



Implications of a data-rich multivariate world. New needs and directions in statistical education

Jim Ridgway*

University of Durham, England - jim.ridgway@durham.ac.uk

Joachim Engel

Ludwigsburg University of Education, Germany - engel@ph-ludwigsburg.de

The data revolution is affecting many aspects of life, giving individuals access to unprecedented amounts of high-quality, up-to-date information about the world. It can also provide organisations and agencies with unprecedented access to individuals' personal data, and about their patterns of internet use, and can also be used to determine the information that individuals see on the web. Power is knowing how to use knowledge, and power relations are often asymmetric – organisations and agencies are better than individuals in accessing data, and knowing how to transform data into knowledge that they can use for their own benefit. Here, we map out some of the key features of the data revolution and relate them to the educational needs of students, citizens, and politicians. The two most obvious features are the emergence of new sorts of data, and the change in citizens' relationship with data. New sorts of data include: Big data – the firehose of data from sensors and other sources; social network data; professional data; and personal data. Our relationship with data has been changed by the open data movement, which seeks to put high-quality data from (for example) government sources into the public domain, and also the open access to internet traffic provided by (for example) *twitter*, along with contested access by governments and commercial organisations to medical records, purchasing histories etc. without the explicit permission of individuals. New data sources and open data pose quite different challenges, and offer quite different opportunities to educators. In many countries, the statistics curriculum ignores both, and so is in need of radical reform. We highlight new educational goals associated with the data revolution, and suggest aspects of the curriculum that can receive less emphasis. We argue that statistics education needs to be more militant in its approach to statistical literacy by challenging policy makers on matters of fact and interpretation (using open data), and by engaging with journalists on stories based on some of the rich data sources we have available that are directly relevant to governance. A further priority is to help everyone understand the strengths and limitations of the statistical tools that are used for description, pattern recognition and prediction from big data, which each impact upon all our daily lives. We describe some recent initiatives that promote our educational ambitions.

Keywords: data revolution; statistics education; statistical literacy; multivariate evidence.