



The population census, methodology, quality control, output and challenges¹

Alphonse L. MacDonald*

General Bureau of Statistics (ABS), Paramaribo, Suriname – fonzhan@gmail.com

Abstract

The census is a universal scientific instrument with well-established methodology. The population census satisfies the need for information on the population and its characteristics for the administration of societies. Census and registers in historical perspectives. The role of scientific organisations, such as the International Statistical Institute (ISI), the Inter-American Statistical Institute (IASI), and the United Nations and its agencies in developing and promoting census methodology is acknowledged. The population census has a well-established generally accepted methodology, with quality control criteria with special emphasis on measurement errors, completeness of enumeration: over- and under coverage. The census has several types of output with scientific, administrative and social utility and use. Recently alternatives have been proposed to use administrative records and registers as substitute for field enumeration. The challenges faced by the execution of the census will be reviewed.

Keywords: population census, methodological requirements, quality criteria; completeness.

1. Introduction

The census as a scientific data collection instrument is usually identified with reference to the population census. However, the census is a scientific data collection instrument used in a wide variety of scientific disciplines from astronomy, to botany, through mineralogy to zoology. In population science three sources of statistical information are generally distinguished: administrative records (and derived administrative registers), the census and the sample survey. Each source is used to obtain specific types of statistical information. Recently efforts are being made to assess the suitability of “Big Data” as a source of statistical data. All societies of a certain complexity require information to manage their affairs. They require information on the size and composition of the population for administrative purposes: taxation, conscription for military or other services, social control, or for social welfare measures, food security. Classical historical societies, such as Babylon, Egypt and China had well developed records on population, land size, agricultural production and trade patterns. These records were probably lists or inventories, kept by specific officials or scribes. The exact procedures used are not well documented. There is evidence that in addition to routine records, occasionally *ad hoc* counts were taken to satisfy the demand for specific information. These activities were carried out when new territories were conquered or in cases of dynastic change. They were generally described as censuses, but the term headcount would be more appropriate.

The coverage of these administrative inventories was restricted to the population under the control of the rulers. It is most likely that these were not always counts of individuals, but counts of aggregates, such as families or households from which the total population size was calculated. However, in the post-medieval period on at least two occasions, individuals and their characteristics were registered. In 1665 the King of France mandated that a census of the population of Quebec (New France) should be carried out (Sulte 1882, pp. 51 – 84) and in 1703 a national census was carried out in Iceland (UNESCO), then a dependency of Denmark. These enumerations are considered the forerunners of the modern census because all persons were registered separately with their individual characteristics.

The scientific basis of the population census was established in the late nineteenth century at International Statistical Congresses between 1853 and 1876, originally organised under the dynamic

¹ This is an abridged version of part of a larger essay on “the population census and its alternatives”, under preparation. The opinions expressed are the author’s and not necessarily those of ABS.



leadership of Adolphe Quételet. At the Jubilee meeting of the Royal Statistical Society in London in 1885 the International Statistical Institute (ISI) was established to continue the tradition of international meetings of statisticians (De Neumann Spallart 1886). The ISI and the League of Nations continued to promote the census methodology until 1940. After the Second World War the United Nations Statistical Commission and the Population Commission became responsible for the promotion of the census methodology. Guidelines were developed that relied, partly, on the experience and activities of the International Labour Organisation (ILO) and Inter-American Statistical Institute (IASI), a regional branch of the ISI established in 1940. One of the key recommendations was to take decennial censuses within the context of the World Population (and Housing) Census Programmes. Thanks to the efforts of the international donor community a large number of newly independent countries were able to carry out population censuses. During the 2010 World Population and Housing Census programme 214 countries or territories conducted at least one census covering 97 % of the estimated world population. Only 21 countries or areas did not participate in the 2010 programme (United Nations Statistical Commission 2015, pp. 2-3). A small number of countries, did not use the standard census methodology (United Nations Statistics Division, pp. 9 – 11), but used alternatives like the “rolling census” as yet only used by France (United Nations 2007, pp. 20 – 21) and a variety of register-based “censuses” (United Nations 2007, pp. 43 – 46) used by 15 countries or areas, 12 in Europe, two in Asia, and one in Northern America. In this paper issues related to the standard census methodology are discussed; those of the register-based “censuses” are treated in a separate paper.

2. The development of the definitions of the population census

Although the census has been used for millennia as an administrative activity one very seldom finds a definition of a census. When the methodology of the population census was established at the first International Statistical Congress in Brussels, Belgium in 1853 no definition was given, but a set of its characteristics and recommendations for its execution were provided. The Congress had a clear idea of the distinction between a census and a register, as it also proposed detailed recommendations for an annual civil register (births, deaths, marriages and divorces) and emigration and immigration registers (Levi 1854, 5 – 7). At the Eight International Statistical Congress in St. Petersburg in 1872 more detailed recommendations were adopted. The censuses should be taken at least once in ten years, in the years terminating with 0; it should be taken in one day, or at least relate to a fixed day and hour. The Congress defined the types of population that could be enumerated: the population “de fait” (actual population), the population “de séjour habituel” (usual residence) and the population “de droit ou légale” (legal population) and recommended that the each individual in the actual population be enumerated; that individuals be identified through the family; that special agents (or enumerators) be used; and that individual questionnaires (= bulletins) be used. It listed the essential variables to be included in the questionnaire: name and Christian name, sex, age, relation to the head of the family and the household, civil state or conjugal condition, profession or occupation, religious affiliation, language spoken, knowledge of reading and writing, origin, place of birth, and nationality, usual residence, nature of the residence, the place where the census is taken, and whether blind, deaf and dumb, cretin, idiot, or of unsound mind (Brown 1872, pp. 443 – 445). The recommendations were accepted, used, and remained more or less unchanged till 1950.

In 1958 the United Nations Statistical Office defined the population census as: “... the total process of collecting, compiling and publishing demographic, economic and social data pertaining, at a specified time or times, to all persons in a country or delimited territory.” It listed as essential features of “an official national census”: Governmental sponsorship, a defined territory, universality, counting every member of the community, simultaneity, the persons should be counted with reference to a “well-defined point of time”, individuality, information of each persons should be separately registered and the results to be compiled and published. (United Nations 1958, p. 3). In 1969 (United Nations 1969, p. 2) and 1980 (United Nations 1980, p.2) minor modifications were introduced. . In the 2015 revision of the *Principles and Recommendations for Population and Housing Censuses* two new elements are



added to the definition of the population census namely “planning” and the “capacity to produce small-area statistics”. The definition of the population census became: “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.” The list of essential elements now also includes the “capacity to produce small-area statistics. It is not known why the capacity to produce small area statistics has been included in the 2015 definition because one of the 1980 recommendations was: “Data from population censuses may be presented and analysed in terms of statistics on persons, . . . , and for a wide variety of geographical units ranging from the country as a whole to individual small localities or city blocks (United Nations 1980, p. 2).

In the Recommendations for the 2020 Censuses of population and Housing of the Conference of European Statisticians a new definition for population and housing censuses is introduced valid for the UNECE region only. The generic definition of the population and housing census is “the operation that produces at regular intervals the official counting (or benchmark) of the population [of all housing stock] in the territory of a country and in its smallest geographical sub-territories together with information on a selected number of demographic and social characteristics of the total population. [of characteristics of housing]” (United Nations 2015, pp. 5 – 6).

3. The census methodology

The census process is a complex operation covering a wide range of activities. Four key activities are of relevance: the cartography, the questionnaire construction, data collection and data evaluation and processing. The current practice of census taking in the majority of countries is based on direct data collection of the individual characteristics of each person through the households. To ensure complete coverage of the population within the territory the census cartography is used in which the whole country is sub-divided in interlocking, non-overlapping, geographical units of increasing smaller size (respecting administrative sub-divisions) in which all buildings, dwellings and households are identified. Ideally the smallest geographical unit is the census enumeration area (EA) or census block consisting of the number of households an enumerator could complete in one day. For each of the EAs a map indicating the key features of the area, the open spaces, each building, dwelling and household is drawn. In its simplest form it is a sketch map of the area, but at present due to enhanced mapping technology, these might be maps with geo-positioning of the identified features.

The variables to be covered in the census are contained in the census questionnaire. Ideally the content of the questionnaire is determined through a process of consultation with all stakeholders in the society. Questions should be carefully tested for validity and reliability before their use.

Traditionally individual households were visited by enumerators who obtained the information directly from the individual concerned, or the head of household or any responsible adult present. Currently, there are several options for the enumeration of the population:

1. The face-to-face personal interview, in which an interviewer visits each household and completes the questionnaire;
2. Self-completion, through the mail-out-mail-in procedure, in which the questionnaire is delivered to each household (address) and the occupants complete the questionnaire themselves;
3. The telephone interview in which the household is interviewed by telephone by an agent, and the details are entered by the agent in a specially designed data-base, and
4. The internet interview, in which the household completes an electronic version of the questionnaire in a specially designed data-base.

Each of these approaches has its advantages and disadvantages. Several modalities can be used in a census, and many countries use multiple enumeration modalities. However each modality has its own



requirements and when a combination of modalities is used it is necessary to establish how the validity and reliability for each will be assessed, and how the overall quality of the census will be established.

Manually collected data and self-completed questionnaires are processed at data processing centres. The data are transformed into electronic formats, using mostly automated electronic scanning techniques, and each question is verified for completeness, and correctness and consistency of the answers, using special data control techniques establishing their validity and reliability.

Increasingly enumerators use electronic hand-held data collection devices. These use specially developed software applications, which are also used by the telephone and internet approaches. The computer-assisted personal interviewing (CAPI) packages are able to verify the correctness of the answers to the main questions, producing quality data instantly, reducing the need for lengthy data management activities after the field operations. This could reduce data processing time and possibly also the total census cost. However, the packages do not contain a facility to record the original answers of the respondents, negating the possibility to assess the validity and reliability of the original answers. The verified data files are used to produce the statistical tables of the census reports.

4. Methodological and quality criteria

The census operations have to comply with the methodological requirements of validity and reliability to ensure that data of good quality are produced. To establish the validity and reliability of the variables a series of standard techniques are used. As the main objective of any census is the complete enumeration of the population, hence it is necessary to establish how well the census population reflects the true population of the country. Several procedures have been developed to assess the completeness, and under and -over- enumeration. The most versatile instrument is the Post-enumeration survey (PES) (United Nations 2007, pp. 60, 85-88, 94; United Nations 2010).

As the census has been defined by the United Nations as a total process a comprehensive system to ensure quality assurance is proposed. On the basis of “methodological soundness and adherence to professional methods and (internationally) agreed standards”; and “efficiency, the degree to which statistics are compiled in such a way that the cost and the respondent burden are minimized relative to output” it is suggested that the census output comply with some or all of eleven quality elements (United Nations 2015, pp. 73 – 84). The technical instruments for quality control of the census operations are well established. There is no international standard for quality assurance, but several countries have established national quality assurance systems for their statistical operations.

5. Census output

The results of the census are normally presented in census reports, and census atlases, showing key results in maps of the different geographical and administrative divisions of the country. Increasingly, statistical offices make depersonalised data files (of the total population or a sample) available on-line for further analysis or deposit such files with international data repositories, such as the IPUMS programmes of the Minnesota Population Center of the University of Minnesota, USA for use by the global research community (Minnesota Population Center; MacDonald 2016). In such cases metadata are required: a detailed and complete documentation on the methods, techniques, procedures, variables, questions, response categories and coding instructions used in the census.

In addition to these final products there are at least two intermediate products that are essential for the statistical research community and the society at large. The updated census cartography serves as the national master frame for sample surveys and other statistical activities such as agricultural or industrial censuses and could also be used for policy interventions and commercial activities. The absence of a cartography based national sample frame may hinder proper designs of surveys and other statistical data collection operations and may lead to unnecessary or inappropriate expenditures.



Similar use might be given to the updated inventories of localities, buildings, addresses and households with the relevant geographical identification information subject to restrictions of privacy legislation. The majority of statistical offices place the census documents and publications on their websites. The public availability of census products is restricted due to confidentiality clauses in the statistical and other legislation and whether the outputs have commercial value, in which case the Government may try to recuperate some of its expenditures.

6. Conclusions and challenges

The census is a powerful statistical data collection instrument that provides an inventory of all the elements of a population (persons, animals, non-organic materials, astronomical phenomena, etc.) and their characteristics that are simultaneous and individually enumerated within an area with reference to a specific time, and ideally with a certain periodicity. It based on well-established scientific principles; it is flexible in incorporating new technologies and is capable of measuring emerging phenomena and populations with new rare characteristics, and has well-established procedures to assess the quality and completeness of the information collected. Although there are global and regional recommendations on the methodologies and procedures to be used, there are no set standards for quality assessment, nor are the results of the different census operations systematically reviewed. Changes in the population dynamics, population growth and movements, complicate the execution of the census, including the type of population to be measured, which also may affect the correct use of the census reference date especially if extended periods of data collection with multiple data collection methods are used.

The essence and the integrity of the concept of the census are under attack, not because of scientific or philosophical deficiencies, but because of technological developments and political, financial and social considerations. The essential characteristics and functions of the different sources of statistical information, administrative records and derived administrative registers and the census are increasingly being ignored or confounded. Increasingly methodological principles and standards are being ignored, compromised or trivialised. The census is said to be a costly operation, but it is flexible and adaptable, has multiple outputs, its cost structure is transparent and comprehensive and expenditures are spread over several years. Its outputs, intermediate and final have multiple and sometimes unique applications. No comparative cost analysis system exists yet for its alternatives.

The expansion of the electronic media has created a large number of commercial entities that collect, archive and process detailed information on individuals, their behaviour, transactions and personal opinions which, some fear, could compete with and supplant official statistics including census data.

Yet, as long as there is a need for verifiable and detailed statistical information on population groups, at different levels of the national territory, and as long as the completeness of registers needs to be established or verified the census will be required.

References

Brown, S. (1872). Report on the Eighth International Statistical Congress, Held at St. Petersburg, 22nd/10th August to 29th/17th, 1872. *Journal of the Statistical Society of London*, 35, 4, 431-457. DOI: 10.2307/2338764, Stable URL: <http://www.jstor.org/stable/2338764>

De Neumann Spallart, M.F.X. (1886) La fondation de l'Institut International de la Statistique: Aperçu historique. *Bulletin de l'Institut International de la Statistique*, I (1-2):1 – 34. Available at <http://gallica.bnf.fr/ark:/12148/bpt6k615433?rk=42918;4>

Levi, L. (1854). Resumé of the Statistical Congress, held at Brussels, September 11th, 1853, for the purpose of introducing unity in the Statistical Documents of all Countries. *Quarterly Journal of the Statistical Society*. 17, I, 1 - 14. DOI: 10.2307/2338350.



Stable URL: <http://www.jstor.org/stable/2338350/>.

MacDonald, A.L. (2016). IPUMS International: A review and future prospects of a unique global statistical cooperation programme. *Statistical Journal of the IAOS*. 32, 715–727. DOI 10.3233/SJI-161022 IOS Press.

Minnesota Population Center, University of Minnesota. *IPUMS*. Available at: <https://www.ipums.org/>

Sulte, B. (1882) *Histoire Canadiens-Français 1608-1880 : Origine, histoire, religion, guerres, découvertes, colonisation, coutume*, Tome IV, Montréal, Wilson & Cie. Available at:

<https://archive.org/details/HistoireDesCanadiens-franais4>.

United Nations, (1958). *Principles and recommendations for national population censuses*. *Statistical Papers Series M No.27*. (United Nations Publication Sales No.: 58. XVII. 8).

United Nations. (1969). *Principles and recommendations for the 1970 population censuses*. Second printing with changes of a non-substantial nature, *Statistical Papers Series M No.44*, New York, (United Nations Publication Sales No.: E.67. XVII. 3).

United Nations. (1980), *Principles and recommendations for the population and housing censuses*. *Statistical Papers Series M No.67*. (United Nations Publication Sales No E.SO.XVII.S 3).

United Nations, (2007) *Principles and Recommendations for Population and Housing Censuses, Revision 2*. *Statistical Papers Series M No.67/Rev. 2*, (United Nations Publication Sales No E.07.XVII.8).

United Nations. (2010). *Post Enumeration Surveys: Operational guidelines. Technical Report*. New York. Department of Economic and Social Affairs. Statistics Division.

United Nations. (2015). *Principles and recommendations for the population and housing censuses. Revision 3*. *Statistical Papers Series M No.67/Rev.3*. New York. (United Nations Publication Sales No. E.15.XVII.10).

United Nations Economic Commission for Europe. (2007) *Register-based statistics in the Nordic countries. Review of best practices with focus on population and social statistics*. (United Nations Publication Sales No. E.07.II.E.11).

United Nations Economic Commission for Europe, Conference of European Statisticians. (2015). *Recommendations for the 2020 Censuses of population and Housing*. United Nations, New York and Geneva. United Nations.

UNESCO, International Memory of the World Register. *The 1703 Census of Iceland*. Available at: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/iceland_1703_census.pdf

United Nations Statistical Commission. (2015). Forty-sixth session, *2010 and 2020 World Population and Housing Census Programmes, Report of the Secretary-General, E/CN.3/2015/6: 2-3*. Available at <http://unstats.un.org/unsd/statcom/doc15/2015-6-Censuses-E.pdf>.

United Nations Statistics Division. (n.d.). Report on the Results of a Survey on Census Methods used by Countries in the 2010 Census Round: 9 – 11, Available at:

<http://unstats.un.org/unsd/demographic/sources/census/wphc/docs/ReportOnSurveyFor2010Census.pdf>