# Characterising labour market self-containment in London with geographically arranged small multiples

**Roger Beecham[1] and Aidan Slingsby[2]**

[1]School of Geography, University of Leeds, UK

[2]giCentre, Department of Computer Science, City, University of London, UK

## Abstract

We present a collection of small multiple graphics that support analysis and understanding of the geography of labour-market self-containment across London's 33 boroughs. Ratios describing supply-side self-containment, the extent to which working residents access jobs locally, and demand-side self-containment, the extent to which local jobs are filled by local resident workers, are first calculated for professional and non-professional occupations and encoded directly through geographically-arranged bar charts. The full distribution of workers into-and out-of- boroughs that underpins these ratios is then revealed via Origin-Destination flows maps (OD maps) – sets of geographically-arranged choropleths. In order to make relative and absolute comparison of borough-to-borough frequencies between occupation types, these OD maps are coloured according to signed chi-square residuals: for every borough-to-borough pair, we compare the observed number of flows to access professional versus non-professional jobs against the number that would be expected given the distribution of those jobs across London boroughs. Our geographically-arranged small multiples demonstrate potential for spatial analysis: a rich, multivariate structure is depicted that reflects London's economic geography and that would be difficult to expose using non-visual means.

## Keywords

Labour market self-containment, travel to work, origin destination, OD map, small multiples

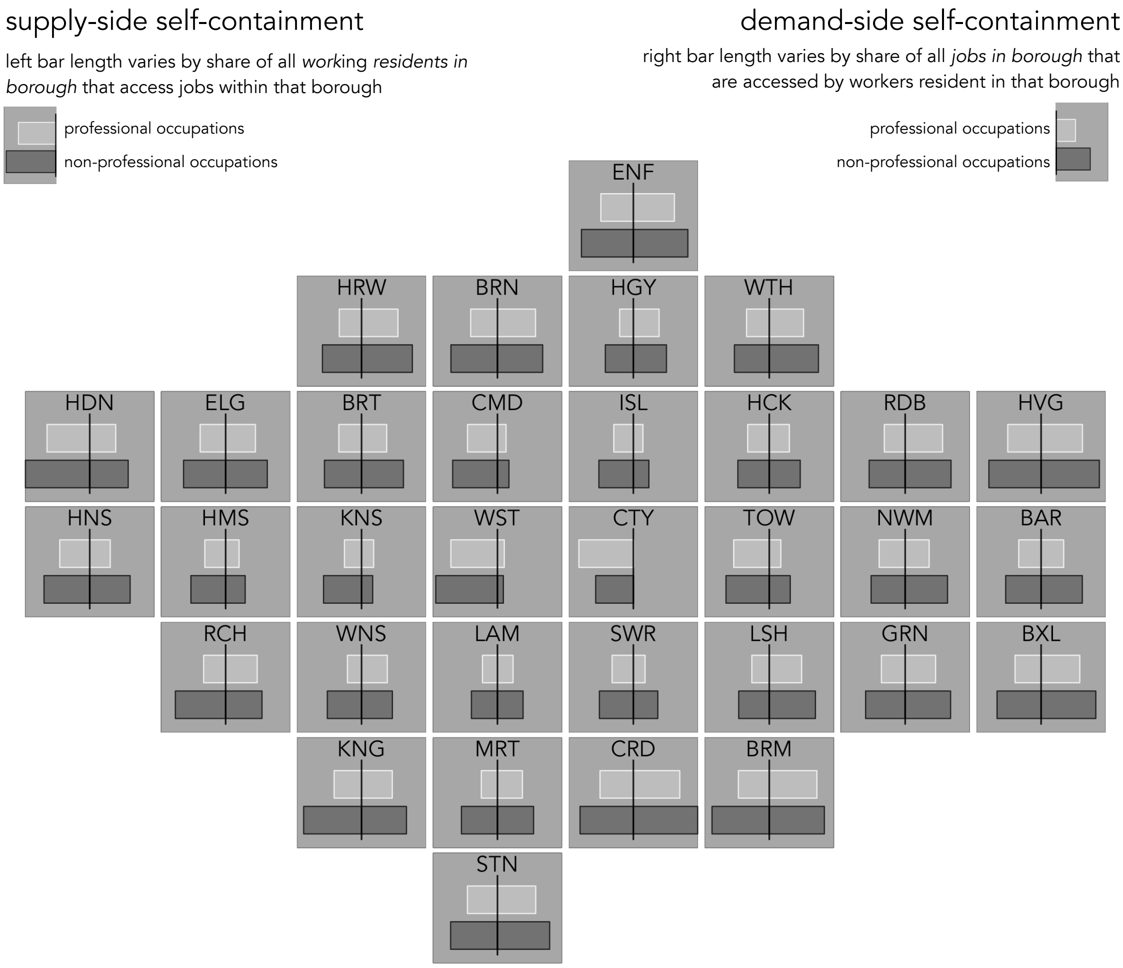


**Figure 1**: Bar charts, geographically arranged, to depict borough-level self-containment by occupation type. [Colour version from: https://github.com/rogerbeecham/visualizing-self-containment].

Of routine interest to economic geographers is the concept of labour market self-containment. Self-containment is typically quantified using two intuitive indicators: supply-side and demand-side self-containment. Supply-side self-containment describes the extent to which local residents of an area access jobs in that area rather than commute elsewhere for work; demand-side self-containment describes the extent to which local jobs are filled by local residents rather than workers commute in from surrounding areas. Behind these statistics there is an implied distribution of Origin-Destination (OD) flows: areas with high self-containment exhibit a power-law type distribution whereby most employed residents (supply) and available jobs (demand) are satisfied internally and this tendency becomes less strong as self-containment scores decrease. Our short paper presents a set of graphics that support analysis and understanding of the geography of labour-market self-containment in London. 2011 Census travel-to-work data are used describing flows between London's 33 boroughs. The graphics distinguish between jobs categorised as professional (light grey) and non-professional (dark grey), as per Office for National Statistics Standard Occupational Classification Hierarchy1.

Figure 1 displays bar charts of supply- and demand- side containment scores. These are small multiples (Tufte 1983): a separate summary chart is created for each reference borough and boroughs are arranged according to their approximate geographic position (Figure 2). Left of the vertical lines supply-side self-containment is summarised: bar length varies according to the share of all workers resident in the reference borough filling jobs within that borough. Notice that the light grey bars are almost always shorter than the dark grey bars: those working in professional occupations tend to fill jobs outside of their borough in greater number than do those working in

---

[1] Professional jobs comprise manager, director and senior officials, professional and associate professional occupations. Non-professional jobs comprise administrative and secretarial, skilled trade, caring and leisure, sales and customer service, process plant machine operative and elementary occupations.

non-professional jobs. The one exception is the City of London, which contains only 3,657 working residents, where the reverse is true. There is a geography to these distributions, with higher supply-side self-containment scores in more obviously discrete, outer London boroughs: Croydon, Bromley and Kingston in the south, Hillingdon in the west and Havering in the east. Right of the vertical lines demand-side self-containment is summarised: bar length varies according to the share of all jobs in the reference borough that are filled by local residents living in that borough rather than workers commuting in from elsewhere. The difference in self-containment scores for professional versus non-professional occupation types generally persists for non-central London boroughs. For central London boroughs containing much employment opportunity – most obviously Westminster and City of London, but also Camden, Islington and Kensington and Chelsea – a different pattern is observed. Demand-side self-containment scores are predictably low with only 6% (Westminster) and 0.4% (City of London) of jobs filled by workers living in those boroughs. However, rates of in-commuting are in fact slightly higher amongst those filling non-professional jobs in these boroughs.

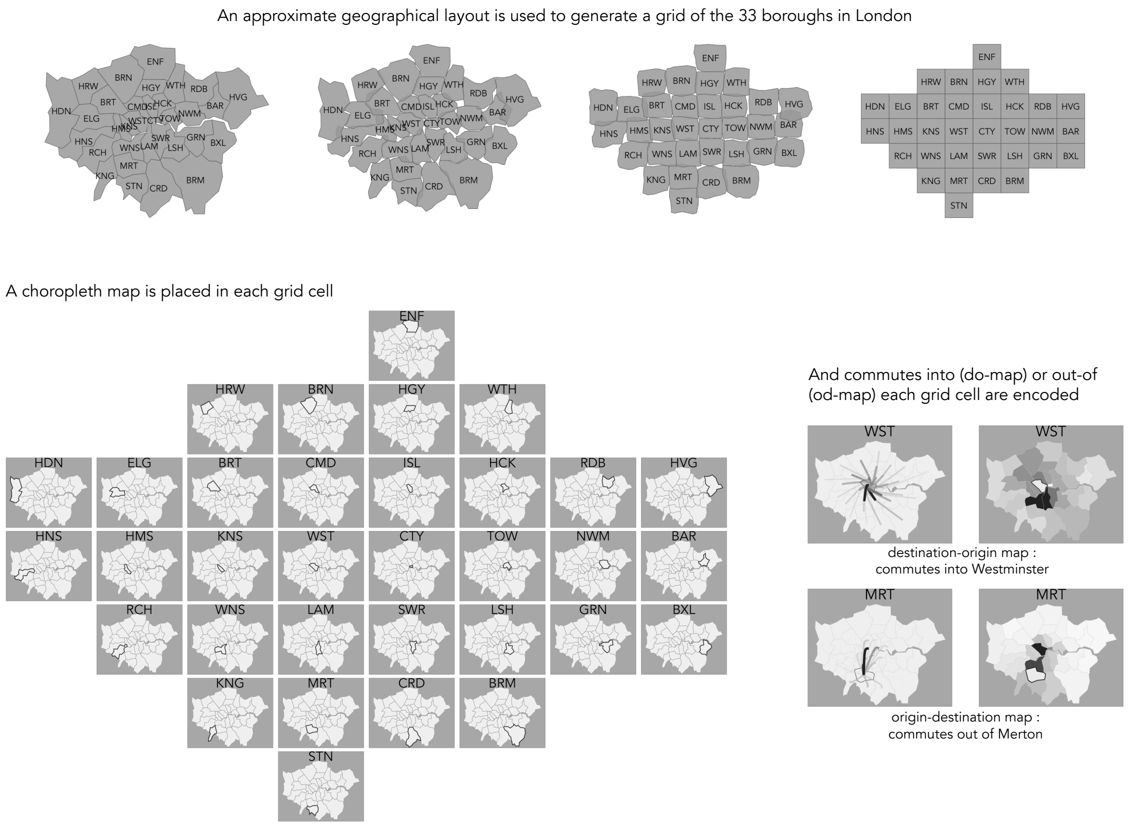


**Figure 2**: An approximate geographical layout is used to generate a spatially arranged grid of London's 33 boroughs. The LondonSquared (After the Flood 2015) arrangement is used here, but see Meulemans et al. (2017) for a generalisable technique for effecting such arrangements. In each grid cell we place a graphic -- a bar chart, as in Figure 1 or, in order to analyse full OD flows, a choropleth map as in Figures 3 and 4. [Colour version from: https://github.com/rogerbeecham/visualizing-self-containment].

We have already identified obvious differences in supply- and demand- side self-containment scores related to London's economic geography and to more general factors around occupation type: that job-rich boroughs contain low demand-side self containment, that outer London boroughs contain generally higher demand-side self-containment and to a lesser extent supply-side scores and that this tendency is greater for those working in non-professional occupations. But how are self-containment scores expressed geographically? From which parts of the city do job-rich boroughs such as Westminster and City of London draw workers and does this vary between occupation type? How different is the geography of supply-side self-containment between occupation types -- to which parts of London do professional and non-professional workers living in outer London boroughs commute?

To explore this, we generate contingency tables describing frequencies of borough-borough commutes for professional and non-professional occupations and calculate signed chi-square residuals (e.g. Visalingam 1981) to compare these frequencies, by OD pair, against what would be expected given the relative number of professional and non-professional jobs and an OD pair's job-count. Contingency tables are generated separately for each reference borough and differently based on whether reference boroughs correspond to origins (homeplaces) or destinations (workplaces).

Expected frequencies are calculated as:

$$\exp_{ij} = \frac{occ_i \times jobs_j}{total}$$

where

- $exp_{ij}$ is the expected frequency of jobs of occupation $i$ (professional or non-professional) in OD pair $j$;
- $occ_i$ is the total number of jobs of occupation $i$ accessed by residents of the reference borough (where reference boroughs are origins), or contained within the reference borough (where reference boroughs are destinations);
- $jobs_j$ is the total number of jobs (across occupation types) in OD pair $j$;
- and $total$ is the grand total of jobs accessed by residents in the reference borough (where reference boroughs are origins), or contained within the reference borough (where reference boroughs are destinations).

Each expected frequency by occupation and OD pair $exp_{ij}$ is then compared with the actual frequency by occupation and OD pair $obs_{ij}$:

$$\chi_{ij} = \frac{obs_{ij} - exp_{ij}}{\sqrt{exp_{ij}}}$$

Positive values of $\chi_{ij}$ indicate a higher than expected number of in- or out- flows in that occupation type, negative indicate a lower than expected number in that occupation type.

Generating contingency tables separately for each reference borough is an important addition. If the City of London is the reference borough and is set to represent a destination or workplace, then the expected values describe whether the number of commutes by occupation for any OD pair is higher or lower than would be expected given the relative number of jobs available in that borough. Our expected values are therefore sensitive to the residential and workplace geography of the city. For example, where reference boroughs are destinations, we have different expectations in the relative number of professional and non-professional jobs based on the known availability of those job types in the borough.

The signed chi-square measure has advantages over alternative measures of effect size (such as risk ratios) since, in considering observed and expected frequencies and dividing by the square root of expected, saliency is given to relative differences that are also large in absolute magnitude. This effect is necessary since in c.10% of the 1089 borough-borough pairs, fewer than 30 job counts are recorded; we might expect large relative differences between professional vs. non-professional occupations here, and so would register large effect size values if risk ratios were being used – these would nevertheless be comparatively small, and possibly idiosyncratic, differences in absolute terms.

These data are presented in Figures 3 and 4 as Origin-Destination flow maps (OD maps; Wood et al, 2010; Slingsby et al. 2014). We attempt a graphical explanation of the OD map layout in Figure 2. Each choropleth map describes flows of resident workers out of (Figure 3) and non-resident workers into (Figure 4) the reference borough. The choropleth maps are arranged according to their approximate geographic position, again using the LondonSquared layout (After the Flood 2015). In Figure 3, each labelled reference cell represents an origin (home) borough and the choropleth map encodes frequencies to destination (workplace) boroughs in London. In Figure 4, each labelled reference grid cell represents a destination (workplace) borough and the choropleth encodes frequencies of origin (home) boroughs from which the destination draws workers. The dark outlines further clarify reference boroughs: the origin (home) borough in Figure 3 and destination (workplace) borough in Figure 4.

Figure 3 reinforces the overall trend implied by the bar charts. Dark grey colours for reference boroughs confirm the higher supply-side self-containment scores for non-professional jobs. By representing individual OD pairs within their geographic position, we can add that commuting patterns are comparatively more localised for non-professional jobs – boroughs neighbouring the reference cell are typically dark grey. By contrast, central London boroughs are always coloured light grey: those working in professional occupations are overrepresented amongst residents commuting into central London for work. Notice that, even for this latter category of commutes – professionals into central London from outer boroughs – geography matters. For boroughs to the south east such as Bromley, Greenwich and Southwark the lightest greys, signifying both large relative differences and absolute frequencies, are for jobs filled in Tower Hamlets and City of London (boroughs in east London). For boroughs to the west, such as Hounslow and Ealing, lighter greys representing jobs filled in Westminster, Camden and Hammersmith also appear.

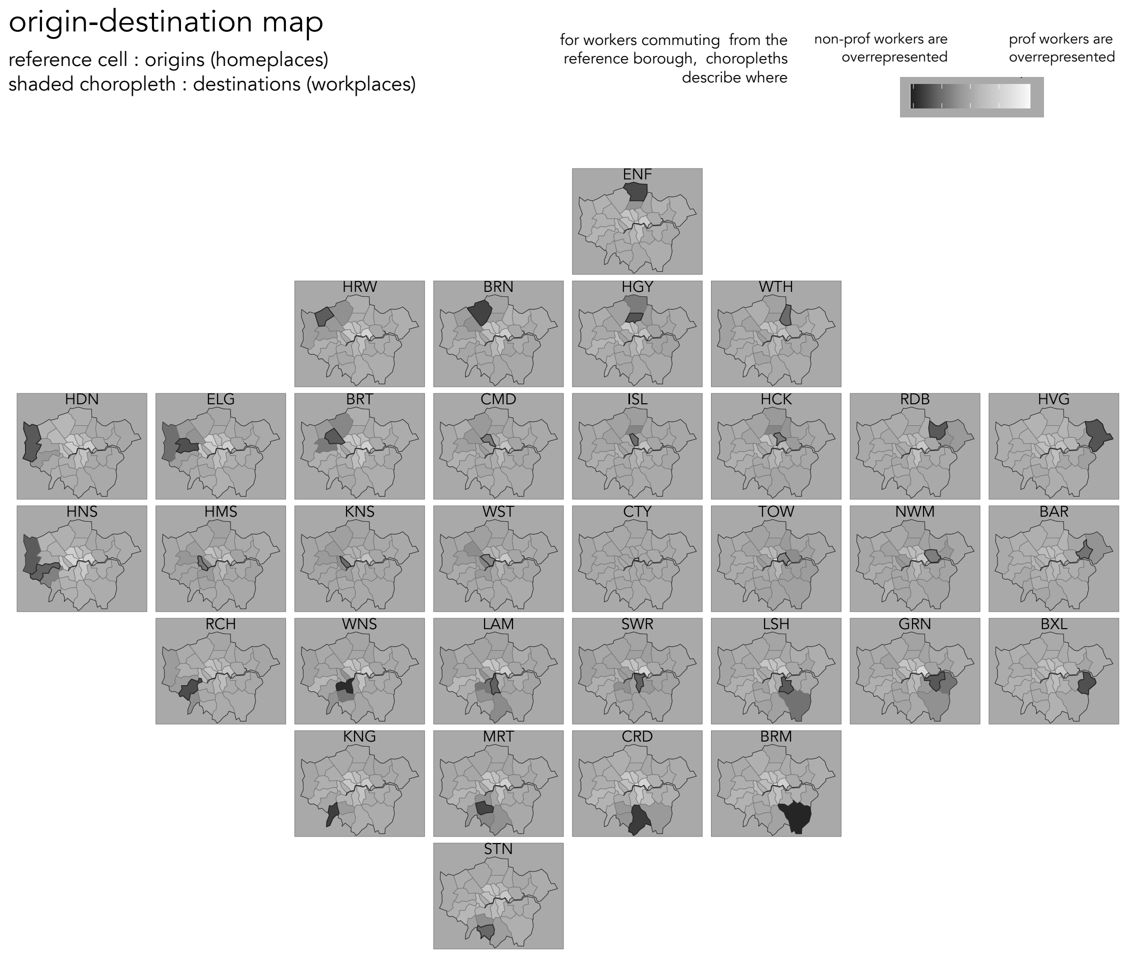


**Figure 3**: Origin-Destination map depicting flows of workers living in the reference cells to fill professional and non-professional jobs in London. [Colour version from: https://github.com/rogerbeecham/visualizing-self-containment].

A key pattern in Figure 4 is that for jobs filled in central London boroughs (City of London and Westminster), non-professionals are overrepresented amongst those commuting in from east London and professionals are overrepresented amongst those commuting in from west London. In this respect, the figure reflects differences in London's socio-economic geography: particularly salient is the light grey shading representing high numbers of professional workers from Wandsworth and Richmond. A close inspection of Figure 4 also supports the earlier observation around higher demand-side self-containment scores for non-professional jobs for all but central

London boroughs: notice that the reference cells are coloured darker grey with the exceptions being City of London and Westminster.

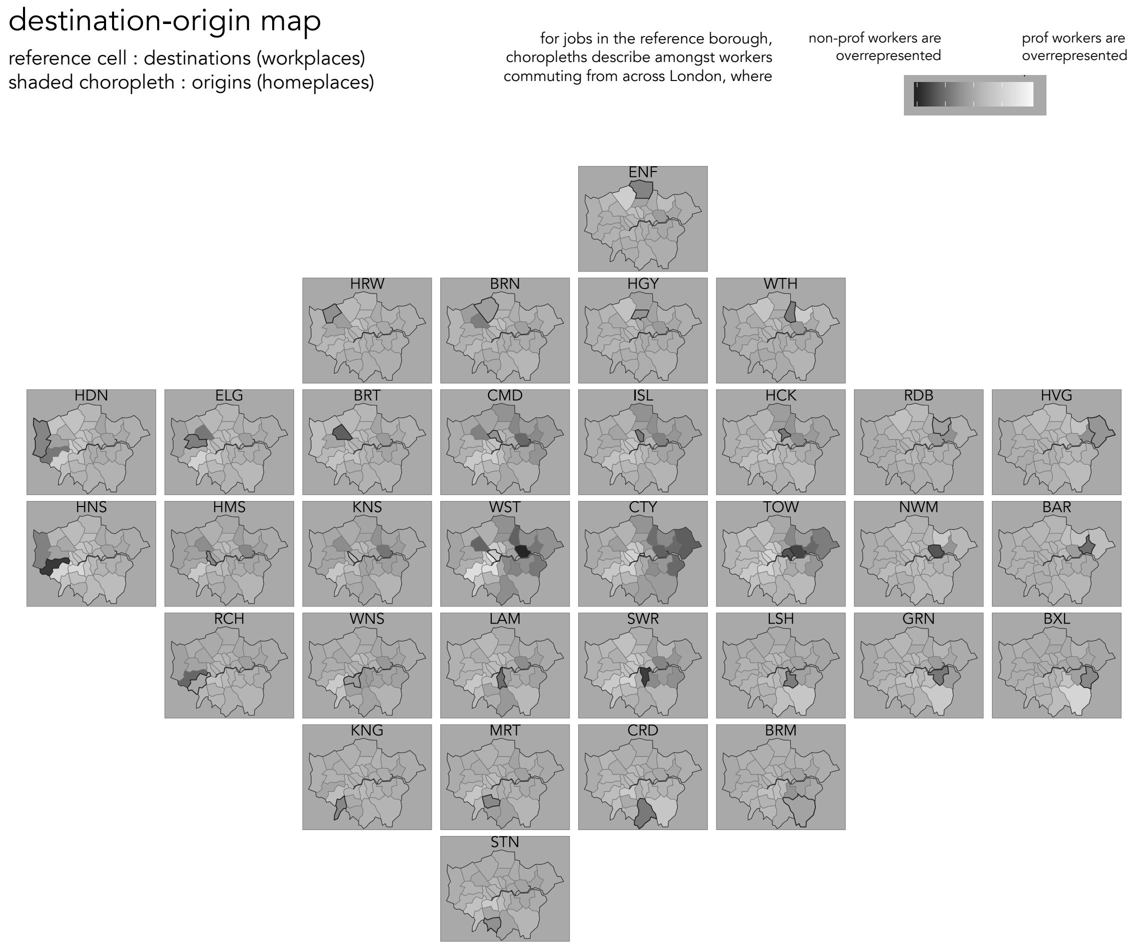


**Figure 4**: Destination-Origin map depicting professional and non-professional jobs filled in the reference cell by commuting workers in London. [Colour version from: https://github.com/rogerbeecham/visualizing-self-containment].

There are limits to the application of geographically-arranged small-multiples: most obviously, that applying the same encoding to a more disaggregated geography could be problematic. Our graphics nevertheless expose rich structure around London's labour market geography that would be difficult to explore using non-visual means. Not only do our designs support geographic comparison of the two self-containment metrics by occupation group but, in our OD maps, a sense of the underlying patterns of resident workers commuting out of boroughs to access wider employment opportunities (Figure 3) and of non-resident workers into boroughs to fill jobs (Figure 4).

## Code and software

All graphics were generated in R using the ggplot2 package. Further discussion and code describing how these layouts are produced is at this github repository: https://github.com/rogerbeecham/visualizing-self-containment.

## References


After the Flood. 2015. City and citizen service development. [web-page]. Available at: https://aftertheflood.com/projects/future-cities-catapult/. [Accessed 12 Mar. 2019].

Meulemans, W., Dykes, J., Slingsby, A., Turkay, C. & Wood, J. 2017. Small Multiples with Gaps. IEEE Transactions on Visualization and Computer Graphics, 23, pp381–390.

Slingsby, A., Kelly, M. & Dykes, J. 2014. . "Featured Graphic. OD maps for showing changes in Irish female migration between 1851 and 1911", Environment and Planning A, 46, pp2795-2797.

Tufte, E. 1983. The Visual Display of Quantitative Information, Graphics Press Cheshire, CT.

Visalingam, M. 1981. "The signed chi-score measure for the classification and mapping of plychotomous data", The Cartographic Journal 18 (1) pp32-43.

Wood, J., Dykes, J., Slingsby, A. 2010. "Visualization of Origins, Destinations and Flows with OD Maps", The Cartographic Journal 47 (2) pp117-129.