



**Economic and Social  
Council**

Distr.  
GENERAL

ECE/CES/GE.41/2009/15  
19 August 2009

ENGLISH ONLY

---

**ECONOMIC COMMISSION FOR EUROPE**

**CONFERENCE OF EUROPEAN STATISTICIANS**

Group of Experts on Population and Housing Censuses

Twelfth Meeting  
Geneva, 28-30 October 2009  
Item 5 of the provisional agenda

**CENSUS OUTPUT TO MEET USERS NEEDS**

**The European Union dissemination of 2011 census results**

Note by the Statistical Office of the European Communities<sup>1</sup>

**I. TRANSPARENCY, QUALITY AND USEFUL DATA FOR EVERYBODY**

1. In 2008, the European Parliament and the Council adopted legislation<sup>2</sup> that, for the first time in history, provides a framework to produce and disseminate comparable census data for all European Union (EU) Member States. This information will be of great value to policy making, administration, businesses and the EU citizens. It will not only support EU-wide activities, but will make it easier for people acting nationally or even regionally to see their situation as part of the larger European picture.

2. Any census operation is a huge challenge to the National Statistical Institutes (NSI). It is influenced by national traditions and sensitivities, different data sources and resources. Each NSI

---

<sup>1</sup> This paper expresses the view of its author which are not necessarily those of the European Commission.

<sup>2</sup> Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing censuses, OJ 218, 13.8.2008, p.14.

strives hard to develop a census methodology and technology that those responsible for the national census can decide to be optimal in the context of their country. Developing census methodology is a complex process both on the political and technical levels. It involves a first hand "turf knowledge" of all parties involved.

3. The EU legislation on censuses is "output oriented". It does not prescribe a specific methodology how the Member States have to conduct their censuses. The responsibility to develop appropriate census methodology and technology remains with the Member States (no "input harmonisation"). However, the EU legislation aims at a dissemination of census data that are comparable between the EU Member States. The data have to:

- (a) follow transparent concepts, definitions and specifications;
- (b) be of sufficient quality and subject to a quality reporting;
- (c) follow a European programme of statistical data (hypercubes) and metadata, and be accessible via a user-friendly dissemination format.

4. To achieve this, the European Commission shall adopt different "implementing regulations".

## **II. TECHNICAL SPECIFICATIONS OF TOPICS AND BREAKDOWNS**

5. When this document was submitted (19 August 2009), the regulation on "the technical specifications of the topics and of their breakdowns" was about to be submitted for adoption. The technical specifications of the topics are a prime condition to achieving comparability. They are in line with those of the Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing (CES Recommendations) that were prepared by the United Nations Economic Commission for Europe (UNECE) in cooperation with the Statistical Office of the European Communities (Eurostat) and the UNECE Member States. The high quality of this document was an excellent starting point for EU-wide harmonisation. Nevertheless, the legal text goes further to provide concrete and comprehensive instructions wherever this is necessary. The breakdowns of the topics are defined in view of an efficient dissemination, but also taking into account aspects of data production raised by the NSI.

6. A special feature of this regulation is that for some topics it contains more than one breakdown, with the distinction being the level of detail. This was introduced to manage the size of the hypercubes (see below).

## **III. QUALITY REPORTING**

7. A regulation on the quality reporting on the census results is currently prepared by Eurostat, in cooperation with a task force consisting of experts from the Member States. The regulation is likely to comprise general information (e.g. history of population and housing censuses, legal background, responsible bodies) putting particular emphasis on the data sources

used to produce census data, and the census lifecycle. Thought is given to quantitative quality measurement. The objective is to report on phenomena like undercoverage, overcoverage, unit-non-response, imputation, and sampling errors. The methods of any quality related treatment of the data shall be described. The quality related information will be disseminated to the public.

#### **IV. PROGRAMME OF STATISTICAL DATA (HYPERCUBES)**

8. The following presents drafted legislation as of 19 August 2009 and might be subject to modification. The 2011 censuses will be a common achievement of the European Statistical System, and it shall lead to a common dissemination platform of census results for the European Union: A common entry point for all looking for regionally detailed data on the people living in the EU and their housing situation.

9. The EU dissemination of census results and the different national publications shall complement each other. Publications of census results on a national level report on all topics that the national census covers (these might be more than covered by the EU legislation). They will often have a structure, level of detail and focus that are of particular national interest or serve particular information requests of national, regional and local users (e.g. ministries, other administration). Each country has its traditions concerning the publication of census results. Often, NSI also provide complex, user-defined extractions of census data.

10. The EU-wide dissemination is limited to the topics (variables) that are covered by the EU legislation ('core topics' of the CES Recommendations). It will emphasise data useful to formulate and assess policies that are of high Community interest like employment and migration. Its main focus will be on NUTS 2 regions given that the EU has many policies for that regional level. Examples for EU policies that require comparable census data:

- (a) Census data help to maintain a high quality of annual population estimates. The democratic process within the European Union calls for annual population estimates of the highest possible quality. Article 11 (5) of the Rules of Procedure of the Council of the European Union defines the process for qualified majority voting within the Council. One criterion is that a qualified majority shall represent Member States comprising at least 62 per cent of the population of the Union;
- (b) The European Union has competency in the area of regional cohesion which is regularly one of the largest budgetary items. This requires population data on a regional level and of the spectrum needed to evaluate the progress of regional cohesion;
- (c) The European Union supports the development of its less prosperous regions. The Convergence objective of the Structural Funds is the main priority of the EU's cohesion policy. The areas lagging behind in their development are identified as those whose Gross Domestic Product per head is less than 75 per cent of the Community average. To establish which regions are eligible, high quality regional level population data are needed. The quality of regional population estimates

depends on a sufficiently detailed regional breakdown of the source on which these estimates are based;

- (d) Social and demographic developments pose a substantial challenge to the cohesion of societies and the sustainability of public finances. The societies in the European Union are subject to demographic ageing during the next decades: The share of the older generations is increasing while the share of the younger generation is decreasing. To formulate, conduct and evaluate related policies, benchmarking becomes ever more important – including benchmarking on the regional level;
- (e) Many European policies relate to areas that have boundaries different from those of the NUTS 2 regions. Examples are "Regional development", "Fisheries and Maritime Affairs" (coastal areas), "Tourism" (touristic areas), "Trans-European networks" (Transport and energy);
- (f) Many important sample surveys which are harmonized within the EU require statistical data on the structure of the population at a regional level to draw their samples and to extrapolate results (e.g. the Labour Force Surveys). The statistical frame has to be harmonized at the European level to ensure the comparability of the survey results;
- (g) The housing situation of the population has substantial implications. Access to decent housing is a major concern of social policy. Buildings that house private households are important consumers of energy, water, and other supplies and services that have an environmental impact. Moreover, housing has important economic significance (offer and demand for housing, finance, construction, refurbishment).

11. The EU dissemination of census results shall provide a systematic and sufficiently detailed overview over the census results, focussing on the ease of use and comparability of the data. It shall grant the user maximum flexibility while at the same time limiting the burden NSI might have take on in order to assemble and transmit the required data to Eurostat.

12. This consideration has lead to the definition of 'hypercubes' (multidimensional tables) that can be retrieved via an electronic dissemination platform. Within the scope of a given hypercube, a user can define his/her request flexibly through a system of breakdowns and categories that can be pre-selected. Such a flexibly defined user request will not entail any additional burden to the NSI: the NSI have to assemble the hypercubes only once.

13. Balancing flexibility of the usage against feasibility of production was a major challenge when the hypercubes were defined. It is clear that a hypercube which cross tabulates more topics (variables) in greater detail provides more flexibility than a comparable cube with fewer and more aggregated topics. On the other hand, the price for an increase in flexibility is a growing size of the hypercubes which can lead to problems concerning data production and the protection of statistical confidentiality. A break-up of a 'big' cube in several smaller cubes increases the total number of cubes in the dissemination programme.

14. Eurostat and the census task force managed the problem by different means:
- (a) Wherever reasonable, breakdowns representing different levels of detail were defined for the topics. For each hypercube containing the topic in question the breakdown with the most appropriate level of detail can be chosen;
  - (b) Consideration was given on the topics that were to be cross tabulated in each cube. The objective was to provide sufficiently detailed and useful information on each census topic, but also to keep the size of the cubes to what was considered feasible;
  - (c) Moreover, given that the legislation does not allow the transmission of confidential data, the size of the hypercubes had to be limited to avoid that too many of their cells are so thinly populated that they will have to be kept confidential;
  - (d) Finally, it had to be decided what size of a hypercube is to be considered as feasible in terms of data production. On this point, the differences of the census production systems in the Member States became very apparent. Some countries will be able to provide very detailed hypercubes, e.g. when they enumerate the population fully. Other countries have much stricter limitations, particularly in if one or more topics are the result of sampling.
15. A lot of detail work was done to merge the different aspects. The result is a system of 62 hypercubes. The overview list can be seen in the annex. It can be summarized:
- (a) NUTS 2 regions (46 hypercubes)
    - (i) Households (5 hypercubes, of which 2 with an employment or education related focus),
    - (ii) Families (5, of which 2 with an employment or education related focus),
    - (iii) Employment (12),
    - (iv) Migration (10),
    - (v) Housing (7),
    - (vi) Single ages (7) (here single ages are reported, whereas for the other cubes the Age is mostly reported for classes comprising five years);
  - (b) NUTS 3 regions (11);
  - (c) Municipalities (5).
16. Principally, Member States have to transmit to Eurostat data for all cells defined for a hypercube. However, cells outside a well defined set of "marginals within the hypercubes" can be blocked if NSI consider the data to be unreliable. The number of cells of the "marginals within the hypercubes" (i.e. the crossings for which data has to be transmitted) is limited to
- (a) 5000 cells for the of the 46 hypercubes defined for the NUTS 2 level;

- (b) 400 cells for the 11 hypercubes defined for the NUTS 3 level;
- (c) 50 cells for the 5 hypercubes defined for the municipality level.

## V. THE EUROPEAN UNION CENSUS HUB

17. The census data have to be transmitted to Eurostat in the SDMX format. How the Census Hub will be implemented depends on further development. Eurostat, in co-operation with a number of NSI, is currently working to develop the EU Census Hub. The objective is to provide the user with an attractive and performing entry point to retrieve the data from the hypercubes.

18. The Census Hub is a conceptually new system to manage the transmission and dissemination of the 2011 Census data. It is based on SDMX standards and the concept of data sharing, where a group of partners agree on providing access to their data according to standard processes, formats and technologies. The Hub will facilitate the data transmission between the NSI, Eurostat and the users of census data.

19. From the data management point of view, the Hub is based on agreed hypercubes, but the hypercubes are not sent to the central site. Instead, the following process operates:

- Step 1: the user defines a dataset and requests it to the Hub;
- Step 2: the Hub sends the resulting query to the NSI using a standard format;
- Step 3: NSI's systems process the query and send the result to the Hub in a standard format;
- Step 4: the Hub puts together the results and presents the dataset to the user.

20. This approach offers several advantages:

- (a) For the user, the Census Hub will allow for a high flexibility and at the same time respect that the NSI maintain complete control over their data. The NSI remain "proprietors" of their census data;
- (b) The NSI can keep the IT platform that they already use for their national purposes. They would not be obliged to send the data in another technical format to Eurostat;
- (c) In the case of revisions or updates the NSI would only have to upload the new data in their own system (something they do anyway) instead of sending a complete new data set to Eurostat;
- (d) NSI having a generic data dissemination system could use the same infrastructure developed for the census to offer other types of data to Eurostat and the outside world. The Census Hub allows for complete decoupling of the NSI's systems from the central one via standard formats and techniques for the exchange of data, metadata and queries.

21. How to participate to the pilot? Participation to the pilot is on a voluntary base.
- (a) As the pilot phase has been planned to keep things as simple as possible, it lets the NSI participate with minimum effort;
  - (b) The project is a good opportunity for all those who want to start using SDMX;
  - (c) Some software, developed during the first phase of the project as open source packages, is available and can be used immediately;
  - (d) The earlier NSI start the easier it will be.

22. On CIRCA a public folder is available with technical documents, presentations and software produced in 2008 and 2009. [http://circa.europa.eu/Members/irc/dsis/x-dis-xensus-hub/library?l=census\\_documents\\_1](http://circa.europa.eu/Members/irc/dsis/x-dis-xensus-hub/library?l=census_documents_1)

## **VI. HYPERCUBES OF THE EUROPEAN UNION DISSEMINATION PROGRAMME FOR 2011 CENSUSES**

The list presents drafted legislation as of 19 August 2009 and might be subject to modification.

### **Households**

#### **People in households**

1. Marital status
2. Education
3. Current activity status, Status in employment
4. Size of locality

#### **Private households**

5. Type, Size, Tenure status

### **Families**

#### **People in families**

6. Marital status
7. Education
8. Current activity status, Status in employment
9. Size of locality

#### **Families**

10. Type, Size

### **Employment**

#### **People where they live**

11. Occupation, Industry
12. Status in employment
13. Education
14. Country of birth, Citizenship
15. Residence one year prior to census

16. Marital status

**People where they work**

17. Occupation, Industry

18. Status in employment

19. Education

20. Country of birth, Citizenship

**People working abroad**

21. Occupation, Industry

22. Country of birth, Citizenship

**Migration**

**Arrival and country**

23. Year of arrival in country, Country of birth, Citizenship, Size of locality

24. Year of arrival in country (detailed), Size of locality

25. National level: Country of birth (detailed), Year of arrival in country

26. National level: Citizenship (detailed), Year of arrival in country

27. National level: Country of birth (detailed), Citizenship

**Employment**

28. Occupation

29. Industry

30. Status in employment

31. Education

32. National level: Occupation, Education

**Housing**

**Population in dwellings and collective living quarters**

33. Current activity status, Residence one year prior to census

34. Size of locality, Residence one year prior to census

**Homeless and Persons in unconventional shelter**

35. Sex, Age, Size of locality

**Living quarters**

36. Type of living quarter

**Dwellings**

37. All dwellings: Occupancy, Age of building

38. Occupied dwellings: Size, Density

39. Occupied dwellings: Water supply system, Toilet facilities, Bathing facilities, Type of heating

**Single ages**

**Population overview**

40. People in households, People in families

41. Current activity status

42. Occupation

43. Industry

44. Status in employment, Education

45. Size of locality

**Country**

46. Country of birth, Citizenship

**NUTS 3 Regions****Population overview**

47. Country of birth, Citizenship, Residence one year prior to census,  
Marital status

**Households**

48. People in households (Household status), Marital status  
49. People in households (Detailed household status)  
50. Private households

**Families**

51. People in families (Family status), Marital status  
52. People in families (Detailed family status)  
53. Families

**Housing**

54. Living quarters  
55. All dwellings: Occupancy, Age of building  
56. Occupied dwellings: Size, Density

**Single ages**

57. Sex, Age

**Municipalities****Population overview**

58. Sex, Age

**Households**

59. Private households

**Families**

60. Families

**Housing**

61. Living quarters  
62. Conventional dwelling.

\*\*\*\*\*