

**Economic and Social Council**

Distr.: General
13 March 2012

Original: English

Economic Commission for Europe**Conference of European Statisticians****Group of Experts on Population and Housing Censuses****Fourteenth Meeting**

Geneva, 24-25 May 2012

Item 4 of the provisional agenda

Address/dwelling listing**Experience with address lists and other preliminary work****Note by Istat, Italy***Summary*

This paper presents methods and technological tools developed by Istat in connection with the 2011 census, for the construction of a database of street lists integrated with house numbers geo-coded to enumeration areas, by the use of different sources. This strategy provides differentiated arrangements for the acquisition and processing of data, depending on the different demographic classes of the municipalities.

I. INTRODUCTION

1. This document has been prepared by Damiano Abbatini and Francesco Di Pede of the Italian National Institute of Statistics (Istat).
2. In the context of preliminary activities to general census, Istat proceeds to the definition of the Municipal Territorial Bases (Basi Territoriali Comunali - BTC) and uses the local street list as support to surveys. The BTC include municipal administrative boundaries, urban localities and the enumeration area that are referred collected data. In previous Censuses have been used as an aid to census operations "road segment", lists of street included in enumeration areas with the references of house numbering (first-last house number of road segments. Therefore were not built complete database of addresses.
3. In the new 2010-2011 census strategy, Istat has developed and tested methods and technological tools for the construction of a database of street list integrated with house

numbers geocoded to enumeration areas, by the use of different sources. This strategy provides differentiated arrangements for the acquisition and processing of data, depending on the different demographic classes of the municipalities.

4. The capital municipalities and municipalities with more than 20 thousands inhabitants, concerned by the first processing stage (late 2010), have made, for the major urban localities: i) a control on the external access which connect directly or indirectly to housing units; ii) census of buildings. The remaining municipalities have provided their street list accompanied by information concerning the house numbering found in its administrative files, but have not made a fieldwork.

II. Experimentation on municipal addresses list

5. In the preparation period of the census, while planning the National data base of house numbers, a feasibility study of the project was conducted by analyzing the existing local data base in order to evaluate the development of new methods of merging information deriving from Istat and Poste Italiane (Post) and also from commercial origin (e.g. Tele-Atlas road net). Particularly Istat carried on recognition and testing activities about the features of the existing data bases at local municipality level.

6. Analyzing existing local data bases in order to include them in national ones achieved from private companies or public institutions, it has been necessary to evaluate their features. A number of bilateral agreement with the municipalities were then concluded in order to exchange and process data. These activities were developed during 2006 and 2007. Agreement of two main forms were established:

- Case 1 – Public Institutions provided with a data base of house numbers geolocated and geo-coded to enumeration areas, (Municipalities of Milano, Cuneo, Verona Bologna, Firenze e Pesaro);
- Case 2 - Public Institutions provided with a data base of house numbers without geo-location or not geo-coded to enumeration areas, or Public Institutions without any house number data base (Province of Pesaro, Municipalities of Palermo and Trapani).

7. With the occasion it has been possible to anticipate and put the best accuracy in the fulfilment to Regulation of the naming roads ordering and the house numbering, compulsory fulfilment for the municipalities in preparation of the census operations. The agreement subjects and the process defined by the experimentation project plan were put in a background where a number of operations were ongoing by local institutions aiming at building comprehensive data bases of house numbers, centralized and geo-located, because of their usefulness either in the management of the territory and in controlling tax fluxes, tariffs and local tributes.

III. Survey on municipal street and addresses lists

8. The mere evaluation of qualities of house numbers databases owned by a few municipalities was not enough to answer about their existence and quality on a National scale. From June to September 2007 Istat performed a survey to enumerate and describe the peculiarities of the municipal data bases of house numbers and sub municipal areas (“Rilevazione sulla presenza e le caratteristiche di archivi comunali di numeri civici e zone sub comunali”). The survey was performed on the provincial capitals and all municipalities with more than 10 thousand inhabitants at the date of 1 January 2006 (i.e. 1166 municipalities). The question form, in a web format and answered by the reference persons

of the 1166 municipalities, put in evidence a general lack of spatial information in the databases (addresses were not geo-reference or geo-coded).

9. To the form, made of six sections (general information, street list, road graph, addresses list, geo-coding and geo-location of house numbers, sub municipal zones) 1083 municipalities answered, namely the 93 per cent of total. Analyzing results was demonstrated that about an half of the municipality maintained a data base of streets or addresses in a digital form. We noticed that there are strong differences among municipalities in separate geographic location or in different demographic classes: the South, the Islands and the municipalities with less than 50 thousands inhabitants were in late with the digitalization of their database. Only 15 per cent of municipalities had addresses geo-coded to the enumeration areas, while 24 per cent of municipalities had the addresses lists partially or totally geo-located.

10. Aiming at a more accurate evaluation and quantization of activities included in the Istat project to establish a National data base of house numbers, we considered that:

- In About 44 per cent of municipalities was necessary to build a digital addresses database [27 per cent of capital municipalities and municipalities with more than 50 thousands inhabitants (type A) and 47 per cent of municipalities with a population between 10,000 and 50,000 inhabitants (type B)];
- In 49 per cent of municipalities a digital list of addresses was built (32 per cent in type A municipalities and 52 per cent in type B ones), while another 29 per cent of municipalities needed a completion of the existing data base (23 per cent among type A municipalities and 30 per cent among type B ones).
- In 84 per cent of municipalities was necessary to quantify the number of addresses from other data bases (e.g. Poste Italiane and/or field surveys).
- In 85 per cent of cases the use of house numbers was not known.
- In 83 per cent of municipalities was necessary to geo-code the house numbers to enumeration areas (62 per cent of type A municipalities and 89 per cent of type B ones).
- In about 76 per cent of municipalities was necessary to totally geo-locate the addresses data base (51 per cent of type A municipalities and 80 per cent of type B ones), while in 6 per cent of municipalities the geolocation of addresses in the database was to be completed.

IV. Calculations on municipal street maps and address list

11. The data received by the municipalities undergone a normalization process and a process that allowed the creation of a database of standardized address.

12. Three lists of municipalities were identified according to three different types of archive held: municipalities with electronic address list, geo-referenced or not, municipalities with street-name electronic list, municipalities with paper archive. For the various cases, a web environment was designed for secure data interchange. The paper provided by the municipalities have been previously digitized and made available via web by Istat.

13. Since the information received was heterogeneous, in the first phase, the archives of house numbers were cleared before they can be included in the Operational Data Base (ODB). The development involved the elimination of duplicate or blank records and the identification of qualifying attributes. Once prepared the data, the standardization of archive

was started up. The operation consisted in the recognition and association of fields to standard information, such as Istat codes, street type, street name, house number, exponent and the creation of an ID for each record. Afterwards, data were subjected to normalization through a hierarchical recognition system applied to the following entities: Municipalities/Localities, street names, streets names in the Municipalities/Localities.

14. At the end of the normalization of the archives, addresses were geo-coded and geo-referenced. For each address in the archive of house numbers, and by using the street vector graph and the coverage of the enumeration area of Territorial Bases, the spatial coordinates were calculated in the coordinate projection system previously fixed, and the corresponding code area. The final product was the creation of an accurate cover of addresses for each municipality with the appropriate information for geo-reference and geo-coding. The generation of the representative point of an address is done using the function of address-matching of the GIS system used in Istat. The exact correspondence address-area, obtained thanks to the geometrical properties of inclusion of a point (the geo-referenced address) in a census area, and GIS functions which analyze these overlay, to produce automatically the street list in a enumeration census area. In this way, the points located on the territory were coded (geo-coded) with reference to the census2011. Later, when the design of the Territorial Bases will change in function of the updates to be applied due to changes occurred in the urban territory, the process can be iterated easily, and the integration of the information contained in the two informative levels, one geographic and one alphanumeric, will be guaranteed.

V. House Number Survey

15. House Number Survey (in italian: RNC - Rilevazione dei numeri civici) is a necessary tool to achieve some of the most important innovations in method for conducting fifteenth Census of Population and Housing. 2011 census strategy, in fact, introduced important innovations to reduce the intervention of the census enumerators to families and, in general, the overall burden in local census structure. In this context of technical and organizational innovations, addresses and house numbers are instruments of great importance, as indispensable for the accurate localization of statistical units.

16. RNC had three objectives: to produce, before the census, a database of house numbers validated by each municipality and geo-coded to the enumeration areas census; allow construction of a list to facilitate the recovery of municipal population registers under-coverage; collecting information on building stock and characteristics, thus anticipating a substantial part of the field work traditionally carried out simultaneously in the census of population and housing.

17. Thanks to RNC each municipality could have an address list updated and shared with ISTAT, a prerequisite for the success of the census of population and housing. Within the same RNC, the municipalities have also detected information on buildings associated with the house numbers under investigation. The final database house numbers has been the basis for the addresses list used by census enumerators.

18. The RNC has also made it possible to establish the relationship between buildings and addresses. Building-address relationship is a rather complex relation, like that “many-to-many” relationship. In fact, in some situations one or more addresses may correspond to a building, while in others the opposite occurs, to one address match more buildings.

19. In order to obtain valid signals of municipal population register under-coverage, for each address has been detected the information on the number of flats, articulate into “residential” and “nonresidential”. This information was then compared with the number of households recorded in the municipal population register at the same address. The

difference between the number of dwellings and the number of households was assumed to signal the presence of unoccupied dwellings or households not registered in the municipal population register.

20. To improve the timeliness of the information production and to facilitate the task of enumerators, Istat provided by informatics to each municipality, the temporary store of house numbers normalized, geo-coded to the enumeration areas date to 2010 and accompanied by additional information derived from different sources. Moreover, to facilitate field operations, Istat has provided maps of each enumeration area composed of several information layers: municipal administrative boundaries, position of the buildings resulting in the national cadastre, road net and aerial photography.

21. Enumerators had to: check the correspondence between the addresses of the list and those on the street; integrate list with any addresses not in the list and geocode them to Enumeration Area; collect information on number of housing units relative to each address; collect information on the buildings characteristics.

22. Field operations were conducted by enumerators selected and trained by each municipality. To collect data each enumerator used two paper enumeration forms, one to recording information on house numbers (Mod. Istat RNC.1), pre-populated with information given by the municipal street list or the address list) and the other one to collect information about buildings (Mod. Istat EDI.1).

23. After the field work, the data were recorded on a website (RNC Portal), which also served to manage all activities for exchange of documentation and monitoring field operations. Data submitted by municipalities were subject to control and validation procedures, and subsequently included in the fifteenth General Census of Population and Housing Survey Management System (in italian: Sistema di Gestione della Rilevazione – SGR). The information of the address list will be indispensable for the localization of the statistical units in the area and quality control of census operations.

24. The RNC has affected all municipalities with at least 20 thousand inhabitants on 1st January 2008 and the capital municipalities, totaling of 508 municipalities which cover approximately 53 per cent of Italian population. Fieldwork took place from 15 November 2010 to 28 March 2011, with a diversified calendar for groups of municipalities.

25. RNC has collected data for approximately 4.3 million buildings and 9.6 million addresses. Through the data of the RNC was possible, by SGR, geo-coding 10,936,717 households and transmit 2,708,407 signals of municipal population register under-coverage to the municipalities.

26. The municipalities that not attended the RNC have conducted Census of building during the operations of 15 General Census of Population and Housing.