



# Economic and Social Council

Distr.: General  
12 July 2017

Original: English  
English and Russian only

---

## Economic Commission for Europe

### Conference of European Statisticians

#### Group of Experts on Population and Housing Censuses

##### Nineteenth Meeting

Geneva, 4-6 October 2017

Item 3 of the provisional agenda

##### Innovations in census technology

## 2020 Population and Housing Census Planning, innovations in the use of technology

### Note by National Institute of Statistics, Geography and Informatics of Mexico

#### *Summary*

For the first time in Mexico, the 2020 Population and Housing Census data collection will be carried out through a CAPI (Computer-Assisted Personal Interviewing) scheme as the main enumeration method, but will also include the CATI (Computer-Assisted Telephone Interviewing) and CAWI (Computer-Assisted Web Interviewing) methods. These innovations, given the census structure size and the rapid changes in technologies, are a significant challenge for INEGI. Progress in census planning will be presented, including the main challenges to be faced, the innovations to be tested for their implementation, as well as the successful experiences in the use of technologies for georeferencing the information both in the data collection stage and for results dissemination.

GE.17-11638(E)



\* 1 7 1 1 6 3 8 \*

Please recycle The recycling symbol, consisting of three chasing arrows forming a triangle.



## **I. Background of Censuses in Mexico**

1. In 1895, the first census exercise of national status was carried out in Mexico, and with it, all the contemporary censuses were initiated. The second national census was carried out in 1900 and since then, the census projects of population and housing have been carried out every 10 years, with the exception of the one postponed until 1921 due the Mexican Revolution. This periodicity has allowed to have statistical information on a regular basis, which facilitates knowledge about the evolution of the living conditions of the country's population, and enables the public and private sectors to make informed decisions. Since 1930, censuses in Mexico have been "de jure" and data collection is achieved through face to face interviews to an adequate informant using printed questionnaires.

2. The 2010 Population and Housing Census (CPV2010) is the most recent census in the country. Its main purpose was to enumerate the resident population of the country, to update information about their main demographic and socio-economic characteristics, and to locate their distribution in the national territory, as well as to enumerate the dwellings and to collect information about their basic characteristics. It was also the first census conducted by the National Institute of Statistics and Geography (INEGI) as an autonomous public institution (2008).

3. For the 2000 and 2010 censuses, two types of questionnaires were used: the basic form and the long form that seeks to deepen the census subject matter and is applied only to a probabilistic sample of dwellings.

4. In the last decades, the users' demands to have updated information in periods shorter than the decennial ones have been increasing, as well as the request that it can have a greater level of disaggregation. In order to provide information between censuses, two population and housing counts have been carried out, the first in 1995 and the second in 2005. In both cases, the main objective was to generate basic socio-demographic information, which allowed to have updated data on population and housing in the country.

5. Based on these experiences, in 2015 INEGI considered that the most viable option to update the demographic and socio-economic information of Mexico in the intercensal period was to conduct a probabilistic survey with a broad thematic coverage that sought to satisfy the demand for information as much as possible. Its design was conceived so it would allow comparability with previous censuses, cost reduction, and could provide estimates for the total population by state, municipality and for the main cities of the country, as well as for the main characteristics of population and private inhabited dwellings. Thus, in Mexico, thirteen censuses, two population counts and one intercensal survey have been carried out to date.

6. The use of technological advances has allowed to reduce the time for publication of definitive census results, for example from almost 2 years in 1990 and 15 months for the 2000 Census, to 9 months in the case of the 2010 Census. Although electronic devices for data collection have not been used in the census operations to date, technology has made possible to speed up processes and to reduce the time of results delivery.

## **II. Planning the 2020 Census**

7. The population and housing census is a project that generates the most important statistical information in the country, since it covers the entire population and dwellings of the national territory in a short time. This implies having a great amount of human, material and financial resources distributed timely and in due form in all the places where they are required; hence, its planning needs years of anticipation to the field data collection.

8. INEGI, as an autonomous organism responsible for generating sociodemographic information of national interest necessary to assist the development of the country, will carry out the 2020 Population and Housing Census (CPV2020). For that year, the population in the country is estimated between 128 and 130 million inhabitants, and the number of dwellings between 45 and 46 million, of which 35 million are expected to be inhabited. This statistical exercise will seek to enrich the historical series of census information, keeping as far as possible the comparability with the operations carried out in Mexico and in other countries. Also, it will generate the inputs for the elaboration of population projections and for the construction of the sampling frames over which the household surveys are carried out. Thus, by 2020 the census operation will look to:

- Perform an exhaustive enumeration of the population and existing dwellings in the country;
- Deepen the knowledge of some priority topics, through the application of a long questionnaire to a probabilistic sample of dwellings;
- Obtain information of optimum quality;
- Deliver results timely.

9. Although the possibility of replacing traditional censuses by a census based on administrative records is not visualized in the medium term in Mexico, by 2020 several technological innovations will be implemented, the main one being the data collection with mobile devices in order to make the processes more efficient and to reduce the time for results delivery. This change in the main method of data collection represents an important challenge for INEGI, given the size of staff structures and the rapid changes in technologies.

10. For the planning of this event, the methodological, conceptual and technical contributions of the previous statistical projects carried out, mainly the 2010 Population and Housing Census and the 2015 Intercensal Survey, will be evaluated, as well as the experiences obtained from the Census of Schools, Teachers and Students of Basic and Special Education 2013 (CEMABE) in matters of the use of technology for data collection. Furthermore, the Principles and Recommendations for the 2020 Census Round, issued by the United Nations (UN), will be considered.

11. The following sections will describe the main characteristics of the 2020 Census, the advances in planning and the technological innovations currently in testing phase for their further implementation.

## **A. Methodology**

12. For the 2020 Census, INEGI considers the development of the project under the following methodological basis:

- Type of census: “de jure”, that is, the population will be registered in their usual place of residence;
- Observation units: “Residents of the national territory and private and collective dwellings”;
- Questionnaires:
  - Short form: For the exhaustive enumeration, about 30 questions.
  - Long form: For a probabilistic sample of dwellings, about 75 questions.

- Social Assistance Accommodations (CAAS): About 145 questions in 4 700 establishments.
  - Proxy respondent: The head of the household, either woman or man, his or her partner or a resident of 18 years of age or older;
  - Data collection period: March 2020;
  - Special field operations: Population in collective dwellings, Foreign Service Personnel, personalities and homeless population.
13. Additionally, taking advantage of the field operations in the whole of the national territory, information of each of the localities under 5 000 inhabitants will be collected, applying a questionnaire of about 100 questions to an appropriate informant (usually some authority of the locality). For the localities of 5 000 inhabitants or more, information on the characteristics of the urban environment of each of the blocks will be registered during the data collection (availability of public lighting, sewage, public transports, wheelchair ramps, etc.).
14. Regarding the data collection methods, the implementation of the following strategies is considered:
- Face to face interview using mobile devices, as the main strategy;
  - Internet self-enumeration;
  - Telephone-assisted interview.

## **B. Census organizational structure**

15. INEGI has a permanent structure in charge of the design, management and monitoring of statistical information generation projects. This central structure is responsible of defining the operation strategies of the census, that are coordinated and executed through the territorial structure of INEGI composed of 10 Regional Directorates (DR) and 34 State Coordinations (CE). In order to carry out the CPV2020, temporary staff will be hired in each CE, and operation offices will be established at municipal level throughout the territory. It is estimated that the operation structure will be approximately of 200 thousand people.
16. The following general activities are established in the general program of the 2020 census:
- By 2017, the approach of census scenarios is contemplated, as well as the public consultation for the questionnaires content and methodology. During this year, the definition of the census project and the tests of operational procedures will begin, among other activities;
  - In 2018 the planning and execution of thematic tests will be performed, as well as the definition of the conceptual framework and the data collection instruments. During this year the pilot census will be carried out, also the bidding activities for the purchase of mobile devices to be used in data collection will begin;
  - For 2019, the statistical design of the census sample, cartographical update in rural Basic Geostatistical Area (AGEB) and the statistical design of the post-enumeration survey will be carried out. In addition, the recruitment and selection of personnel, the training of operational staff and the execution of the communication campaign will initiate;

- Finally, in 2020, performing the data collection of the urban environment characteristics and localities of less than 5 000 inhabitants is contemplated; the Population and Housing Census will be launched, the post-enumeration operation will be undertaken, the information processing will be executed, results will be generated and figures will be released. The publication of results will be done before the end of the year.

## C. Computer Assisted Personal Interviewing (CAPI)

17. Currently, the implementation of Computer Assisted Personal Interviewing (CAPI) schemes has allowed National Statistical Offices (NSO), polling companies and research centers to reduce the time reduction for publication of results. This modality consists of a face-to-face interview in which the interviewer uses a computer system from a mobile device.

18. As mentioned above, the CAPI method will be used as the main strategy for field enumeration in the CPV2020. This section provides a general description of the main considerations for the adoption of this technological innovation.

### 1. Previous experiences

19. INEGI has previous experiences in data collection with mobile computing devices (MCD) under the CAPI scheme, a strategy that has been used in the following projects:

- CEMABE 2013;
- Economic Censuses, 2014;
- Agricultural Censuses and Surveys;
- Household Surveys;
- Economic Surveys;
- Government Censuses.

20. Based on the experience obtained from these projects, the following advantages and disadvantages in the data collection with mobile devices have been determined:

Advantages:

- Reduction in times for data collection and faster publication of results;
- Elimination of paper work and implementation of certain data integrity controls (data validation during the interview);
- Online monitoring;
- Elimination of the traditional data capture structure (data capture centers, staff, logistics, etc.).

Disadvantages:

- High investment in MCD: one device per interviewer and supervisors and spare devices;
- Short life of MCD, as well as rapid changes in technologies;
- Exposure to robberies in high-risk places;
- Sensitivity to extreme weather conditions;
- Risk of technical failures;

- Requirement of additional information security measures.

## **2. Characteristics of computer equipment**

21. For the events mentioned in the previous experiences, mini laptops with a 10.1” screen and technical characteristics sufficient for their use in field were used. The model of mini laptop used has a solid and scarcely showy case, which makes them unattractive for robbery. While this equipment has been useful in the previous operations, the decision to look for cheaper and lighter equipment for the CPV2020 has been taken. Hence, it was decided to plan the data collection of the CPV2020 with tablets of 7” or 8” screens, whose suitable characteristics will be determined from the results of technical and operational tests.

## **3. Subsequent use options**

22. One of the critical points in a CAPI scheme is the subsequent use of the MCD. As mentioned before, about 200 thousand equipment will be required, therefore, the changes in technologies, as well as the useful life and storage costs of equipment make their later use impossible even in the next intercensal event.

23. Although the Bring Your Own Device (BYOD) scheme solves the problem of the subsequent use of the computer equipment, technical complications are derived from having a very wide diversity of computer equipment and implications in terms of information security arise. In this context, INEGI is currently working on the search for alternatives to reuse the census MCD.

24. Around 20% of the MCDs are expected to be used in surveys and other INEGI projects, in addition, agreements with other government agencies are being sought to allow a shared acquisition or a donation that strengthen the production of administrative records for their incorporation into the National Statistical and Geographic Information System (SNIEG).

25. Finally, although the CAPI scheme will be the main method for data collection, a Pen And Paper Interviewing (PAPI) scheme will be maintained for exceptional cases such as risk areas or for allowing that the enumerators continue with their work day in case of equipment failure or battery exhaustion.

## **D. Self-enumeration alternatives**

26. In addition to the CAPI data collection strategy, the Population and Housing Census will also try to use the Computer Assisted Telephone Interviewing (CATI) and Computer Assisted Web Interviewing (CAWI) methods as secondary strategies in the following cases:

- Delivery of an invitation for self-enumeration by Internet or telephone if at the time of the third visit of the enumerator the interview has not been achieved, or if in any of the visits the informant makes the express request for self-enumeration;
- As an initial strategy of self-enumeration in blocks with more than 75% of the dwellings that in the previous event declared to have Internet, in localities of 100 thousand or more inhabitants;
- As enumeration alternatives for zones with restricted access.

27. It is important to note that, for the previous cases, INEGI is currently working on the appropriate operating procedures and tests to guarantee that neither the data collection times nor the coverage levels are affected.

28. The Internet self-enumeration strategy will be tested in October 2017, in the context of the operational and computer equipment test, through the delivery of printed notices with unique identification keys for each dwelling. The design of the Web application to be used, will also allow the respondents to answer the questionnaire (only basic form) from any type of device (PC, laptop, tablet, smartphone).

### **III. Georeferencing statistical information**

29. The production of georeferenced statistical information is increasingly common in the National Statistical Offices, and its use in the public, private and academic sectors has intensified. Recognizing that data has a location context is an important factor in decision making; in order to maximize the potential of this information, it is necessary to have adequate mechanisms of data collection and processing, as well as to promote the use of technological innovations.

30. The following sections will briefly present the processes implemented in INEGI regarding this topic, as well as the innovations planned for the 2020 Census.

#### **A. Geostatistical Framework**

31. The National Geostatistical Framework (MGN for its acronym in Spanish) is a unique system, designed and created by INEGI in 1978, to georeference the statistical information of censuses and surveys, which is updated periodically, and has been evolving up to digital versions available free of charge in the institutional web site. In this system, each geostatistical area has a unique key in the national territory.

32. In the census context, the MGN requires permanent updating, through the activities previous to a census data collection, and will be updated during the census taking. This continuous update guarantees the coverage and correct geographic referencing of the census information across the national territory. Once the data collection is achieved, the MGN is used for the publication of results and geo-statistical systems, such as the 2010 System for the Consultation of Census Information (SCINCE) and the National Housing Inventory.

#### **B. Use of QR Codes**

33. Among the lessons learned from the 2015 Intercensal Survey (EIC2015), is the use of QR codes for the questionnaires control, from their reception and distribution in field offices to their delivery to data capture centers and their storage in regional or state warehouses. In addition, a great advantage in the use of a QR code in each data collection instrument was the possibility of linking the corresponding questionnaire to each dwelling in the building list, allowing the monitoring of each questionnaire in the different stages of the process and ensuring the correct association of the information with the geographical area in which it was collected.

34. In the context of the CPV2020, it is planned that the paper questionnaires, available for the exception cases, will have QR codes, which will be read through the camera of the mobile devices, to carry out the control in a similar way to what was done in the EIC2015. Additionally, it is intended that the labels that will be placed in the dwellings during the data collection will also have a QR code, which will allow assigning a unique key to each dwelling in the country. Another use for QR codes will be the unequivocal identification of the dwellings during the verification and post-enumeration phases, as they are expected to solve the dwellings identification problem that occurred during the 2010 post-enumeration survey due to deficiencies in the addresses (lack of road names and numbering).

35. Finally, self-enumeration notices will have a QR code that contains the URL to the census web system, as well as the user and password linked to the dwelling.

### **C. Use of GPS**

36. In the activities to be carried out for the CPV2020, and taking advantage of the fact that the MCDs have a Global Positioning System (GPS), INEGI is considering to register coordinates of the dwellings during the field operations. Although it is known that the use of GPS may present precision problems, tests on this matter will be conducted during 2017, anticipating that the captured points can be corrected and at least be properly associated to the block to which they belong. This technology is also intended to be used for monitoring the enumeration by verifying that enumerators conduct the interviews in their assigned work areas.

### **D. Geographic Information Systems**

37. Nowadays, INEGI is one of the few official statistical agencies that integrate, in a single government entity, the official statistical and geographical activities. This fact represents an advantage, since by means of the cooperation of both areas and the application of technological innovations, it is possible to generate products for the dissemination of the information, which additionally to statistical data and its geographical reference, include advanced tools for consultation and analysis. Examples of these systems are the following products:

- 2010 System for the Consultation of Census Information (SCINCE):  
Tool that allows to associate the statistical information of the CPV2010 with the geographic space to which it belongs, achieving a contribution of complementary information is obtained to facilitate the interpretation of the sociodemographic phenomena in the country. It offers a series of sociodemographic indicators with disaggregation up to block level and which can be analyzed applying different methods of univariate and multivariate stratification.
- 2016 National Housing Inventory:  
Geographic Information System (GIS) that integrates statistical information with a territorial vision, on dwelling, population and urban environment. It considers the construction of indicators that characterize the dwellings and the population, at different geographical levels. Its information sources are the CPV2010, special field operations focused on new residential growths, and the EIC2015.

## **IV. Conclusions**

38. Traditionally INEGI has sought the incorporation of technological innovations to the processes of the statistical projects, in order to make them more efficient and improve the quality of data.

39. Particularly for the 2020 Census, one of the most important technological innovations planned, is the use of mobile devices during the data collection. This fact implies a series of important challenges for the Institute, which should be analyzed and evaluated. Several tests will be carried out during the current and the following year in order to ensure that the inclusion of this type of innovations lead to a tangible benefit for the project that contributes to achieve the objectives of the Institute in terms of information production: reliability, opportunity and cost reduction. Hence, INEGI seeks the

modernization of its processes and will promote actions that, complementary to the 2020 Population and Housing Census activities, contribute to the use of administrative records for statistical production.

---