United Nations | Department of Economic and Social Affairs

Overview of the people building block: Context setting for population estimates

(P. Gerland, Population Division, DESA, United Nations)

Building modern and resilient population data systems to enhance data quality, improve cost efficiency and policy relevance

Second sprint organized by the Friends of the Chair Group on Social and Demographic Statistics, in collaboration with the United Nations Statistics Division (UNSD): Thursday, 3 October 2024

Outline

- Introduction of the typology of population data systems
- Overview of country data sources and availability for population estimates
- Overview of metadata availability for population estimates
- Global perspective on challenges and opportunities for internationally comparable estimates

Why population data systems matter?

- Crucial for informed decision-making and policy formulation.
- Population data informs resource allocation, service delivery, and policy development.
- Accurate and timely data is essential for monitoring progress towards the Sustainable Development Goals (SDGs).
- Modern data systems need to be resilient to adapt to evolving data demands and potential disruptions.

Fundamental law of demographic dynamics

- **Population change depends on components of demographic change and a balancing equation for demographic accounting**
- This is represented by the following relationship:

- Population estimates (and projections) are derived by combining population, natural change (births minus deaths) and net migration (immigrants minus emigrants) either at the aggregate or individual level.

Introduction about population estimates

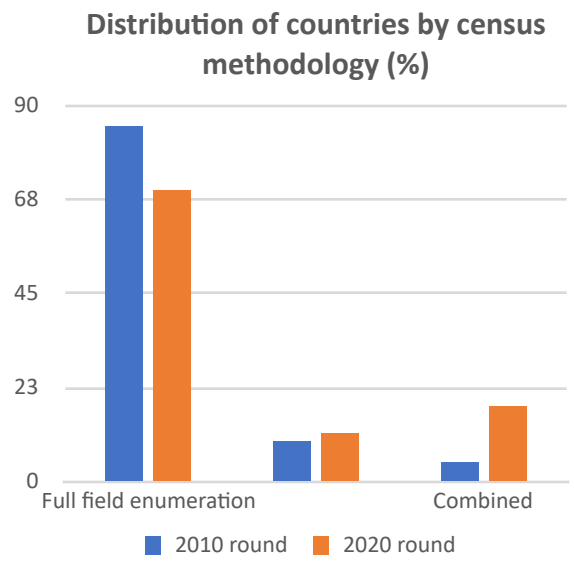
How do we know the current number, characteristics and distribution of people in each country or area?

… it depends on the **type of population data system** available in a given country …

Typology of Population Data Systems

Three main groups based on their primary data source:

1. Population counts (stocks)*
    1. **Full field-based Census systems**: relying on periodic complete field enumeration of the population.
    2. **Register-based systems**: relying on continuously updated population registers and records of vital events (births, deaths, migration) and (a) other administrative data linked at the individual level, or (b) existing sample surveys or integrated administrative sources and existing sample surveys
    3. **Combined systems**: combining data collected from full field enumeration or one or more surveys with administrative data sources or statistical registers.



2. Vital statistics (flows) for intercensal and post-censal periods
    1. Vital registration of births and deaths (full or partial completeness)
    2. Fertility and mortality statistics from censuses and surveys
    3. International and/or internal migration statistics from administrative data

(*) Source: UNSD surveys on 2010 round PHCs (126 responses) and 2020 round PHCs (158 responses)

1. Full field-based Census systems

Description:

- A census aims to count every person in a country at a specific point in time
- Typically, full field enumeration every 5-10 years

Key Features:

- Accurate if conducted well, but costly and infrequent.
- Post-census adjustments often required

Geographical Distribution:

- Predominantly used in countries without developed administrative data systems.

Census role in population estimates:

- Acts as a critical baseline for population estimates and projections.
- Provides benchmark data for other systems (vital registration, administrative data).

Population estimates at aggregate level for different level of geography/groups:

- Inter/post-censal estimates using demographic balancing equation and cohort-component population projection methods
- Relies on vital statistics, and migration data between and after censuses

2. Register-based Systems

Description:

- Continuously updated population register and registration of births, deaths, and migration.

Key Features:

- Linked administrative data sources
- Provides timely and regular data updates on demographic changes
- High accuracy when records are maintained properly

Geographical Distribution:

- Primarily used in Northern Europe and some Asian countries, requires advanced statistical infrastructure

Population register role in population estimates:

- Provide a continuous source of data that supports inter/post-censal estimates

Population estimates at individual level:

- Estimates based on individual level record linkages between multiple data sources, including increasingly additional validations with sign-of-life
- Relies on vital records, migration and residency status from civil registration and admin. sources

3. Combined systems

Description:

- Integrated mix of census, surveys, administrative data such as address registers, social security data, birth and death records, health, education, employment data, tax records, etc.

Key Features:

- Increasing use of administrative data to supplement census
- Cost Efficiency: reduces reliance on full-field enumeration
- Improved Timeliness: provides more frequent updates than traditional censuses
- Data Quality: Helps improve the granularity of data, particularly when linked with other data sources

Geographical Distribution:

- Becoming more common in countries with advanced administrative capabilities (e.g., Europe, North America) and a growing number of countries with the 2020 round of censuses in parts of Asia and Latin America

Population estimates at individual level:

- Estimates based on aggregate and/or individual level record linkages between multiple data sources, including increasingly additional validations with sign-of-life
- Relies on vital statistics, migration and residency status from civil registration, admin. sources and surveys

Overview of Country Data Sources for Population Estimates

Data Sources:

- Population censuses: most countries conduct censuses, but coverage, frequency and quality can differ.
- Civil registration systems (CRS) and vital statistics: birth and death registration systems vary in completeness (from <50% to >90% globally) and accuracy.
- Migration data: often the weakest component
- Population registers: some countries have well-established registers, while others are developing them – but hard to know precisely due to lack of metadata.
- Administrative data (often only for selected subsets of population): Increasingly used, but access and quality can be challenging, and many integration challenges when unique national IDs are unavailable.
- Sample surveys (when other sources are unavailable): provide valuable data, but may have limitations in coverage and frequency.

Global Perspective:

- Differences in data availability and quality across regions.
- Censuses are conducted less frequently in lower-income countries due to costs

Intercensal and Post-Censal Population Estimates

Intercensal Estimates:

- Use data from the two most recent censuses combined with birth, death, and migration.
- Helps to maintain up-to-date population estimates between censuses, and to ensure the internal consistency between demographic components of change (especially migrations) and populations
- Best practice: after each new census, a new set of intercensal estimates should always be recomputed, disseminated and used to recomputed intercensal vital rates and life tables (European Commission 2003, Statistics Canada 2016, Spoorenberg 2020, United Nations 2024)

Post-Censal Estimates:

- Adjustments are made after a census to correct inaccuracies and update for population dynamics
- Post-censal estimates use the demographic balancing equation to project population figures
- Timeliness and frequency of updates: produced shortly after last census or updated annually or more frequently (monthly or near-real time daily updates) to incorporate extra / new data on vital statistics and migrations
- Timeliness and type of input data used for each demographic components
- Projection method: cohort-component (aggregate level) or microsimulation (individual level: Austria, Canada, Mexico based on UNECE Database on Population Projections Metadata)

World Population Prospects (WPP) 2024 – biennial update

- Comprehensive, standardized demographic dataset for all countries/areas Internally consistent prospective annual **population reconstruction** from 1950 to 2023 using all available data (i.e. population by age and sex, mortality, fertility and migration) using the cohort component method
- Population estimates are used to check consistency of WPP estimates (…and WPP allow to detect consistency problems in population estimates)

>> Important to obtain **accurate/consistent** national estimates of population size and composition by age and sex

  - WPP used as “denominator” in several SDG indicators

For further details, see United Nations (2024)

Consistent population estimates

Several methods exist to produce annual population estimates

- Basic requirements:
  - An initial/base population count (recent, evaluated and if necessary adjusted, see P&R for Population and Housing Censuses, Revision 4);
  - A method of time adjustment to update the initial population count

Gold, silver and bronze standard methods

1. Continuously updated population registers
2. Regularly updated combined population systems
3. Cohort-component method applied to census data

→ Population data consistent by age, sex and along cohorts

Official population estimates: Current practice

Base population

- In majority of the regions, full (or partial) population census, or population register serve as base population for population estimates

Method of time adjustment

- Situation varies by region
- Practice of not bringing up to date the initial population still prevalent in few countries
- Not all methods produce consistent population estimates by age, sex and cohort

Adjustment to base population

- In many countries, no adjustment (or only age smoothing) is made to census data
- Under-enumeration, especially of young children, is not systematically corrected

>> Improper method of time adjustment and/or unadjusted base population produce national population estimates that are inconsistent and different from international (WPP) estimates

For further details, see Spoorenberg (2020)

Data Availability for Population Estimates (WPP 2024)

- National statistical sources (either taken as-is or adjusted after in-depth evaluation) for 1950-2023:
  - **2,025 censuses** and 320 post-enumerations surveys
  - **2,970 surveys** (440 since 2015)
  - **vital registration systems** from **169 countries or areas**

Most recent census data:

- 48% of countries or areas (n=114) have data since 2019 (and 54 per cent of the world population)
- 18% (n=43) from 2014-2018, 24% (n=57) from 2009-2013, 10% (n=23) before 2009

Latest available birth statistics (and their completeness) vary significantly by region:

- 83% of countries or areas (n=196) have fertility data since 2019, 14% (n=32) from 2014-18, 3% (n=8) before 2012
- Only 73% of the countries, territories and areas register at least 90% of births occurred (UNSD).

For further details, see United Nations (2024)

Data Availability for Population Estimates (WPP 2024)

Latest available mortality statistics vary significantly by region

- 120 countries or areas with good VR statistics for empirical life tables by age and sex, 117 countries or areas with empirical mortality rates by sex and age too sparse or of insufficient quality to estimate the complete annual time series of mortality rates.
- Only 68% of the countries, territories and areas have at least 90% of deaths occurred (UNSD).

Number of countries with data on age- and sex-specific deaths from complete vital registration for the years 2020 to 2023 (as of March 2024)

	Number of countries with 90 per cent or higher death registration	Proportion of total global deaths covered by complete VR (per cent)
2020	106	35
2021	100	30
2022	51	15
2023	5	1

- Under-five mortality: 78% of countries or areas (n=184) have data since 2019, 15% (n=35) from 2014-18, 3% (n=7) from 2009-13, and 2% (n=6) before 2009.
- Adult mortality: 65% of countries or areas (n=154) have data since 2019, 25% (n=60) from 2014-18, 6% (n=15) from 2009-13, 3% (n=6) before 2009, and 1% (n=2) no data.

For further details, see United Nations (2024)

Data Availability for Population Estimates (WPP 2024)

Great variability in data sources and reliability over time and locations for (net) international migration estimates

- Official annual estimates (e.g., NSOs, Eurostat, n=54)
- Estimates of migrant flows (e.g., Eurostat, OECD, n=56)

n=67 distinct locations

- Foreign-born stocks (UN estimates) and implied annual change
- Administrative data (e.g., work permits issued/renewed)
- Literature review for major forced migration historical events
- Intercensal net residual migration for countries with good VR or residual from cohort-component projection compared to census or population register
- UNHCR estimates of refugee stocks (and implied annual change)

For further details, see United Nations (2024)

Metadata Availability for Population Estimates

Essential information on data sources, estimation methods, adjustments, and assumptions.

Importance of metadata: crucial for transparency, data interpretation, and to enable users to assess data quality and comparability, understand limitations, and make informed decisions.

Global Availability:

- UN DYB questionnaire on data and methods used for official estimates
- Metadata availability and completeness vary significantly by country.
- Some regions provide complete metadata, but others lack details on adjustments, and methods used.

For further details, see Spoorenberg (2020)

Challenges for Internationally Comparable Estimates

Common issues:

- Lack of detailed documentation on population data and methods
- Differences in definitions (e.g., de-facto, de-jure, usual resident), and methods across countries
- Varying data availability and quality gaps
- Differences in coverage of population registers and civil registration systems
- Difficulties in including migration data
- Metadata gaps make it difficult to assess the accuracy of population estimates for international comparison

Opportunities:

- Growing use of administrative data and registers
- Greater use of geospatial data and digital-first census methods to fill data gaps
- Technological advancements, data integration, and improved statistical methods
- Potential for more timely, granular and comparable estimates
- Need for more harmonized approaches while respecting national contexts

Opportunities for Improving Population Estimates

- Countries must modernize data systems to produce reliable, timely, and internationally comparable population estimates.
- Emphasis on the integration of different data sources (censuses, registers, administrative data) for better estimates from individual level to various aggregated level (geographical and subgroups).
- Increased use of technology, administrative data, and geospatial techniques can provide more frequent and accurate population estimates.
- Produce and disseminate more comprehensive and accessible metadata for greater transparency.
- Importance of international collaboration: sharing best practices, capacity building, and harmonization of standards.

References

- European Commission (2003). Basic Methodology for the Recalculation of Intercensal Population Estimates, Luxembourg: Office for Official Publications of the European Communities, Working Papers and Studies (3/2003/E/no 27).
- Johnson, P., and others (2022). *Method protocol for the evaluation of census population data by age and sex*. UN Population Division, Technical Paper, No. UN DESA/POP/2022/TP/No.5, New York, United Nations.
- Pelletier, F. (2020). *Census Counts, Undercounts and Population Estimates: The Importance of Data Quality Evaluation*, New York, United Nations, DESA, Population Division, Technical Paper No. 2020/02.
- Spoorenberg, T. (2020). *Data and Methods for the Production of National Population Estimates: An Overview and Analysis of Available Metadata*, New York, United Nations, DESA, Population Division, Technical Paper No. 2020/01.
- Statistics Canada. Demography Division (2016). *Population and Family Estimation Methods at Statistics Canada*. No. 91-528-X, 1100231609, Ottawa, Statistics Canada.
- United Nations (2024). *World Population Prospects 2024: Methodology of the United Nations population estimates and projections*. Working Paper, No. UN DESA/POP/2024/DC/No. 10, New York, Population Division, Dept. of Economic and Social Affairs