

DataViz: improving data visualisation for the public sector





DataViz: improving data visualisation for the public sector

September 2009

Tom Smith, Stefan Noble, David Avenell and Graham Lally, Oxford Consultants for Social Inclusion (OCSI) Ltd Department for Communities and Local Government The findings and recommendations in this report are those of the authors and do not necessarily represent the views of the Department for Communities and Local Government.

Department for Communities and Local Government Eland House Bressenden Place London SW1E 5DU Telephone: 020 7944 4400 Website: www.communities.gov.uk

© Queen's Printer and Controller of Her Majesty's Stationery Office, 2009

Copyright in the typographical arrangement rests with the Crown.

This publication, excluding logos, may be reproduced free of charge in any format or medium for research, private study or for internal circulation within an organisation. This is subject to it being reproduced accurately and not used in a misleading context. The material must be acknowledged as Crown copyright and the title of the publication specified.

Any other use of the contents of this publication would require a copyright licence. Please apply for a Click-Use Licence for core material at www.opsi.gov.uk/click-use/system/online/pLogin.asp, or by writing to the Office of Public Sector Information, Information Policy Team, Kew, Richmond, Surrey TW9 4DU

e-mail: licensing@opsi.gov.uk

If you require this publication in an alternative format please email alternativeformats@communities.gsi.gov.uk

Communities and Local Government Publications Tel: 0300 123 1124 Fax: 0300 123 1125 Email: product@communities.gsi.gov.uk Online via the Communities and Local Government website: www.communities.gov.uk

September 2009

Product Code: 09RPD06059

ISBN 978-1-4098-1766-6

Contents

Section 1	Introduction	4
	1.1 Introduction	4
	1.2 How we have carried out this project	5
	1.3 Acknowledgements	6
Section 2	Stage 1: What visualisations are being used, both by public sector, and elsewhere?	7
	2.1 Discussion	7
	2.2 Outputs	7
Section 3	Stage 2: Is there a useful typology of visualisation solutions (for public sector users)?	11
	3.1 Discussion	11
	3.2 Outputs	15
Section 4	Stage 3: How effective are particular visualisation solutions in supporting public sector research and decision-making?	16
	4.1 Overview	16
	4.2 Practical steps to good data visualisation	16
	4.3 Outputs	17
Section 5	Stage 4: What are the most appropriate visualisation options for particular purposes?	18
	5.1 Discussion	18
	5.2 Outputs	20
	Appendix A Call for evidence	21
	Appendix B Topic guide for interviews	23
	Appendix C Acknowledgements	25

Section 1 Introduction

1.1 Introduction

Background

- 1.1.1 Effective visualisation of data can help users explore and understand data, and also communicate that understanding to others:
 - *Exploring and analysing*: Data visualisation can be of great use in the analysis process, enabling researchers to explore datasets to identify patterns, associations, trends and so on. Good visualisation should be seen as a central tool in carrying out analysis
 - *Presenting and communicating*: Good data visualisations should help others make robust decisions based on the data being presented: They should provide an effective representation of the underlying data, to help answer a particular question at hand. Communicating data in this way might: support senior decision-makers engaged in strategic *planning*, or *operations* managers needing to understand where services could be improved, or managers wanting to monitor *performance*

Project scope

- 1.1.2 Communities and Local Government (CLG) commissioned this project to better understand the potential for visualisation solutions to analyse and communicate intelligence, and provide guidance for national and local analysts on making best use of visualisation.
- 1.1.3 In particular, CLG were looking to:
 - assess and appraise the current range of visualisation solutions
 - provide advice on the decision making process for selecting visualisations at different levels of the government delivery chain.
- 1.1.4 In this project, we have addressed four central questions in order to provide advice in using visualisation effectively in both exploring/analysing, and presenting/communicating data:
 - 1. What visualisations are being used, both by public sector, and elsewhere?
 - 2. Is there a useful typology of visualisation solutions (for public sector users)?

- 3. How effective are particular visualisation solutions in supporting public sector research and decision-making?
- 4. What are the most appropriate visualisation options for particular purposes?

Key audiences

1.1.5 The primary audience for the project is local, regional and national information and research teams. The website and outputs are also intended to be useful to all researchers engaged in visualising and presenting data – but may not provide examples of more technical visualisations used in specific areas.

Project outputs

- 1.1.6 The primary project output is the data visualisation website DataViz www.improving-visualisation.org. This report acts as project record, and users should view the website for the project outputs. The website contains:
 - examples of visualisations with associated descriptive 'metadata', as well as assessment of benefits and pitfalls to avoid for each visualisation
 - case studies of visualisations supporting public sector research and decision-making
 - a typology of visualisations, based on policy/research questions of relevance to public sector users
 - menu-driven filters, to enable users to identify appropriate visualisations for particular purposes (based on the typology policy/research questions)
 - links to visualisation resources on the web; and
 - functionality for users to add their evaluations of each visualisation.

1.2 How we have carried out this project

- 1.2.1 Our project work has been based on:
 - *Call for evidence*: Contact with wide range of stakeholders to identify good (and bad) examples of visualisation, through emails to user-groups, and forum and blog postings. See Appendix A for the call to evidence text.
 - *In-depth consultations*: Face-to-face and telephone interviews, with a smaller number of selected local, regional and national stakeholders, as well as commercial partners. These covered issues such as extent to which visualisation is used in analysis, and policy/decision-making. See Appendix B for the topic guide.
 - Visualisation literature research.
 - Development of the project website DataViz and database. See www.improving-visualisation.org

1.3 Acknowledgements

- 1.3.1 We would like to thank the CLG team and project Steering Group for their help with the project: James Geehan (CLG), Robert Rutherfoord (CLG), Alistair Edwardes (CLG), Alan Smith (ONS), Professor Mike Batty (UCL CASA)
- 1.3.2 A full list of acknowledgements is provided in Appendix C.

Section 2

Stage 1: What visualisations are being used, both by public sector, and elsewhere?

2.1 Discussion

- 2.1.1 Our starting point has been to identify what visualisations are being used by the public sector and elsewhere. It is clearly not possible to identify each and every example in existence, so we have focused on identifying examples of a wide range of visualisation types.
- 2.1.2 Based on the brief, the project has identified visualisations appropriate for small area data, e.g. Output Area and Super Output Area, as well as higher scales of geography. In addition, we have looked at visualisations that help understand and communicate trends and dynamics over time, and benchmarking comparisons between areas/groups and services both of these issues are critical to local areas wanting to improve service delivery and meet LAA targets.
- 2.1.3 We have also identified examples of "exploratory visualisation"; that is, when analysts are asked to mine a data-set to uncover unknown patterns or trends. The research question or decision-making process maybe unknown, and only through engagement and experimentation with different visualisations can these be uncovered.

2.2 Outputs

- 2.2.1 The main outputs were:
 - the project website, with a wide range of visualisation examples (currently 170) provided as separate resources. Each provides metadata, including what the visualisation shows, common pitfalls to avoid
 - short case study summarising the types of data visualisation reported in three key local strategies and plans for example, Sustainable Community Strategies, Local Area Agreements and Joint Strategic Needs Assessments (this is live on the website).
- 2.2.2 The next pages show screenshots from the website, showing the gallery of all visualisations.





		Last edited by stefan
Crive Contraction Crive Cri	A Strategy (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)	Users Performance management Researchers Questions Which neighbourhoods are faring worse than the average across the area? How vulnerable is an area to external changes e.g. How vulnerable is an area to external changes e.g. How vulnerable is an area to external changes e.g.
Visualisation	Crime mapping - GIS hotspots, and temporal patterns	How is our council performing?
Description	Crime mapping example - using geographic hotspots alongside temporal patterns (by hour of day) to highlight particular patterns and identify appropriate intervention.	Visual types
Source	Home Office, Crime Mapping: Improving Performance	<u>bar chart</u>
Source ref	p8	Data types
Source type	Area Profile	
Source link	http://police.homeoffice.qov.uk/publications/performanc e-and-measurement/crime_mapping_1.pdf	area average
Interactive?	No	
Benefits & pitfalls to avoid	* Bar-charts imply that data is discrete. If your data is changing over time you might consider a line graph instead.	
	* When there are many bars, labeling becomes problematic	
	* Differences between values may be exaggerated if axes do not go to zero	t
	* Shaded maps can emphasise large areas much more than small ones - for example highlighting rural areas over urban, if fixed population areas such as Super Output Areas are shown. Consider providing an alternate mode in which values are represented by circles, or values are scaled by the area size (ie showing density)	

Section 3

Stage 2: Is there a useful typology of visualisation solutions (for public sector users)?

3.1 Discussion

- 3.1.1 In order to provide advice to users on the most appropriate visualisation to use (Stage 4), we need to understand the types of visualisation solution used for different purposes. We have developed a basic typology of visualisation solutions relevant to public sector users.
- 3.1.2 Based on our initial proposal and discussions with the steering group, we have identified that the most appropriate typology is one based on:
 - the purposes that visualisation is being used to help understand; and
 - the audience for the visualisation¹.
- 3.1.3 In addition, we have 'tagged' each of the examples on the website with visualisation type (eg bar-chart, map etc) and have used this to provide an additional way for users to select and find examples based on visual types.

Themes and questions for the visualisation typology

- 3.1.4 We have organised the example visualisations under four themes relevant to the key audience of local, regional and national information and research teams (see figure below):
 - Understanding needs "What do we need to do?": For public services, "understanding needs" requires addressing a set of questions around to get a clear understanding of what communities/users/citizens are like.
 - *Improving service delivery "How do we do it?"*: Improving service delivery is fundamentally about understanding users, and how services can be better tailored and targeted to their needs.
 - *Performance and reporting "How did we do?"*: Covering questions such as whether targets are being met, how performance compares with others, whether there has been improvement in outcomes and whether there has been user satisfaction with performance.

¹ Our classification based on role/purpose is also used by Stephen Few (2006) in *Information Dashboard Design: The effective visual communication of data*, O'Reilly.

• *Data exploration*: Exploring your data with no specific research question in mind; looking at patterns in the data, relationship between datasets, whether findings are statistically significant etc.



3.1.5 Under each of these themes, we have identified a series of policy and research areas. In turn, each of these areas is linked to different types of visualisation. The tables below identifies the policy research areas (and illustrative questions), and the visual types linked to these.

Policy and research areas	Example (illustrative) questions	Chart types (used to link to examples)
What are our neighbourhoods like?	 What is the demographic composition of our neighbourhoods? How do our neighbourhoods fare on a range of indicators? 	 choropleth map bar chart box plot range chart pie scarf chart population pyramid treemap
How do our communities compare?	 How do needs compare with (similar) benchmark areas? Which neighbourhoods/groups are faring worse than the average? Where are the priority areas? 	 bar chart choropleth map univariate chart smooth statistical map bubble map deviation bar graph spider thematic table
How are things changing over time?	 Are deprived areas/groups "narrowing the gap"? What are recent socio-economic trends? How many people are on benefits for short periods? Are people cycling between benefits and low pay? What are projected future trends? How are needs likely to change in future? 	 line chart area chart spider flow map
Who is coming in to (and moving out from) our area?	 What are the levels of inward and outward migration like? What are the commuting patterns (in-flow and out-flow) from our area? 	flow mapcomparative chart
Who is most at risk? What factors are associated with high levels of need?	 Who is most at risk? What factors are associated with high levels of need? What is the distribution of need? How do health inequalities vary by gender? By deprivation level? How do areas with specific socio-economic issues fare economically? What factors are associated with what outcomes? Which groups are appropriate for preventative work? 	 scatterplot histogram box plot] clustered bar chart choropleth map proportional area chart comparative chart quadrant plot
Access to services	What groups and areas have poor access to services?	 choropleth map smooth statistical map bubble map
Are differences significant?	 Are we doing significantly worse/better than benchmark areas on key indicators of need? Are changes over time significant? 	 confidence intervals funnel
What are our citizen priorities?	 What are the key priorities for citizens? Which groups have particular views? 	 qualitative bar chart traffic light

Understanding needs – policy and research areas

Policy and research areas	Example (illustrative) questions	Chart types (used to link to examples)
How do our services compare?	 How do costs of services compare with (similar) benchmark areas? How do levels of service demand and take- up compare with (similar) benchmark areas? Where are services most stretched (highest levels of take-up)? 	 bar chart choropleth map spider
Understanding customers ('insight')	 Who is most/least satisfied with services? What are the key priorities for citizens and users in improving services? Which groups have particular views? What groups and areas have poor take-up of services? Preferred services and communication channels? 	 tag cloud qualitative relationship diagram bar chart proportional area chart
User access to services	 Where are services located? What groups and areas have difficulties accessing services? 	 point location map choropleth map smooth statistical map bubble map
Who are our service users	 Who are our service users? How do service users compare with the overall population? Where do our service users come from? What is our travel to work area? Where are commuters coming from? 	 bar chart proportional area chart flow map

Improving service delivery – policy and research areas

Performance and reporting – policy and research areas

Policy and research areas	Example (illustrative) questions	Chart types (used to link to examples)
How are we performing on targets?	 Are we meeting targets? What targets are we failing on? What service areas are failing to meet targets? 	 scorecard bar chart confidence intervals traffic light range bar chart
How do we compare?	 How does our partnership performance services compare with (similar) benchmark areas? How does performance of services compare with (similar) benchmark areas? Are we providing Value for Money? 	 bar chart choropleth map range bar chart
How are outcomes changing over time?	 What do performance trends look like? Are our programmes working? Are area and individual outcomes improving? Are deprived areas/groups "narrowing the gap"? 	line charttraffic lightarea chart
Satisfaction with services	How satisfied are service users?Who is most/least satisfied with services?	bar chartqualitative

Research questions	Chart types (used to link to examples)
How do we compare (with other areas)?	bar chartchoropleth map
What is changing over time ?	 animations flow maps time line line chart
What is the distribution (of socio-economic indicators)?	 histogram scatterplot bubble map smooth statistical map choropleth map
What are the components of an indicator (for example, how does the workless total split between jobseekers, incapacity benefits, lone parents etc)	proportional area chart
What is the relationship between different socio- economic factors?	 scatterplot quadrant plot clustered bar chart choropleth map comparative chart
How significant are differences?	 confidence intervals funnel
How can I interactively explore a dataset?	Interactive
How can I visualise qualitative (text) data?	qualitative
How can I visualise categorical data (for example, neighbourhood area classifications or service locations)	 categorical (data-type tag) point location (data-type tag)

Data exploration – policy and research areas

3.2 Outputs

- 3.2.1 All examples on the website have been 'tagged' with relevant themes and questions from the typology (note, each visualisation can lie under more than one theme-question, and each question can contain more than one example).
- 3.2.2 Drop-down lists are provided to enable users to select all visualisations belonging under each theme-question.

Section 4

Stage 3: How effective are particular visualisation solutions in supporting public sector research and decision-making?

4.1 Overview

- 4.1.1 There is no objective, reliable, way of assessing the effectiveness of visualisations for supporting public sector research and decision-making (or indeed any other use of visualisations). In addition, the role played by information, including visualisation, in decision-making is a particularly murky area².
- 4.1.2 We have approached assessing the visualisations in two ways:
 - practical steps to good data visualisation, case study identifying series of common pitfalls to avoid (with "before" and "after" visualisations)
 - enabling users to evaluate each visualisation on the project website, based on three dimensions (each on a 5-point scale). The ratings are displayed on the website, along with facility for users to add their own ratings:
 - Impact for the specific audience
 - Accuracy of portraying the underlying data.

4.2 Practical steps to good data visualisation

4.2.1 Based on our review of the visualisation literature, we have identified a series of practical steps for accurately showing the data. We have structured these around three areas – *audience*, *form* (i.e. size and shape), and *colour*:

Audience	• Think about how to present data to decision makers – an example from Leicestershire	
	 Help the audience by being selective in what you show – limit the number of categories in a visualisation 	
Form (i.e. size and	Don't use visualisation effects such as 3D that can hide the data	
snape)	• Do not vary the size of objects in graphs, except to convey difference in values	
	Don't miss zero off the axis scale	
	• For bar charts, set the base of the bars to zero (not the lowest value)	
	Consider using bar charts, not pie charts, to present composition data	
	• Avoid using line charts where data is only available for a small number of timepoints	
Colour	When choosing colours to use, understand the principles of colour design	
	• When using colours in visualisations, ensure that colours are sufficiently contrasting to distinguish between data values	
	Where colour is needed, use solid blocks of colour and avoid fill patterns	
	Avoid using strong or bold colours for the background in a visualisation	
	• When creating choropleth maps, choose colours to help users identify patterns and relationships between areas	

4.2.2 Under each of the practical steps above, we have produced "before" and "after" visualisations, identifying how the practical step can improve the effectiveness of the visualisation.

4.3 Outputs

- 4.3.1 Outputs include:
 - assessment of each of the visualisation solutions based on benefits and drawbacks (added to the project website resources)
 - practical steps to good data visualisation, case study identifying series of common pitfalls to avoid (with "before" and "after" visualisations)
 - website facility added to enable users to rate visualisations on relevance, impact and accuracy (with resources displaying the user ratings).

Section 5

Stage 4: What are the most appropriate visualisation options for particular purposes?

5.1 Discussion

- 5.1.1 The typology identified in Stage 2 above provides the framework for the process by which we identify appropriate visualisation options to users. Linking the typology to the questions asked by users has the benefit that our typology and classification of visualisation solutions is *closely matched to our analysis of most appropriate visualisation solutions for a particular purpose*.
- 5.1.2 Our guidance to users follows the typology set out in Stage 2 (see figure below):



5.1.3 Selecting the step-by-step link on the navigation menu



All levels High

> Intermediate Low

Visualisations that are more technically difficult to produce and require higher levels of statistical expertise to interpret. Visualisations that require longer study or analysis to interpret the findings.

Visualisations that require low levels of expertise to understand the findings



5.2 Outputs

- 5.2.1 The outputs include:
 - toolkit facility for users to select guidance on appropriate visualisation solutions, live on the website
 - guidance on the website to users on how to identify appropriate visualisation solutions.

Appendix A Call for evidence

A.1.1 The call for evidence was sent out to email lists, forums, blogs including the CASA DigitalUrban blog³. This was also presented at the January 2009 Local Information Systems event.

Improving data visualisation for public sector decision-makers and researchers

Call for Evidence

- A.1.2 We are carrying out a review of data visualisation how data is presented and brought to life using charts, graphs, maps, timelines, animations etc – for public sector decision-makers and researchers. The work will identify good practice in visualising data, particularly for:
 - analysing and presenting small area data (eg, below District level)
 - understanding and communicating trends and dynamics over time
 - benchmarking comparisons (between areas, groups and services).
- A.1.3 We are looking for good examples of visualisation that have:
 - made an impact/helped communicate findings to decision-makers
 - helped researchers understand the story behind the data.
- A.1.4 The project/website will be an excellent opportunity to show case your own examples of good visualisation, and help contribute to common standards of good practice in the research community. *If you have any examples, we would love to hear from you!* (contact details below)

Background

- A.1.5 The project has been commissioned by Communities and Local Government, who are looking to understand the potential for visualisations to analyse and communicate data, and provide guidance for national and local analysts on making best use of visualisation.
- A.1.6 Work on the project will explore questions including:
 - What visualisations are being used by the public sector? What can we learn from elsewhere?

³ http://digitalurban.blogspot.com/

- Is there a useful typology (classification) of visualisation techniques for public sector users?
- How effective are particular types of visualisation in supporting public sector research and decision-making?
- What are the most appropriate visualisation techniques for particular purposes?
- A.1.7 The primary output from the project will be a web application, to help users identify strengths and weaknesses of visualisations for particular purposes, also linking to examples and supporting material.

Further information

See www.ocsi.co.uk/visualisation for project details, or contact the project leads below.

James Geehan, (Project manager) Communities and Local Government tel: 0207 944 4857 James.Geehan@communities.gsi.gov.uk Tom Smith (Project lead) OCSI tel: 01273 201345 tom.smith@ocsi.co.uk

Appendix B Topic guide for interviews

B.1.1 This topic guide was sent to all interviewees in advance.

Improving data visualisation for public sector decision-makers and researchers: Topic guide for consultations

About the project

- B.1.2 We are carrying out a review of data visualisation how data is presented and brought to life using charts, graphs, maps, timelines, animations etc for public sector decision-makers and researchers.
- B.1.3 The work will identify good practice and examples in visualising data, particularly for:
 - analysing and presenting small area data (eg, below District level)
 - understanding and communicating trends and dynamics over time
 - benchmarking comparisons (between areas, groups and services).

About this guide

B.1.4 This topic guide covers the areas we would like to discuss. It is NOT a questionnaire – we will explore these questions in our discussion.

We are looking for (short) case studies of visualising information, that have:

- 1. made an impact/helped communicate findings to decision-makers
- 2. helped researchers understand the story behind the data

For each example

- **Background** what is the issue/how is visualisation intended to help
- **How used**: Where in the decision-making process the visualisation was used? What audiences? And how used to help?
- **Impact**: summary of the key ways that the visualisation helped (eg in decision-making etc)
- **Benefits** of the particular visualisation(s) used and any weaknesses with previous visualisations?

- **Details** what was setup? By who? Costs? Technical implementation details?
- Further contact

Thank you for your help

Appendix C Acknowledgements

Name	Organisation	Role
Alan Smith	Office for National Statistics	Data Visualisation Unit
Alistair Edwardes	Communities and Local Government (CLG)	Lead Spatial Analyst
Andrew De'Ath	North East Improvement & Efficiency Partnership	Programme Manager
Andrew Hudson-Smith	Centre for Advanced Spatial Analysis (CASA), University College London	Senior Research Fellow
Andy Hoskins	City Vision Networks Ltd	Director
Christopher Pettit	Department of Primary Industries, Australia	Research Manager – Spatial Information
Dan Jellinek	Headstar	Editor e-Government Bulletin
Dave Tuffery	Bristol City Council	Sustainable City Group
David Burden	Daden Ltd	Director
David Russell	Fluent Technology Ltd	Director
Derrick Johnstone	Educe Ltd	Director
Emma Cunliffe	South Tyneside Council	Information Support Officer
Emma Williams	Linden Labs (Second Life)	UK Operations Manager, Second Life
James Geehan	Communities and Local Government (CLG)	Senior Statistician
James Lawrence	Coventry City Council	Corporate Research, Information and Consultation
John Dawson	marketingQED	Director
John Fisher	Local Futures	Director
John Maslen	Geowise	Local Government Director
Juliet Whitworth	Local Government Association	Head of Commissioning and Research
Ken Feaster	National Institute on Money in State Politics	Systems Designer
Kim Rees	Periscopic Ltd	Partner
Laura Edwards	Local Futures	Local Knowledge Manager
Lincoln Smith	ioLabs Inc	Director

Name	Organisation	Role
Michael Batty	Centre for Advanced Spatial Analysis (CASA), University College London	Director
Mike Franks	Oldham Council	Research Analyst (GIS)
Neil McSweeney	Halton Borough Council	Principal Research Officer
Paul Gutherson	CfBT Education Trust	Consultant
Paul Wilson	Scott Wilson Ltd	Digital Media Manager
Rob Wilson	Newcastle University Business School (NUBS)	Deputy Director of KITE Research Centre
Robert Radburn	Leicestershire County Council	Senior Research Officer
Robert Rutherfoord	Communities and Local Government (CLG)	Head of Neighbourhood Intelligence
Seth Finegan	Informed Solutions	Local Government Director
Simon Rogers	The Guardian	News editor, graphics; Editor, Guardian Datablog and Datastore
Steven Rogers	Office for National Statistics	Data Visualisation Unit
Tommy Davies	Cumbria County Council	Business Intelligence Manager
Wendy Pontin	Norfolk County Council	Principal Information Analyst
Wesley Grubbs	Pitch Interactive, Inc	President
Xavier Sosnovsky	European Central Bank	Statistical Information Systems Expert

