

A new approach to censuses in the Kingdom of Saudi Arabia

Fahad S. Altekhaifi^{a,*}, Fahad Alfahead^b and Arwa Alshanghiti^{c,d}

^a*Chairman of the General Authority for Statistics GASTAT, Saudi Arabia*

^b*Deputy of Mythologies and Operation/Director of Population and Housing Census Project, GASTAT, Saudi Arabia*

^c*Director of Data Analysis and Decision Making Support Centre, GASTAT, Saudi Arabia*

^d*King Saud University, Riyadh, Saudi Arabia*

Abstract. This paper aims to demonstrate the programme of censuses conducted in Kingdom of Saudi Arabia, which includes the Population and Housing Census and Establishment Census.

In order to improve and further develop efficiency and prepare the Kingdom for an efficient move to an administrative based approach in the 2030 Vision, the General Authority for Statistics will use a Combined Census approach between traditional and register-based census. This Combined approach will use administrative data supported by fieldwork. The new approach has required prior testing of administrative data such as the Saudi National Postal Address and data held by the National Information Centre. The paper will further discuss the progress in the preparation and implementation of this new approach. This will include a thorough review of how the National Address system is being used as the framework for identifying buildings and households, as well as reviewing the results of the pilot studies initiated. The paper will also discuss the implications for the fieldwork presented, which include Computer Assisted Interviewing and the publicity and awareness censuses campaigns. Finally, it will also discuss preparations for the release and publication of statistics in the 2020 round.

Keywords: GASTAT, register based census, National Address System, National Information Centre

1. Introduction: Register-based census

The register-based census is a non-traditional data collection approach applied in the Kingdom of Saudi Arabia, and it boosts the reliability of administrative records of different entities. For the utilization of this approach, see, for example, Sahar and Khudadah [1] in 2012; and for the opportunities and challenges of this methodology, refer to Flander [2]. It must be noted that the General Authority for Statistics (GASTAT) will use a Combined Census approach between traditional and register-based census to perform this national task. This Combined approach will use administrative data supported by fieldwork. The new non-traditional ap-

proach of multi-purpose structure required prior testing of administrative data such as the Saudi National Postal Address of Saudi Post and data in the National Information Centre (NIC) of the Ministry of Interior. It is worth noting that GASTAT used to adopt the traditional census method and intends to introduce a more advanced method. This type of census is more efficient in light of the technology advancement that makes the flow of data, the transfer of information, and the comprehensiveness of their scope at the tip of one's finger. Utilization of the formal records in building Registered Census capacity is an indispensable tool in performing censuses nowadays. This also enables the trends of using the available official records, in many aspects of life and activities of the population, to incorporate filed administrative and geographical information at this stage in order to move to the registered-based 2030 census; see Fig. 2 for census governance.

*Corresponding author: Fahad S. Altekhaifi, Chairman of the General Authority for Statistics GASTAT, Saudi Arabia. Tel.: +966 112894444; E-mail: faltekhaifi@stats.gov.sa.

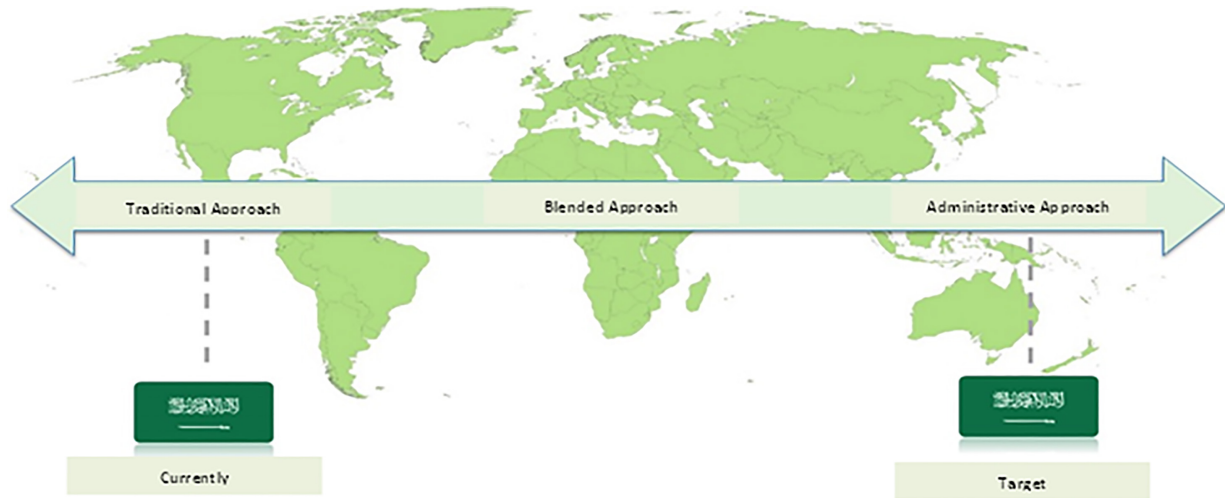


Fig. 1. Development of the census mythology.

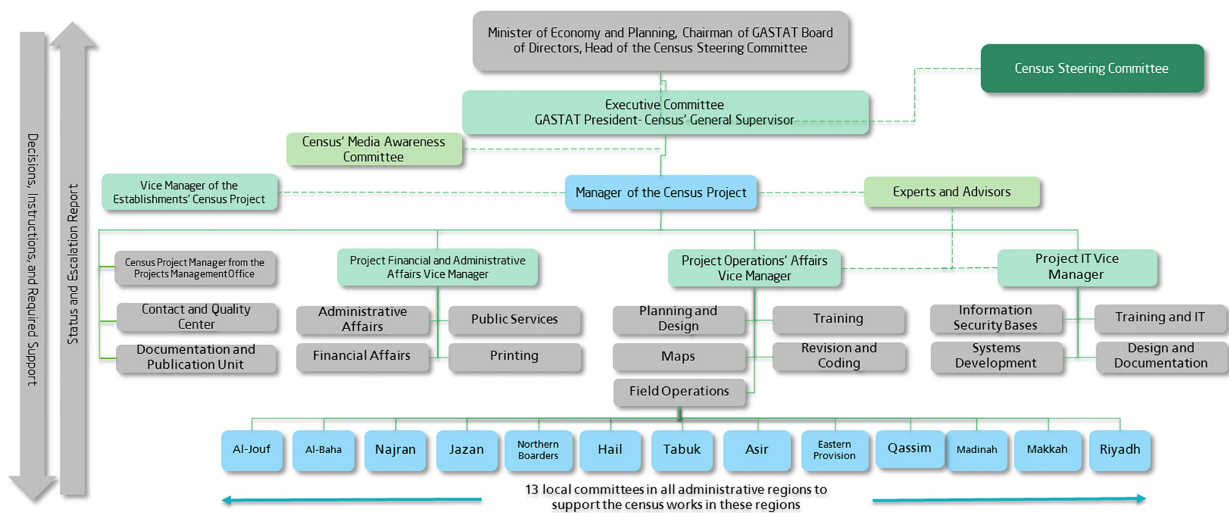


Fig. 2. Governance of the census.

The work of GASTAT for the Census is done according to the following legal framework:

1. GASTAT has been legally authorized to perform population census based on the Population Census Law issued by a royal decree in 1971.
2. A Saudi Council of Ministers decision in 2016, that authorizes GASTAT to perform General Population, Housing and Establishments Census by 2020.
3. A royal Decree has been issued in 2019 that states “the base time line of the General Population, Housing, and Establishments Census has been announced to be on Tuesday March 17, 2010”.

The characteristics of the required approach can be summarized in three main dimensions, as follows:

1.1. Statistical dimension

This dimension consists of many elements that utilize the quantitative and qualitative measures of the census, namely:

- a. Using administrative records and collecting some data in the field.
- b. Using self-enumeration and post-enumeration.
- c. Applying rules that directly detect errors during data collection.
- d. Using addressing system.
- e. Applying the quality system through central op-

- erating rooms.
- f. Sharing the questionnaire with all clients. A number of workshops were held, which all stakeholders attended to discuss the questionnaire and the requirements of the authorities and official entities before the final approval.
- g. The outputs of this census are used as a frame for surveys once the linkage processes are completed.

One of the main quality measures for the register-based census for individual, population household, and housing is associating the census to the government or the official records through the unified National Identity Number. This is also connected to the records of education and commercial activities of population. Another measure of quality is the proper unification of various definitions, concepts, and abbreviations, which simplifies, later on, putting these outcomes into records ready for use and analysis. GASTAT is working with almost all public and private data producers and data users in Saudi Arabia, but it hopes that the legal framework will guarantee data accessibility and transparency. It is noted in this context that information associated with the National Identity Number for Saudi individuals is reliable, precise, and comprehensive. The records of the National Identity Number include name, gender, age, marital status, education level, type of work, and national address.

1.2. Technical dimension

This dimension illustrates the technical issues in the census and particularly includes:

- a. Use of buildings coordinates and addressing systems in the cities and villages with the accuracy of 1 meter.
- b. Data collection systems by using the CAPI method “computer-assisted personal interviewing”.
- c. Synchronizing data from the National Information Centre and GASTAT databases while interviewing the head of the household using tablets.
- d. Following up the enumerator’s progress in the field and calculating all the field quality indicators.

1.3. Data availability and publication

This dimension is the final step that the data go through in order to prepare the finding in a beneficial manner for various types and classes of stakeholders. To be specific, this dimension includes the following:

- a. Instant results during the enumeration period and

after completing the enumeration process.

- b. Flexibility of bulletin at any required geographical level.
- c. Use of all means of bulletin (traditional and electronic), in addition to the use of business intelligence system.

2. Development of the census methodology

While conducting the Population and Housing Census in 1974, 1992, 2004, and 2010, Saudi Arabia relied on the traditional enumeration method, in addition to the Intelligent Character Recognition (ICR) technology. Portable devices were also used in the Establishments Census 2010, whereas tablets were used in the agricultural census, household, and economic surveys, and Hajj statistics. For more details, see GASTAT [4]. Saudi Arabia also seeks to use the register-based method while conducting the Population, Housing, and Establishments Census 2020 to cope with the rapid development of the Information Technology field that facilitates data flow and comprehension. Hence, it is necessary to change the method of enumeration from a traditional one into a more modern method, namely “the register-based method”.

3. Linkage with the National Information Centre

It is an electronic linkage between GASTAT and the National Information Centre through private and secure network for the exchange of data. This linkage means integration of data about the population and housing in GASTAT and NIC. It has four main objectives, as follows:

- a. Accessing the records of individuals, households, housing, and establishments and linking them with the individuals’ national numbers.
- b. Unifying all the definitions and concepts, so that they can be transferred easily into statistical records that can be analysed and processed anytime.
- c. Use of register-based data in all household and establishments surveys.
- d. Developing a unified statistical number for establishments.

4. Household register-based data request

The UN Economic Commission for Europe [3] has

Fig. 3. Household register-based data request.

recently published a guideline on the use of registers and administrative data for population and housing censuses. This guideline has been prepared by the Task Force on register-based and combined censuses. It contains the necessary definitions of the main concepts of registered-based and combined censuses and the requirement for transition from traditional to register-based census; and it introduces a common framework for using registered-based census. In our registered-based census, the census department is aware of other experience elsewhere as described in the above-mentioned reports and references therein.

GASTAT work has emphasized two main factors:

1. Service inputs: This includes two main variables:
 - a. Head of household ID
 - b. Wife ID
2. Service outputs for all household members. The service side has the following seven variables:
 - a. Name
 - b. Age
 - c. Gender
 - d. Nationality
 - e. Relationship to the household head
 - f. Marital status
 - g. Educational status

Figure 3 illustrates a list of household members that appears in tablets.

5. Use of addressing

To develop its coding system of buildings and housing units, GASTAT works on using a unified national coding system for addressing. This system includes more than (7 M) national addresses managed by Saudi Post. It relies on the Geographic Information Systems (GIS), whereas its structure was developed according to the international coordinates system with a 1-meter accuracy. Through this system, two layers might be built:

- Postal codes that are based on geographical factors, population density, borders, and road networks.
- A layer that provides an address for each location inside the postal codes.

It is worth mentioning that Saudi Post Commission has done an excellent work by introducing the national address for Saudi or non-Saudi individuals, housing units, and buildings in various parts of Saudi Arabia. This activity has been done in the past two decades but used heavily by many people in the past 10 years. This is called the Unified National Enumeration System for Postal Addresses. It is based on the geographical information system and international axes. The precision of the location is about one meter, and it includes more than seven million addresses. GASTAT has

utilized this system and added the official administrative division of every location.

6. Preparatory stages of 2020 census

GASTAT adopted three stages for performing the registered-based census; these are:

6.1. Stage one: Updating populated areas' guidebook

- a. A guidebook was prepared for the populated areas according to Saudi Arabia's regions system; the data was processed and the administrative regions of Saudi Arabia were documented as well. The number of the populated areas based on the administrative division is as follows:

Administrative regions	13	Municipalities	1527
Governorates	149	Populated areas	29591

- b. Estimating the number of dwellings for all populated areas and distributing them into statistical working areas.
- c. Confirming the guidebook connection to each national address in all populated areas and setting statistical boundaries through that connection.

Stage one updates the directory of definitions for housing, an important preparatory process that precedes the census works. During this stage, one has to visit all population communities in Saudi Arabia and make sure of various administrative connections. Also, one has to update all national addresses for all buildings and estimate the number of housing units for every building. The labour force needed for the general census is determined in this stag as well.

6.2. Stage two: Updating cities

- a. Identifying cities' boundaries clearly and accurately based on the data held by the secretariats and municipalities.
- b. Estimating the number of dwellings and establishments for each national address.
- c. Distributing the cities and their neighbourhoods and blocks (national addresses) into statistical working areas.

6.3. Stage three: Coding of buildings, dwellings, and establishments

- a. Numbering the buildings, dwellings, and establishments.
- b. Distributing observers' working areas into enu-

meration areas and identifying the actual number of labour force.

- c. Preparing a complete framework for buildings, dwellings, and households connected to the national addresses.

For coding the building, every supervisor of postal area and his team have to visit all national addresses in his area and make complete enumeration of housing units and past numbers on the door of every unit. Housing units in every building are given a serial number starting from one to the end. Also, the questionnaires for housing units has to be fully implemented. This questionnaire includes all housing units whether vacant, occupied with families, occupied by office, used for work and living, or under construction. This work provides good information about the work size and the field enumerators needed to execute the census task.

One of the main steps in the forthcoming census is the point enumeration that intends to reapply the census to a sample from the past, which includes the work of many of the census field enumerators by an independent team to estimate any possible mistakes in the population and housing census. This review is applied to various parts of Saudi Arabia and includes the work of different field enumerators. This process illustrates the comprehensiveness and accuracy of the census. Obviously, there are differences among the self-enumeration and the point enumeration in their style, steps, and management. The self-enumeration is performed during the census data collection, whereas the point enumeration is done just after completing the census.

7. 2020 census pilot tests

2016: Testing the post address.

2017: Testing the data of the National Information Centre.

2018: Testing the fulfilment of the census's questionnaire.

2019 Q2: Testing the fulfilment of the population and establishments census's questionnaire.

2019 Q4: Testing the fulfilment of the population and establishments census's forms questionnaire.

8. Time reference and labour force

The time reference of the General Population and Housing Census was set on 17 March 2020. The Cen-

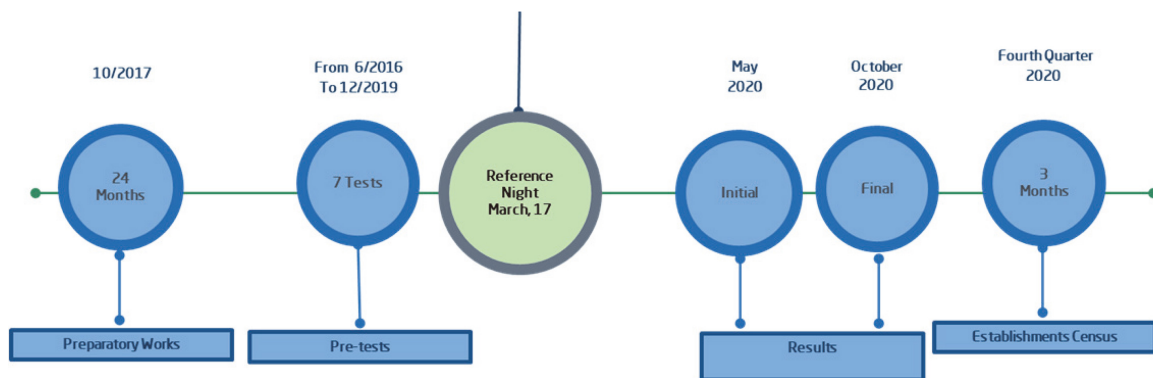


Fig. 4. Time reference.

sus will be implemented with the participation of more than (55000) supervisory or executive entities including (44000) field enumerators. All participants in the census have gone through training and shown their skills through training, workshop, and the appropriate testing of their enumeration and statistical basic skills; see Fig. 4.

9. Data collection

Data collection is usually done by trained data collectors or enumerators, and the process follows one of the following choices:

- Direct personal interview, where field enumerator visits the family in his geographical previously determined area. Every geographical area includes a number of national addresses, which simplifies the accessibility to the intended places. The field enumerator carries his board for collecting and recording data during his interview with the household's head. At this point, the enumerator can check the field data with the retrieved data from the official previously registered ones. The enumerator at this stage can complete the missing or desired data. This process has different measures for assurance of the data quality. Finally, the field enumerator sends the data to the central data storage at the GASTAT's headquarter in Riyadh.
- Self-enumeration, which is a new technological method to be used in the present population census. In this choice, there is no need for the field enumerator's visit; but the family head, through a specified link, fills in all required data concerning his family. The data can be uploaded to the specified website once or in consecutive stages according to the time allocated by every family for this

work. In case of queries, GASTAT has made hot telephone lines for helping self-service enumeration. In fact, the GASTAT makes a wide media campaign through radio, television, and other social media venues. This requires a fewer number of field enumerators and gives a flexible time for the household's head to provide the required data at their convenience.

It is worth noting that all shared information from the National Information Centre at the Ministry of Interior are subject to agreement with the GASTAT to guarantee the individual information protection, where such information are only used for censuses purpose and quality assurance processes.

For the quality assurance of the data collection process and the data collected during the census, the GASTAT has adopted a comprehensive program and procedures that include rules for correction and verification of data items. This is done during the visit by the field enumerator to the household. In addition, a qualified unit for quality and communication ascertains the reliability and quality of the data collected in all geographical locations of Saudi Arabia at a daily basis. Quality control includes repeating the enumeration through telephone for a reasonable number of families and it follows all the remarks given during self-enumeration as well.

10. Sections of the census questionnaire

Figures 5 and 6 illustrate the 9 different sections of the Population and Housing Census questionnaire and the Establishments Census questionnaire, respectively. We can refer to some handbooks in other countries that describe questions that will be asked in 2020 Census; see [5] for example.

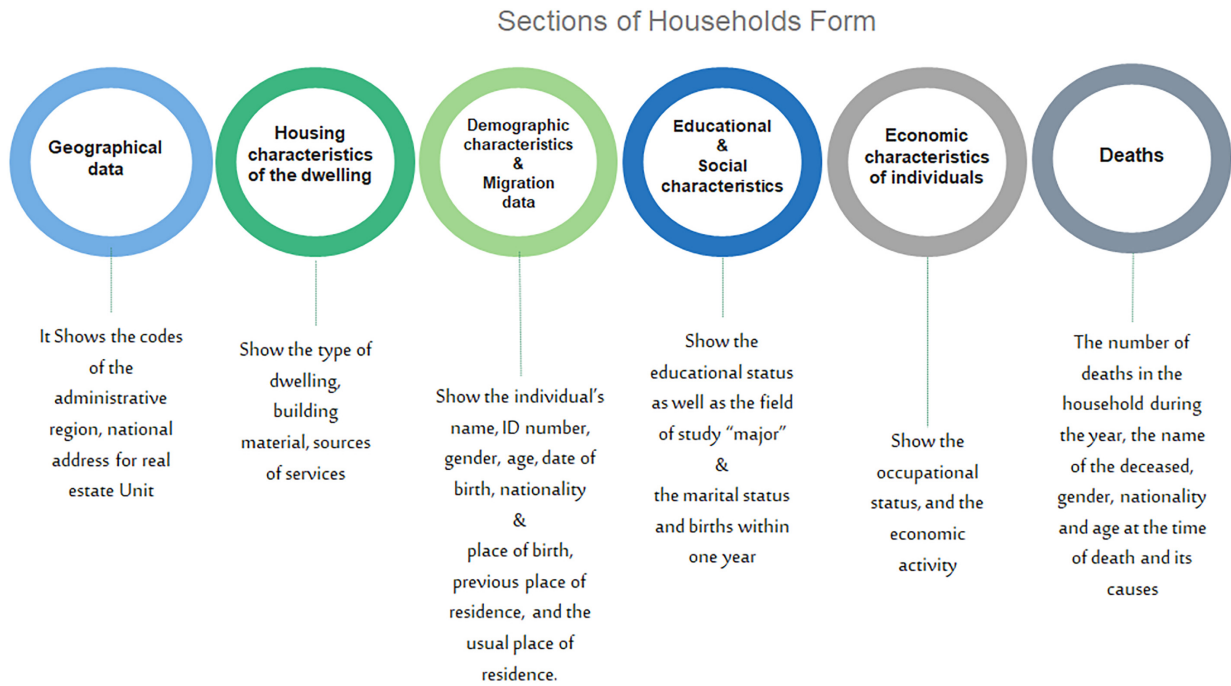


Fig. 5. Population and Housing Census questionnaire.

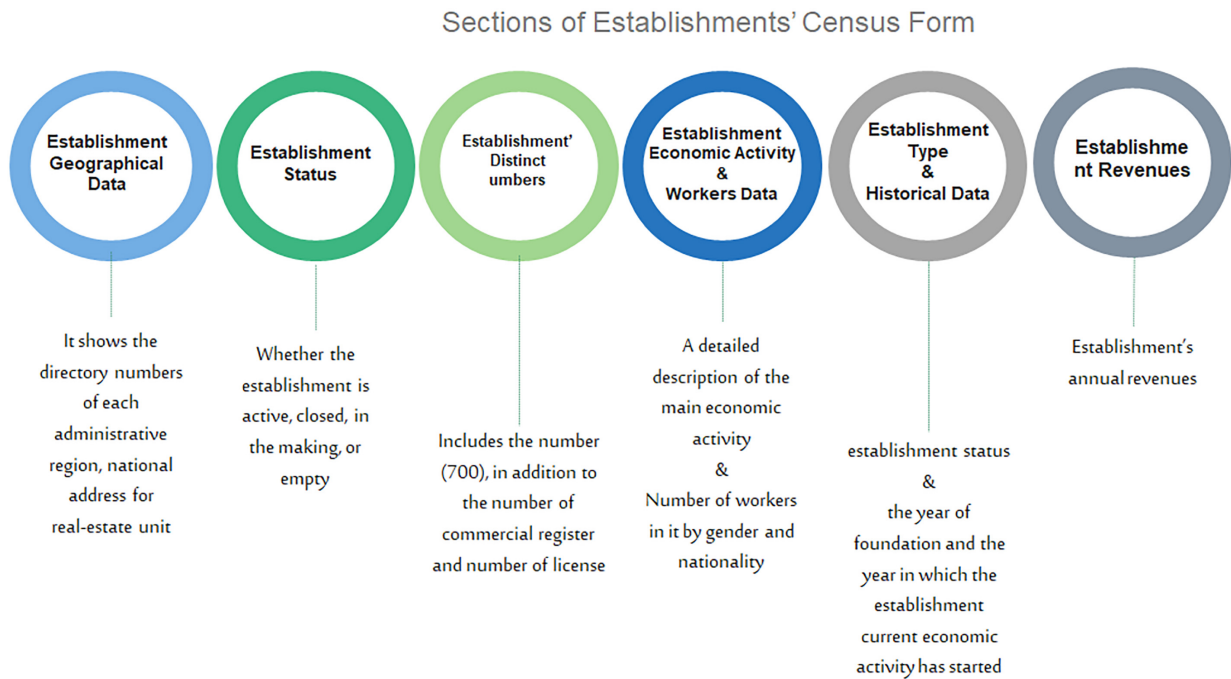


Fig. 6. Establishments Census questionnaire.

11. Results

It may be noted that the experience of GASTAT, see [4], in implementing the registered-based census approach, with extensive knowledge and awareness of all aspects and characteristics of the register-based census, has yielded the following main results:

- a. Easy data synchronizing.
- b. High data security.
- c. Using geospatial technology in the enumeration stage.

Acknowledgments

The authors are very thankful to reviewers for their comments that was helpful to finalize the paper.

References

- [1] Sahar AY and Khudadah A. Non-traditional Data Collection Approaches on 2011 Kuwait Data Census. Country Experiences in the 2010 of the population and Housing Censuses, OIC-StatCom. Directorate of Technical Department and IT, State of Kuwait; 2012.
- [2] Flander AO. Opportunities and Challenges of a Register-Based Census of Population and Housing – the Case of Slovenia, Seminar on Registers in Statistics-Methodology and Quality. Helsinki; 2007.
- [3] Economic Commission for Europe. Guidelines on the use of registers and administrative data for population and housing censuses. Conference of European Statisticians Sixty-sixth Plenary Session; 2018.
- [4] GASTAT. The General Population and Housing Census. General Authority for Statistics, Kingdom of Saudi Arabia; 2019.
- [5] U.S department of Commerce Economics and Statistics Administration, U.S. Census Bureau. Questions Planned for the 2020 Census and American Community Survey; 2018.