Accessibility of primary schools

Project goal

The project target was to dissect the average route distance for families with at least one child of age six to ten years to the nearest primary school. Based on the microcensus[1] and by applying suitable weights the goal was to obtain representative accessibility statistics on family level without increasing the response burden on citizens.

Methodology

For the analysis an experimental data set of the microcensus was merged with additional data that were not covered by the original survey but integrated via the individual observations' geospatial information. German statistics law (BStatG) allows for the permanent linking of statistical data to a geographical quadratic unit with a size of at least 100m x 100m. These data units arranged in a regular grid form a raster which offers great potential for small scale analysis irrespective of administrative units and discovery of geospatial dependencies. Furthermore, the grid cell ID serves as a unique key that facilitates the merging of other data sources via its spatial relation.

As a first step, the primary schools were geocoded by translating their addresses[2] in coordinate pairs exactly describing the geographic location. The households of the microcensus were similarly geocoded by extending the data set with coordinates and the corresponding 100m INSPIRE compliant grid cell IDs appertaining to the location, respectively.

The calculation of the accessibility of primary schools was operationalized by determining the route distance (not Euclidian distance) or required travel time for pre-defined zones (e.g. below 300m, 300m to 500m, and so forth) around each school serving as starting point, applying the routing engine hosted by the Federal Agency for Cartography and Geodesy (Bundesamt für Kartographie und Geodäsie, BKG) based on the "OpenRouteService", which is developed by the Heidelberg Institute for Geoinformation Technology (HeiGIT) of Heidelberg University. As a data basis for the routing the publicly accessible data basis "OpenStreetMap" was used[3]. The polygon shapes representing the accessibility zones were then rasterized to the reference grid mentioned above to allow for the subsequent merging with microcensus data, thus extending the data set with information about distances and travel time to the nearest primary school.

Main message / recommendation

The analysis demonstrates the possibility to extend existing datasets (also samples) with information gained by using geospatial information. High quality statistical information is generated without any burden on respondents. The analysis potential based on official data may be extended. A whole range on new analysis questions may rise and also may be answered by the transfer of the method to other thematic factors.

Open Topics

To ensure that the obtained statistical data are valid, it is first and foremost required that all data used in the analysis – microcensus, school addresses, routing data – pertain to the same reference year. One

---

[1] The microcensus is a representative household survey of the German federal statistic addressing living conditions, which covers one percent of Germany's population every year since 1957. To ensure representativeness of the results participation in the survey is mandatory for the selected citizens.
[2] The addresses are sourced from a publicly accessible school register.
[3] There is currently no ministerial data basis suitable for routing and this type of analyses.

needs to keep in mind, however, that the results remain abstract (and are therefore distorted) in the way that the calculated distances and travel times are mere theoretical, static values, neglecting the daily and hourly changing traffic situation. Still, considering actual route options for different travel modes (because not only motorized private transport but also bicycle riding or walking on foot may be considered) reflects the true distances better than Euclidian distances. Furthermore, in the future results can possibly be improved by applying weights to the network graph edges accounting for these effects during routing.