



Economic and Social Council

Distr.: General
21 April 2010

Original: English

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Population and Housing Censuses

Thirteenth Meeting

Geneva, 7-9 July 2010

Item 2 of the provisional agenda

Census communication and dissemination, including the use of Geographic Information Systems (GIS)

Census results tailored to user needs

Note by the National Statistical Institute of Portugal

Summary

This document presents the development of census products in Portugal since 1981 and the dissemination program for 2011. The Dissemination Program of Census 2011 includes a framework of products and services promoting dissemination via the Web, and some printed publication.

1. The Bureau of the Conference of European Statisticians (CES), at its meeting held in Washington, D.C. (United States) on 19-20 October 2006, approved the renewed terms of reference for the Steering Group on Population and Housing Censuses and the plan for future CES activities on population and housing censuses. The CES Bureau also agreed that the Steering Group would coordinate the work on the diverse types of meetings.
2. The present document was prepared on request by the Steering Group on Population and Housing Censuses, for presentation and discussion at the Thirteenth Meeting of the Group of Experts on Population and Housing Censuses, in Geneva on 7-9 July 2010.

I. Introduction

3. The production of statistical information only makes sense to the precise extent of their usefulness to society in general and for the analysis and evaluation of policies, in particular. The knowledge of the society, in the components of population and housing, is one of the oldest statistical activities of mankind and this is sufficient to demonstrate concern for statistical information, particularly in relation to population.
4. Developments in technology went from a census model based in enumeration rules with relatively vague criteria for more accurate and demanding models, not only as to the enumeration of statistical units, but also in terms of content and quality of information produced. Users are increasingly demanding both in quantity and quality of census information to provide.
5. Thus, a census only truly exists when the results are available to users and respond to their main concerns with the knowledge of this statistical area. The United Nations Economic Commission for Europe (UNECE) international Recommendations for Population and Housing Censuses (2006) are quite explicit in the emphasis they place on the importance of dissemination of results in saying that: *"A census is not complete until the information collected is made available to users in a form, and to a timetable, suited to their needs. Thus in disseminating the results of the Census much emphasis should be put on responsiveness to users and on high standards of quality in the production of statistics. Census results should be disseminated simultaneously to all users, and the greatest care should be exercised to avoid the inadvertent disclosure of information about identifiable individuals"*
6. Other determinant features of census information are: the linkage of alphanumeric information to geographic information, harmonization of methods of enumeration and content of information among countries and metadata.
7. The census information will also be increasingly geographical information, given the large capacity of integration of Geographic Information Systems (GIS). Therefore when we want to retrieve data about a particular region, administrative or statistical area, the combination of geography to the alphanumeric information has a significant added value, thanks to the "preview" of localization and data comparison to neighbourhood and other areas.
8. The harmonization of methods of enumeration and content is a growing concern internationally, both by the global and increasingly integrated world in which we live, as by the unavoidable comparisons describing the international position of each country in the closest set of countries to which it belongs. Globalization also applies to statistical information, with particular emphasis on the census project IPUMS (Integrated Public Use Microdata Series International) led by the University of Minnesota which is a good example of "census globalization".

9. Finally, in order to have "census globalization", there must be also appropriate and clear metadata on the content of information provided. Given the diversity of situations, the recommendations of the United Nations were not sufficient to ensure the harmonization of procedures and results in the European Union (EU). A framework regulation and three implementing regulations helped to build a model of harmonization of content, dissemination and assessment of the quality of census information in 2011, which undoubtedly will make these results more understandable and quite comparable in the EU. Portugal should have no significant problem to adapt to these regulations, once appropriate mechanisms were already introduced in its program for 2011 in order to produce information aligned with these regulations and to ensure comparability with previous census series.

10. In this paper we address the development of census products in Portugal since 1981 and present the dissemination program for 2011.

II. The development of census products in Portugal since 1981

11. The dissemination and availability of census results in Portugal, since 1981, has developed significantly with the purpose of responding as much as possible to user needs. This development focused on the three main components of the dissemination of results: (i) increasing of the number of results both by amount of cells of information available and by the greater geographical breakdown of results; (ii) systematic anticipation of the dissemination calendar and (iii) product structuring in order to respond to the current growing and diversifying demand.

A. Tailoring products to user needs

12. User needs are increasingly demanding and products usually available were not always adequate to the information requested. Moreover, processing capacities in previous censuses were quite limited and therefore it was tried to significantly increase the data available through the regular tabulation program in order to respond simultaneously to the widest possible range of users. This procedure has significantly increased the number of tables and information cells as shown in the table below:

Table 1
Number of information cells released by census year

Census year and n° of cells	1981	1991	2001	2011 (estimate)
Number of information cells released at the time of final results (millions)	20,6	314,9	613,2	648,3

13. Until 2001, the content of publications was organized mostly regarding the dissemination of results in tables and only one publication for the metadata. Since 2001 we have also been designing publications with information on the quality of coverage of statistical units and a large chapter with key indicators, graphic and data analysis, designed primarily to highlight the main variations in the results regarding previous censuses. Within this analysis it was deliberately absent any interpretation of the causes concerning these variations.

B. Structuring the results

14. The census results are awaited with great expectation by all the statistical users whatever their area of expertise. In fact, in addition to the large amount of information produced and to its fine geographical and administrative breakdown, the census results have a strong influence on other statistical indicators which rely heavily on estimates of population and housing. Every household and housing surveys use samples that are built on the latest census results.

15. Thus, since 1981, Statistics Portugal started to structure the census results in 3 phases: preliminary, provisional and final. This sequence of results takes into account that data processing is lengthy and complex and aims to empower users with the latest but sufficiently robust information as it is being produced and avoiding significant differences between the results of the various phases. The final results obviously prevail over all the previous.

16. For 2001 we have released a separate publication with the results on quality assessment of coverage and content, as well as their metadata.

C. The integration of the geographical component (from analogical to digital cartography)

17. The linkage between geographic information and statistical information is a powerful tool to aid understanding and analysis of information released. Most likely, in the near future all the statistical information must also be geographical, or risk losing an important component of their use. Since 1981, the census activities have massively boosted the preparation of a geographic infrastructure supporting the activities of data collection and dissemination of census results. A geographical basis aligned with the division of each basic administrative unit (civil parish) in statistical sections and subsections has been set up since 1991.

18. In 2001 all this infrastructure has been supported on digital cartography and prepared to allow comparison of results from similar geographic areas in the various censuses since 1991.

Table 2

Administrative and statistical structure units

<i>Administrative and statistical units</i>		<i>1981</i>	<i>1991</i>	<i>2001</i>	<i>2011</i>
<i>Unit</i>	<i>Type of unit</i>				
Portugal ^a	Administrative and statistical	1	1	1	1
NUTS I		3	3	3	3
NUTS II		7	7	7	7
NUTS III		30	30	30	30
Municipality		305	305	308	308
Civil Parish		4050	4208	4241	4260
Statistical section	Statistical	36795 ^b	13705	16094	19000 ^c
Statistical subsection			106625	177893	400000 ^c

^a Surface area: 92.207 Km²

^b In 1981 there was no distinction between statistical section and subsection

^c Estimates

NUTS – Nomenclature of Territorial Units (harmonized EU regions)

19. For 2011, in addition to continuing the process of comparing the census results using geography statistics, we will also make the geocoding for each building, which will allow releasing the census information on the basis of geo-referenced coordinate system.

C. Fastening the delivery of final results

20. The freshness of statistical information is a very important element for most users. In these censuses information freshness is even more important because users do not have equivalent information for a long time (10 years for most countries using the traditional census model) and many of these results are only available from censuses.

21. Thus Portugal bet too heavily in anticipation of systematic census results not only in earnings of calendar, but also in using the internet since 2001, according to the following:

(a) For 1981, the final results were available in June 1984, with the census day on 16th March 1981;

(b) For 1991, the final results were released in October 1993, although the existence of a data error in one municipality has forced to make a new edition in 1996; the census day was on 15th April 1991;

(c) For 2001, the final results were made available in October 2002 via the Internet and publications, with the census day on 12th March 2001.

22. Getting the final census results within 19 months after the census day, reached in 2001, and interim information during this period, looks therefore like a good enough and sufficiently cautious "performance", given its requirements of analysis.

D. The integration of new technologies and products in response to the most demanding users

23. Besides the use of GIS technology, the Portuguese census results have sought to integrate other technology tools in order to reply to the most demanding users. Thus, all pre-formatted tables of 1991 and 2001 were made available in digital format on the Internet and later all the final results from 2001 were converted into indicators for using in self-tabulation.

24. In 2001 we also have produced a CD ROM which provided results for 2001 and 1991 with various features and Statistics Portugal has participated in the IPUMS project with a sample of the Censuses 1981, 1991 and 2001. Every census publications were available at the website of Statistics Portugal in pdf format.

III. The Dissemination Program for 2011 Census

25. One of the first tasks to perform in the preparation of the Census is to draft an Action Program, which provides an integrated view of all activities contributing to the implementation of census operations. Alongside the Action Plan, the Program for Dissemination of census results is prepared. These documents aim to respond to the needs of different users.

26. Statistical users are represented at the SEAC 2011 (Ad-hoc Section for Monitoring the Census 2011), a section of the Statistical Council (CSE), providing a forum for discussion, with the role of regulation, advice and guidance of the whole Census 2011

activity, from preparation to the dissemination of results. The SEAC 2011 has permanent representatives from 12 entities, including Statistics Portugal, various ministries, associations and confederations of Employers, Trade Unions and Consumers, the Governments of Autonomous Regions and the Council of Rectors of Portuguese Universities. These ensure a broad base of knowledge and society needs in the census area. The Dissemination Program for Census 2011 got a favourable opinion of this section in February 2009.

27. The program and the products for dissemination of Census 2011 are organized within 3 major purposes:

(a) Easier access to census information – a clear bet in the use of new communication and information technologies in data dissemination, responding to requests from users who want to handle the information directly through the use of dynamic systems, simple and flexible;

(b) Increase the supply of dissemination products - providing a wide range of dissemination products in which, in addition to the traditional tables and publications, will be possible to extract and organize the data according to the preferences and user needs. It will also be possible to release geographical information as well as a sample of micro data.

(c) Fast availability of information - providing the results of Census 2011 in three phases: preliminary, provisional and final results;

28. The Dissemination Program of Census 2011 includes a framework of products and services promoting dissemination via the Web.

29. The Internet being the main means of dissemination of census results, Statistics Portugal's website has available, since March 2009, a specific sub-site to the Census. This web area is the repository of all information about the Censuses carried out up to now and acts as the main means of communication with citizens in the double role of information provider and user of statistics. In the 2011 Census area it is possible to follow the work of the project, with monitoring of all steps before, during and after the census activities.

30. The results of the census are always eagerly awaited by users, in general. It is considered that the fast dissemination of census results should be a central objective of the team in charge of the Census 2011 preparation, organization and tabulation.

31. As it happens in previous censuses, it is expected that the availability of the results of the 2011 Census is conducted in three phases. The first results, designated as preliminary, will be announced in the 1st phase. The provisional results will be released in the 2nd phase. Finally, in the 3rd phase, the final results will be presented replacing all the previous results.

A. Preliminary Results

32. Preliminary results are primarily intended to offer a preview of the evolution of some indicators, compared to the 2001 Census. These data are produced on the basis of data coming from the fieldwork control system.

33. These results are expected to be made available until 4 months after the census day.

34. Preliminary results will be based on counts of statistical units - building, housing unit, household and person - made during the data collection phase, broken down to the level of statistical subsection.

35. The release of preliminary results will be mainly carried out through the Web.

B. Provisional Results

36. Provisional results will be available after the first steps on questionnaire scanning and validation of data entry. The results will be presented in a format that allows comparative analysis with the results of 2001.

37. These results are expected to be made available until 11 months after the census day.

38. The provisional results cover some of the main variables observed in each primary statistical unit - building, housing, household and individual - broken down by municipality or civil parish, since this last has 1,000 or more resident persons, in order to avoid errors for small populations once data processing and analysis is not yet finished.

39. The release of provisional results will be mainly carried out through the Web.

C. Final Results

40. The final results will be available after the processing and analysis of all data is fully completed.

41. These results are expected to be made available until the last quarter of 2012. The following products and services for dissemination will be developed from the database of Census 2011:

1. Publications

42. For some users, the publications are still a privileged means of access to statistical information. It is foreseen the compilation of the most relevant information from the Census 2011, broken down by municipality, including a publication for each NUTS II and one for Portugal, accounting for a total of 8 publications. In addition to the traditional tables, organized according to the content of predefined tables (3.3.2), these publications will include statistical analysis, graphs and maps, in order to facilitate understanding of the data.

43. It is also expected to issue two other publications in bilingual versions (English and Portuguese):

(a) Precedents, Methodology and Concepts – constituting an aid to understanding the metadata used;

(b) Post Enumeration Survey – including the methodology and results of this survey; this takes place, immediately after the data collection phase, in each pre-selected area. These results provide the final quality assessment on coverage and content of 2011 census results.

44. The dissemination of publications will be made through the Web with a limited run on paper.

2. Predefined tables

45. A set of predefined tables (122) will be provided in order to facilitate access by users unfamiliar with the use of new technological tools. The content of these tables was assessed in terms of preparing those with higher demand in the recent past.

46. These tables, broken down by civil parish, are organized by topics of each statistical unit: resume (some very few tables that include variables from every statistical unit), variables of buildings, housing units, households, families and persons.

47. The dissemination of tables will be made through the Web and on DVD.

3. Census 2011 Facts and Figures

48. The information available will contain not only data from the Census 2011 but also from 1991 and 2001.

49. The information is organized according to two hierarchical territorial classifications: administrative structure and nomenclature of territorial units for statistical purposes.

50. The product will present a set of features organized into 5 major groups:

(a) Access to pre-defined tables including comparative tables with the previous censuses;

(b) Construction of statistical tables by selecting variables and geographical units, up to the civil parish. It will be possible simultaneously to access to the metadata associated with the selected variables as well as to export and manipulate the results;

(c) Construction of graphs and maps, after the selection of indicators and geographical units. It will be possible simultaneously to access to the metadata associated with selected variables, as well as to export and manipulate the results;

(d) Metadata Repository to support the understanding of the data. It will include the methodologies, questionnaires, concepts and classifications used in Census 2011 and earlier;

(e) Informative Content - Global information about what are the Censuses, main uses, who responds, why people are asked for, how they are carried out, etc.. It includes documents, images and videos associated with the 2011 Census works, before, during and after the end of the census activities.

51. The purpose of this product is to continue the 2001 edition of the CD ROM "Population and Housing Censuses - Comparative data 1991-2001". This product has been in great demand, especially by students and teachers of secondary and higher education. The release of the 2011 Census in Figures will be made on DVD.

4. CensusStat

52. Continuing the logic underlying the model of the Portal of Statistics Portugal, the CensusStat provide the following features:

(a) Construction of tables crossing a certain number of variables (indicators) that follow a set of conditions, according to the geographical level selected; in principle, every content (statistical units and respective variables) of 2011 census will be converted into indicators available for self-tabulation.

(b) Totals of variables, allowing a graphical representation and thematic maps. Is it possible, in addition to the production of statistical tables, to see the metadata associated with the content of variables, to produce graphs with the result of the table and "mapping" the geographic area requested.

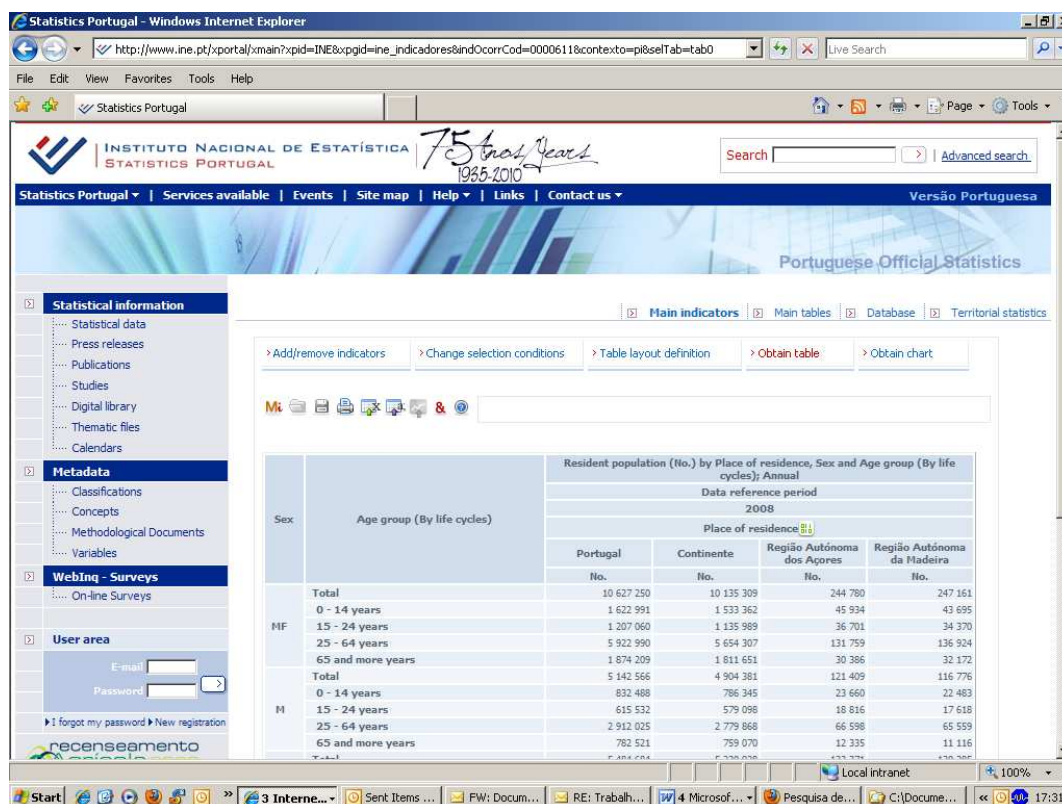
53. This service will also allow access to the hub hypercubes of Eurostat, containing the census results from other EU countries, according to the Community Regulation that has already been approved. The content will be restricted to mandatory variables and to a number of predefined rules.

54. The informational contents include geographic levels where it is possible to make tabulations or only totals for most of the variables: Portugal, NUTS I, NUTS II, NUTS III, municipality and civil parish. For the geographical levels of statistical section and

subsection it is only possible to have totals for a subset of variables that correspond to the Summary File, taking into account the protection of confidentiality.

55. Spreading CensusStat will be carried out mainly through the Web in bilingual version - Portuguese and English. Figure 1 shows the model interface for navigation and presentation of CensusStat results.

Figure 1
Model interface for CensusStat



5. CensusMap

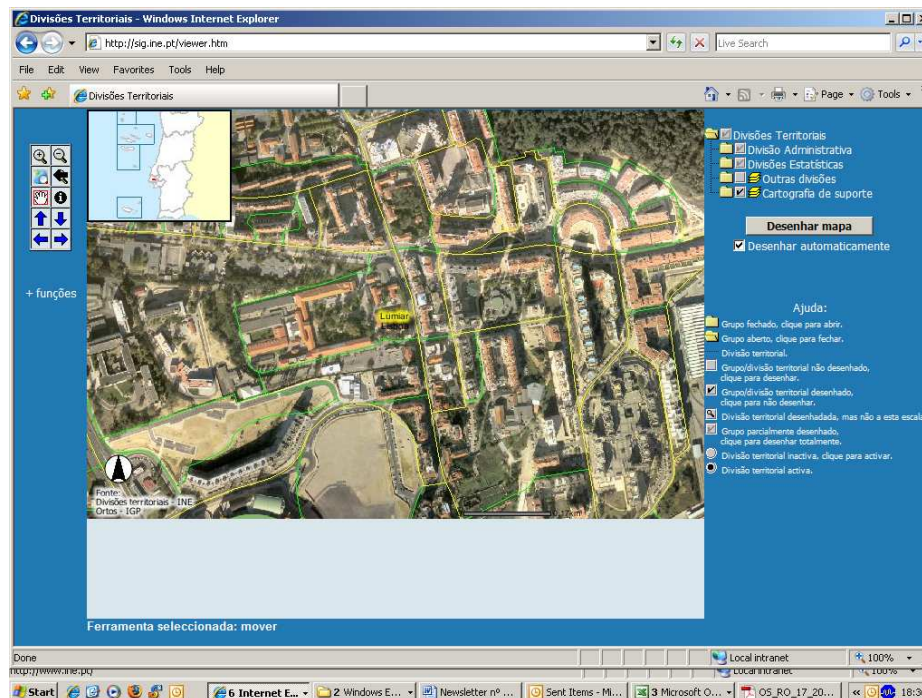
56. This service, to be available on the Web, will enable the user to georeference the data to the level of sub-section (Fig. 2).

57. Is it possible to compare the indicators with the Censuses 1981, 1991 and 2001 at the level of civil parish, statistical section and subsection, whenever geography allows it. The following products will be developed from the databases CensusMap:

(a) BGRI 2011, which corresponds to a digitized map of Census 2011 and includes the administrative boundaries as well as limits corresponding to the statistical sections and subsections;

(b) Summary File 2011 - Is made up of key totals for a set of 147 predefined variables, broken down to statistical subsection and taking into account the protection of confidentiality. It is possible to make the map referencing for these variables. This product continues the series started in 1981, assuring its comparison for the content together with the geography, at least down to the level of parish.

Figure 2
Standard interface for CensusMap



6. CensusMicro

58. The availability of this product is subject to the evaluation of each application, taking into account that its main users are mostly specialized users and researchers.

59. Its design is based on Statistics Portugal experience of participating in IPUMS, a worldwide project that has currently available micro-data from 130 censuses in 44 countries. Portugal provided data from the Censuses 1981, 1991 and 2001.

60. The use of micro data samples allows the user to make self-tabulation and specific analysis to the required level of detail and with a high degree of freedom, ensuring simultaneously the quality and confidentiality of data. The most effective procedures for anonymization, available at the time of the sampling preparation, will be used.

61. This product follows the international practice of dissemination for the 2011 Census and comprises a sample of micro data, consisting of 5% records of resident persons.

62. The maximum geographical breakdown depends on the protection of confidentiality. Spreading CensusMicro will be made through the Web and on DVD.

IV. Conclusions

63. The main conclusions are:

(a) The census results require a dissemination plan that takes into account all the profiles of users. This is due to the fact that censuses produce data whose range of users is the largest in all the statistical information; as such products tailored to each user's profile should be provided;

(b) The Internet and the public contribute to a growing demand for this information; therefore it is essential to adapt Census products to this demand, particularly through standardized models of self-tabulation;

(c) The statistical information must incorporate the models of geographical information in the form of GIS

(d) The dissemination program for 2011 will be oriented mainly to simplify access and increase the supply of products, favouring the dissemination via the Internet.
