# Using hexograms to map areal data

Richard Harris[1], Martin Charlton[2] and Chris Brunsdon[3]

[1,4]School of Geographical Sciences, University of Bristol
[2,3]National Centre for Geocomputation, National University of Ireland Maynooth

January, 2018

**Summary**
This paper introduces hexograms as a method for mapping areal data. It builds on the idea of balanced cartograms that reduce geographic distortion, combining those with hexagonal binning to create non-tessellating tile maps of geographical distributions. The aim is to produce less geographically distorted representations of neighbourhood and other areal data than those resulting from conventional cartograms, whilst also avoiding the problem of invisibility found in traditional choropleth maps. The process behind the method is introduced with examples of its application. The code to reproduce some of the maps is available for R.

**KEYWORDS:** hexogram, cartogram, map design, GeoVisualisation

## 1. Introduction

This paper presents what we call hexograms to produce better maps of areal data. Hexograms are a cross between cartograms, hexagonal binning and tile maps that aim to tackle what has been described as the problem of invisibility in conventional choropleth maps and also the problem of distortion caused by cartograms (Harris et al., 2017b).

The motivation for this work is our interest in mapping socio-economic data at a neighbourhood scale, where neighbourhoods are represented by polygonal or lattice objects. The perennial problem when mapping areal data is the typically variable size of the areas. For a study region that includes places of both high and low population density, it is common for the map to be dominated by the largest areas (which are usually rural and contain the fewest people), leaving the smallest, urban and frequently city central areas as too small to be legible. To address this, a cartogram is often employed, rescaling the areas in accordance with some value other than their physical size; their total count of population, for example, or some measured attribute of the places. However, the cost of rescaling is to replace misrepresentation through invisibility with misrepresentation through distortion because the geography of the locations can be severely warped. Worse, this exchange is conducted on unfavourable terms because the cartogram, when dominated by places with a large scaling term, can also result in areas that are too small to see. In such cases the cartogram has failed to resolve the first representational problem and compounded it with the second.

Harris et al. (2017a) suggest addressing the problem head on. If some areas on the map are too small to be visible, then the obvious solution is to make them bigger. Doing this does not require them to be re-scaled against what is an intuitive yet often arbitrary value like population size, only to ensure that they are big enough to be seen. The result is a cartogram that balances visibility and distortion. An example is shown in Figure 1, which is the log of the average house price of local authorities in England in 2016 (the same data used by Harris et al, 2017b). The cartogram was produced with the

[1] rich.harris@bristol.ac.uk
[2] martin.charlton@nuim.ie
[3] christopher.brunsdon@nuim.ie

cartogram and GISTools libraries for R, and using the Dougenik et al. (1985) algorithm. The spatial clustering in London of high house prices is visible in the balanced cartogram but not the original choropleth map. Critically, this visibility is achieved without a severe amount of geographic distortion across the map overall.

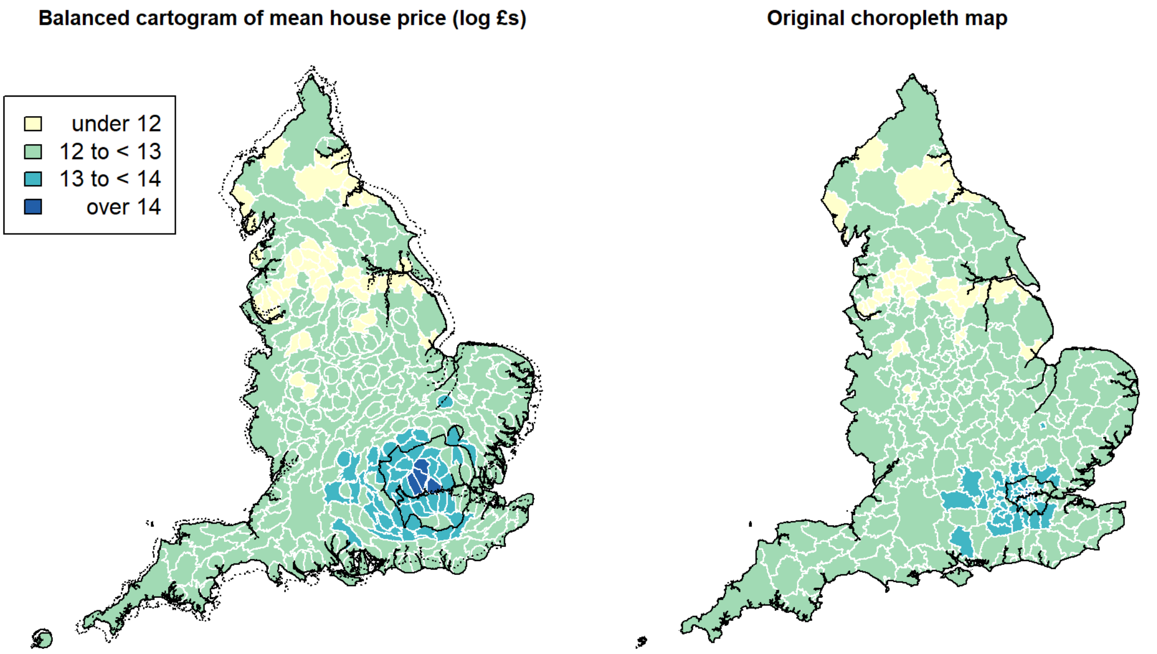


**Figure 1** Comparing a balanced cartogram with the original choropleth map of average house prices for English local authorities in 2016

## 2. Hexagonal binning and hexograms

The idea behind the hexogram is the same as for the balanced cartogram except that the minimum area is that which allows each place to be represented by its own hexagon in a process of hexagonal binning. Hexagonal binning is popular as a method of data visualization that can be used to map the density of point occurrences, such as the density of Walmart stores in the US: https://bl.ocks.org/mbostock/4330486. However, here we are interested not in mapping the density of some point outcome but in visualising the attributes of neighbourhoods in a manner akin to a choropleth map but where all the areas are visible on the map. We do not want multiple neighbourhoods to be represented by the same hexagon because this would obscure the geographical detail in places such as London. Instead, the map should be rescaled sufficiently to produce a cartogram where the centroid of each neighbourhood falls within its own, unique hexagon.

The process begins with an initial cartogram that increases the size of the smallest areas in the map. Hexagonal binning is then applied to the centroids of the cartogram, using R’s fMultivar package, identifying areas that conflict (share a hexagon). To address these conflicts, an attempt is made to move the centroids to different parts of the conflicted areas to see if that will separate them. If not, the areas are enlarged some more. The process iterates through these procedures but as a last resort will attempt a third approach, which is increasing the number of bins.

Examples of the output are shown below. Increasing the number of bins reduces the amount of geographic distortion but reduces the visibility of the hexagons. With 118 bins there is no geographical distortion but the hexagons are too small to be legible. Instead, Figure 2 gives an example with 29 bins and another with 48. With some careful re-arrangement of the hexagons (returning back to their original positions those that are towards the edge), a non-tessellating tile map can be produced and overlaid upon the initial choropleth map, as shown in Figure 3.

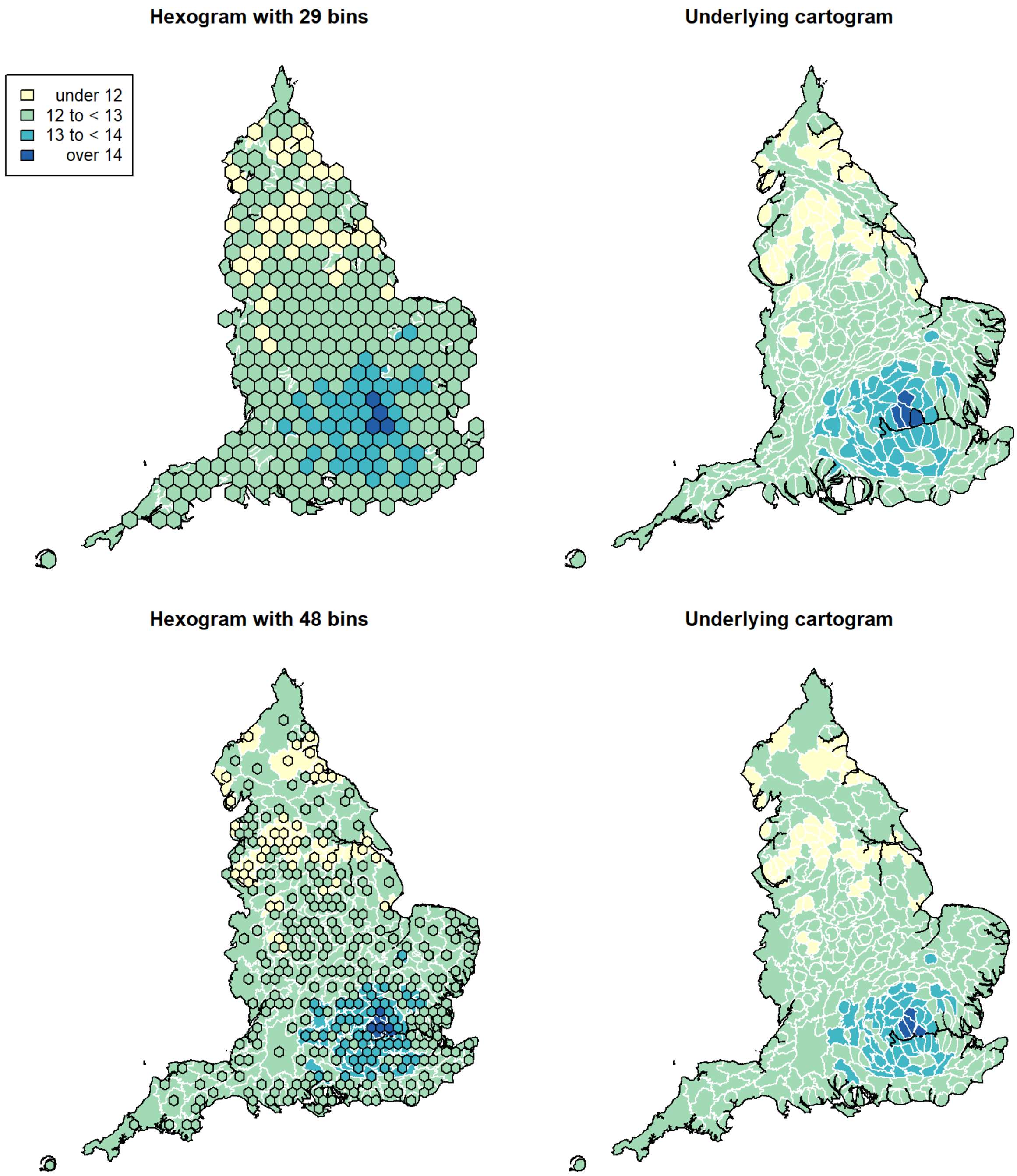


**Figure 2** Changing the number of bins changes the size of the hexagons and the amount of geographic distortion required

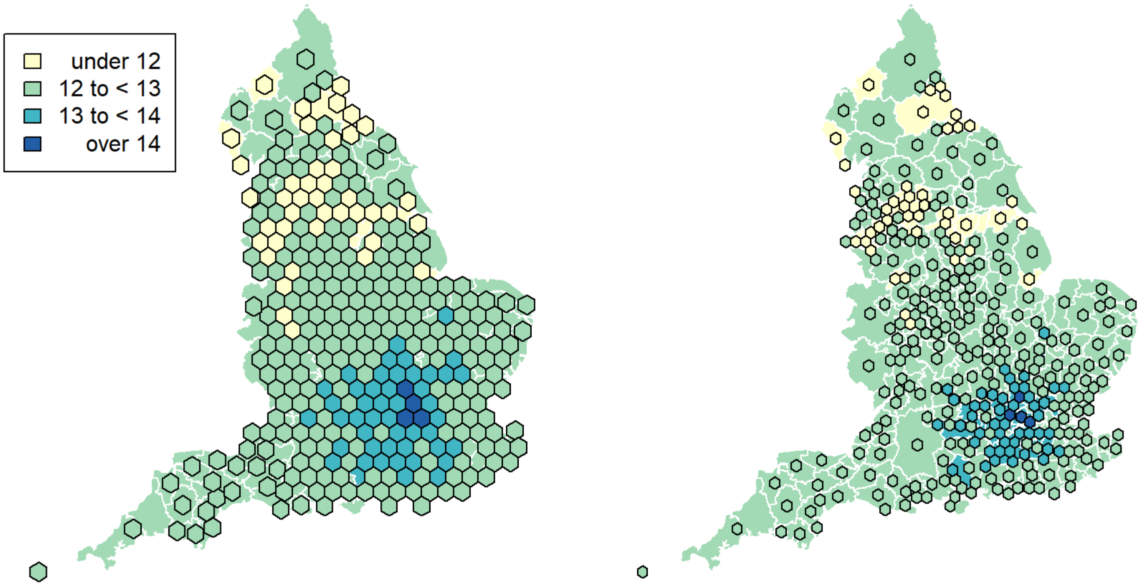


**Figure 3** Non-tessellating tile map created by re-arrangement of the hexogram overlaid upon the original choropleth map

## 3. Further examples

A computationally more demanding example is shown below. It involves mapping the ethnic composition of the residential population of each of 4662 Lower Level Super Output Areas in London, using the 2011 Census data. Taking as an example the percentage of each neighbourhood population that is Bangladeshi, Indian or Pakistani, the standard choropleth map (Figure 4) has the usual problem of invisibility whereas the hexogram (Figure 5) offers a clearer representation of the geography. The map, which looks like a density surface but isn't (there is no interpolation nor smoothing of the data), hints at the patterns of residential clustering more closely studied by Johnston et al. (2016).

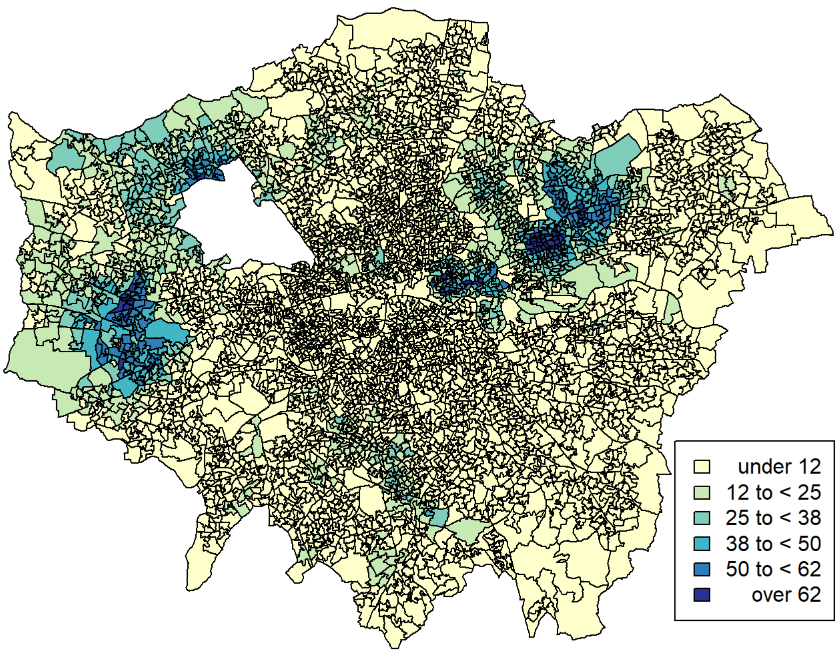


**Figure 4** Choropleth map of the percentage of the population that is Bangladeshi, Indian or Pakistani in each 2011 Census neighbourhood

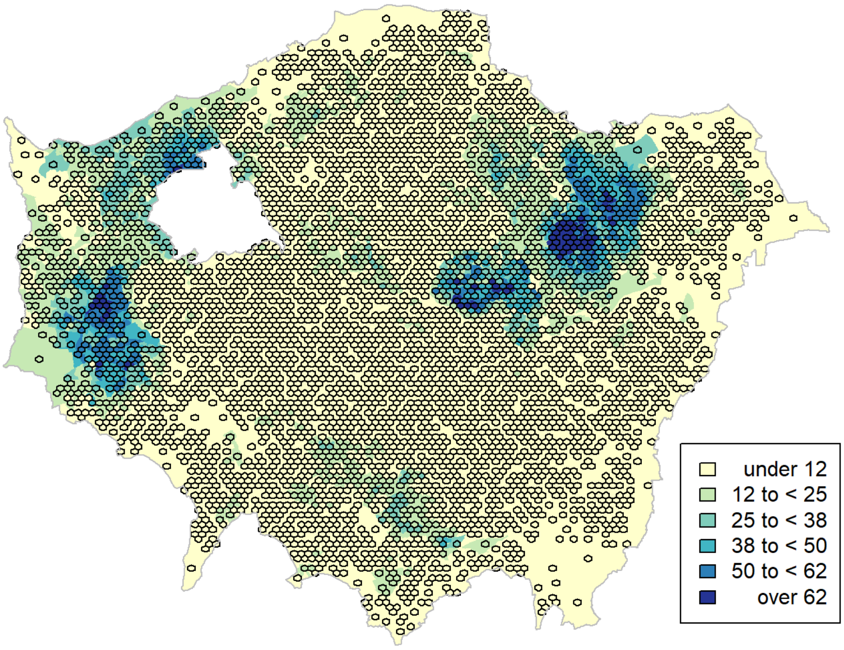


**Figure 5** Hexogram alternative to Figure 4

A final example is the US Presidential results for the 2016 election, which are mapped in Figure 6. The inspiration for this is a map published by the Financial Times, which can be viewed at [https://ig.ft.com/us-elections/results](https://ig.ft.com/us-elections/results). The designers of that newspaper's map took an approach similar to the balanced cartogram whereby only those parts of the map that needed to be enlarged were so. The aim was to avoid the levels of geographic distortion associated with cartograms and tile maps of the election results whilst at the same time giving sufficient space on the map to each state in order to display its result and the number of votes cast in the electoral college. A hexogram version is displayed below.

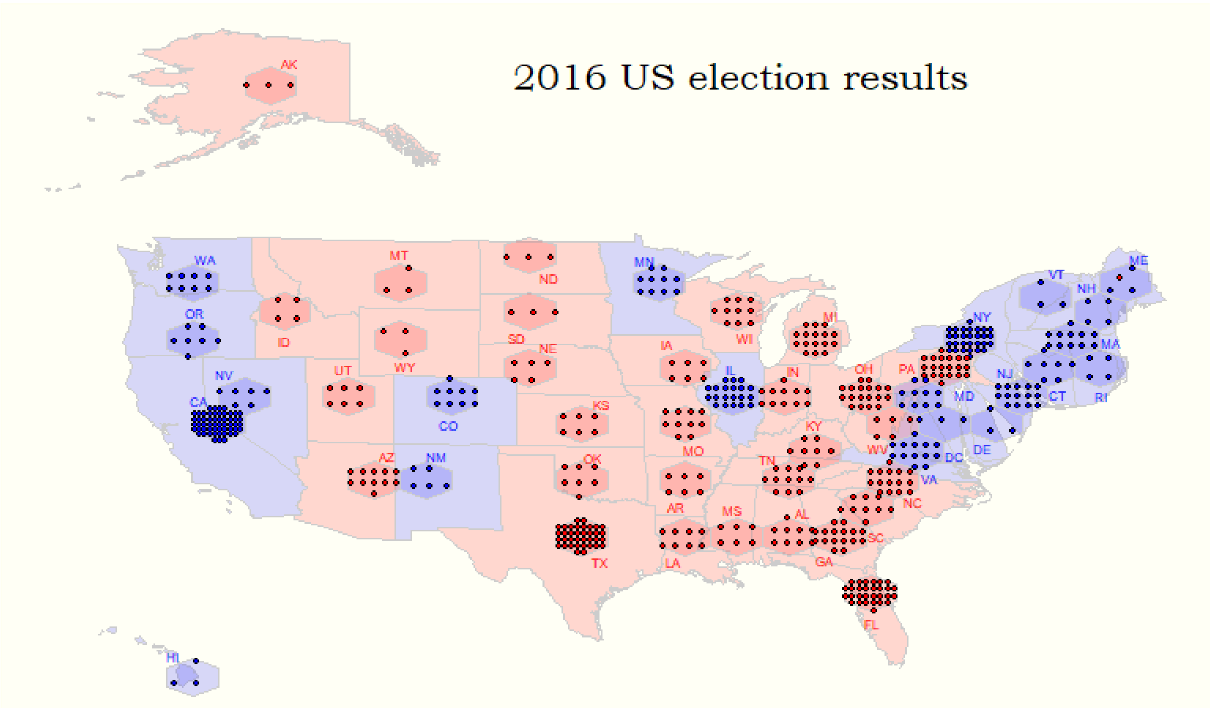


**Figure 6** Hexogram-based map of the 2016 US Presidential results

## 3. Conclusion

Maps of neighbourhood characteristics are a valuable tool for communicating the geographical outcomes of socio-spatial processes but are an arguably taken-for-granted part of urban analytics and of 'visualizing the city' (Singleton et al., 2018). Traditional choropleth maps are often ineffective at visualising geographic detail when dominated by a small number of large areas yet substituting them with cartograms does not automatically resolve the representational issues encountered. What does is to better balance the twin problems of invisibility and distortion they bring. To that end, hexograms have been presented as a hybrid of cartograms, hexagonal binning and tile maps, with the idea that these can produce better and more attractive visualisations of neighbourhood data. The methods have been implemented in R. The code and a tutorial is available at http://rpubs.com/profrichharris/342278.

## 4. Acknowledgements

The research was funded under the ESRC's Urban Big Data Centre (http://ubdc.ac.uk/), ES/L011921/1.

## 5. Biographies

Rich Harris is Professor of Quantitative Social Geography at the School of Geographical Sciences, University of Bristol.

Martin Charlton is Senior Lecturer at the National Centre for Geocomputation, National University of Ireland Maynooth.

Chris Brunsdon is Professor of Geocomputation and Director of the National Centre for Geocomputation, National University of Ireland Maynooth.

## References

Gastner MT and Newman MEJ (2004). Diffusion-based method for producing density-equalizing maps. *Proceedings of the National Academy of Sciences*, 101, 7499–7504.

Harris R, Charlton M, Brunsdon C and Manley D (2017) Balancing visibility and distortion: Remapping the results of the 2015 UK General Election. *Environment and Planning A*, 49, 1945-1947.

Harris R, Brunsdon C, Charlton M and Manley D (2017) Tackling the Curse ofCartograms: addressing misrepresentation due to invisibility and to distortion. Presented at *GIS Research UK conference*, Manchester. Available at http://huckg.is/gisruk2017/GISRUK_2017_paper_29.pdf. [Accessed Jan 3, 2018]

Johnston R, Jones K, Manley D and Owen D (2016) Macro-scale stability with micro-scale diversity: modelling changing ethnic minority residential segregation – London 2001–2011. *Transactions of the Institute of British Geographers*, 41, 389–402.

Singleton AD, Spielman SE and Folch DC (2018) *Urban Analytics*. London: Sage.