Quality of Administrative Records in the
E-Census
for Population, Residences & Establishments
2020

His Majesty Sultan Haitham Bin Tarik

endorses E-Census 2020 results of Population, Residences and Establishments as of 12 December 2020. His Majesty the Sultan expressed his satisfaction with the outcome of this major project which was implemented as scheduled. His Majesty the Sultan underscored the significance of data and indicators provided by the Census in enhancing Oman Vision 2040, as well as the Census's important implications to developmental planning in all sectors of the Sultanate.

14 December 2020

Table of Contents

Introduction

The administrative records available in organizations are one of the most important sources of data in many countries of the world. Rather, they are the main source in periods when there are no censuses or field surveys. Therefore, they are highly relied upon in planning and development processes. Decision makers build multiple policies and procedures in the country depending on those records.

Many countries in the world are seeking to improve the volume and quality of administrative records to be more comprehensive, and continuously updated. They seek to optimize its accuracy, and it is the target of the Sultanate for long periods through many projects that work to develop administrative records including the e-government project and government linkage between the relevant authorities.

When the Sultanate relied on the electronic census based on administrative records, the census project worked on studying those records available in the data-issuing institutions, improving the quality and coverage of those data, completing many shortcomings, correcting a wide range of errors, organizing many classifications and developing many systems in order to use those Records in the electronic census project 2020.

In this version, the quality of administrative records in the electronic census of population, residences and establishments will be addressed. Explanation for some of the statistical terms and concepts will be used. In addition to some statistical quality standards and studies conducted in order to confirm the quality of data, and the procedures undertaken by the project management in order to analyze the shortage sources. Reference to some recommendations related to improving data quality will be mentioned as well.

1 | Records and Electronic Census:

1.1 National Administrative Records:

It is a set of data provided by government and private institutions that represent the source agencies of electronic census data and are linked to each other in an electronic format that can be continuously updated in order to produce indicators and statistical data in real time.

Among the most important advantages of relying on administrative records in the state are the following:

- All government and private institutions can benefit from the data.
- The data is subject to continuous updating.
- Raising the quality of the data in the target records.
- Saving a lot of effort and cost in future censuses.

1.2 Census Based on Administrative Records:

A statistical process aims to publish official statistics, relying entirely on the administrative resources available to government agencies in the Sultanate, or government companies and relevant private sector entities. This is through data collection and preparation so that they are compatible in their use of statistical classifications and best practices for statistical work. The data of administrative records in the census needs in-depth analysis to ensure the accuracy of the sources and related metadata, and variables.

Administrative records data is one of the important tools for evaluating the accuracy and quality of survey and census data by comparing the results of surveys and censuses with them and identifying their consistency and rationality.
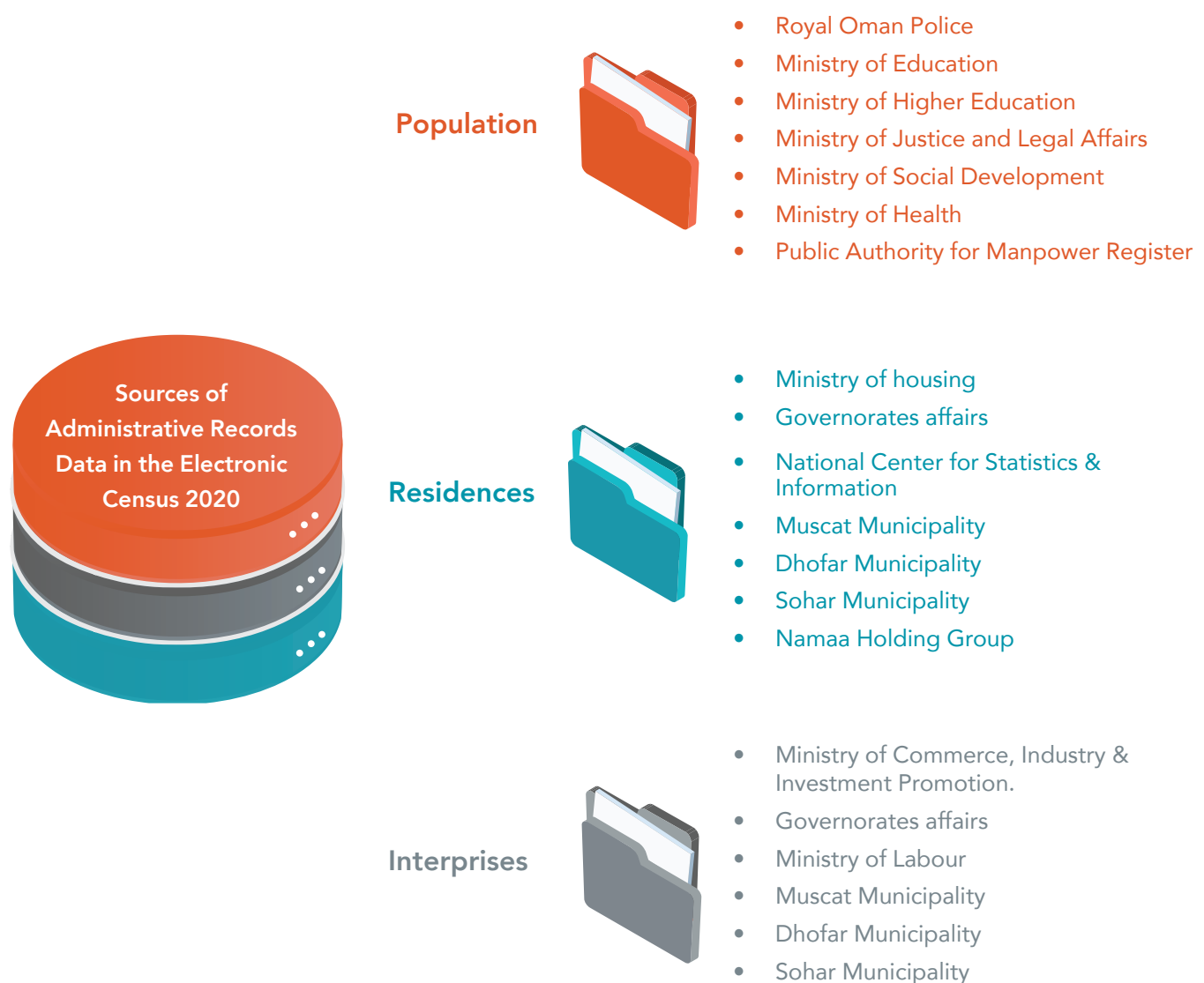
In order for the administrative records to be used in the electronic census, it must pass through a large set of stages, which ensures accurate, comprehensive and timely data:

- Refining and processing data that is missing some variables
- Linking and matching between variables by ID number
- Determining the time reference period
- Application of quality assurance standards

- Contacting record providers to ensure the accuracy of the data
- Knowing the reasons for non-responding
- Knowing the reasons for non-conformity
- Evaluate the quality of the variables
- Report inconsistencies in records
- Report missing metadata
- Update incoming records periodically

1.3 Sources of Administrative Records Data in the Electronic Census 2020

The Census Higher Committee has identified the census data basket, through which the data-issuing institutions that cover that basket were identified. The following figure shows the data sources of administrative records in the Sultanate used in the electronic census 2020:

Population

- Royal Oman Police
- Ministry of Education
- Ministry of Higher Education
- Ministry of Justice and Legal Affairs
- Ministry of Social Development
- Ministry of Health
- Public Authority for Manpower Register

Sources of Administrative Records Data in the Electronic Census 2020

Residences

- Ministry of housing
- Governorates affairs
- National Center for Statistics & Information
- Muscat Municipality
- Dhofar Municipality
- Sohar Municipality
- Namaa Holding Group

Interprises

- Ministry of Commerce, Industry & Investment Promotion.
- Governorates affairs
- Ministry of Labour
- Muscat Municipality
- Dhofar Municipality
- Sohar Municipality

2 | Quality of Administrative Records Data:

2.1 Data Quality Framework:

Data is the lifeblood of every entity. It forms the basis for many critical business decisions. Therefore, the data flow process must be of high quality to ensure the validity of the data. Accurate data can help deliver valuable results. To benefit from the explosive growth of big data, a framework for data quality management should be used first before you start extracting actionable insights from information.

Data Quality Management (DQM) refers to the set of business practices that include employing the right people, processes, and technologies to extract actionable insights from available information, ensuring that the data process flows throughout the data lifecycle with high quality.

2.2 Benefits of Data Quality Framework:

High-quality data improves statistical processes and makes them more efficient and accurate. This helps to improve the results. On the other hand, we find that poor data quality is one of the reasons for the failure of data migration projects. To ensure a high success rate, data quality rules should be used to identify and correct any errors before they are migrated. This helps to run projects faster, more accurately, and reducing data processing time and costs.

The Department of the E-Census 2020 has developed a data quality framework, represented in:

- Create a data quality schedule and methodology
- Checking the quality of the data in the source (records)
- Data quality at the exchange service layer (data purge, ELT, integration...)
- Checking the quality of the data before entering into the statistical records
- Checking data quality before publishing

The importance of assuring the quality of statistical data comes through the development of all necessary measures and procedures that lead to the achievement of the following:

- Implementation of statistical procedures and processes that ensure improvement and development in the quality of work.

- Implementation of all processes that ensure quality assurance on a regular basis.
- All quality assurance processes should be specific and known to all technicians.
- All processes related to quality assurance should be documented and subject to continuous monitoring.
- All the processes and procedures for achieving quality are known to data users, which leads to an increase in the credibility of indicators or statistical data.

2.3 Data Quality Assessment:

The topic of «Quality Assurance» is of great importance to statisticians in general and census operators in particular because of its great impact on the quality and accuracy of data. It is the consequences of strenuous efforts and huge expenditures in order to achieve a successful census.

Ensuring the quality of census data includes specific procedures, which have become of great importance to the competent authorities of the United Nations, which pay special attention to the management of general censuses of population, residences and establishments. It is in order to obtain accurate data, which is a valuable resource for any society, and its way to achieve the desired goals in the comprehensive development of the Country and sustainable well-fare of its population.

The process of "quality assurance" focuses on achieving the best possible results, rather than relying on subsequent data correction processes. Census data must be satisfied with the users of that data. In order to achieve this, it must be of high quality. The quality is mainly related to accuracy, but we realize there are other dimensions of quality, which is the release of results at appropriate times.

Quality standards should be adhered to when conducting censuses, publishing their data, and identifying strengths and opportunities for improvement on a regular basis to develop the quality of operations and statistical outputs, including:

- The existence of clear texts of the commitment of the entity producing the statistical data and the provision of high-quality statistical data.
- Existence of evidence specialized in statistical quality standards and procedures that determine the scope of quality and how to apply it.
- Taking all necessary measures to plan and control the quality of statistical operations through designated persons or a specialized organizational unit to do so.
- Monitoring the quality of statistical products in an organized manner, evaluating and reporting them in accordance with statistical quality standards.

Among the most important conditions that must be met to ensure data quality are the following:

- Available data must meet the needs of users.
- Ease of obtaining data in a timely manner.
- The data should be clear, high-accuracy and credible.
- The available data should enable international comparisons to be made.
- The possibility of the availability and use of metadata, with clarification of the used terms and definitions.
- Taking into account the provision of data at the lowest possible cost.

It is possible to apply internationally agreed quality standards when studying administrative records available at government and private institutions, analyzing them and revising their data in order to improve their quality. The uses of administrative records in the preparation of official statistics are many, and this diversity is due to several factors, including the extent of the accuracy of these data and their compatibility with the statistical standards and the extent of the census of the counted statistical data. Therefore, it is very important to establish a set of quality controls and procedures for administrative records and evaluate them regularly in order to ensure the quality of official statistical indicators.

Recent years have witnessed an increasing demand for the use of administrative records in various surveys, as these records have become available in most statistical institutions in different countries, and their use has provided the following advantages:

- Reducing resource and operational costs
- Reducing the burden of response on institutions, families and individuals.
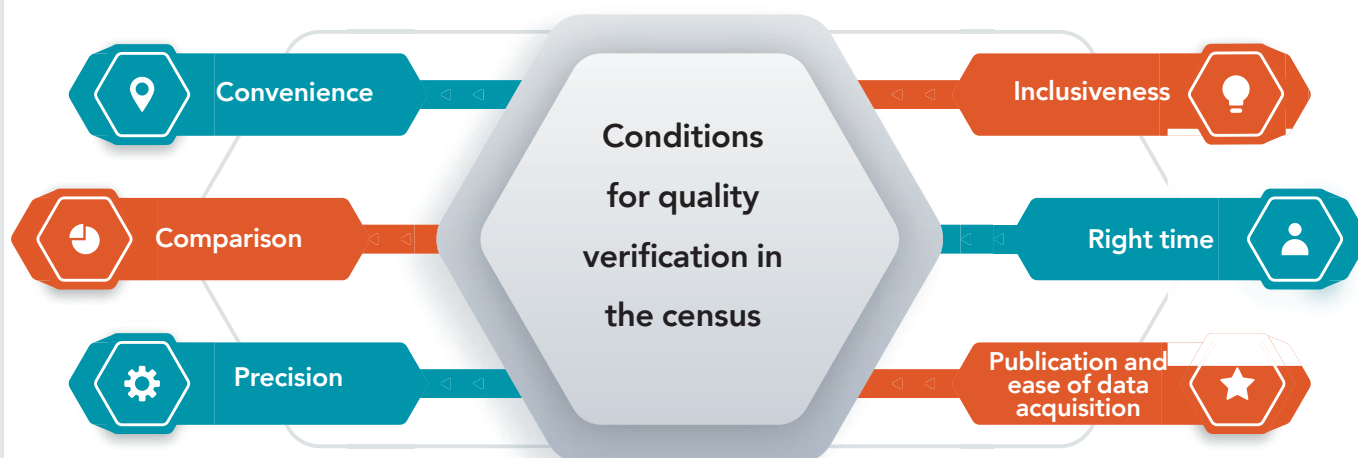- Minimize efforts to review, verify and modify data

The administration of the electronic census in the Sultanate has sought quality in all stages of the census in order to achieve the dimensions of total quality, which are represented in:

| Sustainability | Serviceability | Reliability | Good performance |

Preparation stage

data readiness stage

electronic system building stage

delivery stage

test operation stage of the system
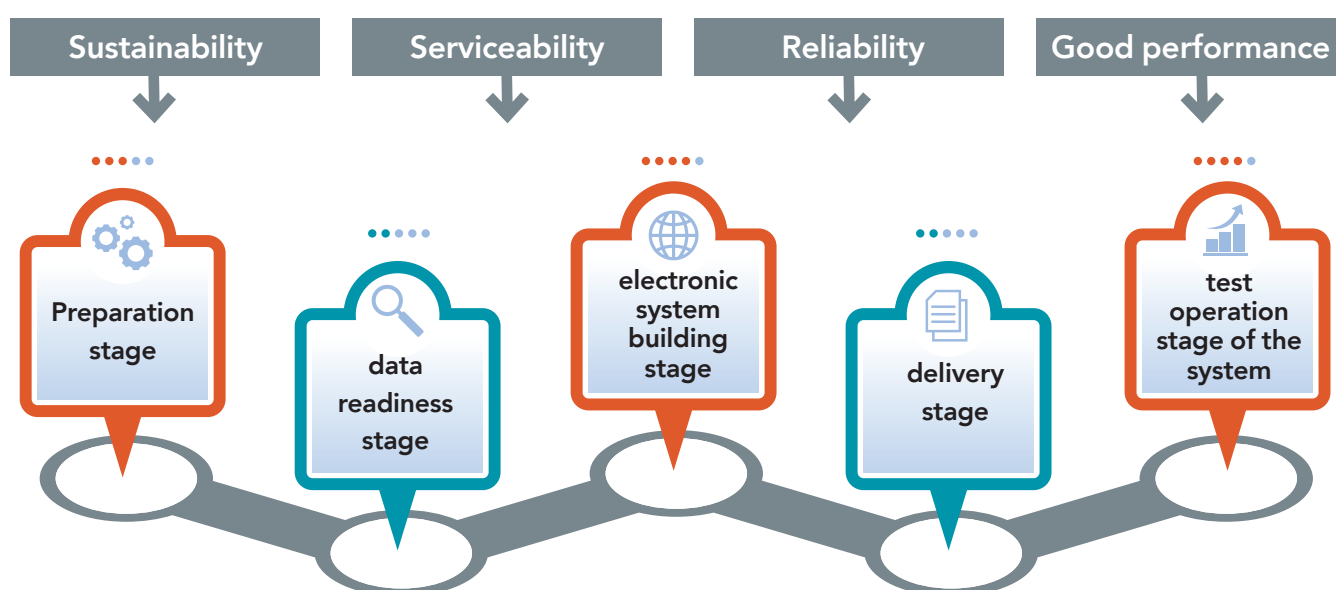
for example, but not limited to, these dimensions were confirmed through the following procedures:

- **Standard of good performance**: Building a comprehensive and integrated information system that links population data with residences and facility data, which is automatically processed through linking with the records of administrative authorities.
- **Reliability Standard:** Processing approximately 560 million records to study the coverage of population, residences and establishments data at the level of census laboratories before uploading them to the system
- **Sustainability Standard:** Complete the coverage of the administrative records of graduates in cooperation with the work team of Ministry of Higher Education, after the coverage was very weak in the range of 140 thousand graduates, while the expected rate was in the range of 320 thousand.
- **Service Standard:** Through coordination between the census laboratories and the work teams of the authorities, more than 9 million records were matched and corrected.

Finally, the quality of the data in the census can be achieved if the following conditions are met:

Convenience

Comparison

Precision

Conditions for quality verification in the census

Inclusiveness

Right time

Publication and ease of data acquisition

Here, a distinction must be made between two types of

data quality

The first type:

Total quality, which is procedures for reviewing data before the process of publishing it officially in order to assess its quality and verify that the published data is free of errors.

The second type:

It is the study of the sources of errors and then quantifying the errors according to the different sources. Often the results of such a study are available after the official publication of the data.

2.4 International Experiences and Consulting Studies in the Strategy of Administrative Records Data Quality:

The project management reviewed some of the international experiences in the field of smart data employment. The project also used some consulting offices specialized in the field of information technology and data governance law in order to provide advice regarding applicable laws and frameworks in the Sultanate of Oman. This will affect the data strategy, through conducting a survey and interviews to determine the feasibility of the data strategy for government agencies, private institutions and the public in the Sultanate. Studying any existing obstacles that might hinder the data strategy and ways to improve it in order to reach the desired quality of that data.

The consulting office, in cooperation with the project, has developed an electronic census model based on the administrative records of the institutions targeted for coverage. The model includes various government agencies to cover the census of population, residences and establishments so the information that has been counted in the statistical records is collected and integrated into an integrated statistical record to create statistical tables, provided that the project is being implemented according to the scientific methodology to meet the requirements of the electronic census project 2020 and in several stages as follows:

Phase one (preparation):

- Confirm general strategy
- Broad identification of risks, including risk mitigation
- Determine the prerequisites for success
- Defining the methodology of the electronic census project 2020

Phase Two (Project Governance):

- - Confirming master tables/basket and associated layout
- - Designing a comprehensive business structure to provide the electronic census
- - Determine the policies required to implement the project
- - Confirm project schedule (data analysis and methodology design)

Phase Three (Launching the Project):

- Preferential analysis and design for the selected centers
- Determine the selected business requirements
- Engineering and design of electronic census solutions
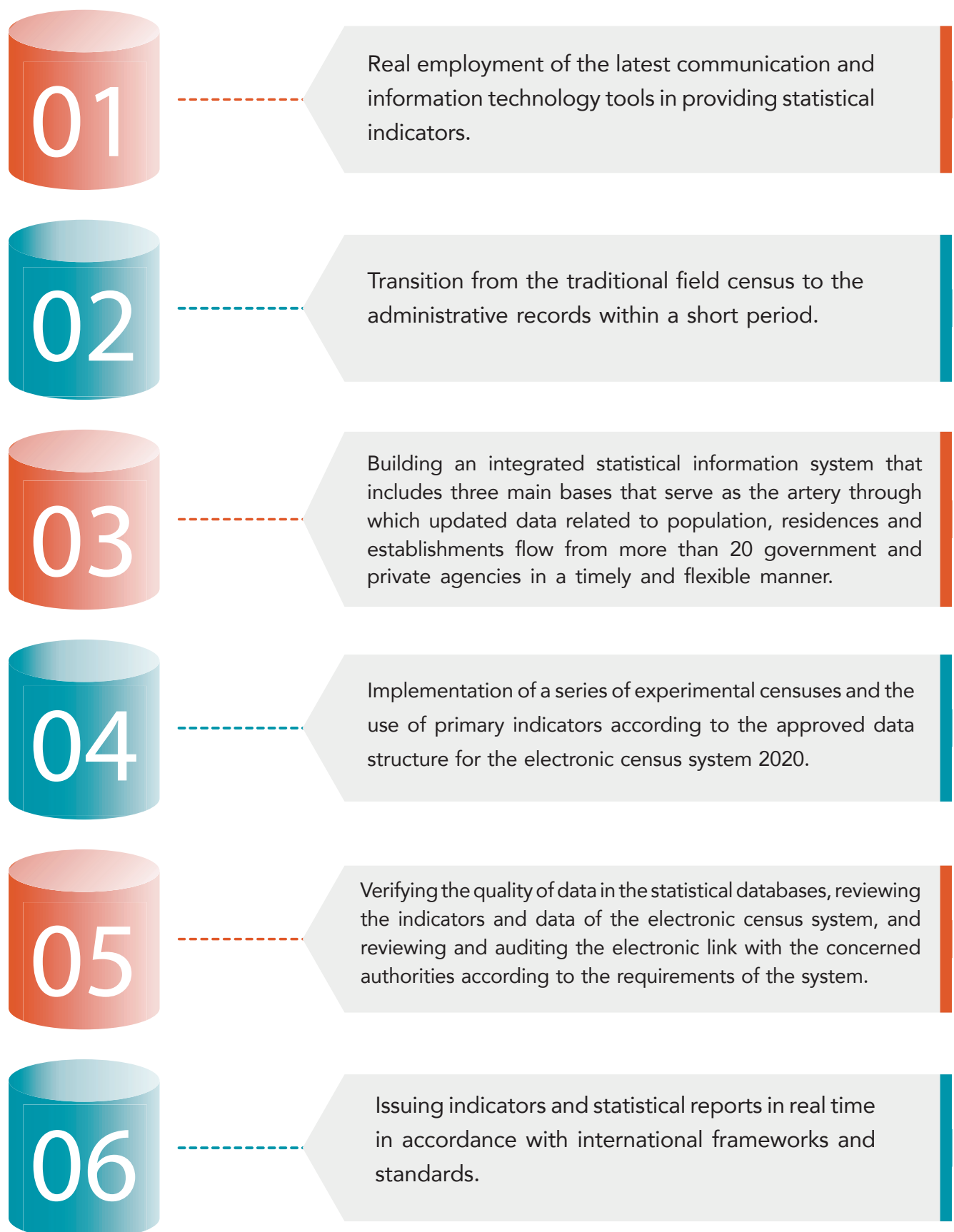- Estimate the implementation of the topic

The project was implemented within a period of 8 weeks, with a joint working team of the census project amounting to 13 members, and about 27 members of the consulting company of different nationalities. The joint team interviewed 15 institutions, 22 interviews with external institutions, and primary stakeholders from the National Center for Statistics and Information. The team also analyzed 76 variables, not including the associated variables and classifications. The joint team prepared a detailed report that addressed:

1. **Project outline:** It includes the project success factors, the work statement of government agencies, and the data quality framework in accordance with a specific quality management strategy in the exchange service and the work plan.

2. **Data Analysis:** It deals with definitions, a list of census variables and a data map, in addition to the structure of statistical records, analysis of census variables, and the methodology of sharing data with statistical records.

3. **Project Governance and Solutions:** The report included emphasizing the technical feasibility of the census, identifying risks and creating a risk reduction strategy, in addition to structuring the work necessary to ensure the success of the census.

The consultative study of the joint working group ended with the following outputs:

- A proposal to restructure the business of the electronic census system.
- Electronic census project plan.
- The most important basic risks of the electronic census based on the analysis that was conducted.
- Summary of the gap analysis for each government agency and the proposed tasks and solutions to fill those gaps.
- The feasibility of the 2020 census project in the field of population, residences and establishments   census.

2.5 The Methodology Used to Achieve Quality in the Census Project:

01 — Real employment of the latest communication and information technology tools in providing statistical indicators.

02 — Transition from the traditional field census to the administrative records within a short period.

03 — Building an integrated statistical information system that includes three main bases that serve as the artery through which updated data related to population, residences and establishments flow from more than 20 government and private agencies in a timely and flexible manner.

04 — Implementation of a series of experimental censuses and the use of primary indicators according to the approved data structure for the electronic census system 2020.

05 — Verifying the quality of data in the statistical databases, reviewing the indicators and data of the electronic census system, and reviewing and auditing the electronic link with the concerned authorities according to the requirements of the system.

06 — Issuing indicators and statistical reports in real time in accordance with international frameworks and standards.

3 | Census Labs and Data Quality Improvement:

In order to analyze the data statistically, correct errors, work on pass tables, and communicate with the competent authorities to correct the data, six laboratories were established within the census system, three of which are statistical, and a laboratory for geographical data with the aim of preparing geographical data and maps, and a database laboratory entrusted with managing databases, examining the system and entering the data received from the authorities manually and all reports related to the databases are made on demand. There is also a laboratory concerned with developing programming processes for the census system.

Census Labs

01 ----- Database Lab
1

02 ----- Geographical Data Lab
1

03 ----- Development Lab
1

04 ----- Statistical Lab
3
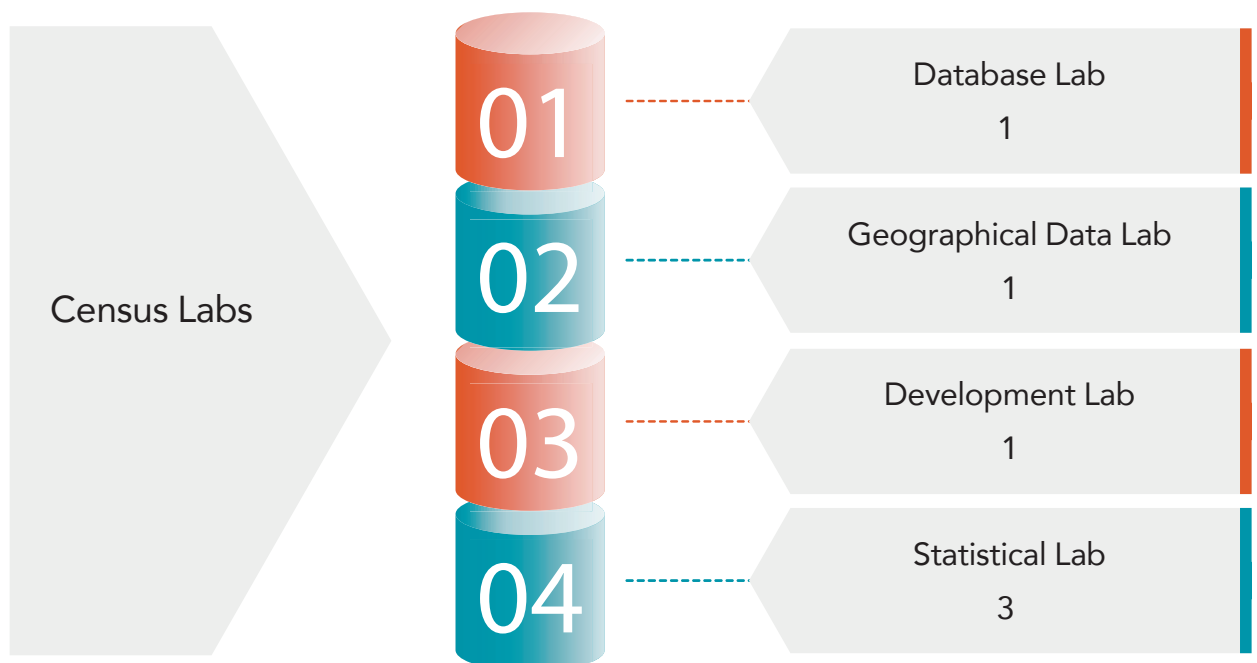
In general, these laboratories study the data and its shortcomings or errors in terms of mandatory fields, coordination of fields and traffic schedules, and some are corrected internally or a special report of errors is sent to the concerned authorities to correct them.

The number of records processed for each quadrant is more than 74 million records located in 38 files received from the authorities, including 30 million records for the identification data of the population, 4 million records for labor force data, 3 million records for education, 6 million for marital status, and one million records for births and deaths. 12 million records related to facilities and 19 million records related to unit data.

3.1 Study Covering Data Deficiency:

It is intended to complete the coverage of all mandatory fields contained in the set of databases (data basket) that are provided to the authorities after being reviewed by the analysis team in the census, according to the following form:
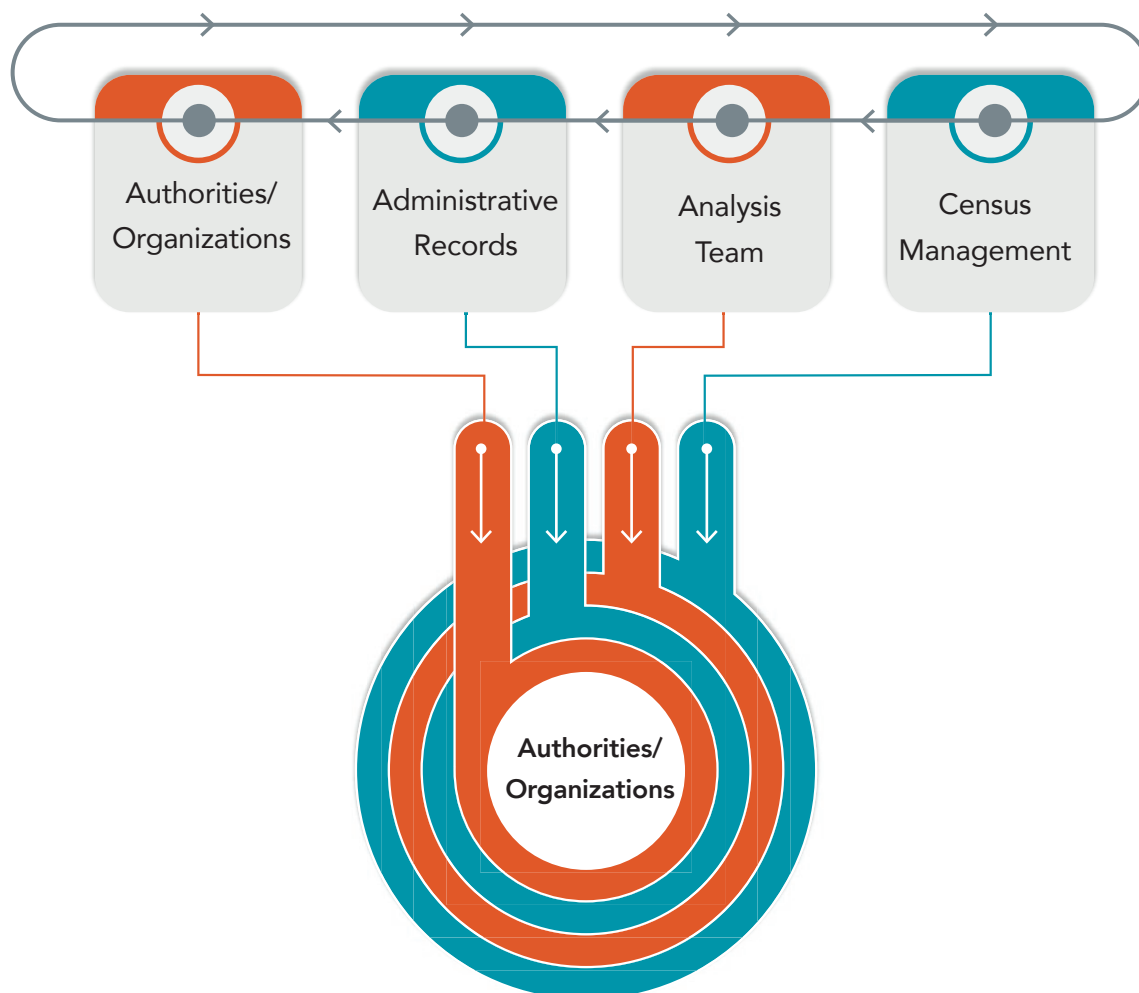
- The authorities prepare their own administrative records according to the required fields and fill in all the mandatory data.
- The authorities send their records to the census management project for study and review.
- The census analysis team works by studying the database, making sure that all data is covered, and preparing a report to the census department regarding the extent to which the fields are full and their data is complete.
- If the field data and the database are complete, this makes the database ready for action, but if the database is incomplete, the census project management returns it to the entity and completes the coverage until it reaches 100%. This was confirmed by the speedy response of the authorities to the recommendations of the technical teams to enumerate.

3.2 The most important achievements in the quality of databases by laboratories:

The most important achievements in the identification database for the population:

- The number of records received was about 30 million records per reference date
- Defining a statistical methodology for building a base of population characteristics and determining the address of residence through administrative records.
- Completion of matching the identification data of individuals among the bodies approved in the census
- Inclusion of the civil registry for the electricity account number to prove the place of residence.
- Inclusion of the account number in the educational portal.
- It is mandatory to include the account number in residential and commercial lease contracts.

The most important achievements in the manpower database:

- Defining a statistical methodology for building a manpower database through administrative records.
- Complete the data coverage of workers through the procedure, such as farmers, fishermen and vehicle owners.
- Setting traffic tables between the classifications of the entities and the statistical classifications of the sector and professions.

The most important achievements in the education database:

- Defining a statistical methodology for building a base of educational characteristics through administrative records.
- Complete alumni coverage.
- Initiating the inclusion of the specializations of foreign graduates in the data level of the Ministry of Labor.
- Initiating ratification and equivalency of academic certificates at the level of the Ministry of Higher Education.
- Building a database for students of community schools.

The most important achievements in the social status database:

- 6 million records were completed for each reference date, and the Ministry of Justice data matched the marital status.

The most important achievements in the disability database:

- The number of records received is about 74 million records for each reference date, and the coverage of disability cases was completed using the data of the Ministry of Health.

The most important achievements in the birth and death database:

- The number of records received is about one million records for each reference date, and the most important achievements are:
- Comparing Ministry of Health data with the Civil Registry.
- Assigning a civil number to the newborn when sending the birth notification.
- Linking Ministry of Health and the Civil Registry to ensure better and faster coverage of deliveries.

The most important achievements in the establishments database:

- Approximately 21 million records were received for each reference date
- Determining a statistical methodology for building a database of establishments and institutions using administrative records.
- Matching and correcting the identifiers of institutions and establishments.
- Circulate the use of the national classification of activities to encode activities at the level of all municipalities.
- Linking the worker to the activity and symbolizing the location of the activity at the level of the manpower.

The most important achievements in the residences database:

- The number of records received was about 91 million records per reference date
- Determining a statistical methodology for building a database of properties of units and buildings using administrative records.
- Matching and correcting account owners' identifiers (civil number / commercial registry)
- Improve the quality of accounts sites.

3.3 Recommendations for improving data quality:

It is worth noting that there are many recommendations that came out of the project team in order to improve the quality of the data as follow:

Population metadata:

- Completion of the inclusion of the electricity account number in the civil registry.
- Unification of the classification of governorates and states, especially gatherings among all administrative authorities
- Listing and coverage of permanent and temporary workers accommodation permits.
- Completion of the addresses of arrivals and monitor the inclusion of the Walayat (Walayat of Muscat in particular)

Manpower Data:

- Completion of base coverage of workers at the level of the manpower.
- Completion of the variable of specialization in the database of expatriate workers.
- Completion of all entities for the Gulf classification of professions to encode professions

Education Database:

- Educational portal coverage for students in community schools
- Educational portal coverage for Omani students studying abroad.

Social status data:

- Recommending the completion of the marital status variable and linking expatriates to the civil registry level.

Disability data:

- Coordination between the Ministry of Social Development and the Ministry of Health to complete the coverage of disability cases

Birth and death data:

- The process of immediately sending birth data to the Civil Registry

Enterprises data:

- Complete linking the municipal license with the activity place code
- Inclusion of the electricity account number in commercial lease contracts and as an identifier for the address
- Assigning a commercial registration number to all government institutions and a place of activity code for its branches

Residences data:

- Completion of correcting account identifiers (10% of accounts)
- Complete the missing coordinates for the accounts (about 40 thousand accounts)

Conclusion:

Data quality assurance is one of the most important pillars in the statistical aspect of countries in the world. The previous traditional censuses in the Sultanate were one of the most important tools for auditing and verifying data in the administrative records in the Sultanate as it is a second source of data that has been updated directly through the field.

With the Sultanate heading in its electronic census 2020 to collect data from the administrative records in the Country, it was necessary to change with it the criteria for ensuring the quality of data and the mechanisms used to verify it.

This is what the Sultanate has achieved through many procedures related to quality assurance, such as reviewing coverage and verifying the classifications used, and verifying the validity and correction of those data by comparing those rules and reviewing the entry and audit rules.

The census laboratories had the largest role in auditing, reviewing, correcting and completing the data, and a large number of decisions issued by the census committees had a role in confirming the quality of data and reforming many databases in line with quality assurance standards.

Perhaps the steps, procedures and results explained in this document confirm that the management of the electronic census project 2020, has achieved data quality standards in accordance with the global frameworks explained by international organizations.

Executive Summary
of the
E-Census
2020

E-Census Documents
Series 2020 | 01

Legal and Regulatory
Framework Project of the
E-Census
2020

E-Census Documents
Series 2020 | 02

Information Security
Management for the
E-Census
2020

E-Census Documents
Series 2020 | 03

Quality of Administrative
Records in the
E-Census
2020

E-Census Documents
Series 2020 | 04

Statistical Methodologies
for Building
E-Census
2020

E-Census Documents
Series 2020 | 05

Statistical Classifications
and Pass Tables for the
E-Census
2020

E-Census Documents
Series 2020 | 06

Electronic Census
System for the
E-Census
2020

E-Census Documents
Series 2020 | 07

The Electronic
Census Portal
for Population, Residences & Establishments
2020

E-Census Documents
Series 2020 | 08

Legal and Regulatory
The Media Aspect of the
E-Census
2020

E-Census Documents
Series 2020 | 09