Merging statistics and geospatial information, 2012 projects - Slovakia

Statistics Explained

This article forms part of Eurostat 's statistical report on *Merging statistics and geospatial information: 2019 edition* .

Undated final report

Problem

Development work on representing census data on the population grid was incomplete, as address points needed to be consolidated and then linked to statistical data from the census.

Objectives

This project was a continuation of a previous project and concerned the development of datasets based on a population grid. The aim was to connect the results of the 2011 population and housing census (SODB 2011) with a population grid to create a dataset for the latter — according to GEOSTAT methodology — with statistics on the total population by sex and by age.

Method

As part of the 2011 census exercise an Oracle database was created. Boundary information was received from the Geodesy, Cartography and Cadastre Authority of the Slovak Republic, while data for residential units were provided by the Slovak Environmental Agency. Information from the population registry from the Ministry of Interior of the

Slovak Republic was linked to individual address points. An application was developed to facilitate the creation of new address points or to modify or delete existing address points during the data collection exercise for the 2011 census.

Results

The result of using this application to add, delete or modify address points was the establishment of a complete layer of address points for the whole of Slovakia, combining a range of geographical information (such as administrative and census districts, postal ad-
dresses, XY coordinates) with information about inhabitants, in other words a geo-based dataset for the 2011 census.
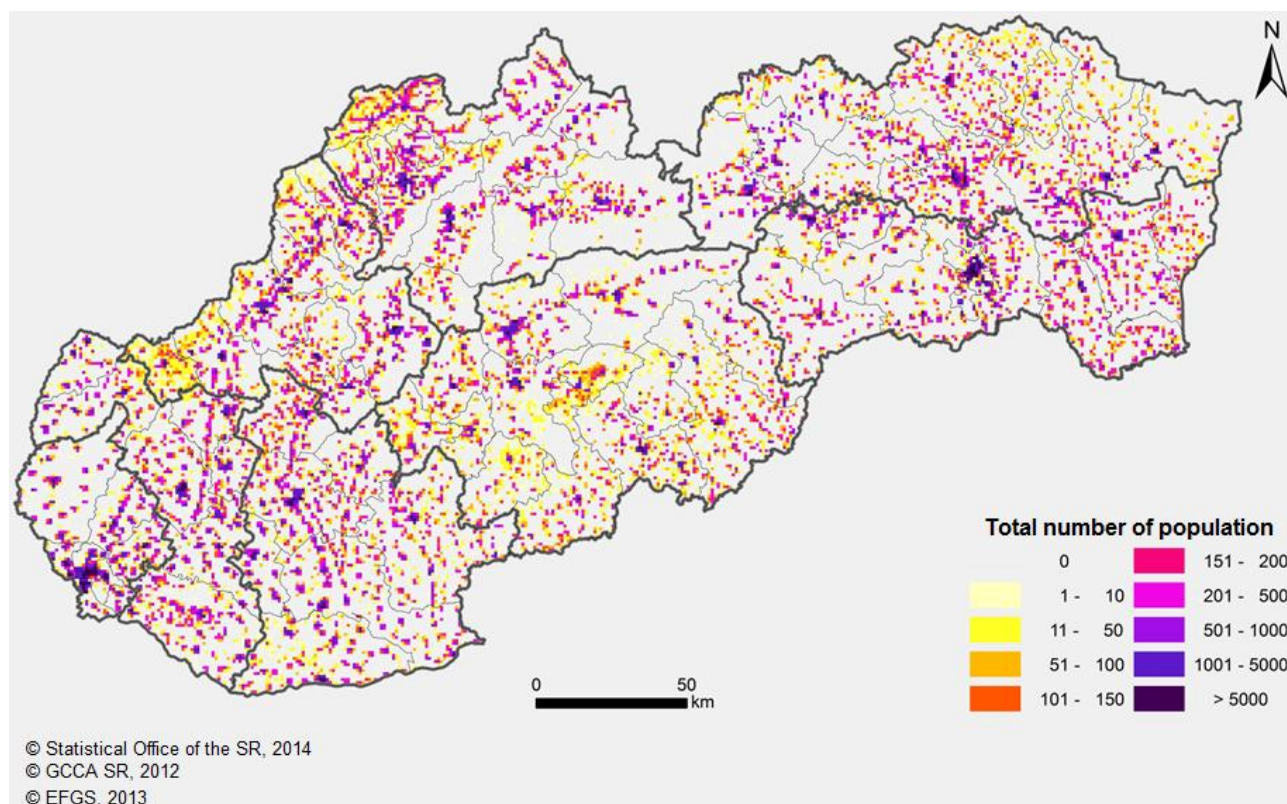

Figure 1: The total number of people in the 1km x 1km raster, Census 2011

From this dataset a range of grid-based statistics was developed containing information on the total population in each grid cell as well as an analysis of demographic information by sex and by age.

While this dataset for grid-based statistics was available for a relatively small range of indicators that were regarded as non-confidential, this was not true for local and regional statistics. Indeed, methods were developed in relation to data presentation and the protection of confidential data, which involved a one-off implementation of a method known as targeted record swapping. The data for individuals were assessed against a number of criteria, resulting in a risk score. A swap rate was calculated for each NUTS level 3 region and from this the number of households to be swapped was calculated for each municipality within that region. Within each municipality a set of households was selected proportional to their risk score. For selected households, addresses in the database were swapped with another household with a similar age-sex structure within the same district (which is a local level situated between a municipality and the NUTS level 3 region). As only the address is swapped the relationship between any other variables in the database are not disturbed, although there might be an impact on aggregated results for some municipalities. An advantage of this approach is that the whole database can be treated once, ensuring that any tabulations made from the database — whether standard or non-standard — are consistent with each other.

The final initiative was to develop a web mapping service for users, incorporating the geo-dataset.

Methodology

- Final report

- My municipality in statistics

- Register of territorial units

External links

- Statistical office of the Slovak Republic