

Merging statistics and geospatial information, 2012 projects - Poland

Statistics Explained

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

Final report 10 February 2014

Problem

The lack of spatial data beyond that traditionally presented for administrative divisions (gminas) led to growing calls from local government, scientists, entrepreneurs and individual consumers for the development of alternative statistics for a broad range of spatial divisions.

Objectives

This project set out to analyse the possibilities to present demographic data in area units other than administrative units that are currently used; it incorporated the manner of presenting statistical data in cadastral units, statistical units and grids. The project was designed to respond to (user) needs for providing data at levels lower than gminas (municipalities; the basic unit of territorial division in Poland), as expressed by local government representatives, scientists, entrepreneurs and individuals.

The project also aimed to develop spatial information for enterprise addresses and to combine this with demographic data to visualise commuting patterns.

Finally, the project looked at the possible use that might be made of spatial data for creating statistical indicators on land use to help the planning work of local government.

Method

The project combined data collected during the national agricultural census in 2010 with spatial data (for cadastral districts, statistical regions and census enumeration areas) from the national census of population and housing (conducted in 2011). As census data are collected with reference to address points, these data can be freely aggregated: the work involved an analysis of the manner of aggregation of demographic data to a number of other spatial classifications. Specifically this was done to produce data for statistical units (statistical regions and census enumeration areas), cadastral units (geodesic precincts) and for a 1 km² grid, although the latter could be modified to reflect any less detailed division of space.

A database of address points representing the location of enterprises was created. Several sources of enterprise addresses were identified (data from the social security system, the Ministry of Finance, the tax authorities, and the statistical business register) and the names and identification numbers of gminas (municipalities) were corrected using the statistical office's register of territorial divisions. Workplace coordinates were assigned to people based on the address identification system of streets, real estate, buildings and dwellings which is part of the statistical office's register of territorial divisions, as well as using the statistical business register and data from the database of topographic objects (of the national mapping agency). In cases where a match could not be achieved between the address identification system and the addresses given in the source for enterprises, a number of simplifications were made in order to try to use as much of the address information as possible.

The Centre for Urban Statistics of the statistical office in Poznan developed a methodology for surveying commuting with the use of data acquired for the purposes of the 2011 census and these showed the possibilities for linking geo-information and statistical information. For the needs of censuses, a spatial database of address points representing the location of residential buildings in Poland was created. Combining this with the new database of address points for enterprises, it was then possible to map commuting statistics. A dataset was prepared containing records for people who had reported in the census that they commuted: the dataset included information on an individual's age and sex, as well as various identifiers for their place of work including the XY coordinates, various identifiers for their place of residence including the XY coordinates, and the type of work performed. From this information the (direct) distance between the XY coordinates of the place of work and the place of residence could be calculated. Two datasets were produced, one focusing on commuter departures (from home) and the other on commuter arrivals (at work).

In addition, the Regional and Environmental Surveys Department conducted an analysis of the possible use of spatial data for creating statistical indicators on land use planning. This involved a conceptual stage (for example, defining survey methods), data collection, data evaluation and then data analysis. For one voivodship (the highest level of administrative unit in Poland that corresponds to a province in several other countries), a database of topographic objects was used, along with an orthophotograph (an image that has been geometrically corrected to give a uniform scale, similar to that in a map); the data were also compared with the land and building register. The data in the database of topographic objects were evaluated to look for errors. This was done by checking the information available for a sample (5 %) of 1 km² grid cells to see if the database correctly contained the objects observed on the orthophotograph. Around four fifths of the grid cells had the same information in the database as observable on the map for roads, with this share rising to 100 % in rural areas; most of the differences in urban areas concerned short sections of roads for groups of new buildings. Concerning buildings, the database was at least 75 % compliant for more than four fifths of all grid cells. Again, grid cells in urban regions were more likely to show higher non-compliance with the database, often related to recent or on-going construction activity.

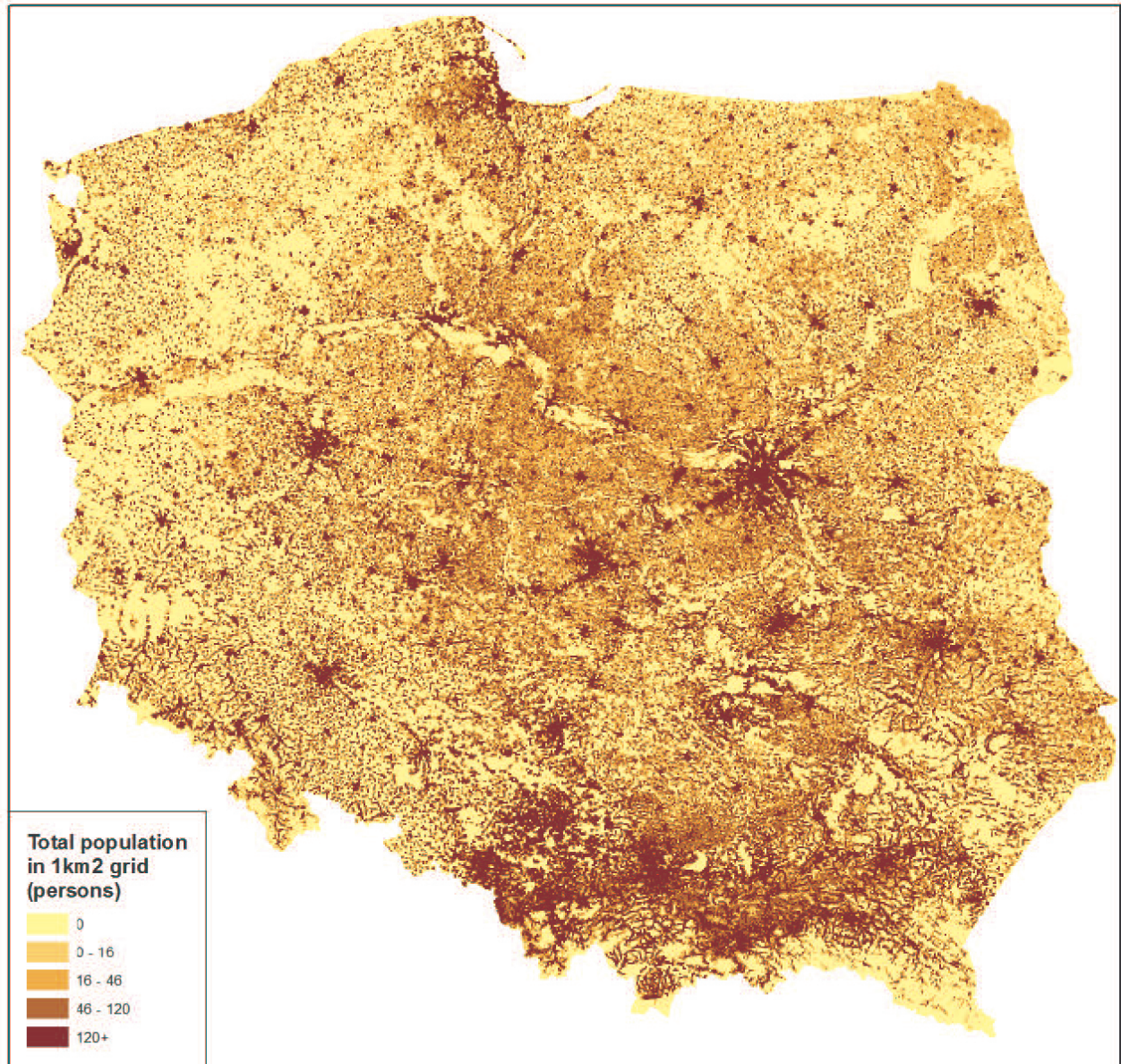


Figure 1: Example of spatial visualisation of demographic data, Poland

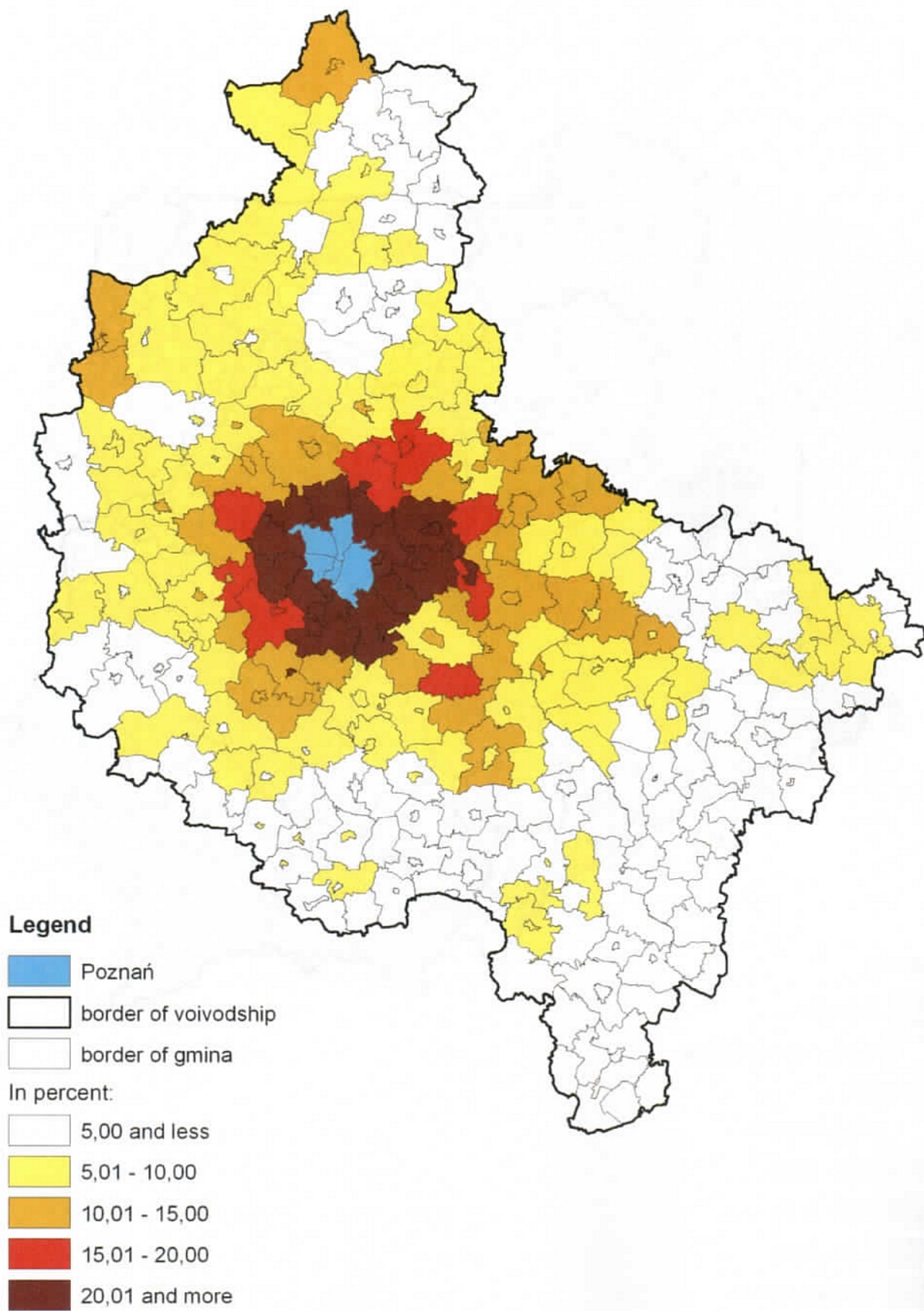


Figure 2: The share of persons arriving to work in Poznań in the number of employees in the gmina of residence in 2011 (in voidvoship scale)

Results

After performing disclosure control on the census data, thematic choropleth maps were prepared and published on a platform developed for the spatial visualisation of statistical data — the Geostatistics Portal. This resulted in the presentation of demographic data that was divided in a different manner to traditional administrative divisions. There was a strong level of demand for the resulting information, especially for demographic information at detailed levels, below that of gminas (municipalities).

The work on assigning address points representing the location of enterprises resulted in a complete dataset with XY coordinates for the place of work for all employed persons, either based on an exact identification of the address of an enterprise or an approximation thereof. The database may be used to visualise many phenomena and statistical data that are related to surveys conducted at the place of work.

Various analyses of these commuter datasets were performed, for example, simply identifying how many people in each voivodship (of which there are 16 in Poland) were commuters, showing at a very detailed level where people worked, or showing the share of commuters among all persons employed. As well as presenting these data according to various administrative and statistical areas, analyses were also performed for 1 km² grid cells, showing where people commuted from and where the numbers of net inward commuters was highest. Further analyses were performed focusing on commuting patterns to particular locations, for example showing the origin of people commuting to a specific city, as well as analysing the structure of commuter flows by age or by sex, and the average distance travelled between a commuter's place of residence and place of work.

The work on land use evaluated the quality of the database of topographic objects as good. The gaps in data were generally in heavily developed areas where new infrastructure emerged and new buildings were constructed (suggesting that identification issues were more likely linked to new building work rather than any to any underlying issue in relation to the methodology for identifying/registering objects). This subproject also developed methods for acquiring information on the density of buildings and road networks, starting from the number of buildings, the extent of the built-up area and the length of roads; this was done for gminas (municipalities) and for 1 km² grid cells.

The conclusions drawn from this work confirmed that it is possible to produce data: i) at a detailed level (for 1 km² grid cells) for buildings which previously were only available for gminas and ii) on road infrastructure which was previously not even available for gminas.

Methodology

Further information and links

- [Final report](#)
- [Annex I to final report](#)
- [Annex II to final report](#)
- [Commuting to work in Poland — results of the national census of population and housing 2011](#)
- [Mapping portal](#)
- [TERYT register](#)

External links

- [Statistics Poland](#)