Detail or Disclosure – Towards a Visualisation of Confidentiality Related Spatial Damage to Demographic Grids.

Sebastian Specht [a,*], Bernd Kramer [b]

[a] OFFIS – Institute for Informatics, Oldenburg, Germany, Sebastian Specht – sebastian.specht@offis.de
[b] regio gmbh, Oldenburg, Germany, Bernd Kramer – kramer@regio-gmbh.de

* Corresponding author

Keywords: cartographic grids, demography, disclosure control, census

Abstract:

Statistical data on demography is the basis for many population-related scientific questions, economic questions of health care and questions of planning public services. Population data in equal-area cartographic grid cells appears to be a good basis, especially for use cases in inter-municipal contexts of administration and planning (Specht et al. 2019). Census results have been used, since the 2011 census made available small-scale population data for the entire Federal Republic of Germany on a 100m grid for the first time. Unfortunately, this data is not updated by the statistical offices.

This presentation describes a use case of demographic grids implemented in a context of inter-municipal cooperation in the region of Bremen. As the calculation of population forecasts was an objective, small-scale data on migration was required. Similar to the approach in the census, demographic data and data on migration are recorded in the residents' registration offices (EMA) of the cooperating municipalities. However, since outside the census other legal frameworks apply, the process cannot be adopted as is. In the EMAs, individual-related micro-data are available, serving as a base file. Under the respective legal framework, the data is anonymised, geo coded and converted into an aggregated tabular form on site.

Aggregated data may still contain individual cases worthy of protection. The higher the number of queried characteristics (region, gender, age, nationality etc.) and their differentiation (100m grid or 1km grid, age years or age groups, etc.), the higher the probability to encounter such cases. A number of procedures for statistical disclosure control are available, of which the SAFE procedure (Höhne 2015) (used in the 2011 census) is currently implemented in the project. As other methods or strategies are up for consideration, how can they be evaluated in a specific regional context?

From the perspective of confidentiality, space is at first just one feature dimension among others, although there are approaches that explicitly take spatial interrelation into account (Young, Martin, and Skinner 2009). From a geographical point of view, however, high resolution data, especially in sparsely populated areas, can generally be expected to show high before-and-after deviations as a result of confidentiality procedures. Depending on the subject matter, these spatial errors can have different degrees of relevance and thus be ultimately relevant for the selection of the confidentiality strategy. To support a decidedly spatial comparison of the effects of different classification, aggregation and confidentiality strategies, a set of indicators together with an interactive visualization for the project area under consideration is presented for discussion.

References

Höhne, J., 2015. Das Geheimhaltungsverfahren SAFE. In: *Zeitschrift Für Amtliche Statistik Berlin Brandenburg* 2: 16–33. https://www.statistik-berlin-brandenburg.de/produkte/Zeitschrift/2007/HZS_070506.pdf. (accessed 2020-09-09)

Specht, S., Blohm, K., Handtke, T. and Wette. L., 2019. Prognosen Im Bevölkerungsraster Für Die Interkommunale Kooperation – Ein Experiment Im Reallabor. In: *AGIT – Journal Für Angewandte Geoinformatik*, 5-2019, no. 5: 284–91. https://doi.org/doi:10.14627/537669027.

Young, C., Martin, D. and Skinner, C., 2009. Geographically Intelligent Disclosure Control for Flexible Aggregation of Census Data. In: *International Journal of Geographical Information Science* 23 (4): 457–82. https://doi.org/10.1080/13658810801949835.