Visualising geographic information: improving interpretation through cartograms, hexograms and regular grids

Samuel Langton[*1] and Reka Solymosi[†2]

[1]Department of Sociology, Manchester Metropolitan University
[2]School of Law, University of Manchester

January, 2019

KEYWORDS: hexogram, cartogram, visualisation

Abstract

Background Thematic maps are an accessible and powerful method for displaying geographic information, but they can also introduce misrepresentation. Irregularly shaped polygons and large differences in the sizes of areas being mapped can mislead readers and result in spurious interpretations. Fortunately, methods are being developed in *R* to enhance the legibility of thematic maps by transforming the original geographic boundaries to shapes better suited for conveying the message of the researcher. Recent methods include area-based cartograms and hexograms (Harris et al., 2018) and uniform hexagonal or square grids (Bailey, 2018). The aim of this study is to examine to what extent these different methods of visualising geographic data affect people's interpretation of thematic maps, compared with original geographic boundaries. In this instance we test the ability of each method to communicate a situation where high values spatially cluster in small areas.

Data & methods Open data on voting totals for the United Kingdom EU referendum held June 2016 were used at Local Authority level in England. Since support for Remain was strongest in London, these data present a clear example of high value clustering in small areas (Greater London Local Authorities are small compared to those in the rest of the country). Thematic maps of voting results were produced using four methods of visualisation; area-based cartograms and hexograms were generated using open source code (see Harris, 2017) and uniform hexagonal and square grids were created using the geogrid package on the cartogram (see Bailey, 2018). These four options (see *Figure 1*) and the original map were then presented to a crowdsourced sample (N = 768) through an online questionnaire. Respondents were asked to rate the extent of agreement with the following statement, which was considered to be an accurate description of the data: "*High values (in yellow) appear to be clustered near one another, with a handful of outliers elsewhere in the country*". Respondents were given on a 5-point Likert scale (Strongly agree, slightly agree, neither agree nor disagree, slightly disagree, strongly disagree) upon viewing each option. Since the statement describes the data presented, more agreement with the statement constituted a more accurate representation of the data, meaning that the visualisation performs better in conveying the message as intended. The survey did not contain any information on the variable being mapped (to avoid any influence of prior familiarity with referendum results) or the method being used to create each map. All visuals were the same size and resolution and had the same aesthetic formatting.

Preliminary results Descriptive findings indicate that the area-based cartogram and hexogram, generated using open source code from Harris (2017), offer an improvement over the original map boundaries (see *Figure 2*). The options available in the geogrid package perform less well compared to the original boundaries. Recoding the Likert-scale to a binary response (agree-disagree), a logistic regression model confirmed these findings (see *Figure 3*). Using an area-based cartogram or hexogram, the odds that people agreed with the (true) statement were 3.8 and 2.6 times that of the original map respectively. Hexagonal and square geogrids reduced these odds by 0.27 and 0.28 respectively. These

[*] samuel.langton@stu.mmu.ac.uk
[†] reka.solymosi@manchester.ac.uk

coefficients were statistically significant at $p < 0.05$ alpha.

Conclusions There is evidence to suggest that thematic maps can lead to viewers drawing erroneous conclusions about geographic data. Augmenting spatial boundaries can offer a way to mitigate such effects. In the example of representing spatial clustering, area-based cartograms and hexograms can enhance message communication, but regular hexagonal and square grids may introduce further misrepresentation. Given that this study is the first to provide evidence either way, we recommend that in application, each method is tested, and a judgement made on the appropriate option, with consideration to the research questions and the data.

Figure 1 Proportion of Leave votes in 2016 EU referendum by Local Authority area. (a) original boundaries, (b) area-based cartogram with 29 bins, (c) area-based hexogram with 29 bins, (d) geogrid package with hexagons, (e) geogrid package with squares.

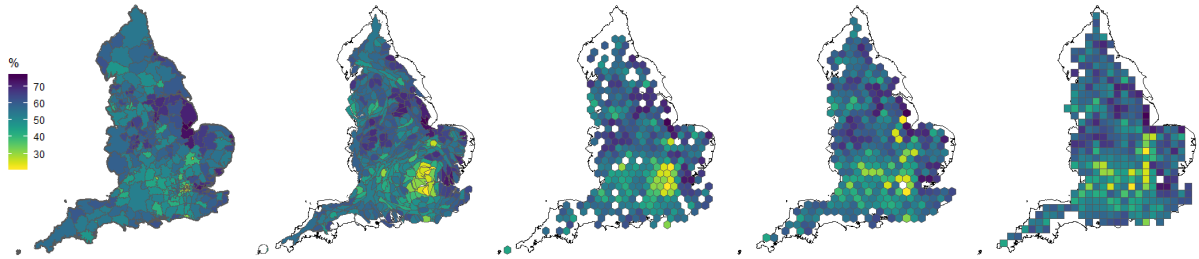


Figure 2 Descriptive results of Likert-scale responses on statement by map type

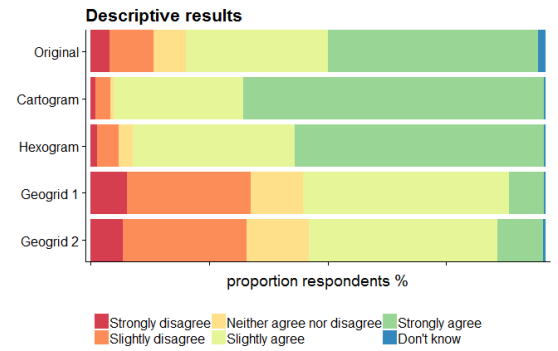
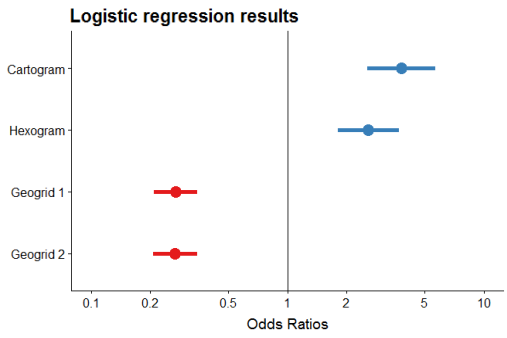


Figure 3 Odds ratios from logistic regression, response variable recoded from Likert-scale to binary agree (1) disagree (0); reference category is original map; 95% confidence intervals reported



References

Bailey J (2018). Geogrid: Turn Geospatial Polygons into Regular or Hexagonal Grids. *R package version 0.1.1*. https://cran.r-project.org/web/packages/geogrid/index.html

Harris R (2017). *Hexograms: better maps of area based data*. Retrieved from https://rpubs.com/profrichharris/hexograms

Harris R, Charlton M and Brunsdon C (2018). Mapping the changing residential geography of White British secondary school children in England using visually balanced cartograms and hexograms. *Journal of Maps*, 14(1), 65-72.