

WP. 14
ENGLISH ONLY

**UNITED NATIONS ECONOMIC COMMISSION
FOR EUROPE (UNECE)**

EUROPEAN COMMISSION

CONFERENCE OF EUROPEAN STATISTICIANS

**STATISTICAL OFFICE OF THE EUROPEAN
UNION (EUROSTAT)**

Joint UNECE/Eurostat work session on statistical data confidentiality
(Tarragona, Spain, 26-28 October 2011)

Topic (iii): Census and other applications

Application for presenting census results in the context of statistical data confidentiality in Poland

Prepared by Amelia Wardzinska, Central Statistical Office of Poland

Application for presenting census results in the context of statistical data confidentiality in Poland

An innovation introduced in both censuses carried out in Poland was GIS technology, which was used at every stage in which digital maps have been prepared. They were essential tools for census enumerators (for mobility in the field, verification of the sampling frame, etc.), municipal leaders and provincial and central dispatchers, who could verify the progress of the census works and the route or location of enumerator on the map. Digital map contains materials derived from the National Mapping Agency (ortophotomaps, borders of voivodeships, counties and municipalities, the names of urban areas), cadastral parcels (ARMA), street network, borders of statistical regions and census enumeration areas and statistical address points created by Polish official statistics.

Digital maps in GIS technology were used in both censuses during the following stages:

1. updating by municipal offices, in which municipalities had the ability to edit address points on-line;
2. during field check out before census:
 - on census enumerators application to allow adding new points, editing or removing existing address points in the field,
 - in the application for municipal leaders and the application for dispatchers, in which they could monitor the course of the census and respond to occurring situations;
3. during the census:
 - on census enumerators application for the interview
 - in the application for municipal leaders and the application for dispatchers, to allow them to monitor and report census progress, as well as planning and managing the work of enumerators in subordinate areas.

The most important aspect of the entire project for preparing statistical address points is their quality. These points have been prepared by the employees of official statistics who for several years were dealing with running and updating of the National Register of Territorial Division of the Country (TERYT). Then their work has been reviewed in the offices of municipalities. The final stage of work was to verify the existence of address points and confirmation of their location in the field via census enumerators. Thanks to the stages of preparing statistical address points they should be considered as reliable.

With such a prepared dataset CSO (Central Statistical Office of Poland) started being independent from changes in the existing system of spatial identifying, moving from spatial assignment (census enumeration area) to point assignment (X, Y coordinates). The change allows more flexible assignment of group data collected from public statistics for arbitrarily small areas - of course while preserving statistical confidentiality. It also allows the creation of a microdata base that will allow analysis of spatial geostatistical various of social phenomena such as:

- demography (e.g. average distance of children from parents living in the country, province, county, municipality, city, neighborhood or street block or any other prescribed area, the average distance from work, school, hospital)
- town and country planning (e.g. to assist in determining the boundaries of urban agglomerations, cities, development of spatial plans),
- agriculture and environment (structure survey crop, environmental pollution),
- the economy (e.g., studying the effects of burdensome road and industrial investments).

Assignment of point with their x, y coordinates will also become independent of burdensome changes in the territorial division of the country, usually resulting in changes in census enumeration and the resulting labor-intensive computation. This will facilitate comparative analysis of time series regardless of the changes taking place in this division.

Spatial data bases prepared for the censuses can be used in cooperation with the management levels in the regions as an extremely helpful tool in making strategic decisions at the municipal, county or province level. With the spatial data it is possible to monitor the site, analyse and forecast the development of the region, to evaluate various risks and develop methods of prevention. A good example of the use of spatial data in crisis management is a flood. The geographic information system can easily determine what terrain may be flooded, and based on data collected during the censuses and surveys, quickly prepare an analysis presenting the number of people and animals that must be evacuated, as well as determine the area of crops which may be flooded.

Official statistics can provide a variety of data and present them at different levels of aggregation, depending on the needs of local communities and local governments at all levels. Information collected during the censuses was collected in a specially protected Operational Microdata Base (OBM). Data collected in OBM, in accordance with both census laws, will be permanently removed no later than after 2 years from the date of completion of each of the censuses. Until then depersonalized data has to be loaded to the Analytical Micro Database (ABM). Dissemination of census data in the form of micro-aggregates, aggregates and result tables will be implemented in an innovative way - through access to ABM. The information obtained will allow to conduct variety of analysis (including spatial analysis) and creation of individual reports. Assessment of contents of the collected data will enable publicly available metadata. Data will be presented in the depersonalized form, preserving statistical confidentiality.

Presentation of the results of censuses

Official statistics data are used widely in almost every unit of public administration. Those uses relate primarily to demographic issues such as forecasting the demand for places in kindergartens and schools.

of Poland has a software and hardware architecture that is going to be used to publish the results of censuses and to implement subsequent surveys. In order to disseminate the results of censuses and their spatial presentation CSO decided to create the "Geostatistic Portal". Portal will serve two functions:

- make aggregated data available in the form of a variety of spatial analysis (pre prepared and individual orders, preserving statistical confidentiality),
- allow address points updates in municipality offices.

The Portal will allow addition of newly created address points as well as edition or deletion of existing address points. Data entered on the server will complement the National Register of Territorial Division of the Country – TERYT.

Portal will be a target of spatial information technology platform for data presentation and CSO geostatistical analysis. In particular this will apply to results of Agricultural Census 2010 and Population and Housing Census 2011 collected in the Analytical Microdata Database (ABM), as well as other micro-aggregated data in analytical databases of the Statistics Public Information System (SISP). The task of the Portal will be spatial presentation of the data collected, analysis of microdata and fulfillment of tasks associated with implementation of the INSPIRE Directive guidelines for subjects 1 and 10 from Annex III of the Directive, including support for metadata profiles and metadata catalogues.

In addition to initially prepared spatial analysis, the individual user will be allowed to order his own thematic maps in the form of a cartogram, based on any of the characteristics of thematic data model. It will be possible to print maps designed by themselves.

It is planned that the Portal will have a feature that allow generating areas of analysis on the basis of: buffering the inserted point, drawing a polygon and generating on the basis of the report summary (tabular report can be supplemented with a map of the selected features, and charts).

It becomes obvious that the demographic, social and economic **surveys** by CSO requires a comprehensive use of GIS technology.

Support for individual orders

1. Support of requirements

Individual users involved in decision-making:

- Main:
 - Needs Manager
 - Product Manager
 - End User
- Support:
 - MS SQL Administrator

Tools and applications used in decision-making

- Main:
 - Inside Users Application
- Support:
 - MS SQL Server

Description of decision-making process

The decision making process starts by the interior user verifying if the product has been already prepared and published. If the appropriate product is in the list of prepared products- it can be downloaded.

If the required product is not found on the list, the user can report a need for the implementation of the product and define its character. The main way to specify the need is to use a graphical interface where the user specifies the horizontal and vertical dimensions (together with the hierarchies and their levels, and attributes), and measure (with an indication of aggregation functions), which are going to be included in the report. An alternative method of the need specification is writing its description in the text box.

After specifying, the need is registered.

Registered need is forwarded to the Needs Manager, who validates the specification. In case of deficiencies in the specifications it will be returned to the applicant for modification. In case of a positive result of verification Needs Manager is checking a list of all the products in the ABM system (public and private). After finding the appropriate product Needs Manager is making the product available for download by the applicant. If the ABM system does not have the appropriate product, the need is targeted for completion and assigned to a particular Product Manager.

After preparing the product Needs Manager provides it to the user.

2. Support of orders

Individual users involved in the ordering-support process:

- Main:
 - Product Manager
- Support:
 - Report Analyst
 - Quality Manager
 - MS SQL Administrator

Tools and applications used in the ordering-support process:

- Main:
 - Inside Users Application)
- Support:
 - MS SQL Server

Support of order process description

The process begins with viewing the list of the public products that can be downloaded by all users inside ABM system. In the absence of the appropriate product, Product Manager appoints a new product in the metadata and describes it. Product Manager also manages the existing products (modification, deletion).

Quality Manager should determine the quality requirements for the product and associate the product with appropriate measures of quality.

Then the Product Manager verifies the product metadata and directs it into execution. The performance of the product corresponds to the process of "reporting, analyzing and sharing data."

After preparing the product- Quality Manager checks the quality requirements of it (through the verification of the actual values assigned to the product quality metrics).

If the product meets the quality requirements it is passed back to the Product Manager who makes the product categorization, attributing it to the proper categories, product groups and subgroups.

Product Manager also manages the categories, groups and subgroups of products.

The process ends in publicizing the product that can be consumed by internal users.