International Cartographic Association
Association Cartographique Internationale

A dashboard application to explore population distribution derived from GPS location data during the COVID-19 pandemic in Kyoto, Japan

James Todd [a,*], Keiji Yano [b], Kazumasa Hanaoka [b]

[a] *University College London Department of Geography, Dr. James Todd - james.todd@ucl.ac.uk*
[b] *Ritsumeikan University Department of Geography, Prof. Keiji Yano - yano@lt.ritsumei.ac.jp, Dr. Kazumasa Hanaoka - kht27176@fc.ritsumei.ac.jp*

* Corresponding author

Keywords: dashboard, GPS, mobile phone, COVID-19, mobility dynamics

Abstract:

The COVID-19 pandemic resulted in unprecedented restrictions on population mobility patterns through the implementation of governmental non-pharmaceutical interventions (NPI). These refer to a variety of policies that aim to mitigate the spread of the virus through measures such as the social distancing, gathering size restrictions or cancellations, mandatory face coverings, work, education and retail space closures as well as lockdown restrictions that forced individuals to quarantine in their homes. Current research surrounding the impacts of NPI on population mobility have been limited to broad estimates of changes in journey volumes (de Palma et al. 2020), durations (Borkowski et al. 2021) and modality (Dingil and Esztergár-Kiss 2021). This lack of granularity is primarily a result of the coarse and aggregated nature of underlying data, such as survey responses or the Apple and Google Mobility data reports.

Individual mobile phone Global Positioning System (GPS) location data are an increasingly abundant and highly informative form of passively produced data that have great potential in deriving novel insights on mobility dynamics. They are typically collected through applications that have location services enabled that are then bought by data brokers and sold onto a wide variety of customers, including commercial firms for purposes such as advertisement and location planning as well as academic research fields including spatial epidemiology, migratory flows and urban planning (Barbosa-Filho et al. 2018). Within the context of this research, GPS location data have been sourced from Japanese data provider (Blogwatcher, Inc. https://www.blogwatcher.co.jp/) and employed to explore the nature of population distribution throughout the pandemic period in the Kyoto city area. The data are incredibly rich, containing GPS location data of over 4.4 million individuals across a two-year study period (1st January 2019 and 31st December 2020). They were provided in Comma-Separated Value (CSV) file formats in excess of 4-terabytes, with an average of nearly 3,500 ping locations per individual per month. They were supplemented with additional user attributes for anonymised identifiers that detail individual's gender, age band (at 10-year intervals) and estimated home and work locations (within 125-metre grid cells) that have been calculated by the data provider. As such, there are numerous potential applications of such data in creating a granular understanding of mobility pattern changes, whether that be for a specific NPI implementation, a certain event or demographic characteristics.

Given the wide-ranging nature of possible implementations, this project aimed to create an interactive cartographic dashboard application that would enable researchers to efficiently conduct preliminary exploratory analysis of these GPS location data. This was developed using *R Shiny,* an *R* package that enables programmatic creation of interactive web applications, depicting counts of individuals across the study region for user specified levels of spatial aggregation (125-metre to 1-kilometre grid cells), timestamps (between 1st January 2019 and 31st December 2020) in addition to aggregations of individual attributes (gender, age and estimated home locations). The dashboard also contains functionality to conduct comparisons to between 2019 and 2020, both in terms of the equivalent day and exact date, as well as download the data that are currently in view on the dashboard.

The creation of the dashboard relied on initially processing individual GPS location data from CSV file formats to a *PostGIS* database. This was essential for two primary reasons. Firstly, the data in their raw state contain a vast number of unnecessary ping locations such as those during transit as well as large concentrations of points whilst individuals are stationary. Since population distributions in the dashboard aimed to identify of those individuals that were visiting particular locations, the first-stage stop point detection algorithm, developed by Hariharan and Toyama (2004), was implemented. Secondly, CSV file formats are inconducive for rapid data querying and storage due to their enormity and unstructured nature. As such, pre-processing the data to identify visit locations at the smallest spatial resolution into a

Abstracts of the International Cartographic Association, 6, 258, 2023.
31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa.
https://doi.org/10.5194/ica-abs-6-258-2023 |

database table considerably reduce the size of these data as well as increasing the efficiency of resultant visualisations. The *PostGIS* database also included tables of individual user attributes that are aggregated on an ad hoc basis dependent on user inputs in the dashboard application.

The *R Shiny* dashboard application is created using an *R* implementation of *deck.gl,* a high-performance WebGL-based visualisation tool for large datasets that uses vector tiles. This has been elected over alternative raster tile mapping platforms, such as leaflet, due to its advantages in enabling much smoother, faster visualisation and interactions such as zooming, panning and rotating. The entire application, including the *PostGIS* database, has been packaged into a docker container to enable its portability to researchers as well as possible future implementations to other web-hosting platforms.

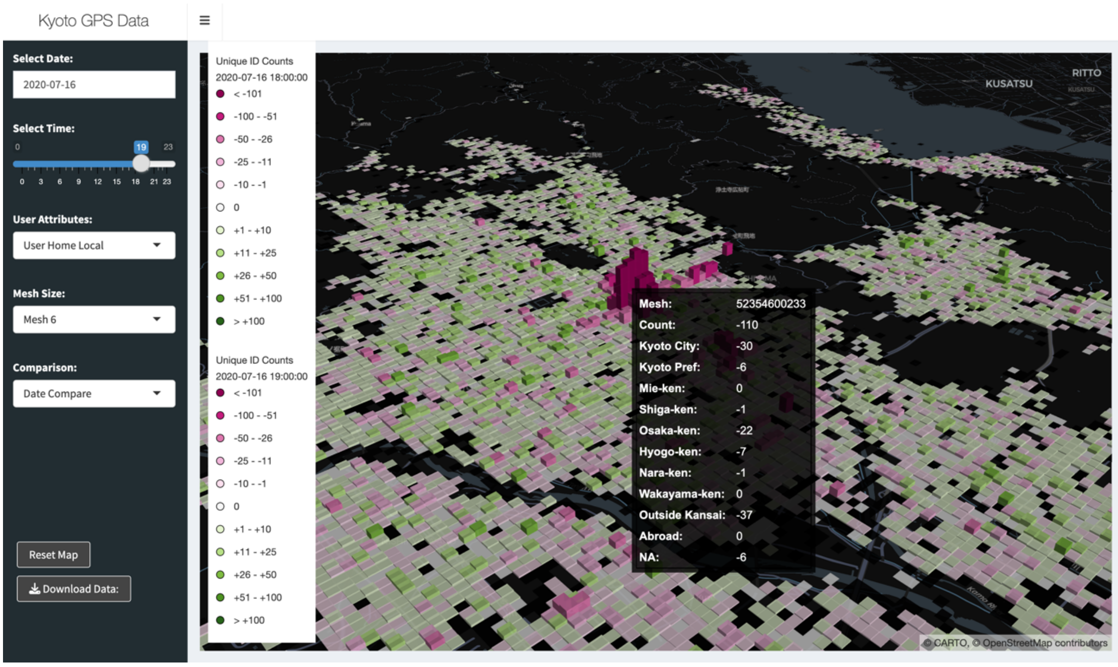

Figure 1. Screenshot of dashboard application exploring the area and time associated with the Gion Festival on 16th July.

In its current state, the dashboard provides a flexible and efficient tool to explore these data in a much more intuitive and visual manner in comparison to typical programmatic techniques. For example, Figure 1 depicts an exploration of the changes in population distribution during the Gion Festival between 2019 and 2020. This annual festival was cancelled in 2020 due to COVID-19, resulting in a significant decrease in population concentrations in the areas where parades and food stalls would line closed streets, particularly of individuals with estimated home locations outside of the Kansai region. Explorative insights such as these are intended to guide and aid researchers in conducting further, more rigorous analyses. This highlights the values of developing dashboards in exploring very large and complex data that have numerous potential applications.

Future work is planned to improve the efficiency of the dashboard visualisation process through the creation of additional pre-processed database tables instead of conducting aggregations on an ad hoc basis per each user input in conjunction with eliminating known bugs such as the persistence of legends that refer to preceding plots. Due to the highly sensitive and disclosive nature of the underlying data, should the dashboard application be extended to be made accessible to public audiences, although the visualisations themselves do not enable the identification of any individuals, the architecture would have to be modified to ensure safe storage of the data.

References

Barbosa-Filho, H. *et al.* 2018. Human Mobility: Models and Applications. *Physics Reports*. 734, pp. 1–74.

Borkowski, P., Jażdżewska-Gutta, M. and Szmelter-Jarosz, A. 2021. Lockdowned: Everyday mobility changes in response to COVID-19. *Journal of Transport Geography*. 90, 102906.

de Palma, A., Vosough, S. and Liao, F. 2022. An overview of effects of COVID-19 on mobility and lifestyle: 18 months since the outbreak. *Transportation Research Part A: Policy and Practice*. 159, pp. 372–397.

Dingil, A.E. and Esztergár-Kiss, D. 2021. The Influence of the Covid-19 Pandemic on Mobility Patterns: The First Wave's Results. *Transportation Letters*. pp. 1–13.

Hariharan, R. and Toyama, K., 2004. Project Lachesis: Parsing and Modeling Location Histories, in: Egenhofer, M.J., Freksa, C., Miller, H.J. (Eds.), *Geographic Information Science*. Springer Berlin Heidelberg. pp. 106–124.

Abstracts of the International Cartographic Association, 6, 258, 2023.
31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa.
https://doi.org/10.5194/ica-abs-6-258-2023 |