Visualising small area data with hybrid cartograms and hexograms

Richard Harris*[1]

[1]School of Geographical Sciences, University of Bristol

GISRUK 2025

Summary

Hexograms, first proposed by Harris (2018), mix hexagonal grids with cartograms, redressing problems of 'invisibility' in conventional maps, and geographic distortion in cartograms. The idea is simple: enlarge areas needing to be observed, but not so much that place geographies become unrecognisable. This paper develops the approach, combining a Dorling-style cartogram (Dorling, 1996) of local authorities with a standard map to create a hybrid cartogram; then using hexagonal grids and spatial optimisation to map neighbourhood-level variations. The resulting hexogram provides a representation of ethnic diversity in England and Wales that improves legibility but limits geographic distortion of and within the authorities.

KEYWORDS: cartogram, hexogram, visualisation, neighbourhood data, ethnic diversity

1 Introduction

In earlier papers, Harris et al. (2017) and Harris et al. (2018a) question the effectiveness of cartograms for mapping spatial distributions. By scaling areas in proportion to population size or some other value of interest, classically the cartogram enlarges places on the map that are small in area but large in population, count, or rate. Usefully, this tends to draw attention to areas of high population density, typically urban, that might otherwise be 'invisible' (too small) on the map to read. However, the resulting distortion contorts the geography, often severely, sometimes rendering areas of lower population density 'invisible' or unrecognisable instead. Employing a cartogram risks trading one visualisation problem for another.

The problem is not the cartogram *per se*. Sometimes the distortion makes a useful and powerful point, emphasising and challenging perceptions about how much of something is found in a particular place. The more specific issue is about using a cartogram when it is not fully required. If the primary need is only to make the measured attribute legible across the map, then it is sufficient to enlarge areas enough for that information to be seen, without recourse to strict proportionality to population size or some other scaling variable.

*rich.harris@bristol.ac.uk

For example, if English and Welsh local authorities (LAs) are mapped in a graphic of size suitable for journal publication, then London's occupy very little space on that map. Consequently, it is hard to discern their colour/shading and therefore what that represents. Scaling by population size enlarges those and other LAs, but if, for example, a contiguous area cartogram is used, it will squeeze and distort rural LAs in the process of making the urban ones bigger (there are examples of this in Figure 2, below). A more visually appealing map is obtained by scaling by the square root of each LA's area (Harris, 2016) – more generally, by applying a suitable transformation to the distribution of area values; or by setting a minimum scaling value to all places that fall below an interpretable unit size on the map. This is what Harris et al. (2017) call a balanced cartogram. In the case of Harris's (2018) original suggestion for hexograms, the size is that sufficient to contain a cell from a hexagonal grid (see Figure 1).

In each of the suggestions above, the result is not *sensu stricto* a cartogram, but it does better balance invisibility and distortion. This paper develops the idea through the application of hybrid cartograms and hexagonal grids to represent the ethnic diversity of neighbourhoods in England and Wales.

2 Hybrid cartogram

Figure 2 plots the ethnic diversity of English and Welsh LAs using 2021 Census data, an entropy index, and maps of various design. The details of the measurements are not important. Of relevance are their visual representations.

The choropleth map preserves the geography (land mass) of England and Wales, but whether this is a useful design feature depends upon the purpose. To a degree, it does show *where* things are: London and other cities are more ethnically diverse than many other parts of the two countries. But does it show *how* things are? Pick an area at random from the data and it may not be especially diverse; pick a person at random and there is a good chance they will be among the many living in diverse places. Arguably, by expanding the higher population areas, each of the contiguous and non-contiguous (Dorling) cartograms better represent the lived experience of a greater number of the population. Of those two, the latter is more readily combined with the original map, creating a hybrid cartogram where some areas are enlarged and shown as circles, but others draw upon their real-world shape.

3 Hexogram

Each of the maps in Figure 2 reveals geographic variation in the ethnic diversity of LAs but at a somewhat coarse scale. What of the variation within them? At the Lower Super Output Area (LSOA) scale there are 35,671 neighbourhoods, a number that is more than 100 times greater than the 316 LAs, and one that is not easily amenable to creating a cartogram of them. Even if it was computationally tractable, the geographic distortion is likely to be severe.

Instead, a hexagonal grid is overlain upon each of the LAs, in turn, within the hybrid cartogram, where the number of cells in each grid is greater than or equal to the number of LSOAs in the LA.

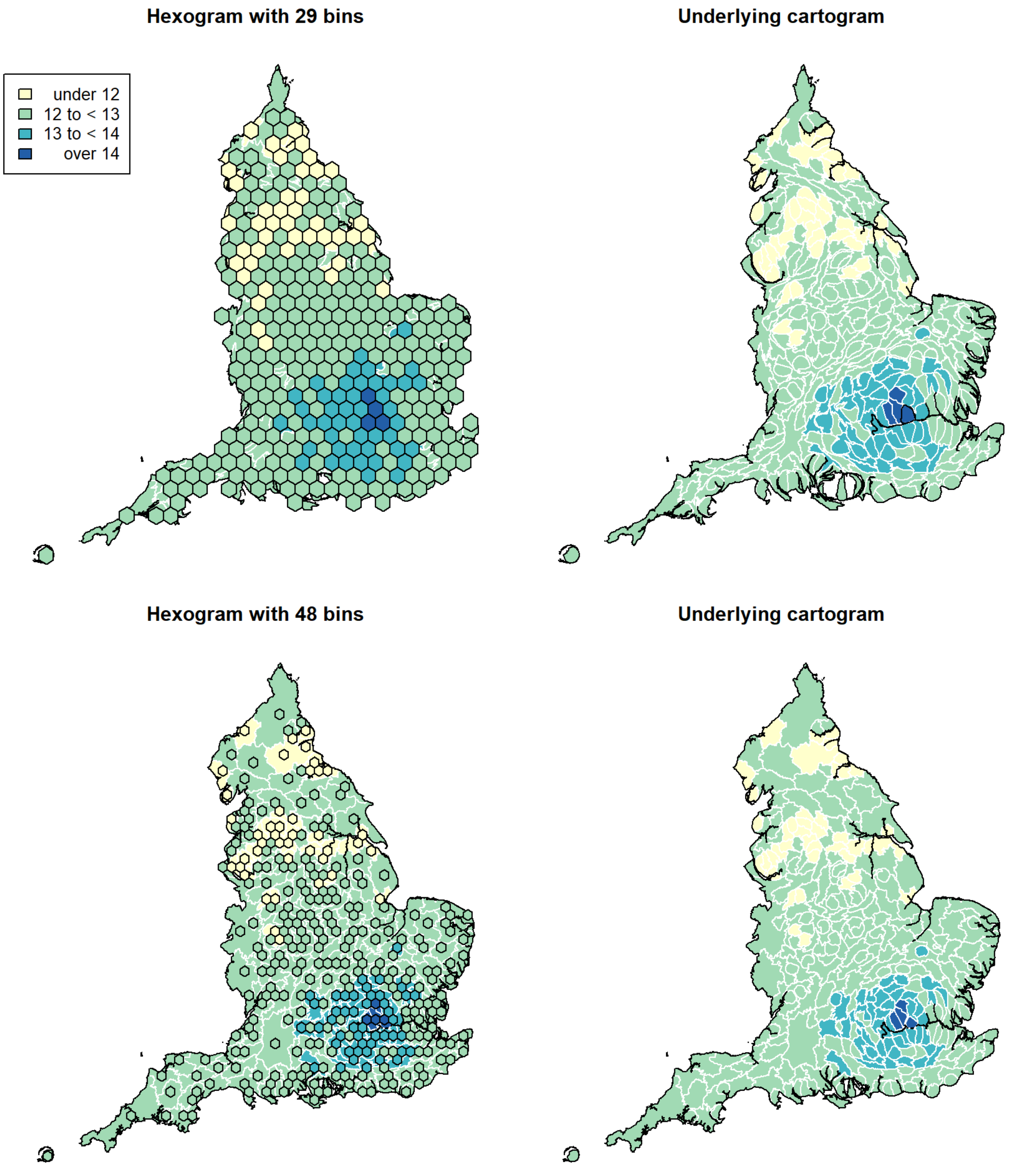

Figure 1: Examples of hexograms from Harris et al.'s (2018b) GISRUK paper (average house price per local authority, in log(£s)).

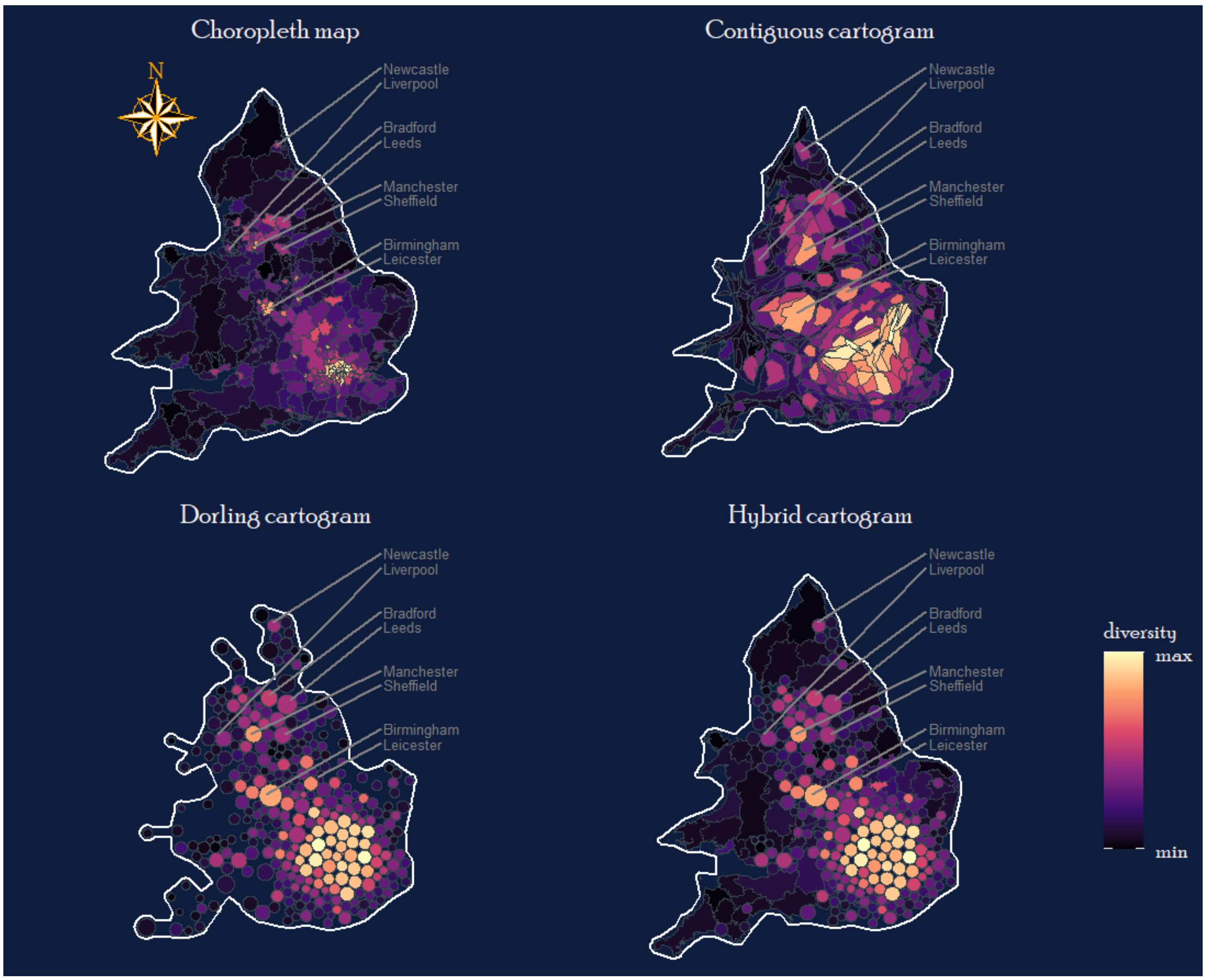

Figure 2: Maps of ethnic diversity, measured for English and Welsh local authorities. The data are the same in all cases.

Spatial optimisation is then used to 'reproject' the centroids of the actual LSOAs to a unique cell within the grid, minimising the cost constraint of the square of the distance between the origin and the destination.

This approach works better for a city such as Birmingham, where the area of the LSOA at the third quartile is twice as large as that at the first, than it does for, say, Cornwall, where the ratio of the third to first quartile is 22. That greater variation in the areas of its LSOAs would matter less if Cornwall was also divorced from its true shape and, as with, Birmingham, represented as a circle in the hybrid cartogram. However, it isn't; the net result being that spatial patterns in the data for Cornwall become displaced and moved along to identifiable but wrong parts of the coastline.

The solution is to return to the principle of the hybrid cartogram, combining the hexagonal cells for the smaller LSOAs with the real-world shapes of the larger, as shown in Figure 3.

Bringing the various processes together, the complete hexogram, which represents the geography of every LSOA, grouped into their LAs, is shown in Figure 4. It is not, of course, the 'true' geography of the choropleth map, but nor is that geography greatly distorted. The Moran's I coefficient of the diversity scores for each LSOA and its six nearest neighbours (similar to first-order contiguity for much of the map) is 0.92 for the choropleth map and 0.90 for the hexogram. The Pearson correlation between the two maps' sets of local Moran values is 0.99. This and a visual comparison of Figure 3 suggest that the key difference in the maps is not in the spatial patterning they display (not in where spatial clustering and spatial heterogeneity can be identified), but only in the attention they give to LAs that are smaller in area but not population size, and often, but not exclusively, contain the most ethnically diverse neighbourhoods.

The 'hot' (H-H: High-High) and 'cold' (L-L: Low-Low) places of ethnic diversity are shown in Figure 5, calculating from the true geography, omitting the local Moran values that are not statistically significant at a 95 percent confidence interval, having adjusted for a false discovery rate, and also omitting the extremely rare cases of significant High-Low and Low-High values. The hot spots of London and Birmingham are evident in both maps, but the hexogram better emphasises that there are other hot spots elsewhere.

4 Conclusion

The paper sets out a method for creating hexograms to visualise small area data that aims to reduce the geographical distortion associated with conventional cartograms and retain geographical cues on the map. It follows Harris et al.'s (2017) and Harris et al.'s (2018a), conception of a hexogram as a visual combination of hexagons and cartograms. However, the method presented here differs from those used previously by employing a Dorling cartogram instead of a contiguous cartogram as the foundation for the hexogram. The reasons for the change include it being computationally faster, more aesthetically pleasing, and relatively straightforward to combine with an original base map to create a hybrid cartogram. However, it does not retain typology. It also extends the original method by using hexagonal grids to map spatial variation within the initial units of analysis (local authorities in the case study presented). The total time to produce the hexogram is about 10 minutes using hacky code and an ageing brain and laptop. That code presently is too crude to share with anyone

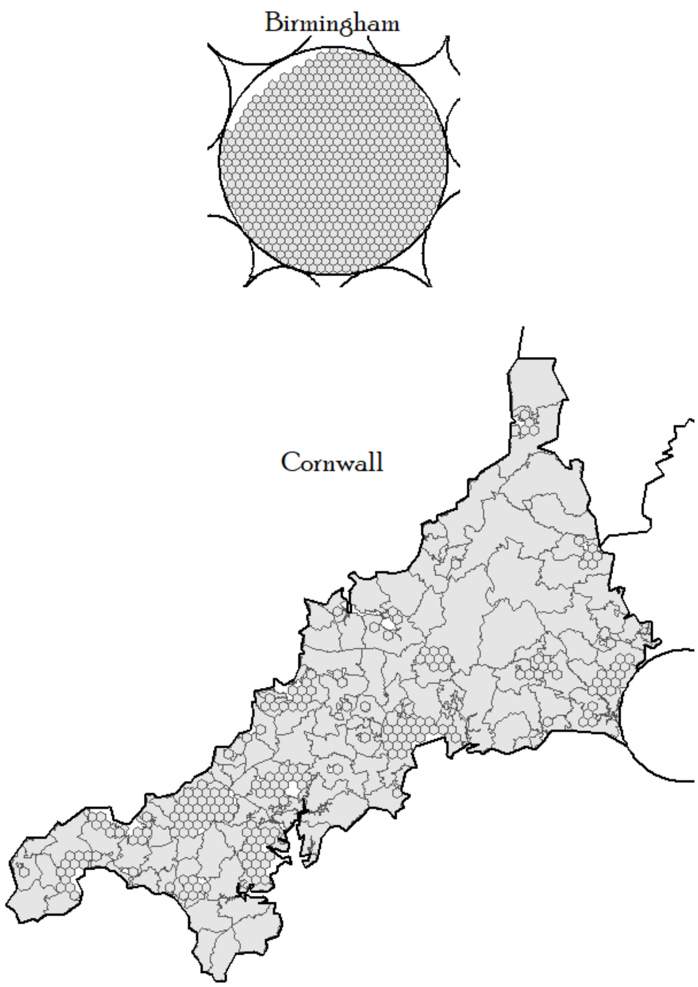

Figure 3: The hexagonal and hybrid representation of LSOAs within Birmingham and Cornwall respectively

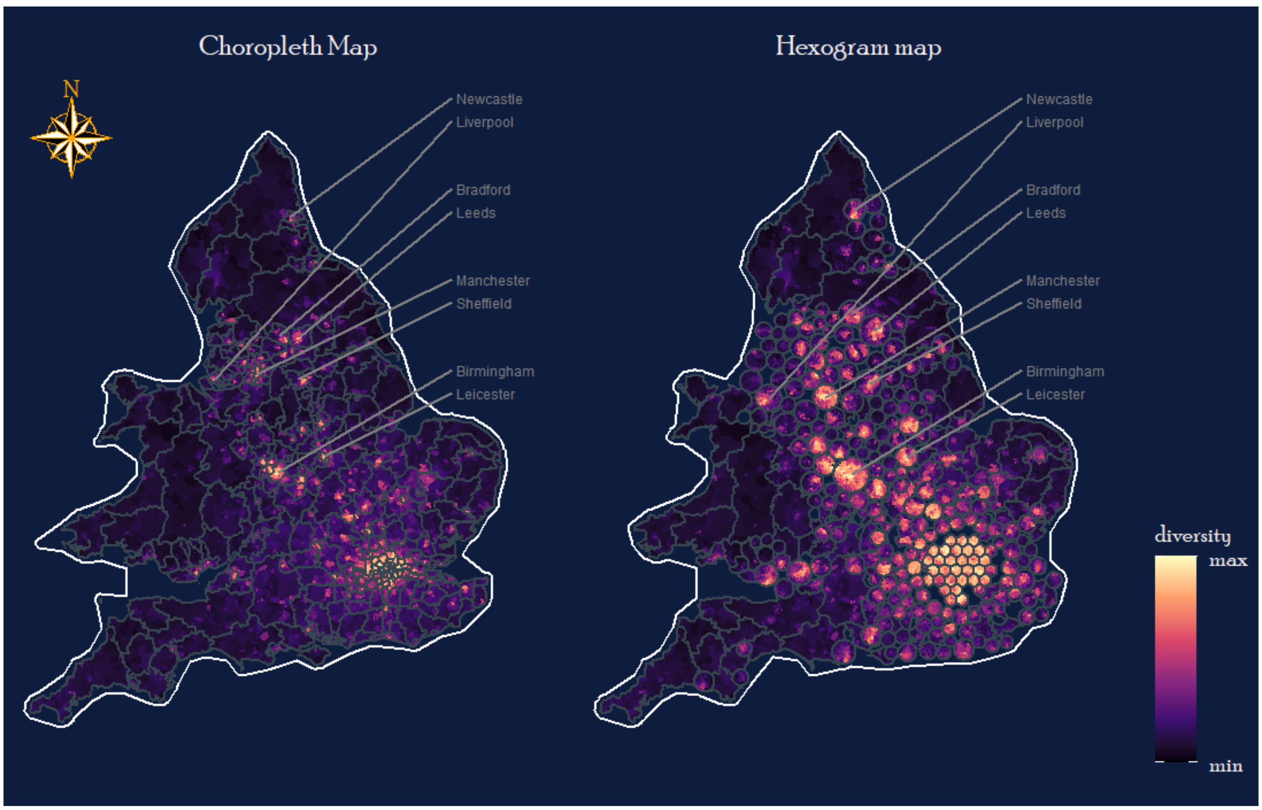

Figure 4: Comparing a standard choropleth map with a hexogram map of neighbourhood level ethnic diversity.

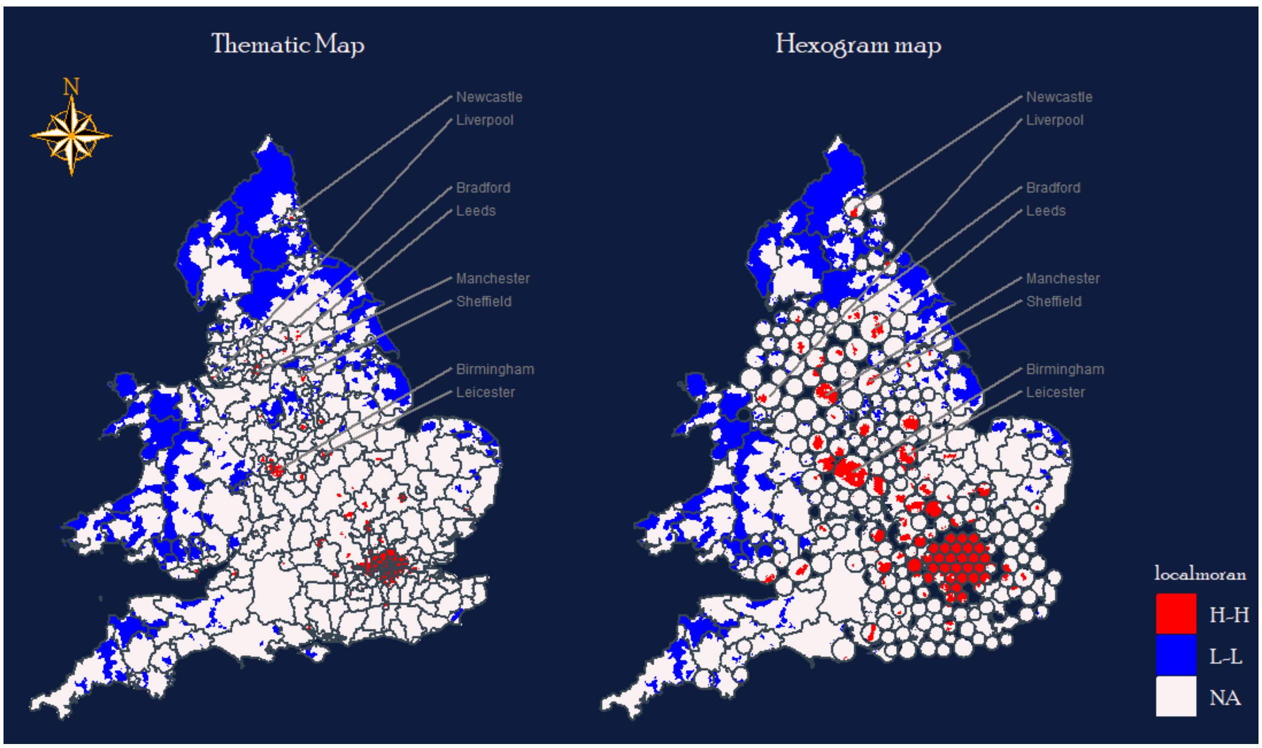

Figure 5: Showing the 'hot' and 'cold' spots of ethnic diversity using a standard thematic map and a hexogram.

but will be developed over time at https://profrichharris.github.io/hexograms.

In a study of how spatial information is understood by readers of standard cartograms, Langton and Solymosi (2021) found that the loss and distortion of the areas' boundaries introduces misrepresentation and misunderstanding. They find that these problems can be mitigated by using balanced cartograms and hexograms. Whether the perceptual and cognitive advantages also apply to the hexogram presented here remains to be tested. However, Moran values suggest it is successful in retaining the detailed spatial patterns in the data it is visualising.

5 Biography

Rich Harris (https://profrichharris.github.io) has been a member of the School of Geographical Sciences as man and boy for 57.8 per cent of his life and rising. This suggests he is not a geographer who gets out much, so it is handy that GISRUK has come to him this year.

References

Dorling, D. (1996). *Area Cartograms: Their Use and Creation (CATMOG 59)*. University of East Anglia: Environmental Publications.

Harris, R. (2016). *Quantitative Geography: the basics*. London: Sage.

Harris, R. (2018). Hexograms: Better maps of area based data `https://rpubs.com/profrichharris/hexograms`.

Harris, R., Charlton, M., & Brunsdon, C. (2018a). 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, `https://doi.org/10.1080/17445647.2018.1478753` `https://doi.org/10.1080/17445647.2018.1478753`.

Harris, R., Charlton, M., & Brunsdon, C. (2018b). Using hexograms to map areal data `https://www.dropbox.com/scl/fi/j20hfgj9zd40wuepy8kh6/hexograms.pdf?rlkey=e1ojn1jglfjfqx16wc2kydqj9&dl=1`.

Harris, R., Charlton, M., Brunsdon, C., & Manley, D. (2017). Balancing visibility and distortion: Remapping the results of the 2015 uk general election. *Environment and Planning A: Economy and Space*, *49*(9), 1945–1947, `https://doi.org/10.1177/0308518X17708439` `https://doi.org/10.1177/0308518X17708439`.

Langton, S. H. & Solymosi, R. (2021). Cartograms, hexograms and regular grids: Minimising misrepresentation in spatial data visualisations. *Environment and Planning B: Urban Analytics and City Science*, *48*(2), 348–357, `https://doi.org/10.1177/2399808319873923` `https://doi.org/10.1177/2399808319873923`.