Good Practice Guide:
An introduction to data visualisation

Client: Big Lottery Fund
Project: Fulfilling Lives: Supporting people with multiple needs
Date: November 2015
Author: Jon Adamson

This guide has been prepared by CFE Research for the local partnerships delivering projects as part of the Big Lottery Fund initiative *Fulfilling Lives (Multiple Needs)*. It provides an introduction to data visualisation and how key principles of this are being used to support the national evaluation. Further useful sources of information are given at the end of this document.

Introduction

For the evaluation of *Fulfilling Lives (Multiple Needs)*, funded projects are required to collect and report information in line with a Common Data Framework (CDF). The information collected feeds into both the national evaluation of all 12 projects and into each project's own local evaluation. With projects running for between 5 and 8 years a huge amount of data will be collected. How that information is analysed and presented back to projects and a wider audience will play a crucial role in both the ongoing development of delivery models and the overall assessment of what impact has been achieved.

What is data visualisation?

Data visualisation (data viz) is the practice of visually presenting mainly quantitative data in a way that facilitates understanding and interpretation. Data viz first emerged in the late nineteenth century when the likes of Charles Minard, William Playfair and Florence Nightingale used statistical graphics and illustrated diagrams to convey a specific message with the intention of bringing about a specific change. More recently, development of new technology and software has facilitated the widespread use of data viz to aid both the analysis and understanding of information to a wide variety of different audiences.

Data visualisation and infographics

It is useful to draw a distinction between 'data viz' and the production of 'infographics', but whilst there are clear areas of overlap, it is a contentious area of debate amongst protagonists of both forms. Data viz is primarily concerned with the concise and accurate representation of data to convey a particular finding or important message. Data viz is included in many different outputs from research, including academic papers, presentations, evaluations and information dashboards. Infographics generally combine several different findings or messages – both quantitative and qualitative – into a particular story. They are often produced by graphic designers or journalists, rather than researchers and are published as separate outputs which stand-alone from other reports. Notwithstanding the aforementioned commonality between the two forms, this guide is primarily focused on data viz rather than infographics.

Principles of data viz

There is much debate around what constitutes good data viz and how to do it. Whilst there are no specific rules for data viz, there are some broadly accepted principles of presenting information well.

1. **Only use tables and charts where they add to the readers' understanding** of the point you want to make. Don't fill up pages with charts showing one or two pieces of information. For example, a sentence describes the difference between male and female for a particular measure well enough; it does not warrant a chart. Use tables to provide precise data and use charts to show patterns in large amounts of data clearly and concisely.
2. **Identify the main point you want readers to understand** from your table or chart and make it as easy as possible for the reader to see your point and understand it.
3. **Simplify**. Think about all the ink on your tables and charts and whether it is necessary for it to be there. For example, don't use thick, dark gridlines on tables; subtle shading is much better.
4. **Think about the best way of ordering the information** in your tables and charts to get your message across. For example, sorting values A-to-Z on their labels is rarely the best order.
5. **Keep comparisons close**. If you want your reader to compare certain information, make it easy for them to do so. Consider the use of 'small multiples' where several charts are placed together which allows for easy comparison.
6. **Show the data**. Don't 'data decorate' charts with an array of different colours and shading. Use colours sparingly and meaningfully – i.e. to represent something. Avoid using pie-charts or other representations that distort the data and make it harder to get the message.
7. **Consider how your main audience will access your tables and charts**: on screen, at a presentation, in a paper report? Integrate charts with other text or speech. Describe and explain the charts, annotate key points with labels where useful to do so without cluttering the chart.

DarkHorse Analytics have produced two good examples of how tables and charts can be improved by de-cluttering and simplifying them (follow the links for animation and slides):

- tables: https://darkhorseanalytics.com/blog/clear-off-the-table/

Before

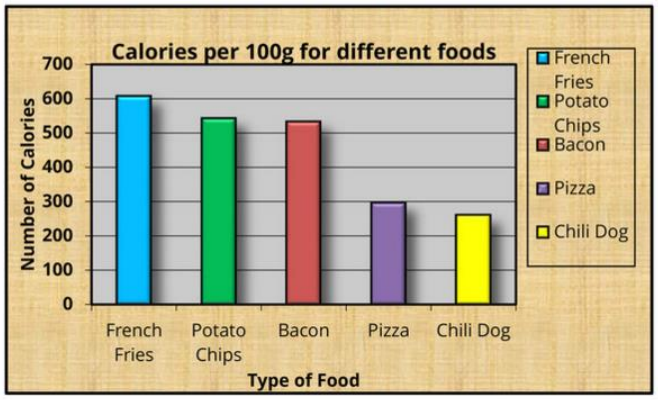
Role	Name	Year of the...	Debut	Number of Fans	Takedown Rate
Face (The Hero)	The Ultimate Warrior	Tiger	May-2011	97320.00	86.2
Face (The Hero)	Hulk Hogan	Oxen	Jan-2008	988551.00	61.978
Face (The Hero)	Macho Man Randy Savage	Monkey	Feb-2008	157618.00	59.29
Face (The Hero)	Hacksaw Jim Duggan	Pig	Mar-2008	30300.00	53.4332
Face (The Hero)	Superfly Jimmy Snuka	Dragon	Mar-2008	12341.00	52.7
Heel (The Bad Guy)	Rowdy Roddy Piper	Rooster	Jun-1968	71645.00	45.4
Heel (The Bad Guy)	The Million Dollar Man Ted DiBiase	Rat	Apr-1975	449342.00	43.7689
Heel (The Bad Guy)	Mr. Perfect Curt Henning	Rat	May-1980	13773.00	38
Heel (The Bad Guy)	Jake the Snake Roberts	Snake	Jul-1975	5609.00	37.99
Jobber (The Unknown)	Brad Smith	Sheep	Aug-2008	1103.00	36.316
Jobber (The Unknown)	Ted Duncan	Sheep	Aug-2008	200.00	33.61
Jobber (The Unknown)	Joey the Uber Nerd Cherdarchuk	Snake	Aug-2008	5.00	21.0196

After

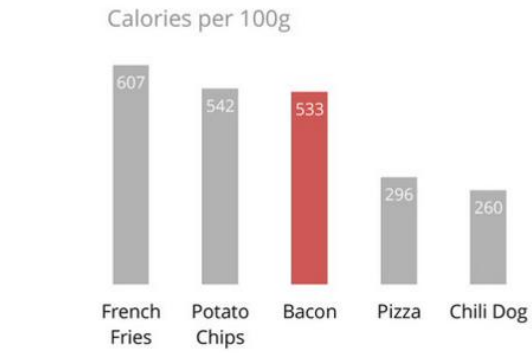
Role	Name	Year of the...	Debut	Thousands of Fans	Takedown Rate
Face (The Hero)	The Ultimate Warrior	Tiger	May-2011	97.3	86.2
	Hulk Hogan	Oxen	Jan-2008	988.6	62.0
	Macho Man Randy Savage	**Monkey**	**Feb-2008**	**157.6**	**59.3**
	Hacksaw Jim Duggan	Pig	Mar-2008	30.3	53.4
	Superfly Jimmy Snuka	Dragon	Mar-2008	12.3	52.7
Heel (The Bad Guy)	Rowdy Roddy Piper	Rooster	Jun-1968	71.6	45.4
	The Million Dollar Man Ted DiBiase	Rat	Apr-1975	449.3	43.8
	Mr. Perfect Curt Henning	Rat	May-1980	13.8	38.0
	Jake the Snake Roberts	Snake	Jul-1975	5.6	38.0
Jobber (The Unknown)	Brad Smith	Sheep	Aug-2008	1.1	36.3
	Ted Duncan	Sheep	Aug-2008	0.2	33.6
	Joey the Uber Nerd Cherdarchuk	Snake	Aug-2008	0.0	21.0

- charts: http://darkhorseanalytics.com/blog/data-looks-better-naked/

Before



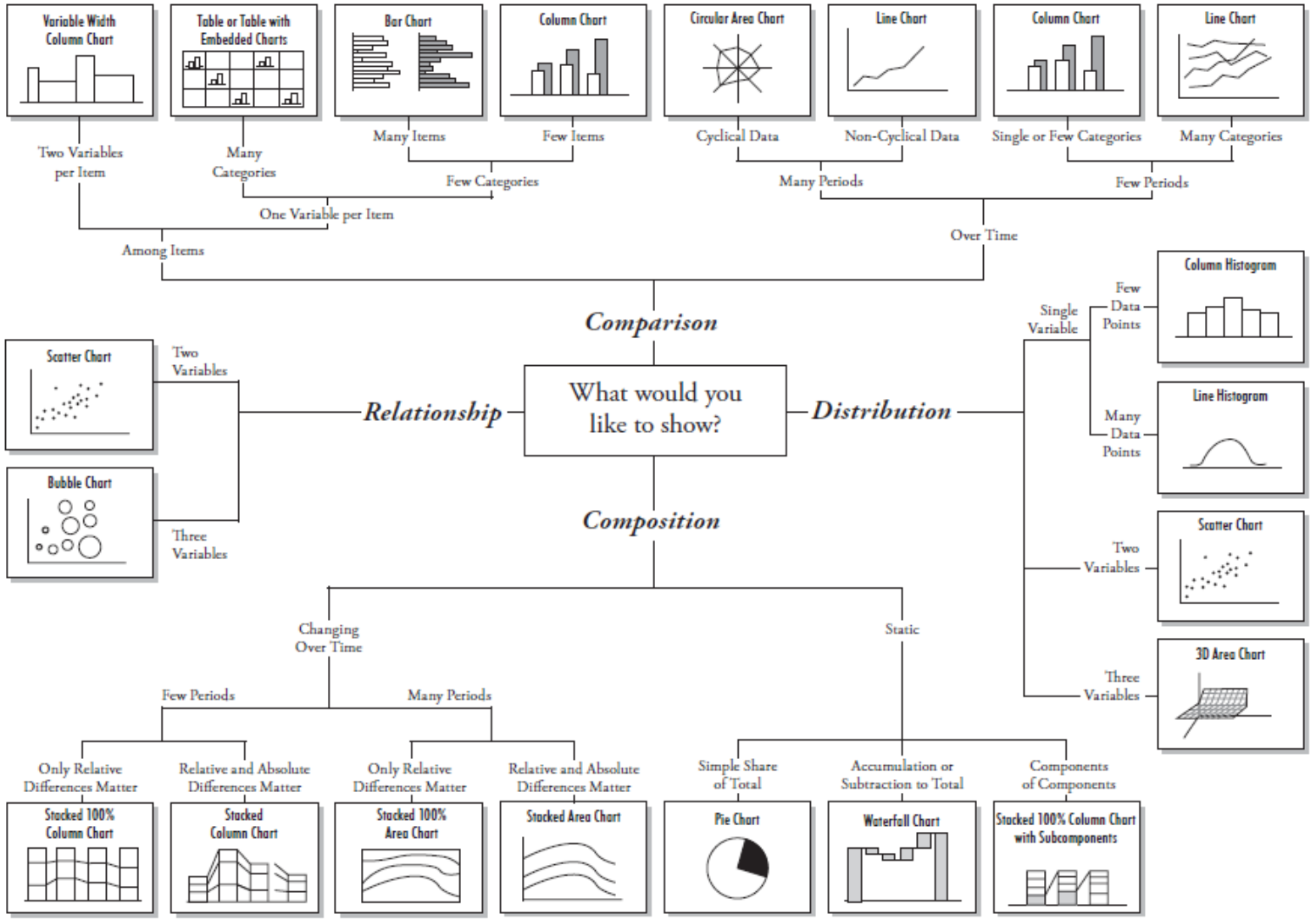
After



Using the right chart

Choosing how to present data requires consideration of what type of data you have and what you would like to show. The flow chart on the next page provides suggestions for which chart to use on this basis.

Chart Suggestions—A Thought-Starter



© 2006 A. Abela — a.v.abela@gmail.com

CFE Research Phoenix Yard Upper Brown Street Leicester LE1 5TE.
Tel 0116 229 3300. Email info@cfe.org.uk. www.cfe.org.uk

Use of data viz in the national evaluation of Fulfilling Lives (Multiple Needs)

An annual report is published for the national evaluation of *Fulfilling Lives (Multiple Needs)* in the Spring. The first annual report (published June 2015) largely covered the inception phase of the evaluation, describing how projects had been set up as relatively little project data had been collated at that point. Subsequent annual reports will contain analysis and data viz using information collected through the Common Data Framework (CDF).

In addition to the annual reports, Big Lottery requires the national evaluation team to produce a quarterly update of key statistics from the common data framework (CDF). To meet this requirement the team at CFE Research have developed an information dashboard using a software package called Tableau.

Multiple Needs Dashboard: Design and Purpose

The primary purpose of the Multiple Needs Dashboard is to provide an update on key data for the Big Lottery Fund and for those working to deliver the 12 funded programmes. As such, the main audience for the dashboard have knowledge of the *Fulfilling Lives (multiple Needs)* programme, which means that it is not necessary to provide lots of explanatory information. Projects provide data for the national evaluation on a quarterly basis in line with the CDF which all 12 projects are signed up to. Data is provided by the third Friday of the month following the quarter-end. For example, for the quarter ending Wednesday 30th September 2015, the deadline for data returns was Friday 16th October 2015.

CFE Research runs an initial sense-check of the data returns, for example to make sure fields are completed. Data is then shared with our partner organisation, 'Get The Data', who clean the data and merge it into a single composite analysis file. This process requires numerous data-checks, including things such as deleting any data sent erroneously for participants who have not given informed consent to share it and converting all 12 project returns into the same date format. This process takes around a week. Once completed, the data is held in SPSS by CFE Research for analytical purposes. Some of the data from the CDF is then transferred into Tableau for the Multiple Needs Dashboards. The whole process currently takes six weeks after the end of the quarter. At this point the dashboards are created in Tableau. Currently, dashboards are shared with projects either as PDFs or, ideally, as a Tableau workbook where projects have installed Tableau Reader. In the latter case this allows projects to explore the data themselves, for example viewing dashboards for other funded projects, re-ordering data and excluding irrelevant or unwanted fields.

The idea behind the Multiple Needs Dashboards is, rather than sending lengthy Word or paper documents, to allow Big Lottery and funded projects access to key parts of the data via an interactive dashboard. Whilst projects will have a thorough understanding of their own data already, sharing it in an information dashboard has two important advantages:

1. It allows projects to see their data in the broader context of that which is also being collected by the other funded projects.
2. It allows the projects to see what data the national evaluation team at CFE Research hold for their project which ensures any discrepancy, omissions or data processing errors are identified and drives up data quality.

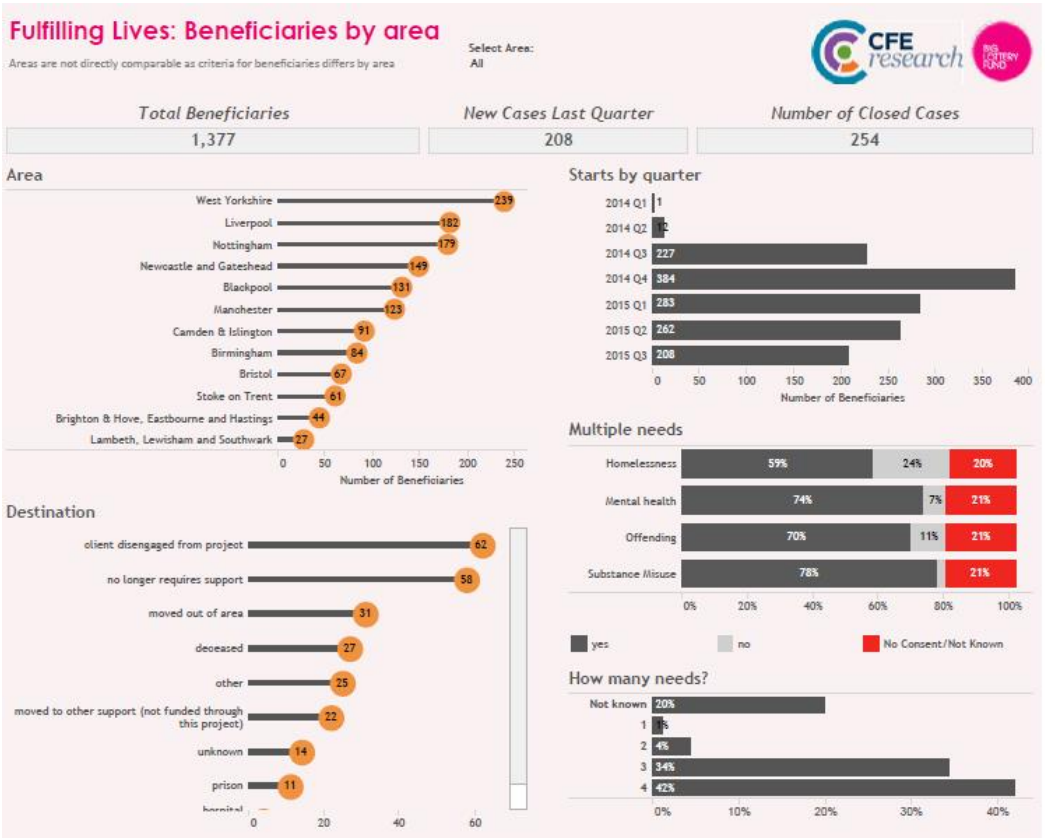
There are also some fundamental principles behind the design and use of Multiple Needs Dashboard which should be borne in mind:

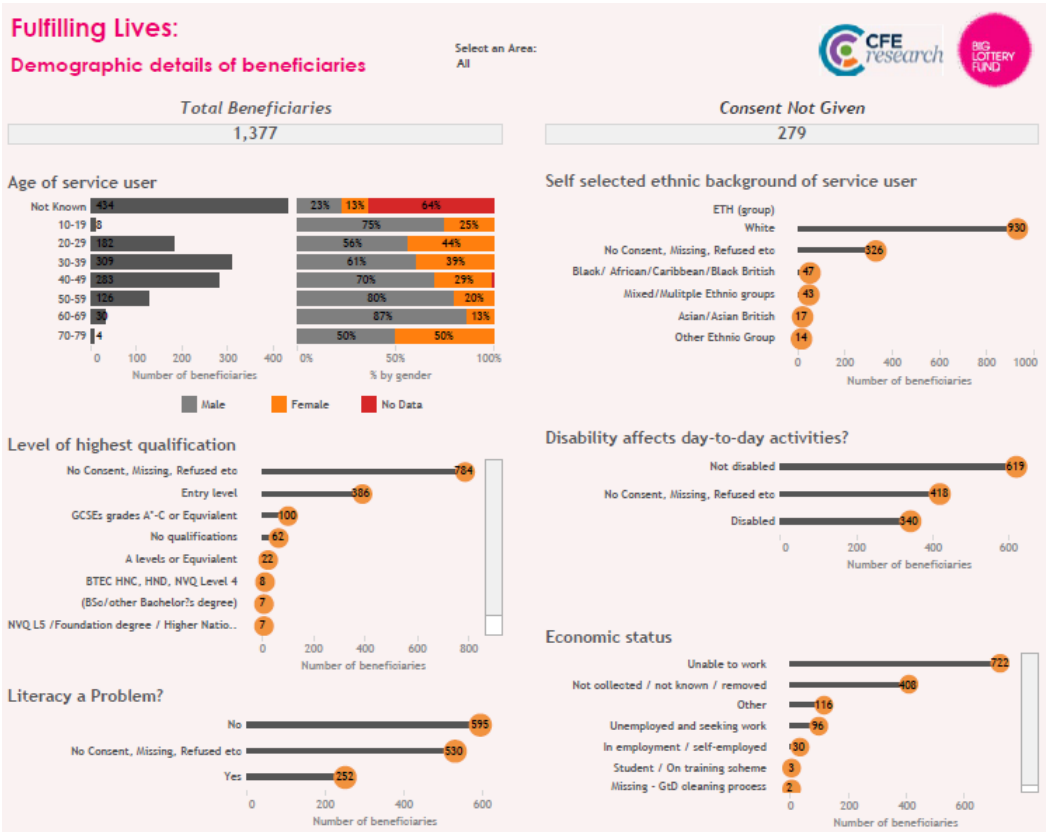
— *It provides largely descriptive, explanatory information about the 12 funded projects. More complex analysis, such as the relative cost effectiveness of the projects, will be conducted for the annual reports.*

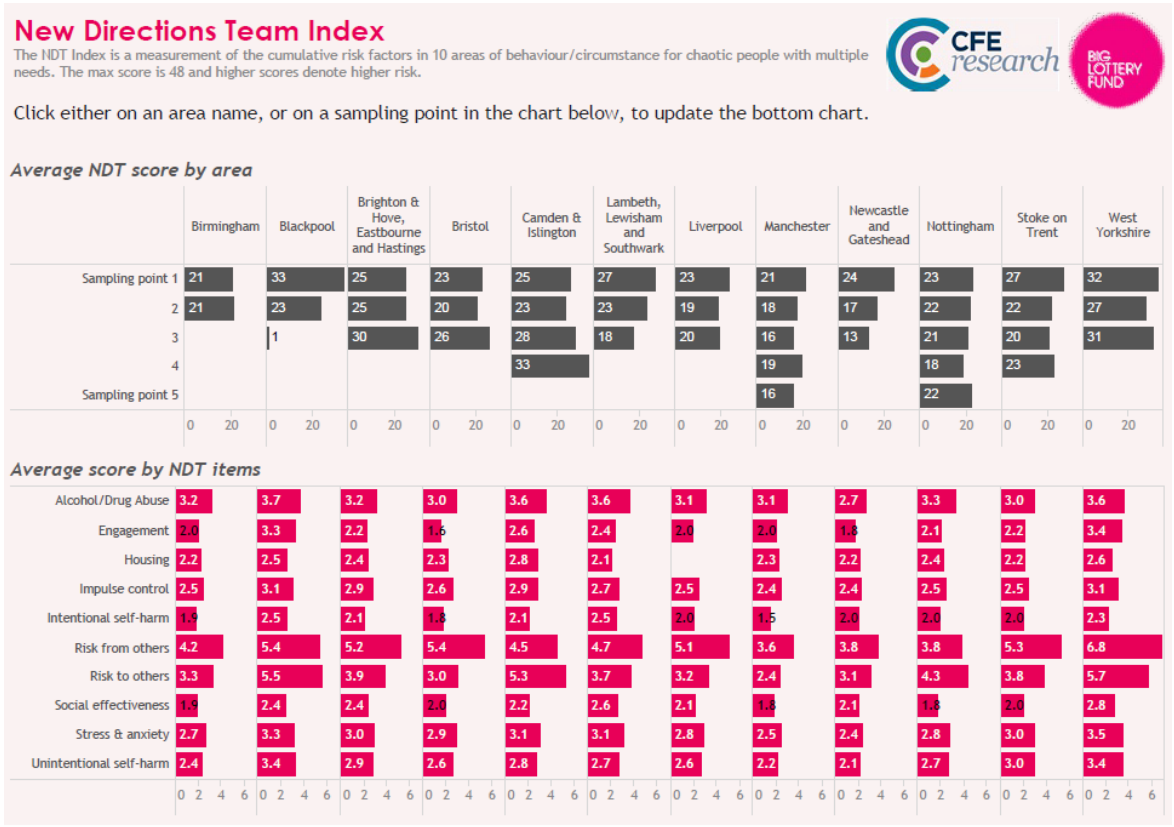
— *Is not intended to answer questions about project delivery, rather it should enable users to ask better questions about the projects, informed by the data.*

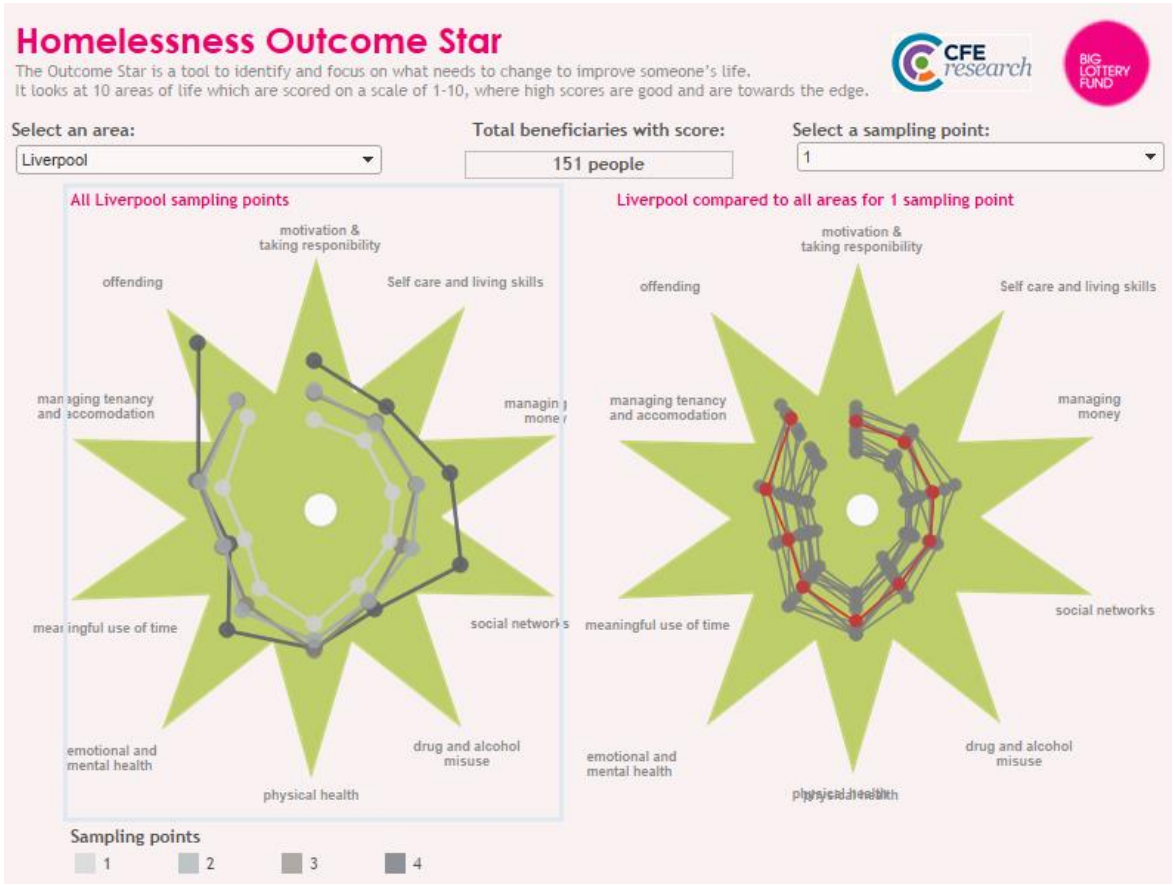
— *It does not replace the grant management processes used by the Fund*

Screen shots of the four Multiple Needs Dashboards currently (November 2015) shared with projects and the Fund are provided on the following two pages.


Fulfilling Lives: Beneficiaries by area
Areas are not directly comparable as criteria for beneficiaries differs by area
Select Area:
All
CFE research
BIG LOTTERY FUND
Total Beneficiaries
1,377
New Cases Last Quarter
208
Number of Closed Cases
254
Area
West Yorkshire 239
Liverpool 182
Nottingham 179
Newcastle and Gateshead 149
Blackpool 131
Manchester 123
Camden & Islington 91
Birmingham 84
Bristol 67
Stoke on Trent 61
Brighton & Hove, Eastbourne and Hastings 44
Lambeth, Lewisham and Southwark 27
0 50 100 150 200 250
Number of Beneficiaries
Starts by quarter
2014 Q1 1
2014 Q2 12
2014 Q3 227
2014 Q4 384
2015 Q1 283
2015 Q2 262
2015 Q3 208
0 50 100 150 200 250 300 350 400
Number of Beneficiaries
Multiple needs
Homelessness 59% 24% 20%
Mental health 74% 7% 21%
Offending 70% 11% 21%
Substance Misuse 78% 21%
0% 20% 40% 60% 80% 100%
yes
no
No Consent/Not Known
How many needs?
Not known 20%
1 1%
2 4%
3 34%
4 42%
0% 10% 20% 30% 40%
Destination
olient disengaged from project 62
no longer requires support 58
moved out of area 31
deceased 27
other 25
moved to other support (not funded through this project) 22
unknown 14
prison 11
0 20 40 60


Fulfilling Lives:
Demographic details of beneficiaries
Select an Area:
All
CFE research
BIG LOTTERY FUND
Total Beneficiaries
1,377
Consent Not Given
279
Age of service user
Not Known 434 23% 13% 64%
10-19 8 75% 25%
20-29 182 56% 44%
30-39 309 61% 39%
40-49 283 70% 29%
50-59 126 80% 20%
60-69 39 87% 13%
70-79 4 50% 50%
0 100 200 300 400
Number of beneficiaries
0% 50% 100%
% by gender
Male
Female
No Data
Self selected ethnic background of service user
ETH (group)
White 930
No Consent, Missing, Refused etc 326
Black/ African/Caribbean/Black British 47
Mixed/Mulitple Ethnic groups 43
Asian/Asian British 17
Other Ethnic Group 14
0 200 400 600 800 1000
Number of beneficiaries
Level of highest qualification
No Consent, Missing, Refused etc 784
Entry level 386
GCSEs grades A*-C or Equvialent 100
No qualifications 62
A levels or Equvialent 22
BTEC HNC, HND, NVQ Level 4 8
(BSc/other Bachelor?s degree) 7
NVQ L5 /Foundation degree / Higher Natio.. 7
0 200 400 600 800
Number of beneficiaries
Disability affects day-to-day activities?
Not disabled 619
No Consent, Missing, Refused etc 418
Disabled 340
0 200 400 600
Number of beneficiaries
Literacy a Problem?
No 595
No Consent, Missing, Refused etc 530
Yes 252
0 200 400 600
Number of beneficiaries
Economic status
Unable to work 722
Not collected / not known / removed 408
Other 116
Unemployed and seeking work 96
In employment / self-employed 30
Student / On training scheme 3
Missing - GtD cleaning process 2
0 200 400 600
Number of beneficiaries


New Directions Team Index
The NDT Index is a measurement of the cumulative risk factors in 10 areas of behaviour/circumstance for chaotic people with multiple needs. The max score is 48 and higher scores denote higher risk.
CFE research
BIG LOTTERY FUND
Click either on an area name, or on a sampling point in the chart below, to update the bottom chart.
Average NDT score by area
Birmingham
Blackpool
Brighton & Hove, Eastbourne and Hastings
Bristol
Camden & Islington
Lambeth, Lewisham and Southwark
Liverpool
Manchester
Newcastle and Gateshead
Nottingham
Stoke on Trent
West Yorkshire
Sampling point 1 21 33 25 23 25 27 23 21 24 23 27 32
2 21 23 25 20 23 23 19 18 17 22 22 27
3 1 30 26 28 18 20 16 13 21 20 31
4 33 19 18 23
Sampling point 5 16 22
0 20
Average score by NDT items
Alcohol/Drug Abuse 3.2 3.7 3.2 3.0 3.6 3.6 3.1 3.1 2.7 3.3 3.0 3.6
Engagement 2.0 3.3 2.2 1.6 2.6 2.4 2.0 2.0 1.8 2.1 2.2 3.4
Housing 2.2 2.5 2.4 2.3 2.8 2.1 2.3 2.2 2.4 2.2 2.6
Impulse control 2.5 3.1 2.9 2.6 2.9 2.7 2.5 2.4 2.4 2.5 2.5 3.1
Intentional self-harm 1.9 2.5 2.1 1.8 2.1 2.5 2.0 1.5 2.0 2.0 2.0 2.3
Risk from others 4.2 5.4 5.2 5.4 4.5 4.7 5.1 3.6 3.8 3.8 5.3 6.8
Risk to others 3.3 5.5 3.9 3.0 5.3 3.7 3.2 2.4 3.1 4.3 3.8 5.7
Social effectiveness 1.9 2.4 2.4 2.0 2.2 2.6 2.1 1.8 2.1 1.8 2.0 2.8
Stress & anxiety 2.7 3.3 3.0 2.9 3.1 3.1 2.8 2.5 2.4 2.8 3.0 3.5
Unintentional self-harm 2.4 3.4 2.9 2.6 2.8 2.7 2.6 2.2 2.1 2.7 3.0 3.4
0 2 4 6


Homelessness Outcome Star
The Outcome Star is a tool to identify and focus on what needs to change to improve someone's life.
It looks at 10 areas of life which are scored on a scale of 1-10, where high scores are good and are towards the edge.
CFE research
BIG LOTTERY FUND
Select an area:
Liverpool
Total beneficiaries with score:
151 people
Select a sampling point:
1
All Liverpool sampling points
Liverpool compared to all areas for 1 sampling point
motivation & taking responibility
Self care and living skills
managing money
social networks
drug and alcohol misuse
physical health
emotional and mental health
meaningful use of time
managing tenancy and accomodation
offending
Sampling points
1
2
3
4

Multiple Needs Dashboard: Future Developments

The Multiple Needs Dashboard is not designed to remain in the same format throughout the national evaluation, which runs until 2021. Rather, the dashboard is intended to change as more data is collected and shared and to reflect changes in the focus of what is most useful for the Fund and projects to focus on. The first iterations of the Multiple Needs Dashboard are driven by two things:

1. **What data is available?** To date, not all fields of the CDF are completed. It takes longer to collect service user information and data at different sampling points, for example.
2. **Proof of concept.** Is an information dashboard the best way of providing a regular update to projects/the Fund? The first version(s) of the dashboard should provide headline descriptive information on beneficiaries – how many people have engaged with each project so far? What is the demographic profile of those beneficiaries? – to test whether the approach works. More detailed information could usefully follow at a later point as the audience becomes familiar with the data and how it is presented.

The intention is to ensure that there is some consistency across quarterly updates of the Multiple Needs Dashboard – i.e. so that it looks familiar/similar to users from one update to the next – while incorporating improvements to the existing dashboards and adding new ones as more data becomes available and the focus of attention evolves. Some improvements currently being explored include:

— *Projects have requested that we look at adding their targets for engaging with beneficiaries into the dashboards and creating a new dashboard summarising service use data.*

— *Providing additional dashboards using the Homelessness Outcome Star data to show progression. For example, how many beneficiaries move from 'stuck' (a score of 1 or 2 on the star) to accepting help (3 or 4)?*

— *Following feedback from the Fund's England Committee, we are investigating changing the NDT Assessment dashboard so that average scores are calculated only for those beneficiaries for whom data is available at all sampling points and potentially supressing averages calculated below a minimum number of beneficiaries (e.g. at least 10).*

— *The Multiple Needs Dashboard is currently shared with the Fund and projects as a workbook to be viewed using Tableau Reader. This is a slightly cumbersome approach and we are exploring how to share the dashboards online instead. Ideally, the dashboards could be accessed via a web browser, providing easy access from any device to the latest version of the dashboards.*

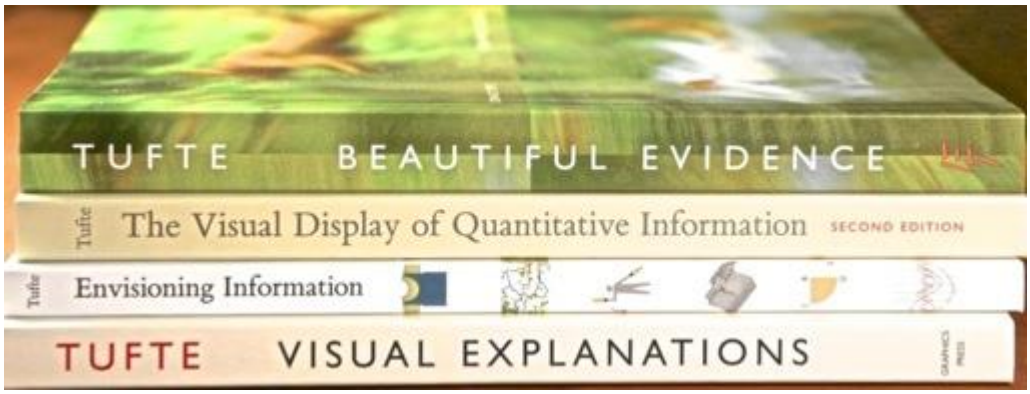
Useful Resources

This guide serves as a very brief introduction to data viz and how it is being used for the national evaluation of Fulfilling Lives (Multiple Needs). There are a plethora of useful resources in the burgeoning discipline of data viz. Some of the more useful sources of information are listed below.

People to follow

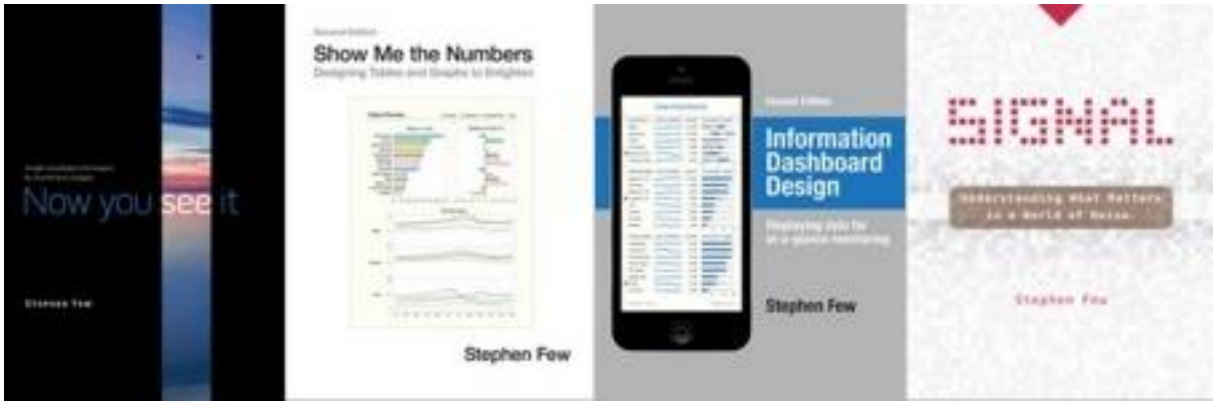
The guru and 'Leonardo da Vinci' of data viz is **Edward Tufte**. He has produced several beautiful books on the topic – the snappily titled 'The Visual Display of Quantitative Information' is a good place to start –and has a useful Q&A section on his website.

http://www.edwardtufte.com/ @EdwardTufte



Stephen Few is another well-published and sometimes controversial commentator on all things data viz. He was produced several instructive books and articles addressing a range of data viz issues. His books are well presented, easy to read and useful to apply. 'Information Dashboard Design' is an excellent guide to designing effective dashboards.

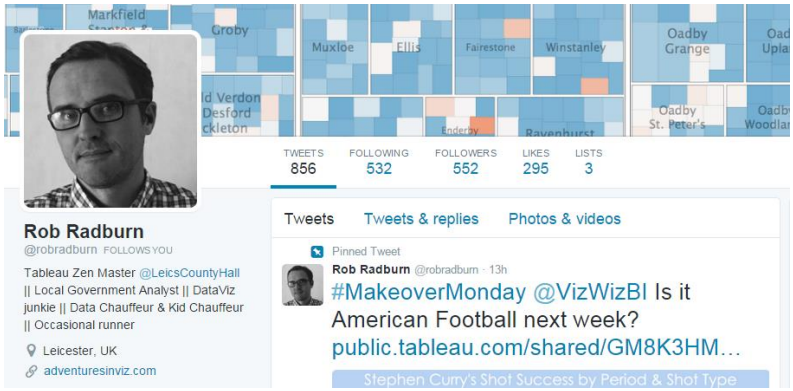
http://www.perceptualedge.com/



Stephanie Evergreen has a great website and produces regular bulletins with easy-to-follow step-by-step guides for creating good charts and graphics in Excel. You can subscribe to her newsletter via her website and follow her Twitter account for regular updates and links to useful free stuff.

http://stephanieevergreen.com/ @evergreendata

A vibrant community of data viz experts using Tableau software has expanded over the last few years, including here in the UK. Exponents of the software share ideas and examples online and meet at regional Tableau User Groups. Those who excel in their use of Tableau for data viz can become a **Tableau 'Zen Master'**. There are currently 21 Zen Masters all of whom are worth following and one of which– **Rob Radburn** – works with CFE Research on the national evaluation of Fulfilling Lives (Multiple Needs).



Software

Most people use **Excel** to create charts for their reports/presentations. Default Excel charts can be reformatted to vastly improve them; this is a good place to start for producing better data viz. Simply removing unnecessary shading and formatting and simplifying the way data is presented makes a big improvement.

The new big-player in the field of data viz software is **Tableau** with Forbes reporting the company growing from 4,400 customers producing $34m revenue in 2010 to 26,000 customers ($413m revenue) by 2014. So its use is becoming much more widespread. This is the software that we are using to create the Multiple Needs Dashboard as part of our national evaluation of Fulfilling Lives (Multiple Needs). Tableau Public is a free version of the software and you can share any data viz you create through Tableau's website but you are not able to protect the data you use (i.e. anyone can download it). Tableau Desktop is the paid-for version of the software (c. £1,250 per licence) which allows you to choose how to share the data viz you create and protect your data from being downloaded. Tableau Reader is free software which works a bit like Adobe Reader in that it allows you to view something (data viz) created by someone else if they share the Tableau workbook with you. In 2015 Tableau started to offer two-year licences of Tableau Desktop for free to non-profit organisations, including those based in the UK. Conditions apply, more info is available here: http://www.tableaufoundation.org/initiatives/tableau-non-profits

Many other online tools to assist with data visualisation, some of which require only a basic understanding of presenting data and some which require a more advanced understanding of coding. **The Data Visualisation Catalogue,** developed by Severino Ribecca, is a brilliant resource which provides a useful guide to which charts to use and what is available to help you create them: http://www.datavizcatalogue.com/

Contact

Jon Adamson, Associate Director, CFE Research
Jon.Adamson@cfe.org.uk 0116 229 3300 / 07793 802 142
www.cfe.org.uk @CFE_Ltd

November, 2015