

Mobility Community Reports data: Geovisual analytics and cartographic synthesis of behaviour changes due to COVID-19 pandemic in Europe

Vít Pászto ^{a,*}, Alena Vondráková ^a

^a Department of Geoinformatics, Palacký University Olomouc, 77900 Olomouc, Czech Republic, vit.paszto@gmail.com, alena.vondrakova@upol.cz

* Corresponding author

Keywords: COVID-19, mobility, geovisual analytics, clustering, synthesis

Abstract:

It was probably more valuable to use user-generated data from location services than ever before during the COVID-19 pandemic. The newly emerged situation caused by the disease's rapid spread found most of the European countries not ready in many aspects. From the geospatial point of view, it was a tremendous lack of open-data about people's behaviour. Many public and private services were closed as a preventive measure to slow down COVID-19 spread, which caused an unprecedented change in citizens' mobility and overall behaviour. It is not possible to acquire such information using traditional ways of data collection (e.g. surveys and questionnaires), or it would be extremely expensive and time-consuming. In order to obtain information about people's change in behaviour, location-based data from mobile phone users could be used. Although individual user-generated data are highly sensitive, they could be aggregated (e.g. for higher administrative units) and thus anonymised; while still possessing valuable insight into the behaviour change. This contribution deals with COVID-19 Community Mobility Reports dataset released in early April 2020 by Google (available at <https://www.google.com/covid19/mobility>), which offers unique information about a change of human activities due to the pandemic. This data is in principle very similar to data from mobile operators; however, more accurate, because the location of the device is determined also by connection to Wi-Fi networks and especially via GPS – besides nearby BTS stations. Information carried by the data gives the average decline in population activity in spatial resolution, either only for the entire country or for the country and its main regions. The data is constantly being updated, but Google provides no information on the date of future dissemination of data. Behavioural changes as such are monitored in six categories according to the location of the activity, which was determined by Google as useful for maintaining social distancing or from the point of view of the availability of basic services. The categories are as follows – a) residential, b) transit stations, c) retail and recreation, d) grocery and pharmacy, e) parks, and f) workplaces. Key information is the percentage decrease or increase in the number of individuals present compared to the usual state of affairs (baseline – 3 January to 6 February 2020) in these six fundamental localities. We analysed and visualized data during the peak of the pandemic (from 5 March to 11 April 2020) by processing the average value from that period. In total, 567 European regions with six pieces of information on activity in the above categories was analyzed and visualized. First, we displayed average values in given categories in separate maps, which allowed us to understand a spatial pattern of data. This procedure is relatively common step in the data visual analytics process. We present resulting maps in the contribution of the conference.

From the analytical part, we applied a cluster analysis, which served as the processing input for the cartographical synthesis. Methodologically, we created a typology based on the results of cluster analysis as multiple variables can be analyzed simultaneously to provide groups of types of regions with common properties. The k-medoid method was chosen for processing data on human activity type due to its lower sensitivity to outliers. By calculating auxiliary numerical statistics (Pseudo F-index, gap statistics, silhouette method) and after expert assessment, we chose to analyse and display results into five categories (Figure 1), although the optimum number of clusters was originally determined to be two, which was cartographically senseless. This also gives readers a finer breakdown of European regions, which more appropriately shows the diversity/similarity of individual types of regions; especially in the case of areas adjacent to the countries most severely afflicted by COVID-19 - for example Portugal, southern Austria, Slovenia, and others.

Besides analytical processing of data (e.g. geovisual and cluster analysis), we also explored a more advance cartographic approach. Cartographic synthesis methods allow the visualization of various types of data with qualitative and quantitative resolution. Therefore, the synthetic processing of the above-mentioned data has been implemented. During the pandemic, a lot of maps were created that are cartographically incorrect, as well as a lot of maps that are

cartographically correct. In most cases, however, it was only a simple visualization of the amount of positive-tested people or any other dealing issues in the form of analytical maps. Thus, synthetic maps can provide a new perspective on the issue and can be useful in unconventionally providing information about Covid-19 pandemic. We show how this dataset can be utilised in terms of cartographic synthesis in form of a regional typology in order to reveal the spatial pattern of such change in citizens' behaviour.

The resulting typology categorizes the relevant administrative units into five clusters (types) in the space defined by the change in behaviour. These clusters can also be named according to the impact of the COVID-19 pandemic on behaviour of citizens: **1) moderate** - regions with a relatively small change in population activity. For this type, it must especially be emphasized that this involves a mild impact of the pandemic compared to other types and that COVID-19 and related restrictions have had a visible effect on human behaviour in these regions as well; **2) substantial (secondary activities)** - regions with a generally statistically average change in population activity with a more significant change in secondary activities (Grocery and Pharmacy, Parks); **3) substantial (main activities)** - regions with a generally statistically average change in population activity with a more significant change in main activities (Residential, Transit Stations, Workplaces); **4) significant** - regions with a statistically significant change in behaviour (values in the fourth quartile outside the range of the box graph); **5) extreme** - regions with an unprecedented change in behaviour in outliers or near their borders.

Five types of regions were identified according to the impact of COVID-19 and related restrictions - from the type moderately impacted by COVID-19 (e.g. Sweden, Latvia, Hungary), to those impacted by the pandemic in a substantial (Ireland, regions of Greece, the Czech Republic, Norway, Switzerland), significant (e.g. regions of France, Belgium, Austria, Slovenia, Portugal), and extreme manner (Spain, Italy, and the Paris region). As a conclusion, The current global situation clearly shows that positional and individual data can also be useful in the case of pandemics or any other situations involving the safety and health of the population. The use of Google Location data proves to be valuable in analysis and evaluation of citizens' behaviour during any pandemic crisis, though, in this case, aggregated into higher territorial units.

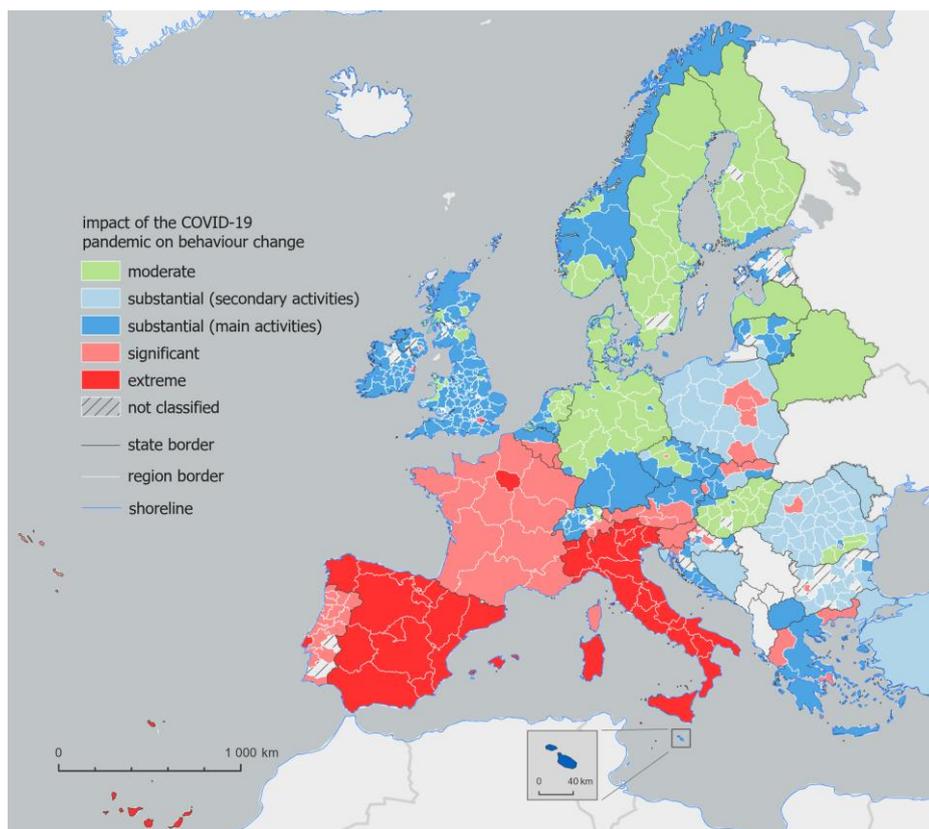


Figure 1. Impact of the COVID-19 pandemic on behavior change during the peak of the first pandemic wave (March 5 to 11 April 2020)

Acknowledgements

This contribution has been supported by the project “Spatial synthesis based on advanced geocomputation methods” (no. GA18-05432S) funded by the Czech Science Foundation.