OPEN ACCESS

University of BRISTOL

Johnston, R., Jones, K., & Manley, D. (2019). Why geography matters. *Significance*, *16*(1), 32-37. https://doi.org/10.1111/j.1740-9713.2019.01234.x

Peer reviewed version

Link to published version (if available):
10.1111/j.1740-9713.2019.01234.x

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

Why geography matters

To understand society, we can analyse administrative data – and these data are often collated by geographical area. The choice of geographical area frames how we see the world, but a poor choice of frame will present a misleading perspective. Ron Johnston, Kelvyn Jones and David Manley outline three problems, and potential solutions

We live in a messy, complex world. To understand its complexities, policymakers and academics increasingly look to collate and analyse information from the vast administrative data sets held by government departments and agencies.

Administrative data typically concern individuals, but the data are aggregated into groups for the purposes of analysis. Researchers are not so much interested in individuals as in groups of people who share some of the same characteristics. For example, a study might look at differences in the income of people of similar ages or educational background. But often, in seeking to understand our societies, data will be aggregated and analysed based on the geographical areas – the "areal units" – in which people live or work.

Analysing data in this way can shed light on important topics, such as differences in the health, wealth and happiness of people living in neighbouring streets, towns, and districts. But it is important to recognise that our understanding of an issue depends on the areal units we choose, or are forced, to use.

These geographical frames determine how we the world. But many administrative data sets have several geographical disadvantages, which impact on their interpretation. Three that were identified in a 1961 book – *Statistical Geography: Problems in Analyzing Areal Data*, by Duncan, Cuzzort and Duncan[1] – remain important today. Tackling these disadvantages presents substantial challenges to data producers and users, but the proper analysis of administrative data to address social, economic, cultural and political questions requires that we do.

Arbitrary areas

> "Many researchers on areal differentiation are forced to work with pre-fabricated areal units which they accept for reasons of convenience and expediency"[1]

Administrative data are often made available for areal units which may be hierarchically organised; in the UK, for example, this might be by local governments within regions. Local governments are not defined on consistent criteria; they vary considerably in size and function. England, for example, is divided into:
- 32 London Boroughs nested within Greater London County;
- 36 Metropolitan Boroughs/Districts, nested within the (now obsolete) six Metropolitan Counties;
- 55 Unitary Authorities, some comprising substantial urban areas but others created from former rural counties whose form of local governance was changed in the last two decades;
- and 27 Shire Counties, divided into 202 District Councils.

Much data produced for this mélange of authorities provide a misleading picture. Some are relevant to the authorities' functions, such as the provision of education and social services; others are not.

Also, comparing life expectancy in a small rural district with an inner London borough is hardly comparing like with like.

Furthermore, many local authority boundaries bear little resemblance to the settlement pattern of homes and other buildings. Some urban authorities are over-bounded; for example, Bradford Metropolitan Borough includes not only the city of Bradford but also different towns, such as Ilkley and Keighley, plus substantial rural tracts. Others – such as Bristol – are substantially under-bounded, with significant parts of their built-up areas covered by two or more different authorities (Figure 1). Some places even have no separate data published about them; for example, major towns such as Huddersfield, Dewsbury and Batley are combined in the Metropolitan Borough of Kirklees. And whereas London is split into its 32 boroughs, no other major city is similarly divided for statistical reporting.

A better set of areal units is needed to portray other aspects of a country's geography. The Office for National Statistics, which makes census data for England and Wales available for built-up areas and for small areas, has produced a hierarchy of reporting units that are internally homogeneous according to their housing characteristics (based on research led by geographer David Martin[2]), which are nested within the local authorities. At a larger scale, Eurostat has produced standard regional and city hierarchies. But much analysis and media reporting of UK data focuses on the incoherent and idiosyncratic system of local governments, and so our picture of measures like national well-being is biased accordingly.

Similar issues arise at other scales. Wards, for which much census and other data are published, are used in the identification of areas deserving receipt of public funds, but these vary considerably in size. In England, the average population of Birmingham's 40 wards in 2011 was 26 826, whereas the London Borough of Tower Hamlets' 17 wards averaged 14 947; the average in Canterbury, England's smallest city was 6 298 across its 24 wards; and in rural Ryedale it was just 2 588.

There is therefore a challenge for administrative data providers to devise – and continually update – hierarchies of areas for reporting data that will ensure not only clearer, more nuanced, pictures of national, regional and local geographies, but also a spatial structure that is better suited for the implementation of spatial policies. They could follow the example of the Australian census authorities, whose TableBuilder function allows analysts to create their own tables from the original data, while protecting the anonymity of individuals and households (bit.ly/2Rzm9ES). Users can, without charge, cross-classify census or other variables – age by sex, by marital status, or by ancestry, for example – and download the resulting table for each area in a selected level of a given hierarchy; tables can be derived for as many levels as needed, with constraints regarding cell size to protect anonymity.

An alternative approach involves the creation of "bespoke neighbourhoods".[3,4] These are based on data for areas at the lowest level in a given hierarchy (the baseline units, such as census blocks in the US and output areas in the UK), which are aggregated into larger areas by the addition of neighbouring baseline units until a threshold is reached – such as the smallest number of contiguous areas containing 1000 persons. At any level above that baseline, therefore, the size and nature of the larger units are determined by the analyst rather than imposed by the data provider. In some countries, access to individual-level data is now allowed through secure-access environments, giving analysts much greater flexibility.

Scale

> "In geographic investigation it is apparent that conclusions derived from studies made at one scale should not be expected to apply to problems whose data are expressed at other scales. Every change in scale will bring about the statement of a new problem, and there is no basis for presuming that associations existing at one scale will also exist at another."
> McCarty *et al.*,[5] quoted in Duncan *et al.*[1]

A pattern or relationship identified at one spatial scale may differ from that at another scale when using the same data – making it potentially unsafe to generalise from results obtained at a single scale only. After all, a country's human geography is the result of processes and decision-making procedures operating at different scales, which need to be incorporated into research designs.

Patterns identified at one scale may only be observed because they incorporate those at a larger scale. Duncan *et al*. calculated an index of population concentration at five nested scales within the United States – from counties up to census divisions. The index value increased as the scale became finer-grained, but

> … if one system of areal units is derived by subdivision of the units of another system, the index computed for the former can be no smaller than the index for the latter. Thus, the index of concentration on a county basis will exceed the index on a State basis, because the county index takes into account intrastate concentration.[1]

Individual classes within a school may be ethnically segregated, for example, with 60 per cent of students drawn from one minority group. But if 60 per cent of the school's students are drawn from that group, there is no segregation at the class scale additional to that at the school scale – and if 60 per cent of all students in the school district come from that minority, there is no segregation at the school scale either; it, and its classes, are merely typical of the wider area. The challenge is thus not only to realise that what is observed at one scale may not be the same at another, even with the same data, but also to develop methods of identifying the scale-specific patterns and, from them, the underlying processes.

Ethnic residential segregation is commonly measured by an index indicating the degree of unevenness in the distribution of two groups across a set of areas; it varies from 0 (the two groups are distributed in the same proportions across the areas) to 1 (the groups share no areas).[6] Using 2011 UK census data, this has been calculated at three scales for the city of Leicester: at the micro-scale, we have 969 output areas that are relatively homogeneous on housing characteristics, with a mean population of 370; at the meso-scale, the output areas are nested into 192 lower layer output areas; and at the macro-scale, the lower layer output areas are nested into 37 middle layer output areas.

Figure 2 shows indexes at each scale for: Indians vs. non-Indians; Indians vs. Pakistanis: Chinese vs. non-Chinese; Arabs vs. non-Arabs; and Black Africans vs. Black Caribbeans. For each, as the scale of analysis progresses from macro- through meso- to micro-scale, the index becomes larger. There is no 'right' measure, only a scale-specific measure. The finer the spatial scale, the more segregated each group in the chosen pair is from the other – though the difference is much greater in some comparisons than others.

These measures treat each scale as independent of the others, whereas Duncan *et al*.'s argument can be explored by a multi-level modelling approach, which identifies the level of segregation at each scale net of its intensity at all of the larger scales within which it is nested.[7,8] The output from such modelling is equivalent to the index of dissimilarity (Dm – which has associated Bayesian Credible Intervals that can be deployed in testing for significant differences between groups and

over time). For the Leicester data, Figure 3 shows very different patterns from Figure 2. For two of the comparisons, segregation is highest at the micro-scale (holding constant segregation at the two larger scales); for two others, it is highest at the macro-scale. In only one case is it highest at the meso-scale.

The modifiable areal unit problem

> "…the results of manipulating areal data often are to some degree dependent on the choice of a set of areal units"[1]

Not only can the results of analysing a data set differ according to the spatial scale employed, analysing the same data at the same scale but aggregated into a different set of areal units can also produce different results. This was identified by Yule and Kendall, recognised by Duncan *et al*., and later named the Modifiable Areal Unit Problem (MAUP) by geographers.[9,10]

The MAUP issue, if not its technical nature, is widely appreciated by those seeking to gerrymander electoral districts to improve a party or candidate's chances of winning. Electoral victory across a set of districts depends not only on number of votes but also – as clearly demonstrated, both theoretically and empirically[11] – on how those votes are distributed across the small areas used to build the districts. For example, when the British Boundary Commissions published initial proposals for 584 new parliamentary constituencies in 2011, estimates suggested that had the new constituencies been used for the 2010 general election, the Conservatives would have won 299 seats, Labour 231 and the Liberal Democrats 46. During the public consultation, the Conservatives presented an alternative set of constituencies that would have given them 312 seats and their opponents 219 and 45, respectively. Labour's proposed set would have given them 242 seats, the Conservatives 289 and the Liberal Democrats 45; and the Liberal Democrats' alternative set would have resulted in Conservatives 295, Labour 225 and Liberal Democrats 56. Using the same building blocks (wards) and the same distribution of (imputed) votes across those wards, within the same set of rules, different configurations produced different outcomes – indeed, they could have determined whether one party had a majority in the House of Commons.[12]

More generally, the MAUP means that any pattern identified at a particular scale must be open to some scepticism. Does it represent the 'true' relationship (between poverty and life expectancy, for example) or is it an artefact of the particular aggregation – of individuals into areas, or small areas into large areas, etc.? Might a different aggregation produce a substantially different outcome and influence our appreciation of a statistical relationship?

The MAUP issue is closely linked to the widely appreciated ecological fallacy of (potentially wrongly) inferring individual behaviour from aggregate patterns.[13] If different aggregations at the same scale, let alone at different scales, produce different relationships between two variables, even more confusion is introduced to the inferences that might be drawn.

MAUP adds further to the uncertainty gained from analysing areal data. Not only are there scale effects, there are also aggregation effects. Is there, ask Duncan *et al*., "a theoretically optimum set of units"? The answer is almost certainly no – there is no 'solution' to the MAUP, and it may be undesirable to think that there could be. Some researchers have thus decided that the problem should be largely ignored, with any results achieved at a particular scale and aggregation treated as 'true' but only for that realisation, with any conclusions – and policies – derived from them treated with a degree of circumspection. Others suggest that the reliability of any one realisation should be tested against the outcome of all realisations.

It is possible, for example, to identify all of the possible ways in which the wards in an English city can be aggregated into a smaller number of constituencies within a set range of electorates and see which realisations of the combinatorial problem produce which results. A study of Sheffield in the 1970s, for instance, identified 15 937 ways that its 27 wards (at that time) could be aggregated into six constituencies, each with an electorate within 10 percentage points of the mean. In 12 327 of the realisations, Labour would have won five seats and the Conservatives one; in 697 versions, Labour would have won all six seats; and in the remaining realisations, Labour would have won only four seats. In actuality, they won five – the most likely outcome.[14]

The reporting of administrative data by areal units reflects this aggregation problem. Are the results of analysing the data typical of the relationships sought, or can different geographies produce different findings, with potentially substantial implications for any policies for which those findings might provide supporting evidence?[15]

A better frame of reference

We can only understand the world, and then seek to change it for the better, if we can accurately portray it – and many portrayals need statistics. Those statistics must be accurate, and within known error terms if they are based on sampled data. They must also be relevant, and one component of that relevance concerns the areas (or places) to which they refer.

Ensuring that statistics are underpinned by a viable geography involves substantial challenges. Those discussed here – the arbitrariness of many subdivisions of a national territory; the scale of reporting and analysis; and the aggregation issues associated with the modifiable areal unit problem – tend to be either ignored or overlooked in many presentations of data, resulting in pictures of the world that at best misinform and at worst mislead. Tackling those challenges is a major task for those who produce, analyse and deploy administrative data.

Author bios

Ron Johnston, Kelvyn Jones and David Manley are professors in the School of Geographical Sciences at the University of Bristol and members of its Quantitative Spatial Science Research Group. Their recent work involves the development of multi-scale models of ethnic residential segregation and electoral polarisation.

Notes and acknowledgement

References
1. Duncan, O. D., Cuzzort, R. P. and Duncan, B. (1961) *Statistical Geography: Problems in Analyzing Areal Data*. Glencoe, IL: The Free Press.
2. Martin, D. J., Nolan, A. and Tranmer, M. (2001) The application of zone-design methodology in the 2001 UK Census. *Environment and Planning A*, **33 (8)**, 1949-1962.

3. McAllister, I., Johnston, R. J., Pattie, C. J., Tunstall, H., Dorling, D. F. L. and Rossiter, D. J. (2001) Class dealignment and the neighbourhood effect: Miller revisited. *British Journal of Political Science*, **31 (1)**, 41-59.

4. Lee, B. A., Farrell, C. R., Reardon, S. F. and Matthews, S. A. (2018) From census tracts to local environments: an egocentric approach to neighborhood change. *Spatial Demography*, doi 10.1007/s40980-018-0044-5.

5. McCarty, H. H., Hook, J. C. and Knos, D. S. (1956) *The Measurement of Association in Industrial Geography*. Iowa City, IA: Department of Geography, Iowa State University.

6. Simpson, L. (2007) Ghettos of the mind: the empirical behaviour of indices of segregation and diversity. *Journal of the Royal Statistical Society, Series A*, 170 (2), 405-424.

7. Jones, K., Johnston, R. J., Manley, D., Owen, D. and Charlton, C. (2015) Ethnic residential segregation: a multilevel, multigroup multiscale approach exemplified by London in 2011. *Demography*, **52 (12)**, 1995-2019.

8. Leckie, G., Pillinger, R., Jones, K. and Goldstein, H. (2015) Multilevel modelling of social segregation. *Journal of Educational and Behavioral Statistics*, 37 (1), 3-30.

9. Yule, G. A. and Kendall, M. G. (1950) *An Introduction to the Theory of Statistics (14th. edition)*. New York: Hafner.

10. Openshaw, S. and Taylor, P. J. (1981) The modifiable areal unit problem. In N. Wrigley and R. J. Bennett (eds.) *Quantitative Geography: a British View*. London: Routledge & Kegan Paul, 60-70.

11. Gudgin, G. and Taylor, P. J. (2012) *Seats, Votes and the Spatial Organisation of Elections*. Chichester: ECPR Press.

12. Johnston, R. J. (2012) Which map? Which government? Malapportionment and gerrymandering UK-style. *Government and Opposition*, **50 (1)**, 1-23.

13. Robinson, W. S. (1950) Ecological correlations and the behavior of individuals. *American Sociological Review*, 15 (4), 351-357.

14. Johnston, R. J. and Rossiter, D. J. (1983) Constituency building, political representation and electoral bias in urban England. In D. T. Herbert and R.J. Johnston (eds.) *Geography and the Urban Environment, Volume 3*. John Wiley, Chichester, 113-156.

15. Goldstein, H., Carpenter, J., Kenward, M. and Levin, K. (2009). Multilevel models with multivariate mixed response types. *Statistical Modelling*, 9 (3), 173-197.

16. Hand, D. J. (2018) Statistical challenges of administrative and transaction data. *Journal of the Royal Statistical Society, Series A*, **181(3)**, 555-605.

FIGURES
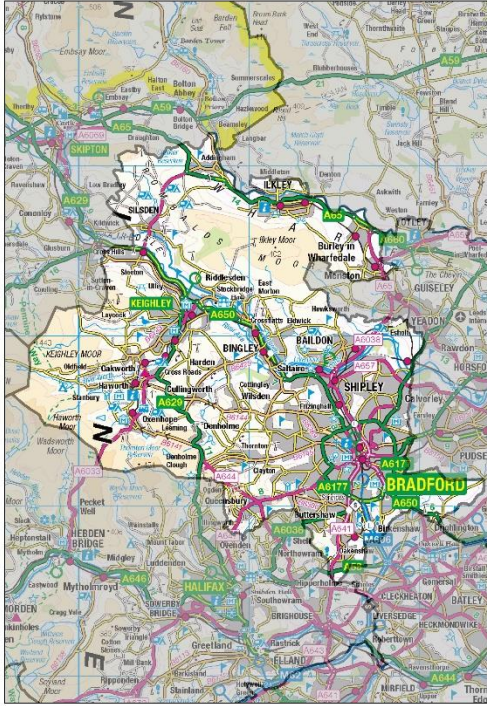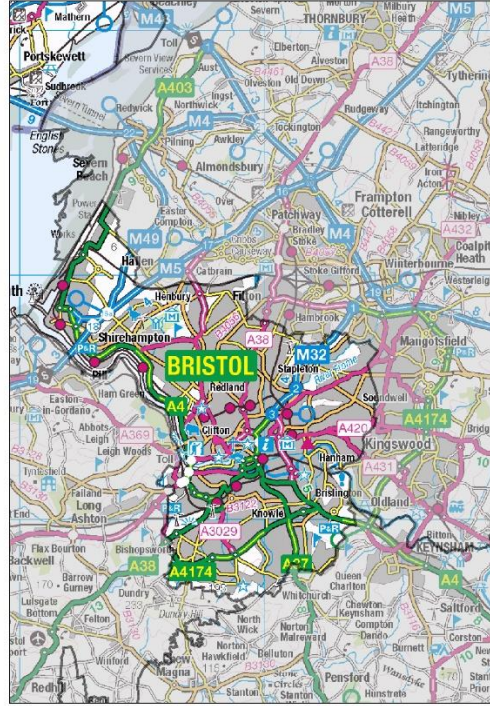
a)   b)

FIGURE 1 Comparing (a) over-bounded Bradford with (b) under-bounded Bristol. © Crown copyright and database rights 2018 Ordnance Survey (100025252). Data accessed via EDINA Digimap.
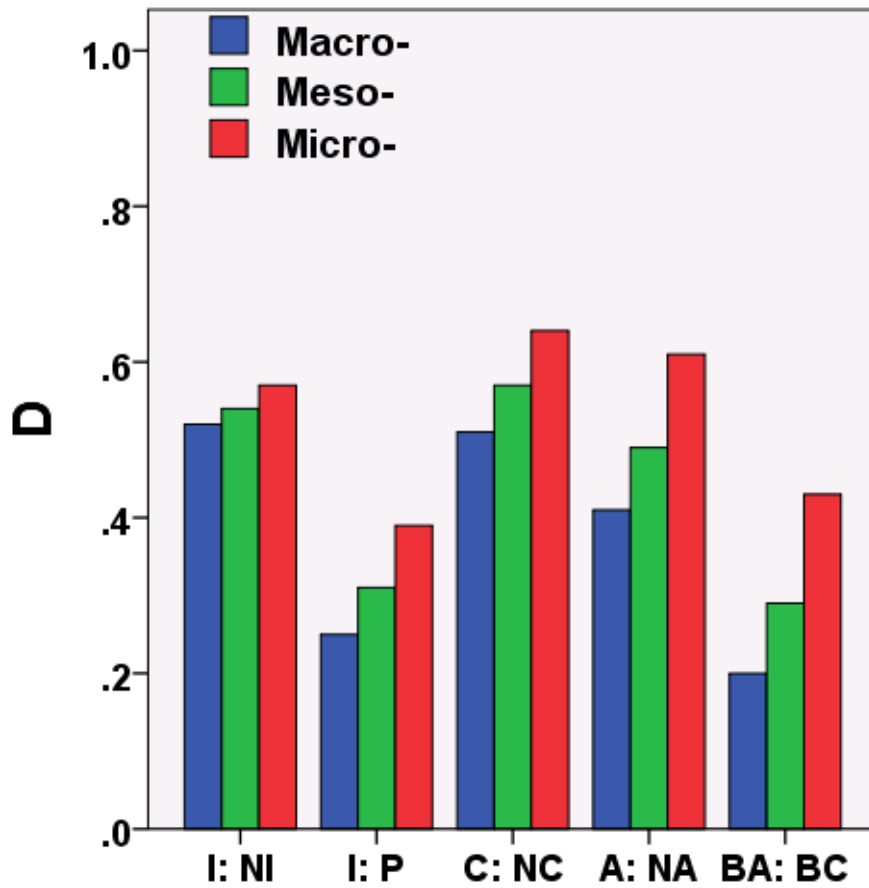
FIGURE 2 The index of dissimilarity (D) for five comparisons of the distributions of pairs of ethnic groups in the City of Leicester at three scales. Key to comparisons: I: NI – Indian: non-Indian; I: P – Indian: Pakistani; C: NC – Chinese: non-Chinese; A: NA – Arab: non-Arab; BA: BC – Black African: Black Caribbean.
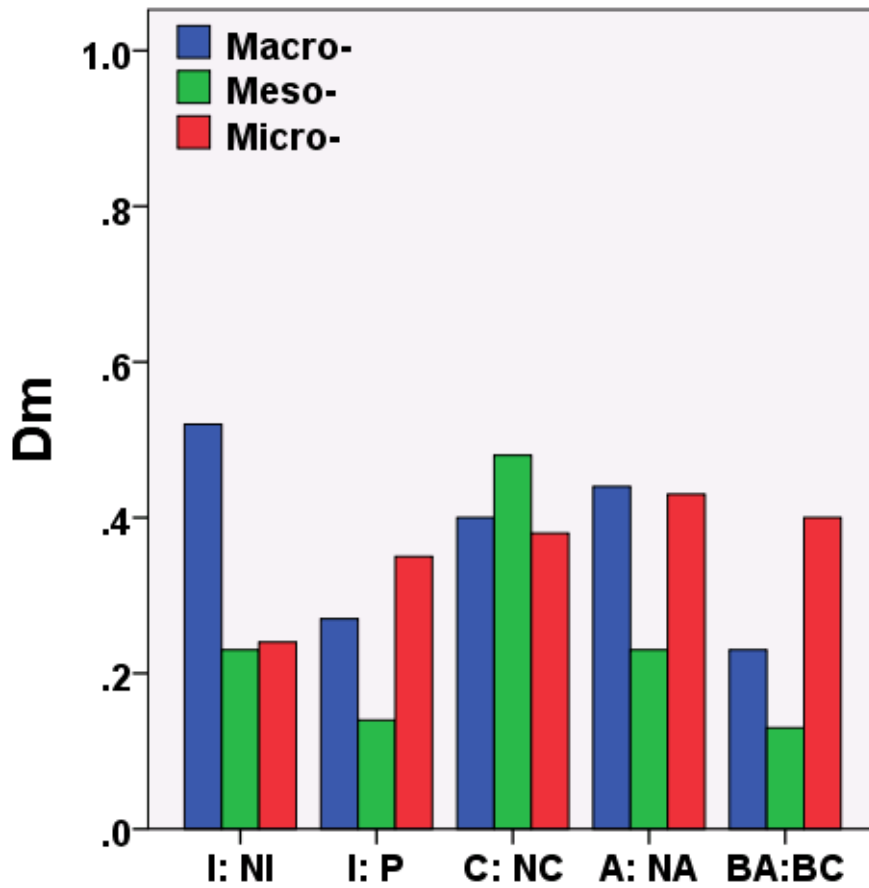
FIGURE 3 The modelled index of dissimilarity (Dm) for five comparisons of the distributions of pairs of ethnic groups in the City of Leicester at three scales. Key to comparisons: I: NI – Indian: non-Indian; I: P – Indian: Pakistani; C: NC – Chinese: non-Chinese; A: NA – Arab: non-Arab; BA: BC – Black African: Black Caribbean.