

Distr.
GENERAL

March 2013

ENGLISH ONLY

**UNITED NATIONS ECONOMIC COMMISSION
FOR EUROPE (UNECE)
CONFERENCE OF EUROPEAN STATISTICIANS**

Work Session on the Communication of Statistics
(Berlin, Germany, 27-29 May 2013)

Session 2: Changes in Communication

Data Visualisation for the Citizen User: Making Better Graphics Quicker

Prepared by Alan Smith OBE, Office for National Statistics (UK)

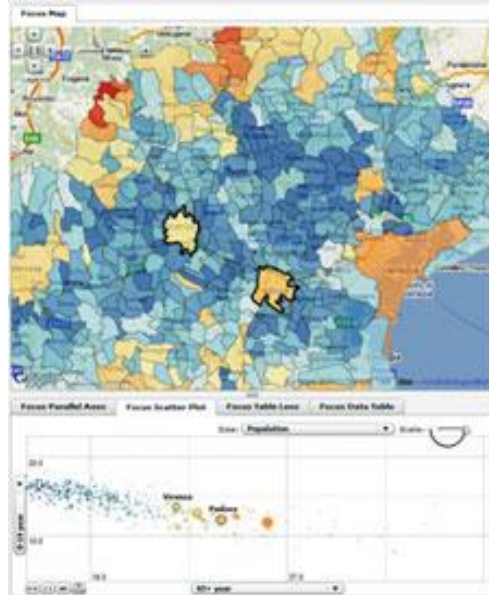
Data visualisation appeals to National Statistics Institutes (NSIs) because of its ability to engage users and increase the potential outreach of official statistics. But data visualisation is a broad field, with content types ranging from simple infographics (Figure 1a) through to sophisticated tools for exploratory data analysis (Figure 1b).

Figure 1 – broad range of data visualisation content types

(a) Infographic:
'Not so cut and dried' – Statistics New Zealand¹



(b) Visual Explorer:
Focus and Context eXplorer – NcomVA²



Data visualisation remains a problem area for many NSIs. There is a perception that the skills required for effective visualisation are in short supply within organisations – and there is a consistent problem with dedicating resource to visualisation, especially given the budgetary constraints faced by most European NSIs. However, *users* are increasingly expecting web content to be visual, engaging and personal, so developing a data visualisation capability remains high on the wishlist of many NSIs.

Data visualisation is increasingly being recognised as an inter-disciplinary activity³. Therefore, the main challenge for an NSI is not just to identify a dedicated data visualisation 'expert' but to assemble a team with the correct mix of appropriate skills and expertise. At a broad level, these skills could be categorised into 3 principal areas:

1. **Data skills** – extracting stories from data and writing about them
2. **Design skills** – concentrating on excellence and consistency in aesthetics
3. **Developer skills** – implementing dynamic web content using appropriate technical solutions

Then, there is the problem of software tools. For some organisations, given the plethora of 'off-the-shelf' tools, it might be tempting to simply choose one or more software products and implement them. But effective visualisations are 'context-dependent' - individual outputs might be linked by a consistent set of underlying principles, but it is difficult to industrialise their production around a single set of software rules. Instead, the aspiration must be for a flexible toolkit capable of dealing with wide-ranging content requirements.

For more complex dynamic and interactive graphics, momentum in the data visualisation community is clearly swinging away from proprietary tools like Adobe Flash towards open-source JavaScript-based frameworks which facilitate rapid content development and reuse. Such libraries can be seen to occupy a middle ground between 'pure code' and 'off-the-shelf' implementation. ONS' Data Visualisation Centre has recently been exploring some of these frameworks. In particular:

HTML5/CSS3 – emerging as the core technical framework of content for the web, offering the possibility for flexible content that ‘adapts’ well to different device displays (e.g. desktop, tablet, smartphone). HTML represents the ‘content’ layer, CSS represents the rules for ‘presentation’.

jQuery/jQuery UI – JavaScript is the open standard for adding interactivity to web applications. jQuery is a library for developing JavaScript applications more quickly and with built-in consistency for a full range of mobile and desktop web browsers. jQueryUI is an extension to this framework allowing for consistent User Interface objects, including many objects difficult to implement using HTML alone.

d3 – ‘Data-driven-documents’. In many ways, D3 is to Scalable Vector Graphics (SVG) what jQuery is to JavaScript: A compact, efficient library for producing web graphics that have data ‘embedded’ in them. D3 promises good performance along with support for sophisticated visualisation techniques and flexible animation/interaction. D3 includes built-in support for ‘layouts’ (standard data visualisations), but crucially allows the development of virtually any 2d graphical forms using code.

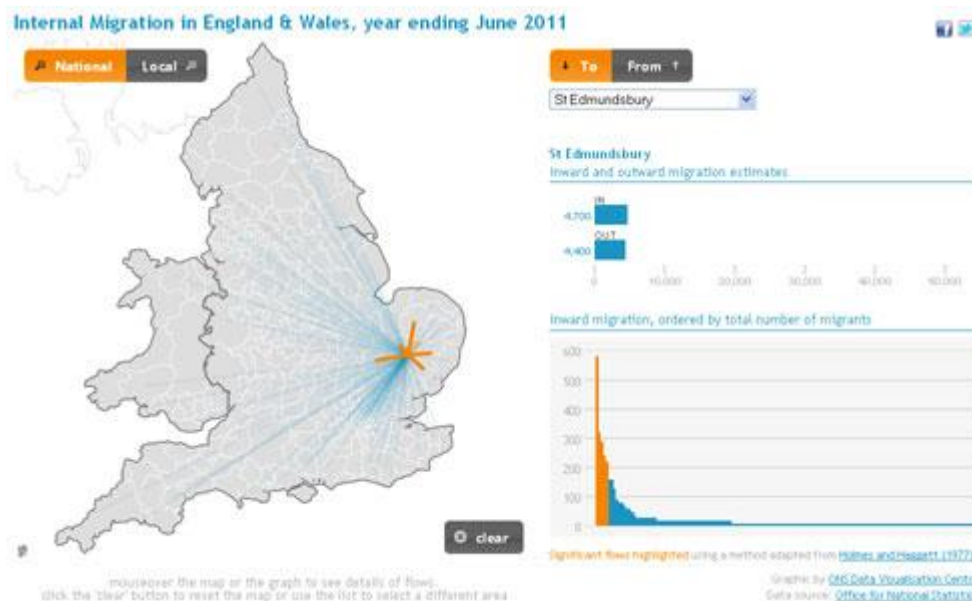
Modernizr – a library which detects the capability of a user’s web browser and offers, thereby offering content authors a framework for optimizing content based on those capabilities. For example, if a user’s device supports touch interactions (as opposed to being cursor-driven), as a visualisation loads, it can assign different rules for interaction optimised for touch, rather than a mouse.

ONS has a dedicated Data Visualisation Centre, located within the Strategy and Standards directorate. As part of its mission to research and define best practices in visualisation, it has recently been exploring these new technologies for interactive graphics. The experience of that work is reflected in 2 recent products:

1. **Dynamic migration flow map**⁴

An implementation of flow-mapping using HTML5/CSS/jQuery/jQueryUI and SVG. This project was a deliberate attempt to test the capability of modern web browsers to handle real-time visualisations of large datasets. The underlying flow dataset in this product contains a matrix of 121,000 cells of data. The visualisation not only maps those flows in real time, but carries out a test to identify significant flows, using a technique identified by Peter Hagget and John Holmes⁵ (Figure 2) – real evidence of the performance improvements made by browser manufacturers in recent years.

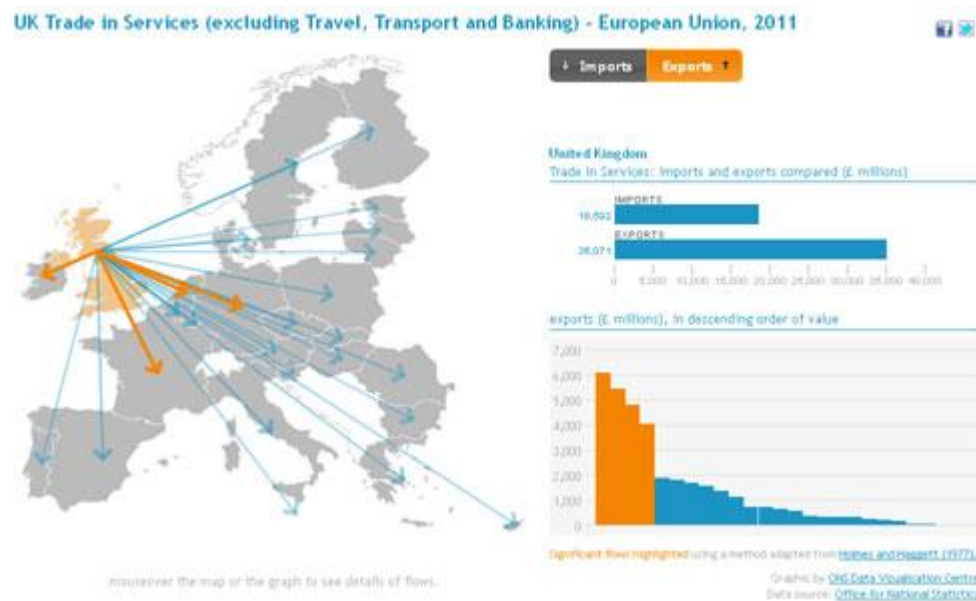
Figure 2 – Internal migration in England & Wales



The map proved extremely popular on release – so much so that it was syndicated by a national newspaper⁶. Syndication of interactive content has been a deliberate target of the ONS Data Visualisation Centre recently and has proven to be an extremely effective way of increasing outreach. Interestingly, most UK media outlets would now consider HTML5 to be a more syndication-friendly format than something like Flash, in part because many users have devices (e.g. iPads) which only support open standards.

Additionally, the flexibility of the underlying code makes it easy to re-use this visualisation with a wide variety of origin-destination matrices. For example, creating a map of significant flows in UK International Trade in Services⁷ required only minimal work to replace the geography and data/meta-data components (Figure 3).

Figure 3 – UK Trade in Services



2. Population pyramid

Currently unpublished, this prototype graphic re-presents ONS' existing Flash-based animated population pyramids into HTML5, using the full capabilities of jQuery/d3/Modernizr to render an interactive graphic which works both on desktop and mobile/touch-enabled web browsers (Figure 4). For example, users of iOS devices (iPhone, iPad, iPod Touch) get a touch-enabled interface which exploits touch interactions (single finger touch to get information for a single age band, a long press and 'swipe' to select a multi-age group).

Figure 4 – Animated population pyramid on Desktop (a) and Mobile (b) browsers



Our experience in both projects helped us to confirm several specific key points:

- The built-in capabilities of the HTML5 family of technologies, when coupled with sophisticated JavaScript libraries, and the performance of modern web browsers, render proprietary technologies like Adobe Flash largely irrelevant. JavaScript libraries allow an NSI to 'stand on the shoulders of giants' and produce better, cross-platform web content more quickly than ever before.
- There remains a 'pre-requisite' entry point in terms of skills to code dynamic and animated content – but these skills are based on industry-standards, many of which will already exist within an NSI. An inter-disciplinary approach makes better use of existing resources.
- If NSIs enter the 'rich content' arena, the content should be good enough to syndicate to national media – if it isn't at that level of quality, users will not be thankful and NSI efforts will be largely wasted.
- Syndication represents a huge opportunity for NSIs to increase the user outreach. Our own experience at ONS of syndicating our interactive content suggests a conservative estimate of around 1,000% increase (compared to hosting on the NSI website alone). This figure will rise as more media platforms are added.

More broadly, data visualisation offers NSIs an opportunity to exploit their expertise in formats which boost user engagement and readership. It also carries with it the highly desirable side effects of boosting relationships with the media and reputational benefits virtually everywhere else. A final note of caution, however, is that these visualisations should be centred on the expertise of the NSI, *not* based on a notion of style over content – others do that better.

REFERENCES

'Not so cut and dried – Statistics New Zealand, retrieved from http://www.stats.govt.nz/browse_for_stats/economic_indicators/CPI_inflation/haircuts-2012-infographic.aspx

NComVA Focus and Context eXplorer, retrieved from <http://www.ncomva.com/products/customized/>

Kirk, Andy (2012) Data Visualization: a successful design process. Packt publishing.

Flow map of internal migration in England & Wales, retrieved from <http://www.neighbourhood.statistics.gov.uk/HTMLDocs/dvc25/index.html>

Hagget, P. and Holmes, J. (1977) Graph Theory Interpretation of Flow Matrices: A Note on Maximization Procedures for Identifying Significant Links, Wiley retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1538-4632.1977.tb00591.x/abstract>

Interactive graphic: 230,000 quit London for the Regions. Daily Telegraph, retrieved from <http://www.telegraph.co.uk/earth/greenpolitics/population/9565594/Interactive-graphic-230000-quit-London-for-the-regions.html>

Flow map of UK Trade in Services, retrieved from <http://www.neighbourhood.statistics.gov.uk/HTMLDocs/ITIS/index.html#UK,nat,from>