

**Title:** Mobile positioning and statistical derivatives – The way forward?

**Short description:**

National statistical institutes (NSI) usually bridge the gap between the interest of both data controllers and the user regarding the open data dissemination policy. Particularly the data controllers managing personal data are subject to the restrictive legislation which precisely determines the possible application of the data. The interoperability of these data is consequently insignificant especially because the data controllers are normally not dissemination-oriented and have no business interest in sharing these data for wider re-use. Being one of the major users of various public or private (personal) data bases, the NSIs are often the only official institution that communicates these data to the expert or general public as open data. Among these, the data managed by the mobile operators are considered to be very desirable, yet also highly restricted. Nevertheless, the Statistical Office of the Republic of Slovenia (SURS) decided to accept the challenge and to explore the possibility of applying the mobile data to official statistical production and to produce applicable datasets that would satisfy the user demands on the one hand and the (personal) data protection rules on the other.

In 2014 SURS launched a pilot project within the Eurostat grant “Merging statistics and geospatial information” with the aim to study the adequacy of the mobile data for statistical purposes and to recognise SURS’s current potential deficiencies regarding the data infrastructures or human resources. The project should also answer some questions regarding the safety of handling these sensitive data (of interest to SURS and the [Information Commissioner](#)) and the possible application of these also by other institutions.

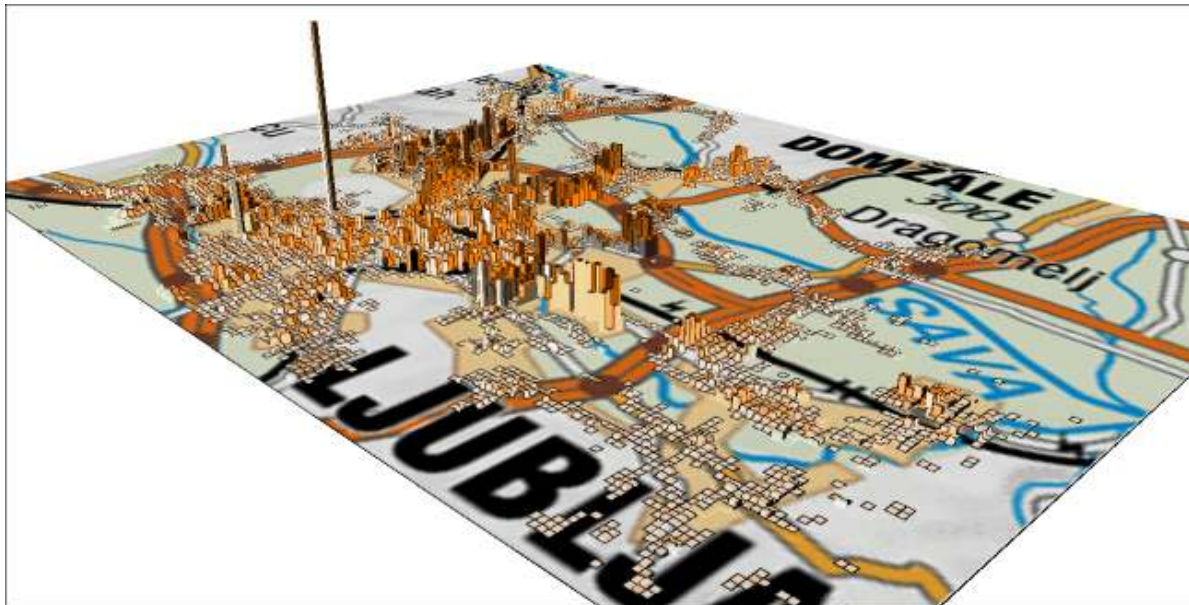
SURS obtained the data from the second largest mobile operator in Slovenia (30% share) on traffic between May and October 2014. The data included XY co-ordinate of the base stations, anonymised ID of the user and exact time of the transactions (call, SMS, etc.).

**Key challenges and opportunities:**

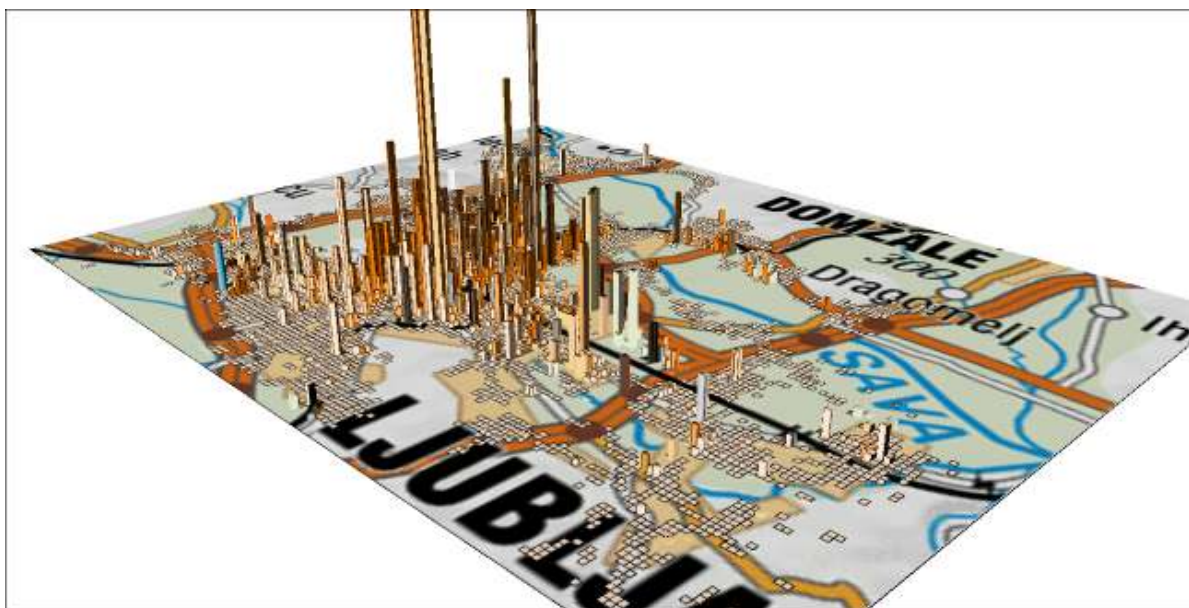
SURS is currently studying the possibility of improving the time-space accuracy of the existing register-based data on population distribution and population mobility (Figure 1 and Figure 2). The next challenge will be to segment various population groups directly from the mobile data. SURS is also exploring the possibility of recognising some patterns that could indicate tourist travels. Besides the option to apply the mobile data for statistical purposes, the Administration for Civil Protection and Disaster Relief has already been recognised as one of the most important users of derived statistical products with direct impact on their efficient emergency planning and response.

SURS is also considering taking steps towards changing the legislation to ensure the access to optional data variables managed by the mobile operators, including personal IDs of the users.

**Figure 1:** “Night time” population distribution by 500m x 500m grid, Ljubljana, 2015. The Central population register containing data on the actual residence of the population is temporarily the only data source that presents the population distribution for the period of the day when people are expected to be at home.



**Figure 2:** “Day time” population distribution by 500m x 500m grid, Ljubljana, 2015. The sources for the “day time” distribution are the Statistical register of employment and data bases on students, secondary and primary school pupils.



SURS intends to compare the mobile data to the register data in order to estimate the accuracy of the latter and also to get a seamless hourly situation of the data presented in Figures 1 and 2.

**How to get started:**

Data access: The key issue regarding the mobile data is data access. Therefore, studying the national and EU legislation was necessary in case of Slovenia. SURS contacted also the Information Commissioner and soon it became obvious that the Electronic Communications Act stands against the National Statistics Act since it precisely specifies the application of the mobile data and does not mention data could be used also for the statistical purposes. Therefore, the Information Commissioner agreed that SURS, which as a rule handles personal data, may obtain these data from the operators on voluntary basis. The second largest operator in Slovenia agreed to co-operate.

Data processing: SURS has already established a secure environment for transferring, storing and handling of personal data and ensured the auditability of the entire process. Thus, many issues regarding these first steps have already been solved. The main challenge was therefore to establish a data lake that could provide more structured data for further analysis.

SURS set up a team of IT, analytics, statistical disclosure control and GIS experts to implement this pilot project.