



Harris, R. (2018). From data to knowledge: teaching data skills in geography. *Geography*, 103(1).

Early version, also known as pre-print

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Geographical Association at <https://www.geography.org.uk/search?q=From data to knowledge> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/pure/about/ebr-terms>

From data to knowledge: teaching data skills in geography

Abstract

This article draws on material produced for the RGS-IBG Data Skills in Geography programme and presented at the 2017 Geographical Association Annual Conference. It argues that the role of data skills in geography goes beyond measurement, analysis and the rote learning of statistical tests to embedding data into the heart of geographical curricula. The twin pillars of statistical and disciplinary knowledge provide the means to tell geographic stories with data and to assess the credibility of the stories told. The skill of reasoning with numeric evidence is a tool for inclusivity and to challenge the propagation of misinformation.

Introduction

In May 2017, three books were published in close succession. The first is called *Post-Truth: How Bullshit Conquered the World* (Ball, 2017). The second (by Newsnight presenter Evan Davis) is *Post-Truth: Why We Have Reached Peak Bullshit and What We Can Do About It* (Davis, 2017). The more conservatively titled, *Post-Truth: The New War on Truth and How to Fight Back* (d’Ancona, 2017) was also published during that expletive-strewn month. All three are a reaction to the circulation of mis-information in social and political discourse, and to what can be done to address it.

Commenting on the subject of free speech, the writer and Harvard Professor, Jill Lepore states that “[a]ll speech is not equal. Some things are true; some things are not. Figuring out how to tell the difference is the work of the university” (*The New Yorker*, 2017).¹ The same comment can be applied to earlier phases of education. Writing on the conference theme of Inclusive Geographies, the 2016-17 President of the Geographical Association, Mary Biddulph, contends that, “now more than ever ... *all* children and young people (not just some) have the right to access geographical knowledge that will help them make sense of, engage with and where necessary challenge a post-truth discourse” (Biddulph, 2017 p.46, original emphasis). This article is about how that access can be achieved in relation to data skills, geographical teaching and the fostering of critical thinking.

Post-truths, geography and evidence-based facts

Post-truth is defined by the Cambridge Dictionary as “relating to a situation in which people are more likely to accept an argument based on their emotions and beliefs, rather than one based on facts.”² It can be linked to the use of social media and the tendency for people to be exposed to, seek out and to circulate attitudes, sentiments and innuendo that fit with their existing beliefs and preferences, and which they are inclined to take at face value without deeper consideration to the veracity of the information conveyed. As such, post-truth is the antithesis of critical enquiry, which is “the process of gathering and evaluating information, ideas, and assumptions from multiple perspectives to produce well-reasoned analysis and understanding, leading to new ideas, applications and questions.”³

¹ <https://www.newyorker.com/magazine/2017/10/09/flip-flopping-on-free-speech>

² <http://dictionary.cambridge.org/dictionary>

³ <http://www.usca.edu/academic-affairs/general-education/critical-inquiry.dot>

If geography is a discipline in which students are “encouraged to understand their role in society, by considering different viewpoints, values and attitudes” (AQA, 2016 p.5) then on what basis can different ideas and opinions be evaluated? The purpose of the evaluation is not to stand in judgement of those who hold the opinions but nevertheless to recognise that not all information is equally truthful, and not every claim to knowledge is correct.

In her article, Buddulph stresses the importance of evidence-based facts, highlighting approvingly the work of the statistician Hans Rosling and his “unique capacity to use numeric data as a mechanism to challenge [misconceptions and ignorance], particularly about global inequalities” (ibid. p.46). The website for which Rosling is most remembered is www.gapminder.org, which has the tag line ‘unveiling the beauty of statistics for a fact-based world view’. The description may sell the website short. The visualisation tools to be found there offer an ability to engage with data in a manner that’s less about (learning) facts and figures and more to do with interacting with data to facilitate inquiry-based learning and the formation of knowledge.

In a similar vein, the role of data skills in geography goes beyond the processes of measurement and analysis for generating results, evidence and ‘facts’ to instead supporting evidence-based *knowledge* in geography, which is a co-construction of what is measured, how it is interpreted and how it relates to disciplinary understandings of the topic at hand.

Data skills, statistical methods and geography

In a recent workshop supporting teachers and data skills in geography, an exams agency circulated a draft exam question with the task to fill in the blanks of a table and calculate a chi-square value. The exercise was well-intended and helpful to the teachers who can expect their students to be assessed on the same. However, therein lies the problem and the concern that this is not what data skills should really be about. It’s not that a chi-square test cannot be useful – for testing samples in a field trip setting for example or, at a more advanced level, for measuring the spatial clustering of events on a map and whether they differ from randomness. It’s just that there is limited addition to the students’ disciplinary knowledge from rote learning of a statistical method that, of itself, generates a value more-or-less uninterpretable without wider understanding of random sampling, statistical significance and of the effect size (see <http://rpubs.com/profrichharris/chisquare> for an example looking at the geography of the Brexit vote).

To be clear, the suggestion is not that chi-square should be avoided, nor that there is no value in learning it and other statistical tests. Instead, the concern is on how they are taught and the kind of teaching ‘fill-in-the-blanks’ assessment might generate. To be specific, if the focus is only on the mechanics of the statistical calculations rather than on the geographical knowledges that both motivate and are generated by the calculation then it risks imparting a rather narrow view to students of what data skills and geography can be about. If the aim is the much-discussed idea of powerful knowledge (Young et al., 2014), and if a geographical education is “powerful if it values students’ everyday knowledge, enables them to see the world in different ways, makes them aware of the value-laden nature of issues studied in geography and equips them with the skills to make sense of geographical knowledge”

(Roberts, 2017 p.9), then of relevance is where quantitative methods and data skills sit in relation to the geography curricula, and what sort of learning they facilitate.

That issue is pertinent because of the recent GCSE and A-level reforms, and the level of numeracy they require of students and therefore of teachers also. These have sought to embed data skills into the heart of what is taught in and as geography. In this, the GCE AS and A Level subject content document is clear:

Competence in using geographical skills should be developed during study of core content and non-core content, not as a separate theme or topic [...] students must be introduced to a roughly equal balance of quantitative and qualitative [methods] across the specification as a whole (DfE, 2014 p.12).

Any temptation to hive off the quantitative components for physical geography only is precluded. For example, the global systems and global governance theme requires the use of both quantitative and qualitative approaches.

These changes have provoked anxiety amongst some teachers, creating a sense of historical déjà vu. In an editorial for this journal, Willis (2016) reflects on the life of Stan Gregory, a leading figure in geography's 'quantitative revolution' of the 1960s and 1970s, commenting on a paper that he wrote that summarised the findings of a questionnaire sent to GA members at the turn of those two decades. Willis writes, after Gregory, "in general, many teachers were unsurprisingly concerned about their lack of training in such quantitative methods and approaches and were worried about the time and resources that would be needed if new elements were introduced into the school curriculum." (p.4) Similar findings were found over 40 years later in a survey undertaken by the Royal Geographical Society (with IBG) (Harris et al., 2013, 2014).

However, there is a key difference between then and now and that is that we are living in the age of much data,⁴ with a large amount of resources readily available that don't need reinventing, just brought into teaching – resources such as Gapminder but also maps of world development (<https://worldmap.canadiangeographic.ca/>), information about how far the world has got in meeting the Sustainable Development Goals (SDGs) (<http://datatopics.worldbank.org/sdgs/>), crime data by postcode (<https://www.crime-statistics.co.uk>), the Aqueduct Global Flood Analyzer (<http://www.wri.org/resources/maps/aqueduct-global-flood-analyzer>), UK Census maps (<http://datashine.org.uk>), and a detailed map of the ethnic geography of England, Wales and Northern Ireland in 2011 (<http://projects.andrewwhitby.com/uk-ethnicity-map/>), to identify just a few.

What matters is how quantitative methods are conceived and, following on from that, how they are practised in teaching. If they are only associated with early nineteenth century statistical tests and the somewhat obscure nomenclature that accompanies them then the anxiety (and antipathy) towards these methods in the classroom will persist and deservedly so – the inclusivity of quantitative methods can be questioned if they raise panic in some

⁴ The phrase 'big data' is commonplace but the point here is not the size but the availability

students and boredom in the rest. But if quantitative methods are about using data to undertake the sort of quantitative 'story telling' with which Rosling was associated and which encourages students to engage visually and creatively with the data for the purpose of geographical enquiry then they have the potential to stimulate, not stupefy, sending the clear message to students that data are useful and relevant to their studies.

Here the aspiration is different to that which has been associated with Gregory, with the quantitative revolution and with spatial science (Johnston et al., 2009). The aim is not to treat the whole of geography as a science nor does it entail the search for laws of spatial organisation and behaviour. It's also not the disciplinary equivalent of a land-grab that would promote only quantitative approaches to students. To the contrary, the upskilling in quantitative skills has been accompanied by an upskilling in qualitative ones, and rightly so given it is the breadth of approaches and of understanding that makes geography both vibrant and distinctive as a discipline. However, it *is* to insist that quantitative skills are a fundamental part of what it means to do geography and to be a geographer because data help us to explore and to explain geographical outcomes and processes, and to lift discussion and debate above mere anecdote, opinion or conjecture. They help to anchor disciplinary knowledge in evidence.

The simple point is that quantitative geography is not about teaching methods, it is about teaching geography. It follows that data skills cannot stand apart from the core themes of geography. Consider a choropleth map displaying unemployment rates for local authorities across the UK. Producing that map is not trivial, requiring a range of data skills: acquiring the data; calculating the unemployment rates; deciding on the map classes (how to group the data); selecting a colour scheme; knowing how to use a GIS if the map is digital. The map requires skills of calculation and of presentation and in it can be identified four of the data skills shown in Figure 1: numeracy, visualisation, statistics and data handling. However, it is the fifth element that takes the quartet beyond method and technique and into geography. It is thinking geographically to explore geographical ideas. In the case of the map, the calculation and presentation are precursors to interpretation and reflection – what is the geography of the map? Why is there a geography to the map? What does it reveal about the economic structure of the UK? How might the map shed light on the social and spatial organisation of the UK and about the different life chances afforded to people contingent upon where they live? How does the map correspond with other maps, such as the ethnic geography of the UK? And if there is some correspondence is it a fluke or is there some social, economic or other connection between the two?

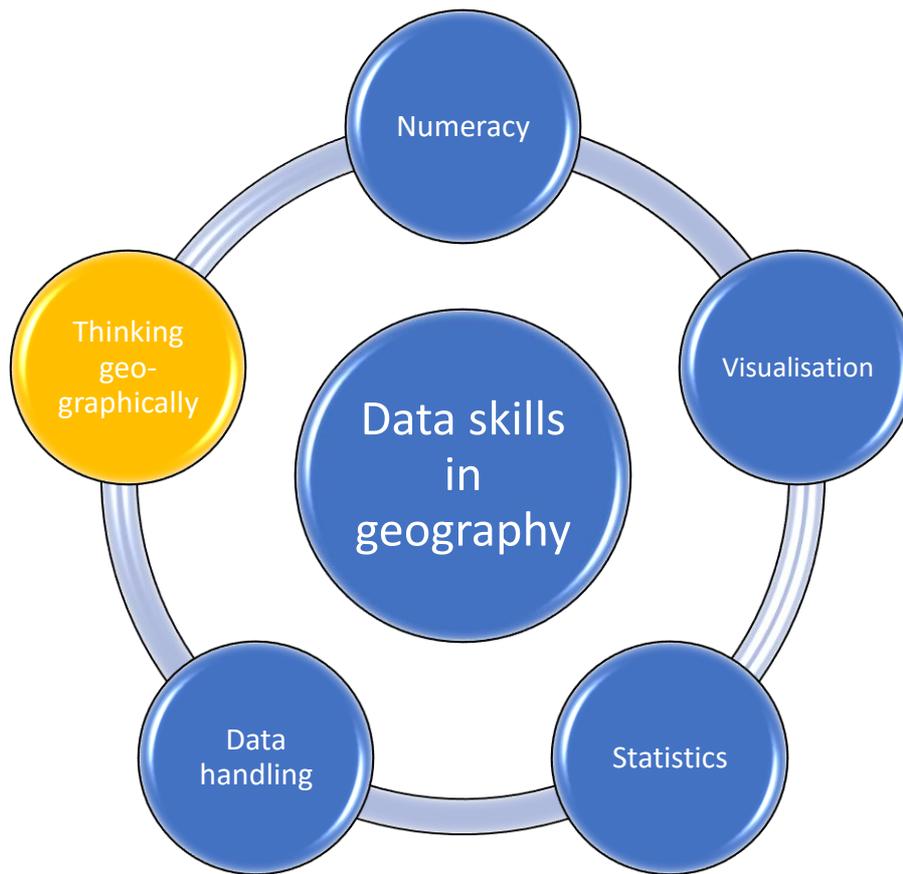


Figure 1. Data Skills in Geography (adapted from Harris, 2016)

Quantitative narratives (telling stories with data)

In practice, there may be no need for a student to create a map herself – it would be sufficient for her to make use of the data, maps and interactive visualisations that are available online to inform her studies. However, this ease of use, which can bypass the often difficult process of gathering data, collating them and making them usable, does not lessen the need for data skills, it increases it. It is a suitable level of statistical literacy that interacts with disciplinary knowledge to help ensure that the meanings that are given to data are trustworthy and well informed. That implies an approach to teaching that operates at the intersections between data and discipline, imparting sufficient statistical knowhow that a student can spot misuses of statistics and avoid making them himself. One way is through quantitative narratives and data-based storytelling.

Quantitative narratives is a phrase used by Milligan, Rose and Harris (2014) to describe the way that data, statistics and graphics are used to ‘tell stories’ in the media, in public policy, in scientific communication, in academic discourse, and so forth. To describe these narratives as stories is not to imply that they are fictitious but to focus attention on how they are presented and crafted, as well as on some of the rhetorical and visual strategies that are employed to make them persuasive. The intention is not to fuel cynicism but in much the same way as Huff (1954) was not actually abetting people in *How to Lie with Statistics*, the

purpose is to facilitate good practice and statistical learning. The idea is that a student learns to tell a good geographical story, where good does not mean definitive but that its plot lines are credible given what is known and can be evidenced about the subject.

Figure 2 shows an example of a deceptive quantitative narrative. It still appears on the Daily Mail website under the heading 'The Great Green Con no. 1: The hard proof that finally shows global warming forecasts that are costing you billions were WRONG all along'. It was published in the Mail on Sunday as "irrefutable evidence that official predictions of global climate warming have been catastrophically flawed."

Except it isn't, and they aren't. All data are uncertain and predictions made from them more so. That is why there are bands of (un)certainty around the predictions and why those bands get wider as they extend into the future. A 95% degree of certainty is not, by definition, correct 100% of the time so had the actual temperature fallen outside of the bands it would still be hasty to discount the predictions. But it didn't, which suggests a better caption would be 'the graph that reveals how 95 per cent certain estimates of the earth heating up were... correct'! In fact, far from "crashing out" of the estimates, 2014, 2015 and 2016 each recorded a record high annual temperature, reaching a change of +0.94 per cent (<https://www.ncdc.noaa.gov/sotc/global/201613>). The newspaper has yet to issue a correction.

The graph is interesting not just as an example of statistical misunderstanding but for how it is linked to on the websites of various organisations debunking the evidence of anthropogenic climate change. It invites discussions on the politics of global warming, as well as offering a means to introduce important statistical principles such as uncertainty, confidence intervals and forecasting.

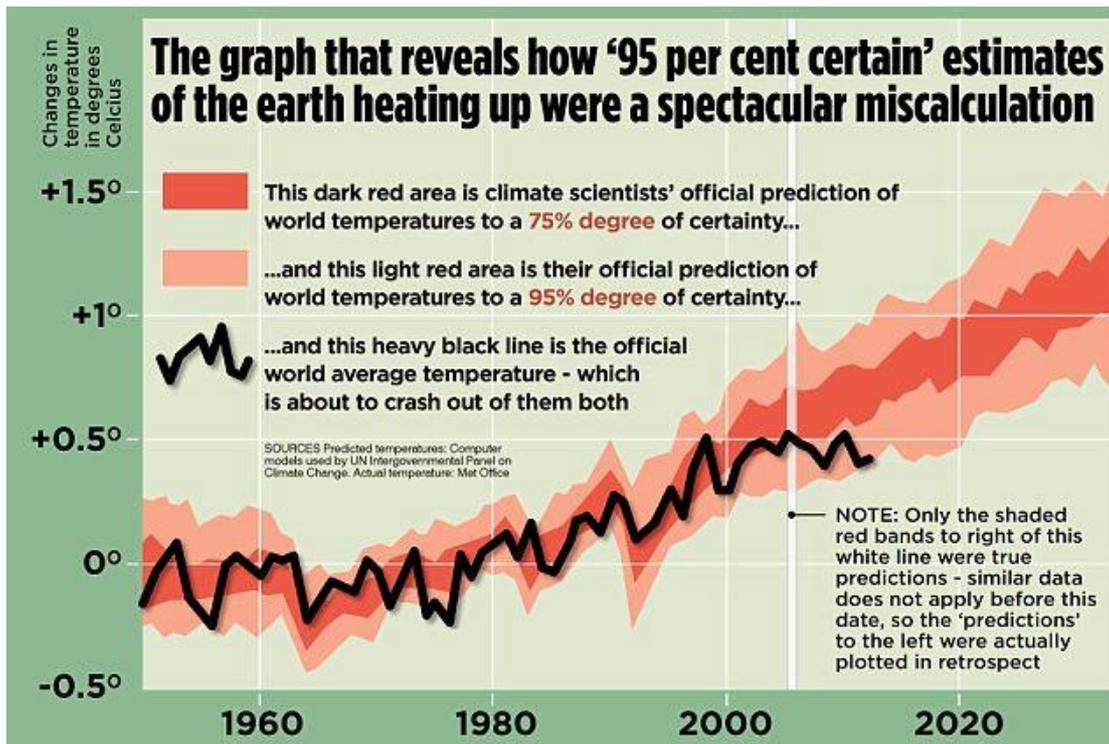


Figure 2. A graphic published on the MailOnline (source: <http://www.dailymail.co.uk/news/article-2294560>)

A second, more nuanced example of a quantitative narrative employing visual persuasion is provided by the book *The Spirit Level: Why More Equal Societies Almost Always Do Better* (Wilkinson and Pickett, 2009). This has been influential in the debate about inequality in (on average) more wealthy countries and its effects; it inspired the film, *The Divide* (dir. Katharine Round, 2015). Central to the book is a series of scatter plots which, by a line of best fit, show the relationship between inequality and a range of associated consequences, including health and social problems, levels of trust and drug use. Those scatter plots can be viewed and downloaded as a slide presentation from <https://www.equalitytrust.org.uk/resources/the-spirit-level>.

Famously, the physician John Snow used a map to highlight the concentration of cholera cases around a water supply in London, lending weight to his theory that cholera is a water-borne disease.⁵ The authors of *The Spirit Level* also employ graphical techniques, specifically scatter plots, to convey the harmful effects of a different kind of malaise, income inequality within countries. The topic is of geographical relevance – differences between countries, and differences between people and places within countries – but there is no reason to accept the claims unquestioningly. Engagement with data skills permits the key arguments of the book to be scrutinised with respect to various statistical ideas, including: scatter plots; lines of best fit (specifically a regression line); modelling the relationship between two variables; sample selection and the difference between purposive and random sampling; ideas of

⁵ See <http://www.bl.uk/learning/histcizen/21cc/publichealth/sources/source13/snow2.html>

statistical significance; whether correlation implies causation; how well the relationship fits the data; and whether there are outliers, which are countries that don't fit the overall trend. Outliers are usually regarded as a statistical nuisance but from a geographical perspective they can be interesting: if a country does not fit with expectation, why? Is there a geographical explanation for the difference?

The process of enquiry is a synergism between data skills and disciplinary knowledge: data generate knowledge and subject-based expertise suggests explanations for what is found in the data. Consider recent concerns about segregation in the UK that have been amplified by the recent publication of the Government commissioned *The Casey Review: a review into opportunity and integration* (Casey, 2016), a report that generated the following headlines:

Segregation at 'worrying levels' in parts of Britain, Dame Louise Casey warns
(BBC News, December 5, 2016)

Casey review raises alarm over social integration in the UK
(The Guardian, December 5, 2016)

and, absurdly,

GHETTO BLASTER Mass immigration to Britain has changed it beyond recognition and turned communities into ghettos, reveals damning report
(The Sun, December 4, 2016)

A flick through the UK Census data for 2001 and 2011 at www.nomisweb.co.uk or <http://infuse.ukdataservice.ac.uk> will reveal that the UK is becoming more ethnically diverse – compare the ethnic composition of London in 2001 and 2011, for example, or use the maps of population change available at <https://popchange.liverpool.ac.uk/>. However, the suggestion of increased segregation surprised many researchers who have shown that residential ethnic segregation is decreasing.

A short article intended for classroom teaching summarises the debates (Harris, 2017). It notes that different people are using the same data to reach opposite conclusions: some say ethnic segregation is increasing; some say it is decreasing. Are some being wilfully misleading?

No. It depends on which groups are looked at. Overall, neighbourhoods are becoming more mixed but there is a slightly increased separation of the White British from other groups as the former have become less prevalent in cities such as London (see Catney, 2013).

It is why they have become so that raises interesting geographical questions. It might suggest a process of separation and of antipathy in the sense of one or more of the groups choosing to avoid each other, which is what the media headlines tend to imply. However, there are other possibilities that can be explored with the relevant data. For example, data on the age structures of the various groups (the White British are, on average, older), their access to labour markets ('minority' groups often face economic disadvantage), social (as opposed to ethnic) segregation (the two are strongly linked) and whether the groups are each equally

willing to live in or have the affluence to move from the increasingly crowded housing associated with the densification of London (Johnston et al., 2016).

Biddulph (2017 pp.47–48) observes that

teachers are well placed to utilize enquiry pedagogies to support young people in interrogating the evidence they encounter. Questions such as ‘Where is the data/information/evidence from? Is it a reliable source and how do I know? Who produced the data/information/evidence? For what purposes? Is there any additional supporting data/information/evidence? What are the counter arguments, ideas/sources? What inferences can I make from what I read/see/hear?’ seem to me more important now than ever.

Creating, challenging and re-creating quantitative narratives through the perspective of geography helps to do this. It fits with Roberts’ (2017) view of a powerful pedagogy that promotes thinking, critical understanding and students’ active involvement in the construction of geographical knowledge.

Conclusion

In his book *The Internet of Us*, the philosopher Michael Lynch (2016) argues that knowledge based on reason plays an essential role in society and that there is more to knowing than the acquisition of information. A similar point can be made about data skills – there is more to these and to quantitative geography than the acquisition of method. Nevertheless, Lynch upholds the scientific method as a tool for inclusivity. His is not a naïve belief in the ‘truthfulness’ of science. Scientific methods are not objective if, by that, it is meant to disconnect them from the social practices and partialities that imbue science with meaning. However, for Lynch there is another issue at stake:

The sceptical argument says, in effect, that we can’t defend fundamental scientific methods as any more rational than other methods. What Kant points out, however, is that we *can* show that they are more democratic, more respectful of basic human autonomy. Why? Because scientific methods use human cognitive capacities such as observation and inference [that] aren’t secret or the province of a few. Observation and logic are strategies that everyone can, at least to some extent, use [...] and that can be made at least a little more effective with training (*ibid.*, pp.59–60).

This paper agrees with that argument, adding that for geographical education to be effective it requires the interweaving of method with disciplinary expertise. Arguments in favour of data skills are often made in terms of employability, to improve students’ numeracy and with appeal to general notions of citizenship, including a push-back against post-truths with evidence-based facts. Each of these is important but also insufficient. What they lack is anything specifically geographical. Ultimately, what is important is not teaching geographers data skills but teaching *data skills in geography* where the “in geography” is the intersection of methods with disciplinary knowledge and the embedding of data in the teaching of core topics.

In the age of much data, statistical ideas and concepts such as descriptive statistics, inferential statistics, the foundations of relational statistics, and issues around measurement, measurement errors, and sampling are important and all appear in the AS- and A-level subject content for geography. This is to be welcomed. What is problematic is when they engender a very literal teaching to the (statistical) test. That is both a shame and a missed opportunity when a wider appreciation of data skills can give students the ability to engage in geographically relevant ideas, to flourish as geographers and to foster critical enquiry with a disciplinary framework of knowledge and understanding. The Data Skills in Geography programme helps to support those goals.

Acknowledgments

The Royal Geographical Society (with IBG) Data Skills in Geography programmes is funded by the Nuffield Foundation. Teaching material is available at www.rgs.org/dataskills. The view expressed here are my own.

References

- d'Ancona, M. (2017). *Post-Truth: The New War on Truth and How to Fight Back*.
- AQA (2016). *GCSE Geography*. Manchester: AQA.
- Ball, J. (2017). *Post-Truth: How Bullshit Conquered the World*. London: Biteback Publishing.
- Biddulph, M. (2017). 'Inclusive Geographies: the illusion of inclusion', *Teaching Geography*, 42, 2, pp. 46–48.
- Casey, L. (2016). *The Casey Review: a review into opportunity and integration*. London: Department for Communities and Local Government.
- Catney, G. (2013). *Has neighbourhood ethnic segregation decreased?* Manchester: ESRC Centre on Dynamics of Ethnicity. <http://www.ethnicity.ac.uk>
- Davis, E. (2017). *Post-Truth: Why We Have Reached Peak Bullshit and What We Can Do About It*. London: Little, Brown. London: Ebury Press.
- DfE (2014). *Geography. GCE AS and A Level subject content*. London: DfE.
- Gregory, D., Johnston, R., Pratt, G., Watts, M. and Whatmore, S. (eds. 2009). *The Dictionary of Human Geography* (5th edn.). Chichester: Wiley-Blackwell.
- Harris, R. (2016). *Quantitative Geography: the basics*. London: Sage.
- Harris, R. (2017). *Population change and ethnic (de-?)segregation in London: evidence from the Census*. www.rgs.org/dataskills
- Harris, R., Fitzpatrick, K., Souch, C., Brunsdon, C., Jarvis, C., Keylock, C., Orford, S., Singleton, A. and Tate, N. (2013). *Quantitative Methods in Geography: Making the Connections between Schools, Universities and Employers*. London: Royal Geographical Society (with IBG).
- Harris, R., Tate, N., Souch, C., Singleton, A., Orford, C., Keylock, C., Jarvis, C. and Brunsdon, C. (2014). 'Geographers Count: A Report on Quantitative Methods in Geography', *Enhancing Learning in the Social Sciences*, 6, 2, pp.43–58.
- Huff, D. (1954). *How to Lie with Statistics*. New York: Norton.
- Johnston, R., Owen, D., Manley, D. and Harris, R. (2016). 'House price increases and higher density housing occupation: the response of non-white households in London, 2001–2011', *International Journal of Housing Policy*, 16, 3, pp.357–375.

- Milligan, L., Rose, J. and Harris, R. (2014). 'Convincing Students? Quantitative Junkies, Avoiders and Converts on a Cross-Disciplinary Course Using Quantitative Narratives', *Enhancing Learning in the Social Sciences*, 6, 2, pp.59–73.
- Roberts, M. (2017). 'Geographical education is powerful if...', *Teaching Geography*, 42, 1, pp. 6–9.
- Wilkinson, R. and Pickett, K. (2009). *The Spirit Level: Why More Equal Societies Almost Always Do Better*. London: Allen Lane.
- Willis, K. (2016). 'Editorial: Geography in uncertain times', *Geography*, 102, 1, pp. 2–4.
- Young, M., Lambert, D., Roberts, C. and Roberts, M. (2014). *Knowledge and the Future School*. London: Bloomsbury.