**International net migration in the context of English, Welsh and Scottish regional population demography, 2002-2013, using a ‘Data Tableau’**

**Dr Jon Minton, University of Glasgow, UK**

**Introduction**

International net migration is a hot political issue in the UK. In the 2015 UK General Election, UKIP, an anti-immigration party, was able to secure more than four million votes. The Conservative government had repeatedly promised to reduce international net migration to the ‘tens of thousands’, and since the 2015 election to restrict in-migration flows by renegotiating the UK’s relationship with the EU.

The aim of this research is to use effective data visualisation to put international international migration in England, Scotland, and Wales, in the context of internal migration, changes over time, population structure, and differences across the UK.

**Data Visualisation**

A necessary but seldom sufficient definition of a data visualisation is as a rule-based graphic, whose rules specify how a series of variables in a dataset should be mapped onto particular types of graphical features or **aesthetics** [Wilkinson 2005; Wickham 2010[2009]]. An example of a set of mapping rules could be: 1) ‘the horizontal position of a graphical object is determined by the variable *time* (unit: time of day)’; 2) ‘the vertical position of this graphical object is determined by the variable *value* (unit: US dollars)’. By additionally specifying that the graphical object (‘**geom’**) whose vertices are determined by the above rules is a *line* (or more precisely a polyline, if it contains more than two points), and that the order in which the vertices of the polyline are joined is in ascending or descending order of the *time* variable, the result of the above specification is a simple time series graph, of the type typically used to show the variation of a share price over the course of a day of trading.

This minimal specification above defines what can be thought of as the ‘data layer’ of a data visualisation. But, for a data visualisation to be effective at communicating to the audience, additional layers are usually required. We can refer to these additional layers by the role they fulfil: we can defining graphics such as axis labels, tick marks, and gridlines, as being part of a ‘support layer’, which allow the audience to infer the values in the data from the rule-bound features of the data layer, allowing the audience to ‘decode’ the data layer. Additionally, we can think about adding an ‘annotation layer’, containing a range of non rule-bound graphical features which the creator of the data visualisation thinks will help the audience to identify important information. An example of an annotation layer may be to add a label nearby particular points, for example to help the audience identify outliers in the data. The choice of what, if anything, to include in the support layer and data layer, in addition to the mapping rules and design aesthetics in the data layer, is the fundamental challenge of producing effective data visualisations.

**Data**

Mid-year estimates of population sizes are produced by the UK Office for National Statistics in order to construct Small Area Population

Estimates (SAPE) for each local authority within England and Wales. These estimates, contained within table MYEB2 of the Components of Change data, are used to construct SAPE, and include estimates of population sizes, and internal and international in-migrations and out-migrations, for each year from 2002 to 2013, and for each single year of age, from birth to 89 years old, with aggregated counts for ages 90 years and above. Internal migration is defined as any movement originating from and terminating at any local authority in the UK; international migrations are those which originate or terminate outside of the UK over the same period. Equivalent data for Scotland was made available by the National Records of Scotland.

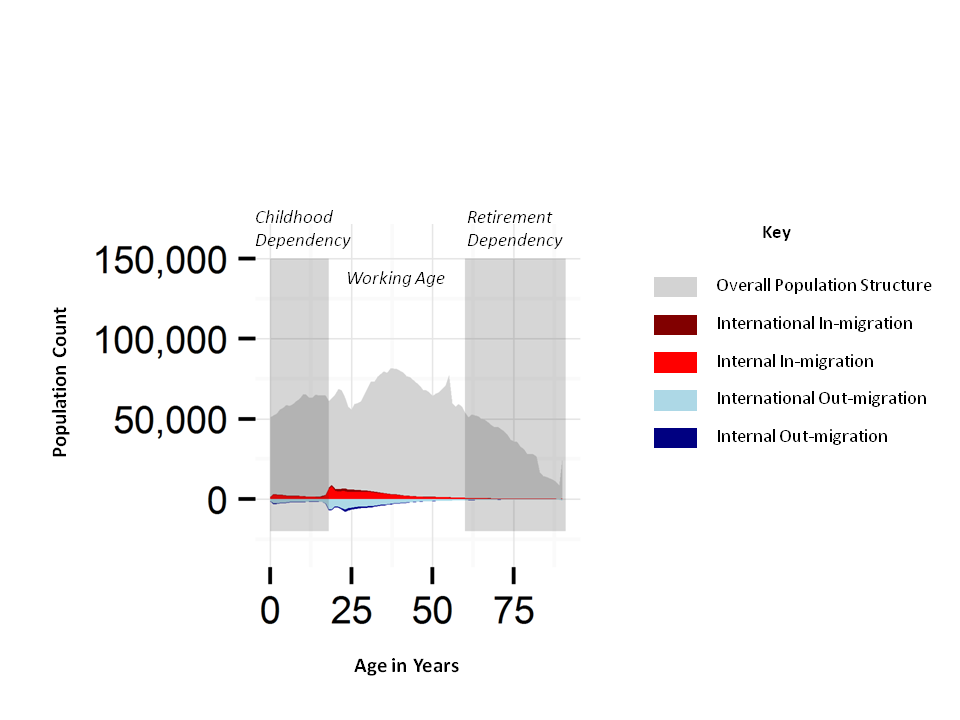
**Figures and Colour Scheme**

The population structure and migrant flows are shown for each year from 2002 to 2013, and for each English region, as well as Wales and Scotland. The figures are arranged as ‘small multiples’ (Tufte 2001 [1983]), also known as ‘trellis plots’ (Cleveland 1993). The vertical axis is population count and the horizontal axis is age in single years. The overall population structure is represented as a grey polygon the zero line. In-migrations are represented as red polygons also above this line, and out-migrations by blue polygons below this line. Darker shades indicate international migration flows, and lighter shades internal migration flows. Grey rectangles are added indicating the start and end of the working age range. The figure was produced using the R package ggplot2 [ref], and the R script used to produce this and other figures is available on Github.

Because of the large number of tiles included in the lattice plot, 12 columns wide by 11 rows deep, the final image is best viewed as a large (ideally A1 size) colour poster on a wall. Because of its size and presentation on a wall, I refer to the final image as a ‘data tableau’

**Anatomy of a tile**

The figure below provides an anatomy of one of the 132 tiles included in the full tableau.



We can see that the data layer comprises the five polygons representing overall population structure and migration counts at specific ages, the support layer comprises the grid lines and (for the left most and bottom tiles) the axis tick marks and labels, and the annotation layer comprises the two translucent grey rectangles indicating childhood and post-retirement ages.

**Results**

The figures highlight a number of important demographic features, such as: the transition of the 1948 ‘Baby Boomers’ from working age to retirement; high rates of young adult international migration in London; the slightly older population peaks in the South East compared with London; and the relatively old and small population sizes in Wales and the North of England; and a broad symmetry between in-migration and out-migration flows. Except for London, they also show internal migration to be much larger at all ages than international migration.

**Discussion**

By presenting these official estimates of annual migration and population counts in this way, a number of complex and engaging patterns emerge. These visualisations allow more knowledge and insight about migration and population processes, and regional variation in the UK, with important ramifications for economics, elder care, school and public service spanning, and regional inequalities.