of data gaps to be addressed through the census.
- Provide administrative data for use in producing or supporting the census
- Promote the dissemination and use of census data.

The media

- Contribute to the identification and prioritization of data gaps to be addressed through the census.
- Promote public education on the value and purpose of the census.
- Promote the dissemination and use of census, including through thematic analyses on dimensions of sustainable development.
- Support public mobilization for participation in the census.

Civil society and other interest-group organizations

- Contribute to the identification and prioritization of data gaps to be addressed through the census.
- Promote public education on the value and purpose of the census.
- Support public mobilization for participation in the census.
- Provide administrative data for use in producing or supporting the census

The general public

- Contribute to the identification and prioritization of data gaps to be addressed through the census.
- Provide complete and accurate responses to the census questionnaire(s).
- Demand accountability for the use of census data to improve population wellbeing.

C. Statistical leadership

2.143. The period of preparation for the population and housing censuses represent a unique opportunity to exercise leadership in promoting the use of statistics in overall development of societies, with the focus on improving service delivery and policy development. Statistical leadership needs to be built and should rely on international standards and guidelines, such as the Fundamental Principles of Official Statistics, on national statistical legislation and on the national statistical code of ethics.

2.144. An important role of the head of the national statistical office or statistical agency responsible for census undertaking is to ensure a successful census programme that delivers results for use in

evidence-based policy development, evaluation and research, and decision-making. It is therefore recommended that heads of national statistical offices or statistical agencies responsible for census undertaking drive the following activities:

i. Facilitate the establishment of the statistical legislation that mandates the undertaking of a census;
ii. Setting policy and strategy by defining targeted outputs and outcomes for the programme;
iii. Strategic engagement with stakeholders by mobilizing participation across government, business and the public at large;
iv. Raising the profile of and commitment to use statistical information, providing an opportunity for engagement on key policy issues and strengthening relationships between information providers, policymakers and opinion leaders;
v. Adhering to the Fundamental Principles of Official Statistics and ensuring best practice is embedded in statistical procedures;
vi. Aligning to international standards and frameworks;
vii. Establishing statistical infrastructure and resources for undertaking the census;
viii. Setting up the census management project structure.

2.145. Preparing and conducting a population and housing census offers the opportunity to exercise statistical leadership by promoting official statistics and their use for development at all levels of society. As the census usually takes place in most countries only once in a decade, there is a need to carefully take advantage of this event in terms of exercising statistical leadership to the fullest extent possible under the understanding that statistics are numerical facts for statecraft.

Comments GB: Add text on statistical leadership in the development and maintenance of administrative data.

X. User consultation, communication and publicity

2.146. The extent of public consultation, communication and publicity can vary enormously according to type of census being undertaken. Most of this section relates to activities associated with a traditional or combined census with a field enumeration component. Counties conducting register-based censuses should implement a communication strategy with messaging that aligns with the register-based approach.[34]

2.147. A comprehensive programme of communications and engagement for a population and housing census covers three distinct audiences: (a) major users of census data, (b) persons and institutions participating in the census operations and (c) the general public, including population sub-groups. Since the census is a national activity that is completely dependent for its success upon the cooperation and support of the general public and many governmental and local organizations, the entire communications and engagement effort should be developed as a coordinated activity in close coordination with the other substantive preparations for the census. These communications and engagement activities are valuable not only for informing others about the census but also for providing census authorities with early and continuing information about the reactions to census plans and

[34] Link to UNECE guidance on Communications for RBC countries

activities of the general public in various parts of the country and of key population groups and institutions. Many National Statistical Offices use design thinking approaches to create better respondent experiences. This involves developing user personas to segment population groups and better understand differing needs, expectations, behaviours.

2.148. Extensive consultation with users and advocacy groups with an interest in census data on topics, on definitions and, particularly, on planned tabulations and other outputs and the development of the census database is an indispensable step in the preparations for the census, which should be taken early. These consultations will assist the census authorities in planning for a census that, within the resources available, is as responsive as possible to user needs in terms of the collection, processing, tabulation, storage and availability of meaningful statistics. Such consultations can also serve to foster a wider and more informed understanding of and support for census plans and activities. The users to be consulted should be from governmental departments, ministries, universities and other research institutions, the private sector, and other organizations (or individuals) representing the economic, social, educational and cultural life of a country. Many countries will want to include among the groups to be consulted organizations or bodies representing culturally diverse communities, ethnic and indigenous communities, and religious groups, those with accessibility requirements and people with disabilities, housing associations and agencies supporting homeless people. Other key stakeholders may include partners with whom the census office collaborates for the provision of specialist services, and donors who may help fund elements of the census operation.

2.149. Taking into account the importance of the census in providing data for local planning and administration, it is also often advisable to have consultations with users in provincial and local governments and institutions in various parts of the country, particularly in remote areas. In large countries or countries where the provincial or local governments have a comparatively high degree of autonomy, consultation with users at the subnational level is essential if the full potential of the census is to be achieved. Strategies should be chosen according to the target group.

2.150. The consultation process can take many forms. If done in the form of meetings, it is often more useful to hold separate consultations with different types of users with common interests, such as administrators, policymakers, planners, demographers, researchers, users in the business community and so forth, rather than a simultaneous consultation with all data users. Consultations involving different types of stakeholders in the same setting frequently prove frustrating to participants because there are substantial differences among users in their technical background and in their concern with the details of census content and operations.

2.151. Meeting data users in person is very informative but imposes physical and budget limitations. Broad consultation can be conducted online and via social media, e-newsletters, and digital stakeholder outreach. The strategy can be used both to collect suggestions from users and also to provide transparency in the census preparation activities. Other forms of technology can be considered for holding decentralized or remote consultations. Users may be sent an electronic questionnaire to collect their priority information requirements, or invited to complete the questions online. Digital channels, including social media engagement, offer the opportunity to reach a substantial number of users if accounts have adequate reach. Video content may be employed to support or promote consultation materials, and webinars can be an effective way to meet with large numbers of users remotely.

2.152. In the context of conducting a census, it is essential to establish two-way relationships with key stakeholders at the lowest territorial disaggregations, such as municipalities. These relationships will

facilitate continuous communication and awareness throughout the entire implementation and development process of the census. Additionally, a meticulously crafted message tailored to the social context of the region must be developed to effectively introduce the census to various interest groups. This lays the foundation for ensuring the entry of field personnel into information sources, which is essential for successful awareness-building and data collection.

2.153. Once these relationships have been established and the census operation has been effectively presented, it becomes necessary to focus on structured dissemination of census information. This becomes a pivotal element in mobilizing society towards the production, understanding, and utilization of statistical information that will benefit the country. To achieve this purpose, efforts to raise awareness must be strengthened, aiming to prevent definitive refusals and ensuring the acquisition of information with the required coverage and quality. In this way, the objectives of collecting accurate and valuable data for the benefit of all can be achieved.

2.154. In order to complete the preparatory work for the census and to carry out the census enumeration itself, the census office of country conducting a census with direct enumeration will have to expand its staff substantially. In addition, numerous governmental and non-governmental organizations outside the census office may be called upon to provide personnel, expertise, equipment, supplies, space, transportation or communications facilities and so on to help in the census work. As a result, large numbers of temporary personnel will have to be recruited and trained (see paragraphs 2.119–2.124) and the contributions of a diverse group of national and local organizations will have to be effectively mobilized. A well-planned recruitment campaign and programme of engagement with external organizations can contribute to both efforts.

2.155. The training and development of field personnel play a pivotal role in the census process and necessitate effective communication, much like the communication with various data users. To achieve this, it is imperative that data collectors comprehend three key aspects. Firstly, they must grasp the significance of statistical information production for the nation and the importance of Censuses, along with a comprehensive understanding of the context in which field operations will take place. Secondly, those responsible for data collection must receive guidance, conveyed through communication, emphasizing the importance of a commitment to data collection quality. Lastly, communication must ensure that census enumerators are well-versed in data collection forms, general statistical operation concepts, and, specifically, those related to cartography and the appropriate use of mobile data capture devices in countries employing this data collection method.

2.156. An effective communications strategy, together with far-reaching publicity and information campaigns, play an essential role in ensuring the success of the census. This is especially so for those countries adopting a field enumeration methodology, either wholly or in part, where the general public is expected or required to participate actively in the census activities as respondents. Particularly in the case of countries that undertake a significant field operation, public acceptance and cooperation is essential to ensure the success of the census. It is crucial to underscore the vital importance of securing active participation from each data source or qualified informant in the census process. This translates into an imperative need to avoid definitive refusals to ensure the acquisition of information with the required coverage and quality. In specific cases where conventional communication channels may not be accessible, establishing a more personalized form of communication with the individuals to be enumerated becomes essential. The aim is to ensure the presence of qualified informants when conducting the questionnaire. It is important to differentiate between a recruitment campaign to hire staff for both office-based and field-based census roles, and a general campaign to drive participation.

2.157. As part of the communications strategy, countries should consider both proactive and reactive media management as well as planning engagements across multiple channels. The use of a call centre to deliver help and support via telephone and providing chat support for online enquires should be considered. Consideration should be given to the appointment and training of census spokespeople to communicate key messages consistently in local media and through community engagement activities.

While there may be occasions when door-to-door communication is necessary, a large-scale publicity and information campaign is recommended to inform the population of the census and to explain its purpose. Opportunities may be taken throughout the campaign to process feedback from the general public, particularly through social media, both to help shape future messaging and where applicable to inform the census operation. Designing and implementing the publicity programme is best undertaken by experts in the field of public relations, advertising and sociology. Such expertise is frequently not found within the national statistical office itself, and it may therefore be appropriate to outsource some or all of this work. The publicity programme may include:

(a) The development of a census slogan and logo)
(b) A public relations campaign;
(b) A community liaison (or outreach) programme;
(c) An advertising campaign;
(d) Monitoring of public opinion;
(e) Media relations, including monitoring of the mass media
(f) A social media strategy.

2.158. There are several main messages that census agencies will need to communicate to the public in order to optimize outcomes for the census. Census publicity campaigns should encompass a wider set of messages, whose components might include (a) making the public aware of the census; (b) educating the public about the benefits (to them and to the country) of the census, in order to improve sentiment towards participation; (c) reminding people about their legal obligation and duty to take part in the census; (d) mitigate the risks of digital exclusion by explaining to the public what to do and when, and providing information on the practical support available to those who may face barriers to responding; (e) informing and reassuring the public that privacy and confidentiality will be protected; and (f) expressing thanks to the public for taking part in the census. Care is necessary in finding the correct balance between these different messages. For example, an overemphasis on the obligatory nature of the census may serve to reinforce negative perceptions that the census is an imposition by the State on the population, rather than an activity for the common good.

2.159. Publicity for a census operation entails an educational campaign, the purpose of which is to ensure the awareness, interest and cooperation of the general public. The aims, as a general rule, are not only to dispel any anxiety regarding the purposes of the census but also to explain the reasons for the various questions in the questionnaire, including the value census information provides by contributing to an informed data driven society, and to offer some guidance on how questions should be answered. The publicity campaign may also be an important tool for increasing the completeness of census coverage, particularly among hard-to-reach groups. Planning for the general publicity campaign should start as soon as the census is authorized. The campaign itself should be closely synchronized with other census activities and full-scale publicity should not begin too many weeks in advance of the date on which enumeration is scheduled to start. Plans for the publicity programme should be closely coordinated with those for the census tests (see paragraphs 3.110–3.114). The programme will have to provide the publicity needed to carry out the census tests. In addition, the programme can use these

tests to study the impact of alternative publicity materials and methods. If either the cartographic or house-listing operations require extensive fieldwork and widespread contacts with the public, it should be recognized that personnel involved in these activities often provide the public with its first impression of the census. Training and publicity programmes should take this into account.

2.160. To extent the census communications budget allows, the general campaign should be directed to all sections of the country and all segments of the population through the use of all available publicity media, with special emphasis on the use of digital media channels. The general campaign may be supplemented by a number of specialized campaigns aimed at specific segments of the population to raise awareness on specific subjects, such as gender, migration, ethnicity and disability, in which the quality of response may depend on the level of prior awareness in the general public or among the specific groups concerned. In multilingual countries, creating multiple campaigns in the various languages used within the country (both official and non-official) is crucial to ensure all communities have the opportunity to understand and participate in the census. Publicity can be secured at low or no cost through social media and through public relations with national and local mass media, such as newspapers, television and radio. It is recommended that a range of spokespeople from the census organization are trained and made available to speak to national and local media, and as well as for community engagement activities. Paid-for advertising through various forms of media can also be effective. Interactive media may also be used such as a toll-free helpline and text messaging. Local events where the public can participate can also improve public awareness and build trust.

2.161. Disseminating information about the rationale of the census and its utility helps allay possible misconceptions among the general public, thus increasing participation and coverage. Outreach campaigns involving a range of organizations and enlisting the support of local leaders and opinion makers to spread the word about the census in their area of influence are also good strategies. Maintaining ongoing relationships with community organisations and representatives in the period between censuses can be of mutual benefit, with the potential to improve engagement not only during the census operation but also during the consultation and dissemination phases. In addition to recruiting such organizations and leaders as partners, the census organization may develop key messages, web content and links, printed material and other material to support them in their activities. The use of publicity may also be considered to support the recruitment of field personnel.

2.162. For countries running a traditional census operation, it is recommended that a census "brand" be established, including a logo and tag line. Census branding is important for establishing awareness and trust. A simple but effective slogan and distinct logo can be used in all national and local advertising campaigns and in all types of media, booklets, posters, brochures and other promotional materials.. A slogan and logo that are well recognized from the initial stages of the publicity campaign may serve to improve "brand recognition" for the census[35]. The aim should be to encourage the respondent to feel more reassured that the census is an inclusive and beneficial activity.

2.163. Special attention should be given to making communications as inclusive as possible in order to improve the quality of responses across the country and across different population groups, as well as to bolster trust in the resulting data. In essence, the aim of these is to engage, educate, explain and encourage, and (if necessary) to enforce participation. Students (particularly older students living away from home), young people (particularly those in urban areas), older people, persons living with

[35] for further guidance see CES guidance (add link)

disabilities, and recent immigrants are among population groups that are generally hard to reach. Other groups that may need to be specially targeted included the homeless, people with literacy and language difficulties, and inhabitants of inner cities, high-rise apartment dwellers, and dense urban areas.

2.164. In rural areas, particularly where digital channels are limited, weekly markets, fairs and public festivals are a good opportunity to publicize the census message among people who may not have much exposure to mass media. An excellent opportunity exists to create widespread awareness of the census through a campaign targeted at schools. Other kinds of local-level publicity, such as wall writing and village announcements, can be planned according to local circumstances.

2.165. For countries taking a digital-first approach to collection, three additional priorities will inform communications and engagement planning across the census operation. Firstly, a website must be built and tested. Design of the website should be informed by user research to ensure users have a straightforward experience in accessing and submitting the electronic questionnaire. The website must be equipped to securely host very large numbers of users concurrently. Secondly, codes permitting access to the electronic questionnaire must be supplied to all households and other residential establishments, with accompanying messaging to facilitate understanding of when and how to complete the electronic questionnaire, and to provide reassurance about the security and privacy of their responses. The general campaign should support these messages. Finally, support should be made available to anyone who may be unable to complete the questionnaire online, for whatever reason. Materials supplied to households should clarify what alternative routes are available to individuals (e.g. completion by telephone, or the option to request a paper questionnaire). Real-time management information (MI) from online responses (and slower MI from posted questionnaires) can be used to deploy the field force more flexibly and responsively towards non-responding households.

2.166. Census organizations should develop a framework for managing misinformation and other incidents that may impact on census operations. Communications staff should be included in strategic discussions during operations to support the organizational response to issues as they arise. During operations public opinion and mass media should be monitored to assess the effectiveness of publicity campaigns and identify such issues. Public opinion could be monitored through surveys that can provide information on public attitudes to the census, or through interaction on digital channels including social media. Media monitoring refers to tracking, analysing, and measuring various media channels. These channels can include both traditional and digital media. Analysing the data from these sources helps NSOs understand public perception and identify emerging challenges related to the census. Such monitoring implies an ongoing accumulation of information, detection and prevention of the development of negative published comments on the census, and preparation of adequate responses to negative reports and information. Increasingly the media has a significant influence on people's behaviour and even minor distractions and mistruths can have a detrimental effect on the outcome of the census. Therefore, in developing their publicity campaigns, national statistical offices should give particular attention to preparing for unexpected events (such as negative attitudes, malicious lobbying, technical difficulties, delays and misleading information). 'Lines to take' should be developed and updated to respond to emerging issues, and may be used on social media directly in response to misinformation on such platforms. Direct communication with social media companies and establishing procedure to work with these companies, as well as points of contact, is also encouraged in advance of a census operation, to develop processes for identifying and managing misinformation. Public relations with mass media can also be used to challenge or alert the public to misinformation. It is also recommended that all official participants involved in census operations know their roles in the communication process both with the media and with the public at large.

2.167. An integral part of census communication and publicity is informing key census data users and the general public about the availability of the census results and their utility (see paragraphs 1.19–1.37). Awareness about the available census data and products should be raised during the intercensal period before the commencement of the next census. This is to make sure that the public recognizes the importance of the census and appreciate statistics that are generated from it. It is critical that such communication strategies be developed as an integral part of census planning and not left as an optional add-on. It has been the experience of quite a few countries that the engagement of professional media and communication personnel adds value to the campaign.

XI.Census calendar

2.168. An indispensable element in the planning of a census is a calendar or timetable indicating the sequence and estimated duration of each of the component operations of the census. At the early stages of census planning, a provisional calendar of selected key dates should be prepared as an overall framework for the census. The calendar must be shared with stakeholders in advance for advice and support. The calendar should be revised and made more detailed as planning proceeds, with the aim of establishing final dates as soon as practicable.

2.169. Such calendars are essential, since they indicate the dates on which each of the numerous operations that make up a census are to be started and completed, and they serve as a guide for measuring the progress of each stage of the census operation. Some census program managers may refer to a census calendar's main elements as the "Census Critical Path", or a series of milestone dates are essential to meet and which significantly impact other elements of the program. Serious delays in work, or errors in time estimates, can be detected by comparing the calendar target dates with the actual dates of each operation. A census calendar is a very efficient instrument not only in the timing control of each census operation but also in the control of the complex of all census operations that are interdependent. Therefore, when modifications in the census timetable are necessary, all related operations should be taken into consideration in order to avoid disruptions in the whole census programme. Obviously, the time schedule will differ for each national census depending upon the general census plan and the resources that are available.

2.170. For censuses conducted with direct enumeration, the census calendar usually shows the various operations grouped into three broad sectors: (a) pre-enumeration, (b) enumeration and (c) post-enumeration. The last-named sector includes evaluation and analysis as well as processing and dissemination. The basic date on which the census calendar and the scheduling of all other operations hinge is the starting date for the general enumeration of the population. For purposes of control, many operations that in fact overlap are shown separately in the calendar. Census calendars sometimes take the form of a chart or graph, in addition to a detailed checklist of operations. Project management software may help in the preparation of the census calendar.

2.171. In establishing the census calendar, it is necessary to consider the relationship of the population and housing censuses to one another as well as to other statistical projects or other large-scale national activities. Although a joint population and housing census operation is likely to constitute, for the period of its duration, the major statistical undertaking of the government, care should be taken that it does

not interfere unduly with the other regular statistical activities that may be going on at the same time. A balanced statistical programme should avoid having too many simultaneous competing enquiries, which might place too heavy a burden on the statistical services and on the public, with a possible resultant loss of both administrative efficiency and public cooperation.

2.172. It is often useful to draw up a comprehensive diagram showing the sequence, interrelationship and timing of all the various steps in the census programme – a Gantt chart and a critical path would be a good example.[51] This type of analysis often reveals the consequences of a delay at one step in terms of delays at other steps in the programme. It can therefore be a useful instrument against which the actual progress of the census preparations may be compared. Indeed, some countries have attempted to use such critical path analyses not only as an aid to census planning but also as a tool for the ongoing management of their census operations. In these instances, it is essential to establish procedures for revising the critical path analysis in response to actual progress. It should be stressed, moreover, that the usefulness of such devices depends on how soundly they are designed, applied and understood. Project management software can be useful in linking the diagrammatic structure of census operations with information about nodes or centres of responsibility for individual broad or detailed operations so as to control the chain of responsibility. Alternatively, event calendars can provide a broad view of the steps of the census programme and allow follow-up. Different tools can be found on the Internet for download or online use. Online versions allow immediate update and make it easier to work in a group but are dependent on Internet access. Other tools, commonly referred to as groupware and collaboration software, as well as Internet and social media forums, can support census operations by providing an environment for exchange of information, files, and data among dispersed teams. A clear monitoring and evaluation plan of the calendar should be in place and regularly reviewed.

2.173. For countries transitioning from a traditional census with paper questionnaires to one with electronic questionnaires, the census calendar should take into account the need for a sufficient period in the pre-census stage of system development. This is because the process of designing and testing the collection systems is more complex and lengthy than that of paper-based questionnaires. Therefore, it is important to allocate enough time for this stage to ensure a smooth transition to electronic questionnaires.

2.174. The census calendar must also allocate sufficient time for the preparation of geospatial information that will support the census operation. The preparation of up-to-date census maps is a critical and time-consuming process that requires careful planning and execution. Sufficient time allocated for map preparation ensures the accuracy, completeness, and effectiveness of the census process, leading to reliable and actionable data for decision-making.

XII.Human resources management

2.175. Traditional census taking requires a large number of people to function properly. Early arrangements are necessary to secure the proper number and type of personnel required for each of the various census operations. While the preparatory and processing work generally calls for office employees possessing or able to learn certain specialized skills (cartographers, coders, data entry operators, programmers and so on), in a traditional censuses the enumeration stage usually demands a large number of persons capable of going to their assigned urban or rural enumeration areas and collecting the information according to specific definitions and instructions. The number of enumerators

required being quite high compared to normal staff strengths, and the period for which their services are needed being rather short, the method of recruiting them needs to be worked out carefully in advance to facilitate quick, simultaneous and transparent hiring, and subsequently compensating them and relieving them of their duties promptly and efficiently. Consideration should be given to computer skills if electronic means of enumeration are going to be used. It is essential that the enumerators and, to the extent possible, their immediate supervisors be conversant with the languages or dialects of the area in which they will be working, and the knowledge of these languages should extend beyond any official languages that the nation provides services in. In addition, attention should be paid to efficiency, economy, experience, ability to read maps and communication skills in general. It is only prudent to recruit and train sufficient reserves to take care of any attrition that may occur in the process.

2.176. Once the cartographic preparations are substantially complete and the questionnaire has been sent for printing, perhaps the single most important means that the census authorities have for influencing the success of the census is the training programme. The contribution that a well-planned and executed training programme can make to the quality of the census results cannot be stressed too strongly. Such a training programme must of course focus on the widely dispersed and difficult-to-supervise field staff (namely, the enumerators and their immediate supervisors) but it must also cover others (for example, the higher-level supervisors, editors, coders and computer operators). Giving all office employees who are working with the census preparations a brief, uniform basic training of all aspects of the census has two prime advantages: first, all personnel understand the importance and the context of their part of the task; and second, since they are conversant with the basics, they can be swiftly deployed in the field for supervision or coordination during the actual census operations whenever and wherever needed. The nature of the training can be traditional (classroom style), e-learning (self-training or live training via internet) or some combination of the two.

2.177. The entire census training programme should be designed to cover each phase of the work and provide an efficient and consistent means of effectively equipping large numbers of fresh employees with the necessary skills. The programme will need to correspond closely to the needs of the various operations and, where appropriate, may include both theoretical and practical instruction, with emphasis on the latter. In the case of the enumerators and their immediate supervisors, the training is most effective if it includes several opportunities for the trainees to participate in practice interviews and role-playing exercises, including the use of adopted IT solutions, if any. (In countries in which multiple languages are used, the method and content of the enumerator training programme will need to be suitably adjusted. For example, if the questionnaire is printed in another language, provision will have to be made for instructing enumerators on the correct formulation of the census questions in the vernacular.) Enumerators and supervisors should be trained as close to the field operations as possible so as to avoid recall lapses. This leaves very limited time for conducting the training. Therefore, the logistics need to be worked out carefully in advance. The training programme for editors, coders, operators of data recording equipment and so forth should also provide opportunities for the trainees to practise under the supervision of the trainers. The intermediate- and higher-level technical staff, such as programmers and system analysts, should also be given special training with emphasis on recent technical developments of relevance to the forthcoming census and on the interrelationships among the various aspects of census plans and operations. Thorough training in census practices is an extremely important component of quality assurance. Detailed and clear documentation of instructions with appropriate illustrations is a basic requirement in this regard. A proper training methodology and a variety of training aids would go a long way in enhancing the training effort.

2.178. The organization and conduct of training courses should be entrusted to those having the necessary qualifications to carry out this task successfully, taking into account not only their professional abilities but also their ability in teaching. This means that staff in charge of training should have certain qualifications that will enable them to stimulate the interest of trainees and to transfer the required knowledge, since otherwise well-qualified technical personnel who are unable to transfer their knowledge to the trainees in a satisfactory manner will be unsuitable as instructors for group training activities. This must be taken into consideration when selecting instructors and it is recommended that objective criteria should be used. In practice, however, it is difficult to find the necessary number of instructors who have both the professional and the teaching qualifications; for this reason, the instructors selected should themselves undergo training in how to organize and conduct training courses. The use of professionally designed training guides can add immense value to the training effort. The involvement of experienced professional experts in the design and delivery of training programmes is also very useful. It should however be noted that the content should be the responsibility of the census authorities and not that of outside experts.

2.179. It is important that training manuals for each training programme are made available to the census organizers and training instructors. Such a manual would be a valuable guide and would help considerably in the efficient training of census staff. It would also contribute to the uniformity of training, which is an essential factor for a successful enumeration, taking into account the great number of census instructors who will be engaged in training. Simple audiovisual aids (for example videos, posters, audio recordings) can also be used to help make the training more effective and uniform throughout the country. If available, new multimedia technologies can facilitate the provision of training at distant locations and be effective and efficient supplementary tools for training. Standardized training may also be provided in e-learning format on the Internet and on handheld devices.

2.180. It is very important to determine the time required to train staff for the various aspects of the census. This depends on several factors, such as the task for which they are being trained, the complexity of the content, the educational level of trainees, the number of instructors available and the funds available. Apart from fixing the number of days for training, it is also important to allocate appropriate time for each subject. Drawing up lesson plans for each session of training is an effective way of ensuring that all subjects are covered, with the right amount of time being devoted to each.

XIII.Logistics management

2.181. A population and housing census differs in many respects from other statistical operations. It requires efficient communication between many different components, including the procurement operation and storage of a large variety of items, most of which have to be distributed to all geographic areas of the country and then recollected.

2.182. Logistics management is a process of planning, implementing and controlling the flow of census materials and equipment needed for implementation of census operations. Logistics planning requires careful coordination between different phases of the census operation, such as mapping, training, field enumeration, data processing and dissemination. The scope of the logistics programme usually differs from one country to another, but mostly covers the following activities: (a) renting central and field offices; (b) installing furniture and equipment; (c) providing help desk support; and (d) delivering and collecting all census materials, including manuals, questionnaires and publicity materials.

2.183. National statistical offices may need to establish a special team for planning, implementation and controlling the logistics programme. The functions of this team should be clearly determined in order to avoid overlap or omission of any activity. During the planning phase of the logistics programme, outsourcing of some activities should be carefully examined as an option.[52] In the context of census logistics, procurement plays a particularly important role throughout the whole exercise.

A. Procurement management

2.184. Developing a strategic approach to procurement is another particular element for successful implementation of a census operation. Taking into consideration the complexity of the process, procurement planning requires logistical coordination with multiple census activities and counterparts. Proper planning contributes to efficient procurement processes, and reduces the risk of confronting problems that may lead to additional costs and delays. Procurement planning is a complicated exercise and emergency work and last-minute operations are usually unavoidable. However, the benefits of procurement planning early in the census operation usually outweigh the disadvantages.

2.185. In the context of censuses, procurement planning entails the process of assessing and projecting the procurement needs of census operations. Needs assessment, cost estimation and requirement definition are the first steps in the procurement process, and are essential components in procurement planning. The purpose of requirement definition is to identify the precise needs of the census operation and to search for the best solution to meet those needs. The needs must be described in the requirement definition in a way that will facilitate the procurement process. The requirement definition is often done in parallel with supplier sourcing and market research in order to let information from the market research influence the requirement definition.

2.186. Procurement practices vary greatly among countries; therefore, no universal system of procurement management can be suggested. However, a few generally accepted procurement principles can be noted. First is the principle of *best value for money*. Best value for money represents an optimal combination of technical and financial attributes – that is, the balance between price and performance that provides greatest overall benefit under the specified selection criteria. This does not necessarily mean selecting the lowest initial price option, but rather represents the best return on the investment, following a proper evaluation of offers under appropriate criteria contained in the solicitation documents. It requires an integrated assessment of technical, commercial, organizational and pricing factors in light of their relative importance. Best value for money can include non-cost factors such as fitness for purpose, quality, service and support, as well as cost-related factors such as price, life cycle costs and transaction costs associated with acquiring, using, holding, maintaining and disposing of the goods or services. The principle of best value for money should be applied throughout the procurement process in order to attract the offer that most effectively meets the stated requirements of the census operation.

2.187. Second is the principle of *effective competition*. Effective competition is best explained as a situation in which at least three independent contractors acting on their own (that is, not in collusion with each other) effectively compete for the same business opportunity and each submit a responsive bid. The procurement processes should foster effective competition as a means of ensuring fairness, integrity, transparency and achieving best value for money. The competitive process should, as

necessary, include (a) procurement planning for developing an overall procurement strategy; (b) market research for identifying potential suppliers; (c) consideration of prudent commercial practices and applicable national regulations, rules and procedures relating to procurement; and (d) formal methods of solicitation, utilizing invitations to bid or requests for proposals on the basis of advertisement or direct solicitation of invited suppliers; or informal methods of solicitation, such as requests for quotations.

2.188. Another important principle is fairness. The manner in which the procurement process is carried out must give all stakeholders the assurance that the process is fair. The concept of fairness includes that the procurement process should be free from favouritism, self-interest or preference in judgment. The assurance of a fair process promotes transparency, a principle that ensures that timely information about existing conditions, decisions and actions relating to procurement activities and about procurement policies, procedures, opportunities and processes are clearly defined and made known simultaneously to all interested parties. A transparent system has clear rules and mechanisms to ensure compliance with those rules. A transparent system of procurement further ensures that procurement records are open, as appropriate, to inspection by auditors. In line with the procurement principles of transparency, every step in the procurement process should be documented and kept on file, preferably electronically as well as in hard copy.

B. Forward and reverse logistics

2.189. The type of census materials would differ depending on census methodologies and technologies used for enumeration and data processing. However, in a traditional census, any kind of materials related to the fieldwork has to be supplied to the field staff and returned. Strategies for distribution and return of the materials should be carefully planned according to the type of materials, volume and final destination of delivery.

2.190. As a first stage in this process, decisions must be made concerning the nature and responsibilities of the centrally controlled distribution and return operation. For example, a decision should be made about what geographic levels the materials will be distributed to – regional office, local census committee, supervisors or other. These decisions must be made by countries, keeping in mind the impact of the amounts of material to be transported, the transport available to field staff and the condition or existence of roads or other means of transport. Once these decisions have been made, the key inputs to the dispatch and return of materials are as follows:

a. Workload estimates from the mapping programme to establish packing volumes for transport requirements;
b. Name and address details for delivery and pickup points.

2.191. During the design of enumeration areas and mapping activities, an estimate will be made of the number of enumeration areas, and the amount of work in each. This information can be used to calculate how much material will be needed by each enumerator, supervisor, manager and so on. This method should provide a more accurate estimate of the total volume of all the materials.

2.192. The majority of these tasks are usually carried out under contract by a government transport service or commercial operator. The contractor will use specifications and consignment details provided by the national statistical offices. The postal service may also be a feasible method.

2.193. A significant task in planning field operations is establishing the specifications for the packing and transport of materials. These specifications need to be developed regardless of whether these activities are carried out by the census agency itself or contracted out to another government agency or private company.

2.194. The role of the national statistical authority with regard to dispatch and return tasks is primarily one of liaison and monitoring. For the most part, the contractor will contact regional managers and supervisors directly about the delivery or pickup of material. The national statistical authority can expect to be involved as a liaison between the contractor and field staff in the early stages of the operation or if there are any particular problems experienced by either field staff or the contractor. The national statistical authority management staff should meet frequently with the contractor to discuss the operation and liaison arrangements. Part of the planning of the operation will include arrangements to enable the national statistical authority's management staff to monitor the delivery and return of materials. In particular, when material is picked up from supervisors, the national statistical authority's management staff should maintain a close watch over what is taking place in the field as the transport of completed census forms is involved.

2.195. Distribution and collection of census materials can be monitored through use of a management information system (see paragraphs 3.145–3.146). Timely information can be produced about the progress in delivering the census materials and questionnaires and collecting the completed questionnaires and the other return materials. Depending on the logistics programme, the information needed for monitoring the logistics activities will differ. In general, the following information would be needed for both forward and reverse delivery: (a) type of materials, (b) timing of delivery, (c) number of delivered materials, and (d) name of the persons involved in the delivery. The periodic reports produced from the management information system will be crucial to ensure the timeliness of the fieldwork across the country by giving an alert if there is any delay or any other problems regarding the delivery of census materials.

XIV.Contracting out

2.196. It is a contemporary practice in many countries to contract out tasks or activities of the population and housing census as a way of increasing efficiency by utilizing the advanced methods and technologies not necessarily available in the national statistical office or public sector responsible for conducting the census. At the same time, costs reductions may be achieved through a competitive selection process. However, not all census tasks are appropriate for outsourcing or contracting out, and doing so will not necessarily bring the desired benefit of strengthening national capacities. Census activities may be broadly classified as core and non-core activities. As a general rule of thumb, core activities should not be contracted out. If for some reason core activities need to be contracted out, then it is essential that the strategic control of such activities should firmly be with the census authorities at all times.

2.197. In the context of contracting out components of census operations, the national statistical authority would need to build the capacity to ensure proper outsourcing. This is of primary importance at the preparatory stages, as outsourcing requires a solid and comprehensive knowledge of contemporary technologies and their advantages and disadvantages, as well as past experiences at home or in other countries. Consequently, the national statistical authority would need to plan and develop a particular unit for the purpose of ensuring adequate and efficient outsourcing well in advance of the census itself, as there would need to be extensive testing of the products and services that were contracted out.

2.198. The terms of engagement (scope of work), the deliverables and the timelines should be clearly laid down with definite dispute redresser mechanisms. Illustrative examples of items of work that may be contracted out in the context of a traditional census are as follows:

a. Layout and printing of census questionnaires
b. Development of electronic questionnaire systems
b. Packaging of census questionnaires
c. Dispatch and delivery of census material
d. Census mapping
e. Publicity and public relations
f. Training
g. Return collection of census questionnaires and other material
h. Inventory and storage of filled-in questionnaires
i. Scanning and data entry
j. Data processing and tabulation
k. Publication and dissemination.

2.199. Time is of the essence in all these activities and it is vital that adequate time is allocated. At the same time, backup plans should be in place in order to deal with any failure on the part of the vendors. Fundamentally, census operations are time critical and commercial compensation is secondary. Depending on whether an activity is on the critical path or not, adequate flags should be provided. Milestones and timelines are also essential. The moment there is a failure in achieving any milestone, alerts should automatically be raised. Risk assessment represents a critical component for outsourcing; the risk of failure, and the costs involved in developing contingencies in case of failure, require particular consideration.

2.200. The appropriateness of contracting out should be determined step by step and after subdividing the overall census tasks into stages. In the context of quality management, the outsourcing of components of census operations still requires the national statistical office to take full responsibility for, and manage the quality of, the census data. Throughout the overall process, activities should be conducted by a method (considering accuracy and timeliness of the results) that can best satisfy the general public. No part of the work tasks should be done by a method that may result in loss of trust of the general public. When outsourcing, the statistical office needs to ensure that it continues to be in a position to understand and manage elements that contribute to final data quality. So, in judging the propriety of contracting out, it is recommended that national statistical offices should carefully consider the following criteria:

a. Strict protection of data confidentiality;
b. Method of confidentiality assurance that satisfies the general public;
c. Guaranteed measures of quality assurance;

d. Ability to manage and monitor the outsourced census tasks or activities;

e. Ability of the vendor to stay within the planned budget

e. Control over the core competence of the national statistical office, and appropriateness judgement, considering the specific situation of each country.

2.201. Confidentiality assurance is the first and most important issue that should be considered by national statistical offices. National statistical offices are responsible for data confidentiality, in terms of both perception and reality. It is extremely high risk for national statistical offices to have to manage leakage or misuse of confidential information. Consequently, contracting out of tasks that have the risk of such an incidence should be avoided. For example, in the phase of data gathering, it is highly recommended that contracting out should be avoided because the task is closely related to the earning of trust from citizens and the strict protection of confidentiality. Where temporary enumeration staff are engaged under contract, this should be done in such a way that they are subject to strict measures of monitoring and control by the national statistical office. These enumeration staff should be engaged in such a way that their activities are governed by the relevant statistical legislation to preserve the confidentiality of the data they collect.

2.202. The second important and related issue that should be considered carefully is conveying confidentiality assurance to the general public. As described in the "Essential roles of the census" (see paragraphs 1.1–1.3), a census should be undertaken by the method that can produce the most reliable results and in a manner that ensures the trust of the general public in terms of both perception and reality. If either one of these attributes is not met, then the method used as well as the results obtained may not meet the approval of the general public and could result in the census itself being questioned. Thus, protecting data confidentiality refers not just to the actual protection of confidential data, but also to protecting the perception of confidentiality among the general public and providing a sense of inward security.

2.203. The third significant issue to be considered in outsourcing is the guarantee of quality assurance in the outsourcing environment. The key point is that the national statistical office is satisfied that the goods or services paid for are provided. Cost should not be the first priority in considering and judging the successful bidder in this respect unless prescribed by procurement rules. Although it is desirable to engage in fair competition among several companies to reduce costs, it is worth mentioning that merely considering low-price bidding as a determinant factor may adversely affect the quality of the job to be done by the successful bidder. Low-quality work could cause a significant loss of trust among the general public. To assess the quality of work, as part of the contract allocation process, potential contractors should be required to provide samples of their work (for example, for printing, manufacturing satchels, and other work), or if this is not possible, to list referees who could be contacted to verify their claims or sites at which previous work can be inspected. The contracting process should state all the key requirements for the services sought and bidders should be measured against these. Although not a guarantee of quality, it will minimize surprises. Once the contract has been awarded, continuous monitoring of the progress of work entrusted to the selected company is necessary and the national statistical office should ensure that a system for monitoring quality is built into the contract. Consequently, in considering the proper contracting-out procedures, national statistical offices should also take into account the costs for constructing a system of surveillance for monitoring progress of the work being contracted out.

2.204. In addition to monitoring the providers of goods and services, national statistical offices need to plan for continuous interaction with vendors. This implies an additional step to monitoring and amounts

to a necessity to work side by side on a regular basis in order to ensure the best quality of the products and services and to meet the standards and needs of the census operations. This coordinated work refers to providing technical and technological advice, as well as following the development of the services and applications from the substantive point of view. While the national statistical office may not have the full capability to develop certain products or applications, it certainly possesses considerable technical experience and understanding of producing statistics on a regular basis. Therefore, planning and implementing for a regular and continuous interface with providers when parts of the operations are outsourced needs to be incorporated in the overall planning from the beginning.

2.205. The fourth major issue in outsourcing census activities is the procedure of assessment and evaluation of the capabilities of the candidate providers. A quality assurance framework (for a detailed discussion on quality assurance, see paragraphs 2.169–2.228) and implementation should be established in a first phase of outsourcing. Through this procedure national statistical offices should fully assess both the capabilities and the disabilities of companies in order to select the winner to which the activities in question are to be outsourced. It is highly recommended that practical and financial peculiarities of providers should be considered after the assessment of their capabilities. Each private company has a potential risk of bankruptcy or of changing the field of its activity. It should be kept in mind that if a selected company is unable to fulfil the assigned tasks, the probable problems might not be resolved by applying penalties. However, a very significant problem that could occur is that users might not be able to make use of accurate and timely census results. In such a case, national statistical offices might lose the trust of the general public in the census and even in future censuses or other routine statistical projects conducted by the statistical office. It is, therefore, very important for national statistical offices to adopt a method in which risks are as low as possible.

2.206. Some approaches to outsourcing put an emphasis on a “turnkey” arrangement, by which contractors deliver the system according to a set of predetermined client specifications with the expectation that the client focuses solely on the outputs and not the internal working of the system. This assumes that the national statistical office completely understands and can fully anticipate all data quality issues that might arise during the census and has included these in the specifications. The client is not expected to have any understanding of how these systems work or how they might contribute to the final outputs. Any changes to the system typically require cumbersome processes to determine contractual responsibilities and heavy financial costs. This sort of approach effectively hands over the quality of the census data to the contractor, while the risks associated with intervention remain with the census agency. It removes any flexibility and greatly restricts the ability of the census agency to react to quality problems that emerge during processing.

2.207. Suppliers should be made fully aware of the quality targets at the outset of the census programme, and the quality requirements of the outsourced components that enable the overall census quality target to be achieved. Operational quality control should apply to outsourced services in the same way as those that are not outsourced. For some contracted-out operations (e.g. printing), national statistical offices may want to consider having staff on-site, to quickly judge and correct any quality issues.

2.208. In addition to managing the outsourced activities or tasks, the ability or the flexibility to cope with sudden or unpredicted change in the situation is also very important. It should be mentioned that contracting out does not necessarily mean lower costs; sometimes the burden of monitoring cost, emergency costs and other matters may jeopardize the census. It is recommended that the national statistical offices themselves should do some tasks or activities that are difficult to manage. National

statistical offices should judge and determine whether to contract out census activities from this viewpoint.

2.209. It should also be recommended that for critical activities, such as the coding of education, occupation and industrial classification, special care should be taken to ensure adequate training of the personnel to undertake the task, particularly when it is contracted out. The same amount of care and training is required even when the task is performed by the national statistical office. This is due to the fact that the coding depends on the minor differentiation and level of coding (general and detailed classifications according to different coding standards), as well as the coding manual and the education of the coders. In the light of such subtle criteria for judgement, it is difficult to prepare a complete coding manual in advance before checking the filled questionnaire. If coding is automated, working in advance with the vendor may be necessary, to ensure that automated coding methods are thoroughly tested in advance of the conduct of census field operations.

2.210. Censuses are large operations with massive quantities of data that require coding and editing. To reduce the staff resources required and to improve timeliness, uniformity and accuracy, automated coding procedures may be employed. Some countries have already implemented automated coding procedures for addresses, countries, education, occupation and industry. The development of the application software could be contracted out although the rules to be followed must be carefully specified by the national statistical office, which should retain responsibility for implementing the system. The software application can often be used for other statistical collections undertaken by the national statistical office. When outsourcing, the staff of the national statistical office should be able to modify the parameters of such operations themselves at little cost and in a timely manner. By having this ability, the national statistical office can manage the appropriate balance between data quality, cost and timeliness.

XV.Use of technology

2.211. Technological advances and expanded access to modern technology have enabled significant streamlining of the way in which the business of a population census is conducted. With modern tools, every stage of the census—from planning and monitoring to implementation and evaluation—can be seamlessly integrated. Many facets of census activities can benefit from the use of technology. For instance, remote sensing and imaging technology can generate live maps with coordinates against which enumeration activities can be tracked. Fieldworkers can be paid their salaries and stipends using mobile device technology. Handheld devices equipped with editing and geopositioning functions can improve the consistency of responses while tracking geo-activities, including those that can report on spatial coverage. Given the breadth and extensive capabilities of today's technology, it is crucial to select carefully at the planning stage the technological innovation elements that will be adopted in the census value chain.

2.212. The comprehensive integration of geospatial technology at every stage of the census process has the potential to bring about a transformative shift in census planning, data collection, analysis and dissemination. In the planning phase, GIS can aid in the delineation of enumeration areas, ensuring that no areas are overlooked or double-counted. This precise demarcation facilitates the efficient allocation of resources and fieldworkers. During data collection, geospatial technology enables real-time tracking of enumerators, ensuring their safety and optimizing their routes for efficiency. The geolocation data

can be cross-referenced with responses to validate the data collected and to ensure full coverage, thereby enhancing the accuracy of the data and aiding in real-time monitoring and evaluation of census activities. In the data analysis phase, the overlay of census data on spatial maps can allow for a deeper understanding of demographic patterns, population densities, and other socio-economic indicators in a spatial context, allowing integration with other geo-coded datasets and providing richer dissemination products for data users.

2.213. Rapid innovation has led to both exponential growth and reduction in costs of technology. These changes have seen an increased adoption of technology across many aspects of the census in a variety of different ways. While on the one hand this is providing considerable benefit, on the other it is increasing dependence on technology providers and introducing new challenges and risks. The key to the successful use of technology in a census is to clearly understand the rationale or the objective of the technology introduction and consider a range of key success factors for technology adoption, which may include suitability, security, scalability, stability, safety and skills.

2.214. Understanding of the true value of the technology is critical to formulating a business case and assessing whether or not to proceed with the project, and if so what technology choices to make. As the introduction of technology can be an expensive and risky exercise, it is important to ensure that there is sufficient value in its introduction for each specific census.

2.215. The most common reasons for the introduction of technology in censuses operations are:

- **Efficiency and reduction of costs.** Technology provides an opportunity to reduce the number of personnel involved in different aspects of the census – for example, scanning and character recognition can reduce manual data entry and increase data accuracy, satellite imagery can reduce manual mapping and Internet self-response can reduce fieldwork. Technology can also reduce other non-labour expenses such as printing, freight and travel. In some cases, technology can simplify business processes and thus reduce cost or risk. For example, fieldworker workflows can be simplified through the use of secure mobile technology that allows census takers to submit their availability, receive assignments and navigation instructions, collect and transmit data, and communicate with their supervisors via their mobile devices.
- **Data quality and consistency.** Technology, and in particular the automation of processes, can increase the consistency of census data and reduce data errors – for example, scanning and character recognition reduces data entry errors, and automated validation rules or edits ensure that data are checked and changed in consistent ways rather than relying on manual, dispersed field processes.
- **Timeliness.** Technology can reduce the time needed to conduct census enumeration, data processing, data analysis and preparation of results for publishing. The faster the census data are released, the more valuable the data are to census users, and thus the use of technology for data scanning, recognition, processing and publishing should be considered for its ability to advance publish dates.
- **Public expectations and confidentiality.** The census relies on the general public to provide data. The public's and user expectations, and in some cases legal rights, may make it advisable or necessary to implement specific technologies to support ensure their convenience and confidentiality. This can manifest in various ways, such as offering online census questionnaires, facilitating online job applications for census roles, or providing census results digitally. Notably, there is some evidence that the provision of online questionnaires has had a positive impact on census response rates.

- **Data integration and dissemination.** For census statistics to hold value, they must be accessible and usable by a diverse group of data users, and when possible meet the principles of findability, accessibility, interoperability, and reusability (FAIR)This includes everyone from the general public, private sector and media to government officials, civil society and researchers, spanning a variety of levels of expertise and data needs. By leveraging various technologies, including GIS, the collection and dissemination processes can be enhanced. This allows for seamless integration of census data with other sectors, offers data aggregation and visualization capabilities, and ensures the availability of data in diverse formats tailored to user needs.
- **Event management.** The coordination and monitoring of census enumeration across the complete nation or area has always been challenging and has relied on dispersed accountability and manual processes. Technology provides better visibility, oversight and capacity to monitor performance indicators and respond to enumeration events.
- **Data retention and utility.** The costs of conducting census enumeration are significant and thus attempts should be made to optimize the value achieved from the census results. The use of technology enables census data to be safely managed, secured and retained, as well as opening up opportunities for statistical data integration and data reuse, such as for time series analyses and other analytical needs.
- **Assurance and anticorruption.** The adoption of technology can reduce the risk of fraud or corruption through providing more standardized, controlled and auditable records of actions, for example for the records of financial expenditure.

2.216. Considering the reasons listed above in a business case for a technology project often reveals inherent tensions and necessitates trade-offs that should be considered openly and transparently. For example, a solution that aims to fulfil a broader range of user expectations might entail higher costs. While the approach may benefit user satisfaction, it may negatively impact overall programme efficiency.

2.217. While the conduct of significant technology projects is becoming increasingly commonplace across government, there is limited evidence to suggest that mere technology introduction would deliver the projects on time and within budget. Careful planning and management is critical to the success of these projects. The unique size, timing and nature of a census throw up unique challenges, which need careful consideration. The following factors should be taken into consideration in the context of optimizing the use of contemporary technology for census operations:

- **Suitability.** The functions and benefits of any new technology need to be assessed thoroughly and objectively. Consideration of whether the technology has adequate user-friendliness for the intended user is another important point.
- **Scalability.** Technology solution should be designed, implemented and tested to be able to handle the number of users and the number of data. Hardware and software used in delivering technology must be scalable to the actual load of each business process during the census project.
- **Security.** The security of census data is critical for guaranteeing the confidentiality of respondents' personal information. Census systems, especially an online census, must be secured to protect privacy of respondent data and to ensure respondent confidence in the system. Systems need to preserve confidentiality, integrity and availability of census information. IT security is especially important in a contemporary census, as cybersecurity risks have greatly increased in recent years. National statistical offices with an existing or

planned digital presence should be prepared to invest properly, to minimize all risk of data loss to cyberthreats.

- **Stability.** A traditional census relies on the successful conduct of enumeration in a tight time period. Key pieces of technology – such as the census call centre, self-response portal or self-help website –may be designed in such a way that it does not have an unrecoverable impact on data response rates, particularly in countries using self-response specially during close to census day.
- **Skill.** The census organization, whether planning to outsource or deliver the solution internally, must ensure that it has the adequate skill, knowledge and capacity to lead and manage this project.

2.218. Census technology projects require strong project management expertise and thus adequately competent, experienced, motivated and knowledgeable staff need to be assigned to this role.

2.219. Detailed business requirements (what the technology needs to do and how it needs to do it) need to be developed by a suitably skilled business analyst in order to ensure that the technical specifications and implementation are aligned to the census need.

2.220. Rigorous testing is essential for the statistical or census office when integrating new technology into its processes. It is crucial to understand the various stages impacted by technological changes. Separate tests should be conducted to prove new technologies and identify potential problems linked to their implementation. Depending on the extent and characteristics of information technology, these tests should include all information technology components related to the fieldwork, data transfer or entry and processing well ahead of the actual census. Testing should evaluate the application systems (software) and the equipment (hardware) under real-world conditions, so as to identify and mitigate any potential malfunctions. Continuity planning should also be integrated into the testing phase, ensuring that there are fallback measures in place to counteract any equipment or system failures that could disrupt a smooth census operation.

2.221. A pilot census ensures that technology can be tested as part of a complete, end-to-end business process and highlight major implementations with functionality or capability. The pilot test should be accompanied by activities to test whether technology solutions are scalable, secure, accessible and robust enough for the real census. Stress tests should be conducted during the testing phase to ensure technology use can handle the maximum load of each module or business case and at the same time sustain data integrity.

2.222. Once the implementation of technology in the census process is completed, a comprehensive evaluation should be conducted. This evaluation serves a dual purpose: to identify lessons learned and define areas for improvement to enhance future census operations and maximize the value of the technology. During this phase, the results will be analyzed, and the efficiency and accuracy of the technology will be assessed, taking into account the costs and benefits associated with the implementation. Feedback from enumerators and citizens who participated in the process will be collected and considered. This will help identify specific issues that require solutions and contribute to the continuous improvement of census operations in the future.

2.223. There is an increasing adoption of technology in every national statistical office across the world. A number of nations have been developing their own technology solutions to support different aspects of the census, with some of these products being made available for statistical organizations in

other countries. When making decisions around technology selection or development approaches there should be strong, favourable consideration of existing products within the international census community, and where products do not already exist the preference should be to undertake co-development with other members of the statistics community.

2.224. There are a number of risks or drawbacks associated with the introduction of technology that should be considered and managed, otherwise these could lead to increases in expenditure, delays to the census timetable or impacts on the quality of the census. These include:

- Incompatibility or other integration issues between different hardware and software applications;
- Solution outage or failure (which could be for many reasons – lack of connectivity, hardware failure, battery life, GPS black spots, software bugs, device theft);
- Lack of skills or knowledge among system users, particularly temporary census staff;
- Insufficient or inadequate communication between technology staff and business staff, particularly leading to misunderstanding of requirements;
- Hacking, online attack or other information technology security event;
- Maintaining, upgrading or decommissioning old or legacy systems;
- Lack of documentation or reliance on a small number of key people;
- Huge amount of digital data available, creating a potential distraction for staff.

2.225. A wide range of technologies, covering all stages of the census from planning to data dissemination, is presented throughout all the chapters of these principles and recommendations. However, the integration of technologies in census operation requires taking into account various specific national needs and the value each technology would bring to a specific aspect of the census.

2.226. The utilization of technology is expected to vary considerably in statistical offices across the world, considering the need and the resources that are available to the national statistical system. Ultimately, technology is an enabler that can enhance efficiency, accuracy, speed and transparency of census operations. It can also lead to optimization of costs, depending on the circumstances of each nation. It is to be reiterated, however, that the choice of technology and its level of deployment should be diligently assessed by each country before introduction.

XVI.Innovation to optimize census processes

2.227. Innovation is an important investment for any major project, and censuses are no exception. The world around us is constantly changing, and organisations must innovate to maintain a modern approach to census activities and interactions with the public. Changes in society bring new public expectations, and advances in technology lead to new opportunities. An ongoing program of innovation provides the platform necessary to keep up with new demands and opportunities.

2.228. A current example of innovation using emerging technology is in the field of artificial intelligence. For example, many organisations are using generative text technology for customer service efficiencies such as ‘chatbots’. National statistical organisations are now exploring the application of such technologies for census use, in responding to the large amount of public queries generated by conducting a national census. This has the potential to create an improved experience for the customer,

while reducing overall costs. Similarly, machine learning technologies have proven application in data coding, with the potential to improve efficiency and quality while reducing the manual effort required.

2.229. Innovation should not take place in isolation. Those conducting a census should look to international counterparts to share innovative ideas and learnings from testing and operational experience. Domestically, it is worth exploring how other organisations undertake similar activities, such as large scale operations or working with big datasets. Establishing a local culture of innovation, and investing in long-term research activities, is important to ensure the ongoing modernisation and future-proofing of the census.

XVII. Quality assurance

A. Plans for quality assurance

2.230. Most countries conduct population and housing censuses once in 10 years, so retaining institutional knowledge from one census to the next becomes challenging. But experience from previous population and housing censuses as well as other censuses, such as agricultural censuses, is very useful to plan for a quality assurance and improvement programme for the current census. It is of critical importance that errors detected from previous censuses or similar activities were documented and used as the basis for developing quality assurance measures for the next census. Each country must have a quality assurance and improvement programme in place to measure the quality of each operation in a census. A quality assurance programme should also be viewed as a quality improvement programme as errors detected can be addressed promptly. If data are of poor quality, decisions based on these data can lead to costly mistakes. Eventually, the credibility of the census and the national statistics office may be called into question.

2.231. Quality assurance should be viewed as an important element of any census, regardless of whether it is a traditional, administrative data or hybrid/combined census. The quality assurance and improvement programme should be developed as part of the overall census project and integrated with other census plans, schedules and procedures. The programme should be established at all phases of census operations, based on the United Nations National Quality Assurance Framework. Establishing a quality assurance and improvement system and setting quality standards at the planning stage is crucial to the success of the overall census operation.

B. Quality assurance components

2.232. Census quality management should be viewed within the broader context of managing the statistical system and managing the institutional environment. At the operational level, quality management should be comprehensive and should cover all census activities including planning, mapping, questionnaire development, data collection, processing, evaluation, analysis, and dissemination.

2.233. The United Nations National Quality Assurance Frameworks (UNNQAF)[36] for Official Statistics provides a framework for addressing quality within the context of official statistics and contains 19 principles that are relevant for quality management in population and housing censuses. These quality principles and associated requirements consist of four levels, ranging from overarching institutional and cross-institutional management through to the management of statistical production processes and statistical outputs:

- Level A: Managing the statistical system.
- Level B: Managing the institutional environment.
- Level C: Managing statistical processes.
- Level D: Managing statistical outputs.

1. Managing the census within the national statistical system.

2.234. A precondition for the efficient production of high-quality census statistics is a national statistical system that supports the coordination of census activities with other parts of that system, the management of relationships with stakeholders, and an enabling legislative environment. Due consideration should be given to the following principles:

i. Coordinating census activities within the broader context of the national statistical system.
ii. Managing relationships with data users, data providers, and other stakeholders.
iii. Managing statistical concepts, definitions, classifications, methods, and procedures.

2. Managing the institutional environment.

2.235. Quality will be better supported by sound institutional arrangements that assure the following principles:

iv. Professional independence. Census results should be developed, produced, and disseminated without any political or other interference or pressure.
v. Impartiality and objectivity. The census should produce and disseminate statistics in a way that is professional, impartial, and unbiased.
vi. Transparency. Census policies and management practices are well-documented and publicly available.
vii. Statistical confidentiality and data security. Information provided by respondents will be kept confidential and will be used for statistical purposes only.
viii. Commitment to quality. Census managers should be dedicated to assuring quality in their work, and the work of their teams, and systematically and regularly identify strengths and weaknesses to continuously improve census processes.
ix. Adequacy of resources. The financial, human, and technological resources should be adequately planned for and adequate both in magnitude and quality, and sufficient to meet their needs regarding the development, production, and dissemination of census-based statistics.

[36] https://unstats.un.org/unsd/methodology/dataquality/un-nqaf-manual/

3. Managing census processes.

2.236. Quality is the outcome of sound statistical processes, and deficiencies in quality (for example, delays in processing or lack of accuracy in the results) are usually the result of deficiencies in the process rather than the actions of individuals working in that process. Census processes should be based on the following principles:

x. Methodological soundness, adherence to professional methods and (internationally) agreed standards, guidelines, and good practices.

xi. Cost-effectiveness. Resources should be effectively and efficiently used, and census objectives achieved at a reasonable cost.

xii. Appropriate statistical procedures. Effective and efficient statistical procedures should be implemented throughout the census production chain.

xiii. Managing respondent burden. Individuals who provide data upon which census products are based are the fundamental contributors to the quality of census data. The requirement to collect data should be balanced against the burden placed upon respondents giving due consideration to the trade-off between the length of the census questionnaire and the quality of responses received.

4. Managing census outputs.

2.237. It is suggested that the outputs of any census should follow all the following principles:

xiv. Relevant.

xv. Accurate and Reliable.

xvi. Timely and Punctual.

xvii. Readily accessible and Clear.

xviii. Coherent and Comparable.

xix. Be supported by Metadata.

These six principles are expanded upon below. [paragraphs 2.177 to 2.184]

2.238. In the census context, some attributes of quality assurance may be emphasized over the rest. The census should produce statistics that are relevant to data users. A census is a particularly expensive exercise to undertake and creates a burden on respondents. Therefore, it is crucial to ensure that any demand for data is met to a minimum acceptable level and that topics for which there is little demand are not included on the census form. Consulting with users of census data as one of the first steps in designing the census process is a positive public relation undertaking and an efficient, transparent means of determining the demand for potential census topics.

2.239. 2.178. The *relevance* of data or of statistical information is a qualitative assessment of the value contributed by these data. Value is characterized by the degree to which the data or information serve to address the purposes for which they are produced and sought by users. Value is further characterized by the merit of those purposes, in terms of the mandate of the agency, legislated requirements and the opportunity cost to produce the data or information. In the context of a census the concept of fitness for purpose as a measure of relevance is important. If it is only necessary that data are available at the broad level (for example, national or major civil division level of geography; broad demographic level), user requirements could be met more cheaply and effectively through a sample survey.

2.179 Statistical agencies should develop, produce and disseminate statistics that accurately and reliably portray reality. Accuracy of data or statistical information is the degree to which those data correctly estimate or describe the quantities or characteristics that the statistical activity was designed to measure. Accuracy has many attributes, and in practical terms there is no single aggregate or overall measure of it. Of necessity these attributes are typically measured or described in terms of the error, or the potential significance of error, introduced through individual major sources of error, for example coverage, sampling, non-response, response, processing and dissemination.

2.240. *Timeliness and Punctuality* of information reflects the length of time between its availability and the event or phenomenon it describes, but considered in the context of the time period that permits the information to be of value and still acted upon. It is typically involved in a trade-off with accuracy.

2.241. *Readily accessible and clear* reflects the availability of information from the holdings of the agency, also taking into account the suitability of the form in which the information is available, the media of dissemination, and whether the user has reasonable opportunity to know they are available and how to access that information. The affordability of that information to users in relation to its value to them is also an aspect of this characteristic.

2.242. *Data coherence* reflects the degree to which the census data can be combined with other statistical information within an integrated framework over time. The use of standard concepts, definitions and classifications promotes coherence. Equally important is *internal coherence of data*, referring to the consistency of information across different topics of the census and census outputs. This coherence is usually addressed through the meticulous development of data edits. Data are most useful when *comparable* across space, such as between countries or between regions within a country, and over time. More and more emphasis is also put on enabling comparison of geography over time, as well as maintaining consistency and comparison of census topics from one census to another.

2.243. *Metadata* should be provided covering the underlying concepts, the variables and classifications used, the methodology of data collection and processing, and indications of the quality of the statistical information to enable the user to fully understand census statistics and their limitations.

C. Need for a quality management system for the census process[37]

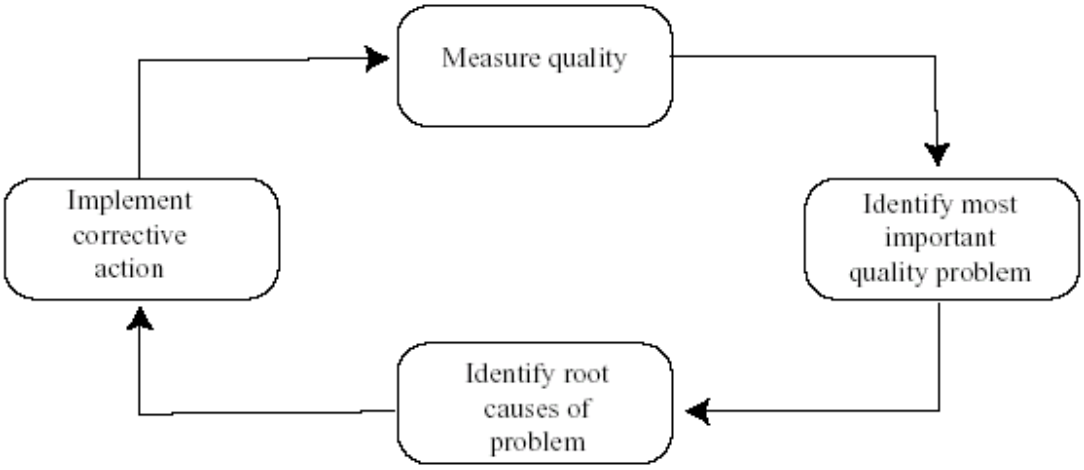
2.244. The essential quality attribute of relevance of census output, and how to assure it, has been discussed above together with the need for consideration of accuracy, timeliness and cost. Quality is relative, and in the end is based on what is acceptable or fit for the purpose, rather than a concept of absolute perfection.

2.245. Deficiencies in quality (for example, delays in disseminating output) are usually the result of deficiencies in process rather than the actions of individuals working in that process. The key to quality assurance and improvement is to be able to regularly measure the cost, timeliness and accuracy of a given process so that the process can be improved when a fall in quality is indicated. The focus of quality

[37] This section draws heavily upon material in *Handbook on Census Management for Population and Housing Censuses*, Studies in Methods No. 83 (United Nations publication, Sales No. 00.XVII /Rev.1), Chapter 1C.

assurance is to prevent errors from reoccurring, to detect errors easily and early and inform the workers so that they do not continue making them. This simple feedback loop is represented in figure 1.

Figure 1. Quality assurance circle



2.246. Being iterative, the quality circle is particularly applicable to tasks that are highly repetitive, such as the processing phase of the census. However, the general principle applies to all processes. For example, there is less opportunity to evaluate performance, identify problems and implement corrective actions in phases such as enumeration owing to time constraints, the once-only nature of some of the processes and communication issues. However, it still can be established with careful planning and documentation in advance of the census.

2.247. It is important that a complete evaluation takes place and is documented at the end of each phase of the census. This is critical to ensure that the organizational learning inherent in the quality circle is carried forward to the next census.

2.248. Since people play a key role in most census processes, they are in a good position to identify problems with quality and provide solutions. Quality is therefore not just the outcome of mechanical application of predetermined measures but relies on a combination of:

a) Established, documented processes, including quality targets (for example response rates, level of error in processing);

b) Systems to monitor the outcomes of these processes;

c) Active encouragement by management to involve staff undertaking the processes in identifying and resolving deficiencies with quality.

2.249. While elements of the quality circle, such as mechanisms to monitor quality, may have some superficial resemblance to some of the elements of traditional quality control approaches, they are

quite different.[38] Traditional quality control is based on correction of error after the event, whereas the emphasis of the quality circle is on improving the process that caused the "error", which may be any of the cost, timeliness or accuracy attributes falling below specified levels. A simple error correction process may suffer from any of the following:

(a) It adds significantly to the cost or length of the operation;
(b) Errors in the inspection process can fail to detect true errors or falsely identify errors;
(c) The correction process can introduce errors into the data;
(d) Operators take less responsibility for the quality of their work, believing it to be the responsibility of the inspectors;
(e) Where a sample of units is inspected, the quality of data is only ensured for those units that are inspected.

2.250. The emphasis should be on process improvement rather than correction. Therefore, an important aspect of quality management may be to not correct errors detected through the quality monitoring process unless they are of a severe nature or are generally applicable. For example, a generally applicable error could be a systems error that miscodes every occurrence of a common event. Resources are thus better focused on improving processes and thus overall quality.

D. The role of managers

2.251. Managers have a vital role in establishing quality and quality management as a priority in the census. The biggest challenge to managers is first to establish a culture within the census agency that has a focus on overall quality management issues and to obtain the commitment of staff to strive to achieve high-quality goals. At the same time, managers need to be aware that to achieve high-quality outcomes they need to give their staff responsibility to achieve these outcomes. Managers who do not delegate responsibility will find it difficult, if not impossible, to establish teams that strive for high-quality outcomes.

2.252. The project manager is responsible for the project work from the initial kick-off through to closure. Only the primary responsibilities are given in the list below, and it can be expanded considerably. The responsibilities of the project manager include:

a. Using quality project management tools and techniques;
b. Identifying and managing the project stakeholders, both those inside the National Statistical Office as well as external groups and organizations that can impact the project outcomes;
c. Creating the conditions for good team working:
    (i) Setting the team norms and behaviours;
    (ii) Deciding responsibilities and coaching team members in new skills;
    (iii) Leading, guiding and directing team members;
    (iv) Controlling the work of the team – input and output;
    (v) Building trust and respect in the team;
    (vi) Encouraging personal growth, development, empowerment and continuous learning of the team members;
d. Establishing a project support office:
    (i) Developing project plans and budgets aligned to the strategy;

[38] *Handbook on Census Management for Population and Housing Censuses*, ibid., contains in Annex IV a case study of a system combining the quantitative components of the traditional system within a conceptual framework of a quality management approach. Annex IV also illustrates the important differences between the two approaches.

(ii) Prioritizing activities within the project;
(iii) Allocating and securing resource (for example financial and technological) commitments;
(iv) Working with established working practices and customs and developing relevant policies, systems and processes to implement objectives and plans;
(v) Developing a project communication plan;
(vi) Coordinating across internal and external stakeholders;
(vii) Developing a governance framework for the project;
(viii) Monitoring and tracking project progress against set objectives and plans and compiling relevant reports, for example progress, variance and status reports;
(ix) Monitoring and controlling expenditure and compiling financial and variance reports;
(x) Solving the problems that interfere with progress;
(xi) Informing internal and external stakeholders of progress and status;

e. Identifying and managing the project risks by developing a risk register, establishing a risk prioritization and risk mitigation plan, and monitoring and managing these risks;
f. Working with the unknown and unpredictable including assuring professional independence of the project staff and demonstrating objectivity through transparency;
g. Implementing improvement and change initiatives to ensure the soundness of methodologies, the cost-effectiveness of resources, the appropriateness of statistical procedures and the proactive management of respondent burden;
h. Delivering the project deliverables and benefits;
i. Leading the project team;
j. Evaluating and closing the project.

2.253. Managers must ensure that staff understand the philosophy behind the approach to quality. As mentioned above, staff involvement is a vital ingredient to quality improvement. Therefore, an environment needs to be established in which staff contributions are expected.

2.254. The second part of their role is to ensure that clients' expectations are known and that these expectations are built into planning objectives and into the systems that are to deliver them. Managers may have to work with multiple clients at once to satisfy project objectives and this complexity should be considered when building systems and during project planning.

2.255. Third, processes need to be documented and understood by the staff implementing them. Systems and processes for implementing the quality circle also need to be documented and put in place. Questions such as how quality is going to be measured, who is involved in identifying root causes of problems with quality, and how the process improvements are going to be implemented need to be answered. These will vary greatly depending on the nature of the process. Appropriate quality assurance techniques for each phase of the census are summarized below. These processes and metrics should also be shared with external clients as appropriate, to ensure their expectations for quality are being met.

2.256. The greatest test of management commitment to genuine quality improvement will occur in how management approaches problem solving. Staff will monitor management responses closely and adjust their own behaviour accordingly. Staff will act in accordance with how they see managers behave rather than what they hear managers say.

2.257. Managers who always react to problems by seeking out people to blame, or who establish systems that focus disproportionately on the merits or demerits of individuals at the expense of the

team, are sending messages that are contrary to the thrust of quality improvement. An environment where the emphasis is on fault finding, rather than on finding solutions to problems, or on excessive competition, will assure that staff cease to be part of the solution and become part of the problem. Managers need to take upon themselves the responsibility for problems, as they are ultimately responsible for the systems or processes that caused the problems. They should not seek to transfer the problems to lower-level staff.

2.258. However, even in the best-managed processes, there are circumstances where individuals can be justifiably responsible for negatively impacting quality. These may be individuals who do not possess adequate skills for performing their duties, and even deliberately flout procedures. These individuals need to be dealt with decisively, first and foremost, primarily by providing additional training and guidance all the way to administering disciplinary measures. Managers must deal promptly with these cases and act in a consistent manner. By doing so, managers will demonstrate to all other staff their commitment to quality.

2.259. To be successful, it is necessary to create a culture in which everyone has the opportunity to contribute to quality improvement. Most of the staff engaged in census operational work undertake routine tasks, and it is up to management to help them see the bigger picture, to motivate them and to encourage them to assume ownership of their work. This can be done by promoting a commitment to quality improvement and by adopting a consistent approach to management.

E. Quality improvement and the census

2.260. The following sections outline the way in which the concept of quality relates to the different phases of census activities. Quality considerations are discussed in relation to:

1. Topic selection;
2. Questionnaire design, development, and processing system testing;
3. Geospatial information and mapping;
4. Development of a data capture system; (including CAPI and CAWI)
5. Field operations;
6. Processing;
7. Analysis;
8. Dissemination;
9. Evaluation;

1. Topic selection

2.261. The first step in managing the quality of the product (namely, census statistics to be produced) is to ensure that the product will be relevant. The key process is extensive consultation with actual and potential users of census information, who may include for example, government ministries and departments, sub-national authorities, universities and other research institutions, the private sector, donors, organizations (or individuals) representing economic, social, cultural or special interest groups, such as religious groups, ethnic groups, persons with disabilities, and older persons. Potential topics for inclusion in the census should be agreed early in the census planning and questionnaire development phases. The key success factor in this process is full, frank and open consultation with users, subject matter and classification experts to ensure that census outputs address well-defined user needs. If their

needs are not met from the current census products, users may be reluctant to propose their needs for a future census. It is important to note that user needs will change over time, and users should be kept informed of changes to variable definitions and output classifications. A review of the extent to which user needs are met should be seen as an evaluation process feeding into the current cycle, the first step of quality management.

2. Questionnaire design, classification and data capture system testing

2.262. The next quality management task concerns the testing of each census question and the testing of the design of the form (paper or electronic version according to instrument(s) used). The development and testing of paper forms and the development and testing of electronic versions (including apps) should be treated as separate, but parallel processes with their own set of quality indicators. Again, the quality circle approach is used, with the results of each test being analysed and evaluated before being fed into further design and testing. The following areas are the key internal stakeholders of the form design process and their requirements need to be taken into account:

(a) The dissemination team (to ensure that the questions asked will deliver the data to meet the needs of users);

(b) The subject matter specialist team;

(c) The team responsible for development of the data capture system. For example, positioning of text and delineation of response areas may be dependent on data capture and the processing methodology to be adopted. It is critical that there is ongoing coordination between the form design and processing areas;

(d) The field operations team, which is responsible for training the enumeration workforce and printing the form.

2.263. The process of testing the classifications employed in the questionnaire is pivotal to ensure their alignment with both national and international requirements. This alignment guarantees that these classifications can be effectively compared and seamlessly linked to other databases. This meticulous testing procedure is imperative to guarantee the accuracy and reliability of the data. By harmonizing the classifications, the foundation is laid for cross-referencing and integration with external datasets. Consequently, this endeavour facilitates a comprehensive analysis and provides valuable insights into various dimensions of the collected data. Furthermore, the rigorous testing of these classifications ensures the apt design of the data structure for the census outputs. A well-structured data framework enhances the usability, accessibility, and interpretability of the census results. This approach ensures that the data outputs are not only accurate but also optimally organized for analytical purposes, enabling more informed decision-making processes.

3. Geospatial information and mapping

2.264. Many of the quality assurance components listed in section B are also applicable for mapping operations, geospatial products, and their downstream uses in the census. For example, processes that improve the completeness or accessibility of geospatial data (e.g., identifying new housing units or enabling user download of geographical areas) support quality. The dual role of geospatial data in providing the operational framework for pre-census operations and the means to associate statistical

observations with recognizable areas during dissemination, leads to some noteworthy quality considerations. These considerations relate specifically to the capacity of geospatial data to:

- Deploy and route field staff and support material efficiently.
- Define assigned collection and supervisory areas unambiguously.
- Associate tabulated census data with administrative and other useful geographic areas.

2.265. Ensuring the quality of demarcation requires rigorous testing during census planning and design that measures the efficiency of statistical unit design for operational purposes. The exact rules used to define physical features (e.g., river, road, ridgeline) priority order or the target number of households or persons per statistical unit will necessarily vary by country. A high-quality demarcation allows enumerators to collect data at predictable rates during census enumeration. These rates are measured during census pre-tests and a full census pilot that use planned census modalities within varied geographic areas.

2.266. The quality of aspatial attributes associated with geospatial data supports quality by improving interoperability and reducing duplication of effort. The use of harmonized classification and coding systems for qualitative characteristics (e.g., road quality, river type, urban status) and the administrative geocoding naming scheme (i.e., a gazetteer) within the NSO and other agencies with mapping functions supports data coherence in cartographic products used to support enumeration and dissemination.

2.267. The sophistication of the technology suite used during a census does not singularly guarantee the quality of geospatial data produced. Progressively more advanced technologies may support improved quality as they allow geospatial data to achieve improved relational and positional accuracy. However, the technological capacity to associate geographic coordinates with housing units using CAPI does not solely ensure the integrity of the census or that a household frame can be generated from the census. Coordinate capture must be part of a coherent, reproducible census methodology with measured positional accuracy and coverage error rates to meaningfully support census data integrity. NSOs must also have the financial and human resources to maintain the frame during the intercensal period, or the potential quality gains for future statistical processes from increased technological sophistication during the census are greatly diminished.

2.268. Programs that reduce ambiguity when interpreting boundaries contribute to quality by improving accuracy and enhancing reproducibility during data collection (i.e., a field worker assigned to a statistical unit based on census geography unambiguously knows which housing units are contained within, or lie outside of, their unit with as little re-canvassing as possible). An ongoing program to harmonize geographic data should coordinate between the NSO, other custodians of geospatial data, local governments, and local statistical offices. Since most NSOs are not the custodians of all administrative boundaries, this coordination ideally takes place within the framework of a national spatial data infrastructure.

2.269. An ongoing and consistent approach to geographic area and boundary maintenance is not always possible in resource-constrained environments. The census may be a rare opportunity to improve the completeness and congruence of administrative/statistical geographies and physical features. Collection and collation of metadata during the census process regarding how cartographic staff and field personnel resolve boundary questions will support the integrity of the census and support data maintenance efforts during the intercensal period. These measures can be as simple as encoding the rules used to prioritize physical features or ensuring that census field notes are digitally captured and saved rather than discarded once an issue is resolved.

2.270. Programs that improve data user familiarity with geographic areas, especially functionally defined areas, improve quality by supporting usability. NSOs may even solicit data user feedback to ensure that the geographic areas used for census tabulation continue to meet their needs.

4. Development of the data capture system

2.271. The transition from paper to digital data collection introduces a new operation- that of developing a data capture system. In this section, specific quality concerns related to the development of Computer Assisted Personal Interviewing (CAPI) applications and Computer Assisted Web Interviewing (CAWI) applications are discussed.

i. Computer Assisted Personal Interviewing (CAPI)

2.272. The development of a Computer Assisted Personal Interviewing (CAPI) application is frontloaded. If the CAPI application is poorly designed, implemented, and tested it will collect bad data. Several software development methodologies can be applied to CAPI development and should include creating a questionnaire specification, managing change control, and applying a freeze to the project. By developing a questionnaire specification which includes CAPI text, response values, consistency checks, and routing the questionnaire design and implementation are separated into two independent activities. Establishing a change control mechanism enables documentation of who, why, and when a specific change is made. Revision control software can help facilitate change control. It is important to select a date when all development will stop and only bug fixes are allowed. Unlike the traditional paper-based approach, CAPI applications tend to be seen as an ever-evolving piece of software. However, even a simple change can have unintended side effects, be incomplete, or simply be wrong. Which in a CAPI environment can lead to a small change having an oversized negative impact.

2.273. The following are several best practices that will improve the quality of the data collected by the CAPI application:

- Consideration should be given to the development of an operational control system to handle tasks that are not directly related to the interview. This will include logging in to the system, the ability to assign and receive assignments, complete and review questionnaires, track progress, and the ability to synchronize data. Once an enumerator is logged into the operational control system and has selected an assignment the questionnaire can be prefilled with data such as the enumerator and geographical codes which will reduce human error. Additionally, it can be helpful to prefill paradata such as device id, interview start time, and software versions.
- At a minimum each CAPI question will be comprised of question text and response options. The question text should include fills customized for each interview. Text boxes which allow open-ended responses should be kept to a minimum and used for fields such as name. More often a discrete number of pre-coded options should be shown as dropdown options. Ideally, the dropdown options include human-readable responses. Numeric codes can be recorded in the background without the enumerator having to be familiar with the numeric codes. Further, responses can be created dynamically, so an enumerator is only presented with valid responses. For instance, if the question is asking what the highest-level of education is and the person is six

years old then responses including higher education can be excluded. Optionally, help and instruction text can be included to assist the interviewer. The help text will include supplemental information like definitions, while the instruction text will include reminders such as probing techniques specific to the question.

- The flow of the CAPI application and use of consistency checks are critical considerations that will impact data quality. The CAPI application must be able to enforce the ordering of questions, so that they cannot be asked out of order. The CAPI application also needs the ability to skip questions that are no longer relevant to the current interview based on previously asked questions. Consistency checks are helpful for identifying grossly inconsistent data while the enumerator is in the field, but should not prohibit the enumerator from entering data in a timely and unobtrusive manner.

ii. Computer Assisted Web Interviewing (CAWI)

2.274. When developing a survey using Computer-Assisted Web Interviewing (CAWI) it is important to consider factors that will maximize survey data quality, user efficiency, user satisfaction, and census branding experience. Additionally, it will be equally important to consider factors that minimize respondent burden, measurement error, and design inconsistencies across data collection.

2.275. Improving survey quality in CAWI involves meticulous questionnaire design, robust pre-testing, and proactive data validation to ensure clear, accurate, and reliable responses from participants. Enhancing user efficiency in CAWI entails optimizing the survey's user interface and providing user-friendly features to streamline the completion process. Elevating user satisfaction in CAWI can be achieved by prioritizing intuitive design and ensuring transparent data handling practices to enhance respondents' overall experience and trust in the survey process. An additional priority is strengthening the census branding experience in CAWI to align the survey platform with the official branding and messaging of the census. This alignment communicates a consistent and trustworthy identity to respondents.

2.276. Reducing respondent burden in CAWI involves designing concise surveys that only request essential information and employing skip logic to personalize the questionnaire. Decreasing measurement error in CAWI involves using clear and unambiguous survey questions, conducting thorough pre-testing and validation checks, and ensuring data quality control measures are in place to enhance the accuracy of collected data. To lessen design inconsistencies across data collection in CAWI, consideration should be given to maintaining a standardized survey template, providing detailed design guidelines, and conducting regular reviews of question order, response scales, and visual elements.

5. Field operations

2.277. The quality management process continues throughout the design of the census field operations. These are tested as far as possible in conjunction with form designs testing. The key internal client of field operations is processing. However, field operations can also impinge on other areas, such as dissemination and classification and subject matter areas where certain concepts, such as what constitutes a dwelling, are implemented during the field operations phase. Several components of field

operations can be subject to specific quality circle mechanisms as these are likely to take some time and involve iterative processes. These components include:

(a) Demarcation of enumeration areas;
(b) Map production;
(c) Form printing, where a sample of forms is rigorously tested for adherence to standards.
(d) Development of applications to ensure data quality, such as
    i. *a collection application* with automatic tablet data storage, the necessary control specifications (consistency and completeness checks, automatic skips and alert messages), and GPS coordinates of the household dwelling to check the enumerator's presence;
    ii. *a control application* for the enumeration supervisor;
    iii. *a data transfer application* at central level, with all the necessary security features (data encryption, data bandwidth fluidity) to avoid data loss in the field;
    iv. *a web application/dashboard* for monitoring data quality via real-time indicators for management teams in the field and at central level;
(e) Installation of a real-time data back-up server at central level and a second replication server at another site to ensure data backup and collection continuity in the event of malfunction;
(f) Installation of a call-center with a toll-free number and a team of teleoperators to answer people's questions about the census in real-time. Teleoperators can collect the various problems encountered by field agents. These problems can be reviewed and analyzed to identify improvements to census operations.

Comment (MS): There is no mention of a pilot census through this chapter. Should a pilot census be regarded as a major QA exercise, especially for Field Operations?

2.278. All systems supporting data collection must be thoroughly tested before collection. This is especially critical when collection technology is used, such as handheld devices, tablets or laptop computers. All data quality benefits of using such technology could be compromised if problems arise during enumeration.

2.279. Quality monitoring should be established for each of these components and mechanisms put in place to ensure that the outcomes of the monitoring are used to improve processes. It is more difficult to implement the quality circle during actual enumeration owing to the very tight time constraints. However, this can be achieved by:

(a) Clearly establishing the aims of the field operations phase;
(b) Applying thoroughly documented procedures;
(c) Ensuring that the enumerators understand their role through appropriate training and providing inspection of corrupted forms;
(d) Providing opportunities for field staff to be observed operating on the job so that feedback can be given and retraining undertaken;
(e) Establishing communication and feedback loops with the general public through helplines, online forums, social media, etc., so that problems in the field can be detected and corrected in real time.

2.280. However, it has to be acknowledged that during the actual carrying out of the enumeration this approach tends to identify “problem enumerators” rather than systemic or process errors. This means

that evaluation following collection is vital. The evaluation should attempt to capture the experiences and suggestions of a range of enumerators and other field staff so that improvements can be made to the subsequent census.

2.281. A general overview of the quality of enumeration can be obtained through:

(a) Use of techniques such as post-enumeration surveys to gauge the level of underenumeration or overenumeration of people and dwellings;
(b) Overall response from the target population or level of non-response at the question level;
(c) Feedback from field staff;
(d) Measures of the quality of any coding undertaken by field staff;
(e) Mechanisms that may be in place to handle queries from the public;
(f) Analysis of administrative data.

2.282. The effectiveness of the public communication strategy may be assessed by the amount of press coverage (positive and negative) of the census and follow-up surveys to test the reaction to particular advertising.

6. Processing

2.283. The key clients of processing are the dissemination area and the areas of the country's statistical agency responsible for maintaining standard classifications and those with specialist subject matter knowledge. The dissemination area depends on the processing team to obtain data in an agreed format and compiled to agreed quality standards. This is necessary so that the data can be used in dissemination systems.

2.284. Since the census is part of an overall national statistical system, data from the census are likely to be used in conjunction with data from other collections. Thus the classification and subject matter specialist areas, which are responsible for those other collections, need to be satisfied that the coding, editing and other data transformation processes are conceptually sound and deliver data of acceptable quality.

2.285. Extensive testing of processing systems must be undertaken in advance of the census. Coding processes and training packages need to be prepared and tested using the type of staff likely to be involved in the operations. The processing phase gives the fullest scope for the use of quality improvement techniques, as many of the processes in this phase are repetitive and take a reasonable amount of time. This enables the quality circle to go through much iteration. It is vital that structures are put in place not only to monitor quality but also to involve processing staff in the identification of problems with quality and in proposing solutions.

2.286. It is generally not possible for processing to improve the accuracy of census data. At best, processes such as editing may reduce some inconsistencies within the data. However, in the end, the data coming from the processing system will not be of any better quality than the information supplied on census forms. Much effort can be expended in correcting apparently inconsistent or inaccurate census data with no real improvement in the fitness for the purpose of the data. It may be a better strategy to educate users to accept slight inconsistencies in census data, rather than developing highly

complex procedures that may introduce other errors and impose heavy costs in terms of delay in release of the data and cost to the community.

2.287. If a strategy of altering census data after collection to eliminate possible errors and inconsistency is adopted, then the following sections provide examples of ***data cleansing***, ***auditing***, and ***output consistency checks*** that can be implemented. Information about the nature of the checks undertaken and the proportion of records subjected to editing should be published to maintain confidence in census results and the methodologies used for adjustments.

2.288. ***Data cleansing*** is an important step in ensuring the quality and reliability of the census dataset. This process involves several key components, including:

a. a data validity review to confirm that all entered responses are accurate. This review ensures that numeric fields contain only numerical data, and text fields do not contain numerical data or other inappropriate information.
b. methods for handling missing data to ensure that all variables pertaining to the characteristics of the census units are complete. In cases where data is missing, appropriate imputation techniques can be used to fill in the gaps and maintain data completeness.
c. the correction of values outside electronic validation rules. It is important to check that any values outside the scope of the electronic validation rules have been rectified. These values may arise due to defects in the electronic survey during data collection or errors made in the survey process.
d. The detection of duplicates and a set of rules to support a decision for deleting duplicates

2.289. ***Auditing*** is an important process that involves the review and validation of various aspects of the data to ensure its quality and reliability. The key components of data auditing include:

a. validation of the coding during the conversion of textual data into numeric codes.
b. validation that the data collected adheres to international standards and classifications to ensure consistency and accuracy in data representation.
c. range checks to ensure that all answer values of variables are examined to confirm that they fall within predefined limits or ranges. Range checking can identify any outliers or values that may be outside of the expected bounds.
d. applying validation rules to review the logical relationships and correlations between variables.
e. dealing with extreme values by applying various reviews and audits to ensure that any illogical values have been addressed and corrected. This involves deciding on the treatment of unrealistic or unreasonable values and making a determination if values should be edited or replaced with more realistic data points.
f. verification that all targeted variables and geographic distributions have been thoroughly and comprehensively covered during censuses. This means ensuring that all relevant data elements have been collected and included in the dataset to provide a comprehensive and accurate representation of the population or sample.

2.290. ***Output Consistency*** and maintaining the consistency of the main census outputs is essential to ensure the reliability and accuracy of data. Maintaining consistency in outputs is vital for making informed decisions, conducting meaningful analyses, and ensuring that the data accurately represents the real-world context it is intended to describe. It helps build trust in the data and usability for various applications. This involves considering various aspects, including:

a. *Historical Consistency* to verify that the main outputs align with historical data and trends. Data should be consistent over time to ensure that there are no significant deviations or irregularities that might raise questions about data quality.
b. *Geographical Consistency* to verify that the main outputs align with location and geographic data and trends data should be consistent and in harmony
c. *Consistency with Other Sources* and cross-reference the initial main outputs with data from other reputable sources so that data aligns with external sources and any disparities are investigated and explained.

7. Analysis

2.291. Processing and analysis can be iterative and parallel. Therefore, some of the quality considerations discussed in this section are also contained in the processing section. Analysis can provide a deeper understanding of the data, which could indicate the need for additional processing. In the analysis step, statistics are produced, validated, interpreted, and made ready for dissemination. Also, during the analysis phase, statistical disclosure control methods are applied to ensure data privacy, and statistical content to support the interpretation of census outputs are prepared.

2.292. Quality assurance should be applied for each of these activities to guarantee interpretability and clarity of the statistical outputs, maintain accuracy, ensure comparability, address coherence with outside sources, guarantee relevance, and ensure privacy.

2.293. The interpretability of the statistical outputs is assessed by checking the existence of metadata, commentaries, technical notes and any other necessary supporting statistical information. In order to interpret and explain census statistics, it is necessary to view them from multiple perspectives and understand the extent to which these statistics meet the needs of the intended audience by conducting in-depth statistical analyses, including consistency and comparability analysis.

2.294. An assessment of statistical confidentiality and data privacy should consider three quality attributes, (a) accuracy, (b) relevance and (c) clarity.

a. Accuracy is assessed by analyzing the impact of statistical disclosure control methods on the quality of the output and by verifying that the statistical disclosure control doesn't affect significant relationships in the data.
b. Relevance is assessed by analyzing how well the output meets user needs after applying statistical disclosure control methods.
c. Clarity is assessed by describing the statistical disclosure control method(s) applied and any restrictions on the use of the dataset.

2.295. After assessing the impact of Statistical Disclosure Control on data quality and determining the appropriate techniques to apply, pre-release discussions and review with appropriate internal subject matter experts with in-depth knowledge of specific statistical domains should be undertaken. In addition, consistency checks should be implemented to ensure relevance and confirm that census outputs are fit for purpose and ready for dissemination. One of the main quality assurance measures is having a thorough review system in place for data products to make sure that the data or indicators being released are of high quality.

8. Dissemination

2.296. Census dissemination can easily be overlooked in the chain of providing a quality outcome for the census as management attention is diverted to the costly and risky enumeration and processing operations. The dissemination area is responsible for the timely delivery of products and services to the census data users. Therefore, insufficient planning and resources for this phase can have the effect of delaying the release of the data and thus compromising the overall achievement of the census objectives. The dissemination phase should also be regarded as an ongoing process that will service the needs of users over a long period of time.

2.297. Management of quality in census dissemination is driven by concerns to (a) deliver relevant products and services; (b) maintain accuracy of the data; and (c) ensure timeliness and predictability of data release within agreed cost constraints.

2.298. The first of these objectives is to provide relevant products and services. This can only be done by reviewing the experiences of the previous census products and services and by user consultation processes with both current and potential users of census data.

2.299. The second objective is to ensure that the data supplied from the processing system are accurately transformed into output products. A quality assurance strategy to ensure that data tabulations and transformations are carried out accurately needs to be documented and followed. Key aspects of a quality assurance strategy are outlined in the previous section on analysis. The quality circle approach to these processes needs to be applied and any gaps identified and corrected through extensive testing prior to the census and ongoing process improvement during the analysis phase before the dissemination of census outputs.

2.300. The third objective is the timely and predictable release of data from the census. While this is the responsibility of all phases of the census programme, the role of dissemination is crucial. The dissemination area needs to be realistic about release dates and ensure that these are communicated to clients early so as to manage client expectations. The involvement of staff actually responsible for the dissemination phase in devising these dates is recommended where this is possible. Dissemination systems and processes need to be available, documented and tested prior to the release of data from the processing and analysis phases.

2.301. A release calendar needs to be prepared to keep the user community informed about the likely month of release so use of data can be planned in advance. A mechanism to provide metadata on census indicators and the geography level at which these are made available needs serious consideration. Every country should assess the requirements and put in place a dedicated team to assist data users. The services of call centres may be used if the number of data seekers cannot be handled in-house.

9. Evaluation

2.302. Evaluation of the overall census operation is vital for identifying strengths and weaknesses of census phases, including planning, enumeration, data processing and dissemination, and also for the purpose of analysing the quality of census statistics, which are the major output of these processes.

With the quality assurance and improvement programme, the main objective is to ensure that quality assessment is consistently incorporated in all phases of the population and housing census, focusing on efforts in controlling the occurrence of errors and taking actions to ensure the highest quality of both the processes and their outcomes. Yet, errors appear to be inevitable in such a complex undertaking – consequently, there needs to be a mechanism put in place to determine the deficiencies and their quantitative impact on the census results.

2.303. Census evaluation with all dimensions of quality (see paragraph 2.175) requires a comprehensive evaluation programme for assessing and documenting the outcomes of each process using appropriate and customized methodologies. Methodologies for evaluation should be planned well in advance, in the planning phase of the census. It should be noted that this is continuous process implemented from the planning to the end of census operations. It is also appropriate to consider it as being the first step in the subsequent census cycle. Similarly, evaluation of one process within a census cycle could be the first stage in the next process of the same census cycle.

2.304. Evaluation of the accuracy of the census data should also be undertaken, to the extent possible, by conducting a post-enumeration survey for measuring coverage and content errors, by comparing the census results with similar data from other sources and by applying demographic analysis. As for other sources, these include surveys and administrative records in a similar time frame, tabulations when a population register is available, and previous census results. The purposes of evaluating the accuracy of the data are to inform users of the quality of the current census data and to assist in future improvements. Future improvement may be achieved by (a) improving processes; and (b) establishing performance benchmarks against which the quality of the data from subsequent censuses can be measured.

2.305. Evaluation of data accuracy may have two parts. Preliminary evaluation will enable the identification of any problem areas that have not been previously detected through the quality management processes in earlier phases of the census. More extensive evaluation should be undertaken on data items where problems have been identified or where new questions or processes have been attempted.

2.306. The census evaluation programme would be undertaken by subject specialist staff according to the agreed goals and methodologies covering all possible dimensions of quality. The following are some examples:

(a) Identification of the deficiencies and achievements in data capture, coding and editing (through mechanisms developed for checking the quality of process and the work of personnel);
(b) Relevance of census data with user needs and satisfaction of users with dissemination tools and products (based on information collected through user consultation);
(c) Achievements and difficulties in use of new technologies and methodologies and identification of possible improvements for the next census;

2.307. The results of evaluations of census operation for both operational aspects and the quality of data should be made available to the stakeholders promptly as they become available.

2.308. Depending on the national context, census results by default may be presented without adjustments. But for analytical and planning purposes, results of the census evaluation should be taken

into account to prepare post-censual population estimates in order to correct deficiencies with the enumeration.

F. Quality for register-based and combined censuses

2.309. There has been a clear trend towards increased use of administrative data in censuses, in line with a more generalized trend towards increased use of administrative data in all statistical domains. Their use can provide more frequent and timely statistics about the population, improvements in accuracy and reliability, and significant reductions in costs and respondent burden. Nevertheless, there are many key quality considerations that must be assessed and evaluated before incorporating an administrative source into a census.

2.310. The quality of data used in the process of compiling census statistics strongly affects the quality of the statistical output produced. Thus, the quality of data from administrative registers and other administrative sources is a key element that should be considered in the decision-making process on the use of administrative data in the production of statistics. It is necessary to prepare and implement a standard method of assessing the quality of administrative data as potential sources for the census. The assessment of the quality of administrative data may be used to evaluate the usability of the data source. To this end a set of indicators should be developed by which each source may be analysed. Based on these indicators, the NSO can decide whether or not to use a specific source.

2.311. In general terms, the NSO should identify the administrative sources that may be relevant to their census. It is important to set out what the expected or required outcomes of using the source would be. This could include improvements to the efficiency of the census operation in terms of reductions in cost and respondent burden, improvements to the quality of the census, or the delivery of new or enhanced census outputs.

2.312. The following key quality considerations should inform decisions about the use of administrative data in a census:

1. Administrative data have, in general, not been collected for the purpose of a census. Therefore, the data sources may have adopted different concepts, classifications and definitions from those required by the census, they may refer to different reference periods, be subject to lags in the updating of information, and may have limited coverage of the census population. Additionally, the accuracy and completeness of the data will be highly dependent on the importance of the data to the data supplier's function. The administrative sources may also be subject to changes over time and inconsistencies in the way the data are collected across segments of the population.
2. The use of administrative data by the National Statistical Offices for purposes other than those for which the data were originally collected raises privacy, security, and legal concerns. An assessment of public acceptability of the use of administrative data for statistical purposes must be performed on a credible legal basis and must be publicly acceptable. It is important that the public and data users understand how and why administrative data are being used in the census. The NSO should be transparent about their use and provide a clear justification of the benefits against any risks and costs. This can be achieved through good communication, including the publication of the procedures and policies that support the effective use and protection of the data.
3. The NSO has limited control over the way the administrative data are collected and processed. Administrative data can be subject to changes in population coverage and

population measures over time due, for example, to legal, policy, procedural, or system changes affecting the data and/or their delivery. This limited control and the associated risks to census processes can be mitigated by working with data suppliers on potential or planned changes, by being flexible and responsive to change, and by reducing reliance on any single data source or item where possible.

2.313. Assessing the quality of administrative sources requires mapping the errors which may occur before and after the data is supplied to NSOs and determining how these errors can be mitigated. At the same time, the quality assessment should not be limited to assessing the error. More generally, it should include the assessment of the impact of the integration of administrative data into the census in terms of the extent to which it adds error or uncertainty to the outputs, vis-à-vis of the advantages gained from integration.

2.314. A framework for assessing the quality of administrative sources for use in censuses can be built in stages: the Source stage, the Data stage, the Process stage, and the Output stage. The first two stages mainly deal with the assessment of the quality of single sources, while the last two stages are concerned with the quality of transformed and/or integrated sources. For each stage, the key quality dimensions should be identified as well as the respective tools and indicators for the quality assessment[39]. A brief description of each stage is included below.

1. ***Source Stage.*** The quality of administrative data sources to be used in the census production should be assessed both at first acquisition and when they are regularly re-supplied to the NSO. This is because the characteristics of any data set that has previously been supplied might differ due to changes in concepts, classifications, collection methods, and so forth.
2. ***Data Stage.*** The quality assessment at the Data stage refers to the quality assessment of raw administrative data as supplied to the NSO, with reference to the expectations and requirements established through the metadata-based assessment at the Source Stage. The quality of administrative data at the Data Stage is to be assessed against several dimensions including readability and validity, accuracy and reliability, timeliness and punctuality, and linkage ability.
3. ***Process Stage.*** Once administrative data are received and the quality assessed, the data will require processing to be usable in the census. The administrative data will need to be integrated into the census design and any quality issues will need to be addressed. Some of the most common processes required for using administrative data in the census are: linking records, assessing coverage errors, resolving inconsistencies from different sources, editing and imputation.
4. ***Output Stage.*** Measuring output quality cannot be reduced to only the estimation of uncertainty (the accuracy of estimates); rather, it should include an assessment across all other quality output dimensions. The introduction of administrative data will likely lead to gains in some dimensions and losses in others. Achieving the right balance across the quality dimensions is the key to best meeting user needs. Updated and new administrative data can be used to improve census statistics by revising previous estimates. However, this needs to be balanced against user needs with respect to revisions.

[39] See UNECE, Guidelines for assessing the quality of administrative sources for use in censuses, 2021

2.315. The results of the quality assessment and assurance should be recorded and published throughout the census production process. This will support an ongoing dialogue between producers and users of census data. and help determine whether the right balance has been achieved across the multiple dimensions of quality.

PART THREE. CENSUS OPERATION ACTIVITIES

I. Introduction

3.1. Part Three of the *Principles and Recommendations* focuses on elaboration of census operations as described in Part One. While the population and housing census is in essence a statistical data collection exercise, it still incorporates components that are not present in a routine statistical survey conducted within the frame of the national statistical system. In consequence, the layout of this part follows the frame and the logic of the Generic Statistical Business Process Model;[40] it also provides particular methodological and operational guidelines relevant to the population and housing census. This part of the *Principles and Recommendations*, therefore, starts with a discussion on developing the census questionnaire; building census infrastructure; mapping; testing; living quarters and household listings; field enumeration; data processing; evaluation of the results; dissemination; analysis; archiving; documentation; and evaluation of overall census operations.

II. Census questionnaires: content and design

3.2. The preparation of the census questionnaire refers to a well-designed process that should start at a very early stage of the census operation. This process includes developing various methods and tools that will respond to national priorities and data user needs from different sectors, such as governmental organizations, research institutions, the private sector, the public, development partners, civil society and non-governmental organizations. In addition to communicating with users, other factors that would have impacts on selection of census topics, include quality of data collected through the previous census, timeliness, national sensitivity for specific topics and available resources. Another factor to consider is the appropriateness of collecting reliable data on a particular topic through the census.

3.3. Taking into consideration that the basic principle is to meet user needs and to make census statistics as useful as possible, the content of the census questionnaire would be determined through consultations with census data users. The consolidation of the results of the consultations with all stakeholders needs to be balanced by factors such as the response burden, respecting respondents' privacy and confidentiality of their information, as well as other critical considerations such as availability of resources (financial, personnel and equipment).

3.4. The final content of the questionnaire will also be the result of rigorous and meticulous testing, in terms of both content and design of the questionnaire. It is necessary that all aspects of the census questionnaire, such as wording, structure of the questions and design, be carefully tested[41] to ensure

[40] *Generic Statistical Business Process Model, GSBPM, Version 5, December 2013,* United Nations Commission for Europe on behalf of the international statistical community.

[41] For more information about census tests, see Chapter IV.

successful application of the questionnaire in the field, and whether or not the census is the most appropriate means for collecting reliable data on the topic. Questionnaire content should also be examined with respect to the modes of data collection to ensure formatting and wording are appropriate to the mode.

3.5. Preparation of the census questionnaire(s) requires most careful consideration, since the consequences of a poorly designed questionnaire(s), and content, cannot be overcome during and after enumeration. Successful implementation of this process will have significant impacts on the quality of collected data and of census outputs.

a) Selection of census topics

3.6. As a first step in determining the content of the census questionnaires, the topics that will be covered in the census should be selected taking into consideration the priorities of national needs, regional and international recommendations,[42] historical comparisons, regional and international comparability, suitability of topics for collecting reliable information and resources available for the census undertaking. Census takers should evaluate the national needs in the light of possible new topics and the needs for continuation of the topics covered in the past censuses and/or surveys. Topics that are not needed in the current census because of changing data needs and availability of alternative data sources, as well as new topics for which data is needed, should be carefully reviewed in this process.

3.7. During the process of selecting census topics, data users and interested parties should be consulted on the type and extent of socioeconomic information they believe should be provided by the census. The results from consultations should be reviewed in conjunction with available resources for the census, the suitability of the topic(s) for inclusion in the census, and the burden to respondents. In deciding what topics should be covered in the census, the length of the questionnaires and complexity of the questions should be carefully examined. More information on the factors determining the selection of census topics is given in Chapters I and II of Part Four.

3.8. Another factor that should be considered in the process of selecting census topics is related to the decision on whether to use a single census questionnaire for all respondents or adopt a two questionnaires approach – short-form and long-form questionnaire. Using a single questionnaire consists of a standard set of questions for all individuals and housing units covered in the census. In the latter approach, countries use a short-form questionnaire with basic questions for enumerating all of the population, while a long-form questionnaire is applied to a sample of the population for collecting more detailed information. Subsequent paragraphs provide a more in-depth elaboration on this approach.

b) Use of short and long questionnaires

3.9. With each new census and the advancement in processing and exploiting census data for informed decision making, there is an increased interest in adding topics to the population and housing census, in addition to those historically covered in the census. Because of additional costs and the burden on the respondents, imposing a long questionnaire on the total population, in many cases, does not seem to be appropriate. Hence, countries often decide to broaden the scope of the census by

[42] Part Four of these *Principles and Recommendations* focuses exclusively on core and non-core topics for population and housing censuses.

covering additional topics through the use of sampling methodology. In this approach, two questionnaires are used: (a) a short questionnaire containing only those questions intended for universal coverage; and (b) a long questionnaire containing questions on the short questionnaire as well as detailed questions on specific census topics for a sample coverage.

3.10. Use of a sampling methodology in conjunction with full enumeration requires careful planning for determining the topics for the long questionnaire. This approach might be cost-effective, considering less duration of data collection for all topics; on the other hand, it may create some complications regarding the field organization.

3.11. The following paragraphs explain how sampling can be integrated with the full enumeration and its possible advantages and limitations.

3.12. The ever-expanding needs in most countries for extensive and reliable data have made the use of sampling a cost-effective part of census taking. Sampling is increasingly being used to broaden the scope of the census through asking additional questions of only a sample of the population and households. The use of sampling makes it feasible to obtain urgently needed data of acceptable precision when factors of timing, respondent burden, and cost would have made it impractical to obtain such data on a complete count basis.

3.13. The suitability of particular questions for a sample enumeration depends on the precision with which results are needed for small areas, and small population groups, and on the enumeration costs involved. Collecting data on particular questions for smaller areas or groups requires careful consideration of sample size. While moving the questions to full enumeration may initially appear as the only solution, exploring the option of increasing the sample size can offer advantages. A thorough cost-benefit analysis will ultimately determine the most efficient and effective approach.

3.14. It is important to bear in mind, however, that national legal requirements may make it mandatory to collect certain information on a complete count basis. Legislation in many countries prescribes complete population enumerations at particular times or makes certain political or administrative dispositions dependent on particular results from a complete enumeration. For example, the apportionment of seats in the legislature among the civil divisions of a country often depends on the number of persons actually enumerated in each division. The data needed for this and similar purposes may not be obtained using a sampling methodology.

3.15. Census information that is collected for only a sample of the population or housing units is usually obtained by one of two different methods. The first predesignates a systematic subset of census households to receive a so-called “long” questionnaire, or the census form that contains the questions on the short form as well as some additional questions on select topics. Depending on the sample requirements, which in turn take account of considerations of cost and precision, the systematic subset that is designated for the long questionnaire may represent, for example, 1 in 4, or 1 in 5, or 1 in 10 of the census households. Under such a sampling scheme, all other households in the census will receive a short questionnaire containing only those questions intended for universal coverage. If countries choose this option, it is recommended that the predesignation of the sample households that are to receive the long questionnaire be carried out at a central location by supervisory statistical staff, since it has been shown that when the enumerators themselves actually identify the sample households the results are often biased. This can be done after the pre enumeration household listing is completed in an enumeration area.

3.16. The second method of sampling often used involves designating a sample of enumeration areas to which the long questionnaire will be administered. In this approach, the long questionnaire will be administered in all households in the designated enumeration areas and the short questionnaire will be administered in all households in the remaining enumeration areas. The advantage of the first method over the second is that the sampling precision of the census results is greater in the former because clustering effects increase the sampling variance when whole enumeration areas are used as sampling units. On the other hand, the advantage of the second method is that different enumerator staff may be trained more easily, since one set of enumerators can be trained only for the long questionnaire and the other set only for the short questionnaire. Sampling the entire universe often requires a significant sample size if there are questions on the long-form questionnaire that target a rare population. To reach these populations, one must weigh the costs of increasing sample sizes against the increased variance of a clustered sample. However, the clustered sample would only be effective if the rare population is known to be clustered in the enumeration areas, which is often not the case.

3.17. It is important to make certain that asking questions that are not asked of all persons does not give rise to legal, administrative or even political issues, since census information is required under statute and often with penalty for refusal.

c) Sections of the questionnaire

3.18. Clearly distinguishing between the units of enumeration is an essential element when designing the census questionnaire, as each question aims to collect data for a specific unit of enumeration. The census questionnaire can be clearly and intuitively structured if it is done based on the unit of enumeration. This structure has a positive impact in better understanding the meaning of the questions and in conducting the interview smoothly. This also facilitates successful completion of the questionnaire for self-enumeration. For more information on units of enumeration, see chapter IV. Section A.

d) Questionnaire design

3.19. Although the majority of countries are still using face-to-face interviews with paper questionnaires, many countries are using other methods of gathering or otherwise generating the required information, including the use of multi-mode methods. Some of these methods are face-to-face interview with electronic questionnaire, telephone interviewing, self-enumeration with a paper questionnaire collected by enumerators, self-enumeration with a paper questionnaire returned by mail, self-enumeration via the Internet, as well as the use of population registers-and other pre-existing administrative records, either alone or in combination with other sources to generate data for the census. The methods of enumeration and technology used for data capture are among the main factors affecting the design of the questionnaires. For example, design of the questionnaires that will be used for the face-to-face interview and for self-enumeration will differ, as the former will be applied through enumerators while the latter will be directly used by the respondents. Whatever methods are chosen, these need to be tested and assessed in advance for data quality and feasibility. It is also important that data security and confidentiality is maintained whatever modes or approaches are used.

3.20. The design of the questionnaire must be based on the type of data collection mode and approach used. Questionnaire design should also be based on the technology for data processing, for

example whether data processing will be done through scanning, manual entry or electronic transmission to the database. It should be noted, however, that regardless of the mode of enumeration used, it is advisable that there is compatibility in the design among the modes in order to enhance data comparability among modes (for more information, see section on designing multi-mode data collection).

3.21. The following paragraphs relate only to those approaches that involve direct enumeration of the individuals covered by the census. While many of the principles of designing a statistical questionnaire will also apply to the design of the administrative instruments underpinning a register-based approach, those instruments may also be based upon specific requirements of the administrative programmes they address.

3.22. Further, where countries utilize the Internet or handheld devices to collect their census information, or a portion of that, the layout and organization of the data collection instrument may differ from that of the paper questionnaire, although it is important that the content be the same for all the modes of data collection used. It is important to ensure the same information is collected and entries are similarly checked and the results are comparable. It is important to note that most often, adopting an Internet approach also means moving from an enumerator-based approach to a self-completion approach. The questions must be designed to be completed by the respondent without outside assistance. Therefore, census management should involve the information technology team, as well as cognitive testing experts right from the questionnaire preparation stage. While many of the same principles (for example clarity of wording, omission of unnecessary material) will apply also to an Internet-based or handheld device-based collection of information, specialized advice should be sought regarding such issues as (a) the technology employed to present the questions to the respondent; (b) the method of capturing the response; and (c) quality assurance checks employed during data capture process.

3.23. A crucial principle is that questionnaire design must be regarded as part of an integrated process of satisfying user demands by collecting, processing and disseminating information provided by respondents.

3.24. Among the many factors that should be taken into account in designing the questionnaire are the method of enumeration, the type of questionnaire (see paragraphs 3.9–3.17), the data to be collected, the most suitable form and arrangement of the questions, technologies used and the processing techniques to be employed.

3.25. It is important that questions and response options are free from ambiguity. Moreover, questions should not be offensive; in many cases this can be avoided by excluding extremely sensitive topics from the census questionnaire, but care must always be taken to consider the reaction of respondents when designing questions. In addition, it should be noted that the quality of information collected in a census will be reduced if the questionnaire is excessively long. These issues should be carefully assessed during the testing programme, including the cognitive test and the pilot census (see paragraphs 3.113–3.114) since poorly worded questions will not only collect poor-quality data, but, by confusing respondents or enumerators, may also impact on subsequent questions in the questionnaire. Skip patters and the form and format of the response options require careful consideration. A question can be perfectly worded, but if the response options aren't presented in a way that is meaningful, unambiguous, and easily understood, resulting data may be of poor quality.

3.26. Special provisions will have to be made if two or more languages are used in the country. Several methods have been used to deal with this situation, such as (a) a single, multilingual questionnaire; or (b) one version of the questionnaire for each major language; or (c) translations of the questionnaire in the various languages available in the enumerators' manual or on the Internet site for the census. In order to ensure consistency and to avoid interviewer bias, it is important that translation of the questionnaire into various languages that are used in the country be centralized and not be undertaken in the field by interviewers. Information on the distribution of languages in the country is important for sound census planning and, if not available, will have to be collected at some stage of the census preparations. Staff recruitment and training procedures (see paragraphs 2.119–2.124) will also have to take language issues into account.

3.27. If the housing census and the population census are to be carried out concurrently, it will be necessary to consider whether a single questionnaire should be utilized to collect information on both population and housing topics. If separate questionnaires are used, they should be uniquely identified in a way that links the component forms so as to permit subsequent matching, both physical and automated, of the data for each set of living quarters with the data that refer to the occupants thereof. This will be particularly important where a single housing form is used to cover separate personal forms for each individual.

3.28. When paper questionnaires are used for data collection, the use of processing techniques, such as optical mark reader or intelligent character recognition, will have a significant effect upon the questionnaire design (see paragraphs 3.175–3.177). In the case of optical mark reader, it is necessary both to provide spacing of response areas and to ensure printing is undertaken to precise tolerances so that the data capture software is able to capture all required data but not any of the material around the designated response areas. With regard to intelligent character recognition, it is crucial to allow sufficient room for response areas and to ensure that these are designed according to the requirements of the processing system so that each response box contains only one character. Enumerators should be trained to print neatly and correctly (usually in upper case). To better handle the entries by self-respondents, the ICR technology has to be built to interpret various forms of manual entries. As noted in paragraph 3.30, where the scanning process requires that a booklet questionnaire is separated into component pages, it is important that some form of linking (for example by serial numbers or barcodes) is employed to ensure that the correct information is amalgamated in the computer records.

3.29. Questionnaire design must be driven by a planning process based upon dialogue between the statistical agency and the data users. Information to be collected should respond to user needs both at national and international levels and therefore user consultation is crucial in this regard. Previous census questions that are no longer relevant should be dropped as these do not add any more value. This is essential if the questionnaire is to be designed to provide the information needed by users, as well as to manage respondent burden.

3.30. The final questionnaire must be drafted in time to allow for printing and developing the data collection application, in the case of using electronic questionnaire (making allowance for the many contingencies, such as industrial action, breakdown of printing equipment or delay of programming activities, that can arise in these processes); undertaking quality assurance checks to ensure the printing is of sufficient quality to be used in the data capture regime and to ensure that the data collection application is running correctly under the data entry rules.

3.31. As some countries are also utilizing Internet portals and handheld devices for conducting the census, sufficient time must be given to design, develop, test and implement bug-free e-questionnaires and related software systems. Last-minute inclusion or changes in the questionnaire may affect the overall quality of the programs, and in turn the census results.

3.32. In view of the many issues to be addressed in designing a census questionnaire, it is not feasible to suggest specific model questions for the census topics covered in Part Two. However, images of all census questionnaires that have been made available to the United Nations Statistics Division have been placed on the Division's website[43] together with research papers relating to questionnaires used to collect information on the various topics recommended for collection and also using the different technologies (Internet, handheld devices). It would be important to review not only questionnaires, but also the validation rules and checks, as well as software programs and scripts used for quality control.

(i) Design of questionnaires for electronic data collection (CAPI, CAWI, CATI)

3.33. The design of electronic questionnaires for data collection via CAPI, CAWI and CATI methods require additional considerations to make the data entry process intuitive for the enumerator or respondent. Some essential functional features that should be used in the design of the electronic questionnaires include:

(i) **Questionnaire navigation** should allow enumerators/respondents to move relatively freely through the questionnaire in order to enter responses in the most effective way, giving the ability to pause and resume at the last answered question with a "save and continue later" functionality. On the other hand, the design should impose some restrictions on navigation, for example, by preventing enumerators/respondents from entering certain questions without having first obtained responses from other, earlier, questions.

(ii) **Skipping/automated routing** is one of the most important error reducing features in electronic questionnaires. It obviates responding to questions that should be skipped. It also avoids the converse - skipping questions that should be asked, thus minimizing the need to impute for missing responses. Basic skips allow the response to a particular question to determine whether or not the next question is relevant, while complex skips are those that either use responses from several previous questions to determine whether the next question is relevant.

(iii) **Pre-coding** allows relevant questions to be answered from precoded drop-down menus. In some cases, drop-down menus could be altered dynamically, depending on previous responses, so that the interviewer is never presented with an impossible response code.

(iv) **Validation.** Real-time data validation checks can correct inadmissible or inconsistent responses that could be the result of either interviewer or respondent error, thus reducing the amount of post-enumeration data edits.

(v) **Support** should be available for different types of issues such as: technical problems; and the actual content of the census questionnaire – in case respondents have difficulties understanding what is being asked of them. Support can be provided in a number of ways including through help buttons and additional instructions and help materials.

[43] See http://unstats.un.org/unsd/demographic/sources/census/censusquest.htm

III. Building census infrastructure

3.34. The population and housing census requires a concentrated effort in building intricate and complex infrastructure that is suitable for conducting massive activities in a simultaneous manner. This infrastructure refers to: the development of the census instruments/questionnaires; logistic schemes for disseminating and collecting returns; facilities and systems for capturing, processing, editing and validation of response data; and the systems and tools for dissemination of statistics and data products. It also requires well-defined workflows and testing of the production system as well as of the whole statistical business process.

3.35. The necessary infrastructure for a census will vary significantly based on the design of the census and the data collection mode(s) being deployed. For a traditional paper-based census, the infrastructure will need to support the paper-based activities such as printing paper questionnaires and materials, space for storage of paper materials, mailing and distribution plans for distributing paper materials to field operational areas and returning completed questionnaires for processing, systems and equipment for capturing data from paper questionnaires, etc. For electronic data collection, on the other hand, the infrastructure will need to include facilities, equipment, and software involved in the electronic data entry, transmission, and storage. In this case, cloud infrastructure can be involved to achieve the necessary flexibility with the performance and reliability of the provisioned resources. The infrastructure for a records-based census will need to also include extensive storage for data files and tools and systems to facilitate the necessary matching and processing of the administrative data. Censuses that employ a hybrid approach using a combination of two or more of these approaches can require a very complicated infrastructure that must support all of the data collection modes being used.

3.36. Census infrastructure must be put in place long before the data collection exercise itself, as all of the components need to be extensively tested in circumstances that are as real as possible and avail reasonable time for any corrections needed as the testing may indicate. The importance of testing cannot be overestimated, especially in the context of the need to build process components for processing and analysing data. The preparation of the infrastructure does not usually translate into immediate procurement decisions, as the technological process results in better performance, richer functionality, and lower cost of the equipment with time. Thus, purchasing equipment or devices (e.g., mobile devices for data collection) several years in advance of the fieldwork brings the risks that the equipment/devices will become obsolete or unsupported by the time of the census. Instead, a limited sample of the equipment can be acquired at the time of building the infrastructure to facilitate testing, with the bulk being acquired later, when confirmed suitable by the tests, and to take advantage of the most recent technology available.

3.37. Building census infrastructure should be based primarily on the experiences and lessons learned from the previous census. If the previous census took place a long time ago and the documentation and institutional memory are not sufficient, putting together census components should rely on statistical practice in conducting large surveys and on the schemes for developing and implementing a statistical business process. Lessons learned from conducting similarly demanding operations (for example, an agricultural census, or census of establishments) can also be considered.

3.38. Particular components of the census infrastructure may already be available within the national statistical office – for example, the statistical network of field offices covering the country's territory and

staff experienced in collecting and producing statistics. Other components will need to be built from the beginning, such as the training synopsis and schedules, recruitment procedures and logistical arrangements. Methodical planning is critical in this process of building census infrastructure and developing detailed project schedules early on would significantly facilitate managing the process in a holistic manner.

3.39. In that context, clear and unambiguous delineation of responsibilities in terms of management is equally important to the building process. Managers and supervisors need to have a comprehensive assignment from the very beginning and at all levels of managing. Along with roles and responsibilities, channels, methods, and cadences of communication must also be clearly defined to ensure the different organizational units are collaborating and coordinating their assignments. While improvising may be necessary in certain circumstances, all efforts should be put into anticipating risks and developing and testing strategies to allow managers to mitigate them and have alternatives in place when needed.

3.40. In building new components of the census infrastructure – for example, outsourcing – it is necessary to liaise with other parts of the government that may have more extensive experience in that respect. In most cases government not only would have extensive experiences – such as subcontracting for defence purposes – but also might have the capacity to provide goods and services that are needed. Therefore, the process of building needs to first assess existing government capacities, and then extend beyond them.

3.41. The process of building has to be completely finished before the pilot census takes place. Certainly, parts or perhaps even whole components of the infrastructure might need to be adjusted and altered as a result of the pilot census. . Taking this process into account, the schedule for the preparations of the census must allocate sufficient time for the modification of systems necessary at the end of the pilot test.

3.42. Once the census takes place, all the steps in building the census infrastructure and its testing need to be meticulously documented and stored for future use. Certain components will certainly remain permanently as part of the broader national statistical infrastructure, such as computers, networks and servers. However, documenting in a comprehensive manner is of paramount importance, not only for future censuses, but also for a better understanding of the one that has just taken place.

IV. Geospatial information and mapping

A. Introduction

3.43. Geospatial information provides one of the most crucial aspects of census taking – location. The UN Statistical Commission at its fifty-first Session in March 2020 endorsed the Global Statistical Geospatial Framework (GSGF) noting the increasing relevance of and the need for the integration of statistical and geospatial information as an important bridge to enable the production of harmonized, standardized and integrated, geospatially enabled statistical data to facilitate data-driven decision-

making, in particular with regard to the implementation of the 2020 round of population censuses and the Sustainable Development Goals[44] .

3.44. A census which can only report aggregated population and demographic data is of limited use; it is the power of understanding the detailed distribution of these statistics within the nation that enables analysis and information for policy and decision making.

3.45. Geospatial information has a fundamental role in the national census exercise: identifying the location of housing units and population; routing enumerators to these locations; tying these basic data attributes to a location (georeferencing or geocoding); monitoring progress by small geographic units; applying location based quality control measures; tabulating and aggregating survey response data into useful geographic units; then distributing the data for geospatial data analysis - at every stage, geospatial information is critical to a census. The organization of the structure of the geographic units (sub-national geographies, enumeration areas, individual enumerator assignment areas, etc.) can play a role in the implementation of the census taking and the geographic organization of this framework can impact resulting quality, or even success of the entire effort.

3.46. To support these efforts, the Global Statistical Geospatial Framework The_GSGF.pdf (un.org) details some of the very basic conceptual phases needed to create a flexible, interoperable national geospatial structure, fundamentally integrated with the statistical data collected during enumeration. And the Integrated Geospatial Infrastructure Framework45 (IGIF) serves to enable the structure needed to support this implementation.

3.47. Geospatial information is the fundamental base of a national census. Providing relevant information for subnational geographic areas is a basic goal of a national census, divided into all basic geographic units for the census, which are then used for the creation of enumeration areas. Knowing that there are unenumerated housing units within the nation is of limited value, while knowing the location of these housing units as well as the nearest available census enumerator is of much greater value.

3.48. The National Spatial Data Infrastructure (NSDI) is a framework of policies, institutional arrangements, technologies, data, and people that enables the sharing and effective usage of geographic information by standardizing formats and protocols for access and interoperability. The development of a NSDI is a multi-agency and multi-partner activity.

3.49. The benefits of using NSDI:

- Providing a unified Standard base map with different scales
- Reduce duplication of efforts among governmental authorities
- Increase the benefits of using available spatial data
- Lower costs related to geographic information while making geographic data more accessible
- Legal arrangements and governance
- Establish key partnerships between Gov. Authorities, states, cities, academic, and the private sector

3.50. The advantages of NSDI

[44] https://unstats.un.org/UNSDWebsite/statcom/session_51/documents/2020-37-FinalReport-E.pdf
45 UNSD — UN-GGIM

- Gains in time for searching and accessing the necessary spatial data
- Access to the latest data and complete metadata
- Flexibility to increase
- Contributes to openness & transparency
- Improved and updated continuously
- NSDI Web mapping Governmental Portal

3.51. All governmental authorities, private sectors and public users can view NSDI data set and map layers based on user rules and privilege to download, print, geo processing services and dynamic queries. To develop an NSDI, the GSGF46 can be used as a conceptual model of the geospatial infrastructure needed. The IGIF provides guidance for building the NSDI's framework of support for the implementation of the GSGF.

B. The role of geospatial information in the census

3.52. The role of geospatial information in the census process is to support enumeration and to present aggregate census results in cartographic form. Very few enumerations during the last several census rounds were executed without the extensive geospatial support.

3.53. In general terms, mapping serves several purposes in the census process, as follows:

(a) Maps ensure coverage and facilitate census operations (pre-enumeration). The census office needs to ensure that every household and person in the country is counted and only counted once. For this purpose, census geographers partition the national territory into contiguous small data collection units. This basic geographic framework creates an essential control device that guarantees complete coverage of the census without any duplication or missing areas. The area of the collection unit considers a balanced workload among enumerators.

(b) Digital mapping applications can support data collection and can help supervise census activities (during enumeration) in near-real time, allowing for detailed supervision of progress. During the census, geospatial applications can be used to generate maps, or support digital platforms to enable the deployment of mobile devices for navigation to assigned geographic areas, in which they will navigate and enumerate households using digital devices using modern computer assisted in person interviewing. Geospatial services support planning to gain insights and efficiencies for field preparation and control tasks, allowing census supervisors to more closely supervise enumerators with real time quality control of census operations in the field, and from the central office, to support enumeration staff. This allows supervisors to strategically plan, make assignments, identify problem areas and implement remedial action quickly.

(c) Maps and other spatial data visualizations are essential to present, analyse and disseminate census results (post-enumeration). The cartographic presentation of census results provides a powerful means for visualizing the results of a census and the demographic data on different levels of maps, administrative boundaries, grid square system (if used) and detailed maps (buildings level).

46 http://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/The_GSGF.pdf

(d) This supports the identification and integration with ancillary information of local patterns of important demographic and social indicators. Digital mapping services and data visualizations are thus an integral part of policy analysis in the public and private sectors. Modern interactive and web-based digital GIS applications should be used to allow for the wide distribution of this information within the enterprise and the public.

(e) Geospatial data analysis can play a supporting role in ensuring data validation at the initial stages of processing the collected census data - visualization of the spatial distribution of potential atypical data (outliers), detection of potential topological errors in the applied territorial divisions, etc.

3.54. The census enumeration team needs to have a set of unique maps covering the entire country that accurately defines the boundaries within which each enumerator has to work during the enumeration phase of the census. Therefore, the quality of maps used in the census has a major influence on the quality and reliability of census data.

3.55. The types of maps required for census management include the following: (a) small-scale reference maps for use in the census agency to manage the overall operation; (b) large-scale topographical maps for use by enumerators; and (c) maps of the subregions or administrative areas, for the use of managers, showing the location of small population settlements and dominant physical features, such as roads, rivers, bridges and the type of terrain. The use of earth observation data to generate these maps is now increasingly common across countries.[47] For implementing agencies with different capacities, it is possible to use readily available free (open data) global human settlement maps (also generated with the processing of Earth observation data).

3.56. Careful consideration should be given to organization and management of geospatial support activities during the census planning and preparation phases. The lead time necessary for the implementation of a geospatial support solution (developing the GIS applications, purchasing digital devices, planning the full enterprise geospatial support infrastructure) for a country will be determined by a wide range of factors, including the number of maps to be produced, the technology available to produce and assign them, the availability of funds to acquire additional resources and the time required for distribution of the maps to field staff. A paper-based solution should only be considered in the most extreme and only in very limited circumstances.

3.57. It is still the case that in many countries there are only a limited range of maps available and these often do not show sufficient detail to enable the boundaries of small areas to be clearly defined. This is particularly likely to apply in areas of unplanned settlement. However, with high resolution satellite imagery of sufficient quality and other current Earth Observation data sources (like open global human settlement layers and maps), this is no longer the case. Unplanned areas can be mapped precisely (especially with localised aerial imagery acquisitions) allowing for efficient field activity planning. These spatial data sources can be further enriched with other material, such as (a) lists of households, preferably compiled by statistical agency staff as part of the process of delineating

[47] It may be noted, however, that hand-drawn maps by the enumerators demarcating the enumeration area with all the houses and other physical features such as roads and rivers continue to have their utility in countries where the enumeration area boundary is provided to enumerators by description rather than through pregenerated maps.

enumeration areas, but on occasion provided by local leaders (see paragraphs 3.115–3.118); or (b) a textual description of the boundary, including roads, railway lines, power lines, rivers and other physical features. This description may also include obvious landmarks on the boundary (school buildings, water points and other reference features). This can create the base of a rich geographic framework for use as a reference moving forward. In this way, census activities can be used to build out a new national geographic framework/base, or as an important source for the update of an existing national geospatial database. The geospatial data collected during planning and field activities should be treated as an extremely important and rich data source for further geospatial processing.

3.58. Previous experience has shown that relying entirely on a list of households, written or verbal descriptions and directions, paper maps, or local knowledge of the area boundaries will often lead to confusion and error because people tend to have mental images of places that may not align with the area as it is really reflected in the design of the enumeration area. The interoperability of such outputs is also limited. Similarly, the supervisor's mental map of an enumeration area may differ markedly from that of an enumerator, and from enumerator to enumerator. To overcome such problems, it is important that the best possible digital geospatial services be the basis for census enumeration operations and that the collection staff receive comprehensive training in the correct use of their devices and the services available and any associated documentation provided.

C. Integrating census statistics and geospatial information

3.59. A nation's geospatially enabled statistical data can serve as one of the most important tools for decision making. Everything happens somewhere, at a given point in time. Abstract datasets with aggregated information at national level are clearly not as relevant as data about a particular event at a precise location. The ability to integrate statistical data collected at a given location at a specific point in time (like a national census) to a spatially explicit layer, for instance containing enumeration areas), provides relevant and actionable data, ready to support decision-making.

3.60. Careful planning and a forward-looking approach can lead to the proper integration of census statistics and geospatial information, delivering interoperable and integrated statistical and geospatial datasets. This requires also to choose the appropriate geospatial representation of different attributes, e.g. GIS features such as polygons to report attributes of an area, or points features to report and geocode specific locations.

3.61. Geospatial information also serves as the organisational structure for statistical data: in order to be useful, it shall be stored as geographies in geospatial datasets, such as a shapefile containing enumeration areas identifiers and their attributes. Data record identifiers must be unique and options like place names are not viable, as oftentimes place names are not unique at country level.

3.62. This integrated statistical-geospatial framework will also allow for the rapid aggregation and disaggregation of national level datasets. The possibility to work across different spatial scales is indeed strategic in order to be able to address each policy or decision-making question at the proper geographical level of analysis with adequate spatial detail. For this purpose, the spatial data infrastructure should be organised with nested hierarchical levels whereas the more detailed geometry is a subdivision of the upper-level geometry with no gaps or overlaps (for example state, region, province, or municipality, district, parish).

D. Planning the census with geospatial information

3.63. Geospatial information is critical to every aspect of a national census, from the earliest planning to the dissemination of tabulated results. It is critical to understand the scope of the plan, which entails having an accurate map of the area to be enumerated, and at least a general understanding of where the nation's urban centres of population are located, as well as towns and semi-dense areas and rural areas, and the rough order of magnitude of the housing and population within them. From this general understanding, enumeration plans can be developed to allocate staff where needed, logistics planning, and plan for the deployment of adequate support and monitoring resources. As enumeration commences, the monitoring of progress in each enumeration area can be tracked with a variety of geospatial services, to monitor progress and allocate additional support to areas if necessary.

3.64. The use of geospatial layers mapping human settlements and showing residential built-up surface maps can support different phases of the census activities:

- **Planning:**
  - Consistent spatial information about the residential built-up areas provides information regarding the amount and distribution of changes in the residential infrastructure since the previous census activity;
  - Independent, up-to-date and accurate information on the estimated distribution of households facilitates an estimate of the cost of enumeration in the census district and in the country as a whole;
  - Geospatial layers of human settlement extents can help the scaling and distribution of the financial and operational resources to regions according to the number of households to be counted.
- **Preparation:**
  - Fine-resolution geospatial maps of human settlements can be used to determine the distribution and density of built-up surface in a given EA and, as a result, to allocate the means of transport necessary for enumerators to reach all the households in a census unit (to tackle concentration or dispersion of settlements to be enumerated);
  - Classification of administrative units in terms of residential development density or substantial changes in the residential built-up area can support the process of establishing census offices and allocating enumerators and budget in line with the resources required to cover the area within available time and resources.
  - Investigation of enumeration areas in relation to the residential built-up surface facilitates identifying units, which no longer meet the census criteria of the accessibility due to the increased amount of residential built-up surface.
  - Field work: geospatial layers mapping human settlements can be used to provide additional information on the estimated number and the current locations of residential buildings in EA, in case the official information is outdated or incomplete, or facilitate the determination of the most convenient route, covering all dwellings in the area.
  - Validation and Verification: The location of dwellings visited by the enumerators during the data collection phase can be compared with the residential built-up surface distribution in order to estimate the completeness of the collected data.

1. Strategic planning

3.65. Mapping has been an integral part of census taking from the first census. The fundamental act of taking a census is inherently a geospatial endeavour; enumerating a population in a given area, at a particular location, in a specific dwelling, is fundamentally geographic in nature. Over the years, census maps have played a critical role in all processes from preparation to dissemination of the census results.

3.66. The geospatial strategy should be developed at a very early stage of census planning, with the geographic enumeration units serving as the fundamental unit for planning. At least 2-3 years before any census activities have commenced, enterprise level geospatial planning should begin, considering the conditions and available resources of the country. Countries should evaluate available options by considering the following factors: (a) available geographic resources; (b) requirements for new technologies and approaches; (c) available funds and the allocated time frame; (d) staff capacity needed for new approaches; and (e) requirements for developing an ongoing geographic system. Evaluation of these factors will determine the best mix of technology and approaches for each individual case.

3.67. Major technological advances include the widespread availability of handheld devices, global navigation satellite system (GNSS) such as global positioning system (GPS) and Galileo high accuracy service (HAS), and geographic information system (GIS) software, and low-cost aerial, drone, and satellite acquired imagery, and open and free earth observation data[48]. These advances have put new tools in the hands of national statistical organizations to collect more accurate and timely information about their populations. Adopting such new methods can require significant planning and lead times for building new enterprise capacity and organizational restructuring but this allows for the development of an incredibly power geospatial data system that should then be leveraged to support a variety of services and applications beyond census activities.

3.68. There are various approaches for geospatial support services for census activities. For example, traditional mapping techniques have been used successfully for many countries and are still relevant in some countries or at least some parts of many countries, particularly in remote areas. However, with the recent developments in technology, countries are now motivated by new technologies to develop digital mapping techniques and improve the quality of census operation. Application of new technologies requires more careful and long-term operational and managerial plans based on a realistic assessment of costs and human resources required, but costs are dropping rapidly for hardware and software solutions. The proliferation of handheld devices, software vendors with off-the-shelf solutions, and the rise of free and open-source software solutions have driven down the costs of the technical infrastructure. In the case of a lack of internal capacity, outsourcing should be examined carefully before deciding the possible option for census-mapping programmes with the cost of these solutions rapidly decreasing, the amount of geospatial data available, and the prior applications of these solutions to the census increasing.

3.69. There is widespread recognition that it is important for national statistical agencies to develop a long-term geospatial data strategy and begin developing the infrastructure, and the human resources, to serve their nation's data needs. Such a capability can make a major contribution to the population and housing census and other elements of the national statistical system. A continuing mapping capability within the statistical agency can also contribute to the analysis and presentation of census results, web-based mapping and data visualizations, and a variety of geospatial data applications and services. This rich data source can be used as the basic framework for the implementation of a national

[48] Open and free data from programmes such as Landsat and Copernicus, are relevant for several applications including the census.

geospatial infrastructure, providing a common geographic base for the organization of location-based services across the government, standardizing geographic units across a variety of agencies from the national down to local levels. The Integrated Geospatial Information Framework[49] and associated supporting materials should be referenced as an integral part of this planning.

3.70. In the process of creating or updating census maps, census organization should collaborate with other relevant agencies. Statistical agencies are usually not mapping agencies and should not, for the most part, try to duplicate the functions of one. Likewise, mapping agencies are not statistical agencies and often may not fully appreciate the statistical value of the information they hold or how best to present statistical information in map-based products. Despite this, undertaking a census can provide a catalyst for the statistical and mapping agencies to work together to the benefit of both agencies and the community. Even more importantly and at both the global and regional levels there is a continuing initiative to ensure complete integration of statistical and geospatial information as a critical piece of national systems for providing a comprehensive overview of many social, economic and environmental phenomena. The ultimate goal is to develop a global statistical–geospatial framework that would make accurate, authoritative, reliable geospatial information readily available to support national, regional and global development.[50]

3.71. While there is a range of techniques and technologies available for use in a census-mapping exercise, the following sections do not make recommendation as to which system would be most appropriate for specific countries. These recommendations aim at presenting and elaborating on essential principles for developing and implementing an effective mapping component of the census infrastructure.

2. Census geography

3.72. Prior to developing the geospatial support strategy for the census, consideration needs to be given to the geographic classification to be used and the geospatial infrastructure necessary to carry out the required tasks. As the geographic unit used for basic census collection will determine the geography on which the census data can be disseminated, a geographic classification should be devised in conjunction with the development of the geographic frame. The publishing of this geographic classification by the national statistical office so that it can be reused throughout the statistical system and for administrative purposes will increase the value of census data as they will be more relatable to other information. The details of designing a general geographic classification, including the definition of the various areas of the geographic classification and their relationship to one another, are more complex than those involved in census geospatial activities and will not be covered further in this chapter.[51] However, the design of enumeration areas and the careful planning of their relationship to other census management areas is of crucial importance for the census and is outlined in the following paragraphs.

[49] Add reference.

[50] The United Nations Global Geospatial Information Management (UN-GGIM) initiative was launched by the Statistical Commission of the United Nations to assist governments in improving policy, institutional and legal frameworks for developing effective strategies to build geospatial capacity in developing countries. For more details, see http://ggim.un.org.

[51] For a full elaboration and details, please see *Handbook on Geospatial Infrastructure in Support of Census Activities,* United Nations publication, Sales No. E.09.XVIII.8, United Nations, 2009.

3.73. If point-based geospatial data is collected in conjunction with each individual enumeration, then the eventual dissemination unit is not as tied to the enumeration geographic unit. To say it another way, the availability of a precise x/y coordinate for the precise location of each household and can tie each individual enumeration response to a precise coordinate where it was collected, then there is a much greater flexibility in the eventual dissemination units, and those points location can be aggregate up and associated data to any geography.

3.74. It is of critical importance to ensure that the boundaries of various administrative units are available in geospatial format (i.e. shapefile), and frozen at least six months in advance of the census date so that no further jurisdictional changes are affected until the enumeration is over. This would be of considerable help in delimiting enumeration areas and minimizing chances of omission or duplication.

(i) Administrative hierarchy

3.75. One of the earliest decisions in census planning pertains to the administrative areas for which census data will be reported. Administrative areas can be any special geographic unit, but mainly they are units of administration, that is, some governmental authority has jurisdiction over the territory. Census preparation involves creating a list of all administrative and statistical reporting units in the country. The relationships among all types of administrative and reporting unit boundaries should be defined. Every country has its own specific administrative hierarchy, that is, a system by which the country and each lower-level set of administrative units (except the lowest) are subdivided to form the next lower level. Census operations should be seen as an opportunity to update these boundaries, adjudicate any differences. This should be incorporated into basic census planning to allow for enough advance planning for the geospatial activities.

3.76. Only some of these hierarchical levels may have actual administrative roles; for example, the province, district and locality levels may have capitals with local government offices that are responsible for those regions. Other units may have statistical roles alone; that is, they are designed for the display of data and not for administering territory. In some instances, however, administrative units may not be completely nested. Especially when considering both administrative and other statistical reporting units, the census office may need to deal with a very complex system of geographic regions.

3.77. In some regions, the establishment of a definitive list of units is a major operation because of difficulties arising from the frequent fragmentation, disappearance or combination of small localities, and from changes in name, variations in spelling, the existence of more than one name for the same place or the use of identical names for different places. This listing should be held as a formal database or as an integral component of the databases forming part of a geographic information system (GIS).[52] The use of unique identifiers in GIS greatly reduces the complexity of dealing with place names. Such identifiers can be numerical. The availability of a formal database that allows to relate such codes with units facilitates several phases of census, including dissemination.

[52] For further details on GIS mapping, see *Handbook on Geospatial Infrastructure in Support of Census Activities*, United Nations publication, Sales No. E.09.XVIII.8, United Nations, 2009.

(ii) Delineation of enumeration areas

3.78. The design of enumeration areas should take various criteria into account. Correctly delineated, enumeration areas will:

a. Be mutually exclusive (non-overlapping) and exhaustive (cover the entire country);
b. Have boundaries that are easily identifiable on the ground;
c. Be consistent with the administrative hierarchy;
d. Be compact and have no pockets or disjoined sections;
e. Have populations of approximately equally size;
f. Be small and accessible enough to be covered by an enumerator within the census period;
g. Be small and flexible enough to allow the widest range of tabulations for different statistical reporting units (if enumeration areas are being used as the basis of dissemination geography);
h. Be useful for other types of censuses and data collection activities as well.

3.79. For the delineation of enumeration areas, population size is generally the most important criterion, but surface area and accessibility also have to be taken into account to ensure that an enumerator can service an enumeration area within the time allotted. Testing should be conducted to determine enumeration capabilities and capacity. Population density will vary between rural and urban areas and have an impact on the size of enumeration areas, since enumeration can proceed more quickly in towns and cities (densely populated areas) than in the countryside (sparsely populated areas). Given these considerations, enumeration areas that are significantly larger or smaller than the national average may have to be defined.

3.80. As a basic reference, before delineation of enumeration area boundaries, the number of persons living in an area and their geographic distribution need to be estimated. Unless there is information from an exhaustive registration system or some other accurate and recent information source, these numbers need to be determined by counting the housing units, the extent of residential built-up surfaces, the number and the size of households through fieldwork, cooperation with government officials, extrapolation from previous census results, by means of aerial photographs, satellite imagery or derived products such as building footprint and settlement data. There are also freely available high-resolution population estimates produced using geospatial data and population from sub-national projections or surveys to estimate current population distribution (than the sub national resolution). But this data should be used only if the accuracy is of acceptable quality.

3.81. Enumeration area boundaries need to be clearly observable on the ground. Even if they do not have considerable geographic training, all enumerators need to be able to find the boundaries of the area for which they are responsible. Thus, population sizes between enumeration areas may be varied in order to produce an easily identifiable delineation. Natural features that can be used for this purpose are roads, railroads, creeks and rivers, lakes, fences, landmarks or any other feature that defines a sharp boundary.

3.82. Procedures for delineation of enumeration areas should be developed that will allow, as far as possible comparability of areas from one census to the next. Change analysis at the local level is greatly facilitated if the units of enumeration remain compatible between censuses. A unique code should be assigned to each enumeration area and the changes (for example in the case of splitting areas with a high level of growth) should be tracked. The statistical office is often the custodian of coding schemes in the country and should also be the focal point for the design of the census-mapping codes. In cases

where this is not possible, the criteria can outline design principles that will allow users to easily compare enumeration area-based data across censuses. When the population and housing census are conducted separately, effort should be made to use the same enumeration area frame for both censuses as far as possible.

(iii) Delineation of supervision areas

3.83. Supervisory areas provide the means for a group of enumerators to be effectively managed. After delineation of enumeration areas, the design of supervisory maps is usually straightforward. Supervision areas consist of groups of contiguous enumeration areas that share some of the same characteristics as enumeration areas. The enumeration areas assigned to the same supervisor should be compact, in order to minimize travel times, and of approximately equal size. They should be included in the same field office area, which usually is defined according to administrative units.

(iv) Delineation of census management areas

3.84. Census management areas will consist of aggregations of supervisory areas brought together for ease of managing the enumeration staff. Where existing government staff and structure are used for enumeration purposes, the census management areas may be the same as the administrative regions. It should be noted that this may be a matter of administrative convenience and the particular hierarchy (or way of combining enumeration areas into larger areas) for this purpose need not necessarily be the same as that for the dissemination phase, which must be driven by the needs of users. However, the greater the congruence between enumeration areas and pre-existing administrative boundaries the easier is the task of conducting the census.

(v) Geographic coding[53]

3.85. The coding scheme needs to be determined on a country-by-country basis, ensuring codes are unique, unambiguous, and should be designed in collaboration with the national statistical office. The most important principles in the design of a coding scheme are flexibility, expandability and compatibility with other coding schemes in use in the country. The statistical office is often the custodian of coding schemes in the country and should also be the focal point for the design of the census-mapping codes. The coding of each housing or population unit r to a specific longitude and latitude, or to the smallest atomic geographic unit, preferably smaller than the enumeration area, allows for a flexible production of different geographic tabulations of outputs, and production of comparable area-based geography over time. Two somewhat different approaches are available for coding the location of housing or population units. The first approach is to code all units to the lowest-level enumeration area, sometimes referred to as the "enumeration district". The second approach, which at greater cost permits finer geographic specificity, is usually based on some coordinate or grid system, such as that of latitude and longitude or a geospatial grid. In either situation, careful planning should allow for maximum flexibility in future aggregation and disaggregation of the data and geographic units.

3.86. A digital geographic database organised in geospatial layers[54] in vector format consists of a structured set of points, lines and polygons, or grid cells. Each geographic feature — point, line or area — has a unique identifier that is used by the system internally. This internal identifier is not usually

[53] This section draws heavily upon material in *Handbook on Geospatial Infrastructure in Support of Census Activities*, United Nations publication, Sales No. E.09.XVIII.8, United Nations, 2009.

[54] Geospatial layers can also be in raster format, with regular grids of defined resolution.

accessible by the user and should not be modified externally. A more meaningful identifier is needed that can be used to link the geographic features to the attributes recorded for them. For the enumeration areas and administrative units, this link is the unique enumeration area or administrative identifier, which is listed in the master file of all geographic areas relevant in the census.

3.87. Indeed, a unique code needs to be assigned to each enumeration area. This code is used in data processing to compile enumerated information for households in each enumeration area and to aggregate this information for administrative or statistical zones for publication. This is the numerical code that provides the link between the aggregated census data and the digital enumeration area boundary database stored in the case of using GIS.

3. Technology for census geospatial information

3.88. Before census mapping commences, the census agency needs to determine the appropriate technology for mapping. Countries need to choose technologies to improve efficiency of census operations, data quality and timeliness, balancing these with cost factors in view of their national needs and circumstances. The application of technology must also ensure that confidentiality of data is maintained.

3.89. In general, countries need to approach the use of technology for mapping as a continuous process rather than merely a sequence of mapping and dissemination operations. Use and application of geospatial technologies and geographic databases are very beneficial to improve the overall quality of census activities at all the stages of the census. Major technological advances include the widespread availability of personal computers, handheld computers and personal digital assistants, GPS and GIS software, and low-cost or open aerial and satellite imagery. These advances would be of interest to national statistical organizations to collect more accurate data in a timely manner. Software/apps that can read both static and interactive maps should be considered as they will help a lot on deciding the type of maps to produce. Interactive maps are very useful in identification of EAs in countries where structures have no addresses. The choice of the technology (especially the software) should not only be based on cost but also take into consideration the existing skill within the cartography team and needs of the enumerators.

3.90. With the rapid growth in freely accessible high/medium resolution earth observation data, specifically satellite and aerial imagery, there are increasingly few, if any circumstances where hand-drawn maps or hand-collected data should be incorporated as an acceptable standard solution for a modern census. Use-cases from previous censuses and surveys should be identified and a mitigation should be put into place to prevent these situations from occurring without a digital alternative. A paper-based collection solution is not subject to the same level of quality assurance and quality control that can be built into a digital capture solution. In the case of capturing geographic data (such as a housing unit location) on paper, the basic geospatial reference data isn't captured, and may require additional time and research in the office, or necessitate additional field visits.

3.91. In circumstances where it has not been possible to acquire appropriate base maps for areas of geography, enumerators (or other enumeration staff) may produce hand-drawn maps, accompanied by a textual description of the boundary features, to enable a successful enumeration. Hand-drawn maps do not possess the level of accuracy or incorporate the geospatial reference data offered by high-quality digital options, and should be considered an option only when digital maps for an area (a) do not exist; (b) are at too small a scale to provide sufficient detail for an enumeration area map; or (c) are seriously

out of date and cannot be updated in the time available. Even in these situations, paper-based maps should only be used for field expediency, and they should be digitized at the earliest opportunity and the resulting data incorporated into the standard digital workflow. A process for quickly digitizing any field drawn maps should be accounted and planned for, and preferably staff in the field will have the ability to quickly utilize their digital devices to update existing enumeration areas, nullifying the need for any paper-based solutions. The primary issue with these hand-drawn maps is the lack of a precise georeference, which is explicitly provided in a GIS platform.

3.92. Where accurate and current maps at relevant scales are not available for a country, or part of a country, the technological alternatives described in the following paragraphs could be employed subject to consideration of the constraining factors described in paragraphs 3.79–3.82:

a) **Satellite images.** A satellite image typically covers a large area and can be cost-effective compared to other sources. Imagery should be pre-processed by the supplier so that it is rectified and georeferenced (a known scale and orientation, with some latitudes and longitudes, is printed on the face of the image). Satellite image data have gained in volume, popularity and ease of use. Satellite imagery, if used pragmatically, can save countless person–hours by focusing attention on critical areas. Remote sensing data and remote sensing derived thematic maps of human settlements and settlement typologies by degree of urbanisation can be used as an independent check on the field verification process.

b) **Aerial photography.** Acquisition of aerial photographs for large areas of a country may be expensive. However, existing archives of photographs can be an excellent resource for preliminary counts of dwellings and as a base for basic maps. In many cases digital images captured from airplane or drone-based platforms are a cost-effective way of incorporating components of a GIS. However, we should not confuse the use of satellite images and aerial photography as end product, because the information contained in them is raw and has to be processed -adding on the burden/cost of their use. Sometimes, where there is no capacity for additional field operations, or no current digital information is available, open data of human settlements can be a rather good asset to fill geospatial data gaps (see point e below).

c) **Global positioning systems (GPS).** Making hand-drawn maps or digital maps from a GIS for use by enumerators in the field can be greatly assisted by GPS. A simple, handheld GPS receiver (or tablets/smartphone with good GPS capability) will give latitude and longitude coordinates with reasonable accuracy of key points. Depending upon the system selected, a GPS may also track linear features and thus be useful for mapping boundaries. Maps printed from a GIS or hand-drawn map can be enhanced by the addition of latitudes and longitudes recorded at key points to provide orientation, scale and absolute position. Such information will be particularly valuable for dissemination purposes or if the work is a component of developing a GIS for later use. The ability to record information directly without transcription has the benefit of removing several intermediate steps. Coordinates are captured and immediately displayed on the portable computer screen, and if a digital base map is available, the coordinates can be displayed on top. Field staff can add any required attribute information and store these data in a geographic database at the home office. Given that notebook computers and other portable computing devices are becoming less expensive, integrated field-mapping systems are becoming a viable option for field data collection for census purposes. Advances in technology, including GPS, wireless communication and computer miniaturization, have made possible numerous new

applications for handheld GIS, particularly the development of specialized software for census fieldwork.

d) **Georeferenced address registry.** A high-quality, comprehensive, updated and georeferenced address registry of each building and dwelling can give great support in planning and organizing a census. A georeferenced list of addresses can play a central role in many fieldwork operations and will provide the key to accurate delivery, collection and follow-up of questionnaires. The best way to associate each address with a location in physical space of a map is to specify its coordinates in a proper geographic reference system. With geographic coordinates addresses can be entered in available maps or into the GIS. If it is not possible to get coordinates it is recommended at least to geocode addresses. Geocoding is the process of finding associated geographic coordinates from other geographic data. For example, as geographic coordinates of an address, the coordinates of the centre point (centroid) of the enumeration areas to which the address belongs could be taken.

3.93. **Open data human settlement thematic maps and building footprint.** For situations where there is no digital map available, or specific acquisition and processing of satellite imagery is unpractical, open and free global human settlement maps can fill data gaps and serve as information baseline for census activities. These products delineate built-up surfaces and human settlements, are generated via GIS processing and based on open satellite imagery data. These resources are well documented and widely used to delineate human settlements. For some regions of the globe there are also building footprints which increasingly available for free and which can provide an initial georeferenced database of buildings. [add reference]

3.94. Where a digital base map is prepared, this may be used in conjunction with a series of geospatial layers in GIS technology as the basis for coding information supplied in the census. This could apply to address of usual residence now and/or in the past, place of work and similar topics.

3.95. The implementation of strategies using such technologies must be thoroughly planned with the guidance of qualified staff or external experts with formal qualifications in the use of advanced mapping technology. It is particularly important that the cost of acquiring and maintaining the hardware required to use this technology is factored into the budget (and a sound cost–benefit analysis undertaken to support such changes), and adequate plans are made to ensure the availability of sufficient quantities of hardware in time for the census.

3.96. It should be noted that there might be additional risks due to the need for equipment to be operated in suboptimal conditions, including poor weather, dusty conditions or poor lighting. Despite its versatility, GPS may not be able to differentiate the coordinates of overlapping or closely located dwellings in multistorey buildings and in this circumstance should only be regarded as providing coordinates for the building rather than the dwelling units within it.

3.97. It is important to ensure that where such systems are employed they are clearly understood by enumeration staff. This should be achieved by ensuring that the staff, whether at the cartographic update (pre-enumeration) stage or enumeration stage, are given adequate training in the interpretation of the maps. Should the maps be incorporated in digital devices such as personal data assistants, the staff should be trained in the use of both the hardware and the software.

3.98. As with all other significant changes to census procedures, it is crucial that census geographic and mapping processes are successfully included in tests prior to being used in the main operation. This is particularly the case where a change in level of technology is being considered.

4. Geographic information systems

3.99. A geographic information system (GIS) can be seen as a system of hardware, software (commercial and/or open source) and procedures designed to support the capture, management, manipulation, analysis, modelling and display of spatially referenced data, and should be understood as operating at the enterprise level, within the framework of the NSDI. The Integrated Geospatial Information Framework has been developed to support the development of this critical modern infrastructure and should be referenced for planning these activities55. In practical terms, a user's interaction with such a system may range from a simple hand-held or desktop mapping facility or a web-based mapping and digital visualization application. However, careful planning at the enterprise level should be undertaken to plan for a complete GIS system that is capable of solving complex census operations planning, providing solutions for managing operations, identifying and resolving enumeration coverage problems, monitor field operations, enforcing quality controls, producing detailed georeferenced inventories, and supporting a variety of dissemination vehicles, products, and formats. The ability to use geographic space to integrate and manipulate data sets from heterogeneous sources makes a modern geospatial system the critical and the fundamental tool for planning and managing the modern census. For example, a GIS provides functions for the aerial interpolation of statistical data in cases where the administrative unit boundaries have changed between censuses. While, the development and implementation of such a repository of georeferenced data can be daunting and take extensive planning, simple desktop mapping systems generating thematic maps from a database of base maps and indicators will satisfy the needs of most census organizations, the census should be seen as an opportunity to make strides and build upon previous geospatial operations. Once collected, a plan for geospatial data maintenance can be developed, and each census can provide the opportunity for an update and advancing the NSDI.

3.100. GIS technology should be considered only at a level appropriate to the skills and resources available, and should constitute an integral part of the overall work of the organization. Cooperative arrangements with other agencies should be pursued particularly with regard to the acquisition and maintenance of base map data, which should not be the responsibility of the statistical organization. Statistical organizations should proceed with GIS development or implementation only where it is feasible to maintain such a system during the intercensal years and where there is no dependence on external support.

3.101. Statistical offices may nevertheless develop GIS applications with population data and other georeferenced data from other sources for more advanced forms of spatial analysis. The task could be shared with other institutions, or be delegated completely to specialists elsewhere. The role of the census office would then consist in supplying census data at the right level and in the right format for such a system. Census offices provide vital information on current demographic conditions and future trends for policymakers in a range of sectors such as health care, education, infrastructure planning, agriculture and natural resources management; and the provision of spatially referenced census databases is an essential prerequisite for facilitation of the use of demographic data in these fields.

55 Microsoft Word - Part 1-IGIF Overarching Strategic Framework 24July2018.docx (un.org)

3.102. In this regard, it should be noted that the GIS should be capable of generating additional geographic delimitations beyond those used in the census, such as school districts, water catchment areas or power service units. These entities will have to be constructed from the smallest geographically identified units available in the census (for example block faces, grid squares or enumeration areas). If (as is the case in most developing countries) enumeration areas are the smallest unit, this will have important implications for the establishment of enumeration area boundaries. Cooperation with the authorities responsible for these geographic entities before the boundaries of enumeration areas are drawn can reduce later problems.

3.103. With the rise of free and open-source software, the cost/benefit analysis for rapid deployment and implementation of a GIS programme in a national census is irrefutable. However, in some developing countries with limited capacity and resources, GIS may need to be introduced carefully and gradually, given that is a rather complex technology and a resource-consuming one. As an alternative to immediately launching full-scale GIS applications, countries may start with a simple and robust design that is likely to be understood and maintained by a wide array of users, transferable to a wide range of software packages and independent of any hardware platform. GIS implementation in a developing country may follow a hierarchical strategy, with the national statistical office employing a high-end commercial GIS with extensive capabilities for handling and analysing large amounts of spatial data. Widespread dissemination of databases can then be achieved by creating a version of the finished databases using a low-end mapping software format for distribution at low cost and through web dissemination of macro information in an online GIS.

3.104. In recent years, many countries have adopted the use of GIS to facilitate census mapping in the production of both enumeration maps and dissemination products. As the cost is declining and the basic technology is now well established, it is expected that this will continue. It is likely that the census could be a useful catalyst for increasing capacity within the statistical office (or the country as a whole). Adoption of GIS should thus be seen as a major strategic decision with impacts beyond the census operation, and many issues need to be considered. A GIS database, a census geographic database built at the enumeration area level or in grid format, is an important element of the national geospatial data infrastructure for the national statistical office to manage, analyse and disseminate census data, and monitor the continual change in geography between successive censuses. It also constitutes a fundamental component of a national geographic information infrastructure (NSDI) that allows the national statistical office and other national organizations to integrate socioeconomic and environmental data for evidence-based decision-making.[56] A prerequisite to the building of a geographic database at the enumeration area level or in grid format is the development of a geocoding scheme, whereby each enumeration has a unique code, an administrative identifier that can be used to link the geographic features to the attributes recorded for them.

3.105. The (potential) benefits and costs of GIS are summarized as follows:[57]

(a) Benefits

(i) Efficiencies and insights for field preparation for data collection

(ii) Closer linkage between maps for enumerators and map-based products for users;

(iii) Enriched dissemination of census data as they can be visualized in geographic areas for easy understanding by users;

[56] See elaboration on GGIM in paragraph 3.49 above.

[57] For more information, see the *Handbook on Geospatial Infrastructure in Support of Census Activities*, United Nations publication, Sales No. E.09.XVIII.8, United Nations, 2009.

(iv) The cost of intercensal updating of the base map will be less with a digital base map, enabling among other things the construction and updating of sampling frames;
(v) Producing duplicate maps may be less expensive with a GIS solution;
(vi) GIS will have increased ability to undertake quality assurance of geographic boundaries;
(vii) The census agency will have a greater ability to perform spatial queries and advanced analysis under GIS, including the application of the degree of urbanisation method;
(viii) Space needed to store input maps for digital purposes will be far less.

(b) Costs

(i) GIS requires additional technical expertise;
(ii) GIS will require a higher level of computing infrastructure;
(iii) A clerical census system can proceed on the basis of basic maps. However, use of GIS in this task requires that a digital map base exists. If it is necessary to create the digital map base, significant lead times are required as well as significant funding. In both cases, more experienced technical staff are required;
(iv) In most cases, the preparation of maps or GIS will not be the core business of a statistical agency.

5. Implementation of census geospatial information programme

3.106. The development of a mapping system within the census agency requires the coordination of a series of complex tasks with relatively long lead times. It is important that project plans are established to manage this process. The main activities to be reflected in such plans are discussed below:

(a) **Establishing a mapping unit.** The census-mapping project requires a specialized project team. Where mapping activities are outsourced, the mapping project teams will be responsible for specifying the requirements of the census for mapping products and coordinating arrangements with the provider of mapping services.
(b) **Developing a timetable.** The critical date is the date that maps must be delivered to the field. The mapping programme must commence early in the census cycle to allow sufficient time to produce national coverage of maps well before the census date and before training of field staff.
(c) **Sourcing of basic mapping and digital geographic data.** A major step in the mapping project is establishing a base map of the country, including digital mapping data if required. If a census-mapping project already exists, the agency may still require updates to their existing map holdings.

(i) Sources and types of hard copy maps

3.107. Where a hard copy base map is to be used, official published maps may be available from national or provincial government mapping agencies, the local government or municipal bodies. Other sources of maps may be other government agencies or private companies. Where the maps are obtained from sources outside the census agency, permission to use the maps collected must first be sought from the original source and any copyright issues addressed.

(ii) Digital mapping data

3.108. When establishing a digital geographic database, a major consideration is the determination by the census agency of data requirements. With increasing amounts of digital spatial data becoming

available, it is also important that standards and a common data specification be produced to ensure data validity and consistency.

3.109. The key rules to be followed in selecting data items for inclusion are to question whether (a) the data item will be useful to enumerators in navigating their way around their enumeration area; and (b) the data item is relevant to users. Assessing the utility of data items to users in a census-mapping context must place significant emphasis on the user needs for small or customized areas. Data items that meet neither of those criteria should not be included in the database.

(iii) Updating *maps or digital mapping data*

3.110. Preparing or updating base maps, or the base map digital data, requires substantial resources. The final content of base maps will have a major bearing on the accuracy and completeness of enumeration area maps and, subsequently, the effectiveness of census enumeration. The updating of base maps should be scheduled according to priorities, based on areas in which changes to the number or characteristics of the people require the maps to be updated. Important features to be updated include (a) accurately named and presented roads and waterways; (b) administrative boundaries; and (c) landmark features, such as schools, place of worship, post offices, parks and large buildings.

(iv) Operational *design for enumeration and supervisory areas*

3.111. Whether a hard copy or digital base is employed, an enumeration area design manual should be produced that contains the design criteria and the procedures to be followed when designing the enumeration area. The manual can be used as a basis of training for those involved in the design process.

3.112. If possible, enumeration area design should be conducted by regional statistical office staff who are primarily responsible for enumeration areas in their province or region. This ensures that local knowledge can be utilized in the design process. A considerable part of the process is the gathering of information on where population and boundary variations have occurred in order to determine the best way to design and update particular enumeration areas. As an output of enumeration area design, a list should be produced that provides the enumeration phase with all relevant field data for each enumeration area, and the dissemination phase with relevant geographic data.

3.113. The design of field supervisor and management area boundaries can be determined at the completion of the process through the aggregation of enumeration areas, and the allocation of geographic identification codes.

3.114. Quality assurance measures should be implemented to ensure that data are correct to a minimum standard, both for field navigation and for technical correctness in cases where a digital base is to be used as an output medium.

(v) Printing *and content of field maps*

3.115. Careful consideration should be given to the (considerable) time required for printing maps when establishing the project plan for census mapping. Digital options should be given primary consideration and accepted as the default format for any static reference maps.

3.116. Maps should be provided to every level of field staff. Operations should refrain from employing paper mapping solutions, and if used, should be kept to a minimum. The overhead of processing paper maps, the lack of precision and accuracy of updates, and the sheer logistics necessary for handling paper should all be taken into consideration during the earliest stages of planning. If paper maps are used, at least one map must be printed for every enumeration area in the country. It is recommended that two copies of the map be produced, one copy to be used by the enumerator and the other by the field supervisor for training and reference purposes (and subsequently retained by the statistical office as input to the following census cycle). The use of computer technology for data collection, such as tablets, laptops and other handheld devices should be considered as the default solution and suitable for displaying maps available to field personnel.

3.117. Other considerations for the preparation of enumeration maps include the following:

(a) Enumerators may be required to navigate in poor lighting conditions and thus details should be easily read;
(b) The maps must be easily interpreted with text and symbols readily identifiable and correctly placed, along with the information being presented in a standard format compared to other source maps;
(c) Boundaries (such as enumeration area boundaries) overprinted on the maps must be clear and unambiguous;
(d) Enumeration areas must be distinguishable when compared to the surrounding area;
(e) Folding or refolding of large paper maps (larger than A2 in size) is inefficient for staff;
(f) Paper and digital maps need to facilitate the addition of written enumerator comments;
(g) Production of the maps should be cost-effective;
(h) The maps should be suitable for reuse to meet dissemination purposes where this reflects user demands.

3.118. Maps for supervisors or regional managers should be of smaller scale, providing sufficient detail to identify major features. When using paper maps those should not be so large as to be difficult to handle easily in the field. In many cases, the use of inset or supplementary maps may be required if the map is to cover a relatively large area. For all levels of senior field staff, the maps should show the boundaries of all subsidiary units for which they are responsible.

6. Contracting out for census geospatial operations

3.119. The development of a mapping project beyond rudimentary clerical systems requires considerable knowledge of mapping, cartography and geographic systems. In the event that a census agency cannot draw on such skills from within the agency, it may be required to contract out some or all of the elements of preparation of census maps.

3.120. Mapping for field purposes under a contract or agreement basis requires the statistical agency to specify its requirements and prepare clear terms of reference to the contractor. These may include the following: (a) acquiring the base map data; (b) creating (or obtaining) the statistical boundaries and aligning them to the base map; (c) providing a process for enumeration area designers to advise on changes to boundaries (and updates to associated spatial data); (d) producing hard copy maps or soft copy using handheld devices such as a smartphone or tablet computer as specified for fieldwork; and (e) ensuring that digital maps are compliant with standards (i.e. file formats ensure compatibility, portability and exchange).

3.121. The statistical agency should undertake the enumeration area design work and validation of the associated spatial data, as well as take delivery of the hard copy maps for quality assurance checks and subsequent delivery into the field. The statistical agency must also accept full responsibility for the quality standards and delivery of the maps to field staff as required. After the census, any feedback received from enumerators about the base map should be communicated to the mapping agency.

3.122. Mapping for dissemination purposes may be more challenging because the outputs will involve representation of statistical information (with, or as part of, a map) and will often be accompanied by analysis or commentary about the information. Advances in mapping software have made it easier for census agencies to produce a wide variety of standard thematic maps. However, advanced mapping products may require the expertise of a contractor. In these cases, it may be better for the statistical agency to focus on the statistics and let the contractor provide the technical skills required to produce the actual products with tight quality assurance procedures, and statistical disclosure control, in place to ensure that the output from the contractor satisfies the end user requirements described above.

V. Census tests

3.123. The success of a census operation depends substantially not only on the effort spent in designing and building the census infrastructure, but also on the effort spent on testing and piloting all components of the infrastructure. Thus, having systematic testing procedures in place is vital for data quality and for the efficiency and effectiveness of the data collection operation. In order to identify problems and to suggest adequate improvements, the use of pre-field tests (testing under laboratory conditions) and field tests (testing under field conditions), including the pilot census, is indispensable.

3.124. The testing of various aspects of a census plan prior to the enumeration is of critical importance for all countries, and an essential one for countries without a long history of census taking and for those in which fundamental changes in census methods or use of new technologies are being considered. Census tests can be designed for different purposes and in different ways. To yield full benefits, tests should be employed for all stages of the census, including enumeration, processing and evaluation of results. Testing should be conducted in phases to build up to the full test of the holistic design, normally done with a pilot census. The testing phases should begin with testing individual components of the census design. For example, individual software component can be testing during a unit testing phase and the questions for the census questionnaire may be tested in specialized content testing. The testing will build to eventually focus on the integration of all of the components together. Testing of the individual components, such as specific technological solutions (e.g., internet data collection or handheld devices), ensures that those solutions are working as intended and allows for efficient identification and resolution of errors specific to the individual solutions. However, the larger scale testing of the integration of all components is of utmost importance to ensure all of the components work together successfully. Testing phases should be planned and conducted with both pre-field testing and field testing. The work to design all testing efforts, whether pre-field or field, should include the development of test objectives and measures. The test design should reflect those objectives and include the ability to collect the necessary measures.

3.125. Much important testing can be conducted without requiring field work. Technological solutions can be tested using data from previous surveys or censuses to conduct simulated data collection activities. If data processing methods are being tested, for example, those methods can be tested against existing data to determine if the processing generates the expected results. Using existing survey or census data for this purpose is especially helpful as the true outcome is known from the implementation of the previous survey or census. In addition, cognitive testing can be conducted in a laboratory environment with small numbers of participants to gain insight into how a respondent interacts with a questionnaire or how an interviewer interprets their instructions. Findings from these cognitive testing events are particularly valuable in probing for weaknesses in the questionnaire, in the instructions, or in enumeration procedures that might affect the quality of the data.

3.126. Field testing is an important stage to ensuring readiness of all components of the census design. Bringing all of the components together in field test creates the ability to identify problems linked to the integration of the data collection application design and architecture, the data transfer system, and the integrity and security of data transferred. Such tests can also give important information on the adequacy of the field organization, training programme, extent of respondent burden, processing plan, budget, and other important aspects of the census. They can be designed to provide information on the relative efficacy of alternative methods of enumeration and technology, and on the average time required for enumerating a single household or a single set of living quarters. Such information is useful in estimating staff and cost requirements. In addition, census tests serve as practical training for the nuclear staff of supervisors and other officials. One critical form of field testing is a final pilot census, or dress rehearsal, that is a final opportunity to bring together all aspects of the census design on a smaller scale to ensure the design is complete and can be scaled as necessary for the production census. The results of the pilot census can be used to ensure accurate coverage of the target populations and to identify gaps in the design that must be addressed before the production census.

3.127. When carrying out census tests, probability samples of geographic areas or units are not always necessary. Since the purpose of the pilot census and pretest is to judge the operational feasibility of a proposed course of action for the main census rather than make population estimates, purposive samples can often be used for such tests. Purposive selection of one or a few geographic areas is generally preferable for such feasibility testing. Purposive samples are also particularly useful when it is necessary to test census questionnaires and methods in areas with particularly difficult conditions. On the other hand, random sampling procedures are often necessary, depending on the objectives of the test. For example, when overall quantitative measures are needed for comparing efficiencies of different procedures (for instance, in examining the anticipated response errors arising from different systems of enumeration), you may need to design the test to ensure adequate representation of the different systems in the resulting data. Or, if you need to develop estimates of the impact of a test objective on particular populations, your sample design must ensure appropriate representation of the population(s) of interest.

3.128. A critically important test to be carried out during census preparations is questionnaire testing. The purpose is to test the suitability of intended census questions, including their formulation and the instructions provided, as well as the suitability of the questionnaire design. Such tests can be particularly helpful in assessing the suitability of the proposed material for enumerating specific population groups, as well as the general public. For example, these tests can determine if the terminology used in a particular question is familiar to potential respondents, or if the flow of questions is easy to follow for respondents of varying levels of literacy. These tests are also used for estimating the time requirements in enumeration. It is practical to carry out questionnaire tests on a small scale in a laboratory

environment or in several purposively selected places. For example, a location that is known to be populated with particular demographic groups whose interaction with questionnaire is particularly important. Because they are relatively inexpensive, repeated rounds of questionnaire tests may be carried out until a satisfactory questionnaire has been evolved. Questionnaire testing can be conducted with a wide range of sophistication. They can be as simple as conducting cognitive testing in a laboratory environment where potential respondents interact with a paper questionnaire and share their thoughts and reactions to the content. Or, more sophisticated testing, especially with electronic data collection modes, can use leading-edge methods such as eye-tracking technology. The full range of methods is beneficial to design questionnaires more scientifically as reading patterns of respondents can be recognized technically through these types of tests.

3.129. A comprehensive test of all census procedures is often called a "pilot census." Such large-scale tests should be designed and managed to thoroughly test all census components and the entire census infrastructure. Essential features of a pilot census are coverage of one or more sizeable administrative divisions and encompassment of the preparatory, enumeration and processing stages of a census, by which it thus tests the adequacy of the entire census plan and of the census organization. In order to best serve this purpose, care should be taken to ensure that conditions in the pilot census are as close to the conditions that would be present during the actual enumeration as possible. For this reason, it is often taken exactly one year before the planned census so as to conform to the expected seasonal patterns of climate and activity and to be conducted as close in time to the production census as possible. It is generally unwise to consider the pilot census a source from which to derive usable substantive data., Apart from the sampling problems involved, such a use inevitably detracts from the central purpose of the pilot, which is to prepare for the main census. However, a pilot census can be an ideal opportunity for testing the analysis and tabulation processes necessary to generate statistical data products. The area(s) where the pilot census is conducted cannot be excluded from the production census, but pilot census participants may be of the opinion that they already responded to the census when they are contacted for the production census. So messaging and notifications around the production census must combat this thinking and include guidance to pilot census participants that their participation in the pilot census did not satisfy their participation in the production census.

3.130. It is critically important to undertake a set of tests of the information and communication technology (ICT) solutions and production systems that are planned to be applied in the census. Depending on the extent and characteristics of ICT, these tests should include all ICT components related to the fieldwork and to data transfer, entry and processing well ahead of the census itself. This is particularly important if a new technology is being introduced, such as the Internet, handheld devices or other electronic collection tools, and scanning the questionnaires as a means of capturing data. Tests should include the testing of applications, systems and the equipment itself, as well as the underlying circumstances necessary to avoid equipment malfunctioning, such as climate, or significant delays due to inadequate quality of paper causing paper jams or unexpected problems in programming activities. In the context of new approaches using electronic data collection, testing should include daily data transfers to the major depository of data and validation of those data deliveries. Testing the efficiency of data entry rules, coding, editing and tabulation applications should be done based on results collected by the pilot census.

VI. Establishing the census frame

3.131. Prior to implementing a census, it is important to establish the frame or universe for the census. In most cases, this frame will consist of all living quarters and/or structures containing living quarters. The frame can be constructed before the census using listing activities in the field, can be developed from existing address and/or person registers, or can be build from available information such as aerial imagery and GIS data files. Field listing activities generally involve listers traveling all roads and neighborhoods in the area for the census to collect the address or location of all living quarters and/or structures. This listing activity can be done with or without contact and interviewing of residents or representatives of the living quarters/structures, but contact is recommended to ensure the most accurate information about the living quarters/structure as possible. Field listing is often conducted using paper address listing tools and paper maps, but can also involve automated tools.

3.132. For automated address listing activities, the enumerator collects data from every household in every concessions of its enumeration area (EA). It is in digital format, with a georeferenced list of concessions and basic infrastructure/equipment and will be loaded onto the tablet. Armed with a tablet containing a pre-loaded list of the EA's concessions, the enumerator begins by identifying the boundaries of his EA and then making it a reality. This means updating the list of concessions and the households they contain. Changes in the list of concessions are recorded on the map of the EA that has been given to him for this purpose. However, they may also be asked to update the infrastructures and addresses located in their EA by specifying their geographical coordinates using the tablet's GPS.

3.133. The frame that is available at the start of the census is an instrument for the control of the enumeration, particularly in the absence of adequate and updated maps. The frame is also useful for estimating the number of enumerators and the number of schedules and other census materials needed in an area, for estimating the time required for the enumeration and for compiling provisional results of the census. It is also very useful for determining the enumeration areas and for establishing necessary links between population and housing censuses when they are carried out separately. Finally, as the frame represents the full universe of living quarters or households that should be enumerated, it can be used as guide to monitor the completeness and quality of the enumeration of the population in a given area.

3.134. Consideration should be given to providing permanent identification to streets and buildings, which can be used for successive censuses and for other purposes. A listing of sets of living quarters, particularly in densely settled places, cannot be made unless streets have names and buildings have unique numbers. Individual apartments in multi-dwelling buildings need to be numbered or otherwise unambiguously identified. Where these prerequisites do not exist, numbering immediately prior to the census would prove useful.

3.135. Having a complete and accurate address frame ensures the ability to provide the enumerators with assistance in the form of lists of addresses to visit. Having the full list enables the enumerators to recognize known structures and identify new structures that may need to be added to the list and enumerated. Address lists will be essential if self-enumeration, whereby questionnaires are sent to the households by mail, is part of the plan. Some countries have and maintain address or population registers that allow more or less complete address lists to be generated relatively simply. The census can then not only use these lists, but also assist in further improving the population register by reporting any

discrepancies found in the field. Where official population registers are not available, or insufficiently complete, it may be possible to obtain additional address lists from postal authorities, utility companies or the private sector (for example, mail order companies). A definitive list for the enumerators could then be prepared by the central office by merging and unduplicating the lists obtained from these various sources.

3.136. Where a functioning population register exists, it may be possible to prefill the household questionnaires with information such as the names of the persons expected to be members of a household, already available from the register. This reduces the response burden, accelerates the information-gathering process, and helps to pinpoint deviations. On the other hand, it might have a negative psychological effect if respondents believed that the authorities were monitoring them too closely. To combat this negative effect, the census organization should develop proactive messaging and notification to the respondent universe about the intent to use register data and clearly articulate the benefits of such an approach. Using one or several registers as the point of departure for a census that still includes full coverage field enumeration is an approach applied in some countries; differences between the register(s) and the field situation will necessarily come to light, and rules will be required to deal with such differences. The differences in the data from registers and the data from field data collection can have various sources. They could be due to a difference in vintage, for example, if the register data are from a previous time period and therefore are not as current as the data collected in the field. Or, the differences could come from definitional differences, such as if the register data represent persons "registered" at an address but do not necessarily reflect the true household members. An understanding of the cause of any difference is critical to ensuring that the merging of the records data and the field-collected data results in accurate and meaningful census results.

VII. Field enumeration

3.137. Recent years have witnessed substantial changes in how countries are conducting field enumeration, resulting mainly from the use of technology in census data collection. While the traditional method of enumerating the population with face-to-face interviews remains widespread, these recent changes show that it can be applied in different ways, including by use of a paper questionnaire, either exclusively or in combination with electronic data collection. Increasingly, more and more countries have been able to carry out their census electronically, making it possible to automatically capture data during enumeration. On the other hand, self-enumeration methods can also be applied in different methods using the Internet. The use of technology during enumeration would be the main challenge for most countries. It should be noted that only countries that have high penetration rates of information technology (including the Internet) have implemented Internet data collection, and mainly in conjunction with more traditional methods. However, these options may never entirely replace face-to-face enumeration, as even where societies enjoy a high degree of information technology use, the entire population cannot reasonably be expected to comply to a mode of self-enumeration.

A. Method of enumeration

3.138. There are two major methods of enumeration: (I) the face-to-face (or enumerator) method, whereby information for each individual (in a population census) and for each set of living quarters and

the occupants thereof (in a housing census) is collected and entered in the questionnaire by a census official; and (ii) the self-enumeration method, whereby the major responsibility for entering the information is given to respondents, such as, a person in the unit being enumerated (usually the reference person of the household), although the questionnaire may be distributed, collected and checked by a census official.

3.139. Traditionally, each household is contacted and enumerated on a face-to-face basis. This approach is still used in most developing countries and for at least part of the population in many developed countries. In circumstances where up-to-date and comprehensive address or population registers exist or can be established and the level of literacy is high, the enumeration process often involves mailing out the census forms or having the public mail back the completed forms. Where telephone and internet services have broad coverage, telephone and Internet data collection can also be used. Approaches for self-enumeration using different modes of enumeration, such as mailing, telephone and internet data collection, may also be used in combination with the face-to-face method. In some countries, Internet data collection and postal distribution of the questionnaire, with or without postal return, are used in conjunction with the self-enumeration method. Both procedures can be used exclusively and/or combined with checking by a census official. Whatever approach is used, the complete enumeration plan should be prepared and adequately tested well before the enumeration begins. This involves (a) the determination of the enumeration method(s) to be used, and whether there are differences by specified geographic area and/or population groups; (b) basic procedures to be followed in the collection of the data and the control of the enumeration; (c) the procedures for quality control of the data; and (d) an estimation of the number of sets of living quarters and the probable size of the population to be enumerated so that the number of questionnaires, handheld devices and other materials required for the enumeration, and the number of enumerators and supervisors needed, as well as the cost of enumeration can be properly ascertained. If more than one enumeration method will be offered, it would have to be decided whether the methods will be offered in phases or simultaneously. A phased approach may be used if self-enumeration is expected to be significant.

3.140. Each method has its own advantages and limitations. The face-to-face method is the only method that can be used in largely illiterate populations or in other population groups that may be unwilling to complete the census forms themselves or find it difficult to do so. On the other hand, in countries where literacy is virtually universal and educational attainment relatively high, the self-enumeration method may often yield more reliable results at substantially lower costs, particularly if Internet data collection or a mail out/mail back procedure can be used. However, the postal services may be used to distribute the census forms only when a comprehensive and up-to-date list of addresses is available or can be prepared. Another consideration is the emphasis to be placed in the census on obtaining responses, whenever possible, directly from the person concerned. The self-enumeration method allows for, and its instructions may encourage, at no extra cost to the census organization, consultations among family members when they complete the census form. In contrast, with the face-to-face method it may be prohibitively expensive to encourage enumerators to go beyond even the "first responsible adult" in each household, that would provide the required information on each member of their household. In the light of these considerations, it may sometimes be desirable to rely on one method for enumerating most of the population and to use another method in certain areas or for special groups of the population. With the advance of information technology, the penetration of the Internet has increased in recent decades. In these circumstances, it is recommended that Internet survey methodology should be explored depending on national circumstance. This method can be cost-effective as the expense of printing questionnaires and wages of field staff can be reduced. Also, self-enumeration through the Internet can secure the privacy of respondents, so it would be welcomed as

more and more people prefer to protect their privacy. However, a combination of a traditional method and Internet survey can result in duplication during enumeration. Therefore, careful consideration of the management of the dwelling/household list is essential. Overly complex designs should be avoided, and adequate quality checks introduced to avoid duplications and frauds.

3.141. The decision regarding the method of enumeration to be employed should be taken at an early stage on the basis of thorough testing of the various alternatives in terms of the costs involved, the quality of the data produced and their operational feasibility. Even where a method has been followed traditionally, it is best to periodically reassess its relative advantages in the light of current census needs, changing techniques, and emerging technologies. An early decision is required because the method of enumeration used affects the budget, the organizational structure, the publicity plan, the training programme, the design of the questionnaire and, to some extent, the kind of data that can be collected.

3.142. Challenges that affect or hinder the ability to achieve a response should be fully considered when developing the enumeration design and methods. The design should, as far as possible, reflect particular activities or actions that seek to increase the likelihood of receiving a response. Challenges can be grouped into two types: people and physical challenges. People challenges tend to focus on particular subgroups of the population that tend to be difficult to enumerate for a number of reasons. Physical challenges are those related to the type of environment in which the people live. The two are not necessarily mutually exclusive, but each require different methods to overcome the challenges.

3.143. The section on groups that are difficult to enumerate (paragraph 4.48) gives a thorough overview of the more common population groups that are difficult to enumerate and how these might be addressed during the enumeration. Population groups that are difficulty in enumerate include people with language difficulties, nomads, migrants, students and the elderly.

3.144. People **with language difficulties.** Not all respondents will speak or understand the language(s) in which the census is being conducted. Therefore, consideration need to be made on translation services and materials, with particular consideration given to understanding the types and concentrations of languages required. To the extent possible, decisions around languages/translations to support and provide should be data driven to avoid perception of bias or discrimination.

3.145. **Nomads.** To successfully carry out the enumeration of nomads, it is particularly necessary to pay full attention to the preparatory work in order to determine suitable enumeration techniques. It should be pointed out that there is no absolute methodology for the enumeration of nomads since conditions vary from country to country. The particular method suitable for a country undertaking to enumerate nomads as part of the census should be determined only after a detailed preliminary study and after undertaking a field testing. Some of the methods used to enumerate nomads and semi-nomads may be classified as follows: (a) group assembly approach, (b) tribal or hierarchical approach, (c) enumeration area approach, (d) water point approach and (e) camp approach. Sometimes a combination of two or more methods may be used.

3.146. In the group assembly approach, the nomads are asked to assemble at particular interview sites on certain fixed dates. This method can be adopted only through the administrative or tribal authorities. The tribal or hierarchical approach is a favourite method, since the nomads usually follow what is dictated by the tribal or hierarchical chief. The enumeration work can be carried out as a kind of administrative census by contacting the tribal chief and collecting, sometimes from memory and sometimes from a register, all the needed information on the chief's followers. The other approach is to

contact those followers with the assistance of the chief or a representative and to collect the necessary data directly from the household. In this case, the unit of enumeration is not a real but tribal. The enumeration area approach presupposes creating conventional census enumeration areas and then contacting each nomadic household that happens to be staying in the enumeration area during the census. In the water point approach, a list of all water points available to the nomads during the period of enumeration is prepared. Since numerous temporary water points are created during the rainy season, a meaningful list of water points may be prepared with reference only to the dry season. The enumerator is given the task of locating and visiting every nomadic household that may be using a certain water point. In the camp approach to enumerating nomads, a list of camps is prepared together with the approximate location of each within the country, and enumerators are sent to visit all the households in each camp.

3.147. **Migrants.** Effective communication is essential for engaging migrants in the census. This can be done by providing materials and assistance in multiple languages, working with community organizations that work with migrants, and addressing concerns about immigration status. For example, materials can be translated into the languages that are most commonly spoken by migrants in the area. Community organizations can help to spread the word about the census and encourage migrants to participate. And enumerators can be trained to answer questions about immigration status and to assure migrants that their data will be kept confidential. Flexible enumeration methods, such as online surveys, phone interviews, and in-person visits, can also be used to accommodate the diverse needs of migrants. And ensuring that enumeration venues are safe and comfortable will further contribute to a positive experience for migrants.

3.148. **Refugees and internally displaced persons (IDPs) in camps**. In many censuses there are several questionnaires, including one for conventional/private households and another for institutional or collective living quarters. These institutions often comprise camps for IDPs or refugees. Institutional questionnaires may not be suitable for camps for the displaced, particularly where these are more than short term arrangements. As the question content is often much shorter in the institutional questionnaire than in the household questionnaire, the information will be insufficient to provide for the requirements of users interested in displaced people. Wherever possible the camps for refugees and IDPs should be enumerated using the questionnaires designed for private households. The census communication plan should include messages to convey that the census will include displaced people, refugees and IDPs and that their responses are important and to assure the confidentiality of data collected. It is important to ensure that communication is available in languages relevant to displaced populations, as refugees and IDPs may speak a different language than the local population.

3.149. **Students.** Tailored approaches are needed to reach students, such as collaborating closely with schools, colleges, and universities. This can help to ensure that students are aware of the census and that they have the opportunity to participate. Clear and concise communication materials that resonate with student life and priorities are essential. For example, materials could be developed that focus on the benefits of the census for students, such as how it can help to improve education funding and opportunities. Appointing student ambassadors can also be an effective way to promote the census within student peer groups. Leveraging online platforms, social media, and mobile applications is crucial for engaging tech-savvy students, who are more likely to respond through digital channels. For example, the census could create a dedicated website or social media page for students. And students could be encouraged to share census information with their friends and classmates through social media. For students residing in large communal living spaces, alternative enumeration methods may be necessary

to ensure accurate data collection. For example, enumerators could conduct group interviews in dormitories or student housing hostels.

3.150. **Elderly people**. Additional support and tailored strategies are essential for engaging the elderly population in the census. This can include providing supplementary materials, such as event calendars, to help them recall or estimate their age. It can also engage involving supporters, such as family members, village elders, or residential home staff, in the enumeration process. For example, family members could help to translate census materials or to answer questions from elderly relatives. Enumerators should be trained to conduct in-home visits for elderly individuals who are unable or unwilling to travel. And enumeration venues should be wheelchair accessible and provide a comfortable environment for elderly participants. Assistance could be provided through call centres as well.

3.151. The more common physical challenges that need to be taken into consideration when developing methods for field enumeration include access-controlled properties and rurality.

3.152. **Access-controlled properties.** Some properties, establishments, communities or compounds have controlled access, presenting difficulties in gaining access to undertake an interview or to deliver or follow up on a questionnaire. Access control mechanisms may include locked gates with an intercom to each individual dwelling, commercial farms where the gates are unmanned and are some distance from the dwelling units or gates or doors managed by a concierge or security guards. When developing enumeration procedures, advice needs to be given as to how to gain access and actions to take if access proves difficult. Some of the activities may include building a relationship with the owner of the properties to approve access to engage with residents; using the postal service to deliver questionnaires to these properties; and additional communication methods (such as a letter informing residents about the census and how to complete their questionnaire or inviting them to arrange a particular time to complete their questionnaire via interview).

3.153. **Rurality.** Understanding the extent of rural populations and the associated logistical and management challenges with running a collection exercise in these areas needs careful consideration. Lack of infrastructure such as access to electricity and Internet connectivity may limit the modes of data collection that could be used in these areas.

3.154. The enumeration of other groups such as the homeless and those in institutions also require special consideration.

B. Multi-mode data collection

3.155. Countries are becoming more innovative in how they conduct population and housing censuses in order to improve response rates while also reducing census costs. One of these innovations is the exploration of alternative approaches to the traditional method of face-to-face canvassing of the whole country to enumerate the total population during a census enumeration. The introduction of these alternative approaches entails also use of a combination or combinations of modes of enumeration or mixed-mode collection of census information. Census mixed-mode data collection involves the use of two or more data collection modes to collect census data. These modes include various combinations of (i) interviewer-administered modes, such as face-to-face interviews or interviews over the telephone, and (ii) self-administered modes, whether online or by mail. Interviewer-based data collection modes include: the traditional face-to-to-face personal interview with paper and pencil (PAPI), computer assisted personal interviewing (CAPI), and computer assisted telephone interviewing (CATI). On the

other hand, self-administered data collection modes include: computer assisted self-interviewing (CASI), and paper questionnaire with self-enumeration (PASI). For more detailed information on these modes of data collections, see Chapter E of the *Guidelines on the use of electronic data collection technologies in population and housing censuses*.[1]

3.156. Application of multimode data collection entails the uses of two or more of the modes of data collection listed above, either concurrently or sequentially. In concurrent mixed-mode data collection, respondents can choose between one of two or more modes, such as completing a paper questionnaire or providing the information online via the Internet. In sequential mixed-mode data collection, all respondents are first requested to provide information in one mode but then offered other modes to increase the response rate. Sequential mixed-mode designs start with the least expensive mode and progress to more expensive and persuasive modes. Both options aim to improve coverage and response rates, for the general population and for special groups, while completing data collection at reasonable costs. It should be noted, however, that while mixed mode data collection methods are effective at reducing non-response, they could lead to measurement error due to measurement differences arising from the use of the various modes.

3.157. Mixed-mode data collection offers several benefits over the traditional single-mode of data collection. First, offering different modes of data collection, to different populations, representing a broader cross-section of the population, has the potential to increase coverage and reduce respondent burden by increasing accessibility to the means of census data collection, and also reduce the cost of the census. For example, some individuals may prefer to complete the census questionnaire online, while others may prefer a telephone call, or a paper-based questionnaire. Also, providing various modes of data collection would enhance the ability to reach respondents who are difficult to reach with the use of a single mode of data collection, such as persons living in remote areas or those with no internet access. This inclusiveness has the potential to improve the accuracy and representativeness of the collected data. Also, mixed-mode data collection can help to reduce census cost particularly when some of the tasks are automated, such data entry through electronic data collection. Additionally, multimode data collection can help to improve data quality by reducing the risk of errors associated with any single mode.

3.158. There are some potential challenges to the use of mixed-mode data collection. It is more complex to implement mixed-mode data collection than it is for single-mode data collection. This is because mixed-mode collection requires more intricate planning and coordination for multiple modes of data collection. Also, once data has been collected from multiple modes, it needs to be integrated into a single dataset, and this can be challenging, as different modes of data collection may produce data that may not be easily comparable. When merging census data from different collection modes, it is important to be aware of the potential for bias. This can be due to a number of factors, such as differences in response rates by mode of data collection. As a result, comparability of data from various modes of data collection may be a challenge due to “mode effect”[1] and has implications for the quality of data that has been collected. Understanding mode effects on coverage, response, and measurement can help develop methods to minimize them, and to minimize mode-associated bias.

3.159. Mixed-mode data collection often relies on technology, such as the Internet and mobile devices, and this can be a challenge in areas with limited access to technology. Use of multiple modes of data collection can be more expensive than single-mode data collection, especially if new technologies are used. It is worth noting, also that not everyone has access to the internet or other technologies needed to participate in online data collection. This can lead to underrepresentation of certain groups in the

census data. Despite these challenges, multimode census data collection is becoming increasingly common. This is because it offers several advantages over single-mode data collection, and offers the possibility to compensate the weaknesses of one mode with the strengths of another. By carefully planning and implementing multimode data collection, census bureaus can improve the accuracy and completeness of their data while also reducing costs and respondent burden.

3.160. The management of mixed-mode data collection operations is a complex task due to the complexity of designing questionnaires for interviews and for self-completion, monitoring of enumeration, as well as the validation and integration of data collected by different modes. The more options that have been employed for data collection, the more complex the operation becomes. Consequently, an integrated management information system should be set up for executing, monitoring, and reporting functions related to operational procedures, including scheduling, contact management, case status management, avoiding duplication of responses, managing switching/transfer of a case between data collection modes, response management, assignment of field staff, and response integration. Furthermore, a dedicated management team is needed to monitor and control the multi-mode collection operation.

Designing a census multi-mode data collection

3.161. A crucial part in the implementation of a multi-mode census data collection relates to decisions on the modes of data collection and how they are to be implemented in a national data collection. It is the case that such decisions should be based on national circumstances given the diversity of national experiences and capacities to successfully implement the various options related to census multi-mode data collection. Hence, there is no “one-size-fits-all” option that can be recommended for international use.

3.162. Regardless of the modes included in the combination for a multi-mode census data collection, however, there are pertinent factors that should be considered in its design. One such issue is the need to have a clear understanding of the target population as well as its characteristics, and how appropriate the modes under consideration can easily be implemented. In addition to national differences in abilities to implement the various combinations of modes of data collection, it is also the case that even within the same country there may be variations in ability based on the level of education of different population groups as well as differences by geographic area of residence of the population. For example, self-enumeration by mail or by use of the Internet may not be appropriate where a large proportion of the population is not literate, as well as for those living in remote areas of a country and in other areas where access to the Internet or to postal services may be very weak or non-existent.

3.163. As an input into the choice of modes to be included in a census multi-mode data collection, the census agency, therefore, would need detailed information about the characteristics of the national population, such as age, education, and location. This would enable the selection, development, and tailoring of multiple modes of data collection to include only those modes that can easily be implemented in the country. A further step would be to develop a mode assignment strategy, or for modes to be concurrently assigned to different sub-groups based on well establish criteria on the ability of the corresponding population groups to easily and accurate provide the information that is being requested. In a sequential design, the census office needs a strategy, not only on the various modes of data collection to be offered, but also on the order that the different modes will be offered in order to make it convenient for the population to take part in the census.

3.164. Another issue that should be paid attention to in designing a multi-mode data collection is ensuring that the questionnaire or data collection instrument is designed in a compatible way across the modes of data collection to be used. For example, question wording and format, response options, as well as the layout and design of the data collection questionnaires should be consistent and comparable across modes to be used. This would enhance the comparability of the collected data across the various modes.

3.165. Designing a census multi-mode data collection entails much more than is the case for a single mode data collection. Designing a multi-mode data collection may entail also an option for respondents to easily switch modes during data collection, such as, data collection initiated in one mode (e.g., Internet) being completed in another mode (e.g., paper or telephone). If this option is offered, the census office should develop a strategy to monitor and manage switching/transfer of a case between different data collection modes during a census. A strategy is also required for handling of cases to ensure completion of entry of information where, for example, a questionnaire may not be completed in one go, particularly when the household has several members.

3.166. It is, therefore, important to pilot test the modes of data collection process using the intended modes in order to identify any problems and make adjustments to the questionnaire and the data collection process. For example, the pilot test would help ascertain the compatibility of the modes of data collection and the collected data, as well as to whether the questionnaires in the different modes are easy to understand and complete, regardless of the mode of data collection.

C. Use of administrative data during enumeration

3.167. The use of administrative data can facilitate the enumeration in a traditional census in a number of ways. One of the ways in which administrative data could be used during enumeration is in prepopulating electronic census questionnaires. Such an approach can reduce the workload of enumerators and the need for respondents to provide basic demographic or information that is already found in administrative data sources.

3.168. Also, the quality of administrative data is sometimes required to be assessed. In such cases, administrative data can be used to prefill the census questionnaire to be updated or verified by enumerators - in field data collections - or individuals - in internet data collections. This will reduce the response burden, accelerate data collection, and assist pinpointing deviations. However, this approach may have negative effect on respondents that have privacy concerns.

3.169. In certain cases, necessary information for particular subpopulations, such as the members of institutional households, like students or workers dormitories, old people's homes, assisted-living facilities and welfare institutions, military barracks, correctional and penal institutions, religious institutions, etc., are available from the administrative data sources. Using these data has the potential to effectively mitigate or substantially decrease the extent and complexity of field operation, especially when direct data collection for a subpopulation is problematic.

3.170. There are also cases where availability of administrative data assist in facilitating the field operation in other ways. For instance, as mentioned earlier, sometimes prior to the field operation households that are likely to be missed by enumerators, such as the ones that have recently relocated, or households who are elderly, migrants, located in remote or hard-to-reach areas, indigenous communities, etc. can be identified to devise proper strategies to cover them. Or administrative data

may be used to determine whether an address is occupied or not, which will assist in planning follow-up visits to non-respondents.

3.171. Furthermore, when administrative data sources provide an accurate aggregate estimate of the number of housing units/households/individuals in a given area, this information can be used as benchmark to monitor actual performance in the enumeration.

3.172. When administrative sources provide data for specific variables that are of high quality, they can eliminate the need for collecting certain information directly from individuals. For instance, information on marital status, employment status, housing tenure, vehicle ownership, year of immigration, and more can be available from administrative sources. This approach becomes particularly relevant in contexts involving culturally sensitive topics like infant mortality or confidential information like income, both of which can be obtained from administrative sources such as civil registration or tax records. This will both reduce the cost and facilitate the field operation and significantly enhance the data accuracy.

D. Timing and length of the enumeration period

3.173. The choice of the time of year in which the census will be taken is of great importance. The main consideration should be to select a period in which the census is likely to be most successful and to yield the most useful data. This may depend on a number of factors. First, it is necessary to avoid those seasons in which it will be difficult to reach all inhabited areas because of rains, flooding, snow and so forth or in which the work will be particularly arduous, as is the case during extremely hot weather. Second, a time should be chosen when most people are staying at their place of usual residence; such a choice will simplify the census operations both in a de jure and in a de facto enumeration, and it can make the results of a de facto enumeration more meaningful. Seasons of peak agricultural activity should be avoided because it is difficult to interview persons who work late every day and who may even stay nights on their land if the land is far from home. Great traditional festivals, pilgrimages and fasting periods are also unsuitable times for census work. Since in many developing countries the bulk of the field staff is recruited among schoolteachers and older students, the conduct of the census may be feasible only during school holidays, though, as already indicated, the days of major festivals should be avoided.

3.174. In a country that includes areas of sharply contrasting seasonal patterns of weather or activity or in which potential census personnel are in very short supply, it may be necessary to enumerate different parts of the country at different times or to enumerate the nomads or other special population groups at a different time from that established for the settled population. This, however, is generally not a very desirable solution because nomads cannot always be clearly differentiated from settled population because there may also be mobility among the settled inhabitants. Furthermore, such a solution creates complications with respect to the use of the census data due to the varying currency of the data. The census design should consider these complications and risks and attempt minimize negative impacts to or biases in the data.

3.175. When a census has been undertaken and the census date is found to have been, on the whole, satisfactory, subsequent censuses should be taken using the same census date, unless there are strong reasons for changing this date. A regular census date enhances the comparability of the data and facilitates analysis. The tradition of a fixed census date in a country also provides administrative discipline, motivating and enabling all those involved in the census to make necessary preparations in a timely manner.

3.176. It is desirable to keep the census enumeration period short in order to avoid double counting and omissions, which can occur in spite of a single reference date. On the other hand, the shorter the enumeration period, the greater the number of field staff that have to be recruited, trained and supervised. This increases the cost and may lower the quality of the data. If the census is using automated methods for in-person enumeration, the increase in costs is further exaggerated by the need for additional devices for the additional staff. How these different considerations should be reconciled depends on the size and nature of the country and on the resources at its disposal. The length of school holidays is sometimes a restricting factor, although governments of several developing countries, recognizing the importance of a census to their nations, have prolonged the school holidays in the census year in order to allow teachers and students to work on the census as long as required.

3.177. While it is desirable to design the census enumeration period to be of optimal length and to conduct the enumeration at the optimal time of the year, the design must also include flexibility and contingency measures to enable adjustment in the case of an unexpected emergency or disaster. If an unexpected event makes it impossible or impractical to conduct the enumeration period as planned, an alternative timing or length must be implemented. In this case, it is imperative to attempt to understand and quantify the possible impacts to data quality and the comparability of the census data over time. Lengthening the enumeration period introduces the risk of recall bias, duplicate enumerations, and/or omissions. Changing the timing of the enumeration period can increase the risk of omissions if some populations (e.g., migratory workers) are no longer accessible to the field staff.

3.178. In recent censuses, most developing countries have allowed about one week to 10 days for the training of enumerators, while the enumeration period has generally varied from a few days to two weeks. In the case of using a self-response method through the Internet, the enumeration period can take longer, and enumerators may require less training if they are simply going to follow up on non-respondents. Short periods are often feasible in small countries while longer periods may be necessary in large countries with poor communications or widely dispersed populations.

3.179. One method sometimes used to allow sufficient time for enumeration and yet make the census simultaneous is first to enumerate the population over a longer period, say a week or more, and then, in one single day, to recanvass all households, deleting and adding persons as needed to update the files. This procedure is, however, not practicable in very sparsely settled areas and can be very costly to deploy a large number of staff on the single recanvass day.

E. Management and supervision

3.180. A comprehensive and elaborate management system is necessary for resource management and providing timely managerial advice to the field staff. There is no unique approach for the management system; however, it is important to develop a hierarchical and geographically dispersed system for building a direct and effective communication mechanism between the managers and the field staff.

3.181. Adequate supervision of the enumeration is essential for ensuring the quality of the field enumeration. Many countries use a field supervisor to enumerator ratio ranging from 1:10 to 1:15 for field enumeration. Periodic control of the quantity and quality of the work accomplished by enumerators and other field staff is recommended, in order to facilitate the correction of inefficiencies and to maintain satisfactory progress during the enumeration period. Periodic and systematic

assessment should be carefully organized to ensure the quality of the work and also for collecting appropriate information about the enumeration progress for management and supervision of the fieldwork.

3.182. Each staff member involved in the management and supervision system should have a clear role and responsibility and should be fully trained for possible problems occurring during field enumeration and their solutions. For an efficient system, it is important to give clear instructions to the field staff on how to perform their own responsibilities. It is important to note that methods and technologies used during enumeration have a direct impact on the roles of managers and supervisors; consequently, a complete understanding of the characteristics and operational aspects of both the enumeration method and the enumeration technology is a prerequisite for efficient supervision of the enumeration component of the census.

3.183. Depending on the communication facilities and other infrastructure available in a country, different mechanisms for exchanging information among managers and field staff need to be developed. These mechanisms are important for ensuring consistent dispatch of field instructions and also sharing best practices, particularly for finding solutions to unexpected problems during the field operation. The use of mobile phones and accompanying technologies, such as SMS, USSD and other enterprise level communication platforms significantly increases communication capabilities, more especially when such communication needs to broadcast simultaneously to allow the field staff getting one message from the source.

a) Management information system

3.184. A management information system for the field operation should be established to collect information needed for timely management and supervision of field operations. To establish this system, the following steps can be considered:

i. Determining information needed for supervising and managing fieldwork;
ii. How and when each piece of information/data would be collected;
iii. How and by whom each piece of information/data will be used.

It is important to collect the amount of information that can reasonably be collected with good quality and used effectively, otherwise every additional topic with low priority will affect the cost of collecting reliable information.

3.185. The advent of real-time data in censuses refers to the technological advancements and practices that enable the collection, processing, and dissemination of information in real-time or near real-time. This concept has gained significant importance of late due to its potential to provide immediate insights, enhance decision-making, and improve overall efficiency.

3.186. The following information can be collected through this system:

i. Information about particular activities that are implemented before enumeration, such as establishment of local census commissions and training of census field staff;
ii. Information about the field staff needed for administrative tasks, such as recruitment and hiring field staff, bank account information for payment, work accomplished, dynamic fieldworker payment advices especially when utilising electronic data collection methods;
iii. Progress of enumeration of population and housing units to evaluate if the field operation proceeds according to schedule;

iv. Information about logistics issues, such as shipment of census materials and questionnaires, timing of receiving and sending materials, and number and types of materials.

3.187. Census operations can be made more efficient through the availability of a management information system and use of this system by field staff for administrative tasks and supervision. It is possible to create a quick communication mechanism for key messages and work allocation. This system should be used for producing and submitting regular reports providing information about the progress of field activities and enumeration. There are several ways of collecting such information. Technology-based solutions include the use of SMS, USSD, websites and portals, and mobile or handheld applications.

3.188. Paradata which is automatic data collected about the survey data collection process captured during computer assisted data collection can become a vital source of information guide management information systems. Examples of these include interviewer observations, keystrokes, data time stamps like between responses and various other data captured during the process. This information is increasingly being used in real time for both monitoring purposes and manage the process of collecting large amounts of data

b) Supervising the enumeration

3.189. A supervision system to monitor the progress of the field operation is important to allow for correction of errors and make necessary adjustments in the course of the fieldwork. In countries where the Internet or handheld devices are used in data collection, a computerized online system can be developed, and some automated procedures introduced for the supervision.

3.190. The key to rapid quality control of enumeration is the fast flow of information from supervisors to the local statistical committees and to the central statistical committee. The most efficient way of exchanging this information is via the Internet. If local and regional supervisors have Internet access, information can even be submitted through a password-protected database interface (a web-based application).

3.191. Close monitoring during the enumeration phase is essential to ensure coverage, quality and compliance with deadlines. It must be ensured that all staff involved in the data collection have access to up-to-date reports with relevant information. These reports should be made available periodically in printed or digital form. Data from previous censuses or other sources can be utilized to improve monitoring and form a database for management indicators.

3.192. As the enumeration is one of the core census processes, each task performed during the enumeration stage must be carefully planned, executed and supervised to achieve qualitative and quantitative targets. For successful monitoring of field enumeration, actual performance should be evaluated against the set targets. The following are recommended to achieve the desired outcomes.

i. Using historical data

3.193. Data from previous censuses and other relevant data sources such as household surveys and administrative registers should be used as a benchmark to determine the data needed for monitoring the performance of enumerators. This information can be used for monitoring actual performance in

the enumeration against the set goals and targets. Examples of historical information include total population of previous censuses; population growth rate; sex ratio; urbanization rate; proportion of vacant dwellings; and occasional use dwellings in relation to those occupied.

ii. Setting goals and targets

3.194. As a population census is a time-bound project, extension can be considered as failure. Setting goals and targets will be very important to measure if the series of activities is under control or not. Goals and targets for measuring the quality of enumeration and for systematic monitoring of enumeration can be set based on experience of previous censuses and other relevant data sources. The following indicators can be used for monitoring enumeration: (a) proportion of occupied and vacant dwellings; (b) average number of residents per dwelling; (c) response rate and refusal; (d) population size; and (e) population growth rate. Significant deviation between the target values and enumerated values may indicate a problem in the collection process. Estimation of housing units and population – if available – based on census maps and the listing of living quarters and households can also be used as information for monitoring the enumeration.

iii. Preparing policies and procedures

3.195. Policies and procedures to be used as the baseline for monitoring during enumeration should be defined at an early stage of the census, with endorsement from the highest levels of decision-making, for proper management support. Therefore, it is important that there is a stage of evaluation of the previous operation in order to identify gaps and improve control procedures and execution of work.

F. Security during data collection

3.196. Census is a monumental undertaking which collects sensitive information from the population. To maintain public trust, it is important to ensure the data gathered in a census is not only accurate but also secure, protected from unauthorized access and misuse. Census data should be safeguarded throughout the entire lifecycle of the census.

3.197. In the **pre-enumeration phase**, plans need to be drawn up for assuring information security prior to the census execution. These safeguards should aim to protect the privacy of individuals and prevent unauthorized access or disclosure of their personal information. Information security measures include robust data encryption methods to secure data both during storage and transmission, strict access controls and authentication protocols to limit access only to authorized personnel, regular security audits and assessments to identify vulnerabilities, and comprehensive training programs to educate census staff such as data ethics on the significance of data privacy and security practices. In parallel, it is essential to establish strong physical security measures, such as limiting access to census facilities and using secure data storage systems, are put in place to safeguard the physical infrastructure and prevent unauthorized tampering or theft of data.

3.198. During **the enumeration phase,** various information security safeguards should be implemented to ensure the protection and confidentiality of the gathered information. First and foremost, robust encryption protocols should be used to secure data transmission between the data collection devices and the central database. Second, data accessibility should be limited to be fit for purpose for census officials, especially in the case where questionnaires are pre-populated with admin data. Strict

confidentiality agreements should be upheld by all personnel involved in data collection, with severe penalties for any unauthorized disclosure.

3.199. Census data collection with electronic questionnaires (whether with handheld/mobile devices or via the internet) should be subjected to a thorough security review to confirm that the overall design of the system is robust to shocks (connectivity disruptions, power failures, etc) and security vulnerabilities (unauthorized access to data). The review process must concentrate on as many risks as it is possible to formulate in advance and include the common risk points: data at rest on a mobile device, data in transit to the intermediate or central location (via networks or via physical media, or both depending on the system), data at store in the central location, data in processing/ collaborating systems.

3.200. In the **post-enumeration phase**, secure physical and/or virtual environments should be used for census data processing, analysis and storage. In this environment strict measures such as encryption, access controls, and firewalls should be put into place to prevent unauthorized access to the data. Additionally, regular security audits and assessments should be conducted to identify and address any potential vulnerabilities. Strict data handling and confidentiality protocols should be followed by trained and authorized personnel, who are required to adhere to ethical guidelines and sign confidentiality agreements. Data anonymization/pseudonymization techniques may also be utilized to further protect the confidentiality of census data.

3.201. The variables for data analysis purposes should be provided on the need-to-know basis, for example, if a report involves construction of the population pyramid, then only the age and gender of enumerated persons should be supplied for such analysis, and variables like location data or contact information should not. Data requirements for each report, table, chart, or analytical product may and should be documented well in advance of the census and accompany the tabulation plan.

G. Use of technology

3.202. Technology, namely the use of computer assisted methods for census planning, organization, data collection and data processing, has become indispensable for improving the cost, quality and efficiency of the population and housing censuses. Rapidly expanding worldwide mobile and internet connectivity coupled with advancements in areas like cloud computing, smart mobile devices, GPS, GIS, natural language processing and AI provide new opportunities for increasing the quality and speed of census data collection. Furthermore, the disruptive effects of the COVID-19 pandemic precipitated more widespread adoption of online tools and methods, as social distancing measures necessitated a shift towards remote solutions for data collection. The technological tools and instruments described below are well documented in national practices in conducting the population and housing censuses in the most recent census rounds.

a) Electronic questionnaire

3.203. Combined with or completely replacing the paper questionnaire, an electronic questionnaire can be used in either the face-to-face or self-enumeration method. Electronic forms can provide improved data quality and operational efficiencies by implementing validation rules on individual questions, cross-validation between questions or with other records, automatic sequencing (leading the operator to the next appropriate question), more options in pull-down lists, capturing more detailed data, providing computer-assisted coding and the ability to ask tailored supplementary questions. Electronic questionnaires can give access to guidelines, illustrations, help material and even videos to provide

instruction to the interviewer or household. Electronic questionnaires can also be prefilled using administrative data and geocodes, which enumerators can validate against GPS information collected from the field and update, if necessary.

3.204. The use of electronic questionnaires provides a great potential towards audit and quality control: in particular, a confirmation of the visit and contact with the household. This can be done with a combination of capturing an image, recording the interview conversation, capturing the geo-location during the interviewing process, and a range of operational information that can be used to monitor operations and analyse responses (often referred to as paradata), including the time taken to complete the form, the date and time the form was completed, and the device used to complete the form.

3.205. Electronic questionnaires can also provide census results more quickly by transferring data to a central database immediately or soon after the enumeration, either using real-time connectivity, periodic/on-demand synchronization, or by transferring using physical media to a local centre. Electronic questionnaires reduce the amount of material (such as paper questionnaires) to be printed, distributed and returned, and reduce data scanning and capture costs and errors. Yet the use of electronic questionnaires may introduce new expenses such as cost of the electronic devices used for data entry in the field, as well as data transmission costs.

i. Electronic questionnaire: interviewer-administered

3.206. Census interviews can be undertaken using an electronic questionnaire on smart phones, tablets, laptops or other portable devices using a computer-assisted personal interviewing (CAPI) methodology. Electronic questionnaires may also be employed in computer-assisted telephone interviewing (CATI) in which the interviewer records the respondent's answers to questions asked over the telephone. Each enumerator can be assigned one or several enumeration areas so that the records are tagged with the respective enumeration area and enumerator/supervisor name. The device may also be able to capture information on the location of the interview, time of day and other metrics that may be useful for quality control. A CAPI system integrated with GIS and operational management applications which captures geotagged data together with operational information (interview date and interview start and end times) can improve the monitoring of census data collection operations, and the coordination of field operations, logistics, and communications.

3.207. For areas with lack of internet connectivity systems with local direct data exchange between the interviewers and supervisors should be considered, to facilitate review of the completed interviews by the teams' supervisors even before such interviews are sent to the central database. Such systems reduce the costs (in terms of the lost time and travel expenses) to the locations where the internet coverage is available.

ii. Electronic questionnaire: self-enumeration method

3.208. Computer-assisted self-interviewing (CASI), whereby respondents access and self-administer the electronic questionnaire by means of an internet browser application, can substantially reduce operating costs when a good percentage of enumeration is achieved using this method. Achieving a good percentage of enumeration using this method can reduce the operation costs substantially. Electronic questionnaires for households should be implemented in a publicly-accessible Internet portal. Households are usually provided with a unique identifier (often delivered by mail, email, by enumerator

or by self-registration system) that is used to initiate their questionnaire or resume a partially complete questionnaire via the Internet. Householders may prefer to respond using an electronic rather than a paper questionnaire for its convenience. If the census is collected in a multimodal approach, for example offering both electronic and paper questionnaire options, a system will be required to track the status of each dwelling (questionnaire) throughout collection to ensure completeness of coverage and ensure non-response follow-up is not conducted with responding dwellings.

3.209. The design of a system implementing self-enumeration should demonstrate effective dealing with the following common problems: duplication of submission of information (e.g. by different members of the same household); fake or bogus submissions/digital vandalism (e.g. submission of data for a non-existing address or in the name of another household by hackers), attempts to obtain confidential data by gaining access to the unique identifier used to access the form.

3.210. Total self-enumeration remains a challenge: access to the internet is not universal, and there are groups of persons in the population who are not frequent users (in particular the elderly) and may have technological barriers, foreigners may have language barriers, and persons with disabilities may experience difficulties reading the questions or supplying the answers to them. Thus, this method appears as a complement, but not entirely replacing the in-person enumeration.

b) Handheld or mobile devices

3.211. Mobile devices such as smartphones or tablets can greatly improve the efficiency of census operations by facilitating canvassing, communication and supervision of fieldwork, data entry and validation, and data transmission. GPS-enabled handheld devices can assist enumerators with navigation, identify buildings for each enumeration area using GPS coordinates, and capture interview location information.

3.212. When deciding to use mobile devices in the census process, it needs to be determined whether the census agency will purchase and provide the device, or whether the field officers will use their own devices. As the availability and proliferation of devices increases there can be financial benefits, as well as reduced training needs, if field officers can utilize their current device rather than be provisioned with a new device. This, however, presents a range of technical, support, security and legal considerations, and for this reason this deployment option should entail detailed regulations on the compatibility/ platform of the device, minimal requirements (in terms of the storage, memory, performance, communication capabilities, duration of autonomous work, etc). Another significant consideration is the fact that mobile devices operate on different platforms and developing applications for different platforms increases the costs and efforts, unless such cross-platform portability is already provided by the electronic system chosen for the data collection and census operations management. If the devices are to be used for data collection, consideration should be given to such factors as screen size, battery capacity, operating system, etc.

c) Geographic information system (GIS) and geospatial information

3.213. Geographic information systems (GIS) may be used for several major uses during a census:

i. to place the enumerator in the context of the environment, by depicting his location, surrounding dwellings (as points or where available as footprint polygons), and their status, whether already enumerated, in progress, or to be enumerated;

ii. routing capabilities to navigate to the household for enumerators (in return visits) and supervisors (in audit visits);

iii. to provide controls for the supervisory staff, especially with respect to the coverage control of the populated areas;

iv. to produce various dashboards and GIS-based analysis of the census progression for the survey coordinators.

3.214. Such uses rely on the location-enabled mobile devices (GPS or analogous positioning system) and availability of digital maps (sourced from the Internet in the areas with connectivity or stored on the device in areas where the connectivity is absent or sporadic).

3.215. The choice of technology for production of EA-level maps – which is dependent on the choice of data capture software - can greatly affect the time it takes to produce maps. These must be taken into account during the planning stage in order to establish an accurate timeline for the census operation. An extensive elaboration on the use of both GIS and GPS is presented in Chapter IV above.

d) Dashboards

3.216. Dashboards present aggregate information about the progression of the census operation overall in a concise, visually compelling and efficient manner. The key indicators to be tracked with the dashboards are to be chosen and defined at the planning stage. A dashboard is not a substitute for detailed reports, but is a supplementary tool, which provides dynamic, possibly real-time updated information.

3.217. A layered approach to construction of the dashboard presents a possibility to visualize different indicators to various members of the team. While the coordinators may be more interested in the overall progression of the census data collection, the supervisors may be getting the indicators stemming from the data collected by their specific team.

3.218. Access to the dashboard information should be well-protected and limited to the team members that need such information for their operational duties. However, a public version of the dashboard can be also conceived, but caution should be taken not to create confusion by presenting intermediate information as final.

e) Contact centre

3.219. The contact center, also known as a call center, is a crucial component of the census process. It can be utilized at every stage of the census to support field operations and ensure that the census is conducted efficiently and accurately. A contact center can employ various communication channels, including phone calls, chats, and web-based FAQs, to provide assistance to respondents. Interactive voice response (IVR) technology can be used to direct calls to specific agents based on the options selected by the caller. It can also be used to provide common answers to frequently asked questions, thereby reducing the workload of agents. Website call-back and chat features can be implemented to assist respondents while they fill out the online e-questionnaire in the portal. Contact centers can perform a wide range of functions, including providing help and support to respondents and collecting data. They can offer support to both the public and census personnel, ensuring that everyone involved in the census process has the resources they need to complete their tasks effectively.

f) Situation room

3.220. A census situation room serves as a centralized operational hub for coordinating and managing the multifaceted operations of a population census. Its primary purpose is to ensure the smooth execution of census activities, swiftly addressing challenges and optimizing resources in real-time. Leveraging geospatial technology, the situation room can provide real-time visualizations of census operations across regions, enabling decision-makers to monitor the progress of data collection, predict potential bottlenecks and allocate resources as needed. In the event of unforeseen incidents or challenges, such as natural disasters, civil unrest, or other disruptions, geospatial data can help the situation room quickly assess the impact on census operations and devise strategies to mitigate the challenges. A geospatially-informed situation room can cross-check enumerator locations with predefined areas, predict potential bottlenecks, ensure enumerator safety, and overlay census data with other geospatial datasets for richer context.

g) Short messaging service (SMS)

3.221. Short messaging service (SMS) is a versatile tool that can be used at various stages of the census project to share information with field personnel and respondents. It can be used to send passwords, guidelines, alerts, marketing messages, reminders, and more. If the SMS gateway is integrated with the central database of the census, alerts can be sent to census management upon various critical business events and when violations occur or attention is required, such as when the monitoring system detects that “coverage is lower than expected”. In addition to SMS, other mediums of communication such as USSD and media platforms can also be used to facilitate communication between field personnel and respondents. These platforms can be used to share information, receive feedback, and provide support to respondents and field personnel alike.

VIII. Data processing

3.222. No matter how thorough and accurate the census enumeration is, the usefulness, quality and timeliness of the census tabulations will suffer unless the collected data are properly processed. An important element of a successful processing operation is the close and continuing collaboration, at all levels, between the data-processing staff, and the subject matter and general statistical staff. At a minimum, the subject matter and general statistical staff will need to become familiar with and take a continuing interest in the processing plans and operations, while the processing staff will need to become familiar with and take a continuing interest in the substantive aspects of the census.

3.223. Plans for data processing should be formulated as an integral part of the overall plan of the census, and those responsible for the processing of the census should be involved from the inception of the planning process. Data processing will be required in connection with the results of census tests, compilation of preliminary results, preparation of tabulations, evaluation of census results, analysis of census data, arrangements for storage in and retrieval from a database, identification and correction of errors, and so on. In addition, data-processing technologies are playing an increasing role in the planning and control of field operations and other aspects of census administration. Data processing has an impact on almost all aspects of the census operation ranging from the selection of topics and the design of the questionnaire to the analysis of the final results. Therefore, data-processing requirements in

terms of personnel skills and knowledge, space, equipment and software (computer programs) need to be looked at from the point of view of the census as a whole and at an early stage in the planning.

3.224. When processing census data collected through various methods and modern practices, several crucial considerations come into play. Firstly, integrating data from diverse collection modes into a cohesive dataset is essential, but it can be challenging due to potential differences in formats and structures. Therefore, establishing a robust integration process is vital for accuracy and completeness. Additionally, assessing the quality of data collected through different modes is imperative. For instance, self-administered surveys may be less reliable, necessitating thorough error identification and correction. Furthermore, ensuring the security and confidentiality of the data, especially when collected electronically, is paramount. Implementing encryption and other security measures safeguards against unauthorized access, while anonymization techniques protect individual identities.

3.225. To address these considerations, employing a versatile data management system is advised, one capable of seamlessly integrating and handling data from various modes. Establishing specific data quality checks and validation rules aids in error identification and rectification. Encryption and other security measures are crucial for safeguarding data integrity, and anonymization techniques help preserve confidentiality.

3.226. To successfully integrate new computer hardware or software into the census, it's crucial to expand and enhance the skills of the existing data-processing staff. This is because unfamiliar technology often demands new proficiencies for operation and maintenance. Given the intricate nature of the census process, it is crucial for the data-processing team to be well-versed in the specific requirements of the census for efficient data processing. Early training is essential to integrate the staff into census planning and operations promptly, enabling them to acquaint themselves with the new technology and contribute valuable insights to the process. This early preparation ensures the team is poised to process the collected data promptly. The training program should encompass a comprehensive range of skills necessary for the staff to execute their roles effectively, including proficiency in operating the new hardware and software, processing census data, conducting quality checks, generating census reports, troubleshooting hardware and software issues, and maintaining the technology infrastructure. To cater to diverse learning styles, the training should be delivered through various formats, such as classroom instruction, hands-on practice, and online learning modules. Additionally, it is crucial to evaluate the effectiveness of the training program through methods like surveys, interviews, and observations, using the results for refining future training endeavors..

3.227. Early training of staff for census software poses the risk of potential departures before the census commences due to various reasons like job opportunities, retirement, relocation, or personal commitments. This exodus could adversely affect census quality and efficiency as remaining staff may lack adequate training and require additional resources. To mitigate this risk, it's advised to train only committed staff, offer attractive incentives, provide avenues for professional growth, foster a positive work environment, and have a contingency plan for replacement training. Factors to consider include the cost of training, availability of qualified trainers, and the importance of maintaining consistency in the census process. Ultimately, the decision to conduct early training should be tailored to the specific circumstances of each case.

3.228. Decisions will need to be made concerning the location of the various data-processing activities within the country, including the extent to which the processing work is to be decentralized. This decision should be partly based on the ability to recruit the required personnel for the processing

operations. Acquisition of both equipment and supplies can require long lead times; estimates of both data capture and computer processing workloads must be made early to enable timely procurement. Closely related to the question of equipment is that of the provision of adequate space. Although the maintenance of most personal computer equipment no longer requires adherence to rigid standards in terms of temperature, humidity, dust and so on, attention to issues related to power supplies is still important. Inevitably, more important is the attention to be devoted to the maintenance of servers (especially heavy-duty servers), where most of the information processing is likely to take place and saved, as well as the data transmission infrastructure.

3.229. Ensuring smooth and uninterrupted internet and web communications is vital for the success of census operations, as timely and accurate data transmission between different units is crucial. To achieve this, it's advised to employ diverse internet connections from different providers, implement load balancing for improved performance, utilize content delivery networks to ease server load, and maintain vigilant network monitoring to promptly address any issues. In addition, for traditional archiving, creating a secure and controlled space to store completed census forms is imperative due to the sensitive personal data they contain. This involves employing physical security measures, logical security measures such as encryption and access controls, environmental controls to protect against damage, and implementing audit trails for accountability.

3.230. When preparing for census data editing and tabulation, decisions regarding the software to be employed are pivotal. This choice hinges on several key factors, encompassing the census's scale and complexity, the proficiency of the data-processing team, budgetary constraints, and the availability of technical support. Various software options are available, falling into three primary categories: census-specific software tailored explicitly for census processing, like CSPro; general-purpose software applicable for various tasks including census processing, such as spreadsheets, databases, and statistical analysis software; and custom software developed specifically to align with the unique needs of a particular census organization. When opting for editing and tabulation software, it is imperative to weigh factors like user-friendliness for the data-processing team, requisite functionality encompassing data entry, editing, imputation, tabulation, and reporting, scalability to manage extensive datasets, robust security features to safeguard data confidentiality, affordability, and the availability of reliable technical support. Subsequently, the selected software will necessitate customization to align with the precise requirements of the census organization, potentially involving code modification, report creation, or procedure development. Allocating ample time for customization is crucial. Additionally, the data-processing team should undergo comprehensive training on the software, covering all essential features. This training should undergo evaluation to ensure its effectiveness.

3.231. Outsourcing some of the predominantly IT-related operations may be considered. Outsourcing should be implemented in such a way as to bring immediate economic and quality advantages to census operations. However, it's crucial to approach outsourcing in a manner that upholds data confidentiality. To safeguard this, national statistical offices should rigorously select reputable contractors known for their commitment to data confidentiality, and contracts should encompass provisions mandating confidentiality protection. Continuous monitoring of contractors' adherence to these provisions is imperative. Employing encryption and other security measures during data transmission, and limiting access to the data, both internally and externally, are essential steps in preserving data confidentiality. Lastly, it's paramount that the national statistical offices retain control over the data, ensuring that outsourcing operations do not entail a relinquishment of data control.

A. Method of processing

3.232. The choice of an appropriate method of processing is determined by the circumstances of each country. Rapid advances in data-processing technology have greatly increased the speed and reliability of producing detailed tabulation, thereby making computer processing the standard method of processing around the world. Furthermore, an alternative to mainframes, whose computational power was necessary before the advent of lighter and more scalable IT hardware solutions and more recently cloud based computing, is the use of a client-server environment.

3.233. Cloud-based computing, often referred to as cloud computing, is a technology model that allows individuals and organizations to access and use computing resources over the internet, as opposed to owning and managing physical hardware and software on their local computers or data centers. In cloud computing, these computing resources are hosted and managed by third-party service providers in data centers located around the world. Cloud computing offers several benefits for conducting a census, making the process more efficient, cost-effective, and scalable. Cloud computing offers census agencies the ability to efficiently and securely collect, process, and analyze data while providing flexibility, scalability, and cost savings compared to traditional on-premises infrastructure. These advantages can lead to a more accurate and timely census, which is crucial for policy-making and resource allocation.

3.234. Several lighter tasks, including editing and tabulation of data files, can very well be done on small-sized desktop systems that can be placed in substantive departments and in field offices. On the server side, most of the heavier computing operations, such as scanning, aggregation and analysis of large sets of microdata, coordination of data transmission, Intranet web hosting and so forth, can be executed more reliably than on microcomputers. However, a client-server environment to handle census data must operate over a robust and secure local area network (LAN) or wide area network (WAN) into the cloud environment. Therefore, computer work is not necessarily dependent on a centralized data-processing facility, provided that a robust LAN or WAN interconnects workstations dispersed over various offices, buildings and different parts of the country.

3.235. In a census office that utilizes a networked computer environment, the central file or database servers allow both data and programme files to be stored in a central location. This system economizes specifications of client computers and removes the need for much physical movement of programs and data on other computer media. Data storage requires frequent backups of the system information to avoid major data loss due to hardware or software faults. Thus, servers have a strategic importance, and their location and administration must be well defined and secure enough to ensure data protection. Also, it is recommended that proper business continuity and security policies duly certified by the competent authorities should be in place.

3.236. In determining the type of equipment to be employed and the advisability of a new machine installation (either complete or partial), or of additions or upgrades to existing equipment, consideration should be given to all the processing requirements of the data collection programme for which the population and housing census is but one part. Only on this basis can a reasonable decision be made. Decisions on the type of data-recording equipment and computer equipment should be made at least one year in advance of the scheduled date of enumeration in order to allow appropriate questionnaire design and proper preparation of instructions to enumerators, development of coding schemes, specification of data-handling controls and procedures, and recruitment and training of data processing personnel. Rapid processing of a pretest or pilot census that covers end-to-end census operation, including enumeration, initial census result, output dissemination, and handing over and closure

procedures, is particularly important for identifying improvements needed in the census questionnaire, instructions to enumerators, computer systems or whatever other preparations may be needed. It is recommended, therefore, that arrangements for using appropriate equipment and software be made well in advance of such tests. It is also recommended that all systems used to support census operation be thoroughly tested in advance of operations to ensure that they function as intended and that they are secure (that is, they will not lead to loss of data).

B. Preparation for data capture

3.237. In the case of paper questionnaires, the most common procedure is to have the census documents arrive in the processing centre in batches by enumeration area. Maintenance of these batches throughout the data processing is recommended, since documents for a given enumeration area reflect the work of one enumerator and may contain a series of errors typical of that person. To ensure the integrity of the batches, the census documents should be stored in a specially designed census document storage facility. The batch for each enumeration area should first be checked for completeness, geographic identification codes and other characteristics of acceptability before being sent to the next stage of data processing. Transcribing all coded data onto another sheet (for example, the coding form) should be avoided since it may add transcription errors. The same considerations apply to the case of electronic transmission of questionnaires or when the first phase of data processing consists of the scanning and text or image recognition of census questionnaires. In the case of questionnaires transmitted electronically (self-enumeration on the Internet or using e-forms), it is appropriate to set up a metadata model where the enumeration area can be recorded. As far as storage is concerned, if paper questionnaires are scanned, secure media storage for their backup, not only the originals in paper, should be planned for.

3.238. If the census is conducted in a multimodal approach, for example using self-response by Internet and field follow-up of non-respondents, it will not be possible to batch questionnaires by enumeration area for processing. A master control system will be required to track the status of each dwelling (questionnaire) throughout collection and processing operations and ensure completeness of coverage. This is closely aligned to the master control system that is being utilised to manage dashboards and fieldwork collection progress.

C. Data capture

3.239. Converting the information obtained in the census using paper questionnaires to a format that can be interpreted by a computer is called data capture. It is possible that several simultaneous and different methods for data capture are being used in a census. There are two types of data capture, that is, data capture during enumeration and after enumeration.

Data capture during enumeration

3.240. Electronic data collection (including collection by Internet or using handheld or laptop computers) has been widely used by countries. Electronic data collection means the integration of interviewing and the data entry process including data capture, coding and consistency checks. Electronic data collection with handheld devices, Internet or telephone allows the capture of information with relevant codes (there might be some exceptional variables that may require a coding in the office, such as occupation and industry). It also allows the identification of potential errors during the interview with pre-programmed consistency checks. Because the consistency checks are performed

during the interview in real-time, errors and inconsistencies can be resolved, and corrective action can be taken by the respondent or the enumerator. However, introducing editing rules into the data collection application has to be carefully examined so as not to affect its performance significantly in the field, not to create a bias on data and not to affect the quality of the interview for questions that may not be answered properly, especially in case of collecting information from proxy respondents.

Data capture after enumeration

3.241. The traditional way of processing census data is after the enumeration phase. They include keyboard data entry, optical mark reading, optical character reading or image-processing techniques, such as intelligent character recognition. Computer-assisted keyboard data entry is usually carried out using personal computer data entry programs with built-in logic controls. Some of the tasks accomplished by the programs are (a) verifying that enumeration area codes are valid, and copying them automatically from one record to the next; (b) assigning a number to each person in a household automatically (and perhaps to each household within an enumeration area); (c) switching record types automatically if the program's logic requires it; (d) checking that variable values are always within predetermined ranges; (e) skipping fields if the logic indicates doing so; (f) supporting keyboard verification of the information entered earlier; and (g) generating summary statistics for the operator and the batch. In order not to delay the data capture task, data entry applications should limit checking to problems that are either very serious (for example, wrong enumeration area code), or likely to be caused by a simple misread or key entry mistake. More sophisticated checking is deferred until the editing stage.

3.242. Optical mark reading (often called optical mark recognition) equipment has been available for many years and has nowadays reached good levels of reliability. Optical mark reading is the simplest of the commonly available form data capture technologies. Owing to relatively stringent requirements for the successful data capture of the paper, countries with very dusty or humid climates and poor transport infrastructures are discouraged from using optical mark reading. It is necessary to heed special questionnaire design restrictions and consider the quality of the paper, and adhere to precise specifications regarding the printing and cutting of the sheets. In some developing countries, this may mean that local production of the questionnaires will be problematic. The need to reserve a relatively large space for marking areas and to adhere to other limitations imposed by optical mark reading equipment sometimes make it difficult to design the best questionnaire from the point of view of the enumeration process.

3.243. Optical mark reading questionnaires can be marked by the respondent or by the enumerator. Marking by respondents is attractive from a cost perspective, but it depends on the presence of a cooperative spirit and relatively universal literacy. A practical problem is that most optical mark reading devices put restrictions on the writing instrument and the colours that can be used in the marking. Assuming the rules are followed, the rejection rate for marked forms is often low, especially if the forms have been inspected visually before being fed into the readers. Converting a manually completed census questionnaire to optical mark reading format after it has been received in the census office is inefficient and becomes a source of errors and should therefore be avoided.

3.244. Optical character reading (also called optical character recognition) and intelligent character recognition consist of the use of special equipment to read characters at specific locations in the questionnaire. These two terms identify very similar technological approaches. Specialized sources tend to identify with optical character reading the capability of recognizing printed characters only, whereas

intelligent character recognition would extend this capability to handwritten text. There is no agreed definition of intelligent character recognition. In the context of censuses, therefore, this would require that handwritten text in the filled-in questionnaire be as standard as possible so as to enable efficient recognition. In general, recognition of numerals is more efficient in an uncontrolled environment, that is to say, where the machine has not been adapted to the writing style of a particular person. Optical character reading and intelligent character recognition technology has matured considerable with sophisticated recognition algorithms and the use of neural networks for self-learning.

3.245. Imaging techniques and scanner devices, together with optical character reading and intelligent character recognition software, have been used by several countries for data capture. Experience shows that significantly low error rates are achieved at an optimum cost using these techniques. The efficiency is greater in the case of numerical and alphanumerical characters written by trained enumerators. However, alphanumerical characters are prone to higher error rates. Extensive testing must be conducted well in advance to determine the best type of equipment and paper. The use of imaging techniques is also dependent on the availability of local maintenance and support capabilities. Whatever methods of coding and data capture are chosen, it is essential that they be carefully tested before final adoption. Recognition engines can be customized to recognize various sets of characters and scripts, but unless good experience is available at the census office, careful planning and preliminary work are needed in conjunction with the optical character reading or intelligent character recognition system providers. A combination of intelligent character recognition (for numerical characters) with computer-assisted coding (for alpha characters) is also an effective method used by some countries.

3.246. In addition to the benefits of the scanning technology for capturing the information, an important by-product of scanning census questionnaires is that this allows for the possibility of digitally filing and naming the scanned questionnaires. This increases the efficiency of storage and retrieval of the questionnaires for future use, particularly during subsequent data-editing operations.

3.247. The quantity and type of data entry equipment required will depend on the method of data capture selected, the time available, the size of the country, the degree of decentralization of the data capture operations, and a number of other factors, such as the use of digital enumeration approaches. For keyboard data entry, the average input rates usually vary between 5,000 and 10,000 keystrokes per hour. Among the factors that affect operator speed are (a) the supporting software and program with easily navigable screens, spell checker on the description fields if any, keyboard shortcuts throughout the program, less utilization of the computer mouse, and so forth; (b) the complexity of the operators' tasks; (c) the ergonomic characteristics, reliability and speed of the equipment; (d) the question whether work is always available; (e) the training and aptitude of the recruited staff; and (f) the motivation of the workers.

3.248. Several options are available to help ensure that data entry operations are completed in a timely manner. They include (a) procuring more equipment; (b) increasing the number of working hours by working double or even triple shifts and during weekends; and (c) applying independent verification to varying extents. In the case of keyboard data entry, with the increasing safeguard of data quality by data entry programs, complete verification has become less necessary. Full independent verification may be applied only in the initial stage of data entry and may be reduced when each worker has achieved an acceptable level of quality. After that, a sample verification plan can be applied. Operators may be assigned to sample verification depending on their observed error rate. The work of reliable operators may be verified only for a small sample of the enumeration areas, while more extensive verification is continued for the more error-prone operators.

D. Coding

3.249. Whenever possible, precoded responses should be used in census questionnaires with numerical or alphanumerical codes. Since computer editing and tabulation of textual material are not practical, verbal responses will have to be replaced by a code. The incorporation of artificial intelligence (AI) and machine learning (ML) technologies would be ideal. These advanced technologies offer the potential to automate critical tasks, including the classification and coding of open-ended responses. Such integration stands to enhance the overall efficiency and accuracy of the data collection process. While the use of pre-coded responses in census questionnaires is encouraged, there will inevitably be instances where respondents provide verbal answers necessitating coding. In such cases, dedicated computer programs or skilled coders can be employed to translate these verbal responses into codes. Although automatic coding is the preferred approach, certain situations may call for manual intervention. While coding directly during interviews can be advantageous, it is often logistically challenging as enumerators may not always possess the requisite training or resources for accurate coding. Therefore, in most cases, responses will be gathered in text format and subsequently coded by proficient coding experts. Given the scale of coding operations in a census, optimizing automatic coding processes is paramount to minimize the need for human intervention (see paragraph 3.186). Techniques like natural language processing (NLP) and deep learning can be employed to facilitate this optimization. Additionally, AI and ML technologies can be harnessed to effectively classify occupations and industries based on text descriptions, a valuable tool for coding open-ended responses pertaining to occupation and industry.

3.250. Automatic or computer-assisted coding will efficiently support the coding activity, reducing coding errors and speeding up the coding process. When required, a coder normally works with one or several codebooks for various items in the questionnaires. This technology expedites coding tasks, liberating coders to concentrate on more intricate responsibilities, such as coding open-ended responses or addressing coding discrepancies. Additionally, automatic or computer-assisted coding provides invaluable feedback to coders, enhancing the overall accuracy of the coding process by aiding in error identification and correction. By automating laborious coding tasks, this technology alleviates coder fatigue, thereby elevating the overall quality of coding work. Lastly, it ensures the consistency of the coding process by enforcing uniform adherence to coding guidelines among all coders, leading to enhanced data quality.

3.251. Computer-assisted coding uses personal computers to assist the coders. The process requires that all the codes be stored in a database file and be accessed by coders during the coding operation. Computer-assisted coding is based on at least two general approaches. In the first one, coded answers are matched to a set of keywords. Textual information from the census questionnaire is parsed and compared to an indexed list of keywords, and then the likelihood of matching between found keywords and coded answers is measured and scored. If the score results are over a certain (high) threshold and there is no ambiguity, a sorted list of coded answers is presented to the coder, who retains the ultimate decision of accepting or refusing the system's proposed answers. In using this method, it may be advantageous to change the order of activities so that the capture of precoded information in the questionnaire occurs first, followed by the capture and computer-assisted coding of the remaining information.

3.252. In the second approach, which is mainly used in image processing of data (intelligent character recognition method) for non-Latin or multilingual countries, owing to the difficulty and existing

problems in character (alphanumerical string) recognition, the procedure is as follows. After the scanning and during the coding operation phase, the image of the text will be shown on the monitor, and at the same time, a pull-down menu from a coding database will present the coder with the ability to enter as few key entries as possible to get to the full textual and coding content of a specific case. When the coder selects a code, it will be allocated and saved in the database for that specific case. Although this approach is more time consuming and costly in comparison to the first approach, the quality of coding is much higher than in the traditional way of coding.

3.253. On the other hand, both techniques have several similar advantages: (a) capturing the precoded information at an early stage leads to some data files becoming rapidly available, which opens up the possibility of generating and releasing preliminary census results; (b) the computer-assisted coding process provides an opportunity for a computer system to alert the operator to problems with data supposedly already captured, for example, missing information for a fully precoded variable; (c) the coder works directly on the computer screen; and (d) information from other variables may be helpful in determining applicable codes for write-ins.

3.254. Automatic coding is a process in which the decision about the code to be assigned is delegated to a computer program. The main difference from computer-assisted coding consists in the automatic acceptance of the answer if its score is over a predetermined threshold and relatively higher than possible identified alternatives. Both computer-assisted and automatic coding systems may exploit self-learning capabilities of neural networks to fine-tune their capacity of detection. A human operator becomes involved only in those cases where the software cannot resolve the issue. Computer coding may use, in addition to the written response for the item in question, other relevant information available in the record or the questionnaire. Therefore, automatic coding is more applicable in cases where the data-capturing process has already been completed, by Internet, handheld devices or other forms of electronic data collection, manually or by some form of automatic reading. Developing computer software for automatic coding is a complex task. Automatic coding methods need to be complemented by computer-assisted or conventional coding methods for unresolved responses.

E. Data editing[58]

3.255. Administrative data and big data serve as invaluable tools for validating self-reported information, enabling the identification of potential errors or inconsistencies within the dataset. Administrative data lends itself to the verification of demographic details such as age, sex, and place of birth, while big data sources can corroborate information regarding employment, education, and housing. Furthermore, these data sources play a crucial role in enhancing data completeness by filling in missing values. Administrative data can contribute to the dataset by supplying previously absent information on income, employment status, or health insurance coverage, while big data sources can complement the dataset with details on geographic location, household composition, or online activity. Additionally, both administrative and big data sources offer enrichment opportunities by providing supplementary information about the population, affording a more comprehensive understanding of the demographic landscape, including insights into social services usage, criminal justice involvement, social media behavior, online shopping habits, and travel patterns.

[58] For further details on census data editing, see *Handbook on population and housing census editing*, *Revision 1,* United Nations publication, Sales No. E.09.XVII.11, United Nations, 2010.

3.256. Adopting a multi-mode data collection approach proves to be a more efficient strategy compared to traditional methods. This approach allows for data to be gathered from various sources, ultimately saving both time and financial resources. Administrative data can be directly sourced from government agencies, while big data can be sourced from publicly available platforms or private companies specializing in data aggregation. Beyond efficiency, the multi-mode approach substantially heightens data accuracy by enabling cross-verification across multiple sources. In cases of discrepancies between information from different sources, the national statistical offices can conduct thorough investigations to ascertain the correct details, ultimately leading to a more robust and reliable dataset.

3.257. Raw data files contain errors of many kinds, some generated by the respondents and others caused by enumerators who misunderstood the respondents' answers. Further mistakes are introduced in the data-processing operations and during coding and data entry, or in the course of the transcriptions that take place. From an operational point of view, such errors are of two types: (a) those that have the potential of blocking further processing (critical errors); and (b) those that introduce distortions into census results without interrupting the logical flow of subsequent processing operations (non-critical errors). Throughout the correction process, backup copies of the original data file should be consistently created. Numerous errors can manifest, including misinterpretation by respondents, enumerator misunderstanding or data entry mistakes, typographical errors during data entry, software-related errors, and transcription errors. Timely error identification and correction are imperative to prevent inaccuracies that could profoundly impact policy-making, resource allocation, and crucial decision-making processes. Various error-checking and correction methods, such as range checks, consistency checks, logic checks, and manual review, can be employed to enhance data accuracy, with manual review offering the highest accuracy albeit with increased time investment.

3.258. Since for large censuses manual correction is rarely economically feasible, the conditions for such corrections are usually specified in specially designed computer programs for automatic error scrutiny and imputation based on other information for the person or household or for other persons or households. Whenever imputation is used, a flag should be set so that analysts are able to distinguish between reported information and that imputed by the editing system. In cases where insufficient information hinders error correction, methods like hot deck imputation can be deployed, leveraging data from previously processed entities with similar characteristics. This technique uses information obtained from previously processed persons, families or households with similar characteristics as the "best suited" value in replacing missing values or values that have failed processing edits. However, this technique requires careful programming work, considering that the search for appropriate information in the census database would slow down computer program execution.

3.259. A prudent approach in data processing involves isolating out-of-range or blatantly inconsistent values prior to editing or classification, mitigating the risk of introducing statistical biases. But precautionary measures should also be defined and set for the fact that overambitious automatic editing programs may cause the so-called "corrected" data to be significantly flawed. In this respect, it would make sense to have an acceptable cut-off value for error rates at the enumeration area level. If a data scrutiny program finds that more than a certain percentage of the records in a particular batch have one or more serious problems, the whole batch should be rejected and subjected to human or fieldwork verification.

3.260. Developing editing and imputation rules should be entrusted to subject matter experts rather than computer programmers. This is because specialists possess a profound comprehension of the data and potential sources of errors, enabling them to identify likely error types and establish effective

correction protocols. An error scrutiny and editing plan should be crafted early in the census, delineating specific rules and their implementation procedure. This plan should also incorporate a mechanism for periodic review and adjustment of the rules. These rules must be meticulously documented and shared with programming staff to ensure accurate and consistent implementation, utilizing clear and straightforward language for easy comprehension. While programmers bear the responsibility of coding the rules, subject matter experts should verify the accuracy of the code. Thorough testing by both subject matter and software testing experts is imperative before applying the editing programs to the census dataset, detecting and rectifying any program flaws under diverse conditions to ensure comprehensive functionality. Adhering to these guidelines yields numerous benefits, including heightened data accuracy, reduced error risks, enhanced transparency in the editing process, improved collaboration between subject matter experts and programmers, and elevated quality of census results. However, challenges persist, encompassing the time-consuming nature of rule development and program testing, and the potential complexities in coordinating the efforts of subject matter experts and programmers, along with the need to mediate disagreements between them.

F. Validation

3.261. The outcome of editing is a set of records that are internally consistent and in which person records relate logically to other person records within the same household. This process does not, however, provide the full range of assurance necessary to accept the data set as the best possible. A range of conditions could cause errors that cause the data to be consistently wrong: for example, perhaps a condition in the editing suite itself is set incorrectly; proportions in an imputation program may be set wrongly; or enumerators may complete a collection control panel incorrectly. To identify such consistent errors, it is necessary to critically review some key aggregate tables to isolate outlier aggregates and identify the cause of the unusual values. These key tables may be a subset of those intended for output or may be tables specifically designed for this purpose.

3.262. It is recommended that a bottom-up approach be used in this process. That is, the tables should first be examined for a selection of enumeration areas, then the next level up and so on up to the first set of national tables. There are two reasons for this:

(a) The first enumeration area will complete the processing cycle well before any other geographic level. Thus, commencing at this level gives the earliest possible warning of a problem, enabling corrections to be made before a large amount of reprocessing is required.
(b) It is far simpler to examine a few hundred records within an enumeration area than to attempt to resolve the problem in the millions of records in a national file.

3.263. A crucial stage in the process is designing the analytical tables. One way of approaching this could be to identify a set of variables that are conceptually consistent with those in the previous census (or a major survey) or administrative records from various authorities in the country, such as expatriate visas issued, national ID programme or number of registered establishments. Thus, a set of benchmark values could be constructed before the census operation commences and compared with those from the current enumeration. The content of the benchmark set will depend upon the content of the enumeration and much of this must therefore be determined by each country. However, any census will include the variables age and sex so a comparison of the age pyramid and sex ratio for each 10-year age cohort would be basic elements of such analysis.

3.264. A second component of the analysis is the compilation of a set of information regarding expected changes since the benchmark survey. For example:

(a) It is possible that in the due time (since the previous collection) improvements in maternal health care programmes have led to an increased survival rate for women. Thus, intercensal cohort survival ratios for females should be higher for younger women than older ones.
(b) If literacy is included in the analysis and government policy has been to strongly support increased school attendance, an increase in the proportion of literate people could be expected.

3.265. There will be a need for careful judgement when the analytical tables show a significant and unexpected difference from the benchmarks. While it may be found that the difference is due to a problem with the current collection, it could also be due to:

(a) A problem in the collection that has generated the benchmarks;
(b) A genuine and previously undetected social change that is being correctly revealed by the current collection.

3.266. In the latter two cases it would be wrong to make any change to the current data set. However, it is crucial that details of the investigation are made known to users (by preparing suitable metadata) so that they would be able to treat and analyse the data correctly. If the analysis indicates that there is a problem with the current collection it will also be a matter for judgement on how to react to it. One proposition is to revise the input processing system to prevent the problem from being continued. After applying such changes, and to avoid introducing further problems, it is essential that they be fully tested and accepted. The second proposition is to decide as to whether to reprocess the records that have already been processed. This decision should be guided by the following:

(a) Significance of the error;
(b) Number of questionnaires that have already been processed;
(c) Time duration for the reprocessing;
(d) Impact of such a decision on other consecutive phases of the census (such as tabulation and dissemination);
(e) Cost and expenditure of that decision.

3.267. The elements involved in data processing require joint and integrated work between analytical (demography, statistics, sociology, economics, geography, etc.) and technological (engineers, analysts, programmers) personnel. Especially for validation where thematic and comparability (with other sources) indicators are generated apart from tabulations to improve data coherence. Before & after comparations (original and validate data) should be included as part of the validation. In the same way, knowing the data dissemination plan helps organize and prioritize both editing and validation.

G. Processing control

3.268. Careful planning and control are required to ensure an uninterrupted flow of work through the various stages from receipt of the census questionnaires through preparation of the database and final tabulations. For questionnaires enumerated through hand-held devices and transferred to a server, a clear data extraction strategy must be put in place with consideration to the scale of the census operation. The plan should provide for the computer edit to follow closely the coding, checking and recording of the data so that errors can be detected while knowledge related to them is fresh and appropriate remedial actions may be taken.

3.269. Countries may wish to establish a computer-based processing management and control system to check individual forms or groups of forms for each enumeration area or for other processing units.

Such a system should link the databases for enumeration areas and other geographic entities with the control information.

3.270. The system would check and manage progress from process to process so as to ensure the completeness of records at each stage of the processing operations. As specified earlier, project management software may support the formal description of different processes and provide an environment to control the execution of all operations connected to an individual phase or status of the census. This system should be fed into the overall quality assurance and improvement system, the management of which is elaborated in paragraphs 2.192–2.200. If a computer-based processing system is established, a close and real-time communication between the headquarters, local offices and field enumerators should also be established. This is beneficial for the control of field staff and the management of logistics of enumeration materials. Also, as any problem occurred in the enumeration field and this solution can be shared through the bulletin board in the system, non-sampling error can be greatly reduced.

3.271. Countries that intend to integrate census operation and census processing should test the whole cycle before conducting the actual census.

H. Master file

3.272. When data editing and validation are in progress, new files consisting of clean data records for each person are produced; these can be assembled to build a master file for later tabulations and indicators (often called the microdata file). This master file, like the raw data files, can have a simple rectangular sequential format. There is usually no need for (but neither should it be discouraged) having the master file organized with a database structure with index files. However, the master file should usually be maintained in geographic order, starting with the lowest geographic entity, sorted by housing unit, household or family. Another method commonly used to generate tabulations involving both the person and the family, household or housing unit is to include in the head of household's record selected characteristics of these latter units. Alternatively, a single hierarchical file can be created involving, for example, person, family and housing unit records. Whatever the chosen structure, the master file must allow for easy checks, controls and computations to be performed. The hierarchical model should be the best shape to generate this master file including all in one.

3.273. One of the most common and problematic errors in census files is that different enumeration areas carry, for one reason or another, the same identification codes, more common now with the mixed capture model. Upon sorting the file, these enumeration areas may have been merged, generating households with abnormal characteristics such as two heads of household, twice the usual number of members, two or more housing records, and so on. To avoid this problem, the enumeration area geocodes should be checked carefully prior to the editing phase. This is best done by keeping a check file of all expected code combinations and marking a code as "used" once an enumeration area using the code has been processed. A module of this functionality can be part of the editing programme. The check file will serve to flag impossible or double identification codes, and towards the end will show which enumeration areas were expected but have not been processed.

3.274. Census master data files are usually very large and require powerful servers to process files of such size or specialized software with those capabilities. Well-equipped desktop systems also have higher computational power and are equipped with much bigger and cheaper mass storage devices. Nonetheless, the hardware infrastructure available to several countries is older, thus two strategies are

applied to reduce file size and to make data management simpler. The first involves working with the next lowest geographic entity as a basis, processing the data on this level and aggregating later to obtain national results. The second remedy is to apply on-the-fly compression and decompression to the storage medium. Census files can be compressed quite significantly to less than 20 per cent of their original size. Since tabulation programs access the data in sequential order, using the compressed data will result in a faster reading process. New compression and storage technologies can be incorporated as well as encryption methods.

I. Methods of tabulation

3.275. Preparing the tabulation plan is the substantive responsibility of the demographers and other subject matter specialists who have the necessary expertise in interpreting the census results. This will require consultation with principal users of the census information (see paragraphs 2.99–2.113). The duties of the data-processing department should be limited to checking the logic of the various accumulations, designing the required programs and producing correct results within the shortest possible time. It is possible that the need for initially unforeseen tables will become apparent, so the census organization should always be prepared to produce additional aggregations. This may involve newly defined classes for certain variables, new types of cross-classifications, differently defined geographic subdivisions, and so on. If the master file is organized according to the principles of relational databases in a relational database management system, original and additional aggregations can be designed according to the relatively easy structured query language statements. Alternatively, if the master file is a list of records with a rectangular structure, online analytical processing (OLAP) tools can be leveraged for creating multidimensional tabulations. It is imperative to ensure that all requisite information for tabulation production is readily available in the master file, as adding it later may prove to be prohibitively expensive. The tabulation plan should meticulously outline the tables to be produced, the variables encompassed in each table, and the cross-classifications to be applied, also specifying the level of detail to be provided in each table. It is essential for the census organization to review and endorse the tabulation plan prior to data processing, with provisions for revisions as necessary throughout the data processing phase. The tabulation plan stands as a pivotal component of the census process, guaranteeing that census data is presented in a format conducive to user needs, while also fortifying the accuracy and reliability of the data.

3.276. Specialized census tabulation software is crucial for census organizations due to its efficiency in processing large datasets, ensuring accuracy through error-checking features, and offering flexibility in generating customized tables and reports. Its user-friendly interface and low cost make it accessible and cost-effective. The software also includes various functionalities like data import/export, data cleaning, table generation, charting/graphing, and report generation, further enhancing its utility. Overall, specialized census tabulation software is an invaluable tool for census organizations worldwide, significantly improving data processing efficiency, accuracy, and cost-effectiveness.

3.277. Tabulation work can also be easily done by software belonging to either of two other classes: statistical analysis and database software. However, these packages have not been designed with large-scale sequential or geographic processing in mind. They may require substantially more computer time than a specialized census tabulation system. In countries with a limited capacity of powerful computers, this can be an important consideration.

3.278. Other factors that should be taken into consideration when selecting software packages for tabulation work include:

(a) The availability of expertise in the census office. It makes no sense to switch to a software system that is only marginally better when this would require a major retraining effort

(b) Opting for a software that the census office is already familiar with is often a more prudent choice; This encompasses factors like the operating system and database software. Incompatibility may lead to the additional expense of acquiring new hardware or software, escalating the overall implementation cost

(c) Compatibility with existing resources at the census office;

(d) Furthermore, the potential need for software customization to execute advanced functions like random perturbation for confidentiality preservation should be evaluated. Customization can be both resource-intensive and costly, underscoring the importance of assessing this requirement prior to finalizing a software choice. Alongside these considerations, factors such as cost-effectiveness, robust vendor support in terms of technical assistance, documentation, and training, scalability to handle large datasets, and robust security features should also be taken into account in the decision-making process. These factors collectively contribute to ensuring the effectiveness and suitability of the chosen software for census tabulation work.

J. Use of administrative data during data processing (for traditional censuses)

3.279. Administrative data may be leveraged during the data processing of a traditional census, mainly for a) imputation of missing or implausible data or non-respondent households or b) validation of results. It may also inform the development of proper edit rules such as checks on the plausibility, or automatic coding systems for specific variables.

3.280. Census data often suffer from undercounting or missing values due to various reasons, such as non-response or errors in data collection. Administrative data can play a crucial role in imputing missing or incomplete census data. When the census form is incomplete, information from administrative sources, such as records of government benefits, healthcare records, or school enrolments, can be used to either replace - when microdata integration is possible - or to impute missing values.

3.281. In cases of non-response, administrative data sources, such as data on electricity consumption from electricity distributors, can be used to identify occupied dwellings. It may also aid in estimating the number of individuals residing in non-respondent dwellings. This will enhance the imputation procedures, resulting in improved census count.

3.282. When administrative data is of high quality, it can serve as an invaluable tool for validating the coverage and content of census data. By cross-referencing the census records with various administrative records such as tax filings, birth and death certificates, or immigration records, discrepancies and inconsistencies can be identified. This validation process helps ensure that the census data are reliable, providing a more accurate reflection of the population. For example, administrative data can be used to verify that individuals listed in the census are indeed living at their reported address and that demographic information aligns with official records. However, as mentioned earlier, this is contingent upon the quality of administrative data. In cases where the quality of administrative at the micro-level is not high yet satisfactory at the aggregates, it may be used in validation of census outputs.

3.283. The use of administrative data in census, either to provide data on specific variables or sub-populations can significantly affect data processing. Among others, it will add a so-called “pre-processing step” that entails cleaning, standardizing and transforming administrative data to assure it suitability for

census purposes. Data-processing staff will certainly need to develop skills to deal with challenges involved in the use of administrative data sources. A close and continuing collaboration, between the data-processing staff, the subject matter staff and relevant administrative agency at all levels is essential in mitigating and tackling the challenges.

K. Security during data processing

3.284. Security concerns during census data processing are paramount due to the sensitive and confidential nature of the information collected. Ensuring the confidentiality, integrity, and availability of census data is critical to maintain public trust and comply with privacy regulations. Here are some of the key security concerns, including disclosure and confidentiality concerns, during census data processing:

i. Data Privacy and Confidentiality:
    i. Data Privacy Laws: Census agencies must adhere to data privacy laws and regulations that govern the collection, storage, and processing of personal data. Violations can lead to legal consequences and damage public trust.
    ii. Anonymization: Techniques like data anonymization and de-identification are crucial to protect individual privacy. Personally identifiable information should be removed or transformed to prevent data from being linked to specific individuals.

ii. Data Encryption:
    i. Data in Transit: Encrypting data while it is being transmitted over networks (e.g., from data collection points to data processing centres) helps protect against interception and eavesdropping.
    ii. Data at Rest: Data at rest, whether stored in databases or backups, should also be encrypted to prevent unauthorized access in case of data breaches or physical theft.

iii. Access Control: Role-Based Access: Implement role-based access control to restrict access to census data based on job roles and responsibilities. Only authorized personnel should have access to specific data sets.

iv. Data Leakage and Disclosure Prevention: Data Leak Prevention : Implement Data leak prevention solutions to monitor and prevent the unauthorized sharing or leakage of sensitive data, either intentionally or unintentionally.

v. Security Auditing and Monitoring:
    i. Logging and Auditing: Maintain detailed logs of data access and processing activities. Regularly review these logs for suspicious activities and potential security breaches.
    ii. Intrusion Detection and Prevention Systems : Deploy intrusion detection and prevention systems to detect and respond to potential threats and attacks in real-time.

vi. Physical Security:
    i. Data Centre Security: Ensure that data centres housing census data are physically secure with access controls, surveillance, and environmental protections to prevent unauthorized physical access or damage.

vii. Incident Response Plan: Develop an Incident Response Plan: Prepare for potential security incidents by creating a well-defined incident response plan. This plan should

outline the steps to take in the event of a breach and how to communicate with affected parties.

viii. Compliance and Audits: Regular Audits: Conduct security audits and assessments to verify compliance with security policies and regulations. This includes third-party audits and assessments for independent verification.

ix. Data Retention and Disposal: Data Retention Policies: Define and enforce data retention policies to ensure that data is retained only for the required duration and is securely disposed of when no longer needed.

x. Employee Training and Awareness:

i. Security Training: Provide regular security awareness and training programs for employees and contractors to educate them about security best practices and the importance of data protection.

ii. Vendor Security: If third-party vendors are involved in data processing, ensure they adhere to stringent security standards and practices to safeguard census data.

xi. Public Communication: Develop a clear communication strategy for addressing security incidents, ensuring transparency, and maintaining public trust.

3.285. In summary, maintaining security and confidentiality during census data processing requires a comprehensive approach that includes technical, organizational, and procedural measures. It's crucial to prioritize privacy and security to protect sensitive information and maintain the integrity of the census process.

IX. Evaluation of the results

3.286. A census evaluation programme should be developed as part of the overall census programme and integrated with other census activities. The scope and objectives of the evaluation programme should be decided well in advance to determine early enough the adequate resources (both financial and human) needed for the evaluation programme. It is important to establish a team responsible for the planning, organization and implementation of the evaluation programme. The cost of evaluation should be covered in the census budget as a separate item.

A. Purpose of census evaluation

3.207 The quality of population and housing census data is very important for many reasons, including building public trust in and understanding of the national statistical system. The purpose of census evaluation is to provide users with an acceptable level of accuracy and confidence when utilizing the data, and to explain errors in the census result. It is therefore important to choose an appropriate way of sending out these messages to the right group of stakeholders.

3.208 The evaluation methods discussed here are those that apply to traditional censuses. To some extent they also apply to register-based censuses and other census methodologies, but these also present their own particular challenges and solutions[59].

3.209 It is universally accepted that a population census is not perfect and that errors can and do occur at all stages of the census operation, but these errors should be measured. Errors in the census results are classified into two general categories – coverage errors and content errors. Coverage errors are the errors that arise due to omissions or duplications of any of the enumeration units – persons, households or housing units – in the census enumeration. The sources of coverage error include incomplete or inaccurate maps or lists of enumeration areas, failure on the part of enumerators to canvass all the units in their assignment areas, duplicate counting for persons who have two or more places of residence, persons who for one reason or another do not allow themselves to be enumerated, erroneous treatment of certain categories of persons such as visitors or non-resident aliens, and loss or destruction of census records after enumeration. Content errors are errors that arise in the incorrect reporting or recording of the characteristics of persons, households and housing units enumerated in the census. Content errors may be caused by several factors, including poorly phrased questions or instructions, or enumerator errors in phrasing the census questions; inability or misunderstanding on the part of respondents in respect of answering specific items; deliberate misreporting; errors due to proxy response; and coding or data entry mistakes.

3.210 Many countries have recognized the need to evaluate the overall quality of their census results and have employed various methods for evaluating census coverage as well as certain types of content error. Comprehensive evaluation should also include assessment of the success of census operations, in each of its phases, including evaluation of activities such as the geographic frame, census publicity campaign, data collection, data processing, data dissemination and data utilization. Countries should ensure, therefore, that their overall census evaluation effort addresses the census process, as well as the results. The present section is devoted to evaluation of the results. However, the section on the quality assurance and improvement programme (paragraphs 2.172–2.228) provides further recommendations relating to controlling and assessing the quality of census operations.

3.211 Evaluation efforts focused on census results should generally be designed to serve the following objectives: first, to provide users with some measures of the quality of census data to help them interpret the results; second, to identify as far as is practicable the types and sources of error in order to assist the planning of future censuses; and third, to serve as a basis for constructing a best estimate of census aggregates, such as the total population, or to provide census results adjusted to take into account identified errors at national or subnational levels if some errors such as coverage error are substantial and the validity of census results is questionable.

3.212 As the decision to adjust census figures is sensitive, it is bound to be decided at the highest levels of the government bureaucracy. There is also some critical statistical consideration that should be very carefully weighed in. Consideration must be given to what geographic domains the adjustment would cover, knowing that such adjustments have an effect on demographic distributions. Transparency in the methodology used for adjustment is critical to maintain trust with stakeholders. In cases where census results are presented without adjustments, the results of the census evaluation should be taken

[59] See, for example, *Census Quality Evaluation: Considerations from an international perspective*, note produced by the UNECE Secretariat for the 11th Joint UNECE/Eurostat Meeting on Population and Housing Censuses, Geneva, 13–15 May 2008, http://www.unece.org/stats/documents/2008.05.census.html.

into account in the preparation of post-censual population estimates and for analytical and planning purposes.

3.213 The final publication of census results should include an estimate of coverage error, together with a full indication of the methods used for evaluating the completeness of the data. The publication should also provide users with some cautions or important notes about the results, in addition to some guidance on how they might use the evaluation results. It is also desirable to provide, as far as possible, an evaluation of the quality of the information on each topic and of the effects of the editing and imputation procedures used.

3.214 The range and quality of editing in regard to the correction of the inconsistent data and imputation possible in a population census are greatly enhanced by the use of computer edit programs that permit inter-record checks (for example, the replacement of missing values based on one or more items on the basis of reported information for other persons or items). If any imputation is made, the topics affected, the methods used and the number of cases affected should be documented and clearly described in the census evaluation report.

3.215 The results of the evaluation of census results should be made available to users with a measure of the quality to help them interpret the results.

3.216 As discussed in the following subsection, a number of methods exist for carrying out a census evaluation. In practice, many countries use a combination of such methods in order to fully serve these objectives.

B. Methods of census evaluation

3.217 The choice of evaluation methods to be used depends upon the evaluation objectives. These, in turn, depend on national census experience in terms of past and anticipated errors, user and public concerns, and the financial and technical resources available for evaluation. The decision needs to be made whether to measure coverage error, content error or a combination of the two. In addition, both gross and net error must be considered in developing the overall evaluation plan. Gross coverage error in a census is defined as the total of all persons omitted, duplicated or erroneously enumerated. Net coverage error takes into account the underestimates due to omissions and the overestimates due to duplications and erroneous inclusions. When omissions exceed the sum of duplications and erroneous inclusions, as is usually the case in most countries, a net undercount is said to exist; otherwise, a net overcount results. Similarly, both gross and net content errors have to be considered in the evaluation design.

3.218 Numerous methods are available to estimate the coverage and content error of censuses. These include simple techniques of quality assurance, such as internal consistency checks. Comparisons of results with other data sources, including previous censuses, current household surveys and administrative records, are also useful techniques. Such comparisons may be made in aggregate by comparing the overall estimates from two sources (net error only). Differences in conceptual frameworks and quality assessment for all these sources should be considered when using them as references to assess the coverage and content error of censuses. Alternatively, record checking, whereby individual census records are matched against alternative sources and specific items of information are checked for accuracy, may be used. Both gross and net errors can be estimated in record checks, which may involve field reconciliation of differences, a costly exercise that cannot be

overlooked. An important but complicating factor in the use of record checks is the requirement of accurate matching. It is essential to plan carefully for this aspect, since the operation can be tedious and costly. It should be noted that record checks are best employed to study the coverage of certain segments of a population, such as children whose birth records are complete, since these checks are, by definition, limited to subpopulations with complete, accurate records.

3.219 Demographic analysis and post-enumeration surveys[68] are two very important methods for evaluating census data, and these are discussed in further detail in the following two subsections.

C. Post-enumeration survey

3.220 The post-enumeration survey can be defined as the complete re-enumeration of a representative sample of the census population and matching each individual who is enumerated in the post-enumeration survey with information from the main enumeration. The objectives of the post-enumeration survey can be summed up as follows:

a. To assess the degree of coverage during census enumeration;
b. To examine the impacts of coverage deficiencies, if any, on the usefulness of the census data;
c. To obtain information for the design of future censuses and surveys;
d. To examine the characteristics of persons who may have been missed during census enumeration.

3.221 While a post-enumeration survey can be designed to provide a comprehensive evaluation of coverage and content error, especially when supplemented by and integrated with detailed demographic analysis of census quality, the methodology of a sound post-enumeration survey is complex, so that countries must accordingly weigh with care the demanding technical requirements and the costs of conducting a successful post-enumeration survey, and elaborate a clear statement of its objectives, before deciding to undertake such a survey.[69] Careful advance planning is crucial. To be valid, a post-enumeration survey has to function within a number of operational and statistical constraints. These include the requirement that the survey be carried out within a few months of the end of the census to ensure that the impact of natural population changes (births, deaths and migration) and lapses in respondent recall do not hopelessly complicate the exercise.

3.222 Another basic property of post-enumeration survey design and execution involves matching and reconciliation. Matching the post-enumeration survey person record or household record against the corresponding census record is an operation whose performance must be of very high quality to ensure that inaccuracies in the post-enumeration survey itself do not effectively ruin the estimate of coverage error. Matching is especially difficult in countries where many surnames are identical or where individuals are known under more than one name, and well-defined street addresses do not exist. Part of the matching operation usually involves a field visit to reconcile differences between the census and the post-enumeration survey with regard to either coverage or content. Reconciliation of course adds another dimension of cost and complexity, since it entails a second visit to the field for purposes related to the post-enumeration survey.

3.223 Clearly defining the objectives of a post-enumeration survey is the first and most crucial step in planning the survey. The objectives might include estimation of coverage error at the national level; estimation of coverage error for major subnational domains or population subgroups, each with its own specified level of precision; and measurement of content error for specific census items.

3.224 As mentioned, the design of a post-enumeration survey is complex and there are various alternatives, primarily depending upon whether single or dual system estimation is to be utilized. A number of references are available that set out highly detailed procedures for designing a post-enumeration survey and the conditions under which they may or should be considered.[70] However, in general, when designing a post-enumeration survey, the following considerations should be taken into account:

a. The time between the census and the post-enumeration survey should be minimized to avoid as much recall error as possible and the impact of population changes (births, deaths and migration).

b. The post-enumeration survey must be independent of the census. Interviewers must not have census information about the areas where they are working. When interviewers have knowledge of census responses, they tend only to confirm what the census recorded.

c. To preserve the independence of the post-enumeration survey, its data collection and processing operations must be completely separate from the census data collection and processing.

3.225 The sample design for a post-enumeration survey must be based upon sound probability sampling methods taking account of the measurement objectives of the evaluation study. These usually include the need to estimate census coverage with a certain degree of reliability. In addition, the estimates of coverage may be wanted for geographic areas such as provinces or states and large cities, for urban–rural comparisons and so forth. Such requirements also greatly affect the sample design of a post-enumeration survey, as the necessary sample size is increased substantially when estimates of subnational coverage (or undercoverage) are wanted.

3.226 Sometimes a post-census survey is designed to measure content error only, in which case it is usually known as a reinterview survey. The advantage of a well-designed reinterview survey is that the results are more accurate than those of the census insofar as the operation is much smaller and can be more effectively controlled. Estimates of relative response bias can be obtained from a reinterview survey, which (rather than the census) is generally taken as the standard in this area on the grounds that the survey, with its better-trained interviewers and more intensive survey procedures, yields superior results.

3.227 As part of the design of some post-enumeration surveys, a sample of the original census enumeration districts, blocks or areas is chosen and recanvassed for the survey. As regards methodology, this constitutes a useful reinterview technique for measuring content error, and such an element in the design is often put into practice because the matching operation between survey and census records is then dramatically simplified.

D. Demographic analysis for census evaluation

3.228 Demographic analysis offers a powerful methodology for evaluating the quality of a census and countries are encouraged to use it as part of their overall census evaluation methodology. A wide variety of demographic techniques have been developed and used, ranging from visual inspection of census data to comparative analysis of two census age distributions. A basic procedure for assessing census quality on age-sex data is graphical analysis of the population pyramid. Age heaping, or the tendency of respondents to report a particular ending digit, is a useful check of the quality of age reporting, as are

sex ratios by age and certain summary indices of age-sex data, including the United Nations age-sex accuracy index, which extends age-sex ratio analysis by observing deviations of the observed age-sex ratios from the ones expected for each five-year age group and combining the results into a single score.[60] Other summary indices are Whipple's index and Myer's blended index, used for judging age heaping.

3.229 Stable population theory has also been used in the past to assess the quality of census distributions by age and sex. It is based upon measuring the reported age-sex distribution against that of an appropriately chosen stable population, assuming that the population is not affected by significant international migration. However, nowadays there are few countries where the other two conditions assumed under the model, namely constant fertility and constant or recently declining mortality, are satisfied. Recent declines in fertility render the technique less useful as an evaluation tool, since the technique is sensitive to changes in fertility levels. Nevertheless, if the population is closed to migration, it can be assumed to have been stable in the not-too-distant past and if approximate estimates of recent fertility and mortality declines and recent growth rates are available, it may still be possible to assess the plausibility of the current age-sex structure in the light of these trends by iteratively fitting projected population structures to the observed numbers.

Comments PG: This section is outdated. The remaining para needs to be revised building on [3.232] and to refer to the intercensal cohort survival approach and evaluation of two or more censuses by birth cohorts, as well as the use of the reverse survival approach to evaluate and potentially adjust the enumeration of children under age 15.
See IUSSP manual + UN/WPP methodological report and census evaluation/adjustment protocol.

3.230 The methods mentioned above, while useful in providing an overall assessment of census quality, cannot differentiate the sources of census error in terms of the relative contributions from undercoverage (or overcoverage) or content error. Better information about coverage error, through demographic analysis, derives chiefly from comparative analysis of data from successive censuses, in which four methods are used.

3.231 The four methods are (a) derivation of an expected population estimate taking account of vital registers of births, deaths and net migrants between censuses, as compared with the latest census; (b) population projections based on the results of the prior census plus data on fertility, mortality and migration from various sources and comparing the projected estimates with the new census results (cohort component method); (c) comparison of two census age distributions based on intercensal cohort survival rates; and (d) estimates of coverage correction factors using regression methods to make the age results from the two censuses mutually consistent (cohort survival regression method). It should be noted that the first two methods would probably have to be restricted to evaluation studies of coverage at the national level, especially in countries that do not have good subnational data on migration.

Comments PG: An extra/newer approach as used by Colombia latest census involves also using remote sensing/GIS as done with WorldPop approach to detect human settlements and identify potential geographical coverage gaps, especially in rural areas. See https://pubmed.ncbi.nlm.nih.gov/36977422/

[60] See *Methods of Appraisal of Quality of Basic Data for Population Estimates: Manual II* (United Nations publication, Sales No. E.56.XIII.2).

E. Comparison with administrative sources

3.232 Comparing census data with administrative data in a country or specific area can provide valuable insights into changes since the last census. By analyzing administrative data (including data on development trends), it is possible to assess how well the new census data captures these changes and whether the census accurately reflects the characteristics of the underlying population.

3.233 Administrative data sources, such as population registers, records of birth, death, migration, etc. provide estimates of population changes. By comparing census counts with these data sources at aggregate level, census under- or over-coverages can be assessed. However, in these comparisons, the alignment of geographical boundaries, reference periods, and the target population covered by administrative data and the census should be considered. The coverage of census at the household and dwelling level can also be assessed if relevant administrative data on households (like family registration systems), building or property, etc. exist.

3.234 Availability of high-quality administrative data enables comparisons at the micro-level as well. Such comparisons can be facilitated if national unique identification numbers exist and have been included in the census questionnaire. This will not only assist with locating coverage errors but also enables evaluation of the content. Administrative data sources, such as healthcare records, school enrolments, and tax filings, social welfare programs, and employment record, etc. can be compared to census data to identify content errors related to demographic or socio-economic information.

3.235 The comparison with administrative data involves the following stages:

1. **Identify the relevant administrative data sources:** Administrative data can include records from government agencies, local authorities, educational institutions, healthcare systems, and other organizations.
2. **Gather administrative data:** Collect the administrative data that aligns with the relevant timeframe and census topics. This data may include population records, building permits, school enrollment figures, healthcare statistics, or employment data.
3. **Analyze administrative data trends:** Examine the administrative data to identify any significant changes or trends that have occurred. Look for relevant indicators such as population growth or decline, changes in employment sectors, infrastructure development, or changes in school enrollment.
4. **Compare census data with administrative data:** Compare the findings from the analysis the administrative data with the census data and evaluate the level of alignment and consistency between these sources. Comparing census data with administrative data can help identify potential discrepancies and provide insights into the accuracy of the census data.

3.236 When comparing census data with administrative data it is important to: assess the quality and reliability of both the census data and administrative data; understand data definitions and methodologies; consider spatial variations by analyzing the data at different geographic levels; conduct additional research to validate the findings; recognize that both census data and administrative have limitations; explore demographic changes; consider socio-economic factors; assess infrastructure and development trends; seek stakeholder perspectives; document findings; and communicate the results of the evaluation.

F. Adjusting census results

3.237 Achieving full coverage of the population is one of the major challenges in carrying out any census. Nonetheless some coverage error is unavoidable. Hence it is important to measure, analyse and report on coverage error.

3.238 Population under coverage is regarded as one of the most important sources of error affecting census data. This results from missing housing units, households, or persons during census enumeration Over coverage, on the other hand, results from duplications. These two sources of error can distort the distribution of population characteristics estimated from census data.

3.239 In some countries the evaluation of census coverage forms an integral part of the census quality assurance and improvement component of the quality management framework, and feeds directly into the published census results. In other countries, it is a separate exercise, used either to measure the census coverage error or to adjust census results.

Adjustments for Traditional Censuses

3.240 In the case of a traditional census, the most common problem is under-coverage. The evaluation is usually done via an independent post-enumeration survey of a sample of enumerations areas (add link to PES paragraphs) or via a Reverse Record Check methodology. In both cases the purpose is to measure the accuracy of the census. By independently surveying a sample of the population, the survey estimates the proportion of people and housing units potentially missed or counted erroneously in the census[61].

3.241 On the basis of net coverage rates, adjustments may be made to population census results. Using results from a carefully designed and implemented PES, under or over counts can be converted into adjustment factors and a correction factor attached to each record. The census population is increased or decreased accordingly by such factors. The adjustment is usually done at the national level or for other large domains.

Adjustments for Register-based Censuses

3.242 In the case of a register-based census, there is also the issue of over-coverage of registers to consider. If, for example, people have not officially declared their emigration, their records may be kept

[61] Usually within a Dual System Estimation model or so-called capture-recapture method (with the PES being the second 'capture' while the census itself is the first 'capture'.

unchanged in administrative registers, and consequently be the cause of over-coverage in the census[62].

3.243 Another possible means of avoiding coverage errors is to create a residency index on the basis of the records held in multiple registers in order to determine a so-called 'signs of life' score. The approach is to define for all possible residents the 'sign of life' as a binary score (with a value 0 or 1) for each record in each register. Using these signs of life as explanatory variables, it is possible to build a model forecasting the size of under and over-coverage[63].

3.244 Another possible way to account for coverage errors in statistical registers is based on the correction of the population register through estimates obtained from two coverage samples for each area. More precisely, a field (area) sample of people living in the area on census day is used for estimating the register under count (the "U sample"), and a sample of people registered in the same area is used for estimating the register over count (the "O sample") The U sample can also be used for collecting socio-economic information.

3.245 Some countries are using a combination of the methods above to investigate and refine methodologies for using of administrative sources as a means to understand coverage issues in the census and make adjustments to census data. This is an emerging area of statistical practice. The availability of administrative data sources is changing rapidly as more data is captured and stored electronically. Therefore, approaches differ from country-to-country and often need to be tailored based on national circumstances.

Adjustments for Combined Censuses

3.246 There are several options to check the coverage errors in the case of a combined census. One is to use the methodology commonly adopted for traditional censuses (for the field enumeration component), by organizing a post-enumeration survey and estimating both over-coverage and under-coverage statistically. The 'signs of life' approach, comparing records from different registers, can then be adopted to provide information on the level of under-coverage and over-coverage of the register-based component of the census.

The decision to adjust, or not to adjust, census figures

62 UNECE: Guidelines for assessing the quality of administrative sources for use in censuses, United Nations, Geneva (2021), http://www.unece.org/fileadmin/DAM/stats/publications/2018/ECECESSTAT20184.pdf

63 UNECE: Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses, United Nations, New York and Geneva (2018), http://www.unece.org/fileadmin/DAM/stats/publications/2018/ECECESSTAT20184.pdf

3.247 Many countries decide not to adjust census figures on the basis of Post Enumeration Survey results. Adjusting the census figures can be politically sensitive. Despite knowing coverage error rates from the Post Enumeration Survey, some countries choose to leave the census results unadjusted. In other countries, census figures are adjusted to take into account the Post Enumeration Survey results.

3.248 Countries may choose to decide about adjustments to the census results when coverage errors are substantial, and the validity of census results is questionable. Consideration must be made to what geographic domains the adjustment would cover knowing that such adjustments have an effect on demographic distributions.

3.249 In those countries where the total population figure has been adjusted for under-enumeration or over-enumeration, both the enumerated figure (the population count) and the adjusted population figure (the population estimate) should be shown and described. The detailed tabulations may, however, be based only on the population that was actually enumerated.

3.250 Whatever decision is made (to adjust or not to adjust census data), it is important that the decision is communicated, and any adjustment fully described alongside the census results.

G. Acceptance of results

3.251 In countries with limited prior census experience and without a well-functioning civil registration system, where population data are based largely on estimates, it is important to inform the users, particularly the governmental authorities, that the census results could differ from such estimates and to explain the reasons for these differences. In some cases, there may be doubts expressed about the census results; usually those doubts focus narrowly on the total population of the country, major subdivisions or population subgroups, rather than on the bulk of the census data relating to characteristics of the population or on the data for local areas. In this situation, it may be possible to take such doubts into account by modifying the census evaluation programme or by adding appropriate qualifications to the text of the census reports or in tabular footnotes. Nevertheless, the government may proceed with the processing and dissemination for official purposes. In any case, every effort should be made to process and evaluate the full census and to make appropriate use of as many of the census tabulations as possible.

X. Census products, dissemination and utilization

A. Introduction

3.252 The population and housing census is a statistical operation of exceptional value to every country. It is the primary source of basic national population data for administration and for many aspects of economic and social planning. Consequently, data from national censuses represent a valuable public good that should be widely promoted by national statistical and census offices in order

to enhance its utilization by the various users. Thus, the census should not be an end in itself but should be backed by the value of the results, in terms of utilization, by the diverse categories of data users.

3.253 While the importance of disseminating census results to the fullest extent possible should be stressed, there are, nevertheless, some things that are essential for the national statistical or census office to keep in mind. Among these is the prerequisite for national statistical and census offices to ensure openness and transparency in the way the results are disseminated. It is equally important that national statistical and census offices maintain professionalism and demonstrate neutrality and objectivity in the presentation and interpretation of the results and are free from real or perceived political interference so that the objectivity and impartiality of the statistics is assured. This in turn will build trust in and acceptance of the results. Furthermore, the disseminated census results should be of sufficient quality to meet user needs, and safeguards should be in place to ensure individual information is kept confidential.

B. Plans for census products and data dissemination

3.254 In order to maximize the utilization of results from their population and housing censuses, national statistical and census offices should have a sound dissemination programme whose objective is to promote the benefits and applications of census data. The statistical or census office should develop and implement an effective strategy for producing and disseminating output products and providing related services based on the demonstrated needs of the diverse users of census data. What follows are some salient issues for an effective dissemination programme.

1. Developing a dissemination strategy

3.255 A census is not complete until the information collected is made available to users in a form and scale suited to their needs. In order to fulfil this requirement, it is essential to develop a strategy for producing and disseminating outputs taking into account all potential users of the data. This in turn requires identifying potential users of census data and their demonstrated needs so that appropriate products and related services can be developed.

3.256 An effective dissemination strategy should have a diversified approach for meeting the requirements of different users. Appropriate technologies and media need to be identified for effective and easy dissemination of census data and information. When developing a dissemination strategy for census data, it is important to take into account geospatial information to provide a spatial context to census data. Geospatial data can be used to create thematic maps that visualize census information more user friendly. Geospatially enabled statistics, in particular at a high geospatial spatial resolution, can allow statistical organisations to produce analytical outputs at more disaggregated levels and conduct a wide range of spatial analysis.

3.257 A number of key elements should be taken into account in the development of a strategy for census data dissemination, including identifying the diverse categories of users and their data needs and uses through (a) stakeholder engagement, (b) data needs assessment, (c) the media of dissemination, (d) metadata to aid in the interpretation of the results, (e) legal and ethical consideration, confidentiality and privacy measures, (f) assessing the required technologies to meet user needs, (g) dissemination policy, (h) quality assurance in terms of accuracy and timeliness, (i) available financial and human

resources, (j) the need for continuous improvement along with a feedback mechanism for dissemination products, and (k) alignment with international standards. The first five elements are covered in subsequent sections of this chapter. The remaining are summarized below.

Technology. Given the widespread availability and use of technology for easier production and access to census products, national statistical and census offices should evaluate which ones are suited to the needs of their data users taking into account budgetary and human resource constraints. These technologies include use of GIS and thematic map generators, new sophisticated data base systems and interactive web access, including client-customized table generators.

Dissemination policy. When planning the dissemination programme, a dissemination policy should be established as well. This policy should cover issues such as ways of marketing the census products, which in most cases means mainly how to inform a wide range of potential users about the availability of the products. Where applicable, a clear pricing policy should also be determined, and a decision made as to the conditions under which external distributors are allowed to disseminate census data. The dissemination policy should be transparent and cover issues connected with the protection of the confidentiality and privacy of personal data, and the measures that will be used for each of the different products.

Quality assurance. Quality refers primarily to user needs and satisfaction. Even if data are accurate, they do not have sufficient quality if they are produced too late to be useful, or cannot be easily accessed, or conflict with other credible data, or are too costly to produce. Therefore, quality is increasingly approached as a multidimensional concept. and the output of any statistical exercise should possess the following attributes: accuracy, relevance, reliability, timeliness, punctuality, accessibility, clarity, coherence, comparability and metadata.[64] Management of quality in census dissemination is driven by concerns to (a) deliver relevant products and services while (b) maintaining accuracy of the data, and (c) timeliness and predictability of data release within agreed cost constraints.

Budget and human resources. Two obvious key elements (usually constraints) in the development of strategies for census data dissemination are the budget that can be allocated and the availability of human resources. With the high relevance of new technologies in all the census stages, and in particular for data dissemination, this is a factor that needs to be carefully analysed when deciding about the specific strategy of census data dissemination. The alternative to the recruitment of human resources may be contracting out some of the dissemination activities, in particular those connected with the development of more sophisticated systems. However, this solution needs to be carefully considered, and it is extremely important to ensure that the contractor is committed to the census project until its very end, and that at least some of the new abilities remain in the organization for further use.

[64] United Nations Statistics Division (2012): *Guidelines for the template for a generic national quality assurance framework (NQAF)* – available at http://unstats.un.org/unsd/dnss/docs-nqaf/GUIDELINES%208%20Feb%202012.pdf.

a. Consultation with data users

3.258 The demand for and use of statistical products and services must drive all census operations. National statistical and census offices should have a sound strategy for developing suitable products and services to respond to the diverse needs of data users so as to promote the utilization of census results. Such strategies should be based on an active dialogue with the users regarding their needs in terms of products and the format of those products.

3.259 The user consultation process in terms of census products is a major factor in the development of a dissemination programme. The type of consultation discussed in this section complements the consultation that is undertaken to determine census content (see paragraphs 2.99–2.102). The work done at this stage of the census is important in achieving the objective that the census is relevant to users, which is a major indicator of the quality of the census. The selection of suitable census data products and related services should be guided by a detailed assessment of user requirements with due consideration given to budgetary and human resource constraints.

3.260 Plans for what and how products will be disseminated should be made early enough in the planning process and shared with potential users in order to get their feedback. Based on this feedback, the national statistical or census office can tailor its data dissemination programme to suit the requirements of the users. Maintaining good communication and obtaining feedback from users is also important for making modifications to products and services, including being able to respond to user requests that become known later in the programme.

3.261 Based on the foregoing, it is important to note that the supply of census products and services goes far beyond the first couple of years after the census. It is important, therefore, that budget and human resources are available for many years after the end of the census collection activities.

b. Plans for outputs

3.262 It is important for census offices to consult stakeholders and identify their needs during the preparatory phase to proactively anticipate the type and format of census products to be produced. This is to ensure that census products are relevant, responsive and add value to the current policy questions and stakeholder needs. It is recommended that census offices include a census products plan and budget as part of the preparatory phase.

3.263 A wide range of statistical products can be made available to the public, the private sector, government agencies, local authorities and the academic and research communities. A detailed plan for producing different census outputs should be guided by early user consultations (see paragraphs 2.99–2.102) to ensure data and information requirements will be met in a format commensurate with user needs and demands; such a plan will also be a useful guide to prioritizing data processing and tabulations. To help communicate the timeline for dissemination, a calendar with the expected release dates for the preliminary results, final results, as well as the timing of the release of the various census products should be published.

3.264 The definition of geographic output products and services and the scheduling of their release needs to be closely coordinated with the timetable for the overall census project. The tabulation of census data may require information from the census geographic unit, and thematic maps and digital geographic databases can only be completed once census data processing has been completed.

3.265 With the rapid development of technology, census data users have an increasing interest in a broad range of products and services from the census organization. The types of output that census offices may produce and disseminate must be current and may include printed products, static electronic products, interactive electronic products, customized products, user-interactive products and special audience products and services. Partnerships with key stakeholders are encouraged in the development of the various census products.

3.266 Some data users will need specialized products that the census organization is not planning to produce as part of the general census programme. It is recommended that the census organization establish a service to meet such specialized requests. Pricing of special products and services may be included in a pricing policy.

3.267 Printed publications, are becoming less popular as many country move towards digital publications . For those countries that produce printed publications, target dates for publication should be determined well in advance and processing and printing programmes should be planned accordingly. In addition to traditional methods of printing, there are various methods of reproduction available that are fast, economical and of good quality, and these should be investigated. For an increasing number of users, environmental concerns and a lack of searchability means online electronic data dissemination is preferred over printed paper. Electronic dissemination reduces cost, and unlike weighty paper publications, supports the direct availability of the data for further computer processing. In addition to the processed tabulations, sample data at the unit level are also provided by some countries for research purposes. In such cases, the sample should be carefully drawn to ensure an adequate level of representation while at the same time ensuring that anonymity is not compromised. Some countries have also adopted very creative techniques for data dissemination, cross-tabulations, infographics, story maps and visualization to allow users to understand relationships and patterns within the data. The development of such data products should be part of the planning process of the census.

3.268 Not all of the processed materials need to be disseminated widely or in a single format. Tabulations required by only a few users can be supplied in unpublished form. Some data may not be tabulated until they are required at a later date. The information stored in the census database allows fast and relatively inexpensive production of additional tables. Countries may offer on-demand services to provide census information to users who require tables or other outputs not produced, or aggregates not available, through other means. If suitable electronic dissemination is available, customized tabulations and applications might also be designed and extracted directly by end users. In this case, the census organization should prepare in advance and then implement an authorization and security policy, so that the risk of breaching confidentiality in data provided to outside users is avoided.

Comments ID: The above paragraphs need to be updated based on current technological developments and practices.

2. **Tabulation programme**

3.269 In **most countries, the tabulation programme represents a compromise between the full range of desired tabulations and the limits imposed by practical circumstances. To ensure that this compromise is made transparently and efficiently it is important that planning the census dissemination task is started at the earliest stage of the census development cycle by a round of user consultations. Once the census-testing programme has identified a practicable range of data items to be included in the questionnaire, data users should again be consulted on the specific cross-**

tabulations required and the relative priority for their production. It is essential that the programme be outlined sufficiently early so that the procedures and costs involved are investigated thoroughly before a final decision is reached. The type of questionnaire and the method of enumeration may limit the kinds and amounts of data that it is possible to collect. Publication time and costs, and the data-processing resources available, will determine the number and complexity of the tabulations that can be produced within a reasonable time. This will enable prospective census data users to make firm plans and the census data processing staff to complete all systems analysis, programming and testing work in a timely manner.

3.270 **The tabulations presented on the website of the United Nations Statistics Division are those fulfilling the most essential or generally required information. The databases of census information can be used throughout the intercensal period to address the needs of specialist users for whom these tabulations are not adequate.**

3.271 **It is important to plan the tabulation programme in such a way that final results can be issued within a reasonable period of time after the enumeration and before the information has become out of date for current needs. It is desirable that the details of the tables be prepared, and the order of their preparation be decided early in the planning so that the processing of the data is not be delayed.**

3. **Geospatial information for analysis and dissemination**

3.272 **Geospatial** information can unlock the full potential of census data and facilitate its use for territorial analysis, to describe and investigate phenomena, such as demographic trends and spatial inequalities in access to services (schools, health facilities, etc.), that unfold across time and space.

(a) Dissemination geography

3.273 Maps, which are now commonly in the form of digital products, play an increasingly important role in the dissemination phase of the census. Statistics compiled from census data can be geographically referenced and provide for methods of analysing the geographic characteristics of those statistics. Maps may then be used effectively to relate statistical data to the geographic area to which the census results refer. This makes the statistics easier to understand and more readily usable by both expert users and the general public.

3.274 If a complete digital census geographic database has been created, then statistical databases for administrative or statistical units can be produced simply through aggregation. For the countries that do not use digital techniques for the production of enumeration area maps, options still exist to develop a digital georeferenced census database at this stage for producing publication-quality maps to accompany census reports, for distribution to outside users who want to analyse census data spatially or for internal applications. This database can be compiled for a suitable level of the administrative hierarchy or for other aggregated statistical regions. At that level of aggregation, the resources required for producing a digital database are much less than those necessary for a complete digital enumeration area map database.

(i) Linking collection to dissemination geography

3.275 An essential feature of the population and housing census is its diversity in terms of the geographic level at which data can be disseminated. This is due to the ability of the census to produce

statistics that can be disseminated at the lowest geographic level (small area), through a geographic hierarchy up to the country level. Consequently, one of the earliest decisions in census planning relates to the administrative and geographic areas for which census data on diverse socioeconomic characteristics of the population will be reported and disseminated in order to satisfy the needs of the various data users, taking into consideration statistical confidentiality.

3.276 In addition to administrative units, most countries will have a number of other sets of areas that are used for different purposes and for which census data will need to be compiled.[65] Such areas, which have special uses, include health regions, electoral districts, urban agglomeration or metropolitan areas, grids, and utility zones (water or electricity supply districts). It should be noted that some of these areas may not fit perfectly into the administrative hierarchy of the country. It is important, therefore, that to the extent possible these reporting units are taken into account when designing enumeration areas in order to facilitate generation of census data for these regions. This draws attention to the fact that when delineating collection geography (enumeration areas), it is essential that dissemination geography is kept in sight.

3.277 Two somewhat different methods are available to provide the census with a flexible capability for generating tabulations in terms of a wide variety of geographic aggregations, including those needed for public and private sector data uses at the local level. The first method simply extends the traditional hierarchical system for coding all major and minor civil divisions so as to cover at the lowest level the enumeration area, sometimes referred to as the "enumeration district". The second method, which at greater cost permits finer geographic specificity, is usually based on some coordinate or grid system, such as latitude and longitude. This method is often referred to as a "geocoding system". Particularly in the absence of a comprehensive system of street names, numbers or geocoded coordinates of housing and building units, the first method, which uses the enumeration area as the key unit to produce small area data, is to be preferred.

3.278 The fact that census data, whether published or unpublished, are available by enumeration area provides for considerable flexibility. Such flexibility can be of value given that the geographic divisions used by various branches of the administration or by other data users do not always coincide and may therefore require different regroupings. Moreover, when changes are planned in administrative boundaries, tabulation of census data by the planned new entities can also be facilitated through the enumeration area approach. However, if these changes cross enumeration area boundaries and it is decided to try to retabulate the census according to the new boundaries, very complex recoding of individual records may be involved. As an alternative, statistical concordances, showing the quantitative relationship between the previous and current classifications, could be used. Further, where buildings or housing units have been geocoded, these geocodes can be used to directly allocate each household to the correct area under either classification.

(ii) Uses of small geographic data

3.279 Censuses provide data from the highest to the lowest geographic levels of aggregation. Tabulations from census results yield relevant statistics for any reasonable combination of characteristics for the country as a whole, regions or provinces, down to small areas such as localities, and even enumeration areas and geographic grids. This important feature of the census makes the data

[65] *Handbook on Geospatial Infrastructure in Support of Census Activities*, Studies in Methods, No. 104 (United Nations publication, Sales No. E.09.XVII.8).

amenable to the development of estimates of variables of interest for small and local areas in two major ways: directly from the production of tables from the microlevel data for the required characteristics, and indirectly from applying estimation techniques by combining other sources, such as sample surveys and administrative statistics to the population and housing census results.

3.280 Census data are typically aggregations of data for many individual small areas and may commonly be used to study large regions or entire nations. Data for small areas enable the user to obtain statistical information about any number of local areas of interest, in addition to showing variations among small areas in individual parts of the country. The increased capabilities of data processing systems greatly facilitates the utilization of census results for analysing the information for small areas, limited only by issues of confidentiality and collection design and statistical disclosure when cell entries in cross-tabulations become very small. For example, the analysis of whether programmes have improved educational attainment for women and girls at a regional level may be carried out by analysing data from the smallest administrative units so as to observe local variations and produce more accurate assessments of cause and effect.

3.281 Implementation of various national social and economic development programmes is a function of the state, province or lower levels of government in many countries. Results of population and housing censuses are useful for planning and monitoring development at the local area, small town level or small area. Small-area data are also important for private businesses in developing their distribution and marketing strategies. For example, information on housing demand from the population and housing census may be used by local authorities, local real estate companies, building and housing development contractors, and manufacturers of construction materials, among others.

3.282 Census data have been traditionally aggregated by various types of administrative units (for example, towns, villages, provinces and electoral units). In addition, other types of small areas are sometimes used in the census that are essentially statistical in nature (for example, census tracts and grid squares that do not change from census to census, and very small units such as city blocks or block faces). There have also been increasing demands for small-area data that cut across the local administrative boundaries. Population and housing censuses provide a powerful tool for assessing the impact of population on the environment, for example on drainage basins and on water resource management systems. The spatial units for such a study may combine a group of local administrative areas. In this situation the availability of census databases with mapping capability (see paragraphs 3.107–3.108) is of great importance.

3.283 Tabulations for small areas may be prepared on the basis of the resident population of each area or on the basis of the population present in each area at the time of the census. Tabulations relating to the resident population are produced for the apportionment of representation in legislative bodies, the measurement of internal migration, the computation of measures of fertility and mortality by place of residence, and the planning and administration of such services as schools and housing, which have relevance only to the resident population. Tabulations based on the population present in the area at the time of the census are useful where this population is considerably larger than the resident population and thus raises the demand for products and services above the level required by the resident population alone. The combined population and housing census may also be used to make comparisons of resident and daytime populations in specific localities, if an item on place of work is included in the population census. It is therefore important that users express their needs for particular data disseminated in a given format, based on the usual residence or place of enumeration, at an early stage of census preparations.

3.284 It was elaborated in Part One, Chapter I, how the population and housing census plays an essential role in the economic and social components of the national statistical system and also serves as a sampling frame for sample surveys. Another significant way in which the census results complement survey statistics is in small-area estimation, whereby models constructed from survey data are applied to census results for any specified geographic area. This estimation approach may be used for generating such indicators as employment, poverty and other economic indicators, for which measurement is required at the local area level.[66] The application of small-area estimation techniques to poverty measurement and mapping is an important extension of the use of census results.

(b) Spatial analysis

3.285 Spatial analysis on census data is key to understanding performances across sectors, and decision makers are able to relate more easily with information presented in maps than with data tables. The SDGs vision of leaving no one and no place behind requires, at a minimum, analyses and presentation of trends at the highest spatial resolution possible – to show existing (spatial) inequalities among places and people. Census data, owing to its high resolution is able to accurately showcase prevailing inequalities, and in turn inform decisions and actions to ensure equitable development. In addition, census data is crucial to producing data disaggregated by location (city, towns and suburbs, villages, rural areas etc), gender, income groups, persons with disabilities (among others) as required within the SDG framework. For example, understanding the share of different groups of urban populations who are within a proximal distance of access to public transport (SDG 11.2.1), open spaces (11.7.1), or the share of rural populations with access to all-season roads (SDG 9.1.1) or the computation of Ratio of land consumption rate to population growth rate (SDG 11.3.1) requires a clear analysis of census data to depict location of populations across the urban-rural continuum.

4. Mode of dissemination of outputs

3.286 As has already been mentioned, a census is not complete until the information collected is made available to potential users in a format suited to their needs (paragraph 3.240). Consequently, meeting the needs of data users means that the data producer should not only provide data products to the users, but should also provide them in formats that are suitable to the needs of the users. The information in the products may be included in published tables and reports for general distribution, produced as tables in unpublished form for limited distribution or stored in a database and supplied upon request, or disseminated online either as static or interactive products.

3.287 It should be noted, however, that regardless of mode, all dissemination is subject to issues of (a) quality assurance; (b) possible disclosure of information about identifiable respondents; and (c) copyright and ownership. In addition, the issue of cost recovery has become important to many statistical organizations. Each medium of dissemination has its advantages and limitations, and the choice of one or more of them depends on the context, and on the intended categories of users. In most instances, these methods complement each other and can provide effective ways to reach out to the public and private sectors.

[66] *Measuring the Economically Active in Population Censuses: A Handbook*, United Nations publication, Sales No. E.09.XVII.7.

3.288 When data are provided in electronic form, special attention should be given to providing users with easy means of data retrieval. The options for obtaining the relevant metadata and the data should be accessible in standard and contemporary formats.

a. Publication of printed tables and reports

3.289 Nealy all countries now disseminate their census results online, and printed publications have become a secondary choice for the dissemination of the main census results.

3.234. When printed publications are still used, the choice of how the actual printing is to be done entails in fact a trade-off involving quality, cost and speed. The best results can usually be obtained by sending the documents in computer-readable format to a professional printing plant. This will allow high-quality typesetting and the use of supporting colours. Alternatively, master printouts can be made in the census office and sent to the printer for cheaper duplication or offset printing. There are also affordable high-speed printing systems that can be directly controlled by the computers in the census office. Publications can also be distributed in electronic format (pdfs) as part of the NSO commitment to reducing paper waste.

3.290 Target dates for publication should be determined well in advance and processing and reproduction programmes should be planned accordingly. In addition to traditional methods of printing, there are various methods of reproduction available that are rapid, economical and legible, and these should be investigated.

b. Online dissemination

3.291 For most users, online is now the preferred channel of to access census outputs. NSOs are increasingly using cloud-based solutions for the storage, security, management, and dissemination of census data to users.

3.292 Online dissemination of all kinds of information, including statistical information, has increased with new innovative formats for displaying census data. The advantages of online dissemination are found primarily in terms of speed, flexibility and cost, and making results accessible to a wide range of data users.

3.293 Online dissemination of data was common well before the Internet gained prominence. The simplest option open to statistical organizations was bulletin board systems, now largely replaced by Internet and intranet websites. The same website could be used for both internal and broad community communication, with the granting of access rights in certain areas to privileged users only. Security measures, including passwords and callback procedures, can be used to exclude unauthorized users from reaching these areas. This is however risky, since resourceful hackers may find their way around the barriers and gain entrance to confidential information. Firewalls are hardware or software security systems that limit the exposure of a computer or network to malicious infiltration from an external location. The census office website is probably the first dissemination medium where Internet-connected users would look for census information. It is recommended that microdata should not be stored on a website in direct contact with the public. It is also recommended that a powerful firewall constitute a security layer between the website that is visible to the public and the working network of the census office. Websites of public administrations are under constant attack from hackers and very sophisticated security measures must be adopted when disseminating interactive census products on

the Internet. Cyber security, despite being an issue of a technical nature, has to be mandated, demanded and resourced by the highest levels of management of the census office.

3.294 Flexible table builder software provides the potential for users to access Census microdata and create their own multivariate output tables at different levels of geography. The software runs statistical disclosure control while the request is being processed to ensure protection of individual identities in detailed statistics, making sure that no individual can be picked out in the tables. It allows NSIs to release billions of anonymised census statistics far more quickly than ever before. A key design choice with profound implications for data security and confidentiality is whether such flexible table builders query the underlying microdata in the background, or a set of very detailed pre-curated queries. Direct microdata queries may lead to additional disclosure risks even with automatic disclosure control methods running in the software.

3.295 A website can be used not only to make information available as soon as it has been cleared, but also for other forms of communication with users. Possibilities include online ordering of publications and one or more receiving areas for questions that would be answered later through the same medium by appropriate specialists. One such area could be the census forum or “chat room”.

3.296 Internet websites may support “door” or “gateway” applications that allow users to run outside programs on the computer on which the Internet web server operates. Interactive access to census outputs can be offered to most types of databases and census products, including reports, publications, tables, maps and graphs. For example, there may be a database of aggregated census data for small areas or a microdata database that users can access in this way. When the required data are not readily available, users could run an on-the-spot query to obtain and retrieve results that satisfy their needs. This can be done by offering to Internet users census microdata samples and an interactive tabulation system. Users can then select records from these data sets that satisfy certain parameters and compute statistical information, such as two-dimensional cross-tabulations of either original or recoded variables. Program execution by users on the outside, however, raises important questions of cost, efficiency and confidentiality, which have to be resolved. For reasons of efficiency, it is recommended that information that is provided or likely to be heavily requested by users accessing the census website be made available in a static format, which is faster to download. Letting the user run data extraction on online databases, which would be a dynamic way of accessing the census information, is more resource consuming and should be the second choice for those users needing more detailed data than those available through static pages.

3.297 Other media such as social media are useful in disseminating census information targeted at different sectors of the population. More generalist media, such as the radio, television programmes, newspapers and press conferences, offer the possibility of reaching out to sectors of the population not otherwise reachable.

3.298 A hybrid solution for data dissemination that appears to combine the advantages of several approaches is one whereby the statistical or census organization makes basic data available to users on a computer-readable medium, usually through a website or optical media, while additional information may be provided by telephone or some other online protocols, such as file transfer protocol sites. This will usually take the form of a package that contains basic data, metadata and data browser software. The basic data may contain existing time series, report files and the like, as well as country and region maps that can be used to generate thematic maps with various indicators. Maps made available to general users need not ensure the same geographic detail as maps used for enumeration areas. Lighter

versions of maps at any subnational level may be provided to the general public, and more sophisticated and detailed ones to those fewer users who would actually need an increased level of detail. It is thus important that the website specify the instructions on how to contact officers responsible for special dissemination needs.

3.299 For some users, if the particular statistical information is not yet available on the physical distribution medium, special access may be granted, provided that adequate screening of their credentials and security checks are performed, to protected areas of the Internet site where up-to-date census information becomes available. Since opening up online resources to users has to be planned carefully and a clear policy established in advance (so that criteria for deciding whether or not to grant access are unambiguous), it is not recommended. Instead, provision of an online data tabulation system for expert end users is advised.

5. Confidentiality and privacy

3.300 According to principle 6 of the Fundamental Principles of Official Statistics, "Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes".[67] Maintaining data confidentiality is an indispensable element for maintaining the trust of respondents. If respondents believe or perceive that a national statistical or census office will not protect the confidentiality of their data, they are less likely to cooperate or provide accurate data. This in turn affects the accuracy and relevance of the statistics.

3.301 All the information stored in the census database allows the production of tables both for very small areas (such as enumeration areas or villages) and for all individual units in these areas. Therefore, when a census database is constructed, not only technical considerations but also the maintenance of confidentiality and the protection of individual privacy – which must be a primary consideration in designing the data collection and -data processing program – must be taken into account. Accordingly, microdata, such as name and local address, or the unique characteristics that permit the identification of individual respondents, must be removed from the database or otherwise altered.

3.302 The same care must be taken if a transcription of information from original questionnaires (that is to say, from a representative sample) is needed for use by qualified agencies and research institutes engaged in special studies beyond the purview of the regular census programme. Such needs have sharply decreased with the almost universal use of computer technology. However, when such a procedure is possible under the census law, individual privacy should be ensured and no exception should be authorized.

3.303 The ever-increasing demand from users for more data, especially microdata and at lower geographic levels, and also with more technological advancement for data linking, particularly over the Internet, have created more challenges for managing data confidentiality. As a result, national statistical and census offices should examine the data and make modifications, when necessary, prior to dissemination of the data. The objective of the modifications is to prevent identification of individual respondents, and also intentional or inadvertent disclosure of their personal information. This is

[67] http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx.

particularly the case when microdata is disseminated and when data are linked to location, such as with the use of GIS.

3.304 Statistical confidentiality methods range from traditional methods like cell suppression or recoding to more modern methods based on noise injection. Depending on the method, the implementation complexity also ranges from a few simple processing steps to elaborate statistical applications for entire databases[68]. The approaches used to limit disclosure are tailored according to the type of data and the product to be disseminated.

3.305 Different techniques are also employed depending upon the type of data product to be released (e.g. microdata files, tables, interactive data explorers, maps).[69] Disclosure methods for microdata include (i) suppression of direct identifiers, (ii) recoding or local suppression for quasi-identifiers, (iii) noise addition, and (iv) suppression of sensitive variables. More recently, differential privacy and synthetic data have been proposed as safer approaches. Regarding **tabular/query data**, more traditional methods aiming to mask confidential information (e.g. cell suppression, recoding) have been widely used over the past decades. However, these methods have meanwhile been found to suffer from additional loss of data security and utility in modern dissemination scenarios with very detailed tables possibly combined with various other data products. Emerging alternatives are predominantly based on adding some uncertainty to the outputs in a consistent manner (noise injection). **Interactive data explorers** are often based on flexible rules for users to submit custom table queries (table builders), where additional disclosure risks are present in implementations where highly flexible queries are evaluated directly from the microdata (as opposed to querying a curated fixed set of detailed tabulations stored in the background). Finally, **maps** are becoming a standard data product from censuses, more and more often with interactive functions. Specific additional disclosure risks may emerge from very detailed geographical outputs when these are not nested (geographic differencing – e.g. between very small administrative or enumeration units and grids).

3.306 The disclosure of confidential information in the context of the dissemination of census data integrated with geospatial information requires careful attention. Geocoded census data are becoming available at increasingly finer resolutions, increasing the risk of data disclosure. Geospatial information adds a new dimension to data with which an individual statistical unit can be more easily identified in combination with other information.

3.307 When an interactive mapping application is used, the tool should be configured in a way that users are not allowed to drill down to the spatial resolution level that unit record data might be disclosed. When an interactive mapping application is used, the tool should be configured in a way that users are not allowed to drill down to the spatial resolution level that unit record data might be disclosed. When microdata are released with geographic coordinates, geomasking techniques should be used to anonymizes their precise locations and prevents identity disclosure. Since aggregated data with spatial attributes provide coarsened location information of the respondents, their privacy can be protected by data suppression on the areas with low values. Geospatial products should be cross-checked with other dissemination products (e.g. tabular aggregates, anonymised micro datasets) before

[68] Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Schulte Nordholt, E., Spicer, K., de Wolf, P. (2012). Statistical Disclosure Control. Germany: Wiley.

[69] de Wolf, V.A. (2003). Issues in accessing and sharing confidential survey and social science data. *Data Science Journal* 2(17): 66–74.

release so that they do not breach confidentiality on their own as well as in combination with other outputs.

6. Metadata

3.308 In order to assist data users to better understand and interpret the data, it is important that there is adequate documentation providing a complete and clear description of the production process, including data sources, concepts, definitions and methods used. This information represents metadata that, it is recommended, should accompany all census products. Providing metadata promotes transparency and credibility of census results. Additionally, dissemination of census products with accompanying metadata ensures harmonization and comparability of census data with other data sets. International metadata standards (such as the the DDI Codebook standard) provide a convenient solution to ensure completeness of the metadata and to make metadata machine readable (which can help make the data more discoverable, visible, and reusable). SDMX is an example of an exchange and data warehouse tool.

a. Definition and content

3.309 Metadata comprise descriptive and structured information or documentation about data that informs users about the content, quality and condition of the data. In this context, metadata offer guidance on the proper usage or interpretation of data by providing information on the processes of production and describing the structure of data sets, thereby making it easier to retrieve, use or manage the data. Metadata constitute a standardized way of organizing data and can be categorized as follows: (a) reference metadata, which allow understanding and interpretation of the corresponding statistical data by describing the concepts, definitions, methodology and quality of data, production and dissemination processes, data access conditions, etc.; and (b) structural metadata, or “data about data”, which provide information about the structure of the data set and act as identifiers and descriptors of the data, making it possible to properly identify, retrieve, browse and further process the data.

b. Uses of metadata

3.310 The need for comprehensive and easily accessible metadata to better understand the statistical data being presented cannot be emphasized enough. Metadata are a key element of census dissemination, clarifying underlying concepts and definitions, and ensuring accurate interpretation of results. Whether utilized by individuals or electronic systems, metadata guide the accurate capture, reading, processing, interpretation, analysis, and presentation of statistical and geospatial information.

3.311 All tabulations should include the following metadata or references to where this information can be obtained: census questions; reasons why they are asked; conceptual definitions (census dictionary); geographic hierarchies used; changes since the previous census with regard to content, operational methods or geographic boundaries; and quality indicators such as coverage rates and item non-response. Data files must also be accompanied with metadata, including names and codes for common variables, personal files and household files. If a long-form sample is used in the census, metadata should also provide information on the sampling variability of the results. When the census tabulations include suppressed data cells due to small numbers, the metadata should also include a methodological note on the rules and methods of suppression. Metadata should be preserved for future reference. With the increased use of technology, properly designed metadata systems for machine-readability and web-based applications are recommended.

c. Geospatial metadata

3.312 Geospatially integrating demographic data necessitates the inclusion of additional metadata to the statistical record, preferably in the form of x- and y-coordinates but other forms of geographic information, through a "geocode" such as an address or locality can be used. Furthermore, additional metadata including details about the coordinate system, projection, datum, the precision of the geocoded locations, as well as data confidentiality measures, such as any spatial aggregation or masking techniques used to protect sensitive locations should be included to facilitate interoperability, appropriate use, and correct interpretation of the data.

3.313 Compared to statistical metadata, there is limited awareness and understanding in statistical organisations on the metadata associated with geospatial information and services. Geospatial information include various types, and this great variety as well as methodologies / technologies involved in producing the data adds challenges for statistical organisations to standardise geospatial metadata and systematically manage it.

3.314 Given the wide scope of geospatial metadata, it is important to first investigate crucial metadata elements (e.g. data type (point, line or polygon), time stamp, coordinate system) needed for different stages of production and determine a core metadata set and standards (e.g. ISO 19115, ISO 19119, GeoDCAT) to follow at the corporate level. After priorities are defined, a continuous improvement process could be put in place to gradually improve the scope covered by the metadata. It is important to have a corporate strategy in place to build a consistent metadata system to avoid compliance issues with existing metadata systems. As for statistical metadata, the geospatial metadata should be managed and updated continuously throughout the production process as the changes affect downstream tasks and influence the final outputs.

3.315 Alignment and harmonisation of geospatial metadata concepts with those of statistical metadata in existing metadata systems is critical, and there has been an increasing effort to connect statistical metadata with geospatial metadata (e.g. technical specification expanded for geospatial metadata in SDMX 3.022).

7. Promotion of, and training on, uses of census data

3.316 The main purpose of a census is to collect, process and disseminate information that will be used as the basis of informed, evidence-based decision-making. The benefits of this approach to decisions are not always apparent to users, especially in situations where other approaches may have been used in the past. It is therefore important to promote such uses of census results among users.

3.317 In other cases, users may be willing to use the information but require additional training to more fully understand the data. Such training may be usefully combined with training in statistical dissemination techniques or uses of more advanced data products. At a very basic level, some users may require training in such mundane issues as how to contact the national statistical office, or how to find the information they require within the systems of that office, or how to use the website and other electronic applications and tools. This training can be delivered via video tutorials, podcasts or e-learning platforms.

3.318 Whichever approach is taken to enhancing promotion and training in the use of statistical data, a number of strategic issues need to be addressed. These include:

a) Ensuring that the needs for training are identified early in the census planning process and that required funds are included in the census budget. In this regard it should be noted that in many cases the courses requested by users will be specific to those users; in such cases it may be desirable to request the user to provide funds to cover the marginal (or full) costs of the course.

b) The proposed courses or materials should be fully integrated into the overall census advocacy or training programme. It is essential that messages about the use of data fully reflect the messages given when initially advocating conducting the census or seeking public cooperation with and participation in the collection phase.

c) If the training facility is itself promoted properly, it is highly likely that the demand for training will far outstrip the ability of the statistical office to deliver it. In this case it will be necessary for the statistical office to have prepared transparent strategies that (a) identify those areas in which the statistical office wishes to participate (for example, dealing with lifeline clients, and topics on which the statistical office has particular knowledge or expertise); (b) establish partnerships with other bodies to provide training; (c) use approaches other than classroom training to provide learning-at-a-distance opportunities (for example, e-learning); and (d) have a pricing regime to cover costs where this is seen as desirable.

3.319 The list of target audiences and topics for such training must be determined by countries. It should be noted, however, that basic training in the use and interpretation of the results of one census is a very strong method of advocating support for future censuses. It is thus recommended that countries consider development of a basic course in (a) potential uses of census data; (b) how to access census data; (c) interpretation of census data at the broadest level, including the interpretation of its completeness and level of accuracy; and (d) spatial analysis. The target audience for such training should be key decision makers in the political and administrative hierarchy of the country. It should be outlined that the uses of census data at the local level (small areas) offer substantial potential for constructive use of census data; spatial distribution of population by age and sex, for example, provides an ideal framework for local officials to address the most pressing issues of their constituents, such as location of schools, utilities, service delivery and so forth.

3.320 A second group of key importance are members of the mass media, such as print, radio and television journalists. A focus on training such personalities is important because they can carry the message to many other people. This will assist in the general raising of awareness in the population at large, as well as in generating an awareness of the census among the governmental, academic and business users who may not have contact with the statistical office on a regular basis. Obviously, such training should be completely integrated with the overall public relations and advocacy work.

3.321 A third group to be targeted are schools, both students and teachers. A focus on training of teachers to use census data in the mathematics and geography curricula creates awareness among children about using statistics in decision-making and allows them to develop numerical skills using real data.

3.322 A fourth group to be targeted are geographers, with the aim of integrating census and survey data with GIS shapefiles in order to perform spatial statistical analysis. This training will enable specialists to better present statistics in space.

3.323 A fifth group to be targeted are non-users of census data. A number of stakeholders are unaware of how census data can be used in their area of work to make evidence-based decisions. Countries need to aim to increase the usefulness of census data by identifying non-users. User segmentation will be a valuable source of information to identify possible non-users to be targeted.

3.324 A sixth group to be targeted is the research community. The focus of the training and demonstrations will be on the application of various statistical techniques to census data. This will improve utilization of census data.

C. Census data dissemination: products and services

1. Provisional and final results

3.325 Some countries release provisional results very soon after enumeration is completed. Subject to change once the full data-processing and verification operations have been completed, they nevertheless provide a general picture of population trends. Provisional census results may be processed manually or by computer. For reasons of efficiency and quality, the use of computers is always preferable. The ability to verify data quality during the enumeration phase with the help of validation programs, quick indicator reports, data consistency reports, and tabulations greatly increases the confidence with which provisional results can be announced. Provisional results will normally cover information only on total population by sex and by major division. The number of households and housing units may also be derived easily from this exercise. The preliminary result of the census can be reported right after the end of census by utilizing the summary of household lists without individual data processing. This can be possible as the summary usually includes the total population, households and housing unit in each major division.

3.326 The final census results will be the output of the main tabulation programme. Tabulations may be based on all of the returns or on a sample. If some of the topics are collected on a sample basis only, proper weights will have to be applied in the tabulation stage to produce valid national estimates. In addition, the census office should be prepared to facilitate the production of tables requested by researchers and users (see paragraphs 3.392–3.398).

3.327 Since provisional and final results may differ (for example, the summaries on which provisional results were based might contain errors), it is important that data users be made aware of and warned about the possibility of such differences. Implications of using provisional population counts must be outlined. It is recommended that quality assurance processes be put in place to minimize variances between the provisional and final results.

3.328 The final census results must be published as soon as possible. Countries may aim to publish the basic, essential results within one year of enumeration. The use of technology may reduce the time between the release of the provisional and final results, which may over time render provisional results obsolete. The dissemination of the final census results must be part of a comprehensive dissemination strategy and plan. The schedule and description of upcoming releases of final results and products

should be made public early in the process to maintain interest by the public in the census (see also release calendar, paragraphs 2.114–2.118). The releases can be staggered, from simple, descriptive one-page summary fact sheets covering a country's major geographic divisions initially, to more comprehensive tabulations and descriptive reports later on.

2. Census reports

a. Basic reports

3.329 Every effort should be made to publish the principal results of a population census (such as those on age, sex and geographic distribution of the population) and of a housing census (such as a geographic distribution of sets of living quarters, households and population by type of living quarters) as soon as possible after the enumeration, otherwise their usefulness and the extent of their interest to the public will be diminished. With technological advancements, the time required for processing and tabulating results has been significantly reduced. As a result, collection restrictions, in terms of cost and accuracy of the data, have a greater relative weight in determining the number and complexity of the tabulations that can be produced and disseminated. The tabulation plan must respond to user needs.

3.330 The population and housing census tabulations presented and illustrated at the website of the United Nations 2020 World Population and Housing Census Programme are intended to provide, in tabular form, the most important census information needed as a basis for programmes of economic and social development and to be used for research purposes. They do not in any way represent all of the tabulations that a given country may publish and certainly not all of the tabulations that may eventually be prepared for special purposes. The tabulations do not take into account the form in which information may be entered into a database, which may be more detailed than that required for these illustrative census tabulations.

3.331 A major goal of these recommendations is to provide a set of tabulations that need to be produced at the lowest geographic level pertaining to the same point in time so that a country or area is able to meet its data needs for evidence-based socioeconomic development planning and monitoring. While the majority of national statistical authorities use a population and housing census as the single most comprehensive vehicle to collect these necessary statistics, others use sample surveys, registers of population and vital events, and other administrative sources or a combination of these methods to derive them.

3.332 Three categories of tabulations are described below: (a) basic or essential, (b) recommended, and (c) optimum tabulations.

Basic or essential tabulations

3.333 These are tabulations that are deemed of top priority for production by countries. They are also regarded as essential for countries in difficult circumstances, such as those that have emerged from a conflict or those that have not carried out a census in a long time, in terms of providing minimum statistics to meet their basic data needs.

3.334 The set of basic or essential tabulations on population and on housing characteristics on the website of the United Nations 2020 World Population and Housing Census Programme. The tabulations include elaborate classifications as well as relevant metadata for each of the tabulations.

Recommended tabulations

3.335 Recommended tabulations are those that are considered adequate for meeting the essential data needs for evidence-based planning, monitoring and implementation of national policies because of their perceived relevance at both the national and the international levels. These tabulations are also designed with the potential for producing statistics at the lowest geographic level and are expected to be produced by each country at least once in the 2020 census decade.

3.336 The recommended set of tabulations also includes the basic or essential tabulations discussed above. Schematic presentations of all tabulations are presented online at the United Nations 2020 World Population and Housing Census Programme.

3.337 Associated with the recommended tabulations are the core topics that go into their production. Core topics are therefore the main variables for the recommended tabulations. There are 31 core topics on population with 25 of them direct topics and six indirect (for a more detailed discussion of direct and indirect topics, see paragraph 4.19).

Comments AS: Update when the work of TT7 is finished.

3.338 As stated in paragraph 4.1, the aim of the recommended tabulations is to permit national and international comparability of data due to use of common concepts and definitions of the core topics. For each of the recommended tabulations, the core topics that it represents are listed as part of the metadata. Other metadata that are presented for each of the recommended tabulations include (a) the source of statistics, that is to say, whether from a traditional census, register-based census, survey or rolling survey; (b) the type of population count, that is to say, whether a de jure or de facto population or a combination of these; and (c) the definition of urban and rural areas used.

Optimum tabulations

3.339 The optimum set of tabulations includes the basic or essential and the recommended tabulations discussed above, as well as additional tabulations, and is designed to meet the needs of most of the users at the national and the international levels. This set can be viewed as being equivalent to the complete set of tabulations that could be generated from a population and housing census.

3.340 In order to avoid producing census tabulations that are overly voluminous or that contain a large number of empty cells, some countries may find it necessary to employ a more restricted geographic classification than that suggested in the illustrations. For example, basic facilities such as piped water or electricity may be almost completely lacking for large areas of some countries. Under these circumstances, tabulation of the relevant data for small geographic areas would not be appropriate. The geographic classification to be utilized needs to be carefully considered, taking into account the type of information being tabulated, it's probable frequency distribution and the uses to which the data are likely to be put. Privacy and confidentiality of individuals and households must always be protected (paragraphs 3.285–3.289, 3.262, 3.394).

3.341 Some countries may also collect data on additional topics in the census questionnaire to address specific concerns, for example, whether the birth of an individual is registered, the age a woman first marries, or vocational and technical skills. In other cases, detailed tabulations for special populations may be required for use in planning or evaluation of programmes. Tabulations for the non-core topics may be done after the basic tabulations are completed. This should be based on user needs. Consequently, consultations with user groups at both the national and local levels may be helpful in determining the most suitable tabulation plan and method of dissemination.

b. Thematic statistical or analytical reports

3.342 Many countries prepare different types of thematic or analytical reports. These reports must be planned and scheduled during the preparatory phase and published according to the release calendar in order to avoid outdated reports. The reports may range from volumes presenting extensive and detailed statistical tabulations, particularly cross-tabulations, to more analytical reports that combine tabular materials with some interpretative or analytical text. This latter group of reports might include, for example, volumes of regional analysis on such subjects as population or housing conditions of urban areas, major metropolitan areas or big cities, and regional distributions; locality reports on infrastructure; and comparisons of key social indicators such as education, living arrangements, housing conditions, sanitation and economic activities. Other such reports might include community profile analysis of, for example, indigenous populations, and profiles of specific population groups, such as families, children, youths, persons with disabilities and older persons. Reports on population growth and distribution that examine changes in the demographic characteristics of the country's population with breakdowns by two or three levels of administrative areas would be very useful. Such reports might focus on the growth, location and mobility of the population at the national and regional levels, and administrative areas. It should be pointed out that it is important that appropriate language is used to correspond to the target audience for each thematic report. It is recommended that multidisciplinary task teams be established, including line ministries and agencies, for the preparation of thematic and analytical reports in line with agreed guidelines. Partnership and external cooperation with academic institutions and other specialists in subject matter, which can facilitate such work and strengthen collaborations, may be sought whenever possible.

3.343 Thematic and analytical reports must be based on user needs and respond to a country's specific development needs and emerging issues. These reports can also be used to show time series and trend analyses of socioeconomic and demographic indicators and may combine census data with other data sources to provide a more comprehensive and current outlook.

c. Methodological reports

3.344 Other published reports may include the census methodology, encompassing, if applicable, sampling design and methodology and a census evaluation report, which may include estimates of census coverage and the methodology used for their preparation.

3.345 It is important that users of census products be provided on a timely basis with as much relevant information regarding the census as possible. A publication that contains information on all types of products that will be available following the census is very useful to users. A brief description of each product should be provided including the estimated timing of release, the level of geographic detail that each product carries and, for products released periodically, the frequency of release. In the case of

large census operations, several such documents tailored to the needs of different sets of users (for example, users in education, health or local government) may be useful.

3.346 Many countries publish a census dictionary, which contains comprehensive definitions of terms and concepts and detailed classifications used to present census outputs. Some countries also publish geographic classifications and codes and the definitions of areas used in the census and their relationships with the administrative areas. Explanations of user-defined areas for specific census tabulations and the type of format available (printed or electronic) may be provided.

d. Administrative report

3.347 One of the most important reports in the publication programme is the administrative report, which is a record of the entire census undertaking, including problems encountered and their solutions (see also paragraphs 3.468–3.472 on systematic recording and documentation of the census experience). The report may include the following topics: a brief history of the census in the country and the legal basis for conducting the census; budget requirements, expenditure and control; source and allotment of funding; census committees and their activities; stakeholder management; census organization and personnel structure; staff management; quality control procedures; census calendar; census cartographic work; development and design of the questionnaires; enumeration methodology; census promotion, publicity and communication; field organization; manual editing and coding; data-processing development and organization; data capture; computer editing and imputation procedure; capital, equipment and infrastructure management; computer hardware and software used; census evaluation; publication and data dissemination programme; and archiving. The census administrative report is very useful both for the users and for the census organization itself. The administrative report is an essential product for the planning of future censuses (see also paragraphs 3.471–3.472).

3.348 With developments in information technology, the census data files and publications have become increasingly available in electronic formats. A description of the procedure in the development of these data files may also be included in the procedural report. Consideration of a separate volume of the procedural report for the processing and dissemination phases may be considered to ensure the completion of the planning and field operations phases immediately after the census enumeration.

3. **Databases**

3.349 To maximize the utility and lifespan of census data, and as a complement to the standard production of tables, national statistical offices are encouraged to store the census data in various computerized database forms to better satisfy the full range of needs of internal and external data users. Census databases assist data users by providing easy access to a wide range of census data.

3.350 The establishment of such databases can enhance the dissemination of the census results as well as increase their usefulness by combining census data together with related information from other demographic enquiries in a common format. (An important special case is bringing together the data from prior censuses into a single database.) In addition, such databases can improve the coherence of the input and output processing systems.

3.351 Given the wide -ranging needs of users, there is no preferred one-size-fits-all approach to setting up a census or population database. Decisions range from the type of data to provide

(microdata, aggregated data or both) to whether to try and incorporate the new census results into existing database structures or establish new ones. In the latter is the case, whether the new database(s) will be exclusively in the form of a census database or constitute instead the nucleus of one or more population databases incorporating data from other sources should be considered. Consideration will also have to be given to such issues as identification of the different types of users, their information requirements, types of information to be stored in the database, sources of information, maintenance and update of information, processing of user queries, identification of the appropriate commercial software or, alternatively, whether it is feasible to develop such software, and selection of the appropriate hardware capable of supporting the current database and its anticipated growth.

3.352 Since building a census or population database requires careful planning and can be time consuming, such implementation should fit within the global statistical framework of the organization, and be seen as an ongoing process both complementing the data dissemination strategy and strengthening the statistical capacity of the organization.

a. Micro database

Comments ID: this section needs to be updated based on current technological devices.

3.353 Microdata (records of individual persons and households) collected in the census can be stored either in their raw form, or in their final edited form, or in a file that combines both raw and edited records. To limit problems of conservation, the data should be stored preferably on a medium of excellent reliability such as an M-DISC optical disk, or a solid state disk (SSD) drive or universal serial bus (USB) flash drive with backup copies. New technologies for mass storage will have evolved. Digital technologies for mass storage present two issues for census managers and technicians: (a) the issue of when it will be appropriate to adopt a new technology as the standard; and (b) that of the need to convert materials stored in older media to the new standard or otherwise provide accessibility to the older materials.

3.354 Cloud-based data storage is becoming increasingly pervasive, and while it offers several advantages, including scalability, enhanced accessibility, disaster recovery, automatic updates, and seamless integration with other services, it also presents significant challenges. Data security remains a primary concern, especially with sensitive census information. Additionally, issues related to data sovereignty, potential cost overruns, data transfer bottlenecks, and the complexity of managing cloud configurations can arise. While the cloud provides a flexible and modern storage solution, its application to storing census microdata should be carefully evaluated.

3.355 With technological advances in mass storage devices and media, it is now feasible to store the full census data file (one character per byte) as a single large rectangular file. After adding a data dictionary that describes the data format and a tabulation module, one obtains a set that could be described as a census database. The microdatabase requires a cross-tabulation program, which can be either part of the package or external. The software normally used for census tabulation still requires some prior training and may be confusing to inexperienced users. More intuitive tabulation software is available but may be either too slow in processing or too limited in its options to be fully satisfactory.

3.356 The organization of the microdatabase may take several formats, for example the software may allow for reorganizing the data in a transposed format (for example, one separate file per variable). This can substantially reduce the need for storage space and increase the speed of tabulations. However, establishing this kind of database is more complex, technically demanding and time consuming. There would be advantages in storing census microdata with standard commercial databases. This approach has the advantage that many users are already familiar with such software and so it is easier to find programmers and system analysts in the labour market. Even though the storage space required would be comparatively larger, today's market for mass storage has made available very large and fast hard disks at much cheaper prices than a few years ago and the hardware market seems to continue to follow this trend.

3.357 One of the main advantages of a microdatabase is that it permits the retrieval of data, at least in principle, at any level of detail. Since microdata could be used to obtain information on individual persons, families, households or family enterprises, privacy concerns must always be taken into consideration. In most countries, the use of census data to identify individuals is prohibited by law. Moreover, the long-term reputation of the national statistical authority may well be jeopardized if such disclosures occur. Full elaboration of principles and protocols for disseminating microdata is elaborated below in paragraphs 3.376–3.391.

3.358 As presented in this subsection, there are methods (such as sampling, introduction of random disturbances, recoding and aggregation) that can be used to make such microdata available while still protecting individuals' rights to privacy. All have in common the fact that they sacrifice some information in order to eliminate or greatly reduce the risk of disclosure. However, it is important that census organizations interested in disseminating microdata to outside users should take the appropriate precautions to protect privacy and confidentiality.

b. Macro database

3.359 Aggregated census data can be stored in many formats, either as the results for one census, as a database covering more than one demographic enquiry, or in a broad database of statistical and geospatial information. Whereas microdata are saved to allow aggregations to be made that were not programmed initially, macrodata are stored to preserve earlier aggregations, offering the general public readily usable information and avoiding redundancy for those who may find that the summary data they require have already been produced.

Publication equivalents

3.360 The simplest form of what could be called a database for macrodata is a straight copy of a publication on the website of the census office. A machine-readable publication-equivalent database may have the advantage of being less expensive to prepare than its hard copy counterpart. In addition, electronic or paper copies can be made quickly, copying only the part of the publication that is required. However, accessing this information requires a computer, potentially with specific software to interpret the census information.

Table-oriented databases

3.361 More advanced users, especially those engaged in spatial analysis, may require a macrodatabase that transcends the scope of the printed publication. They might like to be able to manipulate the tables and shapefiles in various ways in order to obtain views or results that represent their specific requirements more precisely. Associated graphing and thematic mapping capabilities may also be welcome. Several statistical offices have successfully filled this need. However, a major problem often encountered is that there is no generally accepted definition of what constitutes a statistical table and of the rules that should be followed when designing one.

3.362 In a controlled environment, such as that of a given census or national statistical organization, it is possible to standardize table definitions. The most common way is to design a basic layout having a number of attributes that together fully describe a table. Appropriate software will then give users access to a number of operations that process the table or several tables at the same time. Examples of such operations are reclassifying a variable (for example, from 1-year to 5-year age groups), eliminating a dimension from a multidimensional table or joining tables that have a dimension in common.

3.363 The availability of a standard table description language offers important advantages in exchanging tables as data-processing objects among national and international organizations. However, as mentioned before, some statistical tables are not easily pressed into the mould provided by formal descriptions. In this respect, it should be noted that statistical tables have little in common with the structures known as relational tables in popular database management systems.

3.364 Nevertheless, census offices should be aware of the potential offered by extensible markup language (XML). XML is not, as a matter of fact, a language itself but rather a metalanguage system designed to be used on the Internet. With XML, users can define their own “tags” to structure the information within a document. XML thus offers the potential of precisely describing all elements composing a statistical table: title, subtitle, units of measure, indicators, values, the time dimension and footnotes, and in short the metadata. One extension of XML is the Geography Markup Language (GML), which is a modelling language for expressing geographical features. Other solutions, such as EDI/EDIFACT (electronic data interchange for administration, commerce and transport) and SDMX (Statistical Data and Metadata eXchange), are a set of internationally agreed standards, directories and guidelines for the electronic interchange of structured data between independent, computerized information systems.

c. *Time series and indicator databases*

3.365 Databases can also cover more than one demographic enquiry, and census results can be integrated with various other data sets, including the results of earlier censuses. In developing databases that are aimed at serving a heterogeneous user community, the issue of a number of basic trade-offs will have to be addressed. For example, on the one hand, the number of variables should be kept as small as possible to make the database easy to use; on the other hand, it should be as comprehensive as possible to address the broadest possible requirements. A minimum data set of versatile indicators should consist of those variables that are useful for a wide range of applications and consistently available across space and time, and whose characteristics are clearly defined. In developing such a database, not only storage of the key indicators and variables themselves, but also the inclusion of some basic figures (absolute numbers or basic data) as a way of standardizing the basic statistical framework, is recommended.

3.366 With regard to time series and indicators, standardization and interoperability of data and metadata can be achieved via the Statistical Data and Metadata eXchange (SDMX) international standard. In SDMX, Data Structure Definitions (DSDs) define structures that describe the data for each statistical domain, specifying the dimensions, attributes and measures, together with associated code lists. Metadata is similarly structures using the Metadata Structure Definition. Data and metadata adhering to the SDMX encoding standard may be stored in several formats, including SDMX-ML (an extension of XML) and comma separated values. Data and metadata represented in SDMX and conforming to the relevant DSD can be easily read and processed by any system or tool that supports SDMX, as well as stored in traditional databases, file systems, or other storage solutions.

d. Graphing and mapping databases

3.367 Databases equipped with graphing and mapping capabilities, especially those powered by GIS, significantly enhance their utility and relevance. GIS databases are specialized systems that store, retrieve, manage, display, and analyse geographic and spatial information. Unlike traditional databases, GIS databases are designed to handle spatial data, which are associated with geographic locations, and can visualize these data in the form of maps.

3.368 A GIS database typically consists of layers, each representing a particular type of data, such as roads, rivers, population density or land use. Each layer contains features, such as particular roads or rivers, which, in turn, have attributes, which are the specific pieces of information associated with that feature, stored in tables. For example, a road feature may have attributes such as its name, type, length, and surface material. Users can query the database, overlaying multiple layers, including geocoded census data, to see relationships between different types of data and perform spatial analysis. Users can also generate custom maps that highlight trends or illustrate spatial relationships.

3.369 Many users want data for relatively small areas concerning such matters as home ownership, educational profiles and the labour market. While the database may be for one census, some historical information can be included to allow users to observe prevailing trends over time.

3.370 Both microdata and macrodata can be at the basis of these dissemination products. However, owing to disclosure problems as well as in order to increase processing speed, some form of prior aggregation is usually applied, for example by using summary data. Such summary data could also be combined with the general purpose graphing and mapping software. However, this would result in a reduction of the user community to those able to handle rather more complicated processing jobs. Making available a census database with tightly integrated graphing and mapping capabilities (which usually implies a tabulation function) is an excellent way to improve the effectiveness of census information dissemination.

4. **Geographic products**

a. GIS for census data dissemination

3.371 Geographic information systems (GIS) embody hardware and software configurations designed to support the capture, management, analysis and dissemination of spatially referenced data. Applied to census activities and outputs, such systems facilitate census cartography and data capture, and by linking population data (demographic, social and socioeconomic) to geographic areas, GIS provides very

powerful data management functionalities in allowing users to explore, analyse, describe and communicate population census information according to their own data and information demands.

3.372 In practical terms, such systems may range from simple desktop mapping facilities to complete GIS systems capable of solving complex planning and management problems, producing detailed georeferenced inventories and spatial statistical analysis. The ability to use space to integrate and manipulate data sets from heterogeneous sources can make its application relevant to planning and managing the census process itself. For example, GIS provides functions for the aerial interpolation of statistical data in cases where the boundaries of aerial units have changed between censuses.

3.373 Every census office should have organizational competency and staff with competences in geospatial analysis.

3.374 Statistical offices should develop GIS applications with population data and other georeferenced data from other sources for more advanced forms of spatial analysis. The role of the census office should be to supply census data at the right level and in the right format to users. Census offices provide vital information on current demographic conditions and future trends for policymakers in a range of sectors, such as health care, education, infrastructure planning, agriculture and natural resources management; and the provision of spatially referenced census databases is an essential prerequisite of the facilitation of the use of demographic data in these fields.

3.375 To achieve maximum efficiency gains, GIS applications should also be capable of generating additional geographic information beyond those used in the census, such as school and health districts, water and other biophysical catchment areas, and power and utility service units. These entities will have to be constructed from the smallest geographically identified units available in the census, such as census blocks, grid squares or enumeration areas. If, as is the case in most developing countries, enumeration areas are the smallest units, this will have important implications for the establishment of enumeration area boundaries. This requires close collaboration between national statistical organizations and national mapping and survey agencies on the one side, and school, health, water and power authorities on the other, when enumeration area boundaries are drawn or modified, to avoid potential discrepancies later on.

3.376 Apart from providing national statistical organizations with a very effective means/tools to disseminate and increase the utilization of census data, GIS, more than any other data/information management system, provide easy and user-friendly access to census data in user-relevant formats. This allows analysts and planners to undertake policy analysis, planning and research that can more readily identify thematic and geographic priority areas and thus contribute to evidence-based and better-informed policy and decision-making at different levels of geography. Some of the spatial statistical analysis includes clustering, Moran's spatial autocorrelation, Anselin outlier analysis, Getis-Ord hotspot analysis, ordinary least squares regression and geographically weighted regression. It allows governments to effectively monitor development progress across different sectors at village, municipality and subregional levels; it raises awareness about the importance of census and other socioeconomic data; and it increases the institutional capacity of national statistical offices and social and economic planning agencies to engage in more in-depth analyses of social and economic data and deliver information products in even more user-friendly formats.

b. Basic maps

3.377 Census offices should take advantage of emerging GIS technologies to make the census results more understandable and easier to use. The purpose of statistical maps is to present the results in terms of their geographic distribution and also to make it easier for the general public to understand census results than when information is presented in the form of only statistical tables. There is special interest in the current pattern of the geographic distribution and also in changes in the patterns that have occurred over time, particularly since the last census. Harmonization of the boundaries between the censuses is essential for comparability of data.

3.378 The provision of maps serves two purposes: first, census area identification maps locate and show the boundaries of all administrative areas and units for which data are reported in census publications; and second, statistical or thematic maps and graphs present the significant results of the census, thus allowing the general user to visualize the geographic distributions and patterns inherent in the data. Well-designed and attractive maps will interest the users of census reports, and may raise questions that send them to the statistical tables for further details.

3.379 There are three major types of area identification maps that are commonly used in most census publications in printed or GIS shapefile formats: (a) national maps showing the boundaries of the first- and second-order geographic divisions and of the major cities or metropolitan areas; (b) maps of each first-order division showing the boundaries of the second- and third-order divisions for which statistical tables will be prepared; and (c) urban or metropolitan and degree of urbanisation maps showing small sub-area boundaries as well as general streets, roads and rivers.

c. Thematic maps

3.380 A comprehensive map publication programme should be developed as part of the overall population and housing census publication programme in order that the needed resources may be provided within the budget at the initial planning stages. In addition to preparing maps for the census tables and reports, many countries have also found it useful to produce a population atlas (dynamic or static) as a census output. Collaboration with other departments and interested agencies might be sought to facilitate the production of an atlas volume. The atlas would include maps depicting population and housing characteristics, as well as other data influencing the growth, composition and distribution pattern of population and housing (see paragraphs 3.83–3.93).

3.381 As regards thematic maps, priority indicators for a population and housing census are total population and its distribution by sub-area, population density, urban and rural population by degree of urbanisation or metropolitan and non-metropolitan population, and changes in the population totals since the last census. Other important indicators include age, sex, fertility, mortality, migration, educational attainment, employment, household size, type of housing, ownership, number of rooms and sanitary facilities, with a growing demand also for data on communication (telephones, television, computers and Internet access), transport (vehicles), a broad range of household amenities, and recently also population-based development indicators such as household access to safe water, household waste management and multiple sources of household incomes, such as the incidence of remittances. This list of indicators is merely an illustration of the type of thematic maps individual countries might find useful to produce. Producing maps using the same set of indicators enables countries to meaningfully compare their results over time and with international or regional norms.

3.382 Maps are an invaluable aid in meaningfully comparing subnational results with national values or with other international and regional norms. Emerging technologies provide great flexibility in composing informative and visually appealing maps. Often several maps can be combined on a single page to show one indicator, for example, for the urban and the rural population by degree of urbanisation to provide a more nuanced representation in the three classes of the methodology (cities, towns and semi-dense areas, and rural areas). Also, combining maps and statistical charts is an effective means of presenting census information. Dashboards with different methods of presenting census result on maps, charts, tables and infographics can provide users with comprehensive but easy to understand message about presented phenomena.

3.383 By having associated graphing and mapping capabilities, databases will greatly increase their usefulness. Ideally users should be able to generate the graphs or maps for their own needs. Several census organizations have produced this kind of product, sometimes in cooperation with a commercial company. However, it is recommended that census offices develop mapping capabilities as a core competence for statistical production. Many users require small-area data concerning such matters as home ownership, educational profiles, the labour market, and so on. While the database may be for one census, some historical information can be included to allow users to observe prevailing trends over time. As with all time series-type data, it is important to maintain consistency in both definition and spatial representations to ensure comparability.

3.384 Both microdata and macrodata can be at the basis of these dissemination products. However, owing to the need to maintain confidentiality, and in order to increase processing speed, some form of prior aggregation is usually applied, for example by using summary data. Such summary data could also be combined with the general-purpose graphing and mapping software. Making available a census database with codes and names matching the GIS shapefiles with tightly integrated graphing and mapping capabilities (which usually implies a tabulation function) is an excellent way to improve the effectiveness of census information dissemination.

3.385 The following list presents some suggested topics for census maps. The list is not exhaustive: most topics that appear in the questionnaire as well as derived topics covered in Part Two can be presented in cartographic form. In some countries, special topics such as population distribution by ethnic or language group may be appropriate. Conversely, some of the listed maps present information on the same topic in somewhat different form, so that a statistical agency may wish to select the most suitable indicator for the needs of the country.

Illustrative list of thematic census maps

Population dynamics and distribution
Percentage population change during intercensal period(s)
Average annual growth rate
Population density (persons per square kilometre)
Urban population as percentage of total population
Distribution and size of major cities and towns
In-migration, out-migration and net migration rates
Born in country and foreign born
Born in another division of the country

Demographic characteristics

Sex ratio (males per 100 females), possibly by age groups
Percentage of population aged 0–14
Percentage of population aged 15–24
Percentage of population aged 15–64
Percentage of population aged 65 and over
Percentage female population in childbearing ages 15–49
Total dependency ratio (population aged 0–14, and 65 and over, as percentage of population aged 15–64)
Marital status
Birth rate
Total fertility rate
Mean age at first marriage
Death rate
Infant mortality rate
Under-five mortality rate
Life expectancy at birth
Percentage of persons with disabilities

Socioeconomic characteristics
Percentage of children not in primary school
Percentage of youth neither in education training nor employment (15-24) NEET percentage
Adult literacy rate (aged 15 and over)
Mean years of schooling (aged 25 and over)
Illiteracy rate of population aged 15 and over
Illiterate population aged 15 and over (total number)
Educational level of population aged 10 and over
Labour force as percentage of total population
Women's share of adult labour force
Percentage of labour force by economic sector, type of occupation and status in employment
Poverty mapping

Households and housing
Average number of persons per household
Percentage of households headed by women
Average number of dwelling rooms per household
Tenure status (owned, rented, and so forth)
Type of construction material
Percentage of population with access to adequate shelter
Percentage of population with access to safe water
Percentage of population with access to electricity
Percentage of population with access to sanitation
Percentage of population with access to health services

Comments PG: How best to capture the list of suggested outputs? These could be in an annex, along with suggested tabulations. This is becoming more input ant as countries move towards exclusively digital publication of census results. Suggestion to include a Minimum list of recommended tabulations and census outputs that all countries should produce.

3.386 Where appropriate, the indicators can be presented disaggregated by gender as well as by urban or rural area (for example, where the rural population is greater than about 25 per cent of the total population). If information about an indicator is also available from a previous census, it is often very informative to produce maps showing change over time or to present maps for both time periods.

3.387 Also, where appropriate, countries are encouraged to perform spatial statistical analysis by producing maps showing spatial clustering and outlier analysis of the variables of interest, such as electricity and water.

3.388 The development of locality (village, town, city, community, small area) population size maps by region is of particular value. These maps combine two types of information: locality population statistics and locality locations in each region or subnational area. More information can be presented on, for example, the locality location within the district and the region, habitable and non-habitable areas, densely populated localities, areas with no localities, and the proximity of localities. Locality population size maps can also be used as base maps for additional information on locality services and activities, and on location and distribution of localities without specific services, such as primary schools, dispensaries, piped water, and so forth.

d. Grid based census outputs

3.389 Grid-based[70] census outputs are a type of data presentation where census data is aggregated and organized within a grid system, usually a rectangular grid of equal-sized cells. This system replaces the traditional method of presenting census data based on administrative boundaries, offering several advantages and facilitating new applications.

3.390 Grid-based census outputs are an increasingly valuable tool for researchers, policymakers, and the public alike. They provide a powerful way to analyze, visualize, and understand census data, offering significant benefits for decision-making, planning, and resource allocation across various sectors. As technology advances and data integration techniques improve, we can expect grid-based outputs to play an even greater role in future data analysis and research.

3.391 Benefits of grid-based census outputs:

- **Spatial Analysis**: Grids provide a consistent and easily interpretable spatial framework for analyzing census data. This enables researchers and policymakers to identify spatial patterns and trends, visualize data variations across space, and conduct advanced geospatial analyses.
- **Data Integration**: Grids facilitate the integration of census data with other spatial datasets, such as environmental data, infrastructure maps, and land use data. This allows for a more comprehensive understanding of the relationships between population characteristics and other factors in the environment.
- **Data Accessibility and Transparency**: Grid-based outputs can be easily visualized and shared through maps and other interactive tools, making census data more accessible and transparent to the public.
- **Flexibility**: The grid system can be used at different scales, from small local areas to entire countries, making it a flexible tool for analyzing data at various levels of detail.

[70] The grid square system is a method of dividing a geographic area into smaller, regularly shaped units for statistical purposes. These units are typically squares defined by lines of latitude and longitude, forming a grid that covers the entire area of interest.

• **Standardization**: Grid-based outputs often use standardized grid systems, such as the Universal Transverse Mercator (UTM) grid, allowing for data comparison across different regions and time periods.

e. Innovative geospatial products

3.392 Other innovative geographic products for census data which utilize the emerging and still relatively new technologies of artificial intelligence (AI) and machine learning include:

(i) Smarter maps and geoportals

3.393 Smarter Maps is a term used to describe maps that are more interactive, intuitive, and informative than traditional maps. Smarter maps may use artificial intelligence (AI) and machine learning to improve the accuracy and usability of traditional maps. They are designed to help users visualize data in a more meaningful way and make better decisions based on that data. They can be used to analyze large datasets, identify trends and patterns, and create maps that are more informative and engaging.

(ii) GeoAI and GeoBlockchain

3.394 GeoAI is the application of AI technology to enhance the intelligence of GIS software which enables the analysis of e massive amounts of geospatial data like satellite imagery, whereas GeoBlockchain is a combination of blockchain technology and geographic information systems (GIS) that can be used to support the analysis of spatial-temporal trends. When combined, GeoAI and GeoBlockchain can be used to create a powerful platform for analysing geospatial data and generating insights that can drive better decision-making.

5. Interactive electronic outputs

3.395 It is of paramount importance that census data and information produced are widely disseminated and communicated, and that national statistical and census offices involved in this process have a pronounced customer, client and stakeholder focus. That means that national statistical and census offices should place more emphasis on providing a service and creating partnerships than on merely providing products, and should be guided by user-relevance and user-friendliness in all their operations, rather than by tradition in producing tables, graphs and reports that they have always produced.

3.396 Given its importance and widespread use, the web has emerged as the primary means of providing general access to census statistics. Many national statistical and census offices have utilized the Internet as the principal channel for data communication, positioning their websites into comprehensive census data repositories, enabling users to have access to all published data online. When developing new census products, and when reviewing existing products, national statistical and census offices should consider all ways and means of making census statistics accessible, giving high priority to dissemination on the web. The advantages of online dissemination are primarily in terms of speed, flexibility and cost, as well as in providing accessibility to census results to a wide range of data users and allowing the delivery of data to be tailored to the level of sophistication of the user.

3.397 Making a census database available online along with integrated searching, tabulating, graphing, mapping and analysis capabilities is an important way to improve the effectiveness of census data dissemination. Most national statistical and census offices provide user access to electronic databases and data files through their websites, satisfying the full range of needs of internal and external data users. This is a valuable service that allows users to access and display census data instantaneously and interactively. The establishment of such databases can enhance the dissemination of census results as well as increase their usefulness by allowing user interaction with census data. User interaction is a key concept whereby users are enabled and empowered to access and explore census data themselves, and build their own customized tables or spatially configure data outputs according to their own requirements.

3.398 Interactive web-based data tools provide a user-friendly entry point to the entire range of census outputs disseminated by national statistical and census offices. Basic design considerations of web-based interactive tools should factor issues such as identification of the different types of users, their information requirements and the types of information to be stored in the database. Content should be organized so that it can be easily understood and found, with an overview given to provide orienting information to users about the data that can be accessed using the interface. Context should always be provided to all outputs through metadata, links to related information, and cross-referencing to glossaries, publications and other background material.

3.399 In practical terms, interactive web-based data tools should enable users to access census data themselves, and build their own customized tables or spatially configure data outputs according to varying spatial requirements. The tools should allow users to visualize and explore the data in column charts, line graphs, maps and scatterplots. The table-building functionality should also have the ability to sort and order tabular results, and more easily select survey years and indicators. Tools should also be provided for downloading, conducting analysis or retrieval for use in other software. Design considerations to improve the interactivity of data interfaces should include the provision of user support. It is highly recommended to help users to anticipate, interpret and evaluate results. Support to users should include demonstrations and tutorials intended to describe how to perform the various functions related to the interactive web-based tools.

3.400 In addition to the Internet, interactive electronic products can also be accessed through other media, including CD-ROM, DVD and flash drive.

6. **Microdata dissemination**[71]

3.401 The provision of census microdata allows for further analysis and research, especially in the context of analysis of SDGs where further data dissemination by various characteristics is required.

a. Definition of microdata

3.402 In general, when statistical agencies or other data producers conduct surveys or censuses or collect administrative data, they gather information from each unit of observation. Such a unit can be a household, a person, a firm or enterprise, an agricultural holding, a school, a health facility or other. In

[71] The elaboration on the dissemination of microdata is largely based on the *Dissemination of Microdata Files: Principles, Procedures and Practices*, Olivier Duprez and Ernie Boyko, IHSN Working Paper No. 005, August 2010.

the context, microdata are the electronic data files containing the information about each unit of observation. Microdata are thus opposed to macrodata or aggregated data, which provide a summarized version of this information in the form of means, ratios, frequencies or other summary statistics.

3.403 Typically, microdata are organized in data files in which each line (or record) contains information about one unit of observation. This information is stored in variables. Variables can be of different types (for example, numerical or alphanumerical, discrete or continuous). They can be obtained directly from the respondent via a questionnaire or by observation or measurement (for example, by GPS positioning), or imputed or calculated.

3.404 In the context of the population and housing census, microdata refer to electronic files consisting of individual records on persons, households and housing units. More specifically, microdata would typically be organized in multiple files: one with records on households, another with records on individuals, and yet another with records on housing units.

b. Core principles for disseminating census microdata[72]

3.405 The United Nations Fundamental Principles of Official Statistics[73] provide unambiguous guidance in administering official statistics at national and international levels. A particular emphasis of these principles is on confidentiality of information collected for statistical purposes. In the context, the sixth principle, governing international statistical activities, states: "Individual data collected by statistical agencies for statistical compilation, whether or not they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes."[74]

3.406 Any principles for microdata access must be consistent with this recommended principle and the principles contained in the legislation pertaining to the national statistical authority. The following principles should be considered for managing the confidentiality of microdata:

Principle 1: Appropriate use of microdata

3.407 It is appropriate for microdata collected for official statistical purposes to be used for statistical analysis to support research as long as confidentiality is protected.

3.408 Making available microdata for research is not in contradiction with the sixth United Nations Fundamental Principle as long as it is not possible to identify data referring to an individual. Principle 1 does not constitute an obligation to provide microdata. The national statistical office should be the one to decide whether to provide microdata or not. There may be other concerns (for example, quality) that

[72] The elaboration of core principles for dissemination of microdata is quoted from *Managing Statistical Confidentiality and Microdata Access: Principles and Guidelines of Good Practice*, United Nations Economic Commission for Europe, Conference of European Statistician, United Nations publication, Sales No. E.07.II.E.7, United Nations, 2007.
[73] Presented at the United Nations Statistics Division website at: http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx.
[74] Ibid.

make it inappropriate to provide access to microdata. Or there may be specific persons or institutions to whom it would be inappropriate to provide microdata.

Principle 2: Microdata should only be made available for statistical purposes

3.409 For principle 2, a distinction has to be made between statistical or analytical uses and administrative uses. In the case of statistical or analytical use, the aim is to derive statistics that refer to a group (be it of persons or legal entities). In the case of administrative use, the aim is to derive information about a particular person or legal entity to make a decision that may bring benefit or harm to the individual. For example, some requests for data may be legal (a court order) but inconsistent with this principle. It is in the interest of public confidence in the official statistical system that these requests are refused. If the use of the microdata is incompatible with statistical or analytical purposes, then microdata access should not be provided. Ethics committees or a similar arrangement may assist in situations where there is uncertainty whether to provide access or not.

3.410 Researchers are accessing microdata for research purposes but to support this research they may need to compile statistical aggregations of various forms, compile statistical distributions, fit statistical models or analyse statistical differences between subpopulations. These uses would be consistent with statistical purposes. To the extent that this is how the microdata are being used, it could also be said to support research purposes.

Principle 3: Provision of microdata should be consistent with legal and other necessary arrangements that ensure that confidentiality of the released microdata is protected

3.411 With respect to principle 3, legal arrangements to protect confidentiality should be in place before any microdata are released. However, the legal arrangements have to be complemented with administrative and technical measures to regulate the access to microdata and to ensure that individual data cannot be disclosed. The existence and visibility of such arrangements (whether in law or supplementary regulations, ordinances, and so forth) are necessary to increase public confidence that microdata will be used appropriately. Legal arrangements are clearly preferable but in some countries, this may not be possible, and some other form of administrative arrangements should be put in place. The legal (or other arrangements) should also be cleared with the privacy authorities of countries where they exist before they are established by law. If such authorities do not exist, there may be non-governmental organizations that have a “watchdog” role on privacy matters. It would be sensible to get their support for any legal or other arrangements, or at least to address any serious concerns they might have. In some countries, authorizing legislation does not exist. At a minimum, release of microdata should be supported by some form of authority. However, an authorizing legislation is a preferable approach.

Principle 4: The procedures for researcher access to microdata, as well as the uses and users of microdata, should be transparent and publicly available

3.412 Principle 4 is important to increase public confidence that microdata is being used appropriately and to show that decisions about microdata release are taken on an objective basis. It is up to the national statistical authority to decide whether, how and to whom microdata can be released. But their decisions should be transparent. The website of the national statistical authority is an effective way of ensuring compliance and also for providing information on how to access research reports based on released microdata.

c. Microdata anonymization[75]

3.413 When disseminating census microdata files to the public, researchers or other agencies, the national statistical authority faces a conflicting mission. On the one hand, it aims to release microdata files supporting a wide range of statistical analyses; on the other, it must safeguard the confidentiality of respondents' identities. Processes aimed at the latter are referred to collectively as statistical disclosure control or anonymization.

3.414 A disclosure occurs when a person or organization recognizes or learns via released data something they did not know about another person. There are two types of disclosure risk: identity disclosure and attribute disclosure.[76] The former occurs when a respondent's identity is directly associated with a disseminated data record. This can occur easily when the data record includes variables unambiguously identifying the respondent – for instance, the respondent's name, address, passport or identification number, or telephone number. It is essential that such identifying variables be removed from any microdata files before dissemination. Attribute disclosure occurs when attribute values (or estimates thereof) in the disseminated data are associated with a particular respondent.

3.415 A combination of variables in a microdata record that can be applied to reidentify a respondent is referred to as a "key". Reidentification can occur (a) when a respondent is rare in the population with respect to a certain key value; and (b) when this key can be used to match a microdata file to other data files that might contain direct or other identifiers such as voter lists, land registers or school records (or even publically accessible Internet search engines).

3.416 The essential component of dissemination of census microdata files is avoiding both identity and attribute disclosures. In that respect, there is a need to strictly apply statistical disclosure control or anonymization techniques for census microdata files. The first key step in anonymizing a microdata file is to remove all direct identifiers – variables that unambiguously identify the respondent. Thereafter, a microdata file can be anonymized further by applying statistical disclosure control techniques.[77]

d. Protocols for dissemination of census microdata

3.417 Disseminating census microdata may be an unprecedented activity for the national statistical authority. In that context, there is a need to develop particular protocols that would comply with the essential principles for disseminating microdata, as described above, and would also spell out the necessary requirements and components of such dissemination. Hence, such a protocol would need to take into account the following components and requirements:[78]

---

[75] The elaboration of procedures for anonymizing microdata draws extensively from *Handbook on Statistical Disclosure Control, Version 1.0*, Anco Hundepool, Josep Dominho-Ferrer, Luisa Franconi, Sarah Giessing, Rainer Lenz, Jane Longhurst, Eric Schulte Nordholt, Giovanni Seri, Peter-Paul de Wolf, Centre of Excellence for Statistical Disclosure Control, Eurostat project, December 2006.

[76] *Measures of Disclosure Risk and Harm*, Diane Lambert, Journal of Official Statistics, Volume 9, Number 2, Statistics Sweden, 1993.

[77] For full elaboration of these techniques, refer to Chapter 7 of the *Dissemination of Microdata Files: Principles, Procedures and Practices*, Olivier Duprez and Ernie Boyko, IHSN Working Paper No. 005, August 2010.

[78] As presented in OECD Principles and Guidelines for Access to Research Data from Public Funding, Organisation for Economic Co-operation and Development, 2007.

Openness. Access on equal terms for the research community at the lowest possible cost, preferably at no more than the marginal cost of dissemination.

Flexibility. Taking into account the rapid and often unpredictable changes in information technologies, the characteristics of each research field and the diversity of research systems, legal systems and cultures of each member country.

Transparency. Information on research data and data-producing organizations, documentation on the data and specifications of conditions attached to the use of these data should be internationally available in a transparent way, ideally through the Internet.

Legal conformity. Data access arrangements should respect the legal rights and legitimate interests of all stakeholders.

Protection of intellectual property. Data access arrangements should consider the applicability of copyright or of other intellectual property laws that may be relevant to publicly funded research databases.

Formal responsibility. Access arrangements should promote explicit, formal institutional practices, such as the development of rules and regulations, regarding the responsibilities of the various parties involved in data-related activities. These practices should pertain to authorship, producer credits, ownership, dissemination, usage restrictions, financial arrangements, ethical rules, licensing terms, liability and sustainable archiving.

Professionalism. Institutional arrangements for the management of research data should be based on the relevant professional standards and values embodied in the codes of conduct of the scientific communities involved.

Interoperability. Technological and semantic interoperability is a key consideration in enabling and promoting international and interdisciplinary access to and use of research data.

Quality. The value and utility of research data depend, to a large extent, on the quality of the data. Data managers, and data collection organizations, should pay particular attention to ensuring compliance with explicit quality standards.

Security. Specific attention should be devoted to supporting the use of techniques and instruments to guarantee the integrity and security of … data.

Efficiency. One of the central goals of promoting data access and sharing is to improve the overall efficiency of publicly funded (data collection) to avoid the expensive and unnecessary duplication of data collection efforts.

Accountability. The performance of data access arrangements should be subject to periodic evaluation by user groups, responsible institutions and … funding agencies.

Sustainability. Due consideration should be given to the sustainability of access to publicly funded research data as a key element of the research infrastructure. This means taking administrative

responsibility for the measures to guarantee permanent access to data that have been determined to require long-term retention.

e. Dissemination of population and housing census microdata in practice

3.418 It can be expected that the 2030 round of population and housing censuses, taking into account the contemporary development of processing technology and power, brings considerable pressure on national statistical authorities in respect of dissemination of population and housing census microdata. In the context of implementing the principles and protocols described in the preceding paragraphs, national statistical authorities would need to ensure such dissemination in at least two different settings.

3.419 For the purpose of public dissemination, either online or on electronic media, only a representative sample of the individual records should be made available after ensuring the confidentiality or non-disclosure of individual information as elaborated above. The size of the sample would depend on the capacity and resources of the national statistical or census office.

3.420 As for the access to the complete population and housing census master file, it should be made available to users using the model of the data enclave. This is a facility equipped with computers not linked to the Internet or an external network and from which no information can be downloaded via USB ports, CD, DVD or other drives. Users interested in accessing a data enclave will not necessarily have access to the full census data set – only to the particular data subset they require. They will be asked to complete an application form demonstrating a legitimate need to access these data to fulfil a stated statistical or research purpose and be briefed on the legal responsibility and repercussions related to maintaining the confidentiality of individual information. The outputs generated need to be scrutinized by way of a full disclosure review before release and they can contain only aggregates.

7. **Customized products**

3.421 The increasing activity in the field of economic and social planning and the attention of such planning to subnational areas are placing new demands on statistical information in general and on population and housing censuses in particular. There is an increasing need for tabulations and mapping not only by major and minor civil divisions and by other units of analysis such as metropolitan areas but even, beyond these, by small local areas.

3.422 Therefore, it is useful to establish an “on request” service for users who require aggregates not available through other means. This will be especially relevant in situations where outsiders cannot obtain census microdatabases. In essence, the service would require that users provide the census office with the details of the tables or other aggregates requested so that the census office could fulfil the request, normally against payment of a certain compensation fee. Offering and promoting this service, especially online, would place the statistical service in a more desirable proactive position, rather than a static one, and could be a strong catalyst for closer cooperation with census product users.

3.423 The cost of such special purpose tabulations, which require computer programming, could be high, especially for academic institutions and other users who do not have access to a large budget. Some statistical organizations allow the users to do the necessary work using a user-friendly kind of software. A clearly written manual is required to guide the users in using the software, including the

contents of the census data dictionary and other relevant information. The resulting tables are checked for any possible breach of confidentiality, in particular table cells with very small values.

3.424 Many census organizations provide services for special requests for census products, such as thematic databases, tables, and graphic and mapping outputs that can be designed for small, medium and large businesses, communities or special interest groups. These services are normally provided to meet the increasing demand of data users for a wide range of applications, such as monitoring trends, analysing unmet needs, identifying market potential, segmenting markets, identifying service areas and priority zones, determining optimum site locations, and designing and advertising new products and services. Each category of products should also be made available on various media (namely, paper, disk, online) for dissemination according to user requirements.

3.425 Once the databases are created and have served the policy needs, they can serve other data users if they have market value. Since the national statistical organization is normally the only source of many geographic databases related to census data applications, market demand for these products is increasing, particularly in the geographic and population-related areas. In such cases, census products could be governed by a licence. The licence permits the users to use the product without transferring ownership, since the ownership remains with the government agency. Either of two different licensing arrangements may be applied. The first is offered to organizations that use the data for their own needs and the other is offered to organizations that redistribute data or provide analytical services using census data to other persons or organizations for a fee.

3.426 Customized services of data on computer media are differentiated in terms of the forms of the data. Census products may be distributed in their original form, with or without other related information, or they can be distributed after making certain value-added modifications to meet the need of the users. Examples of such value-added activities include converting the data into another format (for use by other software packages), making the data more useful by creating subsets of the original data sets, merging the data from other sources and bundling with software. It is encouraged to disseminate census content in formats that comply with the requirements assessed by the Open Knowledge Foundation.[79] According to these requirements, the openness of the contents is assessed within the Open Knowledge conceptual framework and concerns the possibility to reuse, revise, remix and redistribute data. In cases where copyright laws protect census data ownership, some royalty fees and data usage fees may be charged to the distributors to ensure a minimum return. However, if prices are too high such charges can also be a barrier to the use of the census data.

3.427 Some countries may assist their users by merging selected variables with the GIS shapefiles as a customized product. This has proven to be beneficial for schoolchildren.

8. General interest products and special audience products

3.428 Information generated by a census is by definition of use to a wide range of users with a variety of expertise. With the increase of demand for census products, efforts must be made by census offices to produce a variety of products for various stakeholder groups, including special interest groups. In order to address the various stakeholder needs, census offices must segment stakeholder into groups to better know, understand and respond to their needs. This will form the basis of the various census

[79] https://okfn.org/opendata/

products. In response to the stakeholder needs, census offices may prepare special audience products for key variables such as policy summary reports; thematic and analytical reports; key findings reports; fact sheets; posters, brochures and flyers; basic reports; detailed tables and spreadsheets; articles; and video and social media products. Special audience analytical or thematic reports incorporate a high level of very sound analysis undertaken by staff who have a solid foundation in analytical techniques as well as the topic being analysed. In some cases, countries may undertake the analysis in collaboration with academic institutions or other specialists. Criteria used in establishing the topics chosen will have to be set by the country concerned, and may include particularly interesting facts shown by the census data (perhaps confirming or rebutting conventional theories; confronting census data with material from other sources; or responding to issues raised by the public during user consultations of the collection).

3.429 It should be noted that the following products can only be effective in encouraging the use of census information if they are prepared in a timely and professional manner. This will require specialist skills from people familiar with communicating to the target audiences. These resources are expensive and countries are required to adequately plan and budget for these products and campaigns.

a. Posters

3.430 One of the most common ways to disseminate census information consists of publishing posters highlighting key facts such as: How many are we? Where do we live? and summarizing a profile for the major civil divisions of a country. Posters might also be prepared addressing issues relevant to special population segments: teenagers, adults, indigenous populations, seniors and women's groups.

3.235. Since the objective of a poster is to catch the eye at a distance, relatively few facts should be presented in a way so that the key message is immediately visible. Posters can be greatly enhanced by the addition of well-designed graphs, infographics and maps to increase the readability and comprehensibility of the key message. Posters and banners are short-term communication products and should be used to communicate key findings.

b. Brochures and flyers

3.431 Professionally designed brochures and flyers are another way to disseminate basic census data. These brochures should be written in a very easy and comprehensible language indicating the demographic profile of the country illustrated with suitable graphics and explanatory material. In some countries these brochures might be addressed to specific issues on population. They are particularly suitable for preparation as promotional materials for people attending events and exhibitions, such as the launch of more traditional materials, or for inclusion on display racks in libraries of government offices.

c. Videos, sketches, theatre and online videos

3.432 In order to create a better understanding among certain interest groups, the use of other communication media are recommended, including videos, sketches, theatre and online videos. In order to promote the story behind the numbers and increasing the use of census data, graphics such as charts or maps could be included on videotape, CD or DVD, or memory sticks. These might indicate how census data can assist policymakers, planners and people in general to understand their societies and how census data can assist in identifying the main problems and assist with evaluation of solutions.

d. Instructional materials

3.433 Instructional materials in an easy-to-understand form can be prepared for the general public, indicating the advantages and limitations of census data. Such material can often form the basis of information campaigns as part of the advocacy material for the next census.

3.434 A particular implementation of instructional materials can be the preparation of a kit for use in schools. Not only will this provide high-quality information for the students but, by including exposure to the use of statistical materials in the school process, it will encourage the use of evidence-based analysis throughout society. It should be noted that professional assistance should be sought in ensuring that these materials follow sound educational practices and can be accommodated within the appropriate curriculum.

e. Social media

3.435 Internet-based social media have become an indispensable tool in disseminating information and marketing statistical products. Various social media platforms have been successfully used by countries as part of the dissemination of their census results. Interacting with followers and users on these platforms provides the census organization with an opportunity to disseminate information, build relationships with established and new users, and engage the public on a regular basis.

3.436 Social media platforms such as Facebook, Twitter and online video sites can be used to post all census-related advertisements and other videos.

3.437 During the build-up to the census phase, information about the upcoming census, informing the public about what to expect when fieldworkers visit their homes and also the importance of conducting a survey of this magnitude, can be disseminated using social media.

D. Census data utilization

1. **General uses of population and housing censuses**

3.438 Population censuses are traditionally used for public and private sector policymaking, planning, administrative and research purposes at national and subnational level. One of the most basic of the administrative uses of census data is in the demarcation of constituencies and the allocation of representation on governing bodies, as well as allocation of resources within a country. Certain aspects of the legal or administrative status of territorial divisions may also depend on the size of their populations. Housing censuses are used to develop benchmark housing statistics and to formulate housing policy and programmes, and in the private sector to assist in site selection for industrial, retail and service facilities, as well as for the commercial development of residential housing. (see section III 'Uses of population and housing censuses)

3.439 Information on the size, distribution and characteristics of a country's population is essential to describing and assessing its economic, social and demographic circumstances and to developing sound policies and programmes aimed at fostering the welfare of a country and its population. The population and housing census, by providing comparable basic statistics for a country as a whole and for each

administrative unit and locality therein, can make an important contribution to the overall planning process and the management of national development. The availability of information at the lowest levels of administrative unit is valuable for the management and evaluation of such programmes as education and literacy, employment and human resources, reproductive health and family planning, housing and environment, maternal and child health, rural development, transportation and highway planning, urbanization and welfare. Population and housing censuses are also unique sources of data for producing relevant social indicators to monitor the impact of these government policies and programmes (see paragraphs 3.430–3.432).

a. Uses of population censuses

3.440 The uses of population census results and the associated tabulations described in this volume are listed according to the topics presented in paragraph 4.21.

3.441 The total population and its distribution among major and minor territorial divisions and localities are frequently a fundamental requirement of the census because these results are used for determining the apportionment of representation in legislative bodies, for administrative purposes and for planning the location of economic and social facilities. Internal migration, one of the major sources of population change, frequently affects the trends in population distribution. Data on internal and international migration, together with fertility and mortality, are needed to prepare population estimates for planning purposes and for determining policies on migration and assessing their effectiveness.

3.442 The household, a basic socioeconomic unit in all countries, is often central to the study of social and economic development. The number, size and structure of households and changes in the rate of household formation are useful for planning and for developing special policies formulated for selected groups of the population, such as children and youth, the elderly and persons with disabilities. Therefore, the distribution of individuals within households is used to determine the living arrangements of families, the patterns of family structure observed, the time when new families are formed and changes in family structure due to death, divorce, migration or the departure of children to form their own households. The relationship among household members can be used to determine family structure and the existence of households composed, partially or completely, of unrelated persons.

3.443 Traditionally defined demographic, social and ethnocultural characteristics collected from the population census include sex, age, marital status, religion, language, and nationality, ethnicity and indigenous identification. Sex and age are fundamental to the majority of the characteristics collected in the census. Census data provide more data than any other single source on differences between women and men.

3.444 Depending on national circumstances, cultural diversity may be indicated by language spoken in the home or community, religion and national or ethnic group. For countries that are not homogeneous in terms of one or more of these variables, linguistic, religious and national or ethnic groups provide the basic information for a quantitative assessment of the relative size and age-sex distribution of this diversity.

3.445 Although census data on fertility and mortality cannot serve as a substitute for reliable birth and death statistics from civil registrations, they are particularly valuable for countries where birth or death registration is lacking, or incomplete and vital statistics are therefore unavailable. Even in countries with complete registration of these events, the population census is useful as a supplement to satisfactory registration data because the fertility questions provide data for calculating lifetime fertility of the female population or cohort fertility.

3.446 Education has historically been one of the key factors determining the quality of life, and interest in education continues today in most countries of the world, with emphasis on improving access to education and the quality of education, as well as broadening the scope of basic education.[80] Education is also considered a major tool in closing the gap between women and men in respect of socioeconomic opportunities. Benchmark data obtained from national population censuses will therefore be of considerable importance towards fulfilling this objective. Census data reveal the disparity in educational opportunities between the sexes, age cohorts or generations, urban–rural populations and so forth, and provide important indications of the capacity of the nation for economic and social development. They furnish material for the comparison of the present educational attainment of the adult population with the present and anticipated requirements of educated human resources for various types of economic activities. Such a comparison may serve as a guide both for national policy in terms of the development of the educational system, and for the planning of the economic development programmes that it will be feasible to undertake in view of human resource requirements.

3.447 Census information on the economic characteristics of the population focuses on enumerating the labour force so as to provide benchmark data for current studies of employment and labour underutilization, in particular unemployment and the potential labour force. It provides information on the growth, composition and distribution of the labour force for use in policy formulation and the appraisal of human resource utilization. Economic data from censuses can also provide some input into statistics on the distribution of income, consumption and accumulation of households, and participation in agriculture and non-agricultural activities. Furthermore, the data on the labour force may give an approximate indication of the number of workers who are responsible for the support of dependants.

3.448 Statistics obtained from different sources (for example, labour force surveys, agriculture surveys, establishment surveys and administrative records) rely on the census for sampling frames, and the use of common concepts in the different sources helps in securing comparability when multiple sources for changing patterns of economic activity are being relied upon.

3.449 As interest in the movement of people across national boundaries – in other words, international migration – has grown steadily among countries, and census items and tabulations relative to international migration have grown in importance. Such tabulations are designed to assess the impact of migration on receiving countries, to understand patterns of diversity and develop programmes for the adaptation of migrants to new countries, and to serve as a source of information on emigration from sending countries.

3.450 The census is also an important source of data on persons with disabilities. Census data help to monitor the social and living conditions of persons with disabilities in terms of school attendance,

[80] Education for All Summit of Nine High-Population Countries, New Delhi, 12–16 December 1993: Final Report (Paris, UNESCO, 1994).

educational attainment, employment, marital status and living arrangements. The data also provide a basis for developing policies to meet the needs of persons with disabilities and for evaluating the effectiveness of these policies.

3.451 The census is also an important source of information on household poverty and living conditions. Population censuses also provided historical data for historians, and researchers studying social, economic, and demographic changes over time.

b. Uses of housing censuses

3.452 The primary uses of information from housing censuses include development of a basis for planning housing and human settlement programmes and policies, public and private sector studies of urban and other non-agricultural land use, evaluation of the adequacy of housing stock and assessment of the need and market for new housing, and studies of the living conditions of the homeless and those living in temporary or substandard housing. Information collected on the number of sets, type and characteristics of living quarters and their occupants is crucial from the point of view of monitoring housing conditions and needs of the population. Combined with the information collected by regular annual statistical programmes on housing construction, data from the housing census provide a basis for identifying national, regional and local housing patterns, which are needed for the development of a rational housing market aimed at stimulating various types of housing construction. The type and quality of shelter in which people are housed, that is to say, the space, degree of crowding, facilities, surroundings and available transport, affect their economic activity, health, social intercourse and general outlook. The supply, characteristics and costs of housing are therefore subjects for which the housing census is an important source of information. Housing censuses also provide an address frame for population censuses.

2. **Analysis of the results**

3.453 In order to ensure the fullest possible utilization of census results by national and local governmental authorities, by academic researchers and by others, it is advisable to draw up a comprehensive and coordinated programme of analytical studies, phased over a period of several years. This will help allocate effort and resources in such a way as to ensure that important policy needs are adequately met, undue duplication of research effort is avoided and priorities are observed as far as possible. In these studies, the data of the current census should be examined not only by themselves but also as complemented by relevant data from other sources and from earlier censuses, in order to obtain a broader context, improve the estimates and establish trends.

3.454 The analytical studies to be included in such a programme will vary according to the needs and circumstances of the country. The programme may include descriptive summaries of results, policy-oriented analyses of census results and detailed analytical studies of one or more aspects of the demographic and social situation of the country. Some of these studies may be undertaken by the census organization itself, but others, particularly the more time-consuming studies, can most effectively be carried out in cooperation with specialists in different subjects having experience in in-depth analytical studies from universities or other research centres. In any case, it is desirable to invite specialists from other governmental offices and experts outside government to take part in drawing up this programme of studies, and it is natural that they would play an important part in the execution of various parts of the analytical programme.

3.455 One important aspect to be considered in establishing a programme of analysis is the possible use of census results in achieving the goals and objectives of population, human settlements or similar policies and strategies at the national and local level, and in applying available resources effectively towards the improvement of conditions in these fields. For this purpose, it will be necessary to analyse population and housing census results within the framework provided by other available information so as to achieve an integrated approach to the solutions of population, human settlements and similar problems.

3.456 A permanent census office should be the central repository of all census results; it would thus be equipped with the information needed for comparative studies, which will indicate long-term trends in the phenomena investigated. However, to facilitate the fullest possible use of census results by others, subsidiary depositories should be established that serve different substantive or geographic groups of users.

3.457 Aside from the studies that are part of the overall census programme, additional analyses carried out on their own initiative by research organizations, universities or other experts should be encouraged.

3. Intercensal and post-census population estimates

3.458 Intercensal population estimates are estimates made for the years between two censuses and are produced once the census has been completed. They replace the postcensal estimates produced prior to the completion of the census.

M*ore details to be added.*

4. Cross-cutting and emerging social issues

3.459 Reflecting the concerns and priorities among countries around the world, the United Nations convened a series of global conferences: on children, education, environment and climate change, human rights, population, sustainable development, women and human settlements. Each of these conferences recognized the importance of adequate information for formulating policy and monitoring progress in the achievement of conference goals and called on countries and international organizations to develop and improve the requisite statistics and indicators. These recommendations are reflected for example in the Vienna Declaration and Programme of Action of the World Conference on Human Rights;[81] the Programme of Action of the International Conference on Population and Development;[82] the Copenhagen Declaration on Social Development and the Programme of Action of the World Summit

[81] A/CONF.157/24 (Part I), Chapter III.

[82] *Report of the International Conference on Population and Development, Cairo, 5–13 September 1994* (United Nations publication, Sales No. E.95.XIII.18), Chapter I, Resolution 1, Annex.

for Social Development;[83] the Platform for Action[84] adopted by the Fourth World Conference on Women; the Post-2015 Development Agenda Declaration adopted by the General Assembly; and the Climate Change Declaration. The programmes of action adopted by these international conferences targeted many interrelated areas of concern, and called for improved statistics to monitor progress. In deciding which social groups merit monitoring in regard to measuring the disadvantages suffered by particular groups of people, each country should determine which groups within it need special attention. Some of the common factors leading to social disadvantage are gender, age, physical or mental impairment, race and creed. The disadvantaged are not necessarily small in number; they may constitute the majority of the population.[85]

3.460 To meet the need for statistics on gender, many activities have been undertaken during the last two decades at the national and international levels to improve concepts, definitions and classifications for collection of statistics related to women and men. In the present publication, the importance of the population and housing census as a data source has often been stressed. The population and housing census is also the principal or sometimes the only comprehensive national data source with respect to meeting the need for statistics on children, youths, the elderly and the disabled in the development of policies and programmes at the national and international levels. Therefore, it is important that countries identify data requirements concerning various population groups of particular interest when planning their censuses and ensure that the definitions and classification to be followed in censuses are appropriate and also consistent with those in use for the entire population.

3.461 Furthermore, the census tabulation plan should ensure in advance the inclusion of all relevant details about special population groups and a range of cross classifications for each group, with a view to analysing its social and economic conditions. Concepts and methods for the census and the tabulation plan should be reviewed with users concerned with statistics for each special population group. In the case of some groups, for example persons with disabilities, a special set of questions is required to identify members of the group. In the case of others, standard questions, for example on age, are sufficient to identify groups such as children, youths and the elderly. In both cases, most variables needed for cross-tabulations are already provided for in the international recommendations and many national censuses. In the census operations, however, attention will often need to be given to improvement of coverage, quality-of-data issues and avoidance of stereotypic treatment. The present section deals with gender, a few special population categories such as children and youths, the elderly, and persons with disabilities, so as to assist in the preparing of detailed tabulations and databases according to international standards.

a. Statistics on gender

[83] *Report of the World Summit for Social Development, Copenhagen, 6–12 March 1995* (United Nations publication, Sales No. E.96.IV.8), Chapter I, Resolution 1, Annexes I and II.
[84] *Report of the Fourth World Conference on Women, Beijing, 4–15 September 1995* (United Nations publication, Sales No. E.96.IV.13), Chapter I, Resolution I, Annex II.
[85] *Note by the Secretary-General transmitting the report of the Expert Group on the Statistical Implications of Recent Major United Nations Conferences* (E/CN.3/AC.1/1996/R.4), annex, paragraphs 68–69. Presented to the Working Group on International Statistical Programmes and Coordination at its eighteenth session, New York, 16–19 April 1996.

3.462 The global conferences on women have contributed to an increased awareness of the importance of statistics not only on women but, more broadly, on gender issues. For example, in developing census plans in a number of countries, efforts have been made to review and assess the adequacy of statistics for understanding the diversity of both women's and men's lives. It is now recognized that biases in statistics extend, in the case of women, to their economic roles, and in the case of men, to their roles in the family as husband and father and their roles in the household. Improvement of statistics and statistical methods related to gender should be an important priority in all stages of work on the census, in planning, data collection, analysis and dissemination, and in all topics.

3.463 In addition to the more general problems of the quality of census data, two other types of problem that apply particularly to women and stem from gender based stereotypes and biases have been noted. Similarly, the notion that only men can be heads of the household affects the way questions have been designed and asked in censuses. Such stereotypes also affect the way respondents reply to the questions. A common problem, for example, is classifying women automatically as homemakers without asking whether they perform any work for pay or profit, even as a part-time or secondary activity.

3.464 Another problem relates to biases in the collection, processing, compilation and presentation of data. For example, when census tabulations are prepared for the employed by occupation, they may be prepared either for males only or for both sexes, but only on the assumption that information on the occupational pattern of women is not of much use. Even when tabulations of the employed by occupation are disaggregated by sex, main gender differences in occupations may be missed if the occupation data are presented only at the two-digit level of the classification. Rather, special tabulations showing, for example, the 10 or 20 detailed occupational groups with the highest concentrations of women or men would be needed to render visible gender-based occupational segregation.

3.465 During the past few decades, considerable effort has been devoted, on the one hand, to reviewing such bias and its impact on statistics concerning the situation of women and, on the other hand, to improving the concepts and methods involved in the collection of data in censuses and surveys. Related improvements in the revised System of National Accounts and latest recommendations concerning statistics of work, employment and labour underutilization adopted by the International Conference of Labour Statisticians are also of importance to the population census. They are intended to overcome the above-mentioned conceptual deficiencies and to identify and provide measures for all productive activities (that is, forms of work) performed by women and men, whether paid or unpaid. Similarly, efforts at the national level have been focused, for example, on eliminating biases in concepts, classifications and definitions of head of the household.

3.466 Important statistical series and measures on the status of women can be readily obtained based on the topics in paragraph 4.21 and recommended tabulations for preparation from censuses. Furthermore, in the case of most topics, the primary unit of classification is the individual and therefore a vast array of indicators may be obtained by devising appropriate additional cross-classifications for the female and male populations separately. For an illustration of census topics and tabulations that are useful for developing comprehensive statistics on women, see "Statistics and indicators on women and men",[86] *Handbook for the Development of National Statistical Data Bases on Women and*

[86] Available at: http://unstats.un.org/unsd/demographic/products/indwm/indwm2.htm.

Development,[87] *Integrating a Gender Perspective into Statistics*[88] *Methodological Guidelines for the Gender Analysis of National Population and Housing Census Data.*[89] The household and family status classifications presented in paragraph 4.148 are appropriate for analysing the living situation of women and men, with specific reference to single mothers and fathers and elderly women and men living alone.

3.467 It should be emphasized that while all data collected at the individual level can be presented by sex, this is not always done. Cross-classifications by sex tend to be suppressed when cross-tabulations become complex with multiple-variable tables. In order to satisfy one basic condition for gender statistics, which is that all statistics on individuals should be presented by sex, sex should be considered the overriding variable in all tables, irrespective of the medium of storage or dissemination. This disaggregation by gender should be provided in all publications, databases and computer printouts of census tables on individuals.

3.468 Another important consideration is to broaden the target of dissemination and use of census data by popularizing the statistics that are published. One approach to achieving this wide outreach is to present statistics in the form of charts and simplified tables, with a simple and clear interpretation of the data. Countries planning to issue an analytical report might wish to consider using such innovative techniques and formats. The analytical publication could cover the main census topics or alternatively a few areas that are especially important to understanding the relative position of women and men in the country.

b. Statistics on children and youths

3.469 Extensive data on children and youths are available in censuses but may need improvements in terms of coverage and quality of information on specific characteristics, and on their presentation.

3.470 For statistical purposes, “children” are defined as persons under 15 years of age, and “youths” are defined as those aged 15–24. However, it is useful to further divide these special groups by 5-year age groups (or nationally, by groups of specific school ages) because of the rapid changes in characteristics in this age range, such as in school attendance, marital status and activity status. Also, because of differences by sex in the age at marriage, family or household status and entry into the labour market, data should be classified not only by age but also by sex. To this end, the distribution by single years of age and sex is useful. If single-year age distribution is not feasible for young children under age 5, it would be desirable to distinguish between those under 1 year of age (infants) and those aged 1–4. For youths aged 15–19, it would be desirable to distinguish between those 15–17 years of age and those 18–19 years of age, or to have a distinction corresponding to the age below which the country considers an individual to be a minor.

3.471 For the purpose of developing statistics on children, the principal topics in census recommendations include (a) sex, (b) age, (c) school attendance (for schoolage children) and (d) relationship to head or other reference member of the household.

[87] *Social Statistics and Indicators, No. 6* (United Nations publication, Sales No. E.89.XVII.9).
[88] United Nations ST/ESA/STAT/SER.F/111, Sales No. 13.XVII.9.
[89] UNFPA, 2014.

3.472 Given the priority on the girl child highlighted by the World Summit for Children (1990), the International Conference on Population and Development (1994) and the Fourth World Conference on Women (1995), special attention needs to be given to improving and disseminating statistics on children. Of particular concern is the situation of the girl child with respect to school attendance, mortality, early marriage and other issues of importance. A basic problem with statistics on the girl child is that data on children ever born and children surviving tend not to be disaggregated by sex at either the questionnaire design or the tabulation stage. These data are used for indirect estimates of child mortality.

3.473 The principal topics of investigation identified for children apply also to youths, with the following additions: (a) marital status, (b) literacy, (c) educational attainment, (d) economic activity status, (e) number of children born alive and (f) age at marriage.

3.474 Some of the useful statistics and measures can be readily compiled based on the abovementioned topics, while any additional indicators can also be obtained based on more detailed cross classifications using the existing recommended census topics or tabulations.

c. Statistics on older persons

3.475 For older persons also, extensive data are available in population and housing censuses but may need detailed age-sex classification, as described below.

3.476 Older persons are defined by the United Nations as all persons aged 60 years and over. For purposes of classification, depending on the national situation, it is useful to tabulate data by 5year age groups up to age 100, instead of including them in the single broad age category 60 and over.

3.477 For the purpose of developing statistics and indicators on older persons, the principal topics in census recommendations include (a) sex, (b) age, (c) marital status, (d) economic activity status, (e) income, (f) household (or family) composition, (g) type of living quarters and (h) institutional population.

3.478 The statistics needed for studies of older persons are disparate, depending as they do on national policies and circumstances. Internationally, no illustrative list of indicators is available to ensure appropriate tabulations from the censuses.

d. Statistics on persons with disabilities

3.479 The census can provide a valuable source of information on the frequency and distribution of disability in the population, at national, regional and local levels. Experience shows that although an increasing number of countries ask questions about disability in their censuses, the presentation of disability data has often been limited to tabulations showing the number of specific severe disabilities present in the population. Unfortunately, cross-tabulations with other characteristics are not usually made.

3.480 A great deal of work on concepts, classifications and development of statistics on persons with disabilities has been undertaken in recent years, particularly through the work of the Washington Group on Disability Statistics,[90] and increasing numbers of countries are including disability as a topic in their censuses. For the second time, recommendations on including disability questions in a population census are included in these guidelines. A brief treatment of this topic is given below to highlight issues involved in preparing detailed census tabulations on persons with disabilities.

3.481 For the purpose of developing statistics on the situation of persons with disabilities the principal topics in census recommendations that would be necessary for the assessment of equalization of opportunities include (a) sex, (b) age, (c) place of residence, (d) type of household, (e) marital status, (f) educational attainment and school attendance, (g) labour force status, (h) status of employment, (i) industry and (j) occupation.

3.482 Not only should the tabulation plan for the disability data include prevalence rates by sex and age, but it is also very important that tabulations comparing persons with and without disabilities on key social and economic characteristics be presented. Tabulations based on the topics listed above provide information on prevalence of disability and on the situation of persons with disabilities. In addition, tabulations should be presented in a way that facilitates comparisons of persons with disabilities and those without.

e. Ethnocultural characteristics

3.483 Receiving information about the ethnic composition of the population allows deeper study of the ethnic background of a country's population, especially with respect to indigenous populations, international migrants and other specific groups of population (for example nomads).

3.484 There are some difficulties in collecting this information since some population groups may name their ethnic identification based on its local meaning and in order to correctly allocate these persons to their particular ethnic group it is necessary to compile a list of ethnic groups, sub-ethnic groups and local definitions of small ethnic population groups. This will allow for obtaining accurate data about the ethnic composition of population. It would also be useful if scientists and specialists in the field of ethnography, as well as organizations dealing with indigenous people, would be involved in creating such a list.

3.485 In order to obtain comprehensive information characterizing ethnic composition of population, it would be useful to tabulate data by (a) sex, (b) age, (c) place of living, (d) marital status, (e) birth, (f) death, (g) education, (h) economic activity, (i) employment status, (j) industry, (k) occupation, and (l) type and size of household.

3.486 It is important to obtain comprehensive information on indigenous populations in order to have statistics on the number as well as the demographic and socioeconomic structure of the given population group. These data would be valuable information to support the development of programmes for social support of indigenous peoples.

[90] For more information on the Washington Group on Disability Statistics, see: http://www.cdc.gov/nchs/citygroup.htm.

3.487 Statistics about the ethnic composition of international migrants together with information about country of birth and citizenship will help to more precisely determine the flows and volume of international migration.

3.488 Population censuses are also the sources of information about religious identification of the population. It would be useful to obtain this information by (a) sex, (b) age, (c) ethnic group, (d) place of living, and (d) place of birth. This information would be useful to study distribution of religious affiliations.

3.489 Information about knowledge of languages is widely used. Countries find it useful to study the official language of the country as well as mother tongues or some other languages. In any case it would be useful to have this information by (a) sex, (b) age, (c) ethnic group, (d) place of living, and (d) place of birth.

3.490 Information about knowledge of the official language of the country would be very useful for studying the integration of international migrants and may be used, for example, for development of programmes to learn the language.

3.491 Information about knowledge of the mother tongue of indigenous populations is very important. This information could allow obtaining statistics of "indigenous" languages and would be very useful for development programmes to support the development of those languages.

f. Statistics on poverty

3.492 The census data can provide a valuable source of information on conditions of life of households as a proxy measure of poverty to complement quantitative survey data. Census data provides a quantitative approach to measuring poverty.

3.493 In some cases, countries may compile multiple deprivation indices using census data.

5. **Development indicators**

3.494 Indicators are required by countries to track the progress of various developmental goals and as such efforts must be made by census offices to produce relevant indicators to meet this need. In the 2020 round of population and housing censuses, many countries produced indicators based on the Sustainable Development Goals as was recommended. The type of indicators required to meet international and national reporting requirements need to be taken into account early in the planning phase of the census.

3.495 It should be emphasized that both global and national reporting and monitoring require reliable and comparable national data for the compilation of indicators. In this regard, it is of paramount importance that countries have the statistical capacity to produce, analyse and disseminate the requisite data for these indicators. The availability of reliable statistics and the capacity of governments to systematically measure and monitor indicators is a critical success factor for the achievement of development goals. The lack of statistical capabilities in some developing countries makes it difficult to

obtain good and reliable data. Many countries do not have a sustainable, coherent programme of household surveys, or administrative data systems that can be used to produce basic statistics routinely. Where basic statistical systems are not available, the global monitoring may have to rely on national and international estimates of widely varying quality and reliability. This may lead to misjudgements regarding progress and may undermine the effectiveness of policy interventions at national and subnational levels.

3.496 The 2030 Agenda for Sustainable Development, adopted by all United Nations Member States in 2015, provides a shared blueprint for peace and prosperity for people and the planet, now and into the future. At its heart are the 17 Sustainable Development Goals (SDGs), which are an urgent call for action by all countries - developed and developing - in a global partnership. They recognize that ending poverty and other deprivations must go hand-in-hand with strategies that improve health and education, reduce inequality, and spur economic growth – all while tackling climate change and working to preserve our oceans and forests.

3.497 The 17 SDGs are divided into 169 accompanying targets and over 230 indicators, many of which are subject to disaggregation, which poses a major challenge to the data generation capabilities of National Statistical Offices. In some cases, SDG indicators can be directly generated from the population and housing census which allows for disaggregations of the indicators at the sub-national level and for sub-populations that are harder to report on through sample surveys. (Reference and further detail - Technical Report: Measuring Sustainable Development Goals Indicators through Population and Housing Censuses and Civil Registration and Vital Statistics Data – UNSD Oct 2022).

Comments FWZ: Since these are recommendations for censuses of the 2030 round, SDGs will be a thing of the past. We could either reference 'Our Common Agenda' instead, or perhaps turn this into something more general that doesn't pertain to any specific agenda, but rather to international development frameworks in general.

XI. Documentation of census experience

3.498 The cumulative experience of past censuses in a country is very useful in the preparation of a new census. Because of the lapse of time between censuses (generally 10 years) and the likelihood that experienced staff may leave the census office, it is essential that there is a comprehensive record of how the census was planned, organized and conducted.

3.499 The census office should, therefore, plan for and implement a knowledge management system to assemble complete records on plans, activities, and decisions taken during the entire census operation. This would entail documentation and archiving of information related to plans and their implementation, as well as problems encountered and how they were resolved at each stage of the census cycle. It is recommended that documentation of census experience be undertaken at each stage of the census operation and not be left until the end of the census process. This would include plans, decisions and activities related to preparatory activities, the methodology of the census, fieldwork or

other data collection activity, data processing, cost and implementation of the census budget, and evaluation of performance of each of these activities. Examples of items to track or monitor include implementation of activities, time taken to complete an activity, resources used and cost. All these should be assessed against set goals so that changes to plans can be recorded, including information on what changed and why. Tracking and systematically recording the census experience should also take into account risks encountered and how these risks were managed. For more information see Part Two, Chapter XIV on "Quality assurance".

3.500 Use of knowledge management tools and techniques is thus beneficial for preserving institutional memory in a codified way so that lessons learned from the past may be used for better management of future census planning and execution. Records in the system should be arranged in such a way that information on each aspect of the census operation is found easily.

3.501 Systematic recording of census experience is not an end in itself. It is recommended that every country prepare and, if possible, publish an administrative and methodological report, as a census "historical memory", based on information that has been recorded in the knowledge management system (see paragraphs 3.325–3.326 on administrative report). Depending on the methodology of the census, the administrative and methodological report should contain information on the manner in which the census was planned, organized and conducted, as well as important methodological and other problems encountered at various stages of the programme. As appropriate, the report should provide specimens of the census questionnaires and forms, instructions for the enumeration, and detailed information on the cost of the census and on the implementation of the census budget, as well as points to be considered in future censuses.

3.502 The structure of the report could be similar to the structure of the project plan. It is important that the report be as comprehensive as possible, covering all stages and aspects of census planning and operations, including fieldwork, processing, analysis, dissemination and evaluation. It is important to note that while such a report would be based on items and information in the knowledge management system, it may not necessarily contain detailed descriptions of all the processes or information as some may be for internal use only. This report would both assist the users of the census results in appraising and interpreting the data and facilitate the proper planning of future data collection programmes, including population and housing censuses.

XII. Archiving individual records

A. Purpose of archiving individuals records

3.503 The focus of the following elaboration of archiving is on census individual records irrespective of the format – paper questionnaires or electronic records. The discussion on the complete process and characteristics of archiving all census documents is presented in paragraphs 3.468–3.472 on systematic recording and documentation of census experiences.

3.504 Individual census records refer to either census paper questionnaires in the case these were used for collecting information from the population or to the digital records on each enumerated person and household if the data collection did not involve paper questionnaires, including their direct

identifiers, such as name, address and so forth. In the case of bimodal or multimodal data collection, that is, a combination of paper and non-paper questionnaires as in the case of using Internet forms and mail out/mail back paper questionnaires, the resulting collection of individual records would also be a combination of paper and digital recordings.

3.505 In the case of digital records the accompanying documentation becomes an indispensable part of the archiving process. As a number of variables in the digital record are presented as codes, it is necessary to archive all the codebooks and all the other documentation, such as the data collection instrument, that are needed for fully unlocking the value of each of the variables in the record. While this accompanying documentation is also valuable in the case or archiving paper questionnaires, these are by nature visual, thus requiring only reading skills and the knowledge of the language initially used for filling them to grasp the content, as long as they are in good physical shape.

3.506 The essential purpose of archiving individual census records is to keep them safe for future use, primarily in the domain of genealogical research and longitudinal social and anthropological studies, as well as for use by historians and demographers. The release of archived individual census records is subject to the passage of time as per the census legislation and usually encompasses many decades, thus ensuring that the use of individual information would not endanger the confidentiality and the privacy of the respondents.

3.507 Consequently, the importance of providing detailed guidance on the process of archiving individual census records in the census legislation cannot be overstated. These provisions provide the legal basis for maintaining the archives and procedures related to the release of archived records. The time lag between the data collection and the release of the archived records needs to be clearly indicated – it varies from 72 years (United States of America) to 92 years (Canada). In some cases, the original questionnaires are only temporary stored before being fully disposed of, as in India, one year before the next census takes place.

B. Procedures for archiving

3.508 Archiving a vast amount of records represents a considerable challenge in all circumstances. In the case of individual census records it may be compounded by the sheer number and format. However, in all cases the national statistical authority needs to develop an institutional strategy for archiving, based on three components: organizational infrastructure, technological infrastructure and resources.

3.509 **Organizational infrastructure** refers to the arrangements that need to be put in place within the national statistical office in such a manner as to ensure the efficiency of the archiving and eventual retrieval process. In most cases it is a centralized unit within the office that is put in charge of the archiving, maintenance, secure storage and eventual release of individual records. Once the time lapse mandated by the law for the release of records expires, the actual release to the public is usually implemented by dispatching relevant batches to the libraries covering parts of the country to which the records refer and to a central national library.

3.510 **Technological infrastructure** refers to the actual technology used for archiving. In the contemporary circumstances, storing huge numbers of paper questionnaires would prove not to be cost-effective, as it would require a significant physically secure structure, regulated temperature and humidity, and a host of other requirements, including protection from fire hazards, floods and extreme weather events. Consequently, in most cases the actual questionnaires are scanned and images of them

stored in various electronic storage devices. As an example, the individual census schedules from the 1940 population and housing census of the United States are available from a website[91] in the form of scanned images.

3.511 The technological infrastructure does not refer only to the actual technology used in the archiving process – it also consists of a series of protocols for archiving and establishing cross-references that enable successful retrieval of records. In the example of the 1940 United States census, all the records were archived based on the enumeration district, as the first-level threshold, then county, then district and so forth. Therefore, the technology should be built around a well-developed archiving scheme that enables efficient identification and retrieval of the records.

3.512 In the case of archiving digital records, contemporary technology provides a vast array of possible solutions – however, it also requires a well thought-over archiving scheme that needs to ensure efficient storage and retrieval, as well as access to the accompanying metadata and documentation.

3.513 **Resources** for archiving need to be taken into account at the early stages of planning for the census, in the context of the technological and organizational infrastructure. In assessing the volume of the necessary funds it is necessary to adopt a strategic, long-term approach, as the archiving, maintaining and releasing would essentially constitute a perpetual activity as long as censuses are part of the national statistical systems: there would always be a need to prepare either for the next round of release of records or for archiving the newly acquired one.

C. Archiving individual records and microdata

3.514 Individual census records as described above for archiving purposes differ from the census microdata in a most significant manner: they retain the direct identifiers – name, address, enumeration area – as these very identifiers represent the essential information for genealogical, anthropological, historical and longitudinal social studies. In the case of microdata, these identifiers would be removed, as well as any other that can directly or indirectly identify the respondent. Microdata are defined as electronic records pertaining to each unit of observation; in the case of the population and housing censuses, it would be individuals, housing units and households. This information is stored in variables. Variables can be of different types (for example, numerical or alphanumerical, discrete or continuous). They can be obtained directly from the respondent via a questionnaire or by observation or measurement (for example, by GPS positioning) or imputed or calculated.[92]

3.515 It is expected that the use of anonymized microdata becomes a standard feature of census data dissemination for the 2020 round of censuses. Consequently, paragraphs 3.376–3.388 of these *Principles and Recommendations* present a comprehensive elaboration of principles and protocols for dissemination of microdata files.

[91] See http://1940census.archives.gov/.

[92] *Dissemination of Microdata Files: Principles, Procedures and Practices*, Olivier Dupriez and Ernie Boyko, IHSN Working Paper No. 005, August 2010.

XIII. Overall evaluation of the census

A. Importance of evaluations

3.516 A population and housing census consists of a complex series of interrelated steps, and constitutes perhaps the single most extensive, complicated and expensive statistical operation that a country undertakes. Like any other project, the census needs to be evaluated to ensure that the effort and investment of resources have been worthwhile.[101] Evaluation of the census for coverage and also to assess the quality of the data is covered in Part Three, Chapter IX, and quality assurance for the census in Part Two, Chapter XIV. This section is concerned with the evaluation processes and procedures of the census operation.

3.517 Evaluation of the census is important for a variety of reasons, among which is to provide lessons learned from one census to the other. For this reason, evaluation is generally regarded as the last stage of the census cycle or the first step in the next census cycle. Evaluation assesses the effectiveness of operations, systems and processes and their likely impact on data quality. In this connection, it is particularly important to assess various aspects of the census operation, particularly when changes have been introduced compared to the previous census. Assessments provide valuable information on strengths and weaknesses of past operational procedures, which should be carefully reviewed prior to the development of the next census.

3.518 In addition to evaluation of the conduct and operational elements of the census, it is valuable to evaluate the use and value of census data and products. Evaluation of the outputs of the census can be conducted through the conduct of user surveys to gain subjective feedback or through looking at metrics around product use, for example number of visits to web products or the number of publications requested or purchased. Evaluation of census products can be used to measure whether the data and the selected topics are sufficiently accessible, timely, relevant, accurate, coherent, trusted and interpretable to meet user requirements. The outcomes of the evaluation might provide information that can be immediately acted upon and remedied in the intercensal period, and should certainly feed into the preparations for the subsequent census.

3.519 In recent years, countries have introduced new methodologies and technologies in conducting censuses in order to reduce census costs and also to improve the quality and timeliness of census data. Plans should be made for thorough evaluations on the effectiveness of the new technologies so additional improvements can be implemented for the next cycle. There is use of new methodologies based on administrative registers and combinations of sources to produce census information in all phases of the census. Plans should be developed to assess the quality of potential administrative sources. The added value of including these new administrative sources and the potential improvements in quality should be assessed alongside the risks associated with these new sources (for example, the status of the data supplier and the level of commitment of the data suppliers to the dimensions of quality). Other changes relate to the use of multimodal enumeration methods as well as substantial outsourcing of some aspects of census operations. Each new methodology, technology, and administrative source needs to undergo quality assessment so that a comprehensive risk profile and risk management plan can be established for the census design.

3.520 Given the current state of affairs, evaluation of processes of census operations becomes even more warranted in order to assess how well the innovations have worked. Furthermore, evaluations are necessary so as to provide lessons learned not just for the countries concerned but also for those that want to adopt similar processes for their future censuses.

B. Planning for the evaluation

3.521 Evaluation of census processes should not be undertaken on an ad hoc basis. The evaluation programme should be included in the overall census plan and be appropriately budgeted for. Lack of proper budgeting and planning for activities that come after enumeration, such as the post-enumeration survey, has in some cases led to financial shortages with negative consequences for the activities concerned. The success of the programme of evaluation depends to a large extent on setting, early enough, clear and unambiguous objectives to be served by the evaluation. Having clear objectives would help to design the best way to undertake the evaluation.

3.522 Ascertaining the objectives of the evaluation encompasses other aspects as well. It is important to establish the scope of the evaluation in terms of intended objectives, keeping in mind that the wider the scope the more complex the evaluation is likely to be. It should be noted also that evaluation of census processes and procedures could last several years and may cover different aspects of the census operation, as necessary, thereby resulting in a series of reports. These reports should be shared externally, as appropriate, to maintain transparency about the evaluation results and how those results can be used to improve the future design. As for other census activities, there is a cost associated with the evaluation and the more complex the undertaking or the more processes get evaluated, the higher the cost is likely to be.

3.523 In addition to the financial cost, human resources and skilled staff to undertake the evaluation should also be planned for. Depending on what aspects of the census operation will be evaluated, the national statistical or census office should ensure that they have adequate personnel (in type and quantity) to undertake the evaluation. A related issue that should be taken into account is the feasibility of the staff to undertake the evaluation. Lack of requisite skilled staff may limit the ability of the country to undertake some or all of the planned components of the evaluation. An alternative is to hire temporary staff with the required technical skills to conduct the evaluation or to collaborate with other agencies such as research institutions.

3.524 As has already been stated, plans for the evaluation of census processes and procedures should be an integral part of the overall census plan and be planned for from the start of census activities. In addition, documentation of the census experience should have provision of information for evaluation as one of its objectives. As presented in paragraphs 3.468–3.472 on “Documentation of census experience”, the census organization should have a knowledge management system to document and archive complete records on plans, activities, and decisions taken during the entire census operation, including on problems faced and how they were resolved. Documented evidence on how the census was undertaken provides valuable input for the evaluation programme. It should be noted, however, that depending on what has to be evaluated, some of the information may not be available until the end of the census operation.

PART FOUR. POPULATION AND HOUSING CENSUS TOPICS

Chapter I: Population census topics

I. Factors determining the selection of topics

4.1. In line with the overall approach to revision 3 of *Principles and Recommendations for Population and Housing Censuses*, the selection of census topics is based on the outputs expected to be produced by the census. Therefore, the first step involves clear identification of user requirements for data; the core and additional topics are then decided on that basis. For each of the core topics there is a recommended tabulation. It is recommended that countries collect data on the core topics and also produce the recommended tabulations, as this would improve the international harmonization and comparability of statistics through the use of common concepts, definitions and classifications. Use of an agreed international approach would also enhance the capacity of countries to generate statistics for monitoring the socioeconomic situation of their populations, including for the provision of data for the internationally agreed development goals.

4.2. The topics to be covered in the census (that is, the subjects regarding which information is to be sought for each individual or household) should, however, be determined upon a balanced consideration of:
(a) The needs of the broad range of data users in the country at both the national and local area level (national priority);
(b) Achievement of the maximum degree of international comparability, both within regions and on a worldwide basis (international comparability);
(c) Sensitivity of the topics and respondent burden, that is, the willingness and ability of the public to give accurate information on the topics (suitability);
(d) Technical competence of the enumerators in regard to obtaining information on the topics (suitability);
(e) Total national resources available for conducting the census (resources);
(f) Availability of relevant information held in alternative data sources (alternative sources).

4.3. Such a balanced consideration will need to take into account the advantages and limitations of alternative methods of obtaining data on a given topic within the context of an integrated national programme for gathering demographic and related socioeconomic statistics (see paragraphs 1.63–1.68).

4.4. In selecting the population topics, regard should also be given to the usefulness of historical continuity, which provides the opportunity for comparison of changes over a period of time. Census takers should avoid, however, collecting information that is no longer required by users. Information should not be collected simply because it was traditionally collected in the past, bearing in mind changes in the socioeconomic circumstances of the country. It becomes necessary, therefore, in consultation with a broad range of users of census data, to review periodically the value of even long-standing topics and to re-evaluate the need for their continued collection, particularly in the light of new data needs and alternative data sources that may have become available for investigating topics hitherto covered in the population census. Each of five key factors that need to be taken into account in reaching a final decision on census content are briefly reviewed in the following paragraphs.

A. National priority

4.5. Prime importance should be given to the fact that population censuses should be designed to meet national needs. In defining national data needs for population census data, the full range of national uses (for example, policy, administration and research) and national users (for example, national and local government agencies, those in the private sector, and academic and other researchers) should be considered. The prime consideration is that the census should provide information on those topics that are of greatest value to the country, with questions framed so as to elicit data of maximum utility.

4.6. Each country's decision with regard to the topics to be covered should depend upon a balanced appraisal of how urgently the data are needed and whether the information could be equally well or better obtained from other sources. Experience has shown that national needs will best be served if the census includes topics generally recognized as being of basic value and defined in accordance with regional and global standards. Global and regional census recommendations can help in this appraisal by providing information about standard census topics and related definitions and concepts based on a wide range of national census experience. It is recognized however that counties that rely more on administrative records as their prime data source may be more limited in the precise detail of the information that can be collected on particular topics.

4.7. Many countries may find it necessary to include in the census topics of national or local interest in addition to the topics included in these recommendations. Labour force or household survey data may supplement census data so to obtain information on topics that cannot be included in the census for whatever reason. It is possible that some countries may omit from the census certain recommended topics either because there is no need to collect the data or because there are legal barriers or particular sensitivities in doing so, as for example may be the case for topics such as fertility, ethnicity and religion.

B. International comparability

4.8. The desirability of achieving regional and worldwide comparability should be another major consideration in the selection and formulation of topics to be included in the census. National and international objectives are usually compatible, however, since international recommendations are based on a broad study of country experience and practice, and the definitions and methods contained in international recommendations have successfully met general national needs in a wide range of circumstances. Furthermore, the analysis of census data for national purposes will often be facilitated if, through the use of international recommendations, it is possible to compare the data with those of other countries on the basis of consistent concepts, definitions and classifications. The post-2015 international development agenda, which places increasing demand for expanded data collection, is also another determining factor that countries should take into consideration.

4.9. If the particular circumstances within a country require a departure from international standards, every effort should be made to explain these departures in the census publications and to indicate how the national presentation can be adapted to the international standards.

C. Suitability

4.10. A prerequisite for the inclusion of topics in the census should be the willingness and ability of respondents to provide accurate information on them. It is advisable to avoid topics that could increase

the burden on respondents and those that are likely to arouse fear, local prejudice or superstition or that might be used to deliberately promote political or sectarian causes as these are likely to have a detrimental effect on response rates and support for the census. In an interview-based census or where the collector needs to obtain information through observation, consideration needs to also be given to the level of knowledge and skill of the interviewer or collector and whether they can be adequately trained to collect this information accurately. Topics that are too complicated or difficult for the average respondent or enumerator to answer quickly should also not be included. The exact phrasing of a question that will obtain the most reliable responses may depend on national circumstances and, as described in Part Three of these recommendations, should be well tested prior to the census (see paragraphs 3.110–3.114).

D. Resources

4.11. The selection of topics should be carefully considered in relation to the total resources available for the census. An efficient collection of accurate data for a limited number of topics, followed by prompt tabulation and publication, is more useful than the collection of data for an overambitious list of topics that cannot be properly processed and disseminated in a timely, reliable and cost-effective manner. In balancing the need for data against resources available, the extent to which questions can be precoded is yet another consideration. Information from studies on the capacity of users and on the measurement of data utilization may also be an important factor in determining whether or not it is economically feasible to include certain topics in the census.

E. Alternative sources

4.12. In the selection of topics to be investigated in a census, consideration should be given to whether data are available from other sources, taking into account the relative advantages and limitations of the alternative sources. For example, data may be available from administrative records, or similar data may be collected by household surveys. While household surveys may not be able to collect the detailed information that can be obtained from censuses for small areas or small population groups, there are other advantages associated with interviewers collecting the data rather than, for example, the information being collected on a self-enumeration basis or from administrative records – such as, for example, administering and navigating through probing questions. Those topics for which no alternative sources exist should be given higher priority while those for which alternative sources are readily available should be accorded lower priority.

II. List of topics

4.13. The list of topics included in these recommendations for population censuses are based on the global and regional census experience of the last several decades. The topics included here are, with some minor revisions, generally the same as those included in the previous United Nations population census recommendations.[93] However, the concepts and definitions for some of the topics relating to economic characteristics have been substantially revised to reflect the more recent recommendations of the International Conference of Labour Statisticians.[94]

[93] *Principles and Recommendations for Population and Housing Censuses Revision 2*, Statistical Papers No. 67/Rev2 (United Nations publication, Sales No. E.07.XVII.8).

[94] For more detail see the Resolution concerning statistics of work, employment and labour underutilization, adopted by the 19th International Conference of Labour Statisticians (Geneva, 2013).

4.14. It should be stressed that no country should attempt to cover all the topics included in the list of population topics (see Table 3). Rather, countries will need to make their selection of topics in the light of the considerations discussed in paragraphs 4.1–4.12 above, bearing in mind any regional recommendations currently pertaining to census topics. In using the classifications of different topics presented in this part of the *Principles and Recommendations for Population and Housing Censuses,* it is necessary to outline that all the one- and two-digit classification levels are recommended, while those at the three-digit level are incorporated for illustrative and guidance purposes only.

4.15. Evolving census experience over the past several decades globally and in the regions has demonstrated that a set of topics exist on which there is considerable agreement in regard both to their importance and to the feasibility of collecting data on them in a census. Data on those within this set that are found to be excessively costly are probably best collected through separate surveys of a sample of the population. The exceptions to this consensus occur, at one extreme, among the countries with the most developed statistical systems, where adequate data on a number of the topics listed, including some of the core ones, are available from non-census sources; and, at the other, among the countries in which data collection opportunities are limited and it is felt that advantage must be taken of the possibilities offered by the census to investigate topics that, under more ideal circumstances, might be investigated more suitably by other means.

4.16. Although the set of topics covered in these recommendations is quite comprehensive in terms of those generally considered suitable for inclusion in a population census, it is also recognized that some countries may find it necessary to include one or more additional topics on which information is of particular national or local importance. However, before the final decision is made to include any such additional topics, their suitability should always be carefully tested.

4.17. To assist countries in using the present publication and in determining their own priorities, lists of recommended population topics are summarized in paragraph 4.21, with the core topics shown in boldface. These core topics correspond to those that were included as “priority topics” in the majority of the regional recommendations in previous census decades.

4.18. The topics listed in paragraph 4.21 are grouped into eight categories: geographic and internal migration characteristics, international migration characteristics, household and family characteristics, demographic and social characteristics, fertility and mortality, educational characteristics, economic characteristics, and agriculture.

4.19. Within each category, a distinction is made between topics collected directly (those that appear in the census schedule or questionnaire), and derived topics. Although data for the derived topics also come from information on the questionnaire, they do not necessarily come from replies to a specific question. “Total population”, for example, is derived from a count of the persons entered on the questionnaires as persons present or resident in each geographic unit. Such derived topics may perhaps be more correctly considered as tabulation components, but they are listed as topics in order to emphasize the fact that the questionnaire must in some way yield this information.

4.20. The paragraph numbers in parentheses after each entry in Table 3 refer either to the paragraphs in which the group of topics as a whole is discussed in section IV below or to the paragraphs in which the definition and specifications of individual topics are discussed.

4.21. In the following list of population census topics, core topics are shown in bold and are represented by ♦ for topics that are collected directly, and by □ for those that are derived. Additional topics are represented by ○, and additional topics derived from a core topic are indicated with Δ.

Table 3. List of population census topics

A. Geographic and internal migration characteristics (paras. 4.50–4.100)		
(1)	**Place of usual residence** (paras. 4.52–4.57)	♦
(2)	**Place where present at time of census** (paras. 4.58–4.63)	♦
(3)	**Place of birth** (paras. 4.64–4.71)	♦
(4)	**Duration of residence** (paras. 4.72–4.74)	♦
(5)	**Place of previous residence** (paras. 4.75–4.76)	♦
(6)	**Place of residence at a specified date in the past** (paras. 4.77–4.81)	♦
(7)	**Total population** (paras. 4.82–4.88)	□
(8)	**Locality** (paras. 4.89–4.91)	□
(9)	**Urban and rural** (paras. 4.92–4.100)	□
B. International migration characteristics (paras. 4.101–4.120)		
(1)	**Country of birth** (paras. 4.105–4.109)	♦
(2)	**Country of citizenship** (paras. 4.110–4.115)	♦
(3)	Acquisition of citizenship (paras. 4.116)	○
(4)	**Year or period of arrival** (paras. 4.117–4.120)	♦
C. Household and family characteristics (paras. 4.121–4.148)		
(1)	**Relationship to the reference person of household** (paras. 4.129–4.139)	♦
(2)	**Household and family composition** (paras. 4.140–4.147)	□
(3)	Household and family status (para.4.148)	○
D. Demographic and social characteristics (paras. 4.149–4.213)		
(1)	**Sex** (para. 4.150)	♦
(2)	**Age** (paras. 4.151–4.162)	♦
(3)	**Marital status** (paras. 4.163–4.171)	♦
(4)	Ethinocultural characteristics (paras. 4.172-4.173)	○
(5)	Religion (paras. 4.174–4.178)	○
(6)	Language (paras. 4.179–4.182)	○
(7)	Ethnicity (paras. 4.183–4.187)	○
(8)	Indigenous peoples (paras. 4.188–4.192)	○
(9)	**Disability status** (paras. 4.193–4.213)	♦
E. Fertility and mortality (paras. 4.214–4.257)		
(1)	**Children ever born alive** (paras. 4.228–4.233)	♦

(2)	**Children living** (paras. 4.234–4.236)	♦
(3)	**Date of birth of last child born alive** (paras. 4.237–4.240)	♦
(4)	Births in the past 12 months (paras. 4.241–4.243)	Δ
(5)	Deaths among children born in the past 12 months (paras. 4.244–4.246)	Δ
(6)	Age, date or duration of first marriage (para. 4.247–4.248)	○
(7)	Age of mother at birth of (date or time when) first child born alive (para. 4.249)	○
(8)	**Household deaths in the past 12 months** (paras. 4.250–4.254)	♦
(9)	Maternal or paternal orphanhood (paras. 4.255–4.257)	○
	F. Educational characteristics (paras. 4.258–4.288)	
(1)	**Literacy** (paras. 4.258–4.264)	♦
(2)	**School attendance** (paras. 4.265–4.271)	♦
(3)	**Educational attainment** (paras. 4.272–4.280)	♦
(4)	Field of education and training, and educational qualifications (paras. 4.281–4.288)	○
	G. Economic characteristics (paras. 4.289–4.386)	
(3)	**Labour force status** (paras. 4.307–4.338)	♦
(5)	**Status in employment** (paras. 4.339–4.351)	♦
(6)	**Occupation** (paras. 4.352–4.355)	♦
(7)	**Industry** (paras. 4.356–4.359)	♦
(8)	Place of work (paras. 4.360–4.365)	○
(9)	Institutional sector of employment (paras. 4.366–4.368)	○
(10)	Working time (paras. 4.369–4.375)	○
(11)	**Participation in own-use production of goods** (paras. 4.376–4.381)	♦
(12)	Income (paras. 4.382–4.386)	○
	H. Agriculture (paras. 4.387–4.396)	
(2)	Own-account agriculture production (paras. 4.389–4.392)	○
(3)	Characteristics of all agricultural jobs during the last year (paras. 4.393–4.396)	○
	Legend: ♦ Core topic, collected directly; □ Core topic, derived; ○ Additional topic; and Δ Additional topic, derived from a core topic.	

III. Population count

4.22. One of the main objectives of a population census is to provide an accurate count of the population of a country at a point in time. An accurate population count is essential for the efficient planning and delivery of services, distribution of resources, defining boundaries for electoral representation, policy development and a wide range of other administrative and statistical purposes (cross reference to part I on uses of PHCs).

4.23. A "population count" may be a subset of or the whole of the enumerated population. A country may produce one or more population counts, all derived from the enumerated population.

4.24. Countries are usually most interested in the count and distribution of usual residents because usual residence is generally the best indication of where people will demand and consume services. A count of usual residents is therefore most relevant for many common planning and policy purposes.

4.25. Some countries may supplement the population count from their census with information from other sources, for example on usual residents temporarily outside the country at the time of the census, to produce population estimates. Alternatively, the census itself may collect data on these groups, resulting in an enumerated population that includes individuals who are usually resident but temporarily absent.

4.26. In countries whose censuses use direct enumeration, information about each person can be collected and entered on the census questionnaire either where he or she is (or was) present on the day of the census or at his or her usual residence. Paragraphs 2.46–2.63 describe the basis for deciding the place of enumeration in the census.

4.27. A country may need to produce one or more different counts for different national purposes. Population counts may be required on a population present basis, or for the usually resident population, or for the service population. The choice of population count(s) required will depend on national circumstances; some countries will require more than one. The information collected about each person by the census will need to enable the required population counts to be derived. In some cases, for regional and international comparison purposes, the population count based on the concept of usual residence might need to be produced.

4.28. A census aims to achieve a full and unduplicated coverage of the population. In practice, countries may face a range of challenges in enumerating the population at the place they decide (where present on census day or where usually resident), and in producing the population counts they require. Many of these challenges relate to the difficult-to-enumerate groups of the population and persons for whom the concept of "usual residence" is not easily defined. The latter present an increasing problem as populations become more mobile (nationally and globally) and household and family structures more dynamic.

4.29. In developing strategies for enumerating the population and collecting information to support the required population counts, it is important to consider consistency with the standards for international migration statistics described in paragraphs 4.101–4.104.

A. Population present count

4.30. A population present count is the simplest form of population count from a population census. In a questionnaire-based census where no reference is made to usual residence, people are enumerated at the place where they are found, usually the dwelling where they spend census night. Foreign residents who are in the country at the time of the census will be included but usual residents of the country who are absent at that time will be excluded.

4.31. A population present count removes complications associated with the application of the concept of place of usual residence, and can reduce the incidence of double counting or missing people if the enumeration is carried out in a single day or if reference is made to the same census moment for the whole population. Apart from the benefit of simplicity, a population present count offers a cost advantage because the census does not need to collect additional information about usual residents not at their usual residence at the time of the census.

4.32. The major disadvantage of a population present count is that it does not enable a full count of usual residents to be derived, and may not provide a true geographic distribution of usual residents for effective planning and policy purposes.

4.33. A population present count may be a good proxy for a count and distribution of usual residents, particularly if nearly all the population will be at their usual residence at the time of the census, or if the characteristics of those persons present are very similar to the characteristics of usual residents. However, in many countries significant numbers of people will not be at their usual residence at the time of the census, and the characteristics of absent usual residents will be systematically different from non-residents present. In such cases, a population present count will not be a good proxy for a count of usual residents. Large seasonal movements of people due to weather changes, employment, education, holidays and other factors can result in these systematic differences between the population present and the usually-resident population. The ability to produce accurate information on families and households is also reduced when persons temporarily absent from their usual residence (e.g. seasonal workers or students) are not enumerated with their families or households.

4.34. To produce a population present count, information is required on all persons present and the address at which they are enumerated. It is also very useful to collect information to identify those persons present who are not at their usual residence and those persons who are not usual residents of the country.

4.35. Ideally a population present count should include all the persons present at the census reference moment, regardless of the difficulty of their enumeration. For some of these groups the concept of “at the time of the census” may need to be extended to allow the enumeration to take place. When, however, the enumeration is extended over a period of time, the risk of either overcount or undercount may increase. In fact, persons who are at multiple locations during this extended period may be counted at more than one location, or alternatively they may not be counted at any location. Those risks increase further when reference is made to a *census period* rather than to a *census moment*.

B. Usual resident population count

4.36. Countries increasingly prefer a usual resident population count because this count offers better information for planning and policy purposes on the demand for services, households, families and internal migration.

4.37. A **usual resident population count** is a count of all usual residents of a country at the time of the census. Although countries will determine the definition of a usual resident according to their own particular circumstances, it is recommended that in defining a usual resident and the place of usual residence, countries apply the definition contained in paragraph 2.50. Usual residents may or may not have citizenship of the country, and they may also include undocumented persons, applicants for asylum or refugees. Usual residents then may include foreigners who reside (legally or illegally), or intend to reside, in the country continuously for either most of the last 12 months or for 12 months or more, depending on the definition of place of usual residence that is adopted by the country. Persons who may consider themselves usual residents of a country because of citizenship or family ties, but were absent from the country for either most of the last 12 months, or for 12 months or more, depending on the definition adopted, should be excluded. Conversely, persons who are normally resident in the country but who are temporarily absent should be included in the usually resident population. Countries applying a different definition of a usual resident for national purposes should produce a usual resident population count using the recommended 12-month definition for the purposes of international comparability.

4.38. A usual resident count provides a comprehensive count of the population of a country for long-term planning and policy purposes, and for service delivery purposes at subnational geographic levels.

4.39. Whether enumeration takes place on a "place where present" basis or on a "where usually resident" basis (as described in paragraphs 2.55–2.63), a count of usual residents can be produced provided that the necessary information is collected.

4.40. To produce a usual resident population count, information is required on all usual residents and the address of their usual residence, with sufficient detail to be able to record usual residence at the lowest geographic area level required for tabulation. If the census is taken on a population present basis, then the information collected needs to differentiate clearly between persons enumerated at their usual residence, persons usually resident who were elsewhere at the time of the census, and persons present who are usually resident elsewhere. Information should also be collected to identify those persons present who are not usual residents of the country. If, however, the census is taken on a usual residence basis, then information about all usual residents needs to be collected with respect to their usual residence, regardless of whether they are present at the time of the census or not, to ensure full coverage.

4.41. It may be difficult to obtain information from those usual residents who are absent from the country at the time of the census, particularly where no other person is present at the place of usual residence at the time of the census to provide information about those people. Some countries include estimates or imputations of the number and characteristics of these usual residents not enumerated by the census, and obtained from other sources, in the census population count.

4.42. There can be challenges in applying the concept of a "usual resident" if a person is considered to have more than one residence, sometimes in different countries. This is particularly the case for people who may spend parts of the time in communal establishments or institutions, such as schools or military camps. There may also be those who do not consider themselves to have a usual residence at all, such

as nomadic peoples or persons sleeping rough[95]. In such cases place of usual residence can be considered to be the place where they are enumerated. Countries will need to develop appropriate operational rules for resolving cases where it is not clear whether a person is a usual resident of the country, or where the usual residence of the person within the country is not clear.

4.43. There are population groups for which some uncertainty may arise in defining their place of usual residence within the country. The recommended conventional treatment of these cases is as follows:

a) Persons who work away from home during the week and who return to the family home at weekends should consider the family home as their place of usual residence.

b) Persons of minor age in primary and secondary education who are away from home during the school term should consider their family home as their place of usual residence.

c) Students in tertiary education who are away from home while at college or university should consider their term-time address as their place of usual residence regardless of whether this is an institution (such as a boarding school) or a private residence.

d) The institution should be taken as the place of usual residence of all inmates who at the time of the census have spent, or are likely to spend, six months or more in the relevant institution. Examples of inmates of institutions include patients in hospitals or hospices, older persons in nursing homes or convalescent homes, prisoners and those in juvenile detention centres.

e) Where a person regularly lives in more than one residence within the country during the year, the one where he or she spends the majority of the week or year before the census should be taken as his or her place of usual residence. These persons are not considered to be persons with no usual residence.

f) For the (national) military, naval and diplomatic personnel and their families located outside the country the following classification rules should be applied:
i) If they are residing abroad for less than 12 months and they are intending to return to the place of departure, they should be allocated within the country in accordance with the rules for usual residence. In particular, they could be allocated to (by decreasing order of priority):
- The family home address within the country, if any;
- The duty station within the country to which they were attached before leaving.

ii) If they are residing abroad for at least 12 months or if they are not intending to return to the place of departure (although returning in the country within a 12-month period), they should be attributed to a “virtual region” (extra-region) of the country of departure.

g) The place of enumeration should be taken as the place of usual residence of homeless or roofless persons, nomads, vagrants and persons with no concept of usual residence.

h) A child who alternates between two households within the country (for instance when his or her parents are separated) should consider the household where he or she spends the majority of the year before the census as his or her place of usual residence. Where an equal amount of time is spent with both parents, the place of usual residence should be with the parent or household with whom the child is staying at the census reference time.

[95] People ‘sleeping rough' are those living in the streets or public spaces, without a shelter that can be defined as living quarters.

Comments FWN: As earlier in the usual residence section, we should check with ESTAT's CES task force and ensure these bulleted lists are the same in CES recommendations unless there is some strong justification for any difference.

Comments FWN: We should look for an alternative phrasing than 'family home'. It's not necessarily a person going back to a 'family' at weekends. They could be going back to a bunch of friends/housemates, or indeed to live alone but in a different place. The matter is to do with a person living in different locations on weekdays and weekends, not who they are living with. Similarly, we should look for an alternative for the term 'inmates'. This sounds like prisoners, specifically. Are you an 'inmate' of a college dorm?

C. Supplementary population counts

1. Service population count

4.44. Counts of usual resident and/or population present are the most commonly produced census population counts, since in most cases they provide a good indication of the geographic distribution of demand for and patterns of use of services. However, in some countries or specific parts of a country, the demand for and use of services, as well as other considerations such as exposure to risks and opportunities, is shaped significantly by temporary movements. This is the case, for example, in some agricultural areas which have very few usual residents but large populations of short-term residents during harvest season; in tourist resorts with large numbers of visitors but relatively few usual residents; or in urban centres where few people live but many are present as workers during the daytime. In such cases a country may wish to produce supplementary population counts. Service populations are relevant where a significant proportion of the population providing or using services in an area are not usual residents of that area. Types of service population counts include daytime populations, workplace populations and visitor populations. In some countries there may also be an interest in foreign service populations, consisting of foreign residents who cross the border regularly to provide or consume services. This is particularly important in the planning and provision of transport services.

4.45. A service population count may include some or all of the difficult-to-enumerate groups, depending on the type of service population required. For example, daytime service populations may include civilian foreigners who cross the border daily to work or consume services in the country.

4.46. To produce a service population count, in addition to an estimate of usual residents, information is required about where people provide or demand services. For seasonal populations (holiday, resort), information is needed on the destination and timing of seasonal trips. Some countries will produce service population counts by supplementing the population present count or usual resident population count with information from other sources, such as visitor information from hotels and resorts, to produce visitor population counts. Alternatively, additional information may be collected by the census.

2. Counts of population subgroups

4.47. Accurate population counts, required for the efficient planning and delivery of services, distribution of resources, defining of boundaries for electoral representation, policy development and the design and analysis of household surveys, are required for various population subgroups within a country. These subgroups are typically based on geography, age and sex. There may also be a need to identify other populations such as the school population, working population, indigenous population or

disadvantaged populations to enable more informed policy formation and better targeted service provision. A range of characteristics will be required to identify these populations and population subgroups, depending on the services being planned, the resources to be distributed and so on. The need for population counts for particular subgroups will determine the questions asked in the census.

D. Difficult-to-enumerate groups or difficult-to-reach groups

4.48. Difficult-to-enumerate groups can be defined as those for whom a real or perceived barrier exists to full and representative inclusion in the data collection process[96]. In fact, some populations present special challenges of various sorts that make them harder to enumerate than the general population. The definition of difficult-to-enumerate groups can vary between countries as the reasons why members of a population group are hard-to-reach (or difficult to enumerate) can vary according to the context of each national, geographic, or social environment. Nonetheless, countries should ensure that the following difficult-to-reach groups are included in the population count:

(a) **Nomads and persons living in areas to which access is difficult.** Making contact with these groups to enumerate them can be difficult, particularly as part of a point-in-time count. Enumeration may need to be done at a different time, over an extended period, or by using alternative methods to enable contact with these groups. For example, countries might consider asking those who provide services to these groups to assist with their enumeration. Seasonal movements may be identified in advance and this information can be used by collectors to enable contact. There needs to be planning and consultation, particularly with influential members of these groups, prior to the census to organize for their enumeration. Communications that publicize the benefits of the census and engaging appropriate leaders in support of the census may assist coverage. Awareness of cultural issues relevant to specific groups should also be considered in developing enumeration strategies.

(b) **Civilian residents temporarily absent from the country.** As these persons will be absent from the country at the time of the census, they will be excluded from a usual resident population count. To produce a usual resident count countries may collect information on these people from another family or household member present at the time of the census, but where a complete family or household is outside the country at the time of the census, it may not be possible for the census to collect information about these people. Estimates for usual residents temporarily absent from the country based on other sources may be required to produce reliable estimates of usual residents for planning and policy purposes.

(c) **Civilian foreigners who do not cross a frontier daily and are in the country temporarily.** These include undocumented persons, or transients on ships in harbour at the time of the census. These groups may be in the country at the time of the census and therefore form part of the population present count. It is important to include these groups in the population count if their demand for services is to be considered for planning and policy development purposes. However, these groups may prefer not be counted, either because they fear ramifications from being counted or because they do not identify themselves as part of the population of the country. Language and communication may present

[96] The reference here is to Tourangeau R., Brad Edwards B., Johnson T. P., Kirk M. Wolter K.M. Bates N. (eds), Hard-to-Survey Populations, Cambridge University Press (2014), DOI: https://doi.org/10.1017/CBO9781139381635.003 and to U.S. Census Bureau: Counting the hard to count in a census (2019), Available from: https://www.census.gov/content/dam/Census/library/working-papers/2019/demo/Hard-to-Count-Populations-Brief.pdf

challenges. Countries need to develop strategies, appropriate for their context, to include these groups in their enumeration.

(d) **Refugees, asylum seekers and internally displaced persons.** Refugee populations, asylum seekers and internally displaced persons (in and outside camps) should be enumerated and their numbers presented separately, allowing calculation of country population excluding such groups, when such a population count is required for non-demographic purposes.

(e) **Military, naval and diplomatic personnel and their families located outside the country and foreign military, naval and diplomatic personnel and their families located in the country.** Apart from the difficulties mentioned in (b) and (c) that are common to groups who are absent from their own country, enumeration of these groups is subject to diplomatic protocols. Detailed counts and characteristics of these groups may be considered sensitive on security grounds in some countries. Counts of these groups may be available from administrative records.

(f) **Civilian foreigners who cross a frontier daily to work in the country.** This group should be excluded from a usual resident population count. The practice of counting people where they spend census night removes much ambiguity and reduces possible duplication. The difficulty then is trying to include them in a service population if countries want to consider this group in policy development and in planning service delivery.

(g) **Civilian residents who cross a frontier daily to work in another country.** These persons are usual residents of the country and should be included in the population count.

(h) **Merchant seafarers and fishers resident in the country but at sea at the time of the census.** This group includes those who have no place of residence other than their quarters aboard ship. Identifying that the ship will be at sea at the time of the census may be problematic, so countries will need to develop strategies to ensure inclusion of this group in the population count. This may include providing this group with census forms before their ship goes to sea or enumerating the ship before the time of the census.

(i) **Homeless or roofless persons, vagrants and persons with no concept of usual residence.** These should be included in the population count, and the census office should work with local government agencies, charities and other supporting bodies that provide support for this population group to identify the best method of collecting census information from these people.

(j) **Persons living in buildings with restricted access.** Persons living in gated communities, condominiums and apartment buildings could be difficult to enumerate, particularly as part of a point-in-time count. Enumeration may need to be done at a different time, over an extended period, or by using alternative methods to enable contact with these groups.

(k) **Stateless persons.** These are individuals who are not considered as nationals by any State under the operation of its laws. They are often undocumented and may not wish to be enumerated. However, every effort should be made to include such persons in the census. The census office should work with responsible government agencies, non-governmental organizations familiar with this population group and the United Nations High Commissioner for Refugees (UNHCR) to establish the best method for including stateless persons in the population count. Country of citizenship is generally essential for the identification of this group, though the census office should consult with relevant ministries and agencies, including the UNHCR, to determine whether additional information (such as residence history or identity documentation) may be required to establish the status of a stateless person.

(l) **Migrants in irregular situation or undocumented.** Persons moving outside regular migration channels[97].

Comment FWN: Might we need to add older persons and those with cognitive impairments? Both older people living in institutions (relevant for both enumeration-based and register-based censuses), and those living in private households but without the cognitive capability to complete a questionnaire-based census.

4.49. The same applies to register-based censuses. Administrative data might improve the coverage for some hard-to-reach populations but may also lead to other hard-to-reach or "hidden" populations for different population groups. For example, undocumented migrants will by default not be included in the Population Register, and therefore in statistics derived from administrative registers. Many of the other difficult-to-enumerate groups will instead be included, even if not identifiable as such (e.g. this could be the case of the refugees, but this will not be an issue as the interest for the census is that these subpopulations are included rather than identified).

IV. Definitions and specifications of topics

4.50. The present section contains the recommended definitions and specifications of all topics presented in the order in which they appear in Table 3 above. It is important that census data be accompanied by the definitions used in carrying out the census. It is also important that any changes in definitions that have been made since the previous census are reported in the metadata and, if possible, accompanied by an assessment of the effect of such changes on the relevant data, in order to ensure that users will not confuse valid changes over a period of time with increases or decreases resulting from changed definitions.

A. Geographic and internal migration characteristics

4.51. It should be noted that "place of usual residence" and "place where present at time of census" may be considered alternative topics when countries do not have the resources to investigate both topics for general census purposes. Some countries, however, will want to investigate both topics for general purposes. The relationship between the two topics and their further relationship to the topic of "place of enumeration" are set out in Part Two, Chapter IV (see paragraphs 2.55–2.63).

4.52. It is recommended that countries investigating only "place where present at time of census" for general purposes should also obtain information on "place of usual residence" for all persons who do not usually reside in the household where they were enumerated, to be used in connection with the information on "place of birth", "duration of residence", "place of previous residence" or "place of residence at a specified date in the past" for the purposes of determining internal migration status. If, in the compilation of the population of geographic units, persons are allocated to the place where they were present at the time of the census, information on the four above-mentioned migration characteristics will be irrelevant for persons who were only visiting, or transient in, the place at which they were present. Since such persons must, in any case, be identified in the questionnaire as non-

[97] See: https://www.iom.int/key-migration-terms. Although a universally accepted definition of irregular migration does not exist, the term is generally used to identify persons moving outside regular migration channels. The fact that they migrate irregularly does not relieve States from the obligation to protect their rights. Moreover, categories of migrants who may not have any other choice but to use irregular migration channels can also include refugees, victims of trafficking, or unaccompanied migrant children.

residents so that they will not be erroneously classified as recent in-migrants, information on place of usual residence should be collected, which will make it possible to include the entire population in the tabulation of internal migration characteristics.

1. Place of usual residence (core topic)

4.53. Information on the number of people usually residing in an area is basic to most informed decision-making about the area, whether it be a country, an urban agglomeration or a civil division. The number of residents determines the levels of most services required in an area.

4.54. The *place of usual residence* may be the same as, or different from, the place where the enumerated person was present at the time of the census or his or her legal residence. For a definition of place of usual residence, see paragraphs 2.48–2.50.

4.55. Although most persons will have no difficulty in stating their place of usual residence, some confusion is bound to arise in a number of situations where persons have more than one residential address. These cases might include persons who maintain two or more residences, students living at school, members of the armed forces living at a military installation but still maintaining private living quarters away from the installation, and persons who sleep away from their homes during the working week but return home for several days at the end of each week (see also paragraph 2.53). In some other circumstances, referring to the person's intentions for the future may assist the determination of the place of usual residence.

4.56. Problems may also arise with persons who have (a) been residing at the place where they are enumerated for some time, ~~perhaps for more than half of the preceding 12 months~~, but do not consider themselves to be residents of that place because they intend to return to their previous residence at some future time; or (b) left the country temporarily but are expected to return after some time longer than 12 months from the departure. In such instances, clearly stated time limits of presence in or absence from a particular place must be based upon the 12-month limit and used to determine whether or not the person is usually resident there. The 12-month criterion is necessary for determining whether or not a person is usually resident in the country. It is also required for international comparability for migration purposes. ~~but less so for place of usual residence within the country for measuring internal migration, where a six-month rule might be more appropriate as it will refer more closely to the concept of "most of the time".~~

4.57. If each person is to be entered in the questionnaire only at his or her place of usual residence, the topic need not be investigated separately for each person, because the information will be available from the location information entered for the questionnaire as a whole.

4.58. Information on the place of usual residence should be collected in enough detail to enable tabulations to be made for the smallest geographic subdivisions required by the tabulation plan and to meet the requirements of the database within the cost limits and operational procedures required to code to a fine degree of detail.

2. Place where present at time of census (core topic)

4.59. In cases where the census is taken on the basis of "place where counted", this topic may fulfil some of the functions of place of usual residence.

4.60. The *place where present at time of census* is, in theory, the geographic place at which each person was present on the day of the census, whether or not this was his or her place of usual residence. In practice, the concept is generally applied to the place where the person was present at the moment of the census, because many persons may not be physically present at the place of enumeration during most of the day.

4.61. As mentioned in Part Two, Chapter IV (see paragraphs 2.57–2.58), the concept is sometimes further extended to apply to the night preceding the day of actual enumeration in cases where the enumeration extends over a long period of time and persons are not likely to be able to supply information relating to a single moment in the past. Other departures from the definition may be necessary to deal with individual cases, such as persons travelling during the entire night or day of the census and persons who spent the night at work.

4.62. If each person is to be entered in the questionnaire only at the place where he or she was present at the time of the census, the topic need not be investigated separately for each person, because the information will be available from the location information recorded for the questionnaire as a whole.

4.63. Information on the place where each person was present should be collected in enough detail to enable tabulation to be made for the smallest geographic subdivisions required by the tabulation plan and to meet the requirements of the database within the cost limits and operational procedures required to code to a fine degree of detail.

4.64. For countries that collect information from administrative data sources, the concept of "present at the time of the census" may not be relevant. Further, if the census population is used for the purpose, which has legal implications, it is recommended that census is taken on the basis of usual residence.

3. Place of birth (core topic)

4.65. Information on the place of birth is a major input to development of policies relating to migration and the related issues of service delivery to migrants. For the purposes of measuring internal migration, migrants are defined as those persons who usually are residing in a civil division of the country at the time of the census, but were previously resident outside that division. That is, movements within the civil division should not be regarded as being migratory.

4.66. The *place of birth* for those persons born within the country is the civil division in which the person was born; for those born in other countries, it is the country of birth. For persons born in the country (the native-born population), the concept of place of birth usually refers to the geographic unit where the mother of the individual resided at the time of the person's birth. In some countries, however, the place of birth is defined as the geographic unit in which the birth actually occurred. It is recommended to use the place of birth as where the mother of the individual resided at the time of the person's birth. If this concept is not used, the country should explain the definitions it uses in both the census enumerator instructions and in the census reports to aid the interpretation of the data.

4.67. The collection of information distinguishing between the native-born population and those born elsewhere (foreign-born) is necessary where any enquiry on place of birth is made. Even countries

where the proportion of foreign-born population is insignificant, and who may only be interested in information on the place of birth of the native-born population, must first separate the native-born from the foreign-born population. It is therefore recommended that place of birth be asked of all persons. In countries that combine the questions on place of birth and country of birth (where the latter is used to measure international migration), the guidance on the country of birth (see paragraphs 4.105–4.109) should apply.

4.68. Information on the place of birth of the native population is usually used primarily for the investigation of internal migration. For countries that have been recently formed from parts of previously separate entities, however, such information may be of use in assessing the relative size of the population segments from each of those entities and their distribution throughout the country.

4.69. Information on whether or not a person is "born in the country" captures the population according to the boundaries at the time of the census. Using the "born in the country" concept would account for individuals who may have been affected by changes to a country's boundary.

4.70. For the purposes of measuring internal migration, it is usually sufficient to collect information only on the major civil division (state, province or department, for example) in which the place of birth is located. If desired, more detailed information on the subdivision of a specific locality can be collected and used for accurate coding of the major division or for presenting data for smaller areas.

4.71. However, for more detailed studies of internal migration, data on the place of birth of the native population even in terms of major civil divisions may not be adequate. For better understanding of the movements of people since birth it ~~is~~ may be necessary to collect information at the smallest possible geographic level, bearing in mind that:

(a) The boundaries of administrative units such as cities and other civil divisions will change over time, which may give rise to ambiguity in data reported;

(b) The costs of coding the reported data to these smaller units may be prohibitive, especially where there are many units and the population is highly mobile.

To overcome the first problem, to the extent possible, both national and subnational boundaries should refer to the boundaries applying at the time of the census. Countries must address the second problem in the light of their own circumstances, bearing in mind the reduced value of place of birth as a measurement of internal migration in a very mobile population.

4.72. It is recommended that, for the study of internal migration, the data on place of birth be supplemented by information collected on duration of residence (see paragraphs 4.72–4.74) and place of previous residence (see paragraphs 4.75–4.76) or of residence at a specified date in the past (see paragraphs 4.77 and 4.81).

4. Duration of residence (core topic)

4.73. The *duration of residence* is the interval of time up to the date of the census, expressed in complete years, during which each person has lived in (a) the locality that is his or her usual residence at the time of the census; or (b) the major or smaller civil division in which that locality is situated.

4.74. In collecting information on duration of residence, it should be made clear that the interest is in length of residence in the major or smaller civil division, or the locality, but not in the particular housing unit. The concept of duration of residence also relates to the most recent move to the current place of usual residence.

4.75. Data on the duration of residence have only limited value in themselves because they do not provide information on the place of origin of in-migrants. Therefore, when the topic is investigated, the place of previous residence should also be collected so that the data can be cross-classified.

5. Place of previous residence (core topic)

4.76. The *place of previous usual residence* is the major or smaller civil division, or the foreign country, in which the individual resided immediately prior to migrating into the civil division of present usual residence.

4.77. Data on the place of previous residence have only limited value in themselves because they do not provide information on the time of in-migration. Therefore, when the topic is investigated and included in the census, the duration of residence (see paragraphs 4.72–4.74) should also be included so that the data can be cross-classified. Alternatively, countries may choose to include a question on place of residence as a specified date in the past (see paragraphs 4.77–4.81 below).

6. Place of residence at a specified date in the past (core topic)

4.78. The *place of usual residence at a specified date in the past* is the major or smaller division, or the foreign country, in which the individual resided at a specified date preceding the census. The reference date chosen should be that most useful for national purposes. In most cases, this has been deemed to be one year or five years preceding the census (or both of these time frames in cases where internal migration is of particular importance to users and resources are sufficient to code the data).

4.79. The former reference date provides information for statistics of both recent internal and international migration during a single year, while the latter may be more appropriate for collecting data for longer-term analysis of migration. When selecting the reference date the ability of individuals to recall with accuracy their usual residence one year or five years earlier than the census date should be considered. For countries conducting quinquennial censuses, the date of five years earlier can be readily tied in, for most persons, with the time of the previous census, but it should be noted that a one-year recall is likely to result in more accurate information than a five-year recall.

4.80. Some countries, however, may wish to use a different time reference than either one year or five years preceding the census because these intervals may present recall difficulties. In such circumstances the time reference should be one that can be associated with the occurrence of an important event that most people will remember.

4.81. For foreign-born persons, the collection of information on year of first or last arrival in the country is recommended (see “International migration characteristics”, paragraphs 4.101–4.120).

4.82. However, no matter what previous date is used, provision must be made for the treatment of infants and young children who are resident at the time of the census but were not yet born at the earlier date. Tabulations of the data should indicate the nature of the treatment of this group.

[Suggested New Topic] - Reasons for change of residence (core topic)

The *reason for change of residence* of persons to be enquired who usually are residing in a civil division of the country at the time of the census, but were previously resident outside that division. As different members of the household might change residence for different reasons, reason for change of residence of each member of the household should be ascertained.

In collecting information on reason for change of residence it should be made clear that it is with respect to place of previous residence (see paragraphs 4.75–4.76). If a person had moved from the place of previous residence for the purpose of education at some point of time got employment there only, the reason for change of residence would be education. There are no recommendations on the specific reasons, countries must establish their own reasons in accordance with their needs.

This topic may be especially relevant to countries with recent experience with significant internal displacements, as it may help in the identification of IDPs.

7. Total population (core topic)

4.83. For census purposes, the *total population* of the country consists of all the persons falling within the scope of the census. In the broadest sense, the total may comprise either all usual residents of the country or all persons present in the country at the time of the census. The total of all usual residents is generally referred to as the de jure population and the total of all persons present as the de facto population.

4.84. In practice, however, countries do not usually fully achieve either type of count, because one or more groups of the population are included or excluded, depending on national circumstances. The general term used to describe the total might imply a treatment opposite to the one given to any of these groups. It is recommended, therefore, that each country describe in detail the figure accepted officially as the total, rather than simply label it as “de jure” or “de facto”.

4.85. The description should show clearly whether each group listed below was or was not included in the total. If the group was enumerated and identified as a separate group, its magnitude should be given; if it was not enumerated, an estimate of its size and the method of estimation should be given, if possible. If any group is not represented at all in the population, this fact should be stated and the magnitude of the group should be shown as “zero”. This may occur particularly with groups (a), (b, (d) and (n) described below (see also paragraph 4.48).

4.86. The groups to be considered are:

a) Nomads;
b) Persons living in areas to which access is difficult;
c) Military, naval and diplomatic personnel and their families located outside the country;
d) Merchant seafarers and fishers resident in the country but at sea at the time of the census (including those who have no place of residence other than their quarters aboard ship);
e) Civilian residents temporarily in another country as seasonal workers;
f) Civilian residents who cross a frontier daily to work in another country;
g) Civilian residents other than those in groups (c), (e) or (f) who are working in another country;

h) Civilian residents other than those in groups (c), (d), (e), (f) or (g) who are temporarily absent from the country;
i) Foreign military, naval and diplomatic personnel and their families located in the country;
j) Civilian foreigners temporarily in the country as seasonal workers;
k) Civilian foreigners who cross a frontier daily to work in the country;
l) Civilian foreigners other than those in groups (i), (j) or (k) who are working in the country;
m) Civilian foreigners other than those in groups (i), (j), (k) or (l) who are in the country temporarily;
n) Refugees in camps;
o) Transients on ships in harbour at the time of the census.

4.87. In the case of groups (h) and (m), it is recommended that an indication be given of the criteria used in determining that presence in, or absence from, the country is temporary.

4.88. In those countries where the total population figure has been adjusted for undercoverage or overcoverage (or net undercoverage), both the enumerated figure and the estimated adjusted population figure should be shown and described. In such cases, documentation should be provided for users explaining how the total population figure from the census has been adjusted. Ideally, where possible, the detailed tabulations should be consistent with the adjusted population figures. However, where this is not possible, if, for example, the costs of the methodology for undertaking these adjustments are prohibitive, the detailed tabulations will, of necessity, be based only on the actual enumerated population.

The use of both adjusted and non-adjusted population figures may be confusing for some stakeholders, particularly those unfamiliar with the adjustments made. For example, planners may find it more helpful to receive only the adjusted population data, while researchers are likely to require both figures for their analysis.

4.89. The population of each geographic unit of the country, like the total population of the country (see paragraph 4.82), may comprise either all usual residents of the unit (see paragraph 4.53) or all persons present in the unit at the time of the census (see paragraphs 4.58–4.59).

8. Locality (core topic)

4.90. For census purposes, a *locality* should be defined as a distinct population cluster (also designated as inhabited place, populated centre, settlement and so forth) in which the inhabitants live in neighbouring or contiguous sets of living quarters and that has a name or a locally recognized status. It thus includes fishing hamlets, mining camps, ranches, farms, market towns, villages, towns, cities and many other population clusters that meet these criteria. Localities can be defined using a national methodology and/or the methodology associated with the Degree of Urbanisation[98], particularly for international statistical comparison purposes. The census reports should explain which methodology has been applied, and if it departs from the concept of a distinct population cluster, provide an explanation to aid in the interpretation of the data.

[98] The Degree of Urbanisation is a harmonised methodology to facilitate international statistical comparisons and to classify the entire territory of a country along an urban-rural continuum. The Degree of Urbanisation classification defines cities, towns and semi-dense areas, and rural areas. See https://read.oecd-ilibrary.org/urban-rural-and-regional-development/applying-the-degree-of-urbanisation_4bc1c502-en#page1

4.91. Localities as defined above should not be confused with the smallest civil divisions of a country. In some cases, the two may coincide. In others, however, even the smallest civil division may contain two or more localities. On the other hand, some large cities or towns may contain two or more civil divisions, which should be considered as subdivisions of a single locality rather than separate localities.

4.92. A large locality (that is to say, a city or a town) is often part of an urban agglomeration, which may comprise the city or town proper together with a suburban fringe or heavily populated area lying outside, but adjacent to, its boundaries. The urban agglomeration is therefore not coterminous with the locality but is an additional geographic unit, which may include more than one locality. In some cases, a single large urban agglomeration may comprise several cities or towns and their suburban fringes. The components of such large agglomerations should be specified in the census results.

4.93. To classify localities by size, it is recommended to use a subset of the following size classes and to place an open-ended category at the two extremes:

- 49 inhabitants or less
- 50-99
- 100-249
- 250-499
- 500-999
- 1,000-2,499
- 2,500-4,999
- 5,000-9,999
- 10,000-24,999
- 25,000-49,999
- 50,000-99,999
- 100,000-249,999
- 250,000-499,999
- 500,000-999,999
- 1,000,000-2,499,999
- 2,500,000-4,999,999
- 5,000,000 or more inhabitants

4. 93. For each locality, it is recommended to measure the share of population living in slums. A slum may be defined by the Government or based on the criteria below[99]. Slum households may be defined as one in which the inhabitants lack one or more of the following:

1) Lack of access to clean water;
2) Lack of access to improved sanitation;
3) Lack of sufficient living space;
4) Lack of housing durability; and,
5) Lack of security of tenure.

[99] https://unhabitat.org/sites/default/files/2020/06/indicator_11.1.1_training_module_adequate_housing_and_slum_upgrading.pdf

9. Urban and rural (core topic)

4.94. Urban and rural areas can be defined using a national methodology, which can be established in accordance with a country's own needs. In addition, for statistical comparisons, urban and rural areas can be delineated using the Degree of Urbanisation methodology which was endorsed by the UN Statistical Commission in 2020. For a full elaboration on the Degree of Urbanisation methodology, please see the manual[100].

4.95. A classification of areas as urban or rural should be done at the smallest administrative unit of the country, or the smallest census collection unit. The classification should be made, in addition to the national methodology, first and foremost, on a measure of population density and population size too.

4.96. Some of the information required for classification may be provided by the census results themselves, while other information may be obtained from external sources. The use of information provided by the census (as, for example, the size class of the locality or the percentage of the population engaged in agriculture), whether alone or in conjunction with information from other sources, means that the classification will not be available until the relevant census results have been tabulated. If, however, the census plans call for the investigation of a smaller number of topics in rural areas than in urban areas or for a greater use of sampling in rural areas, the classification must be available before the enumeration takes place. In these cases, reliance must be placed on previous censuses and/or external sources of information, even if only to bring up to date any urban–rural classification that was prepared at an earlier date.

4.97. The usefulness of housing census data (for example, the availability of electricity or piped water) collected simultaneously with, or not too long before, the population census should be kept in mind. For assembling information from more than one source, the importance of a well-developed system of geocoding should not be overlooked.

4.98. The traditional distinction between urban and rural areas within a country has been based on the assumption that urban areas, no matter how they are defined, provide a different way of life and usually a higher standard of living than are found in rural areas. In many developed countries this distinction has become blurred, and the principal difference between urban and rural areas in terms of the living standards tends to be the degree of population concentration or density. On the other hand, the differences between urban and rural ways of life and standards of living remain significant in developing countries, but even here rapid urbanization in these countries has created a great need for information related to different types of urban areas.

4.99. Hence, although the traditional urban–rural dichotomy is still needed, a classification by Degree of Urbanisation can usefully supplement the dichotomy as it captures characteristics related to the rural-urban continuum from the most sparsely settled areas to the most densely built-up localities with seven classes. When using the Degree of Urbanisation, it is recommended to use the three classes at level 1 (i.e. with two classes within urban areas i.e. urban and semi-urban).

100 EC, FAO, ILO, OECD, UN-Habitat and the World Bank (2021). Applying the Degree of Urbanisation: A methodological manual to define cities, towns and rural areas for international comparisons — 2021 edition https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-02-20-499

4.100. The Degree of Urbanisation should be applied first to a population grid and then to small administrative or statistical spatial units. To create a population grid, it may be necessary to combine the total population of census enumeration areas with the distribution of built-up areas derived from remote sensing.

B. International migration characteristics

4.101. Interest in the movement of people across national boundaries, namely, international migration, has steadily grown among countries concomitant with the increase in international migration. The decision to collect and disseminate information on international migration in a census is dependent upon a number of considerations and national circumstances, including, for example, the national needs for such data. Data on international migrants could provide information on the diversity of a population and can serve to identify subgroups of a population. The present section on international migration supplements and expands the topic "geographic and internal migration characteristics", which is covered above. The United Nations Statistics Division, under the guidance of the United Nations Expert Group on Migration Statistics, has initiated the revision of the *Recommendations on Statistics of International Migration*, Revision 1. At the core of the revision work, lie the revised conceptual framework[101] on international migration and mobility and accompanying statistical definitions, which are reflected in this section.

4.102. The revised conceptual framework on international migration statistics deals with both international migration flows and international migrant stock, and identifies population censuses as being one of the main source for collecting data on international migrants and their characteristics. It emphasizes consistency and alignment between international migration flows and immigrant populations (stocks) and international temporary mobility and temporary populations for improved accuracy and accountability in international migration data. This section is concerned chiefly with the topic of international migrant stock as derived from population censuses.

4.103. In the revised conceptual framework on international migration statistics, four subgroups of the resident population critical to understanding immigrant and emigrant populations are defined: native-born citizens, native-born foreign citizens (including stateless or those without citizenship), foreign-born citizens and foreign-born foreign citizens (including stateless or those without citizenship).

4.104. Consequently, for the study of international migration and migrants, all four subgroups of the resident population of the country of enumeration are critical. In order to identify members of those groups, two items must be recorded in the census: (a) the country of birth, and (b) the country of citizenship. In addition, it is important to record the year of arrival in the country of enumeration so as to establish the length of stay of international migrants residing in the country.

1. Country of birth (core topic)

4.105. *Country of birth* is the country in which the person was born. The concept of country of birth may refer to either the country in which the mother was usually residing at the time of the person's

[101] Report on *Conceptual frameworks and Concepts and Definitions on International Migration* (UN Expert Group on Migration Statistics, April 2021): https://unstats.un.org/unsd/demographic-social/migration-expert-group/task-forces/TF2-ConceptualFramework-Final.pdf

birth or the country where the physical birth actually took place. Either concept can be used depending on the information needs of the country; but for the purpose of measuring international migration, it is recommended to use the country of birth as where the mother of the individual resided at the time of the person's birth. Each country should explain which definition it used in the census. It should be noted that the country of birth of a person is not necessarily the same as his or her country of citizenship, which is a separate census topic dealt with below. It is recommended that country of birth be asked of all persons to distinguish the native-born from the foreign-born population. The collection of this information is necessary even in countries where the proportion of foreign-born population is small. For the foreign-born population, the collection of information on the specific country of birth is recommended so as to permit the classification of the foreign-born population by country of birth. For respondents who are born outside the country of enumeration and cannot identify their country of birth, at least the continent or region where that country is located should be ascertained.

4.106. For purposes of both internal consistency and international comparability, it is recommended that information on the country of birth be recorded according to national boundaries existing at the time of the census. Information on the year of arrival in the country (see paragraph 4.118 below) can be used to identify persons who owe their status of foreign-born to changes in national boundaries. It is essential that the coding of information on the country of birth be done in sufficient detail to allow for the identification of all relevant countries of birth.

4.107. For purposes of coding, it is recommended that countries use the numerical coding system presented in *Standard Country or Area Codes for Statistical Use*.[102] The use of standard codes for classification of the foreign-born population according to the country of birth will enhance the usefulness of such data, including an international exchange of foreign-born population statistics among countries. The option "Unknown" should be added to the codes for country of birth. If countries decide to combine countries into broad groups, it is recommended that the standard regional and subregional classifications identified in the above-mentioned publication be adopted.

4.108. Countries with a significant number of international migrants may wish to collect information on the **country of birth of parents** (both father and mother), in which case the information should be asked of all respondents following the same guidelines given for country of birth. The decision to collect and disseminate information on country of birth of parents in a census is dependent upon a number of considerations and national circumstances, including for example the suitability and sensitivity of asking such a question that relates to persons who may not be in the country in which the census is taking place.

4.109. Information on the country of birth of parents can be used, in combination with information on the country of birth of the enumerated person, to identify native-born children of the foreign-born population (the so-called "second generation") and to study the integration processes and outcomes of migrants and their descendants. Moreover, in countries that have experienced return migration, information from this topic allows the identification of foreign-born children of native-born parents.

2. Country of citizenship (core topic)

[102] United Nations, *Standard Country or Area Codes for Statistical Use*, http://unstats.un.org/unsd/methods/m49/m49.htm.

4.110. *Country of citizenship* is defined as the country an individual is a citizen of and with which the individual enjoys a particular legal bond, acquired by birth, naturalization, marriage or some other mechanism. A citizen is a legal national of the country of enumeration; a non-citizen may be a foreign-citizen (that is a citizen of another country), or a stateless person or a person with unknown or undetermined citizenship status. Because the country of citizenship is not necessarily identical to the country of birth, both items should be collected in a census.

4.111. Information on the country of citizenship is particularly important for foreign citizens. It is important to record country of citizenship as such and not to use another concept to indicate citizenship, since some of those concepts may also be used to designate ethnic groups.

4.112. It is essential that the coding of information on country of citizenship be done in sufficient detail to allow for the individual identification of all countries of citizenship that are represented among the foreign population in the country. For purposes of coding, it is recommended that countries use the numerical coding system presented in *Standard Country or Area Codes for Statistical Use*. The use of standard codes for classification of the foreign population by country of citizenship will enhance the usefulness of such data and permit an international exchange of information among countries on their foreign populations. If countries decide to combine countries of citizenship into broad groups, it is recommended that the standard regional and subregional classifications identified in the above-mentioned publication be adopted. Two additional categories should be listed: 1) stateless or no citizenship/nationality status, and 2) unknown citizenship/nationality status.

4.113. The reliability of reported citizenship may be doubtful in the case of persons whose citizenship has recently changed as a result of territorial changes, or among the population of some newly independent countries where the concept of citizenship may have only recently become important. Clear guidelines issued by the national statistical authority can help improve the quality of the data collected. As an aid to the analysis and interpretation of the results, notes on the likelihood of these and other possible causes of misstatement should accompany tabulations based on citizenship.

4.114. Enumeration and processing instructions should provide clear guidance on the treatment of stateless persons, persons with dual nationality, persons in the process of naturalization and any other groups with ambiguous citizenship. The treatment of these groups should be described in the census reports and included in the metadata for accompanying tabulations.

4.115. In cases where people have more than one citizenship and where this information is useful for decision-making, details may be collected on whether the person holds one or multiple citizenship. If this information is to be published, care should be taken to explain how the possibility of people being included in the table more than once affects the marginal totals on the table. Usually, however, it may be more practicable for tabulations by citizenship to refer to one citizenship only. Thus, persons with multiple citizenships should be allocated to a single “primary” citizenship, for example by giving precedence to the citizenship of the “home” country.

3. Acquisition of citizenship

4.116. In addition to collecting information on citizenship, for countries where the population includes a significant proportion of naturalized citizens it may be important to collect information on the method of acquisition of citizenship so as to enable the classification of the population into (a) citizens by birth; (b) citizens by naturalization whether by declaration, option, marriage or other means; and (c) non-

nationals. In such countries it may also be useful to ask questions on previous citizenship and year of naturalization.

4. Year or period of arrival in the country (core topic)

4.117. *Year or period of arrival in the country* refers to the calendar year and month of arrival of a foreign-born person to the country of enumeration. This information enables the calculation of the number of completed years between the time of arrival in the country and the time of enquiry, usually the census date. Information on the month and year of arrival also provides the flexibility of classifying foreign-born persons by period of arrival in terms of any prespecified period, such as 1975–1979, 1980–1984 and so forth. It is thus recommended that the period of arrival be shown, in any tabulation in which the variable appears, in terms of the actual year of arrival.

4.118. It is possible to collect information on either the date of first arrival in the country or the date of the most recent arrival in the country. Each has its own advantages and disadvantages. In making the choice of which information to collect, countries should be guided first and foremost by their policy and user needs.

4.119. Information on time since arrival can also be collected by asking how many years have elapsed since the time of arrival, instead of in what calendar year and month the person arrived. However, use of such a question is not recommended because it is likely to yield less accurate information.

4.120. Note that information on the year and month of arrival is focused mainly on persons born outside the country of enumeration, that is to say, persons who must have arrived in that country at some time after their birth. However, it should be noted that the phenomenon of “international return migration” is becoming increasingly common, and countries that have population groups that maintain links to other countries, migrating to or from another country at different life stages (for example, as students or pensioners), may have an interest in collecting information on returning migrants: in this case, the question on year and month of arrival could also be asked of native-born respondents who have ever lived in another country. In addition, it might also be important to collect information on previous country of residence for persons who have ever lived abroad.

[Suggested New Topic] - Reason for international migration

(From the Handbook on International Migration) Reason for migration is a topic of primary interest to Governments. In the absence of a specialized sample survey on immigrants – the vehicle best suited to eliciting this information – a number of countries have adopted a more modest approach to obtaining this information through their census.

The question to be asked in this regard would be on “reason for coming or returning to the country”. It is typically addressed to persons who have ever lived abroad for at least 12 months. While it is apparent that in some countries, where the target respondents are returning migrants, in others the range of target respondents appears to be broader so as to include foreign-born persons.

This topic may be very complex and challenging for 4 main reasons:
a) The reasons tend to overlap and are generally not mutually exclusive. Example: environmental drivers might overlap with reasons related to security or social and economic factors. The categorization proposed in the Handbook, while useful, is not complete (missing for ex. the category "as an

accompanying spouse or family member") and it is not fully mutually exclusive (see category "Settlement"). We might need to refine it a bit more if we decide to use it. Further, requiring respondents to choose one "main" reason will not improve the utility or quality of responses, in my view.
b) The reason for migration can be considered from different perspectives (ex. immigrants vs emigrants; stocks vs. flows) and can include the reason for admission, the purpose for immigrating/emigrating; the purpose of staying, etc. The question being proposed by UNFPA (from the Handbook para 316) is addressed to persons who have ever lived abroad for at least 12 months, that is who have "ever migrated" and refers to both foreign-born/foreign citizens as well as "returning" native-born persons/citizens. It does not explicitly specify that it refers to the most recent migration spell, as per wording "What was the main reason person XX came to live in this country?". If the intention of the group is to focus only on the most recent spell, the question might need to be reformulated and made more explicit.
c) The time frame for the reason of migration is important since it will likely change over the life course of each person. The reason to "stay", may be different from the purpose of arrival, while the reason to emigrate might be different from that to return. However, if we decide to focus on the most recent migration spell only, this will have important implications. Because the reason for migration is often closely related to age at the time of migration, we would only capture a range of those changes for individuals who have experienced multiple migration spells (either went back and forth between their country of birth/citizenship and country of enumeration or migrated to multiple countries). Therefore, migrants whose reasons change (ex. from "Education and training" to "Marriage", or from "Education and training" to "Employment") while they are abroad would only be reflected if they have additional spells that coincide with these changes.
c) Many countries with large migrant stocks have flow data that capture the reason for admission/type of permit granted, including for asylum seekers and refugees. How will this question add to that information? In countries where the stock of international migrants is small relative to the size of their population is this type of detailed question useful?

This is also a recommendation from EGRISS IRSS (para. 155) The categories in the EGRISS Compiler's Manual (Box A.2, pp. 20) for Reason for (international) migration, with response categories:
a. Employment (including military service)
b. Education and training
c. Marriage, family reunification or family formation
d. Forced displacement (refugees, asylum seekers, temporary protected status, others)
e. Other
The reason should refer to the main reason for the most recent move.

[Suggested New Topics]

- Country of residence five years prior to the census
- Country of residence one year prior to the census
- Ever resided abroad

C. Household and family characteristics

4.121. In considering the topics related to household characteristics, it is important to be aware of the differences between the concepts of household and family as used herein.

4.122. A *household* may be either:

(a) A one-person household, that is to say, a person who makes provision for his or her own food or other essentials for living without combining with any other person to form part of a multiperson household; or

(b) A multiperson household, that is to say, a group of two or more persons living together who make common provision for food or other essentials for living. The persons in the group may pool their resources and have a common budget; they may be related or unrelated persons or a combination of persons both related and unrelated. This arrangement exemplifies the "housekeeping" concept.

Some countries use a concept different from the housekeeping concept, namely, the "household dwelling" concept, which regards all persons living in a housing unit as belonging to the same household. According to this concept, there is one household per occupied housing unit. Therefore, the number of occupied housing units and the number of households occupying them are equal and the locations of the housing units and households are identical. Countries should specify in their census reports whether they used the "housekeeping" or the "household dwelling" concept of a private household.

4.123. A household may be located in a housing unit (see paragraph 4.427) or in a set of collective living quarters such as a boarding house, a residential hotel or a camp, or may comprise the administrative personnel in an institution.

4.124. The *family* within the household, a concept of particular interest, is defined as those members of the household who are related, to a specified degree, through blood, adoption or marriage. The degree of relationship used in determining the limits of the family in this sense is dependent upon the uses to which the data are to be put and so cannot be established for worldwide use. See paragraph 4.141 for a definition of the family nucleus.

4.125. Although in practice most households are composed of a single family consisting of a married couple without children or of one or both parents and their children, it should not be assumed that this identity always exists; census tabulations should therefore clearly indicate whether they relate to households or to families within households.

4.126. From the definitions of "household" and "family", it is clear that these are different concepts that cannot be used interchangeably in the same census. The differences between the household and the family are that:

a) A household may consist of only one person but a family must contain at least two members;

b) The members of a multiperson household need not be related to each other, while the members of a family must always be related.

4.127. A household can contain more than one family, or one or more families together with one or more non-related persons, or it can consist entirely of non-related persons. A family typically will not comprise more than one household. However, the existence of polygamous families in some countries, as well as shared child custody and support arrangements in others, means that individual countries should decide how best to derive and report data on families.

4.128. It is recommended that the household be used as the unit of enumeration (as defined in paragraphs 2.33–2.37) and that the family be a derived topic only. The place of usual residence is recommended as the basis for assigning persons to households where they normally reside. Where the de facto approach is used as the method of enumeration (see paragraphs 2.55–2.63), household lists should, where feasible, also include usual residents temporarily absent. The place of usual residence is where a person usually resides and it may or may not be the person's current or legal residence. The latter terms are usually defined in the laws of most countries and need not correspond to the concept of place of usual residence, which, as employed in the census, is based on conventional usage. In published reports, countries should indicate whether or not household information refers to usual residents and also what the time limits are in respect of being included or excluded as a usual resident. For a more detailed discussion on the difficulty of collecting information on place of usual residence, see paragraphs 4.52–4.57.

1. Relationship to the reference person of household (core topic)

4.129. In identifying the members of a household (as defined in paragraphs 4.122–4.123), it is useful to identify first the household reference person and then the remaining members of the household according to their relationship to the reference person. Countries may use the term they deem most appropriate to identify this person (household reference person, head of household, householder, among others) as long as the person so identified is used solely to determine relationships between household members. It is recommended that each country present, in published reports, the concepts and definitions that are used.

4.130. With respect to selecting the household reference person, it is important to specify criteria for choosing that person in relation to whom household members would be best distinguished, especially in polygamous, multi-family and other households, such as those composed only of siblings without a parent and those composed entirely of unrelated persons. This information should be included in training materials and instructions to enumerators.

4.131. The traditional notion of head of household assumes that most households are family households (in other words, they consist entirely, except possibly for domestic servants, of persons related by blood, marriage or adoption) and that one person in such family households has primary authority and responsibility for household affairs and is, in the majority of cases, its chief economic support. This person is then designated as the head of household.

4.132. Where spouses consider themselves to be equal in household authority and responsibility and may share the economic support of the household, the concept of head of household is no longer considered valid even for family households. In order for the relationship among members of the household to be determined under these circumstances, it is essential that either:

(a) The members of the household designate one among them as a reference member with no implication of headship; or
(b) Provision be made for designation of joint headship where desired.

In any case, it is important that clear instructions be provided in the census as to how this situation is to be handled.

4.133. Even in countries where the traditional concept of head of household is still relevant, it is important to recognize that the procedures followed in applying the concept may distort the true picture, particularly with regard to female heads of households. The most common assumption that can distort the facts is that no woman can be the head of any household that also contains an adult male. Enumerators and even respondents may simply take such an assumption for granted.

4.134. This common sex-based stereotype often reflects circumstances that may have been true in the past but are true no longer, insofar as the household and economic roles of women continue to change. It is therefore important that clear instructions be provided as to who is to be treated as the reference person of the household so as to avoid the complications of enumerator or respondent preconceptions on the subject and the bias that such preconceptions may create. The procedure to follow in identifying a reference person when the members of the household are unable to do so should be clear and unambiguous and should avoid sex-based bias. Where alternative definitions are used, this should be made explicitly in the census questionnaire and in the tabulated census results.

4.135. The selection of one reference person in a household to whom all other persons in the household report, or designate, their relationship requires careful consideration. In the past the person considered to be the "head of the household" was generally used as the reference person, but this concept is no longer considered appropriate in many countries. It has also sometimes been proposed that the person designated as the reference person should be the oldest person in the household or the one who contributes the most income. However, given that the primary purpose of the question is to assign family status and to assign individuals into families, both of these approaches have weaknesses. The automatic selection of the oldest person may be undesirable because in multigenerational households many explicit kin relationships can be reported where the reference person is selected from the middle generation. Similarly, the selection of the person with the highest income may be a person who will not solicit the broadest range of explicit kin relationships.

Comments by MN: Address the instances of child headed households, or instances where children are left in the care of the housekeeper/nanny, where the housekeeper is treated as their own household, even where they share same dwelling as the children.

4.136. Given below is some guidance on the selection of the reference person, which will yield some explicit kin relationships:

(a) Either the husband or the wife of a married couple living in the household (preferably from the middle generation in a multigenerational household);
(b) Either partner of a consensual union couple living in the household where there is no married couple present (where applicable);
(c) The parent, where one parent lives with his or her sons or daughters of any age; or
(d) Where none of the above conditions apply, any adult member of the household may be selected.

Note that these categories are neither comprehensive nor mutually exclusive.

4.137. After identification of the reference member of the household, each of the remaining members of the household should be distinguished in relation to that person, as appropriate, as one of the following:

(a) Spouse,
(b) Partner in consensual union (cohabiting partner), where applicable,
(c) Child,
(d) Spouse of child,
(e) Grandchild or great-grandchild,
(f) Parent (or parent of spouse),
(g) Other relative,
(h) Domestic employee, or
(i) Other person not related to the head or other reference member.

Where this classification is considered too detailed for successful collection of the information, categories (f) and (g) may be consolidated as "Other relative" and (h) and (i) can be consolidated as "Other unrelated person".

4.138. As an aid to the identification of family nuclei (as defined in paragraphs. 4.141–4.143) within the household, it might be helpful if persons were recorded on the census questionnaire to the extent possible in the order of nuclear relationship. Thus, the first person entered after the head or other reference person would be the spouse of that person, followed by unmarried children and then by married children, their spouses and children. For polygamous households, the order of entry could be such that each wife and her unmarried children appeared in succession.

4.139. For estimating fertility by the *own children* method (see paragraph 4.217), the natural mother of each child under 15 years of age should be identified if she appears in the same questionnaire as her child. One way of doing this is to provide the line number of the mother alongside that of the child, if both are living in the same household. The information is not relevant for stepchildren, adopted children or foster children under permanent or temporary care.

4.140. In order to meet increased data needs on households and families, countries may wish, while conducting their population censuses, to collect more detailed information on relationships. In households where the relationship structure is complex, including those with foster children, obtaining accurate information on the relationships between household members may be difficult. Some countries may supplement information on relationship to the reference person of the household with information on direct relationships between household members by, for instance, relating a child to its parents even when neither parent is the reference person of the household. Enumerators should be encouraged to probe for a clear relationship (such as child, niece or aunt). The recording of non-specific responses such as "relative" should be avoided. It is recommended that specific guidance be provided on acceptable responses, that relationships be specified completely in the census questionnaire, and that any precoded categories used should be sufficiently detailed to produce desired outputs.

2. Household and family composition (core topic)

4.141. Household and family composition can be examined from different points of view, but for census purposes it is recommended that the primary aspect considered should be that of the family nucleus.

4.142. A *family nucleus* is of one of the following types (each of which must consist of persons living in the same household):

(a) A married couple without children;
(b) A married couple with one or more unmarried children;
(c) A father with one or more unmarried children;
(d) A mother with one or more unmarried children.[103]

Couples living in consensual unions may be regarded as constituting a family nucleus.

4.143. The concept of family nucleus as defined above limits relationships between children and adults to direct (first-degree) relationships, that is to say, between parents and children. In some countries, numbers of skip-generation households, that is to say, households consisting of grandparent(s) and one or more grandchildren with no parent of those grandchildren present, are considerable. Therefore, countries may include such skip-generation households in their family nucleus definition, and the census report should clearly state whether or not skip-generation households are included in the family nucleus definition.

4.144. The family nucleus is identified from the answers to the question on relationship to the reference member of the household, supplemented where necessary by information on name and marital status. The identification of offspring and their mother and the order in which persons are entered in the questionnaire may be of additional assistance in this respect. The identification of family nuclei is likely to be more complete in de jure than in de facto enumerations, because the latter do not take account of temporarily absent household members who may constitute part of a family nucleus.

4.145. For census purposes, a child is any unmarried individual, regardless of age, who lives with his or her parent(s) and has no children in the same household. Consequently, the definition of a child is primarily a function of an individual's relationship to other household members, regardless of age. In accordance with this definition, a household consisting of a married couple with two never-married children, divorced son, and a married daughter and her husband would be considered to be composed of two family nuclei, with the divorced child being regarded as a member of the parents' family and the married daughter and son-in-law as a second family. As used here, the term "child" does not imply dependency, but rather is used to capture household living arrangements of persons who are in a parent-child relationship. Countries need to be clear in their metadata how they treat foster and adopted children.

4.146. The family nucleus does not include all family types, such as brothers or sisters living together without their offspring or parents, or an aunt living with a niece who has no child. It also excludes the case of a related person living with a family nucleus as defined above, for example, a widowed parent living with her married son and his family. The family nucleus approach does not, therefore, provide information on all types of families. Countries may extend the investigation of families beyond that of the family nucleus, in accordance with their own interests.

4.147. Households should be classified by type according to the number of family nuclei they contain and the relationship, if any, between the family nuclei and the other members of the household. The relationship should be through blood, adoption or marriage, to whatever degree is considered pertinent

[103] In countries where a different definition of family nucleus is used, it should be clearly stated in the census report.

by the country (see paragraph 4.139). Given the complexity of this item, it is important that information on relationship to the household reference person be properly processed. The types of household to be distinguished could be:

(a) *One-person household*;
(b) *Nuclear household*, defined as a household consisting entirely of a single family nucleus. It may be classified into:
  (i) Married couple family:
    a. With child(ren);
    b. Without child(ren);
  (ii) Partner in consensual union (cohabiting partner):
    a. With child(ren);
    b. Without child(ren);
  (iii) Father with child(ren);
  (iv) Mother with child(ren);
(c) *Extended household*, defined as a household consisting of any one of the following:[104]
  (i) A single family nucleus and other persons related to the nucleus, for example, a father with child(ren) and other relative(s) or a married couple with other relative(s) only;
  (ii) Two or more family nuclei related to each other without any other persons, for example, two or more married couples with (or without) child(ren) only;
  (iii) Two or more family nuclei related to each other plus other persons related to at least one of the nuclei, for example, two or more married couples with other relative(s) only;
  (iv) Two or more persons related to each other, none of whom constitutes a family nucleus;
(d) *Composite household*, defined as a household consisting of any of the following: [105]
  (i) A single family nucleus plus other persons, some of whom are related to the nucleus and some of whom are not, for example, mother with child(ren) and other relatives and non-relatives;
  (ii) A single family nucleus plus other persons, none of whom is related to the nucleus, for example, father with child(ren) and non-relatives;
  (iii) Two or more family nuclei related to each other plus other persons, some of whom are related to at least one of the nuclei and some of whom are not related to any of the nuclei, for example, two or more couples with other relatives and non-relatives only;
  (iv) Two or more family nuclei related to each other plus other persons, none of whom is related to any of the nuclei, for example, two or more married couples one or more of which with child(ren) and non-relatives;
  (v) Two or more family nuclei not related to each other, with or without any other persons;
  (vi) Two or more persons related to each other but none of whom constitute a family nucleus, plus other unrelated persons;
  (vii) Non-related persons only;
(e) Other;
(f) Unknown or not stated.

4.148. In the census tabulations, all countries should at least distinguish between one-person, nuclear, extended and composite households. Where feasible, some or all of the subcategories shown above should also be distinguished, although countries may find it appropriate to modify the classification

[104] The subdivisions in this category should be modified to suit national circumstances.
[105] The subdivisions in this category should be modified to suit national circumstances.

according to national circumstances. For example, in countries where almost all households contain only one family nucleus at most, the distinction between nuclear, extended and composite households may be applied only to households containing one nucleus or no nucleus; multinuclear households may then be shown as an additional category without any further classification by type. In countries where multinuclear households are comparatively common, further breakdowns of extended and composite households, distinguishing between those with three, four or more family nuclei, may be helpful.

3. Household and family status

4.149. For purposes of determining household and family status and identifying how a person relates to other household or family members, persons may be classified according to their position in the household or family nucleus. Classifying persons according to household and family status has uses in social and demographic research and policy formulation. Census data could be presented according to both household and family status for a variety of purposes. Although status itself is based on information derived from responses to the item on relationship to the head or other reference member of the household and other items, the classification of persons by their household and family status is a different approach from the traditional one of classifying household members solely according to their relationship to the head or reference person. The following household and family status classifications illustrate how such an approach may be used.[106] Care should be taken at the planning stages to relate this item to the classification of households by type as recommended in paragraph 4.146.

Persons living in households are classified by household status as:

1 Person in a household with at least one family nucleus
- 1.1 Married spouse
- 1.2 Partner in consensual union (cohabiting partner)
- 1.3 Lone mother[107]
- 1.4 Lone father[108]
- 1.5 Child living with both parents
- 1.6 Child living with lone mother
- 1.7 Child living with lone father
- 1.8 Not a member of a family nucleus
  - 1.8.1 Living with relatives
  - 1.8.2 Living with non-relatives

2 Person in a household with no family nucleus
- 2.1 Living alone
- 2.2 Living with others[109]
  - 2.2.1 Living with sibling(s)
  - 2.2.2 Living with other relatives
  - 2.2.3 Living with non-relatives

[106] To date, only the population and housing census recommendations for the Economic Commission for Europe region contain household and family status classifications.
[107] Person living with children, without spouse.
[108] Person living with children, without spouse.
[109] The subdivisions in this category should be modified to suit national circumstances.

Persons are classified by family status as:[110]

1 Spouse
- 1.1 Husband
  - 1.1.1 With child(ren)
  - 1.1.2 Without a child
- 1.2 Wife
  - 1.2.1 With child(ren)
  - 1.2.2 Without child

2 Lone parent
- 2.1 Male
- 2.2 Female

3 Child
- 3.1 With both parents
- 3.2 With lone parent
  - 3.2.1 With lone father
  - 3.2.2 With lone mother

4 Not member of a family nucleus
- 4.1 Relative of husband or wife
  - 4.1.1 Parent of husband or wife
  - 4.1.2 Sibling of husband or wife
  - 4.1.3 Other relative of husband or wife
- 4.2 Non-relative

D. Demographic and social characteristics

4.150. Sex and age are considered to be the most basic of all demographic variables. Of all the topics included in population censuses, *sex* and *age* are more frequently cross-classified with other characteristics of the population than are any other topics. Apart from the importance of the age-sex structure of the population in itself, accurate information on the two topics is fundamental to the great majority of the census tabulations. A very important use of census data on the sex and age composition of the population is the evaluation of the data, especially with respect to coverage. The variables are therefore very crucial and it is important that this information be reported in respect of every person for whom census information has been collected. It is therefore recommended that where this information is incomplete it should be imputed for census purposes rather than being reported as "not stated". Possible difficulties in securing accurate age data are often not recognized because the topic appears to be a simple one. The difficulties associated with this topic are therefore highlighted in paragraphs 4.151–4.162 below.

1. Sex (core topic)

4.151. The sex of every individual should be recorded on the census questionnaire for those countries that collect their census information in this way. The disaggregation of data by sex is a fundamental requirement for gender statistics. For many socioeconomic and demographic characteristics that could be collected through a census, such as education, economic activity, marital status, migration, disability and living arrangements, there are generally variations by sex. The successful planning and

[110] The subdivisions in this category should be modified to suit national circumstances.

implementation of gender-sensitive policies and programmes requires the disaggregation of data by sex to reflect problems, issues and questions related to both men and women in society. Sex, together with age, represents the most basic type of demographic information collected about individuals in censuses and surveys, as well as through administrative recording systems, and the cross-classification of these data with other characteristics forms the basis of most analyses of the social and demographic characteristics of the population, as it provides the context within which all other information is placed.

Comment/suggestion by MN:
We propose adjustment of response categories to incorporate the LGBTQ, however countries need to be guided on testing and structuring these questions as any change to sex question might affect the population structure, for sampling age & sex in the intercensal years.

[Suggested New Topic] Gender Identity

The following is a draft by NM:

Gender Identity
In addition to sex, countries may want to measure gender identity. The terms gender and sex are often used interchangeably, but they are different concepts and should be measured separately. Gender differs from sex in that sex refers to the biological attributes that are used to label individuals as male, female, or intersex (FCSM, 2016). Gender, on the other hand, refers to the socially constructed characteristics of women and men, which includes norms, behaviors and roles associated with being a woman or a man. (FCSM, 2016) Gender identity refers to "a person's deeply felt, internal and individual experience of gender, which may or may not correspond to the person's physiology or designated sex at birth." (WHO, 2016) Gender identity should not be confused with sexual orientation. Sexual orientation pertains to "a person's physical, romantic and/or emotional attraction (or lack thereof) towards other people" (WHO, 2016) and is separate from gender identity.

The decision to include gender identity questions in a census should begin with considering how resulting data will be used. As with the addition of any question to a census questionnaire, the census program must ensure that the needed data are collected without undue burden to the respondents. Data should only be collected if there are plans to disseminate the data. To minimize burden and privacy risk to respondents, only the minimum amount of information needed to meet the planned uses should be collected. Also, the ability to protect confidentiality when collecting data should be considered. (Federal Committee on Statistical Methodology, 2023)

Concepts of gender identity vary greatly across the world. The terminology used to describe gender identity and the way people identify with those concepts is also fluid and evolving. It is important to test terminology with a broad range of respondents or rely on previously tested terminology. Similarly, when translating the terminology to multiple languages, the translations should be tested to convey the intended meaning. (Federal Committee on Statistical Methodology, 2023)

2. Age (core topic)

4.152. *Age* is the interval of time between the date of birth and the date of the census, expressed in completed solar years. Every effort should be made to ascertain the precise and accurate age of each person, particularly of children and older persons.

4.153. Information on age may be secured either by obtaining the date (year, month and day) of birth or by asking directly for age at the person's last birthday.

4.154. The first method yields more precise information and should be used whenever circumstances permit. It also allows for the calculation of age at reference dates other than census day for the purposes, for example, of deriving annual census-based mid- or end-year population estimates. If neither the exact day nor even the month of birth is known, an indication of the season of the year can be substituted if this information can be easily recorded. The question on date of birth is appropriate wherever people know their birth date, whether in accordance with the solar calendar or a lunar calendar, or whether years are numbered or identified in traditional folk culture by names within a regular cycle. It is extremely important, however, that there should be a clear understanding between the enumerator and the respondent about which calendar system the date of birth is based on. If there is a possibility that some respondents will reply with reference to a calendar system different than that of other respondents, provision must be made in the questionnaire for noting the calendar system that has been used. It is not advisable for the enumerator to attempt to convert the date from one system to another. The necessary conversion can be best carried out as part of the data-editing work.

4.155. Where the information is taken from administrative data sources, date of birth is usually more accurately recorded.

4.156. The direct question on age is likely to yield less accurate responses for a number of reasons. Even if all responses are based on the same method of reckoning age, there is the possibility of a misunderstanding on the part of the respondent as to whether the age wanted is that at the last birthday, the next birthday or the nearest birthday. In addition, asking a direct question on age can result in occurrences, with comparative ease, of rounding to the nearest age ending in zero or five, providing estimates not identified as such and deliberate misstatements. Difficulties may arise in the reporting or in the recording of the information for children under 1 year of age, which may be given erroneously as "1 year of age" rather than "zero years of age". These difficulties may be mitigated by collecting information on the date of birth of all children reported as "1 year of age", while using only the direct age question for the remainder of the population. Another possible approach is to obtain age in completed months for children under 1 year of age. This method, however, can give rise to another type of recording error, that is to say, the substitution of years for months, so that a 3-month-old child, for example, might be entered in the questionnaire as being 3 years of age.

4.157. Some countries have made improvements in the quality of age data by asking both questions on age and date of birth.

4.158. An additional complication may occur with the use of the direct question if more than one method of calculating age is in use in the country. In some countries, certain segments of the population may use an old traditional method whereby persons are considered to be 1 year of age at the time of birth and everyone advances 1 year in age at the same fixed date each year. Other segments of the population in the same countries may use the Western method, in which a person is not regarded as being 1 year of age until 12 months after the date of birth, and advances 1 year in age every succeeding 12 months. If there is a risk of different methods of age calculation being used by respondents, provision must be made to ensure that the method used in each case is clearly indicated in the questionnaire and that the conversion is left to the data-editing stage.

4.159. In spite of its drawbacks, the direct question on age is the only one that should be used when people cannot provide even a birth year. As regards persons for whom information on age is unavailable or appears to be unreliable, an estimated age may have to be recorded. This may occur in isolated cases in societies where knowledge of age is widespread or in general in cultures where there is little awareness of individual age and no interest in it. In the latter circumstances, criteria for making estimates should be provided in the instructions for the enumerators.

4.160. One of the techniques that have been used to aid enumerators consists in providing them with calendars of historic events of national or local significance to be used either in probing questions or in identifying the earliest event the respondent recalls. Another technique consists in pre-identifying locally recognized age cohorts in the population and then asking about membership in the cohorts. Enumerators may also ask if the person in question was born before or after other persons whose ages have been roughly determined. Furthermore, use can be made of age norms for weaning, talking, marriage and so forth. Whatever techniques are used, enumerators should be impressed with the importance of securing age data that are as accurate as possible within the amount of time that they can devote to the topic.

4.161. In view of the possible difficulties in the collection of age data, census tests should be used, as appropriate, to determine the difference in results with the use of a question on age as compared with a question on date of birth, what calendar or method of age reckoning most people use, and in what parts of the country age will have to be estimated for the majority of the population and what techniques to use as an aid in estimation. Testing of the calendar or method of age reckoning that most people use is particularly important where an official change from one calendar or method of reckoning to another calendar or method has taken place recently enough so that the new calendar or method of reckoning may not yet be in popular use among some or all of the population.

4.162. Enumerators who are likely to be called upon to estimate age in a substantial number of cases should be given training in the applicable techniques as part of their general training.

4.163. As noted in paragraph 4.149 it is recommended that where this information is incomplete it should be derived or imputed for census purposes rather than being reported as “not stated”.

3. Marital status (core topic)

4.164. Despite the changing nature of marriage, marital status remains a useful demographic variable. The direct relationship between marriage and fertility is still recognized, as is the indirect relationship with other demographic, social and economic characteristics. Numerous variations exist in many countries but it is important that marriage be defined in terms of the laws and customs of individual countries.

4.165. *Marital status* is the personal status of each individual in relation to the marriage laws or customs of the country. The categories of marital status to be identified should at least include the following:

(a) Single (in other words, never married);
(b) Married;
(c) Married, but separated;
(d) Widowed and not remarried;
(e) Divorced and not remarried.

4.166. In some countries, category (b) may require a subcategory of persons who are contractually married but not yet living as man and wife. In all countries, category (c) should comprise both the legally and the de facto separated, who may be shown as separate subcategories if desired.

4.167. In some countries, it will be necessary to take into account customary unions, such as registered partnerships and consensual unions, which are legal and binding under law. In countries with legal provision for registered or legal partnership (for opposite-sex couples or same-sex couples), or where same-sex couples can legally marry, subcategories may either be included in the category (b) Married or in a legally registered partnership, namely (b)(i) "Opposite-sex marriage/partnership", (b)(ii) "Same-sex marriage/partnership".

4.168. The treatment of persons whose only or latest marriage has been annulled is dependent upon the relative size of this group in the country. Where its size is substantial, the group could constitute an additional category; if its size is insignificant, however, the individuals in the group should be classified according to their marital status before the (annulled) marriage took place.

4.169. At times countries have experienced difficulties in distinguishing between (a) formal marriages and de facto unions and (b) persons legally separated and those legally divorced. If either of these circumstances necessitates a departure from the recommended classification of marital status, the composition of each category shown in the tabulations should be clearly stated.

4.170. If complete information on marital status is needed, then this information should be collected and tabulated for persons of all ages, irrespective of the national minimum legal age, or the customary age for marriage, because the population may include persons who were married in another country with a different minimum marriage age. In most countries, there are also likely to be persons who were permitted to marry below the legal minimum age because of special circumstances. In order to permit international comparisons of data on marital status, however, any tabulations of marital status not cross-classified by detailed age should at least distinguish between persons under 15 years of age and those 15 years of age and over.

4.171. The collection of additional information related to customs in particular countries (such as concubinage, polygamous or polyandrous marital status, or inheritance of widows) may be useful in meeting national needs. For example, at times countries may wish to collect data on the number of spouses of each married person. Modifications of the tabulations to take account of such information should be made within the framework of the basic classification in order to maintain international comparability as far as possible.

4.172. The concept of marital status and the marital status categories described above should not be confused with the concept of de facto union status, which describes extralegal unions (including some consensual unions) of varying degrees of stability common in some countries. It should be recognized also that these marital status categories do not adequately describe the prevalence of formal legal marriage combined with the relatively stable de facto union that may exist outside the marriage. Information on these relationships is very useful in studies of fertility, but it is not possible to provide an international recommendation on this matter because of the different circumstances prevailing among countries. It is suggested, however, that countries wishing to investigate these relationships should consider the possibility of collecting separate data for each person on de facto unions and on the duration of each type of union (see paragraphs 4.247–4.248). Information on these relationships can

also be derived from information collected on the relationship to head or reference person or other persons in the household, in order to distinguish between people who are living in either a consensual union or marriage, and those who are not.

[Suggested New Topic] - Marriage registration (non-core topic)

Considerations:

1) To be tested in surveys (Mosidi, Paul) and captured in surveys as done in DHS in Kenya. Marriage registration should not be part of the census, it makes it over-complicated (Vijay).

2) As Marital Status is a core topic, census question on marriage registration may not be best embedded in marital status, but be a separate (non-core) topic (Nobuko, Romesh). In CAPI census space is not an issue. Marital status and marriage registration can be two separate questions. Only the "Married" persons can then answer to marriage registration question.

3) Marriage registration is important for indicators related to women's rights; can be a recommended topic and implementation left to the country [i.e. non-core topic].

4) Some countries considered including marriage registration question(s) in the census recently - South Africa pilot tested in a survey and did not include it in the census, whereas Zambia and Burkina Faso pilot tested and included it in the census questionnaire for the 2020 census round

UNFPA is putting together a review paper of how marriage registration questions have been formulated in past censuses and the data that has resulted. Paper will be circulated and presented at November Expert Group Meeting in NY.

4. Ethnocultural characteristics

4.173. Countries with a culturally diverse population may wish to collect information on the ethnic identity (or composition) of the population, on mother tongue, on the knowledge and practice of languages and on religious communities and denominations. They are all characteristics that allow people the flexibility to express their ethnocultural identity in the way that they choose. Data on such ethnocultural characteristics of the population are of increasing relevance to countries in the context of migration, integration and minority policies.

4.174. Ethnocultural characteristics generally have a subjective dimension, as there is often no common understanding as to what characteristic or concept is really being measured in a particular census. Moreover, different countries will adopt different concepts. Ethnocultural characteristics can also be politically sensitive and may apply to very small, yet identifiable, population subgroups. The free and open declaration of the respondents is therefore of essential importance. Members of certain minority groups may be particularly vulnerable to discrimination on the grounds of ethnic group or religion. Special care, therefore, may be required in census procedures and outputs relating to ethnic group and religion in order to demonstrate to respondents that appropriate data protection and disclosure control measures are in place. In some cases, countries may even wish to collect such data on a voluntary basis if this is permitted by national legislation.

5. Religion

4.175. Each country that investigates religion in its census should use the definition most appropriate to its needs and should display the definition that has been used as part of the metadata in the census publications and dissemination programme.

4.176. For census purposes, *religion* may be defined as either:

(a) Religious or spiritual belief of preference, regardless of whether or not this belief is represented by an organized group; or

(b) Affiliation with an organized group having specific religious or spiritual tenets.

4.177. The decision to collect and disseminate information on religion in a national census is dependent upon a number of considerations and national circumstances, including, for example, the national needs for such data, and the suitability and sensitivity of asking a religion question in a country's census. Owing to the sensitive nature of a question on religion, special care may be required to demonstrate to respondents that appropriate data protection and disclosure control measures are in place. It is important that the responding public be informed of the potential uses and needs for this information.

4.178. The amount of detail collected on this topic is dependent upon the requirements of the country. It may, for example, be sufficient to enquire only about the religion of each person; on the other hand, respondents may be asked to specify, if relevant, the particular sect to which they adhere within a religion. In countries where a large number of sects or denominations exist there will be implications for space on any census questionnaire and implications for data capture, especially in cases where "write-in" responses are required. In an effort to ensure international comparability as far as possible, it is recommended that religion or religious affiliation should be measured directly by a question that asks "What is your religion?" rather than use of a filter question that asks for example "Are you religious?" and if so "What is your religion?" Response categories should include "No religion/religious affiliation" together with a "Religious but prefer no to disclose" or "Not stated" category, in effect making responses to such a question voluntary.

4.179. For the benefit of users of the data who may not be familiar with all of the religions or sects within the country, as well as for purposes of international comparability, the classifications of the data should show each sect as a subcategory of the religion of which it forms a part. A brief statement of the tenets of religions or sects that are not likely to be known beyond the country or region would also be helpful.

6. Language

4.180. There are four types of language data that can be collected in a census, namely:

(a) Mother tongue, defined as the language usually spoken in the individual's home in his or her early childhood;

(b) Main language, defined as the language that the person commands best;

(c) Usual language, defined as the language currently spoken, or most often spoken, by the individual in his or her present home;

(d) Ability to speak one or more designated languages, including the country's official language(s).

4.181. Each of these types of information serves a very different analytical purpose. Each country should decide which, if any, of these types of information is applicable to its own needs. International comparability of tabulations is not a major factor in determining the form of the data to be collected on this topic.

4.182. In compiling data on the usual language or on the mother tongue, it is desirable to show each language that is numerically important in the country and not merely the dominant language.

4.183. Information on language (including any sign language) should be collected for all persons. In the tabulated results, the criterion for determining language for children not yet able to speak should be clearly indicated.

7. Ethnicity

4.184. The decision to collect and disseminate information on ethnic or national groups of a population in a census is dependent upon a number of considerations and national circumstances, including, for example, the national needs for such data, and the suitability and sensitivity of asking ethnicity questions in a country's census. Owing to the sensitive nature of questions on ethnicity, special care may be required to demonstrate to respondents that appropriate data protection and disclosure control measures are in place. It is important that the responding public be informed of the potential uses and need for data pertaining to ethnicity, as this improves public support for the census exercise. Data on ethnicity provide information on the diversity of a population and can serve to identify subgroups of a population. Some areas of study that rely on such data include demographic trends, employment practices and opportunities, income distributions, educational levels, migration patterns and trends, family composition and structure, social support networks, and health conditions of a population.

4.185. Broadly defined, *ethnicity* is based on a shared understanding of history and territorial origins (regional and national) of an ethnic group or community, as well as on particular cultural characteristics such as language or religion. Respondents' understanding or views about ethnicity, awareness of their family background, the number of generations they have spent in a country, and the length of time since immigration are all possible factors affecting the reporting of ethnicity in a census. Ethnicity is multidimensional and is more a process than a static concept, and so ethnic classification should be treated with movable boundaries.

4.186. Ethnicity can be measured using a variety of concepts, including ethnic ancestry or origin, ethnic identity, cultural origins, nationality, race, colour, minority status, tribe, language, religion or various combinations of these concepts. Because of the interpretative difficulties that may occur with measuring ethnicity in a census, it is important that, where such an investigation is undertaken, the basic criteria used to measure the concept are clearly explained to respondents and in the dissemination of the resulting data. The method and the format of the question used to measure ethnicity can influence the choices that respondents make regarding their ethnic backgrounds and current ethnic identification. The subjective nature of the term (not to mention increasing intermarriage among various groups in some countries, for example) requires that information on ethnicity be acquired through self-declaration of a respondent and also that respondents have the option of indicating multiple ethnic affiliations. Data on ethnicity should not be derived from information on country of citizenship or country of birth. The classification of ethnic groups also requires the inclusion of the finest levels of ethnic groups, self-perceived groups, regional and local groups, as well as groups that are not usually considered to be ethnic groups, such as religious groups and those based on nationality.

Countries collecting data on ethnicity should note that the precoding or preclassification of ethnic groups at the time of data capture may have a tendency to lose detailed information on the diversity of a population unless space to record write-in, free-form responses is provided.

4.187. Respondents should be free to indicate more than one ethnic affiliation or a combination of ethnic affiliations if they wish so. Countries should explain in the census instructions and the census documentation how the ethnicity of children from mixed couples is to be reported (for example, explicit instructions to allow respondents to provide multiple responses or to allow for responses such as "Biracial"). Also, to guarantee the free self-declaration of ethnicity, respondents should be allowed to indicate "None" or "Not declared".

4.188. Because the ethnocultural composition of a country can vary widely from country to country and due to the diversity in the approach and the various criteria for establishing ethnicity, it is recognized that there is no single definition or classification that could be recommended that would be applicable to all countries. However, countries should document the basic criteria and classification procedures for ethnicity and inform the data users about the concepts on which they are based.

8. Indigenous peoples

4.189. Facilitating the collection of data on indigenous peoples for national and international needs can serve to improve socioeconomic and active participation of indigenous peoples in the development process for many countries. The sensitive nature of questions pertaining to the indigenous population requires care in assuring the public that the appropriate disclosure and data protection methods are being enforced. The responding public should be informed on the potential uses and need for such data to improve public support for the census exercise.

4.190. Dissemination of census data pertaining to indigenous peoples contributes to research in areas such as the socioeconomic conditions of the indigenous population, trends, causes for inequities, and the effectiveness of existing policies and programmes. Availability of these data can also assist indigenous communities in assessing their conditions of living and give them the information they need to participate and advocate in the development of programmes and policies affecting their communities, such as those impacting health systems, models of economic production, environmental management and social organization. In addition, the development of indicators relevant to the indigenous population and the measurement of such indicators in the data collection process can be used to monitor the human development of indigenous populations.

4.191. Generally, *indigenous peoples* of a particular country are social groups with an identity that is distinct from the social and cultural identity of the dominant society in that country. Questions on indigenous identity should abide by the principle of self-identification. It is important that, where such an investigation is undertaken, multiple criteria are developed to accurately capture identity and socioeconomic conditions of indigenous peoples. Defining the indigenous population can be done in many ways, such as through a question on ethnic origin (that is to say, ancestry) or on indigenous identity. Identifying the indigenous community also requires recognition of the diversity in this subpopulation, including nomadic, semi-nomadic and migrating peoples, peoples in transition, displaced persons, indigenous peoples in urban areas, and particularly vulnerable sects. It is important to point out that there is no single term among countries to describe the indigenous population. Consequently, countries tend to use their own national concepts to identify the indigenous population. For example, in

Australia the terms “aboriginal” or “Torres Strait Islander” are used, while in New Zealand the term “Maori” is used.

4.192. Differing national contexts also imply that enumerating the indigenous population can be done in multiple ways, for example, by way of specific questions on the census form, with specialized questionnaires for the indigenous population, or with follow-up or complementary surveys. In Canada, for example, identification of the indigenous population comes not only from its national census, but also from a post-censal survey. In Australia, in addition to the national census, there is the National Aboriginal and Torres Strait Islander survey, while in Argentina there is a complementary survey after the census targeting indigenous peoples. In addition to a general census, Paraguay also administers a specific census in the same year to identify the indigenous population.

4.193. Involvement of the indigenous community in the data development and data collection processes provides the arena for capacity-building and helps to ensure the relevance and accuracy of the data collection on indigenous peoples. Using local indigenous languages, employing local indigenous people (as interpreters, for example), and training and building the capacity of local indigenous people in data collection processes can facilitate the collection and dissemination of this information. Non-indigenous professionals and technicians should also be informed of the culture and practices of indigenous peoples.

9. Disability characteristics

4.194. A census can provide valuable information on disability in a country. For countries that do not have regular special population-based disability surveys or disability modules in ongoing surveys, the census can be the only source of information on the frequency and distribution of disability and functioning in the population at national, regional and local levels. Countries that have a registration system providing regular data on persons with the most severe types of impairments may use the census to complement these data with information related to selected aspects of the broader concept of disability and functioning based on the International Classification of Functioning, Disability and Health (ICF).[111] Census data can be utilized for general planning programmes and services (prevention and rehabilitation), monitoring selected aspects of disability trends in the country, evaluation of national programmes and services concerning the equalization of opportunities, and international comparison of selected aspects of disability prevalence in countries. While considering measures to determine disability, preferences should be given to legal provisions defining and categorizing disability.

a. Disability status (core topic)

4.195. *Disability status* characterizes the population into those with and those without a disability. Persons with disabilities are defined as those persons who are at greater risk than the general population for experiencing restrictions in performing specific tasks or participating in role activities. This group would include persons who experience limitations in basic activity functioning, such as walking or hearing, even if such limitations were ameliorated by the use of assistive devices, a supportive environment or plentiful resources. Such persons may not experience limitations in specifically measured tasks, such as bathing or dressing, or participation activities, such as working or

[111] Adopted in 2001, the ICF is the international standard for describing and measuring health and disability at both the individual and population levels. More information on the ICF framework is available at: http://www.who.int/classifications/icf/en/.

going to church or shopping, because the necessary adaptations have been made at the personal or environmental levels. These persons would still, however, be considered to be at greater risk of restrictions in activities or participation than the general population because of the presence of limitations in basic activity functioning, and because the absence of necessary accommodations would jeopardize their current levels of participation.

4.196. A comprehensive measure to determine disability would include the following six domains of functioning in a way that can be reasonably measured using a census and that would be appropriate for international comparison:

(a) Walking
(b) Seeing
(c) Hearing
(d) Cognition
(e) Self-care
(f) Communication.

4.197. The first four domains (a) to (d) are to be considered essential in determining disability status. The additional domains (e) and (f) comprise a more comprehensive measure for determining disability.

b. Use of the census to measure disability at the aggregate level

4.198. A census format offers only limited space and time for questions on any one topic such as disability. Since ICF offers several dimensions for use to develop a census measure, it is best to focus on a few of those dimensions, leaving the remaining dimensions for use in more extensive household surveys. Short sets of disability questions, which can be included in censuses and extended sets to be recommended for inclusion in population-based surveys, have been developed and tested by the Washington Group on Disability Statistics.[112] The aim of the recommended sets is to improve comparability of disability and functioning data across countries.

4.199. The definition of disability status (see paragraph 4.194) requires that disability be defined in terms of limitations in basic activity functioning that would place a person at greater risk than the general public of restricted performance of or participation in organized activities (such as educational attendance or work participation). Given the complexity of disability definition and measurement and, in certain cultures, the sensitivity attached to people identifying as having a disability, it is recommended that several functional activity domains be defined whereby people can respond to questions about their difficulty in performing those activities rather than enquiring directly whether or not they have a particular disability.

(i) **Essential domains**

4.200. It is suggested that only those domains that have satisfied a set of selection criteria be eligible for inclusion in a short set of questions recommended for use in censuses. Criteria for inclusion include cross-population or cross-cultural comparability, suitability for self-reporting and space on the census

[112] The Washington Group on Disability Statistics, a United Nations City Group that focuses on proposing international measures of disability, has developed these questions. See www.cdc.gov/nchs/washington_group.htm for updates on the question wording and more information supporting the collection and use of data on disability.

form. Other suggested criteria include the importance of the domain in terms of public health problems. Based on these criteria, the Washington Group on Disability Statistics has developed a Short Set on Functioning (WG SS-F) questions in the six domains for the purpose of measuring disability in a census format. The four basic domains are considered to be essential: walking, seeing, hearing and cognition. In addition, if space permits, two other domains have been identified for possible inclusion: self-care and communication.

4.201. *Walking* fulfils the criteria of cross-cultural applicability and space requirements for comparable data since walking is a good indicator of a central physical function and is a major cause of limitation in participation.

4.202. While *seeing* also represents a public health problem, self-reporting of seeing limitation is more problematic, particularly when individuals use glasses to correct visual impairments. Similar difficulties are associated with asking about *hearing* activity.

4.203. Assistive devices, such as glasses and hearing aids, provide almost complete accommodation for a large proportion of those with impaired functioning. It is often argued that asking about seeing without the use of glasses greatly increases the number of persons with disabilities and makes the group too heterogeneous, that is, the group would include persons at very little risk of participation problems along with those at greater risk. As a result, questions on difficulty seeing or hearing should be asked with the use of glasses or hearing aids if they are usually worn.

4.204. Of the four essential domains, *cognition* is the most difficult to operationalize. Cognition includes many functions such as remembering, concentrating, decision-making, understanding spoken and written language, finding one's way or following a map, doing mathematical calculations, reading and thinking. Deciding on a cross-culturally similar function that would represent even one aspect of cognition is difficult. However, remembering and concentrating or making decisions would probably serve the cultural compatibility aspects the best.

(ii) **Additional domains**

4.205. In additional to the four essential domains, two other have been identified for possible inclusion: self-care and communication. The self-care domain is intended to identify persons who have some problems with taking care of themselves independently. Washing and dressing represent self-care tasks that occur on a daily basis and are considered to be basic activities.

4.206. The purpose of the communication domain is to identify persons who have some problems with talking, listening or understanding speech such that it contributes to difficulty in doing their daily activities. Two aspects of communication are considered: understanding others (receptive communication) and being understood by others (expressive communication). Communicating (understanding and being understood) refers to the exchange of information or ideas between two people through the use of language.

4.207. Beyond the six domains identified above, there are further physical functioning domains that could be included in a set of census questions depending on the space available. The Washington Group Extended Set on Functioning (WG ES-F) includes questions that address functioning in domains such as upper body (functioning of the arms, hands and fingers), affect (anxiety and depression), pain and fatigue.

Comments/suggestions by MN: For countries without administrative data on disability the current domains are limited. For instance, identifying those with mental disability with current set of questions is a challenge. The module is also not practical to apply on children, because the disability indicators for children become exaggerated with these questions. Maybe additional questions for children 0-4 years can be created.

c. Census question wording

4.208. It is recommended that special attention be paid in designing census questions to measure disability. The wording and the construct of questions greatly affect the precision in identifying persons with disabilities. Each domain should be asked through a separate question.[113] The language used should be clear, unambiguous and simple. Negative terms should always be avoided. The disability questions should be addressed to each single household member and general questions on the presence of persons with disabilities in the household should be avoided. If necessary, a proxy respondent can be used to report for the family member who is incapacitated. The important thing is to account for each family member individually rather than ask a blanket question. Scaled response categories can also improve the reporting of disability. The census questions on disability endorsed by the Washington Group include four response categories:

(a) No (meaning no difficulty at all);
(b) Yes – some difficulty;
(c) Yes – a lot of difficulty;
(d) Cannot do [the activity] at all.

Disability prevalence is determined based on any response that is "a lot of difficulty" or "cannot do at all" for any of the questions.

4.209. The information that results from measuring disability status (see paragraph 4.194) is expected to:

(a) Represent a large proportion of, but not all, persons with limitation in basic activity functioning in any one country (only the use of a wider set of domains would potentially cover close to all such persons, but as stated this would not be possible in a census context);
(b) Represent the most commonly occurring basic activity limitations within any country;
(c) Capture persons with similar problems across countries.

4.210. The questions identify the population with limitations in basic activities that have the potential to limit independent participation in society. The intended use of these data would be to compare levels of participation in employment, education, or family life for those with disability as measured by the question set versus those without disability to see if persons with disabilities have achieved social inclusion. In addition, the data could be used to monitor prevalence trends for persons with limitations in the particular basic activity domains selected.

[113] When domains are combined, such as asking a question about seeing or hearing, respondents frequently are confused and think they need to have difficulty in both domains in order to answer yes. In addition, having the numbers with specific limitations is useful for both internal planning and for cross-national comparisons.

4.211. Because disability is a complex concept, it is necessary to adopt an explicit definition based on the ICF domains used when developing census or survey questions that will be used to identify disability status. The recommended set of questions for censuses is based on such an explicit definition (as described above). It is essential that estimates or tabulations based on the recommended set be accompanied by information on how disability is defined and how the questions are asked. This information should be included as part of the metadata associated with the questions and data set and it should be included as a footnote to tables that include these estimates.

d. Use of census to screen for disability and follow-up with other surveys

4.212. Countries that are planning specialized surveys on disability may want to use the census to develop a sampling frame for these surveys and include a screening instrument to identify persons who will be interviewed subsequently. The main purpose of a screening instrument is to be as inclusive as possible in order to identify the largest group of people who could be further studied. The screening question should be designed so that false negatives[114] are minimized, while false positives[115] should be less of a concern.

4.213. The same recommendations highlighted in paragraphs 4.207–4.210 should also be considered when a screening module is designed.

4.214. Before embarking on using the census to develop a frame for a follow-up survey, it is important that the legal implications of using the census data for this purpose are fully considered. Respondents should be informed that the data may be used for follow-up studies and national authorities responsible for ensuring the privacy rights of the population may need to be consulted in order to obtain their approval.

E. Fertility and mortality

4.215. The investigation of fertility and mortality in population censuses is particularly important in countries lacking a timely and reliable system of vital statistics because of the opportunity the data provide for estimating vital rates that would not otherwise be available. Even in countries with complete birth and death registration, some of the topics (such as “children born alive”, “children living”, “age at marriage or union” and “age at first birth”) are equally appropriate because they provide data that are not easily available from registration data but are necessary for the computation of cohort and period fertility tables. The census provides an opportunity to collect data for estimating fertility and mortality at national and subnational levels in a cost-effective manner. The inclusion of these topics in population censuses for the purpose of estimating fertility and mortality rates and other related indicators is both prudent and cost-effective, particularly in countries where civil registration and vital statistics systems are weak, and costs of conducting large periodic demographic surveys are high. Nevertheless, it is important to note that census information is a poor substitute for complete and reliable vital registration data. If countries desire accurate and detailed estimates of fertility and mortality, they must establish, and need to maintain, civil registration systems and ensure their universal coverage.

[114] Persons who have disabilities but are not identified in the census as having disabilities.
[115] Persons who are identified with disabilities in the census but in reality do not have disabilities (as assessed in the largest instrument used in the follow-up survey).

4.216. To obtain information on fertility, information may be collected on "children ever born", "date of last child born alive" and "age of mother at birth of first child born alive". In addition, questions on age, date or duration of marriage or union may improve fertility estimates based on children ever born (see paragraph 4.247). For the collection of reliable data, some of the topics may require a series of probing questions that, because they are time consuming, are more suitable for use in sample surveys than in censuses.

4.217. The universe for which data should be collected for each of the topics included in this section consists of women 15 years of age[116] and over regardless of marital status. It is acknowledged that in some countries certain cultural sensitivities (for example, regarding the collection of information on childbearing from never-married women) exist towards collecting information from all women aged 15 years of age and over without regard to marital status. In such cases, every effort should be made to collect the information. In countries that do not use the data for women 50 years of age and over, it may be appropriate to limit data collection to women under the age of 50, allowing more concentrated effort on data collection for such women.

4.218. In addition to the topics indicated above that are used to estimate fertility, another useful topic that allows the estimation of fertility is the "own children" method[117] and birth history reconstruction.[118] The application of these methods requires the identification of the "natural mother" of each child in the household when the natural mother appears in the same questionnaire as the child. In cases where it is difficult to ascertain the identity of the natural mother, one may use as a proxy the relationship to the reference person of the household (see paragraphs 4.129–4.139) or children living (see paragraphs 4.234–4.236) to establish the identity of the natural mother. In essence, information on the child's age and the mother's age are used to estimate a series of annual fertility rates for years prior to the census. The reliability of the estimates produced depends, among other things, on the proportion of mothers enumerated in the same questionnaire as their own children, the accuracy of age reporting for both mothers and their children and the accuracy of available estimates of mortality for women and children. In the case of subnational estimates, the fertility rates may be affected by migration patterns of mothers who leave their children with other family members while they pursue work elsewhere and also good and timely birth registration of migrant mothers.

Comments by TS: Add a para on the reverse survival method in addition to the own-children method. It provides results very similar to the own-children method, but requires less detailed data. The reverse survival method does not produce age-specific fertility rates as the own-children method does, though.

[116] It may be appropriate in some countries to reduce the lower age limit by several years.

[117] For methodological details, see United Nations (2004). *Handbook on the Collection of Fertility and Mortality Data*. Studies in Methods. Series F, No.92, Chapter V, section B (http://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf); United Nations (1983). *Manual X: Indirect Techniques for Demographic Estimation*, Population Studies, No. 81 (http://www.un.org/esa/population/publications/Manual_X/Manual_X.htm), Chapter VIII, section C; and Cho, L.-J., Retherford, R. D., & Choe, M. K. (1986). *The own-children method of fertility estimation*. Honolulu, HI: Population Institute.

[118] For methodological details, see United Nations (2004). *Handbook on the Collection of Fertility and Mortality Data*. Studies in Methods. Series F, No.92, Chapter V, section C (http://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf); and Luther, N.Y., Cho, L-J. (1988). Reconstruction of birth histories from census and household survey data. *Population Studies*, 42: 451–472.

4.219. Mortality topics include infant and child mortality, obtained from data on children ever born and children living, and adult mortality, obtained from household deaths in the past 12 months and maternal or paternal orphanhood. The extent to which mortality (particularly infant and child) can be adequately measured from population census data is largely dependent on the quality of the training of field staff to minimize non-response to questions on maternal and adult mortality and the reporting of erroneous information. Enumerator manuals should include the measures that are needed to minimize such errors. Accurate responses to the questions described here are often difficult to obtain, thus resulting in faulty data. Nevertheless, it is often possible to derive usable adjusted estimates from this information.

4.220. As far as possible, efforts should be made to obtain information on fertility, child mortality (or survival) and marriage directly from the woman or mother involved, because she is more likely to recall correctly the details of her fertility, the mortality of her offspring and her marital experiences than any other member of the household. Information on household deaths, by date, sex and age, in the 12-month period prior to the census should be collected from the head of the household (or household reference person). Information on maternal orphanhood and paternal orphanhood should be collected for each person in the household regardless of age. As with fertility, mortality questions may be limited to a survey sample.

4.221. A number of countries have restricted the collection of data from fertility and mortality questions in the census to a sample of enumeration areas,[119] entailing the introduction of more vigorous training and permitting the selection of more suitable field staff. When those items are included in the census, certain precautions to ensure accuracy and completeness should be observed. As mentioned above, every effort should be made to collect all relevant information directly from the woman concerned. To reduce underreporting of events and to improve the accuracy of responses to questions on fertility and mortality, enumerators need to receive specific training on probing questions that highlight common errors and omissions.

4.222. The limitations of the data collected and of the estimates based on them should be made clear in the census reports. Furthermore, since some of the estimation procedures are only suitable for use in certain circumstances, it is important that census data producers consult specialists or carefully evaluate the methodologies for estimating the indicators for their appropriateness in a given situation. In general, the data in the basic tabulations resulting from these questions should not be used for the direct calculation of fertility and mortality rates. Reliable estimation of fertility and mortality levels using census data requires adjustment based on methods of demographic analysis.[120]

4.223. As a general guide, only one of the items related to fertility discussed below ("children ever born") is recommended for inclusion in all situations. Even in countries with reliable vital registration of births, census information on this topic can be useful for assessing the completeness of the registration system and for estimating levels of lifetime fertility for older cohorts.

---

[119] For the use of sampling in the enumeration, see Part One, Chapter VI.

[120] *Manual X: Indirect Techniques for Demographic Estimation*, Population Studies, No. 81 (United Nations publication, Sales No. E.83.XIII.2); National Academy of Sciences, Committee on Population and Demography, *Collecting Data for the Estimation of Fertility and Mortality*, Report No.6 (Washington D.C., National Academy Press, 1981), p. 220; *Handbook of Population and Housing Censuses, Part II*, Studies in Methods, No. 54 (United Nations publication, Sales No. E.91.XVII.9), Chapters III and IV; *Step-by-Step Guide to the Estimation of Child Mortality*, Population Studies, No. 107 (United Nations publication, Sales No. E.89.XIII.9); Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/.

4.224. In countries where vital registration of births and deaths is incomplete or unreliable, it is recommended that a subset of the remaining items should be included as well. Among these, one item ("date of birth of last child born alive") is useful for the indirect estimation of current fertility levels. Two additional items ("children living" and "household deaths in the past 12 months") are especially important, as they allow for the indirect estimation of mortality levels.

4.225. The three remaining items have lower priority: "age, date or duration of first marriage/union"; "age of mother at birth of first child born alive"; and "maternal or paternal orphanhood". However, in situations where a country has included one of these items in consecutive previous censuses, it may be useful to collect comparable information to measure changes over time and because cohort analysis, particularly of the prevalence of orphanhood, can be useful in assessing levels of mortality.

4.226. For countries with low fertility and mortality settings, and where the vital registration is not reliable, further consideration should be given to the item "age of mother at birth of first child born alive", since it improves the timeliness of estimates of child mortality based upon children ever born and children surviving.

4.227. In countries with high adolescent birth rates and common child marriages (with large proportion of women married before age 18), it is recommended that the questions on "age or date of first marriage/union" and "age of mother at birth of first child born alive" are included, since it improves understanding of the spatial distribution of extent of early childbearing and marriage that surveys cannot provide.

4.228. It is worth emphasizing that all estimates of fertility and mortality derived from census data are approximate and subject to various sorts of error. Therefore, in the absence of complete and reliable civil registration data, it may be desirable to have more than one type of census information on each topic (for example, both household deaths in the past 12 months and maternal or paternal orphanhood for the purpose of estimating adult mortality). Lastly, it should also be born in mind that while fertility surveys can provide data on current fertility, they cannot provide the small-area data that the census can. Therefore, a fertility question in the census can still be a priority for many countries.

1. Children ever born alive (core topic)

4.229. Information on number of *children born alive* (lifetime fertility) should include all children born alive (that is to say, excluding foetal deaths) during the lifetime of the woman concerned up to the census date. The number recorded should include all live-born children, whether born in or out of marriage, whether born in the present or a prior marriage, or in a de facto union, or whether living or dead at the time of the census. In the event of multiple births (for example twins), each child should be counted as individual birth.

Note: At 2 October 2023 meeting, there was consensus not to prescribe an age below 15 years old for data collection. Countries can choose to ask these questions to under-15-year-old girls depending on national needs and circumstances.

4.230. Data on the total number of live-born children should preferably be collected for all women 15 years of age[121] and over, regardless of marital status. If, from a cultural standpoint, it is not acceptable in some countries to attempt to obtain the information for single women, it should be collected for all women 15 years of age and over who are or have been married or in a union (in other words, all ever-married or ever cohabiting women), a group that also includes all widowed, divorced and separated women. Depending on national needs and circumstances, data on their live-born children can be collected for girls under 15 years old as well. In either case, the group of women for whom the data have been collected should be clearly described in the census report so as to avoid ambiguity in the analysis of the results. In some countries, there can be substantial misreporting of ages or dates in the census, which distorts fertility and mortality estimation based on children ever born and children living cross-tabulated by age or years since the woman's first birth.

4.231. In order to improve the completeness of coverage and to assist the respondent in recalling her children ever born alive, it is recommended that a sequence of questions be included in the following order:

(a) "total number of sons ever born alive during the lifetime of the woman";
(b) "total number of sons living (surviving) at the time of the census"; and
(c) "total number of sons born alive who have died before the census date";

and then

(d) "total number of daughters ever born alive during the lifetime of the woman";
(e) "total number of daughters living (surviving) at the time of the census"; and
(f) "total number of daughters born alive who have died before the census date".

The responses to topics (b), (c), (e) and (f) allow for a checking of the responses to (a) and (d). Inconsistencies in the figures, if any, can sometimes be resolved during the interview. In obtaining this information and ensuring its accuracy, more than one filter question can be asked to serve as a reminder for the respondent. The respondent can be confirmed by using additional tables/lists containing the names and birth years of the children she has given birth to.

Comments/suggestions by PW:
For countries that have intersex as an option for the question asking about a member's sex then the following questions are proposed to be asked.
(g) "total number of intersex ever born alive during the lifetime of the woman";
(h) "total number of intersex living (surviving) at the time of the census"; and
(i) "total number of intersex born alive who have died before the census date".
Intersex is a term used to describe a range of natural variations in sex characteristics that do not fit typical definitions of male or female. Intersex individuals may have physical characteristics that are not exclusively male or female. Collecting data on the intersex helps policymakers understand the unique healthcare needs of intersex individuals, paving the way for improved access to specialized medical and psychological services, surgeries, and hormone therapy.
If adopted this will change para 4.231., 4.235. (children living)

[121] It may be appropriate in some countries to reduce the lower age limit by several years.

4.232. The number of sons and daughters should comprise all children ever born alive whether born of the present or a prior marriage or union[122] and should exclude foetal deaths and adopted children. Also, the number of children, male and female, who are alive at the time of the census should include those living with the mother in the household and those living elsewhere, no matter where the latter may reside and regardless of their age and marital status.

4.233. The collection of data on children ever born specified by sex not only improves the accuracy of information but also provides data for the indirect estimation of sex ratio at birth as well as the calculation of infant and child mortality, in combination with data on children living (surviving) by sex (see paragraph 4.234). If the information on "children ever born alive by sex" is collected for only a sample of women, the data on "children living by sex" should also be obtained for the same sample.

4.234. Collecting data on the "total number of children ever born alive by sex" is desirable because it can enhance the quality of information. This can be achieved by ensuring that sex ratios of births follow an expected pattern and do not exhibit unusual behaviour. While this point does add value by addressing the importance of sex-specific data collection, it could benefit from providing more context on how anomalies in sex ratios might impact demographic analysis and planning.

2. Children living [123] (core topic)

4.235. Data on *children living*, in conjunction with those on children ever born, are used in indirect estimation of infant and child mortality in situations where there are no reliable data from a civil registration.

4.236. It is expected that improved coverage and quality of data on the total number of children ever born will be achieved if more detailed questions about the current residence of children ever born are asked, in terms of the following:

(a) “Total number of sons living in the household”;
(b) “Total number of sons living elsewhere”;
(c) “Total number of sons born alive who have died before the census date”;
(d) “Total number of daughters living in the household”;
(e) “Total number of daughters living elsewhere”;
(f) “Total number of daughters born alive who have died before the census date”.

These questions not only give a more complete and accurate reporting of children ever born alive specified by sex but also increase the questions’ suitability for subsequent analysis.

4.237. The identification of the natural mother of each child under 15 years of age in the same household, to be used in the “own children” method of estimating fertility (see paragraph 4.217), should be made by asking each woman who reports one or more of her children as being born alive and living in the household to identify these children in the census questionnaire. The section of the questionnaire

[122] As indicated in paragraph 2.146, couples living in consensual unions may, where appropriate, be regarded as married.
[123] For methodological details on the uses of the data, together with data on live-born children, see the publications mentioned in footnote 120.

on “relationship to the head of the household or to the reference person in the household” may be used for identifying the natural mother of each child living in the household.

3. Date of birth of last child born alive (core topic)

4.238. Information on date of birth (day, month and year) and sex of the last child born alive is used for estimating current fertility, and data on the sex of the child can also be used to evaluate the sex ratio at birth and to detect potential sex-selective birth omissions, misreporting or coding errors. This information can be useful as a means of deriving both national and subnational fertility estimates. In countries lacking adequate data from civil registration, sample surveys have become a major source of information for estimating national fertility levels, but surveys usually do not permit the derivation of reliable estimates at subnational levels.

4.239. At the data-processing stage, an estimate of the number of live births during the 12 months immediately preceding the census date can be derived from information on “date of birth of last child born alive”. For estimating current age-specific fertility rates and other fertility measures, the data provided by this approach are more accurate than information that may have been collected in earlier censuses from a question on the number of births to a woman during the 12 months immediately preceding the census.[124] Information on the date of birth of the last child born alive provides the number of women who had at least one live-born child during the 12-month period, not the number of births during the 12-month period. However, generally only a very small proportion of women will have had more than one child in a year and hence that omission will not significantly affect the fertility estimate derived from it.

At 2 October 2023 meeting this suggestion to include the birth history in the census was not accepted by the majority of the members. The questions on births history were considered to be best suited for surveys.

4.240. The information needs to be collected only for women between 15 and 50 years of age who have reported having at least one live birth during their lifetime. For countries where childbearing starts early the minimum age can be lowered according to the national circumstances and needs. Also, the information should be collected for all the marital or union status categories of women for whom data on children ever born by sex (see paragraph 4.228) are collected. If the data on children ever born are collected for a sample of women, information on date of birth for the last child born alive should be collected for the same sample.

4.241. A census question on birth history should always be paired with a simple follow-up question about whether the child is still alive, from which information on deaths of children born in the last 12 months can be rederived (see paragraph 4.244 below), and which yields data that can be used for studying child mortality (see paragraph 4.250).

Note: In the same manner as for birth history questions, adding the following questions for births occurring in the last 5 years was not accepted by the majority of the members.

[124] The approach to calculating fertility rates from these data is described in Moultrie T.A. 2013. Evaluation of data on recent fertility from censuses, in Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/content/evaluation-data-recent-fertility-censuses (accessed 06/11/2013).

4. Births in the past 12 months

4.242. Births in the past 12 months refers to the total number of children born alive to the woman concerned during the 12 months immediately preceding the census. The topic is more suitable for investigation in a sample survey than in a census. The information about births in the past 12 months is useful for calculating fertility indicators either directly or indirectly.

4.243. Because of errors and omissions commonly encountered in the reporting of live births within a 12-month retrospective period, this topic by itself cannot generally be relied on to generate accurate estimates of current fertility. It is recommended that an estimate of the number of live births during the 12 months immediately preceding the census date be derived from information on "date of birth of last child born alive" (see paragraph 4.238). For estimating current age-specific fertility rates and other fertility measures, the data provided by this approach are more accurate than information that may have been collected in earlier censuses from a question on the number of births to a woman during the 12 months immediately preceding the census.[125] Only if a country's population is characterized by low levels of date numeracy – that is, there is a strong indication of deficient accuracy of the replies – should a direct question on number of births in the last 12 months be asked. In all other cases, the core question on date of last child born alive should be sufficient.

4.244. If the topic needs to be introduced into the census, information should be collected for all the marital status categories of women for whom data on children born alive are collected. In countries in which current births to mothers at ages below 15 years are statistically significant, the age limit for the investigation of current fertility may be lowered to include mothers at appropriate younger ages.

5. Deaths among children born in the past 12 months

4.245. Deaths among children born in the past 12 months refers to the number of deaths that occurred among the live births within the 12 months preceding the census reported for the woman concerned. Hence, the topic should be investigated only if live births within this period are also investigated.

4.246. It is recommended that information on the topic be derived from a pair of census questions on "date of birth of last child born alive" and the follow-up question about "whether the child is still alive" (see paragraph 4.240). Although this pair of questions does not produce a valid estimate of the infant mortality rate (since the numerator excludes infant deaths occurring below age 1 in the past 12 months among children born 1–2 years before the census date), it can provide useful information on differences in child survival by age of mother or other socioeconomic characteristics.

4.247. The information needs to be collected only for women for whom at least one live birth during the reference period was reported.

[125] The approach to calculating fertility rates from these data is described in Moultrie T.A. 2013. Evaluation of data on recent fertility from censuses, in Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/content/evaluation-data-recent-fertility-censuses (accessed 06/11/2013).

Comments by DK: Consider extending the reference period for to 5 years instead of just the last 1 year and more observations can be obtained (like DHS).

6. Age, date or duration of first marriage

4.248. *Date of first marriage* comprises the day, month and year when the first marriage took place. In countries where date of first marriage is difficult to obtain, it is advisable to collect information on age at marriage or on how many years ago the marriage took place (duration of marriage). The information should relate to all types of marriages such as contractual first marriages and de facto unions, customary marriages and religious marriages.

4.249. For women who are widowed, separated or divorced at the time of the census, information on the "date of/age at/number of years since dissolution of first marriage" should be collected. Information on dissolution of first marriage (if pertinent) provides data necessary to calculate "duration of first marriage" as a derived topic at the data-processing stage. In countries in which duration of marriage is reported more reliably than age, tabulations of children ever born by duration of marriage yield better fertility estimates than those based on data on children born alive classified by age of the woman.[126] Data on duration of marriage can be obtained by subtracting the age at marriage from the current age, or directly from the number of years elapsed since the marriage took place.

Comments/suggestions by MN: The question on first marriage be asked from all individuals 12 years and older in the household, to accommodate countries with issues of child marriage.

7. Age of mother at birth of (date or time since) first child born alive[127]

4.250. Date of first birth comprises the day, month and year when the woman's first live birth took place. In countries where date of first birth is difficult to obtain, it is advisable to collect information on age of mother at first birth or on how many years ago the first birth took place (time since first birth). In countries in which time since first birth is reported more reliably than age, tabulations of children ever born and children surviving by time since first birth yield more timely child mortality estimates than those based on data on children born alive classified by age of the woman.[128] If the topic is included in the census, information should be obtained for each woman who has had at least one child born alive.

---

[126] See United Nations (1983). *Manual X: Indirect Techniques for Demographic Estimation*, Population Studies, No. 81 (http://www.un.org/esa/population/publications/Manual_X/Manual_X.htm), Chapter II, section D; and United Nations (2004). *Handbook on the Collection of Fertility and Mortality Data*, Studies in Methods Series F, No. 92 (http://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf), Chapter V, section D.

[127] Ibid., Chapter II, section B.3.

[128] See Hill K. 2013. Indirect estimation of child mortality, in Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population (http://demographicestimation.iussp.org/content/indirect-estimation-child-mortality); and Rajaratnam, J.K., L.N. Tran, A.D. Lopez, & C.J.L. Murray (2010). Measuring under-five mortality: validation of new low-cost methods. *PLoS Med*, 7(4), e1000253. doi: 10.1371/journal.pmed.1000253http://demographicestimation.iussp.org/content/indirect-estimation-child-mortality (accessed 13/12/2013).

8. Household deaths in the past 12 months[129] (core topic)

(suggested to extend to 24 months or five years prior to census date, or any other time period as necessary up to five years)

4.251. Information on household deaths in a specified period prior to the census date classified by sex of deceased and age at death is used to estimate the level and pattern of mortality in countries that lack satisfactory continuous death statistics from civil registration. In order for estimation derived from this item to be reliable, it is important that all deaths to household members occurring during the 5 years preceding enumeration be reported as completely and as accurately as possible. Typically, reports of deaths in censuses underestimate the overall number of deaths if only because some deaths result in the disintegration of households so that household survivors, if any, may not report their occurrence (in particular, deaths of persons living alone at the time of death are unlikely to be reported). Nevertheless, provided that there are no serious errors in the reporting of age at death, estimates of completeness of death reporting can be derived via indirect estimation and adequate mortality estimates obtained.[130,131]

Note: At 2 October 2023 meeting there was discussion whether the period can be 2 years or 5 years prior to the census.

Comments by TS: The EGM on crisis mortality recommended a period of 2 years (to be confirmed).

4.252. Ideally, information on mortality should be collected for each household in terms of the *total number of deaths in the 5-year period* prior to the census date. For each deceased person reported, name, age, sex and date (day, month and year) of death should also be collected. Care should be taken to clearly specify the reference period to the respondent so as to avoid errors due to its misinterpretation. For example, a precise reference period could be defined in terms of a festive or historic date for each country. It is also recommended for a country to use specific reference period e.g between DD/MM/20YY and DD/MM/20YY.

[129] See United Nations (2004). *Handbook on the Collection of Fertility and Mortality Data*. Studies in Methods, Series F, No.92 (http://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf); United Nations (2002). *Methods for Estimating Adult Mortality.* ESA/P/WP.175 (http://www.un.org/esa/population/techcoop/DemEst/methods_adultmort/methods_adultmort.html); and Dorrington, Rob E. (2013). The Brass Growth Balance Method and the Preston-Coale Method for one census, and the Generalized Growth Balance Method and Synthetic Extinct Generations Methods upon the availability of deaths from two censuses, in T.A. Moultrie, R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus & B. Zaba (eds.), *Tools for Demographic Estimation.* International Union for the Scientific Study of Population. http://demographicestimation.iussp.org.

[131] See Chapter 4 on methods for data evaluation and adjustment in WHO (2013). *WHO guidance for measuring maternal mortality from a census*. World Health Organization, Geneva (http://www.who.int/reproductivehealth/publications/monitoring/9789241506113/en/index.html); and Moultrie, T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus & B. Zaba (eds.) (2013). *Tools for Demographic Estimation.* International Union for the Scientific Study of Population (http://demographicestimation.iussp.org).

4.253. When information is collected on household deaths in the previous 5 years (or some other reference period), countries may wish to ask a pair of follow-up questions concerning cause of death. After ascertaining the name, age and sex of the deceased person and date of death, two additional questions could be asked:

(a) Was the death due to an accident, violence, homicide or suicide?

(b) If the deceased was a woman aged 15[132] to 49 at the time of death, did the death occur while she was:

b1. Pregnant? Yes/No

b2. During childbirth? Yes/No

b3. During the six weeks after the end of pregnancy? Yes/No

See: WHO, 2013 and US Census Bureau, 2017.

4.254. Data derived from such questions can help to assess trends in levels, and some causes, of adult mortality and maternal mortality. According to the Tenth International Classification of Diseases, maternal death is the death of a woman "while pregnant, or within 42 days of termination of pregnancy, irrespective of the duration and the site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management, but not from accidental or incidental causes" (ICD-10, Class XV, codes O00–O99). There are several concept regarding maternal death, such as:

- Maternal death: a maternal death is the death of a woman while pregnant or within 42 days of the termination of pregnancy, irrespective of the duration and the site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental causes.
- Pregnancy-related death: a pregnancy-related death is the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of cause.
- Late pregnancy-related death: a late pregnancy-related death is the death of a woman while pregnant or within 12 months of termination of pregnancy, irrespective of cause.

4.255. There is no universal agreement about the feasibility of collecting reliable cause of death information as part of a population and housing census. More research is needed on both the feasibility and methods of collecting cause of death information as part of a national census.

Proposed as new paragraphs:

To guarantee precise data collection in CAPI when the number of occurrences in a record hinges on a preceding variable—such as when a survey/census inquires about the number of children a respondent has, followed by specific questions about each child—certain strategies can be employed. Post-enumeration audits and sampling techniques should be deployed to pinpoint errors. Additionally, it's crucial to prioritize supervisor presence during pivotal phases and conduct surprise field visits for effective monitoring. Encouraging enumerators to perform peer checks can also be highly beneficial. Furthermore, fostering community engagement plays a vital role in providing support and overseeing the data collection process. By implementing these measures, CAPI data collectors can significantly enhance the accuracy and completeness of the data they gather.

[132] It may be appropriate in some countries to reduce the lower age limit by several years.

In a population census, asking about the relationship of the deceased to the head of the household holds significant importance and provides valuable information for several reasons. In a population census, inquiring about the relationship of the deceased to the head of the household serves as a fundamental source of information with wide-ranging implications such as illuminating demographic dynamics by revealing household composition and structure, from nuclear families to extended and multi-generational arrangements. It can guide policymakers in understanding caregiving responsibilities, dependency ratios, and family support systems, thus aiding in the development of social and economic policies tailored to specific household needs. It can also provide insights into migration patterns, gender dynamics, and cultural practices, offering a holistic view of the population and facilitating informed decision-making in areas spanning healthcare, housing, education, and social services.

Proposed as new paragraph:
At the data-processing stage, reported deaths can be tabulated according to broad categories of cause of death: external, pregnancy-related, other and unknown. Ignoring the "unknown" responses, both external and pregnancy-related deaths can provide valuable information in countries where no other sources of information to systematically obtain causes of death are available. Of course, such information is approximate and must be interpreted with caution after careful evaluation and often adjustment. Nevertheless, using these simple questions should make it possible to derive some useful information about major trends in mortality that are otherwise difficult to obtain.

9. Maternal or paternal orphanhood[133]

4.256. Some countries may also wish to collect information on maternal or paternal orphanhood in another attempt to ascertain the level and patterns of mortality in the population. Census data from these two topics are intended for indirect estimation of mortality by sex. Estimates are based on the proportion of persons classified by age whose natural mothers or fathers are still alive at the time of the census.

4.257. For the collection of information on orphanhood, two direct questions should be asked, namely:

(a) Whether or not the natural mother of the person enumerated in the household is still alive at the time of the census;
(b) Whether or not the natural father of the person enumerated in the household is still alive at the time of the census, regardless of whether or not the mother and father are enumerated in the same household.

The investigation should secure information on biological parents. Thus, care should be taken to exclude adopting and fostering parents. It should be kept in mind, however, that overcounting may occur in the case of parents with more than one surviving child among the respondents, particularly in high fertility societies.

[133] For methodological details on the uses of the data, see Timæus, I.M. (2013). Indirect estimation of adult mortality from orphanhood, in T.A. Moultrie, R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus & B. Zaba (eds.), *Tools for Demographic Estimation.* International Union for the Scientific Study of Population (http://demographicestimation.iussp.org/content/orphanhood); and United Nations (2002). *Methods for Estimating Adult Mortality*. ESA/P/WP.175 (http://www.un.org/esa/population/techcoop/DemEst/methods_adultmort/methods_adultmort.html).

4.258. It is preferable for these questions to be collected from every person in the household regardless of age (not just children under 18, which would otherwise make the information useless for estimating adult mortality). Not only is this important for estimating mortality at older ages, but also for estimating the extent of age exaggeration at the older ages. Whenever the context allows, the date of death should be collected to help to improve knowledge of the timing of death, and in other contexts a simple follow-up question about whether the parent was still alive five years ago can help to narrow down the timing of death and to improve adult mortality measurement for recent years by analysing these data as successive cross-sectional enquiries.[134]

F. Educational characteristics

1. Literacy (core topic)

4.259. *Literacy* has historically been defined as the ability both to read and to write, distinguishing between "literate" and "illiterate" people. A literate person is one who can both read and write, with understanding, a short, simple statement on his or her everyday life. An illiterate person is one who cannot, with understanding, both read and write such a statement. Hence, a person capable of reading and writing only figures and his or her own name should be considered illiterate, as should a person who can read but not write as well as one who can read and write only a ritual phrase that has been memorized. However, a more modern understanding referring to literacy as a continuum of skills, levels, domains of application and functionality is now widely accepted.

4.260. The notion of literacy applies to any language insofar as it exists in written form. In multilingual countries, the census questionnaire may also enquire into the languages in which a person can read and write. Such information can be essential for the determination of educational policy. This item would, therefore, be a useful additional subject of enquiry.

4.261. It is preferable that data on literacy be collected for all persons 10 years of age and over. In a number of countries, however, some children may only become literate through school between the ages of 10 and 14 years. The literacy rate for this age group may be misleading. Therefore, in an international comparison of literacy, data on literacy should be tabulated for all persons 15 years of age and over. Where countries collect the data for younger persons, the tabulations on literacy should at least distinguish between persons under 15 years of age and those 15 years of age and over.

4.262. Straightforward operational criteria and instructions for collecting literacy statistics should be clearly established on the basis of the concept given in paragraph 4.258, and applied during census taking.[135] Accordingly, although data on literacy should be collected so as to distinguish between persons who are literate and those who are illiterate, consideration should be given to distinguishing broad levels of literacy skills. Simple questions with response categories that reflect different levels of literacy skills should be used. In addition, since literacy is an applied skill, it needs to be measured in

[134] See Timæus, I.M. (2013). Indirect estimation from orphanhood in multiple inquiries, in T.A. Moultrie, R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus & B. Zaba (eds.), *Tools for Demographic Estimation.* International Union for the Scientific Study of Population (http://demographicestimation.iussp.org/content/orphanhood-multiple-inquiries).

[135] Depending on the need for small-area data and the circumstances in a country, literacy may best be measured through surveys.

relation to a particular task, such as reading, with understanding, personal letters and newspapers or magazines, or writing a personal letter or message. Respondents may be able to do so easily, with difficulty or not at all, reflecting the different levels of literacy skills. Reading and writing may be measured separately to simplify the questions.

4.263. It would be preferable to use standardized questions, harmonized across countries to ensure comparability. The United Nations Educational, Scientific and Cultural Organization (UNESCO) has developed a reference database of model questions. In addition, UNESCO recommends that literacy tests should be administered, in order to verify, as well as improve, the quality of literacy data. Nevertheless, administering a literacy test to all household members in the course of enumeration may prove impractical and affect participation, therefore limiting the utility of the results. Instead, administering such a test to a sample of respondents may be considered either in the census itself or in a targeted post-enumeration follow-up survey. Some countries have regularly used simple self-assessment questions within a census to provide an indication of literacy rates at the small-area level. An evaluation of the quality of statistics should be provided with census statistics on literacy.

4.264. The collection and tabulation of statistics on literacy during the population census should not be based on any assumed linkages between literacy, school attendance and educational attainment. In operational terms, this means systematically enquiring about the literacy status of each household member irrespective of school attendance or highest grade or level completed.

4.265. The literacy question currently varies across countries and, as a result, the data based on it are not always internationally comparable. Literacy should not be derived as an educational attainment proxy because although the two are related, there are substantial differences. For example, there are numerous cases where people leave school with only partial literacy skills, or lose them because of a lack of practice. Therefore educational attainment is not a good proxy measure of literacy skills.

2. School attendance (core topic)

4.266. *School attendance* is defined as regular attendance at any regular accredited educational institution or programme, public or private, for organized learning at any level of education at the time of the census or, if the census is taken during the vacation period, at the end of the school year or during the previous school year. According to the International Standard Classification of Education (ISCED), education is taken to comprise all institutionalized, intentional and planned activities designed to meet learning needs. Instruction in particular skills that is not part of the recognized educational structure of the country (for example, in-service training courses in factories) is not normally considered “school attendance” for census purposes.

4.267. Information on school attendance should, in principle, be collected for persons of all ages. It relates in particular to the population of official school age, which ranges in general from 5 to 29 years of age but can vary from country to country depending on the national education structure. In the case where data collection is extended to cover attendance in pre-primary education or other systematic educational and training programmes organized for adults in productive and service enterprises (such as the in-service training courses mentioned in paragraph 4.265), community-based organizations and other non-educational institutions, the age range may be adjusted as appropriate.

4.268. Data on school attendance should be cross-classified with data on educational attainment, according to the person’s current level and grade (see paragraph 4.272). This cross-classification can

provide useful information on the correspondence between age and level or grade of educational attainment for persons attending school.

4.269. The issue surrounding the number of out-of-school children has grown in importance within recent decades, particularly within the context of the UNESCO Education for All goal with regard to achieving universal primary education. The target year for Education for All was 2015 and new goals for the post-2015 period were not yet defined at the time this document was prepared. The census offers an opportunity to measure the number of "out-of-school" (reciprocal of attendance) or "ever-in-school" children.

4.270. There is a difference between "attending school" and "enrolled in school", thus results from censuses and administrative data may differ. A child can be enrolled in school but not necessarily be attending. It is recommended that these concepts be clearly defined so that countries can determine which variable they wish to collect via the census.

4.271. It is also recommended that Member States consider the need for internationally harmonized questions in order to measure school attendance and school enrolment.

4.272. For purposes of international comparison, data on school attendance should be presented by the ISCED-P (or ISCED-Programmes) levels listed below, which are used for the classification of education programmes in ISCED 2011. Correspondence between a national education system and ISCED can be established through mapping of national education programmes to the ISCED classification.[136]

ISCED level 0: Early childhood education
ISCED level 1: Primary education
ISCED level 2: Lower secondary education
ISCED level 3: Upper secondary education
ISCED level 4: Post-secondary non-tertiary education
ISCED level 5: Short-cycle tertiary education
ISCED level 6: Bachelor's or equivalent level
ISCED level 7: Master's or equivalent level
ISCED level 8: Doctoral or equivalent level

3. Educational attainment (core topic)

4.273. The recommendations on "educational attainment" (see paragraph 4.273) and "educational qualifications" (see paragraph 4.287) make use of categories of the 2011 revision of ISCED, issued by UNESCO.[137] In accordance with national conditions and requirements, many countries can continue to apply national classifications of levels and grades of education and of fields of education in collecting and tabulating statistics from population censuses. Special attention needs to be paid to establishing appropriate level or grade equivalence for persons who have received education under a different or foreign educational system. These national classifications, however, should be able to be converted or

[136] The UNESCO Institute for Statistics maintains a database with ISCED mappings at http://www.uis.unesco.org/ISCED.
[137] UNESCO Institute for Statistics (UIS). 2012. *International Standard Classification of Education: ISCED 2011*. Montreal: UIS. http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf.

mapped to the ISCED 2011 classification system, this typically being achieved during post-census processing.

4.274. *Educational attainment* is defined as the highest ISCED level successfully completed by an individual. Educational attainment is usually measured with respect to the highest education programme successfully completed, which is typically certified by a recognized qualification. Some countries may also find it useful to present data on educational attainment in terms of the highest grade completed. For international purposes a "grade" is a specific stage of instruction usually covered in the course of an academic year. Information on educational attainment should preferably be collected for all persons 5 years of age and over.

4.275. To produce statistics on educational attainment, a classification is needed that indicates the qualifications certifying the successful completion of primary, secondary and post-secondary education. Since the educational structure may have changed over time, it is necessary to make provisions for persons educated at a time when the national educational system differed from that in place at the time of the census. In addition to focusing attention on the collection of educational attainment data, enumerator instructions, coding and data processing need to be designed in a way that will take account of any changes in the educational system of a country over the years and of those educated in another country, as well as those educated in the current system.

4.276. Information collected on the highest level of education successfully completed by each individual, typically certified by a recognized qualification, facilitates flexible regrouping of the data according to various kinds of aggregation. Recognized intermediate qualifications are classified at a lower level than the programme itself. Information on intermediate qualifications or on the highest grade completed can be used to distinguish between persons who did and persons who did not complete each level of education.

4.277. For international comparison, data from the population census are needed for all levels of education defined in ISCED. To the extent possible, countries should classify statistics on educational attainment by the individual ISCED-A (or ISCED-Attainment) levels listed below, which are used for the classification of educational attainment in ISCED 2011 (or by their equivalent as set forth according to the national classification of levels of education):

ISCED level 0: Less than primary education
ISCED level 1: Primary education
ISCED level 2: Lower secondary education
ISCED level 3: Upper secondary education
ISCED level 4: Post-secondary non-tertiary education
ISCED level 5: Short-cycle tertiary education
ISCED level 6: Bachelor's or equivalent level
ISCED level 7: Master's or equivalent level
ISCED level 8: Doctoral or equivalent level

4.278. For the classification of educational attainment, ISCED level 0 has a different meaning in ISCED 2011 than for the classification of education programmes (see paragraph 4.271): it means not having successfully completed ISCED level 1. This includes individuals who have never attended an education programme, who have attended some early childhood education (defined as ISCED level 0 in the classification of education programmes), or who have attended some primary education but have not

successfully completed ISCED level 1. Any differences between national and international definitions and classifications of education should be explained in the census publications in order to facilitate comparison and analysis.

4.279. Countries could consider asking a question that captures levels of education not successfully completed, should this be of interest to policymakers or other users. This could be in the form of a direct question asking if a person has some education at the relevant level or via a question asking the last grade or year completed from any given level of education.

4.280. Data on school attendance, educational attainment and literacy status should be collected and tabulated separately and independently of each other, without (as elaborated in paragraph 4.263) any assumption of linkages between them.

4.281. In order to ensure continued and improved international comparability of census data by level of education, it is recommended that countries continue to ensure that the educational attainment variable can be mapped into the ISCED 2011 classification. This is typically achieved in post-census processing.

4. Field of education and educational qualifications

a. Field of education and training

4.282. Information on persons by level of education and field of education and training is important for examining the match between the supply and demand for qualified workers with specific specializations within the labour market. It is equally important for planning and regulating the production capacities of different levels, types and branches of educational institutions and training programmes.

4.283. A question on field of education and training needs to be addressed to persons 15 years of age and over who attended at least one grade in secondary education or who attended other organized educational and training programmes at equivalent levels.

4.284. The ISCED Fields of Education and Training 2013 (ISCED-F 2013) distinguishes between broad fields (two-digit codes), narrow fields (three-digit codes) and detailed fields (four-digit codes) of education and training.[138] The broad fields are listed here:

00 Generic programmes and qualifications
01 Education
02 Arts and humanities
03 Social sciences, journalism and information
04 Business, administration and law
05 Natural sciences, mathematics and statistics
06 Information and communication technologies (ICTs)

[138] UNESCO Institute for Statistics (UIS). 2014. *ISCED Fields of Education and Training 2013 (ISCED-F 2013).* Montreal: UIS. http://www.uis.unesco.org/Education/Documents/isced-fields-of-education-training-2013.pdf.

07 Engineering, manufacturing and construction
08 Agriculture, forestry, fisheries and veterinary
09 Health and welfare
10 Services
99 Unknown

4.285. Countries may wish to consider collecting data on narrow and detailed fields of education and training, not only on the broad fields. For this, countries should make use of the classification and coding of fields of education and training of ISCED.

4.286. Countries coding field of education and training according to a national classification should establish correspondence with ISCED, either through double coding or through conversion from the detailed national classification to ISCED. A problem may arise in identifying the exact fields of education and training of persons with interdisciplinary or multidisciplinary fields of specialization. In these cases it is recommended that countries follow the procedure of identifying the major or principal field of education and training of those with multidisciplinary specialization.

4.287. In order to ensure continued and improved international comparability of census data by field of education and training, it is recommended that the classification structure for the fields of education and training continue to be based on the most recent version of ISCED.

b. Educational qualifications

4.288. *Educational qualification* is the official confirmation, usually in the form of a document, certifying the successful completion of an education programme or a stage of a programme. Qualifications can be obtained through (a) successful completion of a full education programme; (b) successful completion of a stage of an education programme (intermediate qualifications); or (c) validation of acquired knowledge, skills and competencies, independent of participation in an education programme.

4.289. According to national needs, information on qualifications may be collected from persons who have reached a certain minimum age or level of educational attainment. Such information should refer to the title of the highest certificate, diploma or degree received.

[Suggested New Topic] Ownership of communication devices, access to Internet, and ICT skills

Topics:
1. Individuals using the internet
2. Individuals owning mobile cellular phones
3. Individuals with ICT skills

Questions on ownership of communication devices and access to internet were previously asked at household level, however, developments in information and communication technologies and proliferation of mobile devices at individual level indicates that countries might be missing on developments in the digital landscape, level of mobile phones penetration and use to inform policy.

The proliferation of mobile phones and their use beyond communication for among other things accessing internet, streaming (showing a movement away from use of such devices as radio and television in a population), etc is motivation for proposal to move these questions from household to individual level. Limiting these questions to household level only gives proxy responses, while the likelihood is that individual members of the household might access internet differently.

It is recommended that the question on ownership of communication devices and access to internet be moved from household section to demographic and social i.e., this should be asked at an individual level as follows:

Does (*name*) own any of the following devices ……in working order?
Desktop Computer
Laptop
Mobile phone
Tablet
The types of devices options included could be left to individual countries.

OR

We can opt to adopt the questions with three months reference period, but exclude option [c], because it is already included in access questions below:
Has <Name> owned a Mobile Phone in the last 3 months? Yes/ No
Has <Name> used a mobile telephone in the last three months? Yes/No
Has <Name> used the internet from any location in the last 3 months? Yes/ No
Has <Name> used a computer/ Laptop/ Tablet from any location in the last 3 months? Yes/No

How does (*Name*) MAINLY access internet?
The following options are recommended:
Use mobile phone or any other device
Access internet from work place
Access internet from school/College/University
Use internet connection from home (Wifi in dwelling)
Use community internet (public Wifi)
Use internet Café
No access to internet

This question asks about the MAIN method of internet access connection that the individual uses. If an individual has several ways of accessing internet, then the most frequently used option should be chosen.

Individuals with ICT skills (ITU recommendations)

ICT skills can be measured for individuals by examining the activities they have carried out on digital devices. This is important to link ICT usage and impact. ICT skills are categorized by area of activity (Information and data literacy; Communication and collaboration; Digital content creation; Safety; and Problem solving). Regarding the indicators of ICT skills several clarifications may be required for respondents. Reading or downloading newspapers, magazines, or books refers to those in a digital

format. Health information includes health-related information on injury, disease, nutrition, etc. Participating in social networks includes creating user profiles, posting messages or other contributions to Facebook, Twitter, Instagram, Snapchat, etc. Making calls refers to telephoning over the Internet/VoIP using Skype, WhatsApp, Viber, iTalk, etc. This also includes video calls via webcam. Sending messages refers to e-mail, messaging service, SMS among others. Attached files refers to documents, pictures, or video for example. Consultations on civic and political issues include urban planning, signing a petition and other such activities. Electronic presentations include those with text, images, sound, video or charts. Writing a computer program using a programming language refers to programming or coding in digital environments (e.g. computer software, app development). Using copy and paste tools refers to duplicating or moving data, information and content in digital environments (e.g. within a document, between devices or on the cloud). Self/user-created content includes text, images, photos, videos, music, software, etc. Editing includes text documents, spreadsheets, or presentations. Transferring files or applications includes via cloud-storage. Internet banking includes electronic transactions with a bank for payment, transfers, etc. such as M-Pesa, or for looking up account information and excludes electronic transactions via the Internet for other types of financial services such as share purchases, financial services, and insurance. Purchasing or ordering goods or services refers to purchase orders placed via the Internet whether or not payment was made online and excludes orders that were cancelled or not completed. It includes purchasing of products such as music, travel and accommodation via the Internet. Online courses can be in any subject. Connecting and installing new devices refers to connecting and installing devices such modems, cameras, printers through wired or wireless technologies. Changing privacy settings is for the purpose of limiting the sharing of personal data and information (e.g. name, contact information, photos). Security measures include strong passwords, log-in attempt notification among others.

Information and data literacy
Individuals verifying the reliability of information found online
Individuals getting information about goods or services using the Internet
Individuals reading or downloading newspapers, magazines or electronic books using the Internet
Individuals seeking health information using the Internet

Communication and collaboration
Individuals making calls using the Internet
Sending messages with attached files
Participating in social networks
Taking part in consultations or voting via the Internet to define civic or political issues

Digital content creation
Using copy and paste tools within a document
Creating electronic presentations with presentation software
Using basic arithmetic formulae in a spreadsheet
Writing a computer program using a programming language
Using software run over the Internet for editing
Uploading self/user-created content to a website to be shared

Safety
Changing privacy settings on your device, account, or app
Setting up effective security measures to protect devices and online accounts

Problem solving
Finding, downloading, installing, and configuring software and apps
Connecting and installing new devices
Transferring files or applications between devices
Internet banking
Doing an online course
Purchasing or ordering goods or services using the Internet

G. Economic characteristics

(ILO will update with its latest resolutions from the 21st ICLS held in October 2023)

1. Introduction

4.290. Statistics on the economic characteristics of persons are needed from population censuses for many reasons. Information on the productive activities of persons is vital to establish a comprehensive picture of the economic structure of a country, and the work patterns, labour market participation and extent of labour underutilization of its population. This information, when combined with other personal, household and dwelling characteristics collected in the census, enables assessments of the socioeconomic situation of persons and households, which are essential to inform the planning of a wide range of economic and social policies and programmes related to such areas as employment creation, poverty reduction, work–life balance, vocational education and training, provision of social security and other social benefits, gender justice and social inclusion.

4.291. Such statistics can be obtained from other sources, such as household-based surveys or administrative records, but these other sources have certain limitations. Household surveys, especially labour force surveys, are particularly well suited for generating a broad range of statistics on the economic characteristics of the population at aggregate levels, such as national and broad regional groupings. Data obtained from labour force surveys, however, are subject to sampling error and, therefore, rarely provide reliable estimates for small areas, small population groups, or detailed groups of industries and occupations. In contrast, population censuses can provide certain core statistics at the lowest levels of aggregation for such small population groups and for detailed occupation and industry groups. Administrative records may not have the same quality of occupational and industry coding, nor have the same comprehensiveness in population or activity coverage, generally excluding productive activities that are informal or unpaid.

4.292. The population census also provides benchmark information to which statistics from other sources can be related. Population censuses likewise provide the sample frames for most household-based surveys, including labour force surveys. In countries with a limited or infrequent household survey programme, the population census may represent the main or only source of information on the economic characteristics of the population.

4.293. In deciding which topics relating to the economic characteristics of the population to include in the population census, countries will need to assess the existence of other sources of the statistics and their complementary uses. The aim should be to cover the core topics needed as benchmark information for the preparation of sample frames, and to provide essential statistics for small areas and small population groups, and for detailed occupation and industry groups, as relevant in the national context.

4.294. International resolutions and guidelines to produce statistics relating to the economic characteristics of the population are adopted by the International Conference of Labour Statisticians and endorsed by the Governing Body of the International Labour Organization.[139] Recommendations on topics amenable for inclusion in population censuses are discussed in general in Box 1. The complete labour statistics standards are available at http://www.ilo.org/global/statistics-and-databases/standards-and-guidelines/lang--en/index.htm.

Box 1. New international recommendations concerning statistics of work, employment and labour underutilization

In October 2013, the Nineteenth International Conference of Labour Statisticians adopted the *Resolution concerning statistics of work, employment and labour underutilization*.[140] This resolution replaced the previous international recommendations relating to the measurement of the economically active population, employment, unemployment and underemployment dating from 1982 (Thirteenth International Conference of Labour Statisticians) and related guidelines.

These new standards introduced a number of important revisions, among which are a conceptual framework for work statistics consistent with the System of National Accounts; guidelines for separately measuring different forms of work, including a more targeted definition of employment as *work for pay or profit*, and for expanding the range of measures of labour underutilization beyond the traditional unemployment. New terminology was also introduced, as relevant, and terms considered to be out of date, particularly "economically active/inactive", were replaced with "labour force/outside the labour force."

Important elements from the previous standards essential to the internal consistency of the statistics remain unchanged. The refinements to the definition of employment and new measures of labour underutilization may result, however, in breaks in the historical series of statistics of the economically active population, employment, unemployment and underemployment. In particular, productive activities carried out without pay, such as those listed below, are no longer included within the scope of employment:

- Production of goods intended for own final use by the household;
- Unpaid work by apprentices, interns and trainees;
- Organization-based volunteer work;
- Direct volunteering to produce goods for other households.

Participation in these activities is now to be measured separately through the newly defined forms of work: own-use production work, unpaid trainee work and volunteer work, respectively.

Countries are encouraged to develop their statistical system so as to cover work statistics, including statistics on the labour force, based on their specific national needs and resources. In the case of the measures affected by the Nineteenth International Conference of Labour Statisticians Resolution, the

[139] See http://www.ilo.org/stat.

[140] The Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, October 2013), is available at: http://www.ilo.org/global/statistics-and-databases/meetings-and-events/international-conference-of-labour-statisticians/19/WCMS_230304/lang--en/index.htm.

updated international standards would ideally be implemented over time, in a way that is feasible for national statistical systems. During the transition period it is of utmost importance that the institutions and persons responsible for planning and managing the production of statistics on the economic characteristics of the population develop a strategic and coordinated approach that takes into account all official sources of statistics, including the population census, labour force survey and other household-based surveys and administrative records. Data users will need to be kept well informed of the process, including by widely disseminating the relevant metadata and by maintaining parallel series for a specified period following their implementation.

2. Conceptual framework for work statistics

a. Work

4.295. Measurement of the economic characteristics of the population is based on the conceptual framework for work statistics (see Box 1). In this framework, work is defined for reference purposes as "any activity performed by persons of any sex and age to produce goods or to provide services for use by others or for own use".

4.296. The concept of *work* is aligned with the general production boundary as defined in the System of National Accounts 2008 (2008 SNA), enabling full integration between work statistics and production statistics. All *work* or *productive* activities are thus included, irrespective of their formal or informal character or the legality of the activity. Excluded are activities that do not involve producing goods or services (for example begging and stealing), self-care (for example personal grooming and hygiene), and activities that cannot be performed by another person on one's own behalf (for example sleeping, learning and activities for own recreation).

4.297. *Work* can be performed in any kind of economic unit, including market units (for example corporations, quasi-corporations and household unincorporated market enterprises); non-market units (for example government and non-profit institutions serving households); and households that produce goods or services for own final use by the producers.[141]

4.298. The conceptual framework for work statistics identifies five mutually exclusive *forms of work* for separate measurement (see Figure 3):

(a) *Own-use production work*, comprising production of goods and services for own final use;

(b) *Employment work*, comprising work performed in exchange for pay or profit;

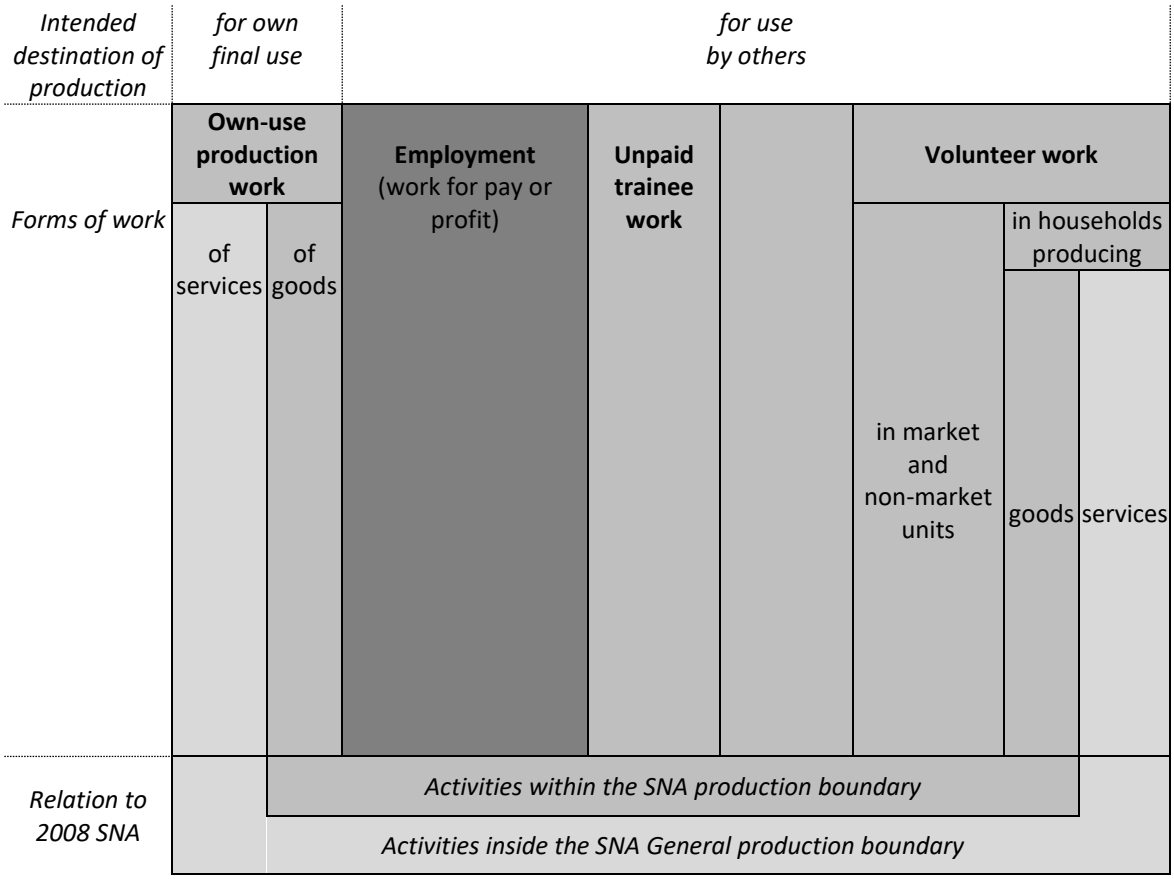
(c) *Unpaid trainee work*, comprising work performed for others without pay to acquire workplace experience or skills;

(d) *Volunteer work*, comprising non-compulsory work performed for others without pay;

[141] United Nations, *System of National Accounts 2008* (New York, 2008).

(e) *Other work activities*, including unpaid compulsory work performed for others, such as community service and work by prisoners, when ordered by a court or similar authority, and unpaid military or alternative civilian service.

Figure 3. Forms of work and the System of National Accounts 2008



Note: in figure, in the column next to the unpaid trainee work column, there will be "other work activities".

4.299. During a given reference period, persons may engage in one or more forms of work in parallel or consecutively, that is, persons may be employed, volunteering, doing unpaid trainee work or producing for own final use, in any combination.

4.300. To meet different objectives, countries may measure the economic characteristics of the population with respect to their participation in one or in several forms of work. In particular, in the population census, this may include measurement of the following:

(a) *Persons in employment* is essential as part of the preparation of labour force statistics that include unemployment and other measures of labour underutilization. It is needed to assess the labour market participation of the population and to classify the population according to their labour force status in a short reference period (see paragraphs 4.307–4.333).

(b) *Persons in own-use production of goods* is especially important in countries where particular groups of the population engage in agriculture, fishing or hunting and gathering for own final consumption, including for subsistence (see paragraphs 4.376–4.381), and to enable integration of the population census with the agricultural census (see paragraphs 1.44–1.50).

(c*) Persons in unpaid trainee work* may be advisable where unpaid apprenticeships, internships and traineeships may be a main mechanism of labour market entry for particular groups such as youths or for specific occupations such as mechanics or tailors, given their likely overall small size in the country and limited availability of alternative statistical sources.

4.301. Given the need for detailed probing, measurement of participation in *own-use provision of services, unpaid trainee work* and *volunteer work* is more appropriate through household surveys or, if desired, through the population census by means of a long form applied to a subset of the population.[142]

4.302. Additional information may also be collected in the population census in order to classify the population according to their main form of work based on self-declaration, in a short or long reference period.[143]

b. Working time

4.303. The concept of working time comprises the time associated with productive or work activities and the arrangement of this time during a specified reference period.[144] Working time relates to each form of work.

4.304. The number of persons engaged in a given form of work provides only a very rough estimate of the volume of work performed, particularly when the work is performed on a part-time, casual or occasional basis. Information on working time is necessary to prepare estimates of the volume of work or labour input for complete national production accounts. It is also essential to support the design, monitoring and evaluation of economic, social and labour market policies and programmes targeting labour market flexibility, work–life balance and conditions of work, including situations of underemployment due to insufficient working time (that is, time-related underemployment) and of excessive working time.

4.305. The population census can serve to provide information on two measures of working time in particular: *hours usually worked* and *hours actually worked*. Where the census is the only available data source it may as a minimum incorporate a single question on *hours usually worked* for persons in employment (see paragraphs 4.369–4.375) and for persons in own-use production of goods, as relevant (see paragraph 4.381).

c. Population coverage and age limits

[142] For more details see Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, 2013), paragraphs 22(c) and 37–39.
[143] Ibid, paragraph 17.
[144] Resolution concerning the measurement of working time, adopted by the Eighteenth International Conference of Labour Statisticians (Geneva, 2008).

4.306. Information on the economic characteristics of the population should in principle cover the entire population, regardless of country of origin, citizenship or geographic location of their place of work. In practice, a lower age limit is usually set in accordance with the conditions in the country. Where national programmes of statistics on the working-age population or on child labour exist, the statistics derived from the population census will serve to complement those bodies of statistics. For purposes of compiling statistics on the working-age population, international standards recommend that countries set the lower age limit taking into consideration the minimum age for employment and exceptions specified in national laws or regulations, or the age of completion of compulsory schooling.[145] For compiling child labour statistics, the relevant international standards identify the target population as all persons in the 5 to 17 years age group.[146] Countries in which many children participate in employment or in other forms of work, including in agriculture, will need to select a lower minimum age than countries where work of children is uncommon. Census tabulations of economic characteristics should at least distinguish between persons under 15 years of age and those 15 years of age and over.

4.307. In general, an upper age limit is not recommended, so as to permit comprehensive coverage of work activities of the adult population and to examine transitions between employment and retirement. Many people continue to be engaged in employment and in other forms of work beyond retirement age and the numbers involved are likely to increase as a result of factors associated with the "ageing" of the population. Countries may, however, wish to balance the cost of collecting and processing information relating to the productive activities of elderly persons (those aged 75 years or more) and the additional response burden imposed on them against the significance and reliability of the information provided.

3. Labour force status (core topic)

4.308. A classification of persons by their *labour force status* provides important information about their relation to the labour market, in particular to *work for pay or profit*, in a short reference period.

4.309. Persons may be classified in a short reference period according to their *labour force status* as being employed, unemployed, or outside the labour force as defined below in paragraphs 4.312–4.333. The three categories of *labour force status* are mutually exclusive and exhaustive. While even during a short period persons may be engaged in multiple activities, to establish their *labour force status*, priority is given to employment over other forms of work, and over unemployment; and to unemployment over outside the labour force (see Figure 4). Thus, a volunteer worker who also has a part-time employee job should be classified as *employed*, a student who is also seeking and available for employment should be classified as *unemployed*, and a person who has a part-time job working only a few hours for pay and who is also seeking another job should be classified as *employed*. The sum of persons in employment plus persons in unemployment comprises the *labour force*.

4.310. The *labour force status* of persons is established with regard to a short reference period of seven days or one week, which may be the last seven days prior to the specified census day, the last completed calendar week or a specified recent fixed week. For comparability purposes, it is particularly useful to apply the same short reference period for the census as for the national labour force survey, if any. This short reference period serves to provide a snapshot picture of labour market participation in the country

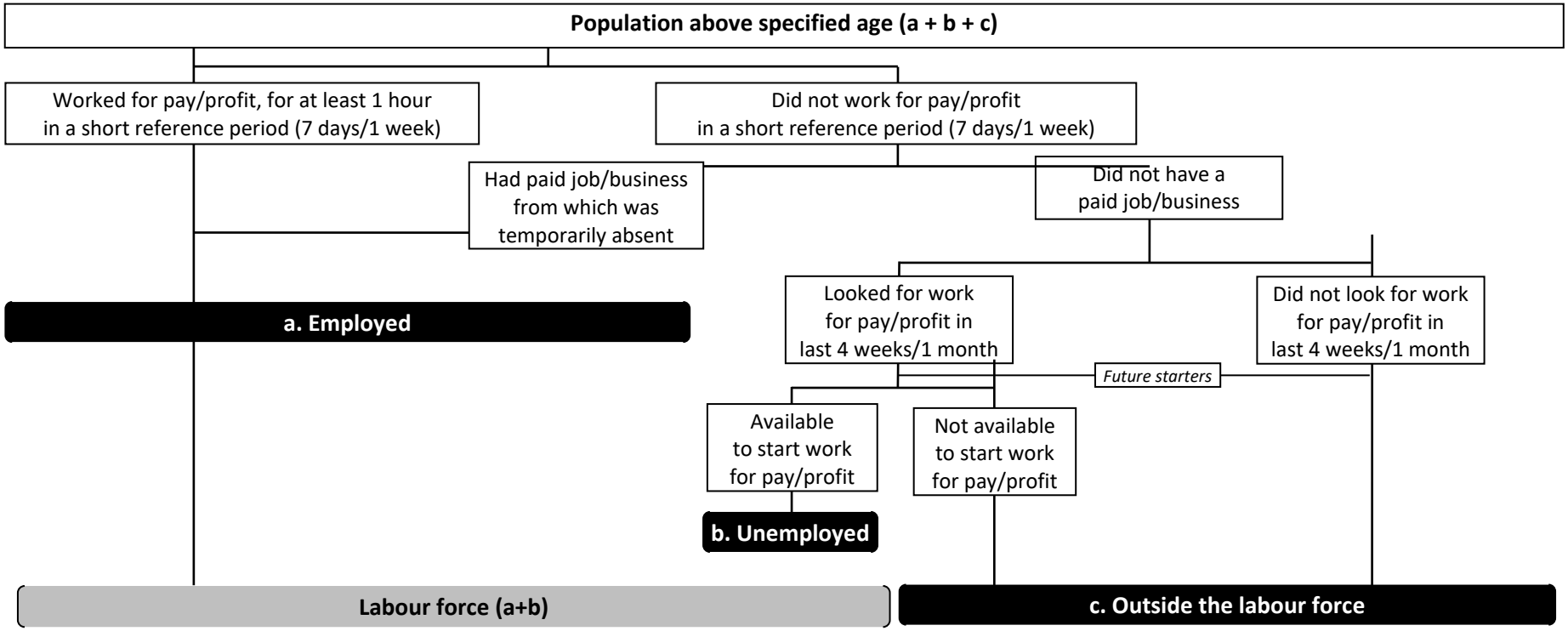
[145] Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, 2013), paragraph 65.
[146] Resolution concerning statistics of child labour, adopted by the Eighteenth International Conference of Labour Statisticians (Geneva, 2008).

around the time of the census. As such, the *labour force* (that is, persons in employment plus persons in unemployment) reflects the supply of labour for the production of goods and services in exchange for pay or profit at a specified point in time. Seasonal variations in employment and unemployment levels, which may be significant both in industrialized and in developing economies, will not be captured. Assessments of such temporal variations in work patterns are more adequately captured through sub-annual household surveys (for example monthly, quarterly).

4.311. Depending on the way the relevant parts of the census questionnaire have been constructed, the determination of the labour force status of a person may be influenced by respondents' or enumerators' subjective understanding of the concepts of employment and unemployment. In this regard, particular attention should be given to special groups for which the determination of the labour force status may be difficult. These groups include youths, women and elderly persons after the normal age of retirement, in particular those working as contributing family members. Their participation in employment and job search activities is frequently

Figure 4. Classification of working age population by labour force status

Note: Figure 4 will be reformatted.



overlooked and needs close attention. In particular, the common notion that women are generally engaged in homemaking duties, or cultural perceptions relating to gender roles, can result in serious omissions with respect to measuring women's participation in employment and job search activities. To reduce underreporting, enumerators need to be explicitly instructed, or the questionnaires specifically designed, to ask about the possible jobs, including part-time, casual, temporary and informal jobs, or job search activities of every woman and man above the specified age in the household.

4.312. The addition of probing questions in an interview, or more detailed questions in a self-administered questionnaire, may lengthen the time required to complete the questionnaire and increase the cost of the census. Accordingly, it will be necessary to balance the gains in terms of minimizing response errors when such questions are used against the added costs associated with their inclusion. Given the importance of reliable data on labour force status, however, serious consideration should be given to minimizing classification errors. To this end, the training of enumerators should highlight likely sources of omission or gender bias leading to underestimation of participation in employment.

a. Employed persons

4.313. *Employed persons* are all those above the specified age who during a short reference period of seven days or one week were engaged in any activity to produce goods or provide services for pay or profit. The notion "for pay or profit" refers to work done as part of a transaction in exchange for remuneration payable in the form of wages or salaries for time worked or work done or in the form of profits derived through market transactions from the goods and services produced. It includes remuneration in cash or in kind, whether actually received or not, payable directly to the person performing the work or indirectly to a household or family member.

4.314. Two categories of persons in employment are (a) employed persons "at work", that is, who worked for pay or profit for *at least* one hour; and (b) employed persons "not at work" due to working-time arrangements (such as shift work, flexitime and compensatory leave for overtime) or to "temporary absence" from a job for pay or profit.

4.315. Use of the one-hour criterion serves to ensure coverage of all types of jobs engaged in, including part-time, temporary or casual jobs, thereby supporting identification of all persons in employment and analysis of their working conditions. This criterion is also essential in order to ensure that unemployed persons refer to those without *any* employment, who are seeking and available for work for pay or profit. Moreover, this criterion is a prerequisite for the consistency of employment statistics with national accounts data on production. When information on working time is also collected (see paragraphs 4.369–4.375), it is recommended that employed persons be classified by specified bands of working time so as to enable identification of persons with both short and excessive working time.

4.316. Persons on "temporary absence" from a job, including as employees or self-employed, should be considered as in employment, provided that they were "not at work" for a short duration and maintained a job attachment during the absence. The existence of a job attachment should be established on the basis of the reason for the absence and, in the case of certain reasons, the continued receipt of remuneration or the total duration of the absence (in general not greater than three months).

(a) Reasons for absence where job attachment is generally maintained and thus do not require further probing include sick leave due to own illness or injury (including occupational); public holidays, vacation or annual leave; and periods of maternity or paternity leave as specified by legislation.

(b) Reasons for absence requiring further assessment of continued receipt of remuneration or total duration include parental leave, educational leave, care for others, other personal absences, strikes or lockouts, reduction in economic activity (for example temporary lay-off, slack work), disorganization or suspension of work (for example due to bad weather, mechanical, electrical or communication breakdown, problems with ICT, and shortage of raw materials or fuels).

Treatment of specific groups

4.317. According to international standards, the following groups of persons should be classified as in employment:

(a) Persons with a job for pay or profit who, during the reference period, were on training or skills enhancement activities required by their job or for another job in the same economic unit;
(b) Apprentices, interns or trainees who work for pay in cash or in kind;
(c) Persons who work for pay or profit through employment promotion programmes;
(d) Persons who work in their own economic units to produce goods intended *mainly* for sale or barter, even if part of the output is consumed by the household or family;
(e) Persons with seasonal jobs during the off season, if they continue to perform some tasks and duties of the job, excluding, however, fulfilment of legal or administrative obligations (for example pay taxes);
(f) Regular members of the armed forces and persons on military or alternative civilian service who perform this work for pay in cash or in kind.

4.318. Persons who either work in a market unit operated by a family member living in the same or in another household (that is, contributing family workers) or perform tasks or duties of an employee job held by a family member living in the same or in another household should also be classified as in employment. These groups of workers are included in employment, regardless of the number of hours actually worked, as they contribute their labour to produce goods and services for pay or profit, payable to the household or family.

4.319. In accordance with the priority rule to establish their *labour force status* (see paragraph 4.308), persons who during the reference period were primarily students, homemakers, pensioners, registered unemployed, or engaged in other forms of work, and at the same time were engaged in employment, as defined above, should be classified as in employment.

4.320. Excluded from employment are:

(a) Apprentices, interns and trainees who work *without* pay in cash or in kind (that is, unpaid trainee work);
(b) Participants in skills training or retraining schemes within employment promotion programmes, when *not* engaged in the production process of an economic unit;
(c) Persons who are required to perform work as a condition of continued receipt of a government social benefit such as unemployment insurance;
(d) Persons with seasonal jobs during the off season, if they *cease* to perform the tasks and duties of the job;
(e) Persons who retain a right to return to the same economic unit but who were absent for reasons specified in paragraph 4.315(b), when the total duration of the absence exceeds the specified threshold or if the test of receipt of remuneration is not fulfilled;

(f) Persons on indefinite lay-off who do not have an assurance of return to employment with the same economic unit;
(g) Persons who work to produce goods intended *mainly* or *exclusively* for consumption or use by the household or family, even if a surplus or part of the output is sold or bartered (that is, own-use production of goods, see paragraphs 4.376);
(h) Household members who provide *unpaid* services for consumption or use by their household (that is, own-use provision of services);
(i) Persons who work voluntarily and *without* pay to produce goods or services through or for other economic units, including market, non-market units and households (that is, volunteer work).

4.321. Information should be given in the census reports describing how the above-mentioned groups and other relevant groups were treated. Consideration should also be given to the desirability of identifying some of the groups (for example paid apprentices, interns and trainees) separately in tabulations.

b. Unemployed persons

4.322. *Unemployed persons* are all those above the specified age who (a) were not in employment, (b) carried out activities to seek employment during a specified recent period and (c) were currently available to take up employment given a job opportunity.

4.323. To be classified as unemployed, a person must satisfy all of the three criteria, where:

(a) "Not in employment" (that is, not engaged in work for pay or profit) is assessed with respect to the short reference period for the measurement of employment as defined in paragraph 4.312.

(b) To "seek employment" refers to any activity when carried out, during a specified recent period comprising the last four weeks prior to enumeration or calendar month, for the purpose of finding a job or setting up a business or agricultural undertaking. This includes also part-time, informal, temporary, seasonal or casual employment, paid apprenticeships, internships or traineeships, within the national territory or abroad. Examples of such activities are arranging for financial resources; applying for permits or licences; looking for land, premises, machinery, supplies or farming inputs; seeking the assistance of friends, relatives or other types of intermediaries; registering with or contacting public or private employment services; applying to employers directly, or checking at worksites, farms, factory gates, markets or other assembly places; placing or answering newspaper or online job advertisements; and placing or updating résumés on professional or social networking sites online.

(c) "Currently available" serves as a test of readiness to start a job in the present, assessed with respect to the same short reference period that is used to measure employment. Depending on national circumstances, the reference period may be extended to include a short subsequent period not exceeding two weeks in total, so as to ensure adequate coverage of unemployment situations among different population groups.

4.324. Unemployment has been one of the most widely used measures of labour underutilization. However, it only captures persons in situations of complete lack of work for pay or profit, and where opportunities for job search exist. In circumstances where there are few channels for seeking employment or where labour markets are limited in scope, or when labour absorption is inadequate, unemployment will not capture fully all persons with an unmet need for employment, as persons will take any available jobs, create their own jobs, often as own-account workers, become discouraged, or

engage in production of goods for own final use (for example, subsistence agriculture or fishing). Additional measures of underutilized labour include the *potential labour force*,[147] defined in paragraph 4.330, and persons in *time-related underemployment*.[148] Separate identification of these groups of persons supports better assessment of the different types of underutilization affecting labour markets across settings, and for more targeted policymaking.

4.325. It may be useful to distinguish first-time jobseekers, who have never worked before, from other jobseekers in the classification of the unemployed. Such a separation would be useful for policy purposes as well as in improving the international comparability of employment statistics. To do so, however, may require an additional question regarding previous work experience, which may impose too much of a burden for a population census.

Treatment of specific groups

4.326. Also classified as unemployed according to international standards are:

(a) *Future starters*, defined as persons "not in employment" and "currently available" who did not "seek employment" because they had already made arrangements to start a job within a short subsequent period, set according to the general length of waiting time for starting a new job in the national context but generally not greater than three months;

(b) Participants in skills training or retraining schemes within employment promotion programmes, who, on that basis, were "not in employment", not "currently available" and did not "seek employment" because they had a job offer to start within a short subsequent period, generally not greater than three months;

(c) Persons "not in employment" who carried out activities to migrate abroad in order to work for pay or profit but who were still waiting for the opportunity to leave.

4.327. In accordance with the priority rule to establish their *labour force status* (see paragraph 4.308), persons who during the reference period were mainly students, homemakers, pensioners, registered unemployed or engaged in forms of work other than employment (for example, own-use production work, volunteer work), and who at the same time were "not in employment", carried out activities to "seek employment" and were "currently available", as defined above, should be classified as in unemployment. Information should be given in the census reports on how persons in these and any other specific groups were treated.

c. Persons outside the labour force

4.328. *Persons outside the labour force* comprise all those who in the short reference period were neither employed nor unemployed as defined above, including persons below the minimum age specified for the collection of economic characteristics.

4.329. Different classifications of persons outside the labour force may be used for analytical purposes. Particularly useful for informing labour market and social policies and programmes are classifications by *degree of labour market attachment* and by *main reason for not entering the labour force*. These alternative classifications can be derived from the same questions used to identify the unemployed and may be used separately or in combination to enable further analysis.

[147] Replaces the optional relaxation of the "seeking work" criterion in the previous International Conference of Labour Statisticians standards.

[148] See Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, 2013).

4.330. Persons outside the labour force may be classified by the *degree of labour market attachment* into the following groups:

(a) *Unavailable jobseekers*, that is, those "seeking employment" but not "currently available";

(b) *Available potential jobseekers*, that is, those not "seeking employment" but "currently available";

(c) *Willing non-jobseekers*, that is, those neither "seeking employment" nor "currently available" but who want employment;

(d) *Others*, that is, persons neither "seeking employment" nor "currently available" who do not want employment.

4.331. The classification of persons outside the labour force by degree of labour market attachment allows identification of the *potential labour force*, computed as the sum of (a) *unavailable jobseekers* plus (b) *available potential jobseekers*. Together with unemployment, the *potential labour force* is a key measure of labour underutilization, relevant both in more and less developed settings, especially where the conventional means of seeking employment are of limited relevance, where the labour market is largely unorganized or of limited scope, when labour absorption is, at the time, inadequate, or where persons are largely self-employed.

4.332. Although not a part of the potential labour force, the group (c) *willing non-jobseekers* represents another group of persons outside the labour force with an expressed interest in employment and is particularly relevant for social and gender analysis in specific contexts.

4.333. Persons outside the labour force may also be classified by their *main activity or reason for not entering the labour market* into the following groups. Some persons may be classifiable in more than one category. In such situations, priority should be given to the possible categories in the following order:

(a) *Attending an educational institution* refers to persons outside the labour force, who attended any regular educational institution, public or private, for systematic instruction at any level of education, or were on temporary absence from the institution for relevant reasons corresponding to those specified for employed persons "not at work".

(b) *Performing unpaid household services* refers to persons outside the labour force engaged in the unpaid provision of services for their own household, such as spouses and other relatives responsible for the care and management of the home, children and elderly people. (Domestic and personal services provided by domestic employees working *for pay* in somebody else's home are considered as employed in line with paragraph 4.312 above).

(c) *Retiring on pension or capital income* refers to persons outside the labour force who receive income from property or investments, interests, rents, royalties or pensions from former employment.

(d) *Other reasons* refers to all persons outside the labour force who do not fall into any of the above categories (for example, children not attending school, those receiving public aid or private support and persons with disabilities).

4.334. Additional reasons for not entering the labour force that are considered particularly important at national or regional level, such as "engaged in own-use production of goods" (for example, foodstuffs from agriculture, fishing, hunting or gathering) should also be taken into account in the classification of population outside the labour force.

4. Characteristics of jobs and establishments

4.335. Once the labour force status of persons has been established, additional important topics regarding the labour market participation of the population relate to the characteristics of their jobs and of the establishments in which they work. These include in particular status in employment, occupation, place of work, industry, institutional sector, working time and income.

4.336. A *job* is defined as the set of tasks and duties performed or meant to be performed by one person for a single economic unit. Persons in employment can and do sometimes have more than one job in the reference period. In such cases the *main job* is defined in the international standards as that with the longest hours *usually* worked even if the employed person was not at work in the reference period.[149]

4.337. Job-related characteristics are generally collected in reference to the main job for persons in employment, and may also be collected in reference to the *last main job* (if any) for persons not in employment (that is, unemployed or outside the labour force). This allows for classification of the labour force (that is, employed persons and unemployed persons) and of persons outside the labour force by characteristics of their (last) main job. Once the (last) main job is identified, it is essential that all subsequent questions refer to that same job, even if the respondent was not at work in the reference period. The census questionnaire or the census information taken from registers should be designed in a way that will ensure that the variables "status in employment", "occupation", "industry", and "institutional sector" are measured for the same job. This should be a central concern also for countries that rely on the use of administrative registrations for the capturing of the correct values of these variables.

4.338. The collection of data on characteristics of the *last main job* of unemployed persons, especially occupation, industry and status in employment, may be useful in order to inform policies aimed at promoting employability and job creation. To serve this purpose, it is generally recommended to set a time limit for past employment experience (for example, during the last five or ten years) and only collect information on the characteristics of the last main job if it was held within the time limit.

4.339. When secondary jobs held in the reference period are also identified, the questionnaire should be designed so as to enable clear and separate identification of characteristics relating to main and secondary jobs. Identification of secondary jobs is particularly important in countries where multiple job holding is commonplace, particularly in agriculture, and when collecting information on income from employment and working time, in order to support analysis of the relationship between employment, income and poverty.

5. Status in employment (core topic)

(ILO will update with its latest resolutions from the 21st ICLS in October 2023)

[149] Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, 2013), paragraph 12(b).

4.340. *Status in employment* refers to the type of explicit or implicit contract of employment with other persons or organizations that the employed person has in his or her job. The basic criteria used to define the groups of the classification are the type of economic risk, an element of which is the strength of the attachment between the person and the job, and the type of authority over establishments and other workers that the person has or will have in the job. Care should be taken to ensure that an employed person is classified by status in employment on the basis of the same job used for classifying the person by "occupation", "industry" and "sector".

4.341. For purposes of international comparison, it is recommended that the main job of employed persons be classified by status in employment in accordance with the latest standards for statistics on this topic. At the time the present set of census recommendations was approved, a revision of these standards was under way and expected to be completed by 2018.[150] The latest standard was the International Classification of Status in Employment (ICSE-93) adopted by the Fifteenth International Conference of Labour Statisticians in 1993.[151] Based on ICSE-93, jobs may be classified by status in employment as follows:

(a) *Employees*
(b) *Self-employed*
    (i) *Employers*
    (ii) *Own-account workers*
    (iii) *Members of producers' cooperatives*
    (iv) *Contributing family workers*
(c) *Persons not classifiable by status*

4.342. An *employee* is a person who works in a job where the explicit or implicit contract of employment gives the incumbent a basic remuneration that is independent of the revenue of the unit for which he or she works (this unit can be a corporation, a non-profit institution, a government unit or a household). *Employees* are typically remunerated by wages and salaries, but may be paid by commission from sales, or through piece rates, bonuses or in-kind payment such as food, housing or training. Some or all of the tools, capital equipment, information systems and premises used by the incumbent may be owned by others, and the incumbent may work under the direct supervision of or according to strict guidelines set by the owner or persons in the owner's employment.

4.343. A *self-employed* person is one who works in a job where the remuneration is directly dependent upon the profits (or the potential for profits) derived through market transactions from the goods and services produced. The term "self-employed" refers to all the subcategories under (b) in paragraph 4.340: employers; own-account workers; members of producers' cooperatives; and contributing family workers.

4.344. An *employer* is a person who, working on his or her own account or with one or a few partners, holds a *self-employment job* and, in this capacity, has engaged on a continuous basis (including the reference period) one or more persons to work for him or her as employees. The incumbent makes the operational decisions affecting the enterprise, or delegates such decisions while retaining responsibility for the welfare of the enterprise. Some countries may wish to distinguish among employers according to the number of persons they employ.

[150] Information about the issues that may be addressed in the revision of these standards may be found in *Revision of the International Classification of Status in Employment (ICSE-93)*, Room Document 8, Nineteenth International Conference of Labour Statisticians, 2–11 October 2013 (Geneva, 2013).
[151] For more details see Resolution concerning the International Classification of Status in Employment (ICSE), adopted by the Fifteenth International Conference of Labour Statisticians (Geneva, 1993).

4.345. An *own-account worker* is a person who, working on his or her own account or with one or a few partners, holds a *self-employment job*, and has not engaged any employees on a continuous basis. (Note, however, that during the reference period an own-account worker may have engaged one or more employees on a short-term and non-continuous basis without being thereby classifiable as an employer.) Persons engaged in agriculture (including livestock care), fishing, hunting and gathering, intended mainly for own consumption by their households, are no longer included in employment on the sole basis of that activity and should therefore not be considered as "own-account workers". Instead, participation in these productive activities is to be measured through the separate concept of "own-use production of goods" (see paragraphs 4.376–4.381).

4.346. A *member of a producers' cooperative* is a person who holds a self-employment job in an establishment organized as a cooperative, in which each member takes part on an equal footing with other members in determining the organization of production, sales or other work, investments and the distribution of proceeds among the members. Note that employees of producers' cooperatives are not to be classified as in this group but should be classified as "employees".

4.347. A *contributing family worker* is a person who holds a self-employment job in a market-oriented establishment operated by a related person living in the same or in another household, and who cannot be regarded as a partner (that is to say, an employer or own-account worker) because the degree of his or her commitment to the operation of the establishment, in terms of working time or other factors to be determined by national circumstances, is not at a level comparable with that of the head of the establishment. Where it is customary for young persons, in particular, to work without pay in a market-oriented enterprise operated by a related person who does not live in the same household, the requirement that the person lives in the same household may be relaxed.

4.348. *Persons not classifiable by status* include those persons with jobs about which insufficient information is available, or who cannot be included in any of the preceding categories (for example, persons assisting with the tasks or duties of an employee job held by a family member living in the same or in another household).

4.349. When members of the armed forces paid in cash or in kind are counted among the employed, they should be included in the category of employees. However, because of the wide range of national practices in the treatment of the armed forces, it is recommended that census tabulations and related notes provide an explicit indication of the status in employment category in which they are included.

4.350. There are several groups of workers that are on the margin between employee and self-employed, such as owner-managers of incorporated enterprises (see following paragraph), outworkers, contract workers and commission workers. [152] Consultations between national accountants and labour market analysts will be necessary to make decisions about the treatment of these groups in a consistent manner, and depending on the descriptive and analytical purposes of the statistics.

4.351. *Owner-managers of incorporated enterprises* are workers who hold a job in an incorporated enterprise in which they (a) alone, or together with other members of their families or one or a few partners, hold controlling ownership of the enterprise; and (b) have the authority to act on its behalf

[152] For a discussion of the treatment of these groups, see Resolution concerning the International Classification of Status in Employment (ICSE), adopted by the Fifteenth International Conference of Labour Statisticians (Geneva, 1993); and *Revision of the International Classification of Status in Employment (ICSE-93),* Room Document 8, Nineteenth International Conference of Labour Statisticians, 2–11 October 2013 (Geneva, 2013).

as regards contracts with other organizations and the hiring and dismissal of employees, subject only to national legislation regulating such matters and the rules established by the board of the enterprise.

4.352. In most census questionnaires, the information concerning status in employment will be captured through precoded alternatives where only a few words can be used to convey the intended meaning of each category. This may mean that classification of some of the situations on the borderline between two or more categories will be carried out according to the subjective understanding of the respondent rather than according to the intended distinctions. This should be kept in mind in designing the questionnaire and also when presenting the resulting statistics. Countries that rely on the direct use of administrative records for the classification of persons according to status in employment may find that the group "contributing family workers" cannot be separately identified. Those who would have been classified as being in this group when using a questionnaire may either be classified as part of one of the other groups or excluded from persons in employment.

6. Occupation (core topic)

4.353. *Occupation* refers to the type of work done in a job by the person employed (or the type of work done in the last job held, if the person is unemployed), irrespective of the industry or the status in employment in which the person's job should be classified. Type of work is considered in terms of the main tasks and duties performed in the job.

4.354. For purposes of international comparison, it is recommended that countries make it possible to prepare tabulations involving occupations in accordance with the latest revision available of the International Standard Classification of Occupations (ISCO). At the time the present set of census recommendations was approved, the latest revision was the one adopted by a Tripartite Meeting of Experts in Labour Statistics in 2007 and endorsed by the Governing Body of the International Labour Organization in 2008[153] and generally known as ISCO-08. Countries coding occupation according to a national standard classification should establish a correspondence with ISCO either through double coding or through mapping from the detailed groups of the national classification to ISCO.

4.355. Countries should code the collected occupational responses at the lowest possible level of ISCO or a related national classification supported by the information given in each response. In order to facilitate detailed and accurate coding, it would be useful for the census questionnaire to ask each employed person for both the occupational title and a brief description of the main tasks and duties performed on the job. Information provided in response to the industry questions (see following section) may also be used to assist in the coding of occupation data, where the occupation response on its own is insufficient to assign a detailed occupation classification code.

4.356. In preparation for the coding of the occupation responses, the organization responsible for the census should prepare a coding index reflecting the type of responses that will be given by the respondents. The coding index should be constructed by occupational classification experts on the basis of responses to similar questions in other data collections, such as previous censuses, census tests and labour force surveys, as well as input from job placement officers of the employment service and the content of newspaper advertisements of vacant jobs. The coding index should clearly distinguish between responses belonging to "not elsewhere classified" categories and responses that do not provide enough information to determine an occupational group.

[153] International Standard Classification of Occupations (ISCO-08), Volume 1, Structure, Group Definitions and Correspondence Tables (Geneva, International Labour Office, 2012).

7. Industry (core topic)

4.357. *Industry* (branch of economic activity) refers to the kind of production or activity of the establishment or similar unit in which the job(s) of the employed or unemployed person was located during the time reference period established for data collection on economic characteristics.[154]

4.358. For purposes of international comparison, it is recommended that countries compile information on industry according to the most recent revision of the International Standard Industrial Classification of All Economic Activities (ISIC) available at the time of the census. At the time this present set of census recommendations was approved, the fourth edition of ISIC, adopted by the United Nations Statistical Commission at its thirty-seventh session in 2006, was the latest revision available. Countries coding industry according to a national standard classification should establish correspondence with ISIC either through double coding or through mapping from the detailed groups of the national classification to ISIC.

4.359. Countries should code the collected industry responses at the lowest possible level of ISIC or a related national classification supported by the information given in each response. In order to facilitate detailed and accurate coding, for each job to be coded the census questionnaire should ask for the main products and services produced or the main functions carried out at the establishment or enterprise in which the person was employed. It is recommended that the name and address of the establishment should also be collected (see also paragraph 4.363). Countries with business registers that are complete and up to date can then use this response as a link to the register in order to obtain the industry code given there to the establishment.

4.360. In preparation for the coding of the industry responses that cannot be matched to a precoded register, the organization responsible for the census should create a coding index that reflects the type of responses that will be given on the census questionnaire. This coding index should be constructed by industry classification experts on the basis of available lists of enterprises, establishments, businesses and so forth, as well as from responses to similar questions in other data collections, including previous censuses, census tests and labour force surveys. The coding index should clearly distinguish between responses belonging to "not elsewhere classified" categories and responses that do not provide enough information to allow for the coding of a detailed industry group.

8. Place of work

4.361. Two main topics related to the place of work of persons in employment are the *type of workplace* and its *geographic location*. The type of workplace refers to the nature of the place where the person performed his or her *main job* and distinguishes between the home and other workplaces, whether fixed or otherwise.

4.362. Three main categories, or a variation thereof necessitated by national circumstances, are recommended for classifying the *type of workplace*:

[154] For those persons who are recruited and employed by one enterprise but who actually work at the place of another enterprise (called "agency workers" or "seconded workers" in some countries), there would be user interest in gathering information about the industry of the employer as well as the industry of the place of work. However, the collection of both would be more appropriate in a labour force survey rather than in a population census. The industry of the actual place of work may provide more reliable reporting of the "industry" variable in a population census. Any such choice should, however, be consistent with the treatment of this group in the System of National Accounts.

(a) *Work at home*. This category includes those who perform the tasks and duties of their main job from within the home, such as farmers who work and live on their farms, homeworkers, self-employed persons operating (work)shops or offices inside their own homes, and persons working and living at work camps.

(b) *No fixed place of work*. This category should be restricted to persons who, in performing the tasks and duties of their main job, travel in different areas and who do not report daily in person to a fixed address as a work base, for example, travelling salespersons, long-distance commercial vehicle drivers, seafarers, fishers and own-account taxi drivers. It also includes ambulant vendors, operators of street or market stalls that are removed at the end of the workday, construction workers working at different sites during the reference period and push-cart operators.

(c) *With a fixed place of work outside the home*. All other persons in employment should be included in this category, including persons who move around in their job but have a fixed-base location to which they report daily, such as bus and taxi drivers (with a base), train and airline staff, and operators of street and market stalls that are not removed at the end of each workday. This group may also include individuals who travel to work, on a regular basis, across the national border to a neighbouring country. Fixed place of work outside the home includes: client's or employer's home; employer's workplace or site; own business premises; client's workplace or site; no single type of location.

4.363. It is likely that for some jobs, performance is at more than one location (for example, at home some of the time or season and in a fixed location outside the home at other times) or the category cannot be clearly distinguished. One approach, in the case of the former, would be to select the place where the individual spends or spent a major part of his or her working time. Where the distinction between categories is blurred, as is the case for work done, for example, on a rented plot of land adjacent to one's home, it would be useful to identify borderline cases, according to national circumstances. Specific instructions should be given to the enumerators on how to select between two or three possible responses to classify borderline cases.

4.364. The *geographic location* of the place of work can provide useful information for planning when used together with information on place of residence. To this end, countries may collect, for employed persons with a fixed place of work outside the home, information on the location of the place of work (or the reporting place) during the reference period. The information collected should relate to the smallest civil division in which the job is performed, for example, in order to establish commuter flows from the place of residence to the place of work. Some countries investigating this topic in the population census have recorded the actual address of the place of work, allowing detailed tabulations and mapping of place of residence by geographic location of place of work. Information on actual address of the place of work can also be useful for industry coding (see paragraph 4.358) in countries where a business register has been developed that shows the industry code of each recorded establishment.

4.365. In some countries there may be concerns about the sensitivity of questions on the address of place of work owing to fears that there may be follow-up to a respondent's employer. In many developing countries, it may not be possible to gather information on actual address of place of work because street addresses do not exist, and for proxy responses, the address may not be known. In those situations, it would be useful to consider collecting information on the village, suburb, or similar low level of geography.

4.366. Additional questions may also be asked on the method of travel to work in order to produce statistics on travel-to-work patterns, valuable as basis for transportation planning. Persons not travelling to work should be classified as "no travel to work."

9. Institutional sector of employment

4.367. The *institutional sector of employment* relates to the legal organization and principal functions, behaviour and objectives of the enterprise with which a job is associated.

4.368. Following the definitions provided in the System of National Accounts, distinction should be made between the following institutional sectors:

(a) *Corporation*, comprising non-financial and financial corporations (in other words incorporated enterprises, private and public companies, joint stock companies, limited liability companies, registered cooperatives, limited liability partnerships, and so forth) and quasi-corporations (that is to say, an unincorporated enterprise that is managed as if it were a corporation, in that a complete set of accounts is kept), as well as non-profit institutions, such as hospitals, schools and colleges that charge fees to cover their current production costs;

(b) *General government*, comprising central, state and local government units together with social security funds imposed or controlled by those units, and non-profit institutions engaged in non-market production controlled and financed by government, or by social security funds;

(c) *Non-profit institutions serving households* (for example, churches, professional societies, sports and cultural clubs, charitable institutions and aid agencies) that provide non-market goods and services for households (that is to say, free or at prices that are not economically significant) and whose main resources are from voluntary contributions;

(d) *Households* (including unincorporated enterprises owned by households) comprising unincorporated enterprises directly owned and controlled by members of private and institutional households (made up of persons staying in hospitals, retirement homes, convents, prisons and so forth, for long periods of time), either individually or in partnership with others. Partners may be members of the same household or from different households.

4.369. In most census questionnaires, the information concerning institutional sector of employment will be captured through precoded alternatives where only a few words can be used to convey the intended meaning of each category. This may mean that classification of some units on the borderline between two or more categories will be carried out according to the subjective understanding of the respondent rather than according to the intended distinctions. This should be kept in mind when presenting the resulting statistics.

10. Working time

4.370. The number of employed persons provides only a very rough estimate of the volume of work performed, especially when such persons have non-standard working hours. Inclusion in the census of an item on time worked helps to ensure a more accurate measurement of the concept by capturing the full contribution of persons who were in and out of the workforce or who worked only for a brief time during the year (for example, women).

4.371. To provide a comprehensive measure of working time in employment that will best inform policy and analytical needs, it may be preferable to collect information about the total hours worked in all jobs rather than to limit the information to hours worked in the main job.

4.372. Information on two distinct concepts of working time can be collected in a population census: *hours actually worked* and *hours usually worked*.

4.373. *Hours actually worked* is defined as the time spent in a job for the performance of activities that contribute to the production of goods and/or services during a specified reference period. It covers the time spent in "direct hours", in "related hours", "down time" and short "resting time". "Direct hours" is the time spent carrying out the tasks and duties of the job – and may be performed in any location. "Related hours", while not leading directly to goods produced or services provided, is the time spent maintaining, facilitating or enhancing productive activities, including upkeep of the workplace, changing time or decontamination of work clothes, purchasing or transporting materials, waiting for business, customers or patients, on-call duties, travelling between work locations, and work training or skills enhancement required by the economic unit. In practice, "down time" includes unavoidable, temporary interruptions to work (for example machinery or Internet breakdown, lack of supplies). "Resting time" is inactive time for short rest or refreshment in the course of performing job-related activities, (for example coffee breaks). Longer breaks for meals, time spent not working because of vacation, holidays, sickness, industrial disputes, etc., commuting to work (if not also performing job tasks or duties) and educational leave even if paid, are excluded from hours actually worked.

4.374. Measurement of hours actually worked in employment, in the context of the population census, is usually collected using one direct question; it is optimally measured using a set of questions, requesting hours separately for each day of the week. For employed persons not at work in the short reference period, it is possible to have a value for hours actually worked of zero (for persons away on leave) or reduced (if a part of the reference period was taken off for sickness, holiday, or other purpose).

4.375. *Hours usually worked* is defined as the typical value of the hours actually worked in a job per short reference period (for example one week) over a long observation period (month, quarter, season, year) that comprises the short reference period itself. This "typical value" of time worked during a normal or typical week may be the modal number of the hours actually worked in the short period as distributed over the long period. This would include overtime hours regularly worked whether paid or unpaid. Days and hours not usually worked and unusual periods of overtime are not included.

4.376. Measurement of hours usually worked in employment relating to the short reference period of one week can be done with one direct question: how many hours do you usually work per week (in your main job or in all jobs)? For persons with more than one job during the reference week, to record both working time in the main job (for which the other descriptive variables are collected) and total working time (sum of working time in all jobs) the questionnaire would require at minimum two questions.

11. Participation in own use production of goods (core topic)

4.377. Countries where production of goods for own final use (such as foodstuffs from agriculture, fishing, hunting and gathering, water, firewood and other household goods) represents an important component of the livelihood of a part of the population, whether as a main or secondary activity, will need to consider collecting information in the population census on the number of persons engaged in this form of work (previously included within the concept of employment). Such information is essential for benchmarking purposes, especially where household surveys are not frequent, for comprehensive sectoral analysis, particularly of work in agriculture, forestry and fishing,

and to enable integration of the population census with the agricultural census (see also paragraphs 4.387–4.396).

4.378. *Persons in own-use production of goods* are all those above the specified age who, during a specified reference period, performed "any activity" to produce goods for own final use. The notion "for own final use" is interpreted as production where the intended destination of the output is *mainly* for final use by the producer in the form of capital formation, or final consumption by household members, or by family members living in other households.

4.379. According to international standards, "any activity" to produce goods (within the 2008 System of National Accounts production boundary) covers work performed for at least one hour in the following activities, when the intended destination of the output is *mainly* for own final use, as specified above:

(i) Producing and/or processing for storage agricultural, fishing, hunting and gathering products;
(ii) Collecting and/or processing for storage mining and forestry products, including firewood and other fuels;
(iii) Fetching water from natural and other sources;
(iv) Manufacturing household goods (such as furniture, textiles, clothing, footwear, pottery or other durables, including boats and canoes);
(v) Building, or effecting major repairs to, one's own dwelling, farm buildings, etc.

4.380. For measurement purposes, the intended destination of the output is established in reference to the specific goods produced, based on self-declaration (that is, mainly for own final use). In the case of goods from agriculture, fishing, hunting or gathering intended mainly for own consumption, a part or surplus may nevertheless be sold or bartered.

4.381. Persons may engage in own-use production of goods as a main or secondary activity, throughout the year or on a seasonal basis. To ensure complete coverage, the census questions on participation in own-use production of goods should be applied to all persons above the specified age for collecting information on the economic characteristics of the population, *irrespective* of their labour force status. The reference period may refer to the last 12 months, calendar year, agricultural year or season, as relevant to national circumstances. Where pertinent, the choice of reference period should promote coherence with the agricultural census (see also paragraphs 1.44–1.50).

4.382. For assessments of the volume of work performed by persons in own-use production of goods, particularly when using a long reference period, it may be useful to include a question on working time, in particular hours usually worked (see paragraph 4.374), or based on broad categories such as part time or full time, part year or full year, number of months, as feasible and relevant to the main uses of the statistics.

12. Income

4.383. Countries may wish to collect information on the amounts of income received by individual persons or households during a specified reference period, from any source. If this topic is included in the census it is recommended that data be obtained for all persons above a specified age, whether they are employed or not. Income may be measured at the household level, or for each individual in the household.

4.384. *Income* may be defined as all receipts whether monetary or in kind (goods and services) that are received by the household or by individual members of the household at annual or more frequent intervals, but excluding windfall gains and other such irregular and typically one-time receipts. Household income covers (a) income from employment (both paid and self-employment); (b) income from the production of goods for own final use; (c) income from the provision of household services for own final use; (d) property income; and (e) current transfers received.[155]

4.385. The collection of reliable data on income, especially income from self-employment and property income, is extremely difficult in general field enquiries, particularly population censuses. The inclusion of non-cash income further compounds the difficulties. Collection of household income data in a census, even when confined to cash income, presents special problems in terms of burden of work, response errors, and so forth. Therefore, this topic is generally considered more suitable in a sample survey of households or from administrative data sources such as tax or social security records. Depending on the national requirements, countries may nonetheless wish to obtain limited information on personal or household income, by covering only some of the income components (such as income from employment), for shorter reference period (such as one month), and cover only cash income. As thus defined, the information collected can provide some input into statistics that have many important uses.

4.386. According to international standards on the subject, the income from employment of employed persons should include wages and salaries of employees, income of members from producers' cooperatives and the mixed income of employers and own-account workers operating business and unincorporated enterprises. In addition to the income from employment of employed household members, the total income of the household should include, for example, the interest, dividends, rent, social security benefits, pensions and life insurance annuity benefits of all its members. The *Handbook on household income statistics*[156] provides further guidance on concepts and methods related to this topic.

4.387. The concepts involved in determining income are not simple to grasp and respondents may be unable or unwilling to provide exact information. For example, income should include social security benefits, pension fund contributions and direct taxes withheld from employees' salaries, but some persons will undoubtedly not include these amounts in reporting their salaries. Significant items of total household income may also be excluded or misstated. Despite instructions given to enumerators, the data collected can therefore only be expected to be approximate. Accordingly, in the presentation of results it is usually appropriate to use broad income or earnings size classes. As an aid to the interpretation of the results, tabulations of the data should be accompanied by a description of the items of income assumed to be included and, if possible, an estimate of the accuracy of the figures.

H. Agriculture

1. Introduction

[155] See Resolution concerning household income and expenditure statistics, adopted by the Seventeenth International Conference of Labour Statisticians (Geneva, 2003), paragraphs 4–5. Available at http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087503.pdf.

[156] *Canberra Group Handbook on Household Income Statistics*, second edition, 2011, available at http://www.unece.org/fileadmin/DAM/stats/groups/cgh/Canbera_Handbook_2011_WEB.pdf.

4.388. In this section two non-core topics on agriculture are presented. These two alternative topics could be considered by countries that would like to collect information in their population census that would facilitate the preparation of the frame of agricultural holdings in the household sector, for a subsequent agricultural census (see also paragraph 1.44–1.50).

4.389. With the first topic, at the household level, information is collected on whether any member of the household is engaged in own-account agricultural production activities at their place of usual residence or elsewhere. This information can be restricted to limited key items or may include a more comprehensive agricultural module. With the second topic, at the individual person level, information is collected to identify persons involved in agricultural activities during a long reference period such as a year before the census, in contrast to most of the topics included in a population census for which the information generally relates to the shorter reference period of a week before the census.

2. Own-account agriculture production

4.390. Some countries may want to use the population census to identify households engaged in own-account agricultural production. This information is useful for agriculture-related analysis of the population census and for use as a frame for a subsequent agricultural census or other surveys. In this case, information should be collected for all households on whether any member of the household is engaged in any form of own-account agricultural production activities.

4.391. Where possible, information should be collected to identify whether the household is engaged in any form of own-account agricultural production, covering the main agricultural activities important to the country (which can include crops, livestock and related activities). Information may also be collected on forestry, fishery and aquaculture activities in case they are important for a country. Additional information should also be collected giving a measure of farm size – such as the area or number of plots used for agricultural purposes. Countries wishing to collect more comprehensive agricultural data may wish to include an agriculture module with the core data items recommended by the FAO World Programme for the Census of Agriculture 2020 (WCA 2020) and the FAO/UNFPA Guidelines for Linking Population and Housing Censuses with Agricultural Censuses (FAO and UNFPA, 2012).

4.392. Where aquacultural production is important at the household level, information can also be collected on whether or not any member of the household is engaged in any form of own-account aquacultural production activities.

4.393. Agricultural production activities refer to groups 011, 012, 013, 014 and 015 of ISIC (Rev. 4.0), namely:
Group 011: Growing of non-perennial crops
Group 012: Growing of perennial crops
Group 013: Plant propagation
Group 014: Animal production
Group 015: Mixed farming.

Aquacultural production activities refer to group 032 of ISIC (Rev. 4.0), namely:
Group 032: Aquaculture

3. Characteristics of all agricultural activities during the last year

4.394. The population census normally collects information about a person's main job or work activity during a short reference period, which may not cover all persons working in agriculture

because of the seasonality of many agricultural activities and because agriculture may not be the main activity of the person. To overcome this problem, information should be collected on all persons that carried out agricultural activities during the year preceding the population census day. The information to be collected should include the occupation and status of employment of all agricultural jobs, and could be expanded to cover working time and whether the job was performed as a main or secondary activity. Given the newly adopted conceptual framework for work statistics, information should also be collected on participation in own-use production of agricultural goods, particularly in countries where subsistence agriculture is practised by part of the population (see paragraphs 4.376–4.381).

4.395. Information on occupation and status in employment of all agricultural jobs (main and secondary), and on participation in own-use production of agricultural goods, can be used as an alternative way to facilitate identification of households engaged in own-account agricultural production activities (see paragraphs 4.389–4.392). Status in employment and participation in own-use production of agricultural goods could be used to distinguish between households that are engaged in own-account agricultural production activities and households with members engaged in agricultural activities only as paid employees, which would not qualify them as households with own-account agricultural production.

4.396. Where aquacultural production is important in a country, similar information on occupation and status in employment of all aquacultural jobs, and on participation in own-use production of aquacultural goods, during the year preceding the population census day can also be included. The information to be collected could be expanded, as required, to include working time and whether the job was performed as a main or secondary activity.

4.397. An agricultural job or work activity is defined as a job or work activity in the agricultural industry as defined by groups 011, 012, 013, 014 and 015 of ISIC (Rev. 4.0); namely:
Group 011: Growing of non-perennial crops
Group 012: Growing of perennial crops
Group 013: Plant propagation
Group 014: Animal production
Group 015: Mixed farming.

An aquacultural job or work activity is defined as a job or work activity in the aquacultural industry as defined by group 032: Aquaculture of ISIC (Rev. 4.0).

Chapter II. Housing census topics

I. Factors determining the selection of topics

4.398. In line with the overall approach to revision 3 of *Principles and Recommendations for Population and Housing Censuses*, the selection of housing census topics, as with the population topics described in Chapter I, is based on the outputs expected to be produced. Therefore, the first step involves the clear identification of expected outputs; the core and additional topics are then decided on that basis. For each of the core topics there is a recommendation. It is recommended that countries collect data on the core topics and also produce the recommended tabulations, to improve the international harmonization and comparability of statistics through the use of common concepts, definitions and classifications. Use of an agreed international approach would also enhance the capacity of countries to generate statistics for monitoring the socioeconomic situation of their populations, including for the provision of data for the internationally agreed development goals.

4.399. In reference to the selection of topics to be included in a housing census, limiting statistical enquiries to the collection of data that can be processed and published within a reasonable period of time is very important. Such cautions are especially applicable to a housing census, since it is customary to conduct housing and population censuses simultaneously or as consecutive operations. There is a high probability that the amount of data required from a census may be beyond the capacity of enumerators to collect or the census agencies to process. It may be sufficient in some developing countries, for example, to ascertain only the number of housing units and other sets of living quarters of various types, the number and characteristics of the occupants thereof and the availability of a water supply system.

4.400. In this context, countries should not attempt to collect housing data that are so incomplete that they fail to serve the principal purposes for which they are required. It is important, therefore, for census takers to consult closely with the principal users at an early planning stage in order to identify the data that are of highest priority and the means of supplying them in the most useful formats.

4.401. The topics, therefore, to be covered in a housing census (that is to say, the subjects regarding which information is to be collected for living quarters, households and buildings) should be based on a balanced consideration of:
(a) The needs of the broad range of data users in the country at both the national and local area level (national priority);
(b) The achievement of the maximum degree of international comparability, both within regions and on a worldwide basis (international comparability);
(c) The sensitivity of the topics and respondent burden, that is, the willingness and ability of the public to give accurate information on the topics (suitability);
(d) The technical competence of the enumerators in regard to obtaining information on the topics (suitability);
(e) The total national resources available for conducting the census (resources);
(f) The availability of relevant information held in alternative data sources like sample surveys and administrative registers.

4.402. Such a balanced consideration will need to take into account the advantages and limitations of alternative methods of obtaining data on a given topic within the context of an integrated national programme for gathering housing statistics.

4.403. In selecting the housing topics, regard should also be given to the usefulness of historical continuity, which provides the opportunity for comparison of changes over a period of time. Census takers should avoid, however, collecting information that is no longer required by users. Information should not be collected simply because it was traditionally collected in the past, bearing in mind changes in the socioeconomic and housing circumstances of the country. It becomes necessary, therefore, in consultation with a broad range of users of census data, to review periodically the value of even long-standing topics and to re-evaluate the need for their continued collection, particularly in the light of new data needs and alternative data sources that may have become available for investigating topics hitherto covered in the population and housing census. Each of five key factors that need to be taken into account in reaching a final decision on census content are briefly reviewed in the following paragraphs.

A. National priority

The priority of designing a housing census should be to meet national needs. The prime consideration is that the census should provide information on those topics that are of greatest value to the country, with questions framed so as to elicit data of maximum utility. Each country's decision with regard to the topics to be covered should depend upon a balanced appraisal of how urgently the data are needed and whether the information could be equally well or better obtained from other sources.

4.404. Some countries may omit from the census certain recommended topics because there is not a need to collect the data. For example, a particular amenity, such as electricity or toilet facilities, might be available virtually everywhere in a country, and, consequently, there may be no need to collect such information in a census at all. Conversely, some topics may not be included in a census because of the almost total absence of certain amenities, particularly in the rural areas of some developing countries.

4.405. The importance of involving stakeholders in the process of identifying priorities and policy needs has to be taken into consideration early in the process of designing the housing census. The topics that are of particular interest to policymakers need to be carefully assessed in terms of applicability, reliability of data and census limitations (number of questions, and so forth). More detailed information on involvement of stakeholders is presented in Part Two, in Chapter VIII, on "User consultation, communication and publicity" (paragraphs 2.98–2.113), and also in the *Handbook on Census Management for Population and Housing Censuses*.[157]

B. International comparability

4.406. The desirability of achieving regional and worldwide comparability should be another major consideration in the selection and formulation of topics to be included in the census. National and international objectives are usually compatible, since broad studies of countries' experiences and practices are the basis of international recommendations.

4.407. If particular circumstances within a country necessitate a departure from international standards, every effort should be made to explain these departures in the census publications and to indicate how the national presentation can be adapted to the international standards.

[157] United Nations publication, Sales No. E.00XVII.15 Rev.1.

C. Suitability

4.408. A prerequisite for the inclusion of housing topics in the census should be the willingness and ability of respondents to provide accurate information on them. It is advisable to avoid topics that could increase the burden on respondents and those that are likely to arouse fear, local prejudice or which might be used to deliberately promote political or sectarian causes, as these are likely to have a detrimental effect on response rates and support of the census. In an interview-based census or where the collector needs to obtain information through observation, consideration needs also to be given to the level of knowledge and skills of the interviewers or collector and whether they can be adequately trained to collect this information accurately. Topics that are too complicated or difficult for the average respondent or enumerator to answer quickly should also not be included. The exact phrasing of questions that will obtain the most reliable responses may depend on national circumstances and, as described in Part Three, should be well tested prior to the census (see paragraph 3.28).

D. Resources

4.409. The selection of topics should be carefully considered in relation to the total resources available for the census. An efficient collection of accurate data for a limited number of topics, followed by prompt tabulation and publication, is more useful than the collection of data for an overambitious list of topics that cannot be properly processed and disseminated. In balancing the need for data against resources available, the extent to which questions can be precoded is yet another consideration. This may be an important factor in determining whether or not it is economically feasible to include certain topics in the census.

E. Alternative sources

4.410. In the selection of topics to be investigated in a housing census, consideration should be given to whether data are available from other sources, taking into account the relative advantages and limitations of the alternative sources. Those topics for which no alternative sources exist should be given higher priority while those for which alternative sources are readily available should be accorded lower priority.

II. List of topics

4.411. The units of enumeration for housing censuses are buildings, living quarters, households and occupants. The building is often an indirect but important unit of enumeration for housing censuses since the information concerning the building (building type, material of construction of external walls and certain other characteristics) is required to describe properly the living quarters located within the building and for the formulation of housing programmes. In a housing census, the questions on building characteristics are normally framed in terms of the building in which sets of living quarters being enumerated are located, and the information is recorded for each of the housing units or other sets of living quarters located within it.

4.412. The principal direct enumeration unit in a housing census is the living quarters. Only by recognizing this as such can data be obtained that will provide a meaningful description of the housing situation and a suitable basis for the formulation of housing programmes.

4.413. The second direct unit of enumeration is the households occupying the living quarters. For each household, it is often useful to collect information on the characteristics of the head or reference person, tenure in the housing unit, and other relevant characteristics.

4.414. The final units of enumeration are the occupants within households. However, the detailed characteristics of each individual household member are collected in a population census and are covered in Chapter I.

4.415. The list presented below (Table 4) is based on the global and regional census experience of the last several decades. The topics included are those on which there is considerable agreement on their importance and feasibility for inclusion in a census for the purpose of measuring and evaluating housing conditions and formulating housing programmes. Those that are likely to present difficulties and require time-consuming questioning can probably best be investigated in a separate housing survey of a sample of living quarters.

4.416. Core topics are those of common interest and value to countries and also of importance in enabling comprehensive comparison of statistics at the international level. Other topics refer to data that need to be collected in order to meet the additional requirements of national users.

4.417. It should be emphasized that the topics or variables on housing contained herein are for tabulation and production of outputs as this is the overall orientation of these guidelines. Issues that pertain to data collection are addressed in other parts of the *Principles and Recommendations for Population and Housing Censuses* and other relevant United Nations handbooks.

Table 4. Housing census topics by unit of enumeration

No.	Topic	Living quarter		Building	Household
		Housing unit	Collective living quarter		
1	**Living quarters – type of** (paras. 4.421– 4.462)	□	♦		
2	**Location of living quarters** (paras. 4.463– 4.470)	□	♦	□	□
3	**Occupancy status** (paras. 4.471– 4.475)	♦			
4	**Ownership – type of** (paras. 4.476– 4.481)	♦			
5	**Rooms – number of** (paras. 4.482– 4.484)	♦			□
6	Bedrooms – number of (paras. 4.485– 4.486)	○			○
7	Useful floor space – (paras. 4.487– 4.489)	○	○		○
8	**Water supply system** (paras. 4.490– 4.493)	♦	○		□
9	**Drinking water – main source of** (paras. 4.494– 4.495)	♦	○		□
10	**Toilet – type of** (paras. 4.496– 4.499)	♦	○		□
11	**Sewage disposal** (para. 4.500)	♦			□
12	**Solid waste disposal – main type of** (paras. 4.501– 4.502)	♦			□
13	**Bathing facilities** (paras. 4.503– 4.505)	♦	○		□
14	**Kitchen – availability of** (paras. 4.506– 4.509)	♦	○		□
15	**Fuel used for cooking** (para. 4.510)	♦			□
16	**Lighting and/or electricity – type of** (paras. 4.511– 4.512)	♦	○		□
17	Heating – type and energy used (paras. 4.513– 4.514)	○			○
18	Hot water – availability of (para. 4.515)	○			○
19	Piped gas – availability of (para. 4.516)	○			○
20	Use of housing unit (paras. 4.517– 4.518)	○			○
21	**Occupancy by one or more households** (paras. 4.519– 4.523)	□			♦
22	**Occupants – number of** (paras. 4.524– 4.525)	♦	♦		□
23	**Building – type of** (paras. 4.526-4.534)			♦	
24	Year or period of construction (paras. 4.535– 4.539)	○		○	
25	Dwellings in the building – number of (para. 4.540)	○		○	
26	Position of dwelling in the building (paras. 4.541– 4.543)			○	
27	Accessibility to dwelling (para. 4.544)	○			
28	**Construction material of outer walls** (paras. 4.545– 4.547)	♦		♦	
29	Construction material of floor and roof (para. 4.548)	○		○	
30	Elevator – availability of (paras. 4.549– 4.550)	○		○	
31	Farm building (para. 4.551)	○		○	
32	State of repair (paras. 4.552– 4.553)	○		○	
33	**Age and sex of the reference person of the household** (paras. 4.554– 4.555)				♦
34	**Tenure** (paras. 4.556– 4.559)				♦
35	Rental and housing costs (paras. 4.560– 4.562)				○
36	Furnished/unfurnished (para. 4.563)	○			○

No.	Topic	Living quarter		Building	Household
		Housing unit	Collective living quarter		
37	**ICT devices – availability of** (paras. 4.564– 4.571)				♦
	E-waste disposal (paras. 4.564– 4.571)				□
38	Cars – number of available (para. 4.572)				○
39	Durable household appliances – availability of (para. 4.573)				○
40	Outdoor space – access to (para. 4.574)				○
	Legend: ♦ - Core topic □ - Core topic, derived ○ - Additional topic				

III. Definitions and specifications of topics

4.418. Paragraphs 4.421–4.574 below contain the recommended definitions. It is important that census data be accompanied by the definitions used in carrying out the census. It is also important that any changes in definitions that might have been made since the previous census be indicated and, if possible, accompanied by estimates of the effect of such changes on the relevant data. In this way, users will not confuse valid changes over time with increases or decreases that have occurred as the result of changed definitions.

1. Living quarters – type of (core topic)

i. Definition of living quarters

4.419. *Living quarters* are structurally separate and independent places of abode. They (a) may have been constructed, built, converted or arranged for human habitation, provided that they are not at the time of the census used wholly for other purposes and that, in the case of non-conventional housing units and collective living quarters, they are occupied at the time of the census; or (b) though not intended for habitation, were in use for such a purpose at the time of the census.

4.420. In any census with a field enumeration, instructions should be issued to field staff so that it is clearly understood at what stage of completion living quarters should be in order to be included. Living quarters being demolished or awaiting demolition should normally be excluded. The system used should be consistent with that employed for the system of current housing statistics and should avoid double counting where construction statistics are used to bring the census data up to date. Special instructions will need to be issued concerning “core dwellings” in countries where these are provided within a preliminary phase of dwelling construction (see paragraphs 4.438–4.441).

ii. Classification of living quarters

4.421. Living quarters are either housing units or collective living quarters. Normally, the collection of information concerning buildings and housing units located within buildings is of prime importance in a housing census, since it is in buildings and housing units that the majority of the population permanently lives. Furthermore, housing units are intended for occupancy, or are occupied, by households, and it is

with the provision of accommodation for households that housing programmes and policies are mainly concerned. However, certain types of "collective living quarters" are also of significance with respect to the housing conditions of households; these include hotels, rooming houses and other lodging houses and camps occupied by households. Housing units should be classified so as to distinguish conventional dwellings from other types of housing units. It should be emphasized that without an adequate classification of living quarters, no meaningful analysis of housing conditions based on housing census data is possible.

4.422. The classification outlined below (see also Figure 5) and a system of three-digit codes have been designed to group in broad classes housing units and collective living quarters with similar structural characteristics. The distribution of occupants (population) among the various groups provides valuable information about the housing accommodation available at the time of the census. The classification also affords a useful basis of stratification for sample surveys. The living quarters may be classified into the following categories:

1 Housing units
- 1.1 Conventional dwellings
  - 1.1.1 Has all basic facilities
  - 1.1.2 Does not have all basic facilities
- 1.2 Other housing units
  - 1.2.1 Semi-permanent housing units
  - 1.2.2 Mobile housing units
  - 1.2.3 Informal housing units
  - 1.2.4 Housing units in permanent buildings not intended for human habitation
  - 1.2.5 Other premises not intended for human habitation

2 Collective living quarters
- 2.1 Hotels, rooming houses and other lodging houses
- 2.2 Institutions
  - 2.2.1 Hospitals
  - 2.2.2 Correctional institutions (prisons, penitentiaries)
  - 2.2.3 Military institutions
  - 2.2.4 Religious institutions (monasteries, convents, and so forth)
  - 2.2.5 Retirement homes, homes for elderly
  - 2.2.6 Student dormitories and similar
  - 2.2.7 Staff quarters (for example, hostels and nurses' homes)
  - 2.2.8 Orphanages
  - 2.2.9 Other institutional places
- 2.3 Camps and workers' quarters
  - 2.3.1 Military camps
  - 2.3.2 Worker camps
  - 2.3.3 Refugee camps
  - 2.3.4 Camps for internally displaced people (IDPs)
  - 2.3.5 Other camps and worker's quarters
- 2.4 Other collective living quarters

4.423. Not all the categories in the above classification are of importance under all circumstances. For example, in some countries certain categories may not need to be considered separately, while in others it may be convenient to subdivide them. However, some of the categories are of special significance for

assessing the housing situation and should be distinguished even where a simplified classification is employed. The distinction between conventional and informal housing units is referred to particularly.

iii. Definitions of each type of living quarters

4.424. A description of the categories listed in paragraph 4.424 is given below.

1. Housing units

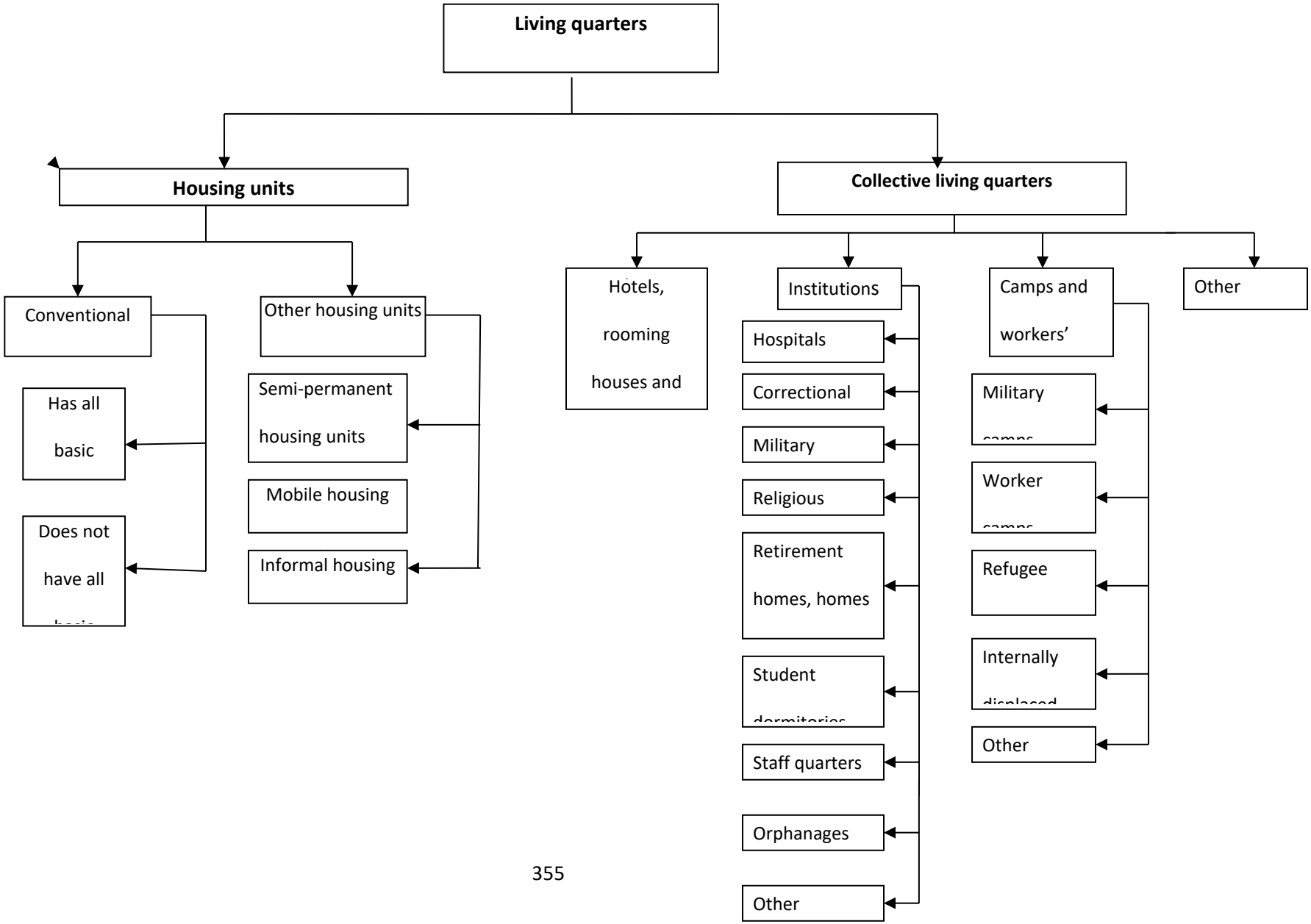
4.425. A *housing unit* is a separate and independent place of abode intended for habitation by a single household,[158] or one not intended for habitation but occupied as living quarters by a household at the time of the census. Thus it may be an occupied or vacant dwelling, an occupied non-conventional housing unit or any other place occupied as living quarters by a household at the time of the census. This category includes housing of various levels of permanency and acceptability and therefore requires further classification in order to provide for a meaningful assessment of housing conditions.

4.426. The essential features of housing units are separateness and independence. An enclosure may be considered separate if surrounded by walls, fences, and so forth, and whether or not covered by a roof so that a person or group of persons can isolate themselves from other persons in the community for the purposes of sleeping, preparing and taking their meals, and protecting themselves from the hazards of climate and environment. Such an enclosure may be considered independent when it has direct access from the street or from a public or communal staircase, passage, gallery or grounds, in other words, when the occupants can come in and go out of their living quarters without passing through anybody else's premises.

4.427. Attached rooms having an independent entrance, or detached rooms for habitation that clearly have been built or rebuilt or converted for use as part of living quarters, should be counted as part of the living quarters. Thus, living quarters may comprise rooms or groups of rooms with independent entrances, or separate buildings.

[158] Although intended for habitation by one household, a housing unit may, at the time of the census, be occupied by one or more households or by a part of a household.

Figure 5. Classification of living quarters



4.428. It should be noted that housing units on the grounds or within the buildings housing an institution, camp, and so forth should be separately identified and counted as housing units. For example, if, on the grounds of a hospital, there is a separate and independent house intended for the habitation of the director and his or her family, the house should be counted as a housing unit. In the same way, self-contained apartments located in hotel buildings should be counted as housing units if they have direct access to the street or to a common space within the building. Similar cases will need to be identified and described in the instructions for the enumeration.

1.1 Conventional dwellings

Note: Section may be updated subsequent to UNHABITAT EGM on Housing to be held in November 2023.

4.429. A "conventional dwelling" is a room or suite of rooms and its accessories in a permanent building or structurally separated part thereof which, by the way it has been built, rebuilt or converted, is intended for habitation by one household and is not, at the time of the census, used wholly for other purposes. It should have a separate access to a street (direct or via a garden or grounds) or to a common space within the building (staircase, passage, gallery and so on). Therefore, there are four essential features of a conventional dwelling:

(a) It is a room or suite of rooms;
(b) It is located in a permanent building;
(c) It has separate access to a street or to a common space;
(d) It was intended to be occupied by one household.

4.430. Examples of conventional dwellings are houses, flats, suites of rooms and apartments. Although a conventional dwelling is a housing unit intended, that is to say, constructed or converted, for habitation by one household, it may, at the time of the census, be vacant or occupied by one or more households. It may be noted that the terms dwelling, dwelling unit, dwelling house, residential dwelling unit, family dwelling, house, *logement*, *vivienda*, *unidad de vivienda* and so forth have been used indiscriminately to refer to housing units of any type. The referent of the term "dwelling" is here limited to a housing unit located in a permanent building and designed for occupancy by one household.

4.431. A "permanent building" is understood to be a structure that is not intended to be moved and that may be expected to maintain its stability for 15 years or more, depending on the way countries define durability. It is recognized that the criterion of permanency or durability (particularly with respect to the significance of materials and methods of construction) may be difficult to apply either in the field or from information held in administrative records and that its adaptation to local conditions would require considerable study and experimentation by the national offices. In some cases, it may be of greater significance nationally to apply the criteria of construction materials and methods of construction directly in order to establish whether or not the building containing the housing unit is of permanent construction, rather than to translate these criteria into a time period.

1.1.1. Conventional dwelling – has all basic facilities

4.432. A conventional dwelling that has all basic facilities refers to a unit that meets all the needs of the household within its confines, such as protection from elements, cooking and maintaining hygiene. Thus, in addition to the four essential features of a conventional dwelling described in paragraph 2.431, all of the following facilities must be available for a dwelling to fall in this category:

(a) Piped water within dwelling;
(b) Toilet within dwelling;
(c) Fixed bath or shower within dwelling;
(d) Kitchen or other space for cooking within dwelling.

1.1.2. Conventional dwelling – does not have all basic facilities

4.433. The conventional dwellings that fall in this category are dwellings that have the essential features of a conventional dwelling (see paragraph 4.431) and some, but not all, of the basic facilities described in paragraph 4.434.

4.434. With increased urbanization, the need for building low-cost housing units within the city limit has been increasing in many countries. This housing most frequently consists of buildings containing a number of separate rooms whose occupants share some or all facilities (bathing, toilet or cooking facilities). Those units do not meet all the criteria of a conventional dwelling with all basic facilities available within the dwelling, especially from the point of view of maintaining health standards and privacy. For example, these units are known as *casa de palomar* in Latin America.

1.2 Other housing units

Note: Section may be updated subsequent to UNHABITAT EGM on Housing to be held in November 2023.

1.2.1. Semi-permanent housing unit

4.435. The term “semi-permanent housing unit” refers to a structure that, by the way it has been built, is not expected to maintain its durability for as long a period of time as a conventional dwelling, but has some of the main features and facilities of a conventional dwelling. As discussed earlier, durability needs to be specifically defined on the basis of national standards and practices. The number of these units in some countries and areas may be substantial. Semi-permanent housing is not to be confused with informal housing units.

4.436. For example, in some countries “core” or “nuclear” dwellings around which a dwelling will eventually be constructed are provided as part of the housing programmes. In others, a significant proportion of the housing inventory is composed of dwellings that are constructed of locally available raw materials and may be less durable than conventional dwellings.

4.437. Many countries with insufficient resources to meet their housing needs have attempted to alleviate the housing conditions of the population living in squatter areas by providing core or nuclear dwellings. Under these programmes, the households move their improvised shacks from the squatter area to a new location, the idea being that gradually, and generally with government assistance, the households with core or nuclear dwellings will keep adding to the nucleus until they can abandon their shacks entirely.

4.438. A core dwelling is sometimes only a sanitary unit containing bathing and toilet facilities, to which may be added, in subsequent phases, the other elements that will finally make up the completed dwelling. Such units do not fall within the definition of a conventional dwelling. However, although the household obviously continues to occupy its original shelter (which would probably be classified as an “improvised

housing unit"), its housing situation is a vast improvement over that of households remaining in the squatter areas and the provision of the cores is a significant step towards the alleviation of housing shortages.

4.439. The problem is thus one of reflecting in the statistics the improvements brought about by programmes such as those described above without distorting the data that refer to fully constructed conventional dwellings. It is recommended, therefore, that core dwellings should be counted as dwellings in the census if at least one room[159] in addition to the sanitary facilities, is completed, and also that those dwellings that have not reached this stage of completion should be recorded as cores. Arrangements should be made so that the facilities available in the core can be related during data processing to the households for whose use they have been provided.

4.440. In other countries and areas, the population has developed, over time, a traditional and typical type of housing unit that does not have all the characteristics of conventional dwellings but is considered somewhat suitable from the point of view of climate and tradition. This is especially the case in many tropical and subtropical rural areas where housing units have been constructed or built with locally available raw materials such as bamboo, palm, straw or any similar materials. Such units often have mud walls, thatched roofs and so forth, and may be expected to last only for a limited time (from a few months to several years), although occasionally they may last for longer periods. This category is intended to cover housing units that are typical and traditional in many tropical rural areas. Such units may be known, for example, as cabins, *ranchos* or *bohíos* (Latin America), *barastis* (Bahrain), or *bahay kubo* (the Philippines).

1.2.2. Mobile housing units

4.441. A "mobile housing unit" is any type of living accommodation that has been produced to be transported (such as a tent) or is a moving unit (such as a ship, boat, barge, vessel, railroad car, caravan, trailer or yacht) occupied as living quarters at the time of the census. Trailers and tents used as permanent living quarters may be of special interest in some countries.

4.442. Although mobile housing units are significantly different from other housing units in that they can be readily moved or transported, mobility in itself is not necessarily an indicator of low quality. For the assessment of housing conditions in countries with a substantial number of mobile units, it may be useful to classify them further as tents, wagons, boats, trailers or other unit types.

1.2.3 Informal housing units

4.443. The term "informal housing unit" refers to a unit that does not have many of the features of a conventional dwelling and are generally characterized as unfit for human habitation, but that are used for that purpose at the time of the census. Therefore, it is neither a permanent structure nor one equipped with any of the essential facilities. Depending on national circumstances, countries should develop detailed instructions to distinguish between informal and semi-temporary housing units.

4.444. Informal housing units comprise three subgroups: "improvised housing units", "housing units in permanent buildings not intended for human habitation" and "other informal housing units". These units are characterized by the fact that they are either makeshift shelters constructed of waste materials and generally considered unfit for habitation (squatters' huts, for example) or places that are not intended for

[159] For the definition of "room", see paragraph 2.482.

human habitation although in use for that purpose at the time of the census (barns, warehouses, natural shelters and so on). Under almost all circumstances, such places of abode represent unacceptable housing and they may be usefully grouped together in order to analyse the housing conditions of the population and to estimate housing needs. Each subgroup is defined below.

1.2.3.1 Improvised housing units

4.445. An improvised housing unit is an independent, makeshift shelter or structure, built of waste materials and without a predetermined plan for the purpose of habitation by one household, which is being used as living quarters at the time of the census. Included in this category are squatters' huts, *poblaciones callampas* (Chile), *hongos* (Peru), *favelas* (Brazil), *sarifas* (Iraq), *barong barong* (the Philippines) and any similar premises arranged and used as living quarters, though they may not comply with generally accepted standards for habitation, and may not have many of the characteristics of conventional dwellings. This type of housing unit is usually found in urban and suburban areas, particularly at the peripheries of the principal cities.

4.446. There is a wide variation in the procedures and criteria used in classifying these units. There are many borderline cases, and countries will need to make decisions and issue detailed instruction on how to enumerate and classify such housing units.

1.2.3.2 Housing units in permanent buildings not intended for human habitation

4.447. Included in this category are housing units (in permanent buildings) that have not been built, constructed, converted or arranged for human habitation but that are actually in use as living quarters at the time of the census. These include housing units in stables, barns, mills, garages, warehouses, offices, booths and so forth.

4.448. This category also may cover units and their occupants in buildings initially built for human habitation but later abandoned with all services cut because of deterioration. These dilapidated buildings can be found, especially in large cities, still standing, although marked for demolition. They should be included in this category if inhabited.

4.449. Premises that have been converted for human habitation, although not initially designed or constructed for this purpose, should not be included in this category, but classified as "other informal housing units".

1.2.3.3 Other informal housing units

4.450. This category refers to living quarters that are not intended for human habitation or located in permanent buildings but that are nevertheless being used as living quarters at the time of the census. Caves and other natural shelters fall within this category.

Slums: UN-Habitat defines a slum household in operational terms, as lacking one or more of the following indicators: a durable housing structure; access to clean water; access to improved sanitation; sufficient living space; and secure tenure. The first four rely on conventional definitions; the last is the most difficult to assess and is not currently used in slum measurement (UN-Habitat, 2003).

Informal Settlements: are residential areas where: 1. Inhabitants have no security of tenure vis-à-vis the land or dwellings they inhabit, with modalities ranging from squatting to informal rental housing, 2. The neighbourhoods usually lack, or are cut off from basic services and formal city infrastructure, 3. The housing may not comply with current planning and building regulations, situated in geographically and environmentally hazardous areas, and may lack a municipal permit.

Inadequate Housing: A housing unit is considered adequate if at a minimum it meets the following criteria:
1. Legal security of tenure, which guarantees legal protection against forced evictions, harassment and other threats;
2. Availability of services, materials, facilities and infrastructure, including safe drinking water, adequate sanitation, energy for cooking, heating, lighting, food storage or refuse disposal;
3. Affordability, as housing is not adequate if its cost threatens or compromises the occupants' enjoyment of other human rights;
4. Habitability, as housing is not adequate if it does not guarantee physical safety or provide adequate space, as well as protection against the cold, damp, heat, rain, wind, other threats to health and structural hazards;
5. Accessibility, as housing is not adequate if the specific needs of disadvantaged and marginalized groups are not taken into account (such as the poor, people facing discrimination; persons with disabilities, victims of natural disasters);
6. Location, as housing is not adequate if it is cut off from employment opportunities, healthcare services, schools, childcare centres and other social facilities, or if located in dangerous or polluted sites or in immediate proximity to pollution sources; and
7. Cultural adequacy, as housing is not adequate if it does not respect and take into account the expression of cultural identity and ways of life.

2. Collective living quarters

4.451. *Collective living quarters* include structurally separate and independent places of abode intended for habitation by large groups of individuals or several households and occupied at the time of the census. Such quarters usually have certain common facilities, such as cooking and toilet installations, baths, lounge rooms or dormitories, which are shared by the occupants. They may be further classified into hotels, rooming houses and other lodging houses, institutions and camps.

4.452. Housing units in the grounds or within the building housing an institution, camp, hotel and so forth should be separately identified and counted as housing units.

4.453. The criteria established for the identification of collective living quarters are not always easy to apply and it is sometimes difficult for an enumerator to decide whether living quarters should be classified as a housing unit or not. This is particularly true in the case of a building occupied by a number of households. Enumerators should be given clear instructions as to when the premises occupied by a group of people living together are to be considered a housing unit and when collective living quarters. This may be less of a problem where census information is collected from administrative data sources and where such buildings are registered as being for communal living.

2.1 Hotels, rooming houses and other lodging houses

4.454. This group comprises permanent structures that provide lodging on a fee basis and in which the number of borders or lodgers exceeds five.[160] Where there are less than five, the living quarters should be classified as a housing unit. Hotels, motels, inns, boarding houses, pensions, lodging houses and similar structures fall within this category. If there is any accommodation within a hotel or similar establishment that is occupied by a household and which fulfils the requirement of a conventional dwelling it should be classified as such.

2.2 Institutions

4.455. This group covers any set of premises in a permanent structure or structures designed to house (usually large) groups of persons who are bound by either a common public objective or a common personal interest. Such sets of living quarters usually have certain common facilities shared by the occupants (for example baths, lounges and dormitories). Hospitals, military barracks, boarding schools, convents, prisons and so forth fall within this category (see the categories in paragraph 4.424).

4.456. It may be useful, depending on national needs, to require that an institution be used as the principle usual residence of at least one person at the time of the census.

2.3 Camps

4.457. Camps are sets of premises originally intended for the temporary accommodation of persons with common activities or interests. Included in this category are military camps, refugee camps and camps established for the housing of workers in mining, agriculture, public work, seasonal works, other types of enterprises, or victims of disasters like earthquake and floods.

2.4 Other

4.458. This is a residual category for collective living quarters that may not conform to the definitions of those included in groups 2.1 through 2.3. It should be used only when the number of units in question is small. Where the number is substantial, additional groups of living quarters that have common characteristics and that are of significance for an improved appraisal of housing conditions should be established.

4.459. In some countries, it seems that certain types of multi-household living quarters have emerged in response to the particular needs of the population, and that the characteristics of these quarters enable them to be readily identified by an enumerator. It may be useful in these countries to provide a separate subgroup for any such special types.

4.460. It should be stressed that the types of living quarters to be included in this category are those intended for communal habitation by several households, that is to say, constructed or converted for this purpose. Housing units intended for occupancy by one household, but those at the time of the census are occupied by several households, are not to be included as collective living quarters because this obscures the identification of households doubling up in dwellings (an important element in estimating housing needs). It is suggested that, in carrying out the census, a strict distinction be maintained between a

[160] The threshold of five lodgers is the one most used. However, depending on national circumstances, this number might be adjusted accordingly.

housing unit occupied by more than one household and living quarters constructed or converted for communal habitation by several households.

2. Location of living quarters (core topic)

4.461. A great deal of information relevant to the location of living quarters is contained under the definition of "locality" and "urban and rural" (see paragraphs 4.89–4.100). It is important for those concerned with carrying out housing censuses to study this information, because the geographic concepts used in carrying out a housing census to describe the location of living quarters are extremely important both for the execution of the census and for the subsequent tabulation of the census results. When the housing census is combined with, or closely related to, a population census, these concepts need to be carefully coordinated so that the geographic areas recognized in carrying out the two censuses are of optimum value for both operations.

4.462. Information on location should be collected in sufficient detail to enable tabulations to be made for the smallest geographic subdivisions required by the tabulation plan. To satisfy the requirements of the geographic classifications recommended in the tabulations as a virtual part of this publication, information is needed on whether the living quarters are located in an urban or rural area, the major civil division, the minor civil division and, for living quarters located in principal localities, the name of the locality.

4.463. Where a permanent system of house or building numbers does not already exist, it is essential for the census to establish a numbering system so that the location of each set of living quarters can be adequately described. Similarly, in cases where streets do not have names or numbers properly displayed, such identification should be provided as one of the pre-census operations. Adequate identification provides the basis for the preparation of census control lists (see also "living quarters and household listing" in paragraphs 3.115–3.118); it is required in order to monitor and control the enumeration, and to identify living quarters for possible callbacks and post-enumeration evaluation surveys (PES) as well as for other post-censal enquiries that use the census as a sampling frame or other point of departure. Ideally, each building or other inhabited structure should be provided with a number, as should each set of living quarters within buildings or structures. In preparing a census control listing, it is the practice to identify further each household within the living quarters. In register-based censuses, it is very important that the frame used does not include under coverage or over coverage and is deduplicated as much as possible. Therefore, periodic quality controls must be carried out.

4.464. Living quarters that are not located in areas with a conventional pattern of streets, such as those in squatter areas or in some places not intended for habitation, may require special identification. Since it may not be possible to describe the location of these units in terms of a formal address, it may be necessary to describe them in terms of their proximity to natural or created landmarks of various kinds or in relation to buildings that are located in areas where a formal address is possible. If possible, coding "slums" separately is of critical importance for countries with prominent slum populations.

4.465. The various geographic designations that together define the location of living quarters are discussed below.

i. Address

4.466. Information that describes the place where the living quarters are to be found and distinguishes them from other living quarters in the same locality falls within this category. As a rule, the information

includes the name or number of the street and the number of the living quarters; in the case of apartments, the building number and the apartment number are required.

ii. Locality

4.467. For the definition of "locality", see paragraphs 4.89–4.91 of the current revision of the *Principles and Recommendations for Population and Housing Censuses.*

iii. Urban and rural

4.468. For the definition of "urban and rural", see paragraphs 4.92–4.100 of the current revision of the *Principles and Recommendations for Population and Housing Censuses.*

3. Occupancy status (core topic)

4.469. *Occupancy status* refers to whether or not a conventional dwelling is occupied at the time of the census. For those dwellings not occupied (because they are vacant or in secondary use), the reason for not being occupied should be classified.

4.470. Information should be obtained for each conventional dwelling to show whether the dwelling is occupied or vacant at the time of the census. For vacant units intended for year-round occupancy, the type of vacancy (for rent or for sale, for example) should be reported. Occupancy status applies only to conventional dwellings, since all other types of living quarters are required by definition to be occupied in order to fall within the scope of the census.

4.471. The enumeration of vacant conventional dwellings is likely to pose difficult problems, but at least a total count should be made for purposes of controlling the enumeration. The type of vacancy is frequently indicated by "for sale" or "for rent" signs posted on the dwelling. Although it may not be feasible to investigate all of the topics included in the census for vacant units, as much information as possible should be collected, including information on whether the living quarters are vacant seasonally or non-seasonally.

4.472. Vacant units intended for seasonal or secondary occupancy may represent a substantial proportion of the housing stock in resort areas and in areas where large numbers of seasonal workers are employed. The separate identification of such categories may be necessary for the correct interpretation of the overall vacancy rate, as well as for an evaluation of the housing situation in the area concerned.

4.473. Whether or not living quarters whose occupants are temporarily absent or temporarily present should be recorded as occupied or vacant will need to be considered in relation to whether a de jure or de facto population census is being carried out. In either case, it would seem useful to distinguish as far as possible conventional dwellings that are used as a second residence. This is particularly important if the second residence has markedly different characteristics from the primary residence, as is the case, for example, when agricultural households move during certain seasons of the year from their permanent living quarters in a village to rudimentary structures located on agricultural holdings. The recommended classification of occupancy status for conventional dwellings is as follows:

1 Occupied
2 Vacant / not occupied

2.1 Seasonally vacant
    2.1.1 Holiday homes
    2.1.2 Seasonal workers' quarters
    2.1.3 Other
2.2 Secondary residences
2.3 For rent/sale
2.4 For demolition
2.5 Other

4. Ownership – type of (core topic)

4.474. This topic refers to the type of ownership of the housing unit itself and not of that of the land on which it stands. Type of ownership should not be confused with tenure, which is a characteristic of the household and is covered in paragraphs 4.556–4.559.

4.475. Information should be obtained to show:

(a) Whether the housing unit is owned by the public sector (central government, local government, public corporations);

(b) Whether the housing unit is privately owned (by households, private corporations, cooperatives, housing associations and so on). The question is sometimes expanded to show whether the housing units are fully paid for, being purchased in instalments or mortgaged. The classification of housing units by type of ownership is as follows:

1 Owner-occupied
2 Non-owner-occupied
    2.1 Publicly owned
    2.2 Privately owned
    2.3 Communally owned
    2.4 Cooperatively owned
    2.5 Other

4.476. Housing units are defined as owner-occupied if used wholly or partly for own occupation by the owner. In principle, if a housing unit is being purchased in instalments or mortgaged according to national legal systems and practices, it should be enumerated as being owned. Instructions should also cover other arrangements, such as housing units in cooperatives or housing associations.

4.477. The information on ownership may be classified, as a minimum, into two main groups, namely "private ownership" and "other ownership". Depending upon the prevalence of various types of ownership and their significance with respect to housing conditions and the formulation of housing programmes, it may be useful to dissect the category "other ownership" into the relevant examples of the subgroups shown. The categories used should be consistent with those employed in the system of national accounts of the country concerned and in accordance with the recommendations contained in the *System of National Accounts, 2008*.[161]

[161] *System of National Accounts 2008*, United Nations publication, Sales No. E.08.XVII.29.

4.478. It has been observed that the collection of information on type of ownership in a general census may be hampered by the fact that the occupants might not know who the owner of the property is and that the owners or their representatives may be situated outside the enumeration zone. Furthermore, there are numerous cases of borderline and mixed ownership, which make the topic difficult for nationwide enumeration. This is one of the topics for which more accurate information might be obtained through a housing survey.

4.479. In countries where there is a substantial amount of employer-issued housing, it would be useful to include the subcategories "issued by the employer" and "not issued by the employer" under the category "privately owned" (or publicly owned where the employer is a public sector entity). It is important that such information be known from the point of view of assessing the impact of job loss, in order to gauge the magnitude of the population whose loss of a job would include loss of housing as well.

5. Rooms – number of (core topic)

4.480. A *room* is defined as a space in a housing unit enclosed by walls reaching from the floor to the ceiling or roof covering, or to a height of at least 2 metres, of an area large enough to hold a bed for an adult, that is, at least 4 square metres. Usually only information on rooms in housing units is collected in a census. The total number of types of rooms therefore includes bedrooms, dining rooms, living rooms, studies, habitable attics, servants' rooms, kitchens, rooms used for professional or business purposes, and other separate spaces used or intended for dwelling purposes, so long as they meet the criteria concerning walls and floor space. Passageways, verandas, lobbies, bathrooms and toilet rooms should not be counted as rooms, even if they meet the criteria. Separate information may be collected for national purposes on spaces of less than 4 square metres that conform in other respects to the definition of "room" if it is considered that their number warrants such a procedure.

4.481. Rooms used exclusively for business or professional purposes should be counted separately, as it is desirable to include them when calculating the number of rooms in a dwelling but to exclude them when calculating the number of persons per room. This procedure allows density levels to be studied according to the number of rooms available for living purposes in relation to the number of occupants. In any event, each country should indicate the procedure that has been followed.

4.482. It is recommended that kitchens be included in the count of rooms provided they meet the criteria concerning walls and floor space. Kitchens or kitchenettes that have an area smaller than 4 square metres or that have other characteristics that disqualify them should be excluded. For national purposes, countries may wish to identify and count kitchens within a separate group that may be analysed with respect to size and utilization, and to consider separately those used exclusively for cooking.

6. Bedrooms – number of

4.483. A *bedroom* is defined as a room equipped with a bed and used for night rest.

4.484. In addition to enumerating the number of rooms, some countries may wish to collect information on the number of bedrooms in a housing unit in order to provide a measure of overcrowding.

7. Useful floor space

4.485. This topic refers to the useful and liveable floor space in housing units, that is to say, the floor space measured inside the outer walls of housing units, excluding non-habitable cellars and attics. Information on this topic is often collected to supplement that on the number of rooms. In some countries, however, such information is collected in preference. In multiple-dwelling buildings, all common spaces should be excluded.

4.486. For collective living quarters, it would be more useful to collect information on the useful floor space per occupant of the set of collective living quarters. Data should be derived by dividing the total useful floor space by the number of occupants who are living in the space.

4.487. Collecting information on the floor space available to occupants of housing units may prove to be difficult; occupants often may not know the exact or even the approximate area of the housing unit they occupy, and training enumerators to calculate the floor space would be complicated and costly, and would result in inaccuracies. In this context, and taking into account the importance of the information concerned, countries should take into consideration developing detailed instructions on proper procedures for assessing these data (for example, a request for information on floor space from the official documents available to the occupants, such as the rental agreement and the title, which are supposed to include such information).

8. Water supply system (core topic)

4.488. Basic information to be obtained in the census is whether housing units have or do not have a piped water installation, in other words, whether or not water is provided to the housing unit by pipes from a community-wide system or a private installation, such as a pressure tank or pump. The unit of enumeration for this topic is a housing unit. It is also recommended that countries should indicate whether the unit has tap water inside or, if not, whether it is within a certain distance from the door. The recommended distance is 200 metres, assuming that access to piped water within that distance allows occupants of the housing unit to obtain water for household needs without being subjected to extreme efforts. Besides the location of the tap water relative to the housing unit, the source of water available to households is also of interest. Therefore, the recommended classification of housing unit by water supply system is as follows:

1 Piped water inside the unit
  1.1 From the community scheme
  1.2 From an individual source
2 Piped water outside the unit but within 200 metres
  2.1 From the community scheme
    2.1.1 For exclusive use
    2.1.2 Shared
  2.2 From an individual source
    2.2.1 For exclusive use
    2.2.2 Shared
3 Other (see category 3 of the classification in paragraph 4.495 for more details)

4.489. A community scheme is one that is subject to inspection and control by public authorities. Such schemes are generally operated by a public body but, in some cases, they are operated by a cooperative

or private enterprise. An individual source of water refers to a source of water that is not part of a community scheme, such as an individual or shared water reservoir.

4.490. As noted above the unit of enumeration for this topic is the housing unit. However, some countries may find it useful to collect information on the availability of piped water for the use of occupants in collective living quarters. Such living quarters are usually equipped with multi-facilities for the use of large groups, and information on the water supply system in relation to the number of occupants would be significant in respect of analysing housing conditions. The water supply system in collective living quarters constitutes an additional topic.

4.491. The most significant information from a health point of view is whether the living quarters have piped water within the premises. However, a category may be added to distinguish cases where the piped water supply is not within the living quarters but rather within the building in which the living quarters are situated. It may also be useful to collect information that would show whether the water supply is for the sole use of the occupants of the living quarters being enumerated or whether it is for the use of the occupants of several sets of living quarters, as indicated in the above classification at the three-digit level. Where there is a large proportion of housing units with no piped water, this category may be expanded to specify sources commonly used in a country. Additional information may be sought on the availability of hot as well as cold water and on the kind of equipment used for heating water.

9. Drinking water – main source of (core topic)

4.492. Having enough water for drinking and personal hygiene is essential, but quantity by itself is not sufficient. The quality of the water is also a crucial health issue. Consequently, one of the targets of the water supply, sanitation and hygiene (WASH) post-2015 recommendations[162] proposed by the WHO/UNICEF Joint Monitoring Programme on Water Supply and Sanitation (JMP) is "universal access to basic drinking water, sanitation and hygiene", assessed in part by having access at home to safely managed drinking water. A safely managed drinking water service is defined as one that reliably delivers water that is sufficient to meet domestic needs and does not represent a significant risk to health. This implies a system that delivers water to the household and includes measures to prevent risks and to verify water quality through compliance monitoring. An improved water source (piped water, public tap or standpost, tubewell or borehole, protected dug well, protected spring, rainwater) can be safely managed. Unimproved sources, which by definition are not safely managed, include unprotected dug well, unprotected spring and surface water from a river, stream, dam, lake, pond, canal or irrigation channel. Delivered water (for example, through trucks, carts, sachets or bottles) can potentially be safely managed, but if these are the primary drinking water sources, other improved sources of water must be accessible at the household for other domestic uses (for example, washing, bathing).

4.493. Countries are encouraged to collect the information on the main source of drinking water for the household, particularly where there is considerable difference between sources of water for general household use and for drinking. For those countries wishing to collect this information, the following categories of main source of drinking water are recommended:

1 Piped water inside the unit

[162] See WHO/UNICEF Joint Monitoring Programme on Water Supply and Sanitation (JMP) report entitled *WASH post-2015: proposed targets and indicators for drinking-water, sanitation and hygiene* (April 2014), accessible at http://www.wssinfo.org/.

1.1 From the community scheme
1.2 From an individual source

2 Piped water outside the unit but within 200 metres
2.1 From the community scheme
2.1.1 For exclusive use
2.1.2 Shared
2.2 From an individual source
2.2.1 For exclusive use
2.2.2 Shared

3 Other
3.1 Borehole/tubewell
3.2 Protected dug well
3.3 Protected spring
3.4 Rainwater collection tank
3.5 Delivered water – bottled, sachet
3.6 Delivered water – tanker trucks, carts
3.7 Unprotected dug well/spring/river/stream/lake/pond/dam/canal/irrigation channel

10. Toilet – type of (core topic)[163]

4.494. A *toilet* may be defined as an installation for the disposal of human excreta. A flush toilet is an installation provided with piped water that permits humans to discharge their wastes and from which the wastes are flushed by water. The unit of enumeration for this topic is a housing unit.

4.495. For housing units reported as having a toilet, additional information may be sought to determine whether the toilet is used exclusively by the occupants of the living quarters being enumerated or is shared with the occupants of other living quarters. For living quarters reported as having no toilet, it would be useful to know if the occupants have the use of a communal facility and the type of facility, or if they have the use of the toilet of other living quarters and the type, or if there is no toilet of any kind available for the use of the occupants.

4.496. Some countries have found it useful to expand the classification for non-flush toilets so as to distinguish certain types that are widely used and indicate a certain level of sanitation. The recommended classification of housing unit by toilet facilities is as follows:

1 With toilet within housing unit
1.1 Flush/pour flush[164] toilet
1.2 Other

2 With toilet outside housing unit
2.1 For exclusive use
2.1.1 Flush/pour flush toilet

[163] It is also necessary to distinguish between conventional dwellings with all main facilities and other conventional dwellings.

[164] A pour flush toilet uses a water seal, but unlike a flush toilet, a pour flush toilet uses water poured by hand for flushing (no cistern is used).

2.1.2 Ventilated improved pit latrine[165]
2.1.3 Pit latrine without ventilation with covering
2.1.4 Holes or dug pits with temporary coverings or without shelter
2.1.5 Other
2.2 Shared
2.2.1 Flush/pour flush toilet
2.2.2 Ventilated improved pit latrine
2.2.3 Pit latrine without ventilation with covering
2.2.4 Holes or dug pits with temporary coverings or without shelter
2.2.5 Other
3 No toilet available
3.1 Service or bucket facility (excreta manually removed)
3.2 Use of natural environment, for example, bush, river, stream.

4.497. As noted above the unit of enumeration for this topic is the housing unit. However, some countries may find it useful to collect information on the availability of toilet facilities for the use of occupants in collective living quarters. Living quarters of this type are usually equipped with multi-facilities for the use of large groups, and information on the number and type of toilets in relation to the number of occupants would be significant in terms of analysing housing conditions. The availability of toilets for collective living quarters represents an additional topic.

11. Sewage disposal (core topic)

4.498. Information on toilets should be combined with the sewage disposal system to which they are connected in order to determine the adequacy of sanitation facilities of the housing unit. To be considered adequate sanitation, toilets or latrines have to be connected to non-clogged sewage disposal systems. The information on housing units by type of sewage disposal system may be classified as follows:

1 Empties into a piped system connected to a public sewage disposal plant
2 Empties into a piped system connected to an individual sewage disposal system (septic tank, cesspool)
3 Other – toilet empties into an open ditch, a pit, a river, the sea, and so forth
4 No disposal system

12. Solid waste disposal – main type of (core topic)

4.499. Securing sustainable development and, in this context, the usual manner of treatment of solid waste (garbage) generated by the household, has prompted the incorporation of this topic in a number of national housing censuses.

4.500. This topic refers to the usual manner of collection and disposal of solid waste or garbage generated by occupants of the housing unit. The unit of enumeration is a housing unit. The classification of housing units by type of solid waste disposal is according to the following guidelines:

[165] A ventilated improved pit latrine (VIP) is a dry pit latrine that uses a hole in the ground to collect the excreta and a squatting slab or platform that is firmly supported on all sides, easy to clean and raised above the surrounding ground level to prevent surface water from entering the pit. The platform has a squatting hole, or is fitted with a seat.

1 Solid waste collected on a regular basis by authorized collectors
2 Solid waste collected on an irregular basis by authorized collectors
3 Solid waste collected by self-appointed collectors
4 Occupants dispose of solid waste in a local dump supervised by authorities
5 Occupants dispose of solid waste in a local dump not supervised by authorities
6 Occupants burn solid waste
7 Occupants bury solid waste
8 Occupants dispose solid waste into river, sea, creek, pond
9 Occupants compost solid waste
10 Other arrangement

13. Bathing facilities (core topic)

4.501. Information should be obtained on whether or not there is a fixed bath or shower installation within the premises of each set of housing units. The unit of enumeration for this topic is a housing unit. Additional information may be collected to show whether or not the facilities are for the exclusive use of the occupants of the living quarters and where there is a supply of hot water for bathing purposes or cold water only. In some areas of the world the distinction proposed above may not be the most appropriate for national needs. It may be important, for example, to distinguish in terms of availability among a separate room for bathing in the living quarters, a separate room for bathing in the building, an open cubicle for bathing in the building and a public bathhouse. The recommended classification of housing units by availability and type of bathing facilities is as follows:

1 With fixed bath or shower within housing unit
2 Without fixed bath or shower within housing unit
- 2.1 Fixed bath or shower available outside housing unit
  - 2.1.1 For exclusive use
  - 2.1.2 Shared
- 2.2 No fixed bath or shower available

4.502. Alternatively, and in line with the elaboration in the preceding paragraph, the following classification may be more appropriate in certain circumstances:

1 Separate room for bath or shower within the housing unit
2 No separate room for bath or shower but bathing space available within the housing unit (for example, in an open area around the well within the housing unit, in the courtyard)
3 Bathing room available but outside the housing unit for exclusive use
4 Shared bathing room outside the housing unit
5 No specific bathing room available

4.503. As noted above the unit of enumeration for this topic is the housing unit. However, some countries may find it useful to collect information on the availability of a bath or shower for the use of occupants in collective living quarters as well. Living quarters of this type are usually equipped with multi-facilities for the use of large groups, and information on the number of fixed baths or showers in relation to the number of occupants would be significant in terms of analysing housing conditions. The number of fixed baths or showers in collective living quarters would represent an additional topic.

14. Kitchen – availability of (core topic)

4.504. Information should be obtained on whether the housing unit has a kitchen, whether some other space is set aside for cooking, such as a kitchenette, or whether there is no special place set aside for cooking. The unit of enumeration for this topic is a housing unit.

4.505. A *kitchen* is defined as a space that conforms in all respects to the criteria for a room, and is equipped for the preparation of the principal meals of the day and intended primarily for that purpose.

4.506. Any other space reserved for cooking, such as a kitchenette, will fall short in respect of possessing the attributes of a room, although it may be equipped for the preparation of the principal meals of the day and is intended primarily for that purpose. The collection of data on the availability of a kitchen may provide a convenient opportunity to collect information on the kind of equipment that is used for cooking, for example, a stove, hotplate or open fire, and on the availability of a kitchen sink and a space for food storage so as to prevent spoilage. The recommended classification of housing units by availability of a kitchen or other space reserved for cooking within the housing unit is as follows:

1 With kitchen within housing unit
- 1.1 For exclusive use
- 1.2 Shared

2 With other space for cooking within housing unit, such as kitchenette
- 2.1 For exclusive use
- 2.2 Shared

3 Without kitchen or other space for cooking within housing unit
- 3.1 Kitchen or other space for cooking available outside housing unit
  - 3.1.1 For exclusive use
  - 3.1.2 Shared
- 3.2 No kitchen or other space for cooking available

4.507. As noted above the unit of enumeration for this topic is the housing unit. However, some countries may find it useful to collect information on the availability of kitchen facilities for the use of occupants in collective living quarters. Living quarters of this type are usually equipped with multi-facilities for the use of large groups, and information on the number of kitchens or kitchenettes in relation to the number of occupants would be significant in terms of analysing housing conditions. It represents an additional topic.

15. Fuel used for cooking (core topic)

4.508. The proportion of households using solid fuels is one of the indicators used in monitoring internationally agreed development goals. There are important linkages between household solid fuel use, indoor air pollution, deforestation and soil erosion and greenhouse gas emissions. The type of fuel and participation in cooking tasks are important predictors of exposure to indoor air pollution. It is thus recommended to collect information on the fuel used for cooking by each housing unit. Fuel used for cooking refers to the fuel used predominantly for preparation of principal meals. If two fuels (for example, electricity and gas) are used, the one used most often should be enumerated. The classification of fuels used for cooking depends on national circumstances and may include electricity, gas, oil, coal, firewood and animal dung. It would also be useful to collect this information for collective living quarters, especially

if the number of sets of collective living quarters in the country is significant. The classification of fuel used for cooking is as follows:

1 Gas
2 Electricity
3 Liquefied petroleum gas (LPG)
4 Kerosene/paraffin (petroleum-based)
5 Oil (including vegetable oils used as fuel)
6 Coal
7 Firewood
8 Charcoal
9 Animal dung
10 Crop residues (for example, cereal straw from maize, wheat, paddy rice, rice hulls, coconut husks, groundnut shells)
11 Other

16. Source of energy for Lighting and/or electricity – type of (core topic)

4.509. Information should be collected on the main source of energy used for lighting in the housing unit, such as electricity, gas or oil lamp. If the source of energy for lighting is electricity, some countries may wish to collect information showing whether the electricity mainly comes from a community supply, private generating plant or some other source (industrial plant, mine and so on). In addition to the type of lighting, countries may assess the information on the availability of electricity for purposes other than lighting (such as cooking, heating water and heating the premises). If housing conditions in the country allow this information to be derived from the type of lighting, there would be no need for additional enquiry.

4.510. As noted above the unit of enumeration for this topic is the housing unit. However, some countries may find it useful to collect information on the availability of electricity for the use of occupants in collective living quarters. Such living quarters are usually equipped with multi-facilities for the use of large groups, and information on electricity would be significant in terms of analysing housing conditions. The availability of electricity in collective living quarters is defined as an additional topic. No classification is specifically recommended.

17. Heating – type and energy used

4.511. This topic refers to the type of heating of housing units and the energy used for that purpose. The units of enumeration are all housing units. This topic may be less relevant for a number of countries where, owing to their geographic position and climate, there is no need to provide energy for heating.

4.512. *Type of heating* refers to the kind of system used to provide heating for most of the space: it may be central heating serving all the sets of living quarters or serving a set of living quarters, or it may not be central, in which case the heating will be provided separately within the living quarters by a stove, fireplace or some other heating body. As for the energy used for heating, it is closely related to the type of heating and refers to the predominant source of energy, such as solid fuels (coal, lignite and products of coal and lignite, wood), oils, gaseous fuels (natural or liquefied gas), or electricity. No classification is specifically recommended.

18. Hot water – availability of

4.513. This topic refers to the availability of hot water in housing units. Hot water denotes water heated to a certain temperature and conducted through pipes and tap to occupants. The information collected may indicate whether there is hot water available within the housing units, or outside the living quarters for exclusive or shared use, or not at all. No classification is specifically recommended.

19. Piped gas – availability of

4.514. This topic refers to whether piped gas is available in the housing unit or not. Piped gas is usually defined as natural or manufactured gas that is distributed by pipeline and whose consumption is recorded. This topic may be irrelevant for a number of countries where there is either a lack of sources of natural gas or no developed pipeline system. No classification is specifically recommended.

20. Use of housing unit

4.515. *Use of housing unit* refers to whether the housing unit is being used wholly for habitation (residential) purposes or not. The housing unit can be used for habitation and for commercial, manufacturing or some other purposes. In a number of countries, houses are used simultaneously for more than one purpose. For example, the lower floor is used as a store or workshop, and the upper floors for habitation.

4.516. The recommended classification of the use of the housing unit is as follows:
1 Used solely for habitation
2 Used for habitation and economic activity

21. Occupancy by one or more households (core topic)

4.517. For the purpose of a housing census, each household must be identified separately. With respect to housing programmes, the use of the separate concepts of “household” and “living quarters” in carrying out housing censuses permits the identification of the persons or groups of persons in need of their own dwellings. If the household is defined as a group of persons occupying a set of living quarters, the number of households in the living quarters and the number of sets of occupied living quarters will always be equal and there will be no apparent housing need as reflected by the number of “sharing” households that require their own living quarters. If living quarters are defined as the space occupied by a household, the number of households in living quarters will again be equal to the number of sets of living quarters, with the added disadvantage that there will be no record of the number of structurally separate living quarters.

4.518. Occupancy by more than one household is a useful topic for assessing the current housing situation and measuring the need for additional housing. For countries relying on the housekeeping concept (see paragraph 2.34), the number of households occupying a housing unit is needed to understand the extent of shared housing. For countries relying on the dwelling unit concept of households (see paragraph 2.35), the household is equivalent to the dwelling unit.

4.519. In countries where it is traditional to count families, the family in the broad sense of the term may be adopted as an additional unit of enumeration; in the great majority of cases the composition of this unit will coincide with that of the household.

4.520. A household and family should be defined in the same way for housing census purposes as for population censuses (see paragraphs 4.121–4.127 and 4.140–4.145).

4.521. For the definitions of "household", "reference person of household" and "persons living in institutions", see paragraphs 4.121–4.148 and 2.39–2.40 in the current revision of the *Principles and Recommendations for Population and Housing Censuses*.

22. Occupants – number of (core topic)

4.522. Each person usually resident in a housing unit or in collective living quarters should be counted as an occupant. Therefore, the units of enumeration for this topic are living quarters. However, since housing censuses are usually carried out simultaneously with population censuses, the applicability of this definition depends upon whether the information collected and recorded for each person in the population census indicates where he or she was on the day of the census or whether it refers to the usual residence (see paragraphs 4.52–4.63). Care should be exercised in distinguishing persons occupying mobile units, such as boats, caravans and trailers, as living quarters from persons using these units as a means of transportation.

4.523. Depending on their national requirements for information, some countries may wish to distinguish between those occupants that are usually resident and those that are not usually resident in the living quarters for the purposes of better understanding the housing conditions and living arrangements of non-residents.

23. Building – type of (core topic)

a. Definition of building

4.524. A *building* is any independent free-standing structure comprising one or more rooms[166] or other spaces, covered by a roof and usually enclosed within external walls or dividing walls[167] that extend from the foundations to the roof. However, in tropical areas, a building may consist of a roof with supports only, that is to say, without constructed walls; in some cases, a roofless structure consisting of a space enclosed by walls may be considered a "building" (see also "compound" in paragraph 4.534).

4.525. In defining a "building", particular care should be given to differentiating this from "type of living quarters" (see paragraph 4.421). Type of living quarters refers to structures that are designed for residential habitation or are being used for residential habitation. A building could be a number of living quarters, a commercial premises not meant, or being used, for habitation, or a mix of the two.

4.526. A building may be used or intended for residential, commercial or industrial purposes or for the provision of services. It may therefore be a factory, shop, detached dwelling, apartment building, warehouse, garage, barn and so forth. In some exceptional cases, facilities usually provided by a set of living quarters are located in two or more separate detached structures, as when a kitchen is in a separate structure. In the case of living quarters with detached rooms, these rooms should be considered separate

[166] For the definition of "room", see paragraph 2.482.

[167] The term "dividing walls" refers to the walls of adjoining buildings that have been so constructed as to be contiguous, for example, the dividing walls of "row" houses.

buildings. A building may therefore contain several sets of living quarters, as is the case for an apartment building or duplex; it may be coextensive with ~~a~~ single detached living quarters; or it may be only part of the living quarters, as is the case, for example, for living quarters with detached rooms, which are clearly intended to be used as part of the living quarters.

4.527. The concept of a building should be clearly defined and, in a census with a field enumeration, the instructions should indicate whether all buildings are to be listed and enumerated or only those used in whole or in part for residential purposes. Instructions should also indicate whether buildings under construction are to be recorded and, if so, at what stage of completion they are to be considered eligible for inclusion. Buildings being demolished or awaiting demolition should normally be excluded.

b. Classification of buildings by type

4.528. The following classification of buildings (or of living quarters) by type of building is recommended:

1.0 Residential buildings
- 1.1 Buildings containing a single housing unit
  - 1.1.1 Detached
  - 1.1.2 Attached
- 1.2 Buildings containing more than one housing unit
  - 2.2.1 Up to 2 floors
  - 2.2.2 From 3 to 4 floors
  - 2.2.3 From 5 to 10 floors
  - 2.2.4 11 floors or more
- 1.3 Buildings for persons living in institutions
- 1.4 Other residential buildings

2.0 Non-residential buildings

4.529. It should be noted that, for the purpose of the housing census, the above classification refers to the building in which the sets of enumerated living quarters are located and that usually it will be the living quarters, not buildings, that will be tabulated according to the classification.

4.530. Category 1.1 provides separate subgroupings for “detached” and “attached” buildings because, although most single-unit buildings (suburban homes, villas, and so forth) are detached, in some countries a substantial number may be attached (row or terraced houses, for example) and in such cases it may be useful to identify these separately. According to the definition of “building” in paragraph 4.526 above, a group of, for example, three row or terraced houses that are attached is considered to be three separate buildings if their “external walls or dividing walls” extend from “the foundations to the roof”. Buildings containing more than one housing unit (category 1.2) will usually be apartment buildings, but they may also be other types of buildings, for example, buildings that are structurally subdivided so as to contain more than one housing unit. Buildings under the latter category should be subdivided into the following: up to 2 floors, from 3 to 10 floors and 11 floors or more. Category 1.3, “buildings for persons living in institutions”, includes hospital buildings, prisons, military establishments, and so on. On the other hand, a structurally separate housing unit (a house or apartment intended for the occupancy of staff of the institution) or one that is either within a building of the institution or detached but within the grounds, belongs in category 1.0; if the housing unit is coextensive with a building, it belongs in category 1.2.

4.531. In addition to the above, and for subsequent analysis of housing conditions, each country will find it useful to provide for separate identification of the special types of buildings that are characteristic of the country concerned. These can be classified as category 4.

c. Compound

4.532. In some countries, it may be appropriate to use the "compound" as a unit of enumeration. In some areas of the world, housing units are traditionally located within compounds, and the grouping of sets of housing units in this way has economic and social implications that need to be studied. A compound, in these circumstances, becomes a distinct unit of enumeration, on a par with a housing unit. For purposes of international comparability, a compound should be classified according to the main features and facilities it displays and classified with housing units.

24. Year or period of construction

4.533. This topic refers to the age of the building in which the living quarters are located. It is recommended that the exact year of construction be sought for buildings constructed during the decennial period immediately preceding the census. Exact year of construction refers to the year the building was considered complete and not the year construction started). For buildings constructed before that time, the information should be collected in terms of periods that will provide a useful means of assessing the age of the housing stock. Difficulty may be experienced in collecting data on this topic in a field enumeration because in some cases the occupants may not know the date of construction. However, more accurate information is more likely to be available where countries use housing registers or other administrative data sources for the census.

4.534. The collection of data for single years during the most recent intercensal period is seen as a method of checking construction statistics for deficient coverage and of more closely integrating the housing census with current housing statistics.

4.535. Instead of collecting single years of construction, if this is seen to be too burdensome on the respondent, periods of construction should be collected. The periods could be defined in terms of events that have some special significance in the country concerned, particularly with regard to the effect on the condition of the housing stock; examples would be the period since the Second World War; the period between the First World War and the Second World War; and the period before a major earthquake, flood or fire. Alternatively, the response ranges could be equal to intervals from one census to the next, such as ten- or five-year age groupings depending on the frequency of census collection. This allows for comparisons across the same periods and across censuses. Narrow periods of construction are most important in the first few decades of a dwelling when the dwelling is undergoing changes, such as foundation setting, or when defects in dwelling systems, such as electrical or plumbing faults, reveal themselves. Afterwards the ranges could widen, but should be as homogeneous as possible to allow for cohort analysis. The total period covered by the age groups and the number of groups distinguished will depend upon the materials and methods of construction used in the country concerned and the number of years that buildings normally last.

4.536. Where parts of buildings have been constructed at different times, the year or period of construction should refer to the major part. Where living quarters comprise more than one building (living quarters with detached rooms, for example), the age of the building that contains the major part of the living quarters should be recorded.

4.537. In countries where a significant number of households construct their own living quarters (countries with large non-monetary sectors, for example), it may be useful to collect additional information that will distinguish the living quarters according to whether or not they were constructed by the households occupying them. The information should refer only to living quarters constructed during the preceding intercensal or 10-year period, and it should be made clear in formulating the question that it refers to living quarters constructed mainly by households (with or without the help of other households in the community) and not to construction executed by enterprises on behalf of households.

25. Dwellings in the building – number of

4.538. This topic refers to the number of conventional dwellings in the building. This topic is applicable in cases where there is a possibility to have unique identifier for the building itself. If a census established such an identifier (building number, for example, linked to the address) then it would be possible to introduce this topic.

26. Position of dwelling in the building

4.539. Some countries may want to collect information on the position of the dwelling or housing unit in the building. This information can be used as an indicator of accessibility to dwellings, possibly in conjunction with information on the accessibility to the dwellings.

4.540. The following classification of dwellings by position in the building is recommended:

1.0 Dwelling on one floor only
- 1.1 Dwelling below the ground floor
- 1.2 Dwelling on the ground floor of the building
- 1.3 Dwelling on the 1st or 2nd floor of the building
- 1.4 Dwelling on the 3rd or 4th floor of the building
- 1.5 Dwelling on the 5th floor of the building or higher

2.0 Dwellings on two or more floors
- 2.1 Dwelling on the ground floor of the building or below ground level
- 2.2 Dwelling on the 1st or 2nd floor of the building
- 2.3 Dwelling on the 3rd or 4th floor of the building
- 2.4 Dwelling on the 5th floor of the building or higher

4.541. For dwellings on two or more floors, information should be provided with reference to the lowest floor level of the dwelling.

27. Accessibility to dwelling

4.542. The following classification of accessibility to the front door of the dwelling or housing unit is recommended, based on the presence of ramps, steps and lifts:

1 Access with no steps or ramp
2 Access by ramp
3 Access by disabled stair lift
4 Access using lift only (though the building may have staircases as well)

5 Access by using only steps
6 Access only by using both lift and steps

Note that these categories are not necessarily mutually exclusive.

28. Construction material of outer walls (core topic)

4.543. This topic refers to the construction material of external (outer) walls of the building in which the living quarters are located. If the walls are constructed of more than one type of material, the predominant type of material should be reported. The types of materials distinguished will depend upon the materials most frequently used in the country concerned and on their significance from the point of view of permanency of construction or assessment of durability. The following classification of construction materials is recommended:

2.2.4 Burnt clay (bricks, blocks, panels), stone, concrete
2.2.5 Unburnt clay (bricks, blocks) mud, earth
2.2.6 Wood
2.2.7 Bamboo, trees, grass
2.2.8 Corrugated sheets
2.2.9 Prefabricated units
2.2.10 Other materials

4.544. In some countries, the material used for the construction of roofs or of floors may be of special significance for the assessment of durability and, in such cases, it may be necessary to collect information on this as well as on the material of the walls. Durability refers to the period of time for which the structure remains habitable, subject to regular maintenance. A durable structure is one expected to remain sound for a considerable period of time. Countries may wish to define the length of the period, for example, 15 or 20 years. Durability does not depend solely on the materials used in construction, since it is also affected by the way the building was erected, that is to say, whether it was built according to construction standards and regulations. Technological developments in treating traditional building materials, such as bamboo, have extended the durability of those materials for several decades. Construction material of outer walls may be considered an indicator of the building's durability. Therefore, in order to assess quality of the national housing stock, durability may be measured in terms of material used together with adherence to construction standards. Specific instructions for enumerators at the national level should be developed on the basis of national building construction practice.

4.545. While the material of construction is a useful addition to data collected on the type of living quarters, it should not be considered a substitute for the latter topic. Wood, for example, may be the material of both a poorly constructed squatter's hut and a durable and well-constructed dwelling. In these cases, information on the type of living quarters adds significantly to the value of the census in assessing the quality of a country's housing stock.

29. Construction material of floor and roof

4.546. In some cases the material used for the construction of roofs and floors may be of special interest and can be used to further assess the quality of dwellings. This topic refers to the material used for roof and floor (although, depending on the specific needs of a country, it may also refer to other parts of the building in which the housing unit is located, such as the frame or the foundation). Information on the

predominant material only should be collected. The following classification of construction materials is recommended:

1. Tile
2. Concrete
3. Metal sheeting
4. Wood
5. Bamboo
6. Palm, straw
7. Mud
8. Plastic sheeting
9. Other materials

Where the materials used on floor and on roof differ, classification specific to each should be used.

30. Elevator – availability of

4.547. This topic refers to the availability of an elevator (or lift) in a multi-storey building (categories 2.2.3–2.2.4 of the classification of type of buildings). It is recommended that the information should be collected on the availability of an elevator that is operational for most of the time, subject to regular maintenance.

4.548. This topic can be useful for providing further information for indicating the accessibility to the building or the housing unit. This is of particular relevance for older persons and persons with disabilities. In this context it could also be useful to collect information on the size of the lift (for the handicapped persons and ambulance transport), if the lift goes to the ground floor, and whether or not the lift stops on the same floor as the dwelling.

31. Farm building

4.549. Some national censuses may collect information to identify if a buildings or dwelling is located on a farm. A farm building may be considered as being one that is part of an agricultural holding whether it is residential or not, that is, whether it is used for agricultural or housing purposes. All the information that is relevant to other buildings and dwellings should also be collected.

32. State of repair

4.550. This topic refers to whether the housing unit or the building in which the housing unit is located is in need of repair and to the kind of repair needed. The following classification is recommended:

1. Repair not needed
2. In need of repair
3. Minor repair
4. Moderate repair
5. Serious repair
6. Irreparable

4.551. Minor repairs refer mostly to the regular maintenance of the building and its component housing units, such as repair of a cracked window. Moderate repairs refer to the correcting of moderate defects such as missing gutters on the roof, large areas of broken plaster or stairways with no secure handrails.

Serious repairs are needed in the case of serious structural defects of the building, such as missing shingles or tiles on the roof, cracks and holes in the exterior walls, and missing stairways. The term "irreparable" refers to buildings that are beyond repair, that is to say, with so many serious structural defects that it is deemed more appropriate to demolish the building than to undertake repairs; most usually this term is used for buildings with only the frame left standing or without complete external walls or roof.

33. Age and sex of the reference person of household (core topic)

4.552. From among the topics recommended for inclusion in the population census, age has been selected as being of most significance in relation to housing conditions. For the housing census, the data usually relate only to the housing units or building in which the housing units are located, but some characteristics of households that are related the housing condition can usefully be presented by the age and sex of household head or other reference person in the household.

4.553. While this information will usually be collected in a country's population censuses and, if the population and housing censuses are conducted simultaneously, as is the practice in the majority countries, then information on age of the head or other reference member of the household will be collected together with other relevant demographic characteristics in the population part of the census. If, however, the housing census is collected independently of the population census, then there should be a separate provision for collecting this information.

34. Tenure (core topic)

4.554. Tenure refers to the arrangements under which the household occupies all or part of a housing unit. The unit of enumeration is a household occupying a housing unit. The classification of households by tenure is as follows:

1 Household owns housing unit
2 Household rents all or a part of housing unit
  2.1 As a main tenant
  2.2 As a subtenant
3 Household occupies housing unit partly free of rent
4 Household occupies housing unit wholly free of rent
5 Household occupies housing unit under some other arrangement

4.555. National circumstances can dictate the need to assess the number of households occupying the housing unit free of rent to further distinguish whether such arrangement is with or without the consent of the owner. However, this information regarding the consent of the owner is subject to special scrutiny in terms of reliability. Furthermore, in countries where communal ownership is significantly represented, this topic on tenure might be further expanded in order to capture tenure arrangements of communally owned housing. Likewise, the category "other arrangements" can be extended to capture forms of tenure specific to some countries.

4.556. The information on tenure collected in the census needs to be clearly distinguished from the information on ownership (see paragraphs 4.476–4.481) and is one that should be asked of all households, otherwise there is a danger that it may be omitted in cases where more than one household occupies a single housing unit. Under some circumstances, it may be useful to indicate separately households that, although not subtenants in the sense that they rent from another occupant who is a main tenant or

owner-occupant, rent part of a housing unit from a landlord who lives elsewhere. These households and subtenant households may be of special significance in formulating housing programmes. On the contrary, in countries where subtenancy is not usual, information on subtenants may not be collected in the census or, if collected, may be tabulated only for selected areas.

4.557. In countries where the land and the living quarters are frequently occupied under separate tenure, the topic may be expanded to show separate information for the tenure under which the household or households occupy the living quarters and for the tenure of the land upon which those living quarters are located.

35. Rental and housing costs

4.558. Rent is the amount paid periodically (weekly, monthly, and so forth) for the space occupied by a household. Information may be obtained on the basis of a scale of rents rather than on that of the exact amount paid. The data may be considered in relation either to household characteristics or to the characteristics of the living quarters. In the latter case, where more than one household occupies a single set of living quarters, the rents paid by all the households will need to be summed in order to obtain the total rent for the living quarters. In the case of living quarters that are partly rented and partly owner-occupied, it may be necessary to impute the rent for the owner-occupied portion.

4.559. In countries where rent for the housing unit is paid separately from rent for the land upon which the housing unit stands, separate information may need to be collected reflecting the amount of ground rent paid.

4.560. In addition to the amount of rent paid by renting households, it may be useful to collect information on the housing costs, which could include information on monthly mortgage payments (for owner-occupiers), taxes and cost of utilities.

36. Furnished/unfurnished

4.561. Countries may wish to make some provision for indicating whether the housing units covered by the rent are furnished or unfurnished and whether utilities such as gas, electricity, heat and water are included.

37. Information and communication technology devices – availability of (core topic)

Note: The updates are based on text provided by ITU (August 2023).

4.562. The importance of the availability of information and communication technology (ICT) devices is increasing significantly in contemporary society. Modern technology offers a range of devices that is ever changing the structure and pattern of major social media and communications. The census provides an opportunity to assess the availability of these devices to the household and usage of these devices by individuals. The choice of devices should be sufficient for understanding the place of ICT in the household, as well as for use for planning purposes by government and the private sector to enable wider and improved delivery of services, and to assess their impact on the society. The recommended classification is:

1. Household having a computer
2. Household with Internet access

4.563. Availability of ICT devices in the household and the use and ownership of ICT devices by individuals is a very relevant topic for inclusion in a modern census. For instance, a category on the "Internet and mobile phones" would be concerned with determining the status of access to the Internet by households, the use of the Internet by individuals, the ownership of mobile phones by individuals, and individuals' ICT skill levels for a country, in relation to other socioeconomic or geographic classificatory variables. Indicators on individual Internet use, individual mobile phone ownership and individual ICT skills are each referenced through SDG indicators (17.8.1, 5.b.1, and 4.4.1, respectively).

4.564. In designing the questions, census planners should differentiate between two distinct viewpoints, namely (a) the availability of ICT devices to the households; and (b) use of ICT devices by the household members/individuals. The distinction is important, since households need not own, but may still have access to computers, mobile phones and the Internet through devices provided by school, university or work to household members. It also means that countries interested in collecting information on ICT use, particularly of the Internet and mobile phones, would need to include a relevant question topic in their census individual form. The rational for adopting either viewpoint, or even a combination of both, is not necessarily only technical, but rests more on the prevailing conditions in the society, and on how the information will be used to characterize the socioeconomic profile of households of a country. Usage statistics, including the intensity (frequency) of use and the range of activities performed, are preferably obtained using household surveys.

4.565. Computer refers to a desktop computer, a laptop (portable) computer or a tablet (or similar handheld computer). They exclude terminals connected to mainframe computers for data processing, and midrange multi-user systems that are primarily intended for shared use. Devices with some embedded computing abilities, such as smart TV sets, and devices with telephony as their primary function are also excluded. 'Household with a computer' means that the computer is generally available for use by all members of the household at any time, regardless of whether it is actually used. The computer may or may not be owned by the household but should be considered a household asset.

4.566. Internet access from home refers to the ability of the household to connect to the public Internet using TCP/IP protocols. Internet connections may be classified according to the technology employed, devices used, communication medium, or connection bandwidth (speed). Internet access at home is meant to include both narrowband and broadband connections. Access to the Internet is measured irrespective of the type of access, device used to access the Internet, or the method of payment. 'Households with Internet access' means that the Internet is generally available for use by all members of the household at any time, regardless of whether it is actually used. The connection and devices may or may not be owned by the household but should be considered household assets. If one member of the household has a mobile phone with connection to the Internet and makes it available for all members, then it should be considered that the household has access to the Internet.

4.567 Use of the Internet by individuals is a key indicator tracked by policy-makers and analysts as an indication of a country's progress towards becoming an information society. The reference period for individuals who used the Internet should be in the last three months and include use by individuals from any location. Classificatory variables for individuals using the Internet – such as age, sex, education level or labour force status – can help identify digital divides in individuals using the Internet. This information can contribute to the design of targeted policies to overcome those divides.

4.568 Mobile cellular telephones have become the predominant method of communication in many countries. Indicators related to mobile telephony are therefore fundamental indicators of the information society. Mobile cellular telephones refer to portable telephones using cellular technology that provides access to PSTN. An individual owns a mobile cellular telephone if he/she has a mobile cellular phone device with at least one active SIM card for personal use. It includes mobile cellular phones supplied by employers that can be used for personal reasons (to make personal calls, access the Internet, etc.) and those who have a mobile phone for personal use that is not registered under his/her name. It excludes individuals who have only active SIM card(s) and not a mobile phone device.

[Suggested New Topic] E-waste disposal

E-waste has become an important environmental and public health issue, not only at global scale but also in low- and middle-income countries as well. Many of Electrical and Electronic Equipments (EEE) contain hazardous chemicals and materials. For this reason, inadequacies in recycling and disposing of electronic waste (e-waste) can cause serious health and environmental pollution problems. Great care must be taken to prevent unsafe exposure in recycling operations and prevent contamination of the environment. In recognition of the potential environmental and health impact of e-waste, many countries have embarked on programs for e-waste management. The e-waste management programs are implemented as either specific e-waste policies or in a form of legislations. Availability of data on e-waste is therefore of paramount importance in these endeavours.

Classification:

- Temperature exchange equipment: More commonly referred to as cooling and freezing equipment. Typical equipment includes refrigerators, freezers, air conditioners, heat pumps.
- Screens, monitors: Typical equipment includes televisions, monitors, laptops, notebooks, and tablets.
- Lamps: Typical equipment includes fluorescent lamps, high intensity discharge lamps, and LED lamps.
- Large equipment: Typical equipment includes washing machines, clothes dryers, dish-washing machines, electric stoves, large printing machines, copying equipment, and photovoltaic panels.
- Small equipment: Typical equipment includes vacuum cleaners, microwaves, ventilation equipment, toasters, electric kettles, electric shavers, scales, calculators, radio sets, video cameras, electrical and electronic toys, small electrical and electronic tools, small medical devices, small monitoring and control instruments.
- Small IT: Typical equipment includes mobile phones, Global Positioning Systems (GPS), pocket calculators, routers, personal computers, printers, telephones.
- Electrical and electronic tools: Drills, electric saws, sewing machines, lawn mowers, large stationary tools, machines.
- Security and healthcare equipment: Surveillance & control equipment, medical instruments & equipment
- Mixed waste electrical & electronic equipment: Various e-waste

38. Cars – number of available

4.567. This topic refers to the number of cars or vans normally available for use by members of the household. The term “normally available” refers to cars and vans that are either owned by occupants or are under some other more or less permanent agreement, such as a lease, and includes those provided by an employer if available for use by the household, but excludes vans used solely for carrying goods or other commercial purposes.

39. Durable household appliances – availability of

4.568. The unit of enumeration is a household occupying a housing unit and information may be collected on the availability, within the housing unit, of durable appliances such as washing machines, dishwashing machines, refrigerators, deep freezes and microwave cookers, depending on national circumstances.

40. Outdoor space – access to

4.569. This topic refers to the reasonable access to an outdoor space intended for the recreational activities of the members of a household occupying a housing unit. The classification can refer to any outdoor space that is available:

1 As part of a housing unit (for example, a garden or backyard)
2 Adjacent to the building (for example, playgrounds placed next to the apartment building)
3 As part of common recreational areas within a walkable distance from the housing unit (for example, parks, lakes, sports centres and similar sites)
4 Beyond a 10-minute walk.