



Interactive web mapping of geodemographics through user-specified regionalisations

Tian Lan & Paul Longley

To cite this article: Tian Lan & Paul Longley (2021): Interactive web mapping of geodemographics through user-specified regionalisations, Journal of Maps, DOI: [10.1080/17445647.2021.1912667](https://doi.org/10.1080/17445647.2021.1912667)

To link to this article: <https://doi.org/10.1080/17445647.2021.1912667>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group on behalf of Journal of Maps



[View supplementary material](#)



Published online: 23 Apr 2021.



[Submit your article to this journal](#)



Article views: 607



[View related articles](#)



[View Crossmark data](#)



Interactive web mapping of geodemographics through user-specified regionalisations

Tian Lan and Paul Longley

Department of Geography, University College London, London, UK

ABSTRACT

The analysis of spatial distributions is possible using a broad spectrum of new and existing digital data sources. Challenges can arise with respect to use of areal units that are both appropriate and compatible. In addition, regional statistics are prone to scale and aggregation effects that manifest the modifiable areal unit problem (MAUP). This paper introduces a web mapping system that allows users to experiment with standard and bespoke zonal schemes in the geodemographic analysis of regional patterns. We describe the architecture and design of the platform and its associated data processing techniques before demonstrating its value through user case scenarios. Using the segregation index as an example, we demonstrate how the use of interactive maps can assist in revealing the scale-dependent nature of the index. Our web mapping system can be employed to help geography students, policymakers and researchers better understand the underlying geodemographic structure of functional regions.

ARTICLE HISTORY

Received 7 August 2020
Revised 29 March 2021
Accepted 31 March 2021

KEYWORDS

Interactive web mapping; geodemographics; scale; aggregation; MAUP; data-informed policy making

1. Introduction

Recent years have seen a proliferation in both the number of available data sources and the range of geographic aggregations that analysts use to frame policy formulation and analysis. New data sources (e.g. Yue et al., 2014), methods of data visualisations (e.g. van Dijk & Longley, 2020) and data-driven analytics (e.g. Engin et al., 2020) are thus reshaping the conception, representation and analysis of geographic phenomena, creating new challenges for spatial decision support systems (Armstrong & Densham, 1990). In this data-rich and data-driven environment there are few, if any, natural geographic units to frame policy analysis, yet increasing visual and analytic capacity to experiment with different geographic scales and aggregations, and to reconcile the areal geographies of different data sources using bespoke or hybrid geographies.

Some existing web mapping tools go certain way to responding to these challenges. LuminoCity3D¹ employs both map visualisations and infographics to explore UK urban dynamics (Smith, 2016). CDRC Maps² and DataShine (O'Brien & Cheshire, 2016) provide interactive mapping platforms for various demographic information and map-based research (Smith, 2005). However, these systems are only designated for extracting census statistics or for aggregating other open or consumer data to standard geographies used in the UK censuses (e.g. Output Areas) or for

other administrative geographies – but not for ad-hoc areal classification.

In addition, most social, economic or demographic datasets are aggregated to predefined administrative units, for reasons of disclosure control for personal and sensitive data. How a spatial distribution is framed invokes issues of scale and aggregation that underpin the modifiable areal unit problem (Openshaw, 1984). Electoral districting provides perhaps the best known example of how purposeful ‘gerrymandered’ aggregation of electoral wards can be used to deliver partisan outcomes (Longley et al., 2015, pp. 320–321). Similarly, scale issues have been explored by Fotheringham and Wong (1991) to describe how different spatial granularities can highlight positive, negative or no apparent relationships between statistical variables.

There is no analytical solution to these issues, which must ultimately be addressed through a clear hypothesis formulation (Goodchild & Longley, 2018): but where requirements of disclosure control restrict analysis to aggregations, multiscale and multiple aggregation analysis can shed light on the effects of founding analysis upon preordained and artificial building blocks. Thus, in this paper, we create a web mapping system (see [Main Map](#)) drawing upon administrative and commercial data that have already been pre-aggregated for disclosure control purposes. It aims to highlight and mitigate the effects of scale and aggregation via different ways of summarising

CONTACT Paul Longley p.longley@ucl.ac.uk Department of Geography, University College London, London, UK

Supplemental data for this article can be accessed <https://doi.org/10.1080/17445647.2021.1912667>

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group on behalf of Journal of Maps

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

populations for calculating segregation indices, whilst also allowing users to generate their own bespoke areas from the most widely used geographic units in the UK, which may be more consistent with underlying research hypotheses.

We shall firstly introduce the data sources and data pre-processing that are used in our interactive web mapping platform and then illustrate the cartographic design and implementation of the web mapping platform, focusing on map design decisions made. We shall subsequently present three map use cases to demonstrate the usability of the geodemographic web mapping system.

2. Data sources

The web toolkit makes it possible to visualise various demographic data sources. Here, for illustration purposes, we choose sex, age structure, ethnicity composition and a residential segregation index as the variables of interest derived from the 2011 UK Census of Population data and Consumer Register data.

2.1. 2011 Census population data

Census data are one of the most commonly used resources used to understand population and society (Lan & Longley, 2019; Walford, 2019) in order to contribute towards evidence-based policy making. Population counts are broken down by age bands and sex and are aggregated to Middle Layer Super Output Area (MSOA)³ levels.

2.2. Consumer Registers and modelled ethnicity data

We also make use of data extracted from Consumer Registers (Lansley et al., 2019), which are modelled to record individual names and residential addresses, with near-complete population coverage of the UK between 1997 and 2016. The most probable 2011 Census ethnic group is inferred for each individual present in the Consumer Registers using the Ethnicity Estimator software developed by (Kandt & Longley, 2018).

2.3. Boundary data

In addition to the demographic data, the tool includes different administrative boundaries that might be used to frame analysis of all or part the entire UK and to function as spatial analysis units. We include a range of widely used geographies⁴, including 2011 UK Census Output Areas and their LSOA and MSOA aggregations, Local Authority Districts (LADs) and electoral ward boundaries.

3. Web mapping system design and implementation

3.1. The workflow for creating map-driven charts

A video demo of the basic interactions between user operations and system outputs is embedded into the website. It is introduced in the welcome dialog box when users open the website. Figure 1 illustrates the workflow developed for geodemographic profiling of either standard or bespoke functional regions.

The user starts from a high-level UK geographical point by specifying either a standard region, such as a Government Office Region, Local Authority District or Combined Authority or creates a bespoke area. When using pre-defined geographies, the user selects the relevant region from a high-level menu before choosing the desired geographical analysis units such as Output Areas, Local Authority Districts or Wards from a dropdown list in order to partition the entire selected region. Bespoke regions are created by dragging a rectangle over the area of interest or clicking on the constituent areal units in order to devise user defined areas. Once the study region and area of interest have been determined, a menu of available datasets and variables are presented as a dropdown list. The final step of the workflow is to extract and formulate data tables of these variables for visualisation. Up to two additional reference areas can be added for comparison by following the same steps.

3.2. Map design

The layout of the webpage on the browser side is a map-centric user interface, consisting of a self-hosted Leaflet map, a user control panel and a foldable pop-up panel. The user control panel on the left provides step-by-step guidance through the workflow of creating the geodemographic profiles demonstrated in Figure 1, which is implemented using the responsive Bootstrap user interface framework. Users can not only select from the dropdown menu but also operate the interactive map to devise their own study regions.

The maps and the regional data table for statistical chart visualisation are pulled from the tile server and PostGIS database on the server side, respectively. We use a self-hosted tile service to render the base map and overlays. Raster tiles and vector tiles are chosen for different layers to make the best use of both techniques based upon their advantages and disadvantages respectively. Charts of the chosen variables such as population structures by age bands and by ethnic groups and segregation index over time are plotted on the foldable pop-up panel.

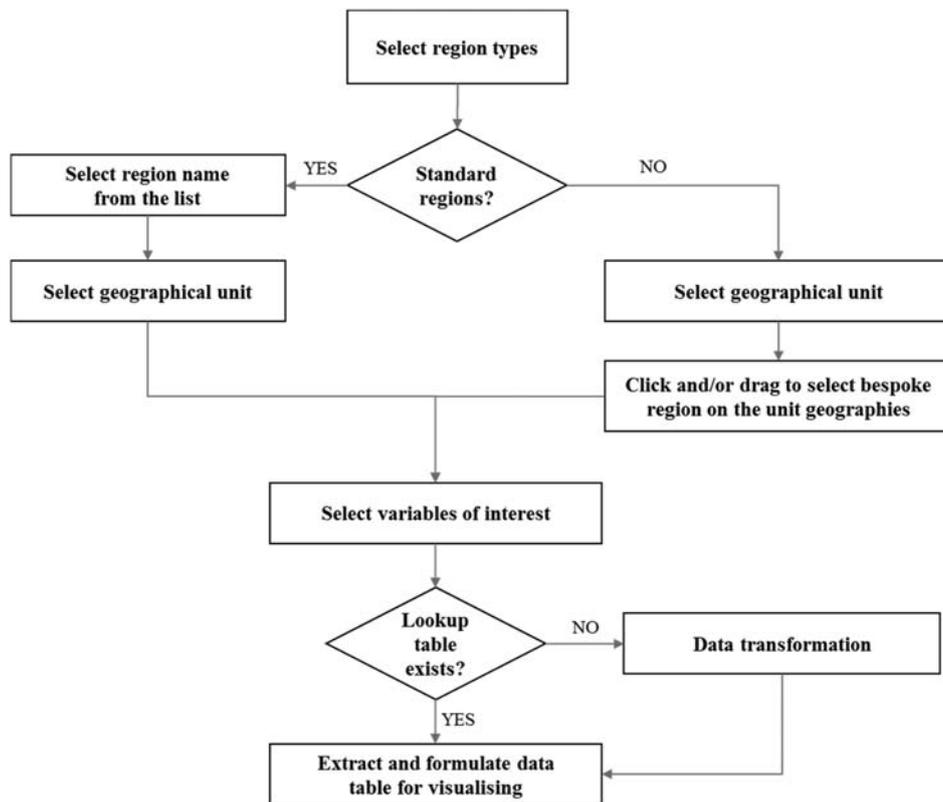


Figure 1. The workflow for creating geodemographic profiles of standard or bespoke regions

3.2.1. Base map

Raster tiles are employed to show static features in the background map, which provides users with geographical context information such as land mass, streets, water bodies, and labels. These geographical features are extracted from OpenStreetMap (OSM) data. We feed the OSM data into Mapnik to generate the raster tile images, based on cartographic rules such as colour, size, font and symbol that we specified in a Mapnik stylesheet. We only keep the minimal and relevant geographical features and make the base map in a low contrast, neat and non-obtrusive style that is ideal for a background map.

3.2.2. Overlays

Vector tiles are adopted to display the elemental analysis units of Great Britain, which are administrative and census polygon boundaries such as the OA, LSOA, Ward and LAD. The primary consideration is to enable spatial queries and interactions, which allows users to click or drag to select and unselect on the vector layers and to formulate their bespoke regions. The fill colours of polygons are changed dynamically to highlight the selected ones. It is also a more flexible and efficient way of map rendering. Vector tiles are normally smaller in file size compared to raster tiles covering the same geographical area. Additionally, the client side only needs to download tiles available in the viewport rather than having to fetch an entire vector data layer, which reduces the data volume

transferred from map servers. For instance, a full set of tiles of 227,759 Output Area polygons can be quickly loaded and displayed in a parallelised way.

3.2.3. Map-driven visualisation

The web mapping system provides geography students, researchers and policymakers analytical interactivity (Smith, 2016) by integrating maps with geodemographic charts. We use d3.js and dc.js to filter out subsets and create various statistical charts. Compared to the widely used traditional charts, there exist many novel ways of data representation and visualisation in the literature such as a Bivariate Map showing spatial variation and multiscale attribution view to explore scale-independent comparison (Zeng et al., 2020). However, our focus here is to enable users with different levels of background knowledge to create areal profiles and experiment with scale and aggregation effect.

Hence, we purposely keep the visualisations in their simple and popular forms that are readily intelligible to the widest range of audiences. For example, we follow the usual practice in demography, employing population pyramids to show the distribution of different age groups. Similar considerations have been made to use pie charts for ethnic compositions and line charts for time series of segregation index over time. With respect to the colour choices of the charts, we mostly make them consistent with the theme colour blue in lighter or darker tones. We

adopt colour palette advice from ColorBrewer⁵ where multi colours are needed to distinguish several qualitative categories.

On-the-fly calculations of a segregation index are implemented as an example to highlight the effects of scale and aggregation. Using individual level ethnicity data modelled from the Consumer Registers, we generate a series of population grid raster images from 1997 to 2016 for each of the selected Census ethnic groups. We use zonal statistics tool to get the group counts in each user-specified geography units and calculate the Dissimilarity Index (Lan et al., 2020b).

4. Visualisations and map use cases

We demonstrate the utilities of the web mapping system through three user case scenarios. Scenario 1 shows the comparison of geodemographic profiles of

different areas using an areal taxonomy. Taking the segregation index as an example, Scenario 2 illustrates how geodemographic profiles of the same region can be investigated across a range of geographic scales, using this web mapping tool. Scenario 3 depicts the creation of geodemographic profiles using bespoke, user-defined, regions.

4.1. Scenario 1: comparing profiles using elements of an areal taxonomy

The age structures and ethnic compositions of Christchurch and Oxford City as represented using MSOA geography are juxtaposed in the screenshot of the web toolkit shown in Figure 2. Population pyramids present the percentages of males and females by age bands respectively using blue and red bars, with the national average proportions in hatched bars in the background. It shows Christchurch as a silver

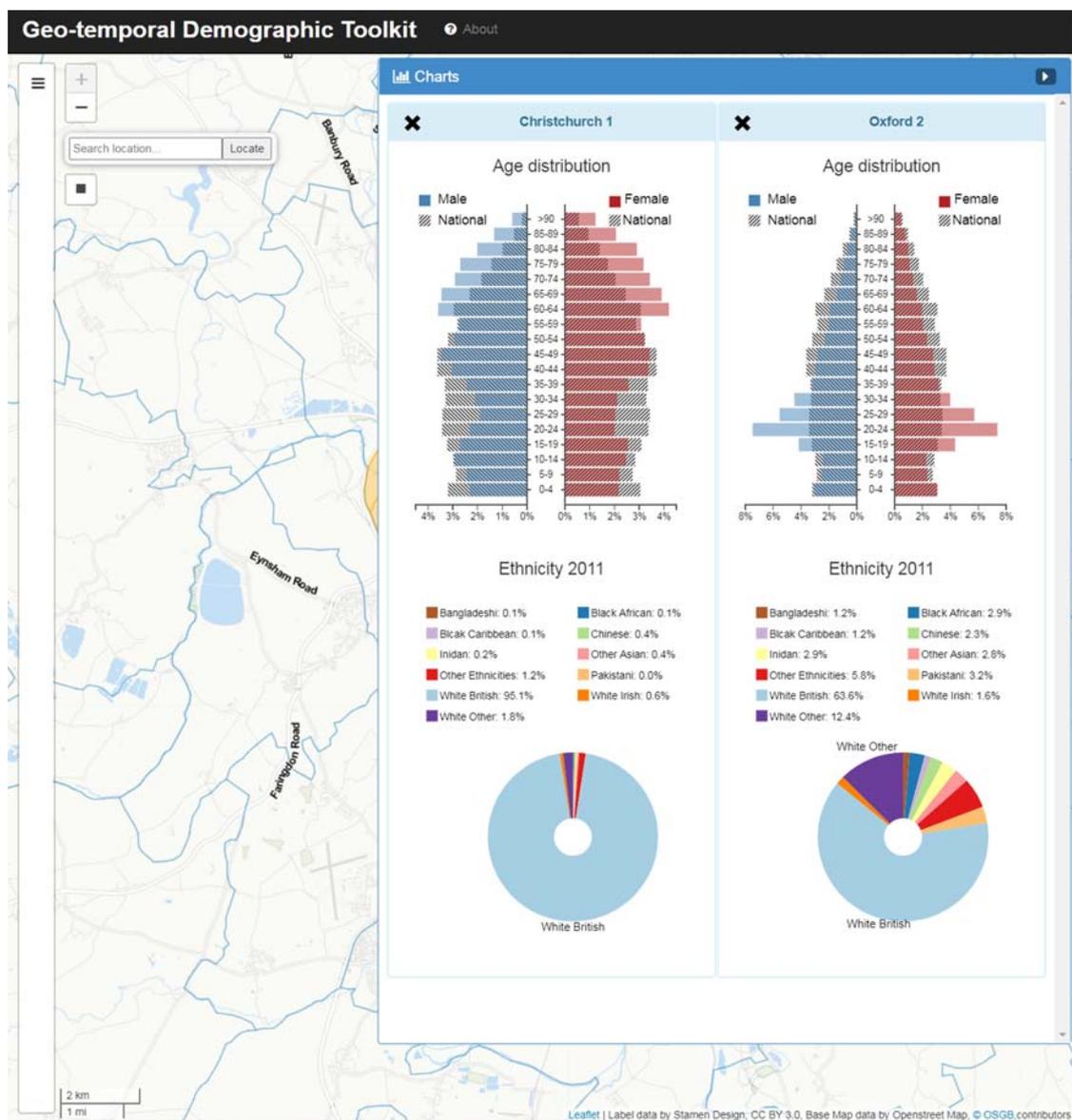


Figure 2. Geodemographic profiles of Christchurch (left column) and Oxford (right column) using MSOA aggregations

town of predominantly White British ethnicity, which has a notably larger share of senior residents than the national average, while the population of studentified Oxford City has a much younger population and more diverse ethnic structure.

4.2. Scenario 2: examining the scale and aggregation effects

The Greater Manchester Combined Authority (GMCA) area can be used to demonstrate the effects of scale and aggregation upon the recorded measurement of ethnic segregation using the dissimilarity index of Equation (1). The bar charts summarise the proportions of ethnic groups in 2016, and the line charts display changes in the dissimilarity index over time for individual ethnic groups at Ward level and at Local Authority District level (see Figure 3).

Recorded levels of segregation vary markedly between different geographical scales, with segregation more apparent at more scales such as Electoral Wards than for more extensive areas such as LADs – an observation that is consistent with the reported research findings (Lan et al., 2020a). The ordering of the degrees of segregation of different ethnic groups also changes with scale: for example, the third most segregated group at Ward level, Black Africans, surpasses the Pakistani group and becomes the second most segregated group at District level. Similar changes may be discerned for the Chinese and White British groups.

4.3. Scenario 3: creating profiles for bespoke regions

The UK Local Enterprise Partnership⁶ (LEP) zones are designed to boost regional development and improve

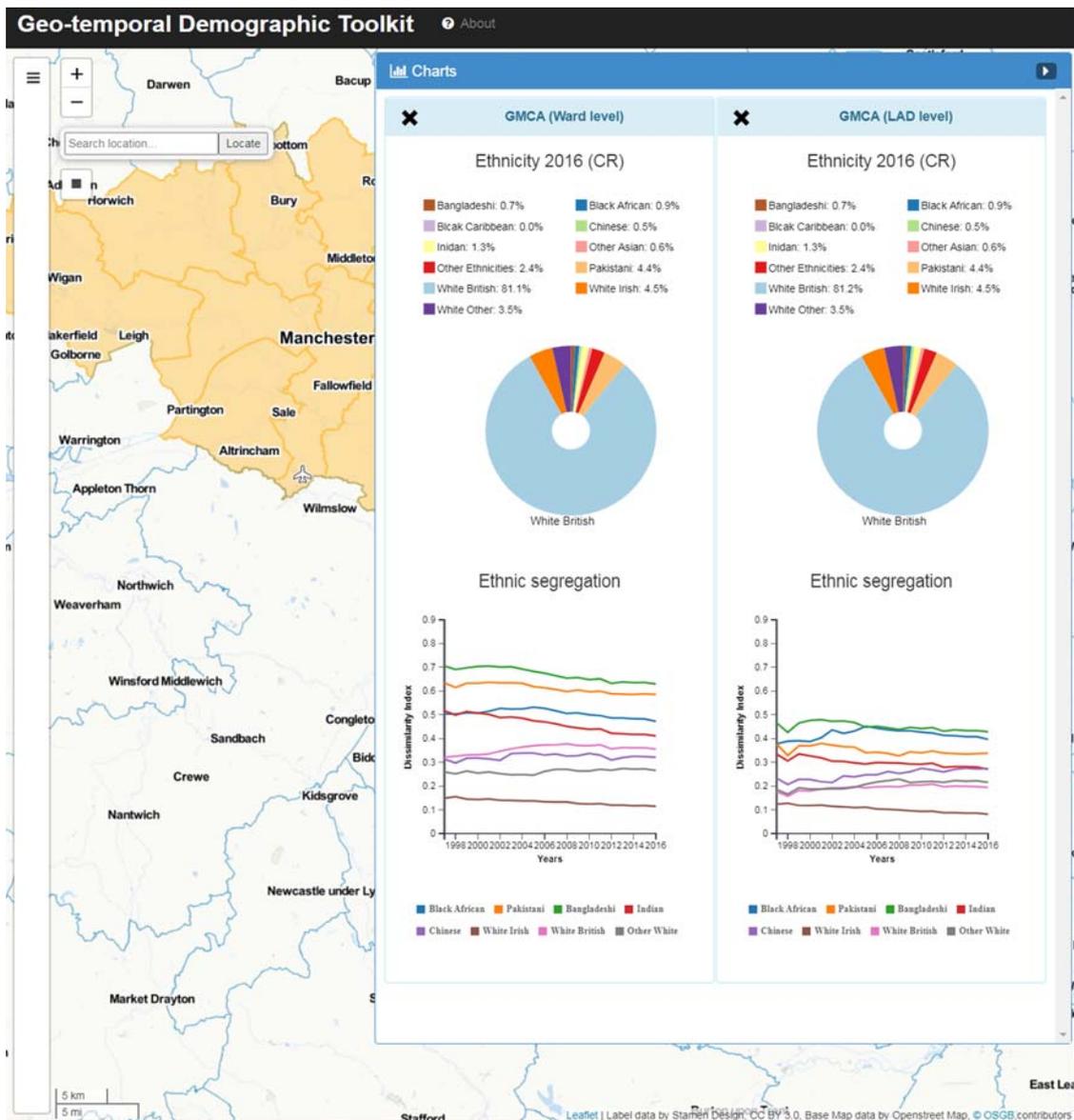


Figure 3. Ethnic compositions and segregation indices of Greater Manchester Combined Authority at Ward level (left column) and District level (right column)

infrastructural links between local authorities and local private sector businesses. They are ad hoc policy jurisdictions created in the UK for economic policy implementation. As such, LEP zones do not match established administrative boundaries, so a customised tool for aggregating lower level zones is useful for policymakers and researchers.

We use the creation of LEP zones as an example of profiling bespoke regions. The highlighted areas in the map in Figure 4 depict the geographical coverage of the West of England LEP zone, which includes the City of Bristol and the wider surrounding functional region such as Bath and Weston. The broad coverage suggests the city boundary of Bristol, which is one of the aggregated conventional statistical units, underbounds its current sphere of influence. This figure shows the population structure and ethnic segregation for the West of England LEP zone. It has a very balanced population in terms of its age and gender

profile, indicative of a high potential labour force participation rate.

5. Discussion and conclusion

Our web mapping system offers analytical insights through visualising and comparing geodemographic profiles of regions in Britain and illustrates a clear and strong relationship between areal indicators (e.g. segregation index) and spatial scales of investigation. Maps in this case serve as a spatial spine to integrate various demographic data sources and insightful socio-economic indicators. An emergent limitation of policy analysis is that datasets are insufficiently granular and analytical tools are not adequately flexible. Thus, for example, a UK government Green Paper on community integration (Ministry of Housing, Communities & Local Government, 2018) has called for more timely updated data and indicators

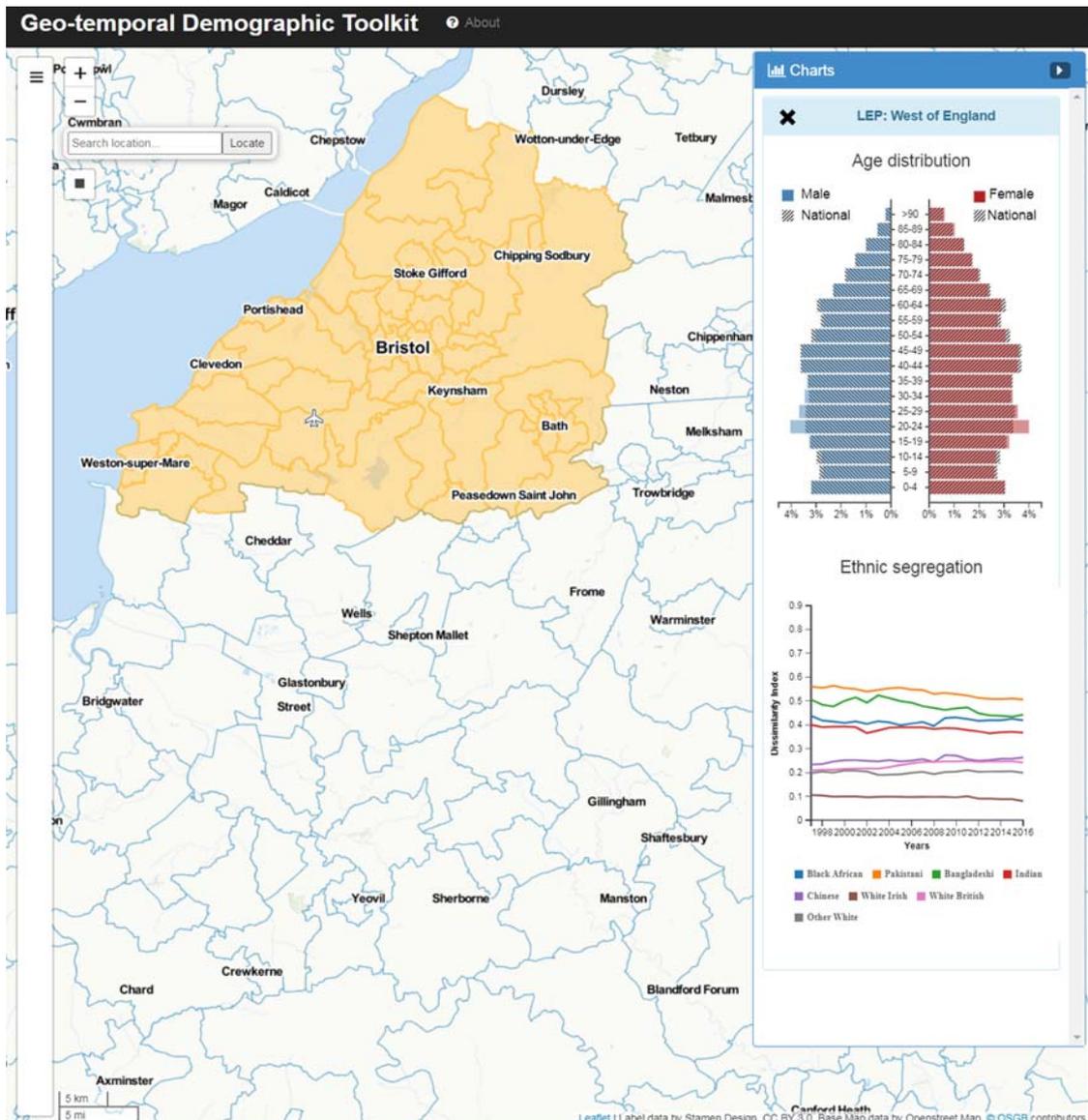


Figure 4. Geodemographic profile of the West of England LEP using MSAO level data

in order to evaluate the implementation of local policies and plans.

In this context, we have demonstrated the utilities of a novel geodemographic web mapping toolkit in addressing three interrelated problems – direct comparison, experimentation with scale and aggregation effects and creating bespoke regionalisation. We have illustrated that the web toolkit presents policymakers or researchers with geodemographic profiles of both standard and bespoke functional regions. We have shown how it allows users to compare relevant indicators between regions and to gain more insights on policy issues.

In addition, a highlighted feature of the web mapping system is a useful facility to examine the scale and aggregation effects inherent in areally-averaged statistics. To a certain degree, the capability of manifesting the MAUP by selecting different zoning and aggregating strategies is attributed to novel individual Consumer Registers, although this is rarely possible when data are personal. With the assistance of this web mapping tool, policymakers can better understand the effects of geographical scale in policy formulation. From a policy point of view, patterns of residential segregation do not remain constant across locations and scales. Therefore, scale-dependent policies of desegregation likely need extensive cooperation on land use planning, job market opportunities, and housing management among the government at multiple levels (Lan et al., 2020a). The changing ordering of ethnic groups with scales suggests group-specific strategies should be customised for the Black, Asian and minority Ethnic (BAME) groups at different scales.

As useful and versatile as it is, this geodemographic web mapping system nevertheless has several limitations that can be improved in future research. For instance, dual- or even multi-map comparisons can be a useful future iteration of the tool to enable users to draw comparisons among different areal unit maps together with their corresponding regional profiles. Availability of data at the level of the individual may not be realistically attainable, but the methods developed here might be used in tandem with microsimulation (Lomax & Smith, 2017) or other synthetic methods for grounding representation at the level of the human individual, in order to allow experimentation across a full range of spatial scales. This, however, requires consistency of assumptions used in both inductive and deductive modelling and the ever-present need for transparency of assumptions used throughout the workflow of database creation and maintenance. If these challenges can be addressed, we believe that the geodemographic web mapping toolkit can provide an informative and holistic supporting system for policymakers and researchers to make data-informed yet hypothesis led decisions.

Software

The front end of the web mapping system is built with HTML/CSS/JavaScript for the website, Leaflet.js for the mapping framework, PHP and AJAX for fetching data from servers and D3.js for the data visualisation. On the back end, Postgres and PostGIS are employed to store and manage the demographic data. Raster tiles are produced using the Python bindings of Mapnik C++ and vector tiles are generated with the Mapbox Tippecanoe. Both tiles are self-hosted using web server NGINX and TileServer GL. The population grid images are created using the ArcGIS 10.5 Point to Raster tool. The segregation indices are calculated using Python scripts.

Notes

1. <http://luminocity3d.org/>
2. <https://maps.cdrc.ac.uk/>
3. <https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeography>
4. <http://geoportal.statistics.gov.uk/>
5. <https://colorbrewer2.org/>
6. <https://www.lepnetwork.net/about-leps/location-map/>

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is supported by the EPSRC grant EP/M023583/1 (UK Regions Digital Research Facility) and the UK ESRC Consumer Data Research Centre (CDRC) grant (ES/L011840/1).

ORCID

Tian Lan  <http://orcid.org/0000-0002-3961-4742>

Paul Longley  <http://orcid.org/0000-0002-4727-6384>

References

- Armstrong, M. P., & Densham, P. J. (1990). Database organization strategies for spatial decision support systems. *International Journal of Geographical Information Systems*, 4(1), 3–20. <https://doi.org/10.1080/02693799008941525>
- Engin, Z., van Dijk, J., Lan, T., Longley, P. A., Treleaven, P., Batty, M., & Penn, A. (2020). Data-driven urban management: Mapping the landscape. *Journal of Urban Management*, 9(2), 140–150. <https://doi.org/10.1016/j.jum.2019.12.001>
- Fotheringham, A. S., & Wong, D. W. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A*, 23(7), 1025–1044. <https://doi.org/10.1068/a231025>
- Goodchild, M. F., & Longley, P. A. (2018). The practice of geographic information science. In M. M. Fischer & P.

- Nijkamp (Eds.), *Handbook of regional science*. (2nd ed, pp. 1107–1122). Springer.
- Kandt, J., & Longley, P. A. (2018). Ethnicity estimation using family naming practices. *PLOS ONE*, 13(8), e0201774, 1–24. <https://doi.org/10.1371/journal.pone.0201774>.
- Lan, T., Kandt, J., & Longley, P. (2020a). Geographic scales of residential segregation in English cities. *Urban Geography*, 41(1), 1–21. <https://doi.org/10.1080/02723638.2019.1645554>
- Lan, T., Kandt, J., & Longley, P. (2020b). Measuring the changing pattern of ethnic segregation in England and Wales with Consumer registers. *Environment and Planning B: Urban Analytics and City Science. Advance Online Publication*, <https://doi.org/10.1177/2399808320919774>
- Lan, T., & Longley, P. (2019). Geo-Referencing and mapping 1901 Census addresses for England and Wales. *ISPRS International Journal of Geo-Information*, 8(8), 1–14. <https://doi.org/10.3390/ijgi8080320>.
- Lansley, G., Li, W., & Longley, P. A. (2019). Creating a linked consumer register for granular demographic analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(4), 1587–1605. <https://doi.org/10.1111/rssa.12476>.
- Lomax, N., & Smith, A. (2017). Microsimulation for demography. *Australian Population Studies*, 1(1), 73–85. <https://doi.org/10.37970/aps.v1i1.14>
- Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2015). *Geographic Information Science and Systems* (4th ed.). John Wiley & Sons.
- Ministry of Housing, Communities & Local Government. (2018). *Integrated Communities Strategy green paper*. London <https://www.gov.uk/government/consultations/integrated-communities-strategy-green-paper>
- O'Brien, O., & Cheshire, J. (2016). Interactive mapping for large, open demographic data sets using familiar geographical features. *Journal of Maps*, 12(4), 676–683. <https://doi.org/10.1080/17445647.2015.1060183>
- Openshaw, S. (1984). *The modifiable areal unit problem geobooks*. University of East Anglia.
- Smith, D. A. (2016). Online interactive thematic mapping: Applications and techniques for socio-economic research. *Computers, Environment and Urban Systems*, 57, 106–117. <https://doi.org/10.1016/j.compenvurbysys.2016.01.002>
- Smith, M. J. (2005). The Journal of maps: An electronic journal for the presentation and dissemination of map based data. *Journal of Maps*, 1(1), 1–6. <https://doi.org/10.4113/jom.2005.39>
- van Dijk, J., & Longley, P. A. (2020). Interactive display of surnames distributions in historic and contemporary Great Britain. *Journal of Maps*, 16(1), 68–76. <https://doi.org/10.1080/17445647.2020.1746418>
- Walford, N. S. (2019). Bringing historical British population Census records into the 21st century: A method for geocoding households and individuals at their early-20th-century addresses. *Population, Space and Place*, 25(4), 1–16. <https://doi.org/10.1002/psp.2227>.
- Yue, Y., Lan, T., Yeh, A. G. O., & Li, Q.-Q. (2014). Zooming into individuals to understand the collective: A review of trajectory-based travel behaviour studies. *Travel Behaviour and Society*, 1(2), 69–78. <https://doi.org/10.1016/j.tbs.2013.12.002>
- Zeng, W., Lin, C., Lin, J., Jiang, J., Xia, J., Turkay, C., & Chen, W. (2021). Revisiting the modifiable areal unit problem in deep traffic prediction with visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 839–848. <https://doi.org/10.1109/TVCG.2020.3030410>