

Data Journalism Visualisation

Group 5 Ari Rauhala, Maja Savanovic, Christine Pichler

Institute for Information Systems and Computer Media (IICM),
Graz University of Technology
A-8010 Graz, Austria

706.057 Information Visualisation SS 2013
Graz University of Technology

20 May 2013

Abstract

Data journalism visualisation is a important part of modern journalism. Visualisations of data have become the backbone of relaying information throughout all media. One might say that visualisations take the viewer on a journey through the previously so complex and incomprehensible data. Visualizing journalism is an user-friendly way to present raw data. Mute data evolves into storytellers. Narrations can make data more interesting and easier to follow. Several types of presenting data can be identified within narrative visualisation. These visualisations may let the viewer sit back and passively study the data or encourage the viewer to explore what lies beneath the shell of various charts and graphs. This survey aims to cover the field of data journalism visualisations focusing especially on narrative visualisation, its forms and its features. The tools used to visualize data by journalists are also addressed.

Contents

- 1 Introduction** **1**

- 2 Data Journalism Visualisation** **4**
 - 2.1 Passive form 7
 - 2.2 Explorative form 7

- 3 Narrative Visualisation** **13**
 - 3.1 Different visualisations 13
 - 3.1.1 Magazine Style 13
 - 3.1.2 Annotated Chart 13
 - 3.1.3 Partitioned Poster 13
 - 3.1.4 Flow Chart 16
 - 3.1.5 Comic Strip 17
 - 3.1.6 Slideshow 17
 - 3.1.7 Film, Video, Animation 17
 - 3.2 Data Sources 17
 - 3.3 Social Components in Narrative Visualisation 17

- 4 Tools** **21**

- 5 Concluding Remarks** **25**

Chapter 1

Introduction

Visualisation and visualizing has been a common tool and a way of passing information since the dawn of man. First visualisations were painted on the cave walls, explaining a certain moment or an event that had some meaning in the life of our ancestors. Since then visualizing different aspects of life has become even more commonly used method. Nowadays visualisation of data is crucial in information transmission - there is too much information and data everywhere and mere numbers are hard to understand. Thus by visualizing the data, it is easier to see what the data implies. There is a great rivalry in the world of media and as media delivers information and data to the common man, the visualisation has a big role.

Chapter 2

Data Journalism Visualisation

Terms like "information visualisation" ["Readings in Information Visualization: Using Vision to Think"], "data visualisation" ["Toward a Perceptual Science of Multi-dimensional Data Visualization: Bertin and Beyond"], and "scientific visualisation" ["Visualization: Expanding Scientific and Engineering Research Opportunities"] are all used to describe the same: The visualisation of data in a user-friendly and easily understandable way. In the case of journalism, the data might be a financial, business or technical report, as well as political or sports news. Nevertheless, all data is some kind of a story. Data journalism visualisation means the preparation of those stories in a visual form, such that the viewer can easily explore and understand the presented information. Everybody knows visualisations in the form of 2-dimensional-graphs and diagrams. But nowadays tools and other techniques provide an easy transformation of data into interactive, dynamic and 3-dimensional representation, which naturally make "the story" more interesting for the viewer. Newspapers, like the Guardian, the New York Times, the BBC, the Washington Post and many more use visual integration in their reports. Now we will present few examples of visualisations in the following figures [2.1](#) and [2.2](#).



RAF fighters attacked the Dorniers at 4,000 metres (13,000ft) over Deal in Kent before they had reached their intended target. The RAF battle report claimed at least six Dorniers destroyed and one damaged for the loss of three RAF aircraft and two air gunners killed.

Next »

Figure 2.1: Shows a glimpse of a step by step story by "The Guardian" that tells a interactive story about a second world war aircraft which is raised from the sea by images, animations and small text blocks. [Image extracted from [*German bomber to be raised from the sea - interactive*].]

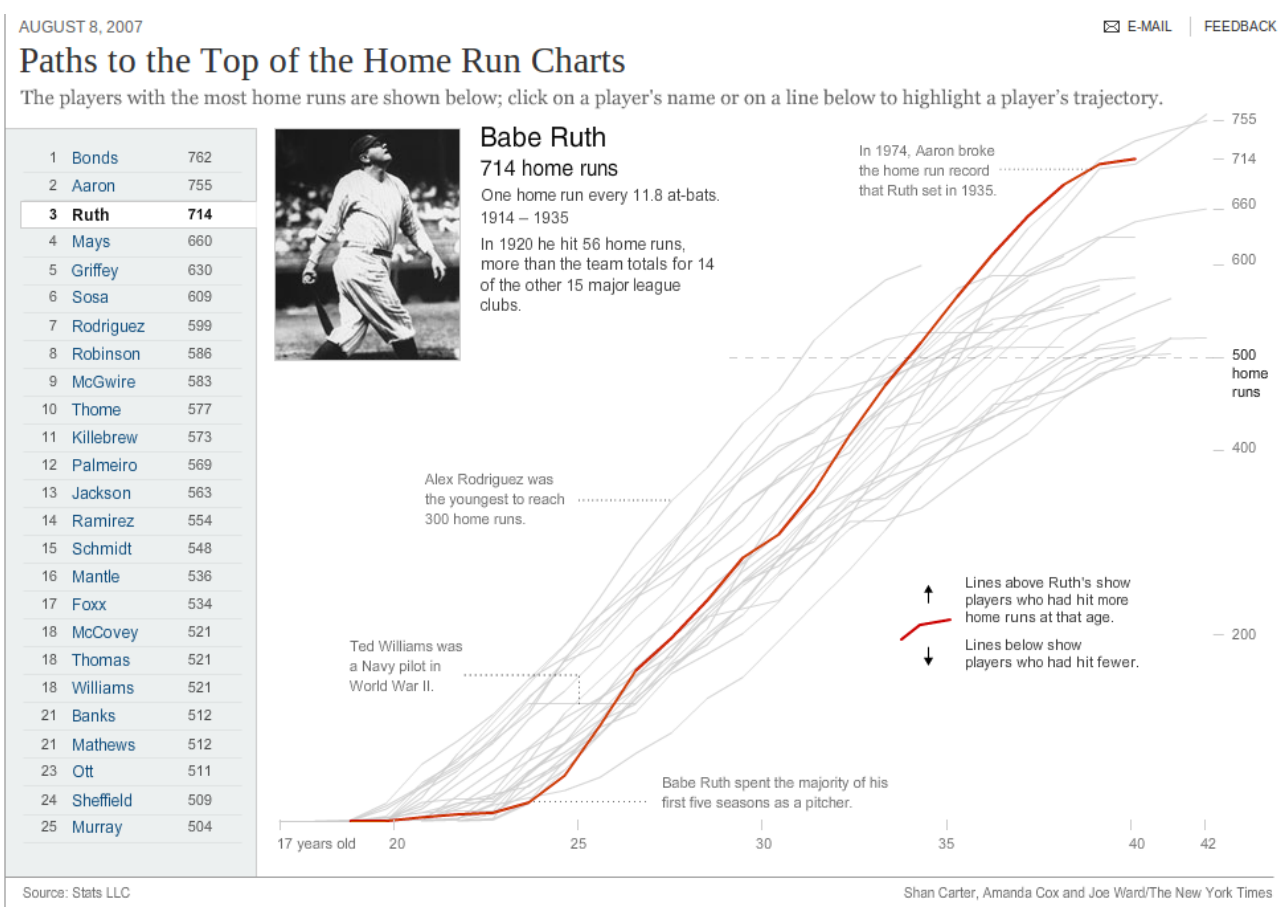


Figure 2.2: The "New York Times" shows the players with the most home runs. Detailed information is shown by clicking either on a players name at the left or a line in the chart. [Image extracted from [Paths to the Top of the Home Run Charts].]

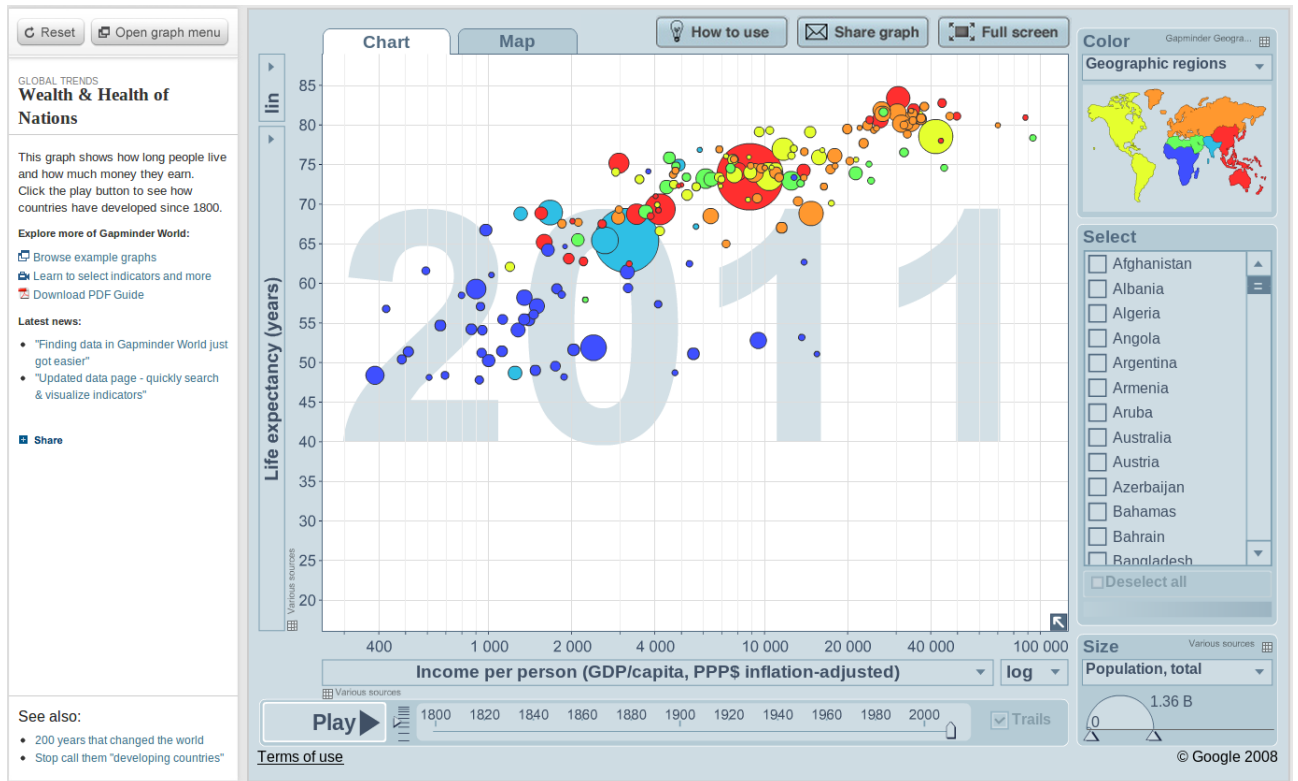


Figure 2.3: Gapminder "Wealth and Health of Nations" [Image extracted from [Wealth & Health of Nations*].]

2.1 Passive form

Visualisations can be divided into two different types depending on the activity of the viewer. Both are described in this chapter. A passive visualisation can be easily expressed as a guided tour through some data. The data is prepared as a narrative visualization and presented to the viewer. The viewer can interrupt this presentation, view the current data more carefully and then continue with the remaining presentation. The important thing is, that the user does not interact with the data himself. There are many ways that such a guided tour can look like. Following an example from gapminder.org is presented (see figure 2.3). This organisation has provided more than 500 different data sets as narrative visualisation in passive and explorative form. The visualisations have been done with Trendalyzer, which is intuitive and powerful at once. Trendalyzer is now available as a free Google feature. We picked a data set about the "Wealth & Health of Nations". The circles represent the countries, their size is an indicator for the total population. The axis represent the life expectancy and the income per person. At the right side, there is a small legend, for the different colors. If the user clicks on the play button, the circles move and grow depending on the dataset. The animation is like a small video.

Other example is the "Riot Rumours" from "The Guardian". This particular visualisation was one of "the winners of the 2012 edition of the Data Journalism Awards" Guardian Interactive team and Voss [*How riot rumours spread on Twitter*]. We picked one of the several examples. See figure 2.4.

2.2 Explorative form

Visualisation can also take an explorative form. Explorative form normally uses the same techniques and same media as narrative visualisation, but with this form the viewer is not only a passive bystander, the viewer can explore the data and focus on the parts that are most interesting. One example of explorative visualisation is the "Top Secret America" project by the Washington Post [*Growth of domestic counter terrorism organizations*]. The project shows how the field of counterterrorism in American governmental agencies have changed since

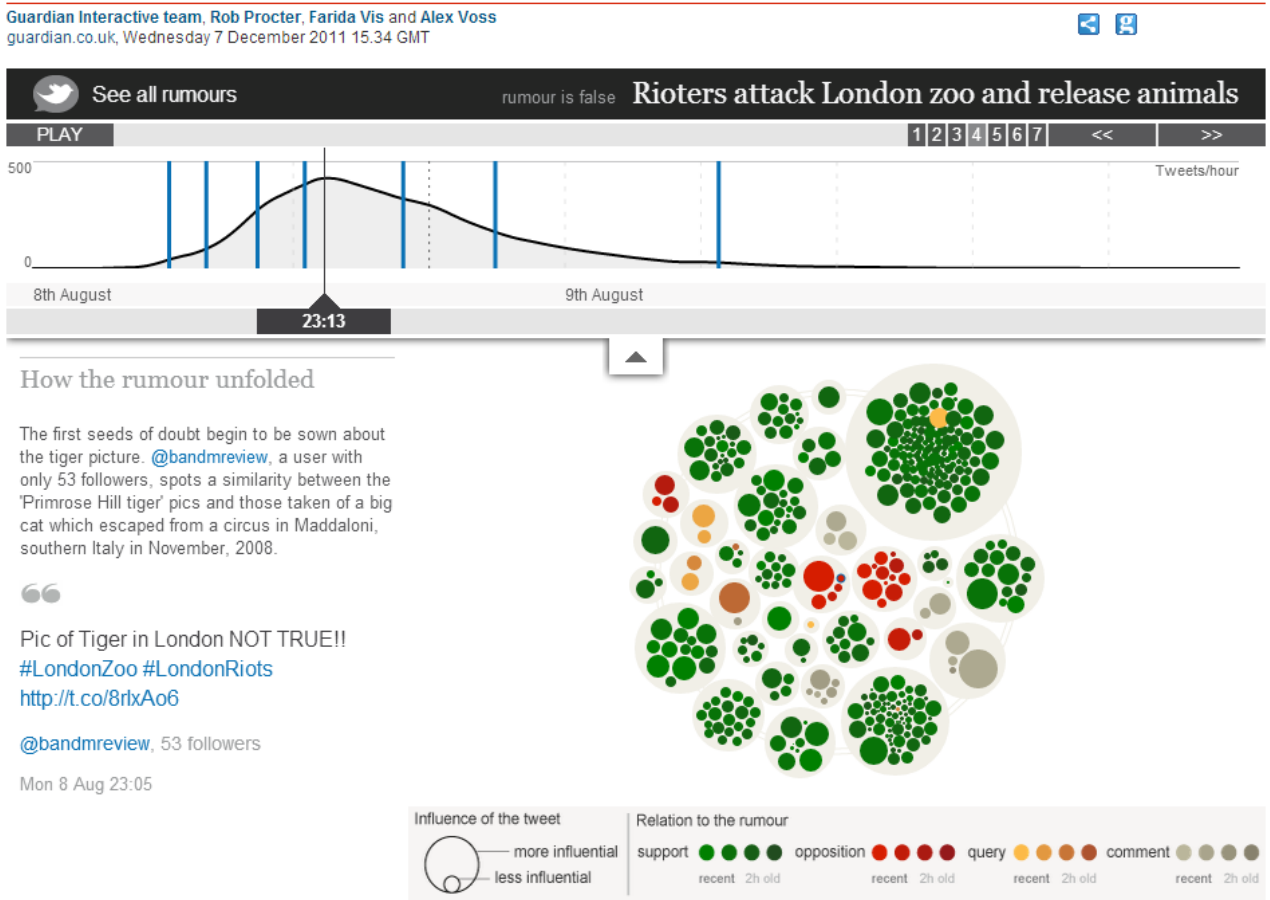


Figure 2.4: "Rioters attack London zoo and release animals" [Image extracted from Guardian Interactive team and Voss [*How riot rumours spread on Twitter*].]

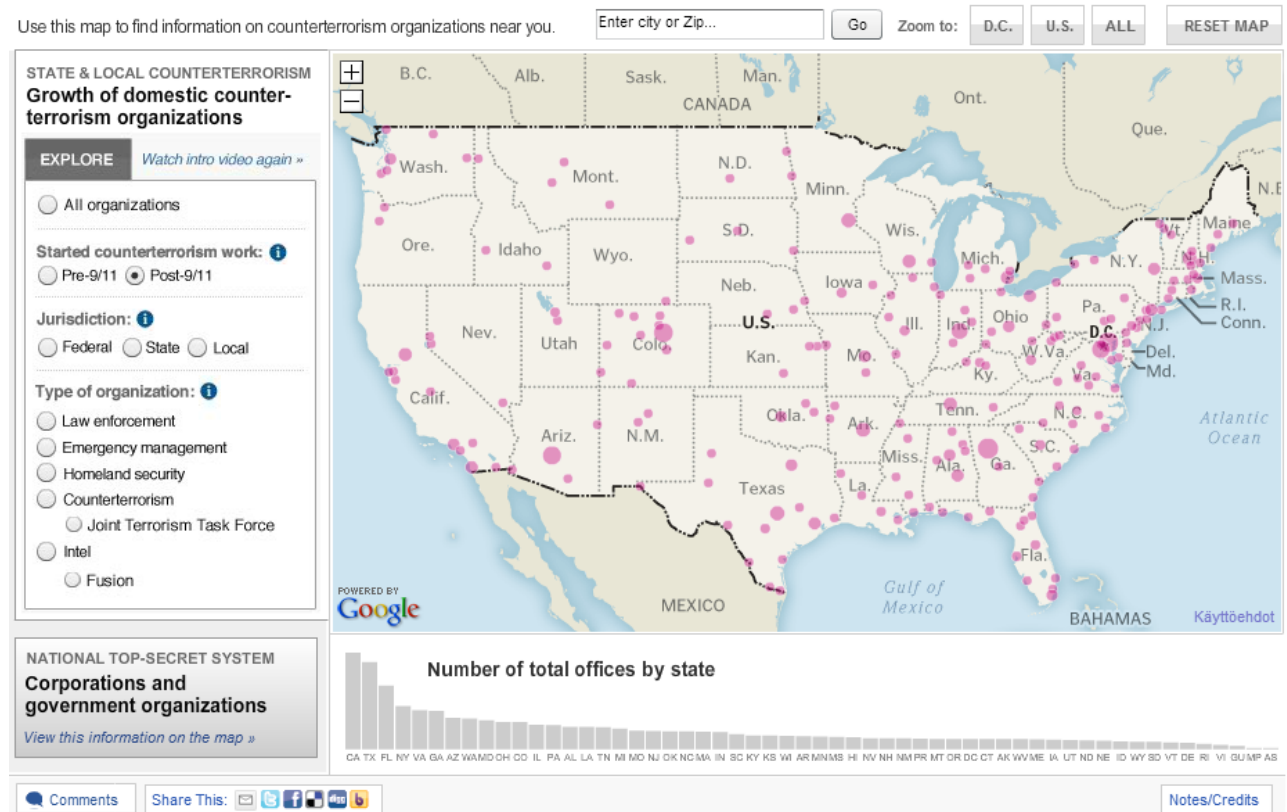


Figure 2.5: [Project "Top Secret America" as map. Image extracted from [*Growth of domestic counter terrorism organizations*].]

9/11. The project includes many types of visualization, not just explorative, but here the viewer can explore on a map, figure 2.5, where organizations have been set up. It is possible to select different criteria to find out for example which organizations existed before 9/11 and which have been founded only after 9/11.

There is also a interactive semicircle, figure 2.6 which shows government organizations and their relationships with their contractors and with other organizations. The information here is color coded showing which type of work the organization does and the visualization also shows which organizations are most active and which are the least active.

To show the explorative possibilities on the Gapminder visualisations we took the same data set like before. But in difference, we have chosen the map view. The circles for the countries are now located on their geographical position (figure 2.7a). The user has different ways to interact explorative with the data. The name of the country is shown on mouseover. By clicking on the countries or selecting them from the combo box on the right, a more detailed view is generated. See figure 2.7b.

In the following, some other examples are listed. See figures 2.7a, 2.7b and 2.7.

The Top Secret Network of Government and its Contractors

Explore the relationships between government organizations and the types of work being done in Top Secret America. [Need help?](#)

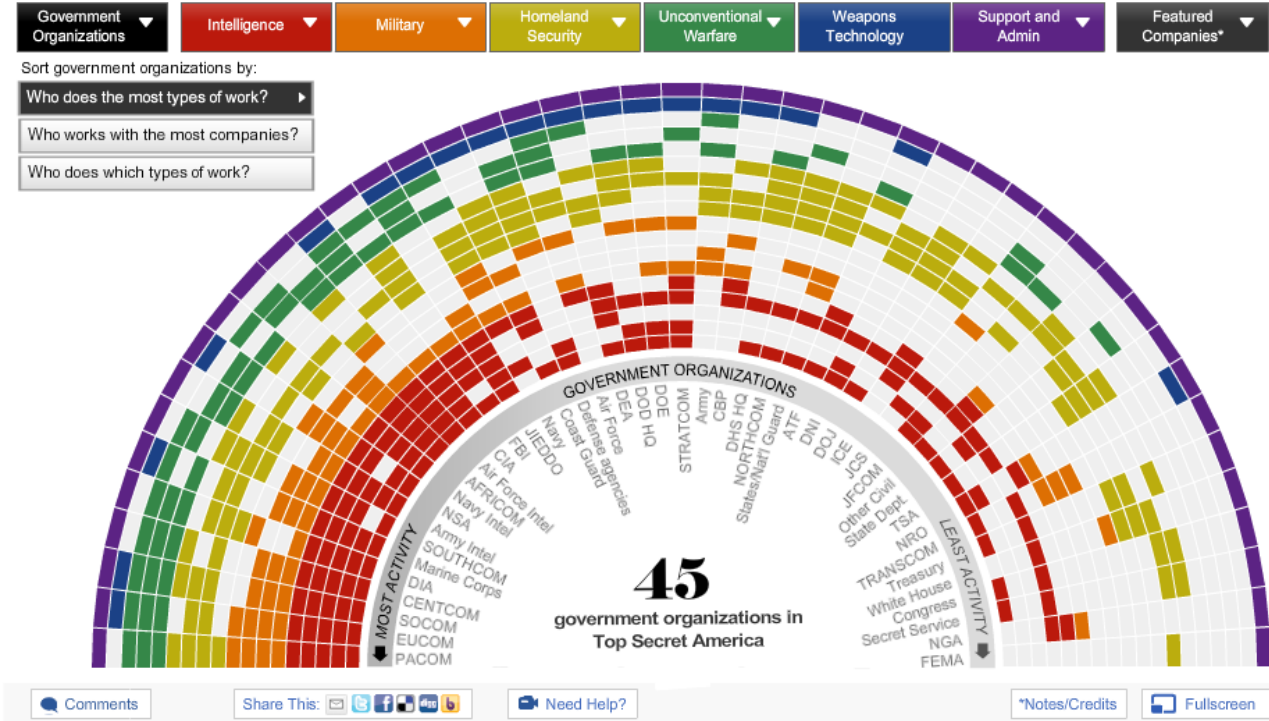
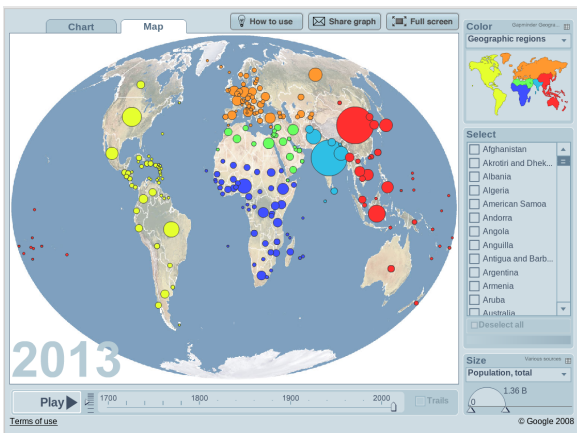
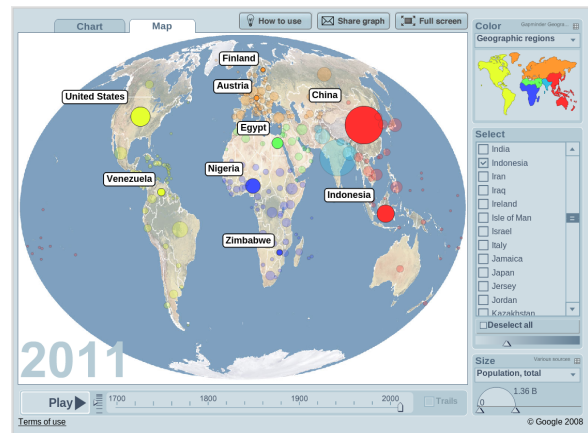


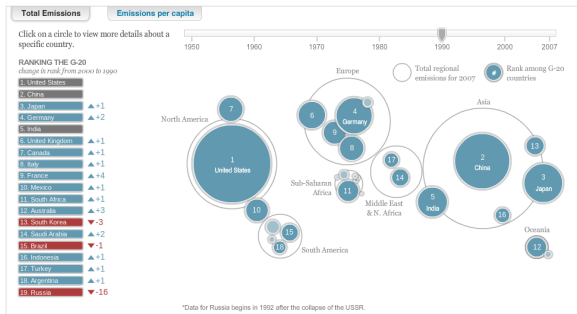
Figure 2.6: [Project "Top Secret America" as semicircle. Image extracted from [*Growth of domestic counter terrorism organizations*].]



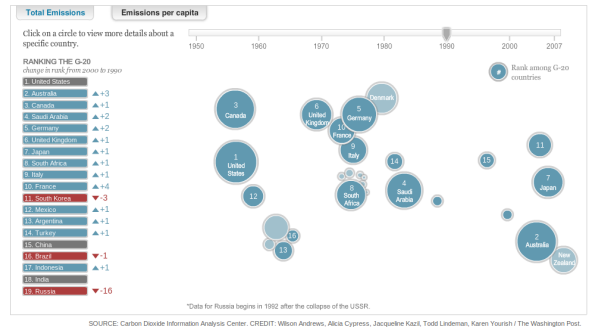
(a) Explorative map



(b) Same map with selected areas. Images extracted from [*Wealth & Health of Nations*]*



(a) Washington post: total global emissions



(b) Washington post: global emissions per capita

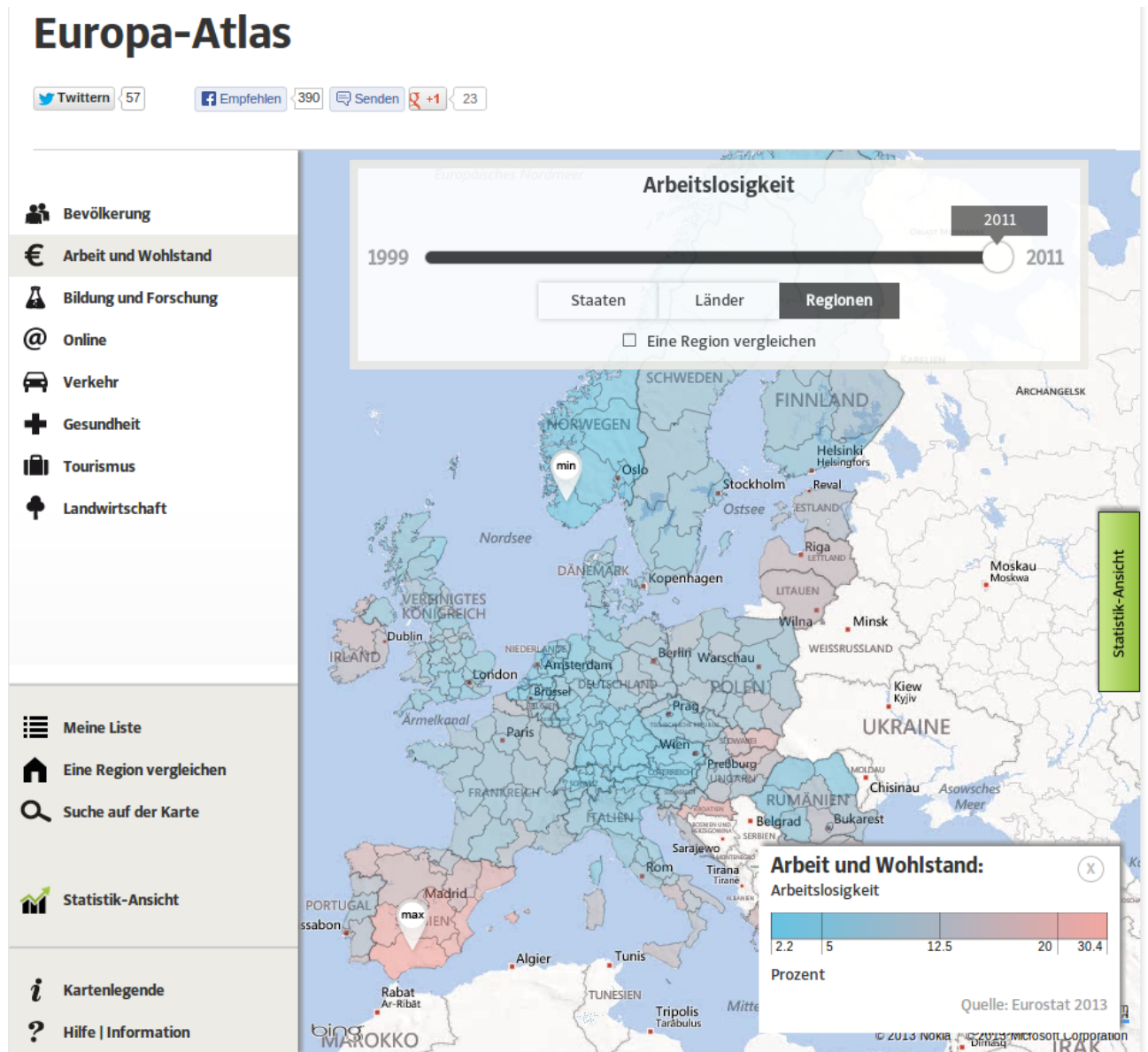


Figure 2.7: [German project about unemployment. Image extracted from Salcher [Euripa-Atlas].]

Chapter 3

Narrative Visualisation

Narrative visualisation is a type of visualisation that has a certain structure. Narrative is defined in the Oxford English Dictionary as "an account of a series of events, facts, etc., given in order and with the establishing of connections between them." Narrative is like a train trip from A to B via C, D and E. It has a starting point and an ending and these two points are connected together and there's also something in between.

3.1 Different visualisations

In fact narrative visualisation is telling the stories with data. Stories have always fascinated people living all around the world. It is easier to follow the idea of something, if it is presented in the form of a story: getting only bits and pieces here and there can be confusing. In most stories the order of events has a meaning. Probably the most common meaning of order is chronology. "Artists, designers, and psychologists have all explored ways in which visual media can be organized to engender a narrative experience. They have developed nuanced techniques for sequentially directing a viewer's attention and keeping viewers oriented across transitions." According to **Segel and Heer** 7 main genres can be identified within narrative visualisation: magazine style, annotated chart, partitioned poster, flow chart, comic strip, slideshow and film/video/animation.

3.1.1 Magazine Style

A good example of magazine style visualisation can be seen on figure 3.1. It shows a visualisation by Susken Rosenthal.

3.1.2 Annotated Chart

Annotated chart is basically a chart that has annotations explaining the data the chart portrays. Annotated charts are very popular and an effective way to display data and show differences between subjects. There are several different chart types, but the most common charts are the histogram, the bar chart, the pie chart and the line chart. Here we have an example of an annotated chart 3.2. It is a cartogram, a map of the world, showing which countries are currently in recession, which have just come out of recession, which have dodged recession and which countries are currently in runaway growth. These different states are depicted by color coding and the annotation is situated on the edge of the chart. The chart doesn't have data from all countries, but of the countries it has, it is clearly visible that most of those countries in the world have come out of recession and only some countries in Europe are currently in recession. It also shows that the only countries with growth are in Asia.

3.1.3 Partitioned Poster

A partitioned poster is a picture or a chart that is divided into partitions. Every partition has a resemblance of its own. Our example is from the New York Times 3.3 and it shows faces, names and hometowns of the American

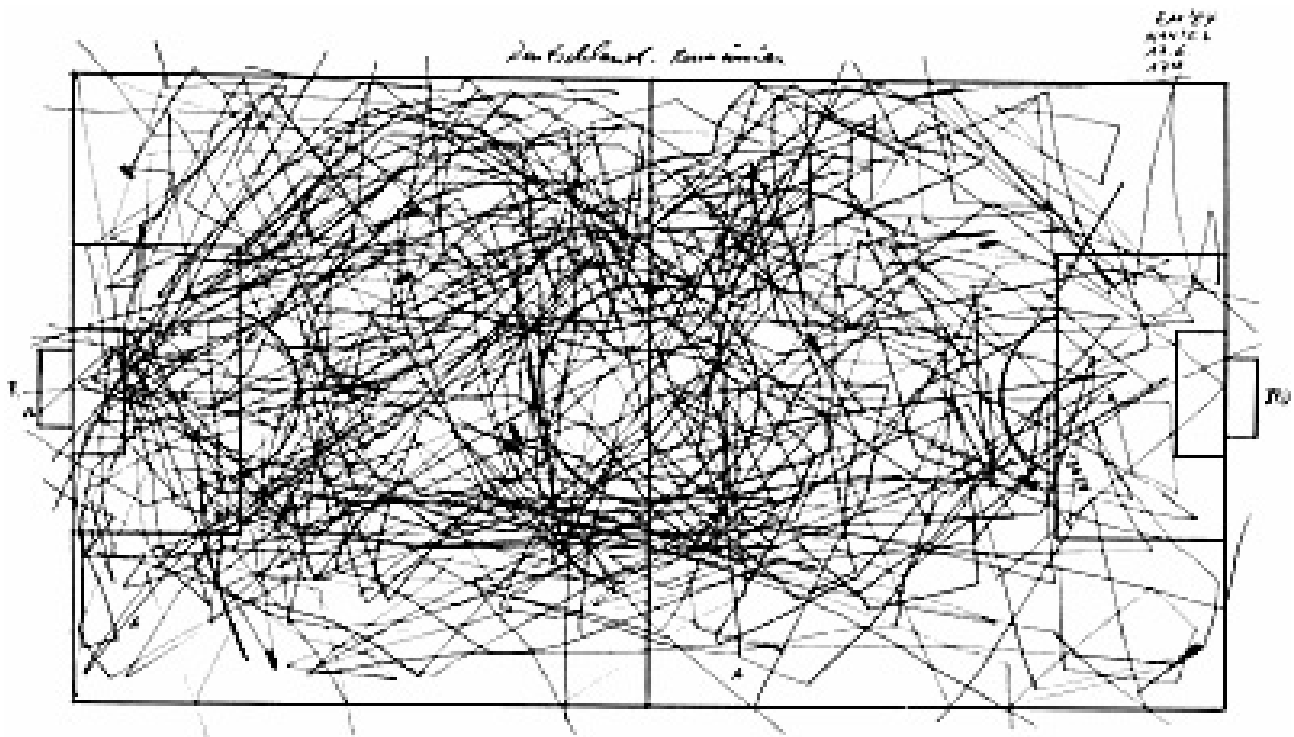


Figure 3.1: Magazine style: Drawing of a football game [Image extracted from [*Football Drawings*].]

The world economy turns the corner

Countries that are in or out of recession, based on the most recent quarterly data

Christine Oliver
 guardian.co.uk, Tuesday 26 January 2010 09.30 GMT

Roll over the countries

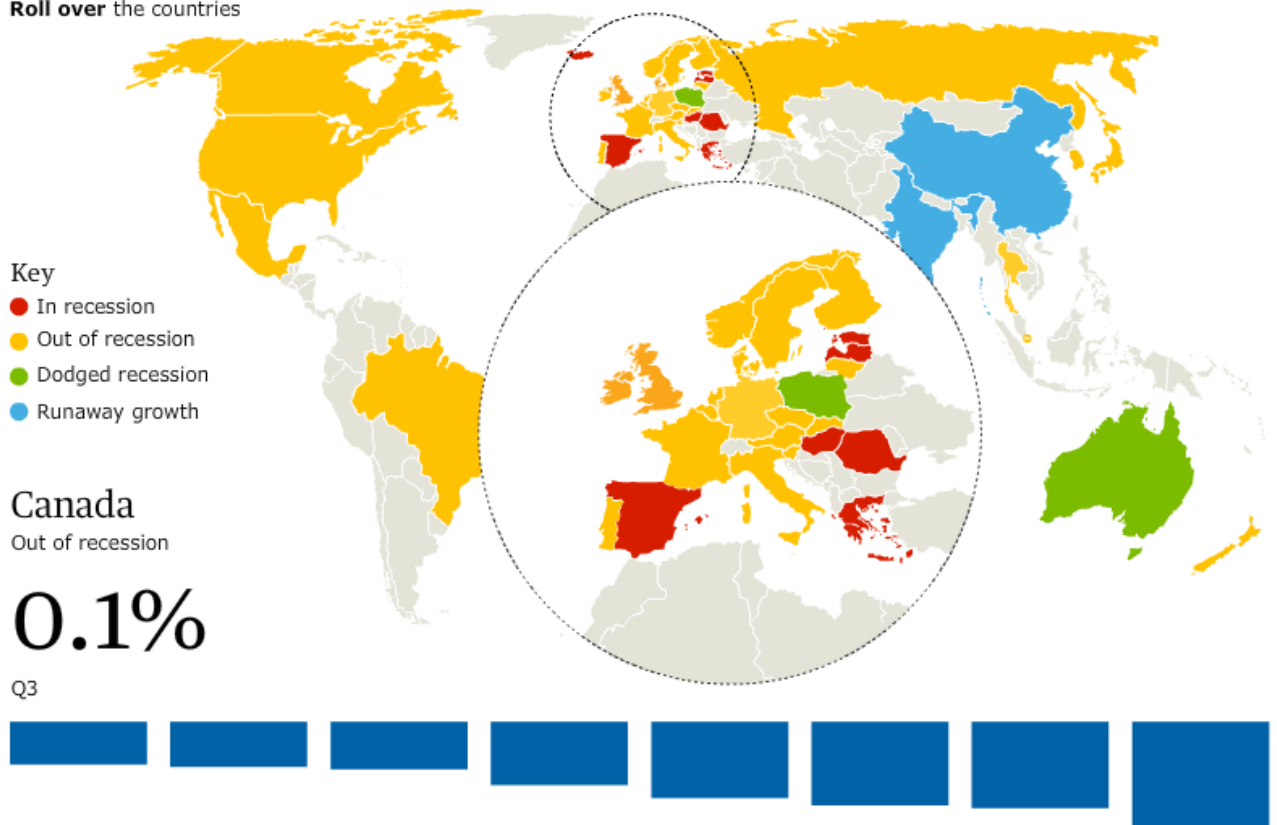


Figure 3.2: Annotated chart: Countries and recession [Image extracted from *The world economy turns the corner*.]



Figure 3.3: Partitioned poster: Faces of the American casualties [Image extracted from *[Faces of the Dead]*.]

casualties of war in Afghanistan and Iraq. This interactive article contains two almost identical partitioned posters, each on its own page. They both contain the same partitions (small squares) in the same order, but they look different. All these squares are arranged in a specific order by the date of death, the most recent ones are situated in the upper left corner. We have combined both pages into one picture. On the left hand side ("Photos" page) the partitions form a picture of the soldier whose partition is selected. On the right side ("Chart" page) all partitions are color coded depending of the country in which the soldier of the square died in. While the small squares on the left side doesn't really tell anything about the whole material, the chart page on the right side illustrates effectively the amount of casualties in both countries and it is easy to tell in which country most of the deaths occurred.

3.1.4 Flow Chart

Flow chart is diagram that represents the flow of a process in a visual way. Flow charts have a starting and ending point (or points). Flow charts makes it possible to have decision points or options in the visualisation. An example from the New York Times 3.4 illustrates what is the chance of getting a free upgrade on a flight. The starting point can be seen on the upper left corner and from there the flow chart takes the reader through multiple options ending in two different results: either the passenger gets an upgrade or ends up packed like sardines. At first glance a flow chart like this can look exhausting, but it is actually rather easy to follow it

through. This flow chart is a static visualisation, there is no guided tour through it. It could have animation in it and the visualisation could automatically focus on the next step depending on the previous answer. We think this flow chart works well even as it is and the apparent complexity just brings out the amusing prerequisites and the humour in getting an upgrade.

3.1.5 Comic Strip

Comic strip is a tangible way to narrate visualisation. Like regular comics, there is a story that goes forwards with every image. In our example 3.5, "Life Cycle of a Beetle through a Year" by Edward Tufte, the comic strip visualizes the life cycle of a beetle in one year. The year is divided into 12 panels, each representing a month. The comic strip in this visualisation is a continuous set of images, a nice way of demonstrating how the beetle ecloses, lives and eats on the ground and lays eggs for the next generations.

3.1.6 Slideshow

Slideshow is a common tool for presenting data and information in a sequence. Figure fig:historyoftoyota is an example from the Financial Times. This interactive slideshow tells the story of a company called Toyota. The story is divided into 13 parts, into 13 slides. Each slide represents a certain year or era in the history of Toyota and they tell something about what happened at that time. Each slide is reachable at any point and slide can be chosen by clicking the circle under a year.

3.1.7 Film, Video, Animation

When talking about visualizing a story, film, video and animation are an everyday media for that. Film, video and animation is also a common instrument of modern journalism. Our example is an animation explaining how drone strikes, see figure 3.7. This animation is a good example of passive visualisation, it is only necessary to sit back and watch as the drone strikes are animated in the browser. After the animation has finished, this is also an excellent example of explorative visualisation as it's possible to find more information about the strikes.

3.2 Data Sources

When working with data visualisations, it might be useful to not only have the visualisation to interact with them, but also to have the data on which the visualisation bases. For this reason download as spreadsheet or at least pdf should be provided. Nowadays at most of the visualisations the data can not be accessed so easy, but at some visualisations there is a download link provided.

3.3 Social Components in Narrative Visualisation

Social components within visualisations aim to enrich the social interaction between different viewers and also to animate them to share the visualisation. Such components can be social network integrations, eg Twitter, Google +, Facebook or to send an e-mail. The viewer can recommend the visualisation and engage others to have a look. An example of this can be seen on Figure 3.8, which shows a social component bar on an article on New York Times webpage.

Another and maybe more important part is to give the viewer the possibility to leave a comment. Due to Segel and Heer ["Narrative Visualization: Telling Stories with Data"] "A list of comments associated with the current view enables journalists and readers to share observations and discuss trends."

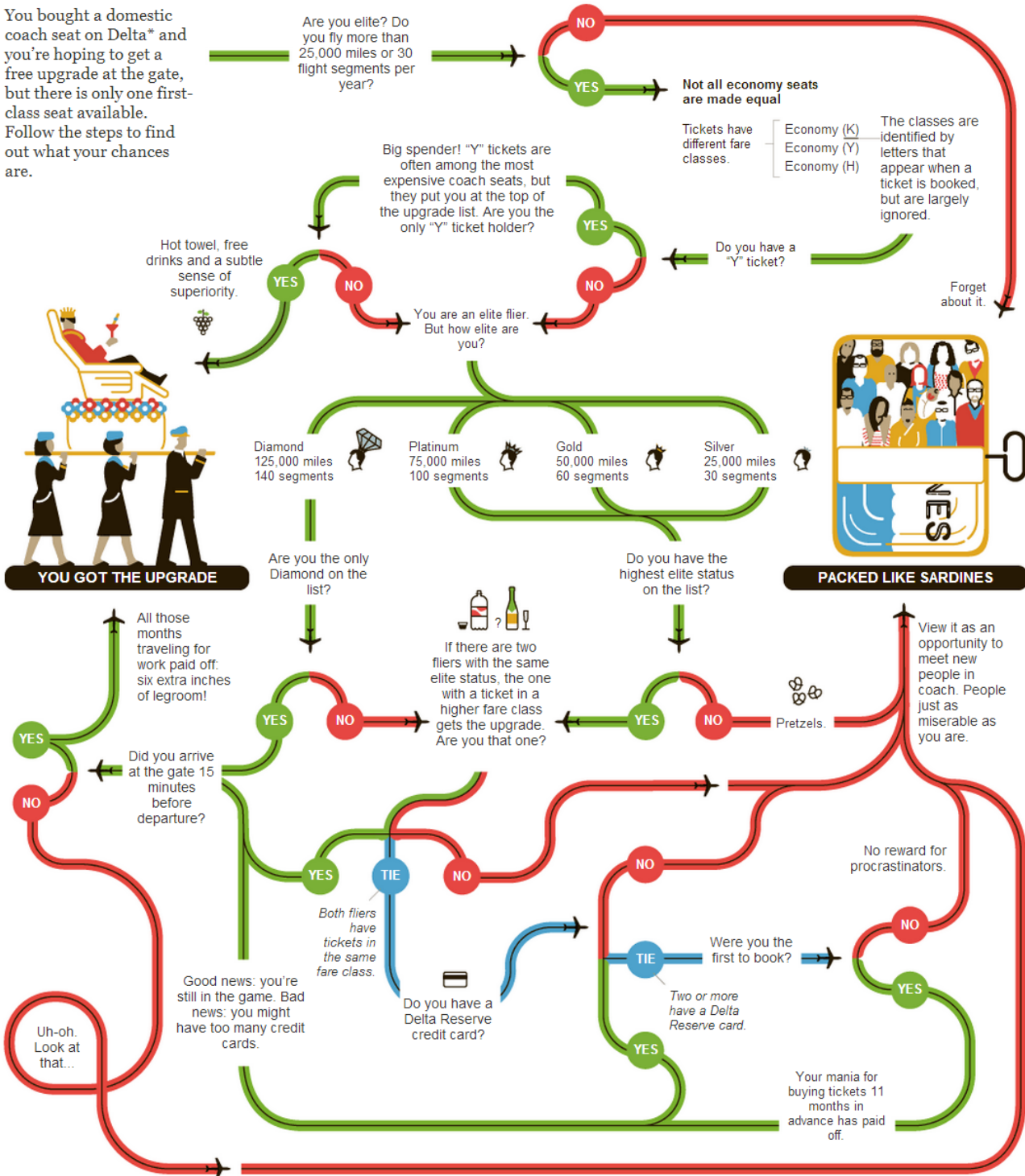
It is kind of personalization, that makes people more emotional and increase engagement.

Published: June 9, 2012

RECOMMEND TWITTER LINKEDIN SIGN IN TO E-MAIL SHARE

Pampered Like Royalty, or Canned Like Sardines?

You bought a domestic coach seat on Delta* and you're hoping to get a free upgrade at the gate, but there is only one first-class seat available. Follow the steps to find out what your chances are.



* Different airlines have different upgrade criteria. The graphic is based on Delta's.

Figure 3.4: Flow chart: chances of getting an upgrade [Image extracted from [The Chances of Getting a Free Upgrade on a Flight].]



Figure 3.5: Comic strip: lifecycle of a beetle [Image extracted from [*The Life Cycle of a Japanese Beetle*].]

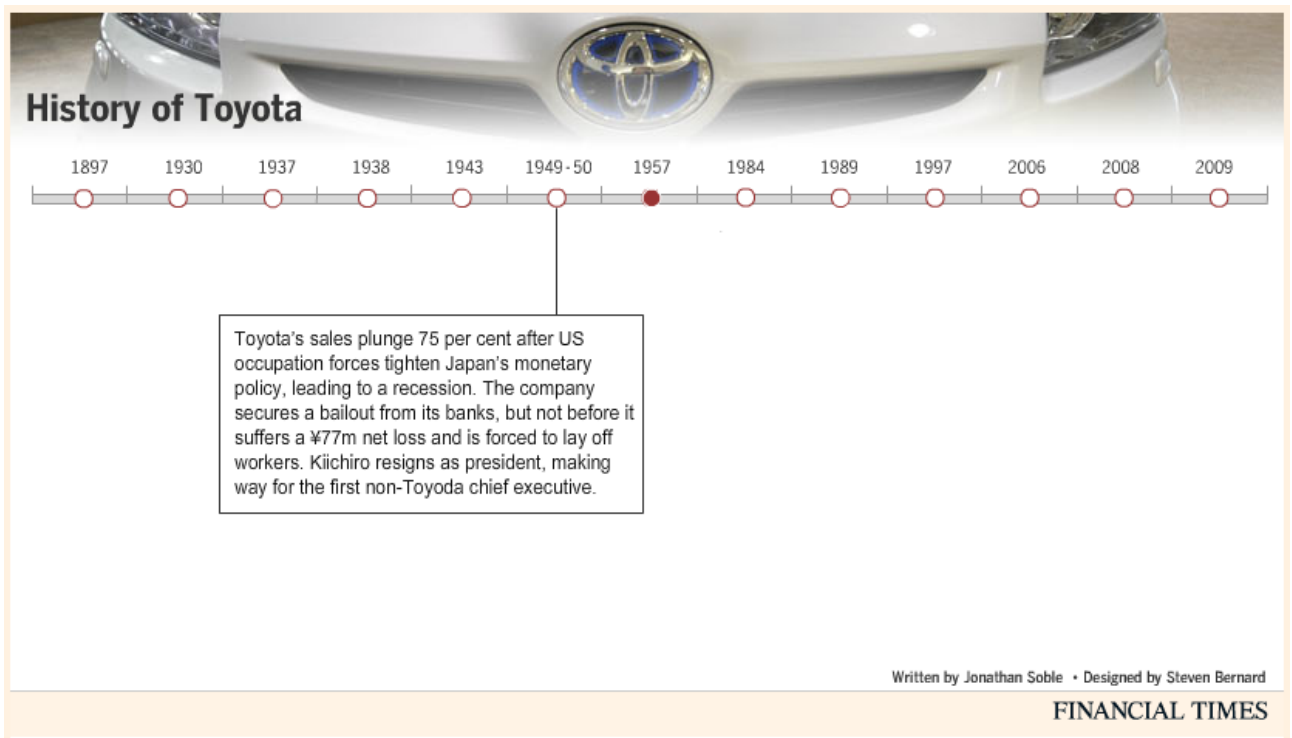


Figure 3.6: Slideshow: history of toyota [Image extracted from [*Toyota timeline: a company history*].]

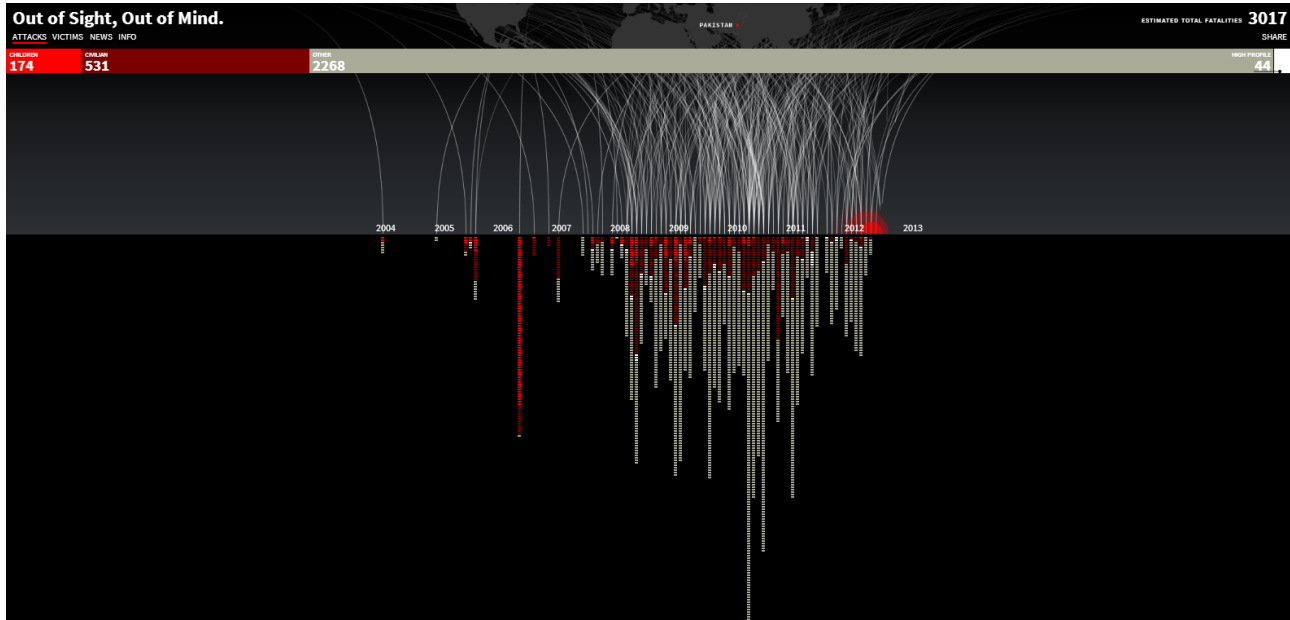


Figure 3.7: Video & Animation: drone strike [Image extracted from [*Out of Sight, Out of Mind. A visualization of all documented drone strikes in Pakistan since 2004.*].]



Figure 3.8: Example of social components [Image extracted from *The social component bar on New York Times article webpage.*]

Chapter 4

Tools

Rogers [*Data visualization DIY: Our Top Tools*] from "The Guardian" described which visualization Tools that are most frequently used by the guardian data store.

- Adobe Illustrator: it is just used if there is enough time. Day to day work is not done with it. For example: [4.1](#).
- Google Fusion: is used to create quickly and detailed maps. The user can afterwards zoom in and open lots of data.
- Tableau and Tableau public: this tool is used to bring different charts together
- Google spread sheet: from a simple line, bar or pie chart up to a powerful animations like e.g. Hans Rosling's Gapminder (which had been presented in the previous chapter) can be created
- Doctor Who Villains [4.3](#) and others

Batchgeo is also sometimes used, and the italian newspaper "Il Fatto Quotidiano" made a huge project with this tool. Also worth to mention is, that Argentinas newspaper "La Nacion" makes all their visualisations with Tableau.

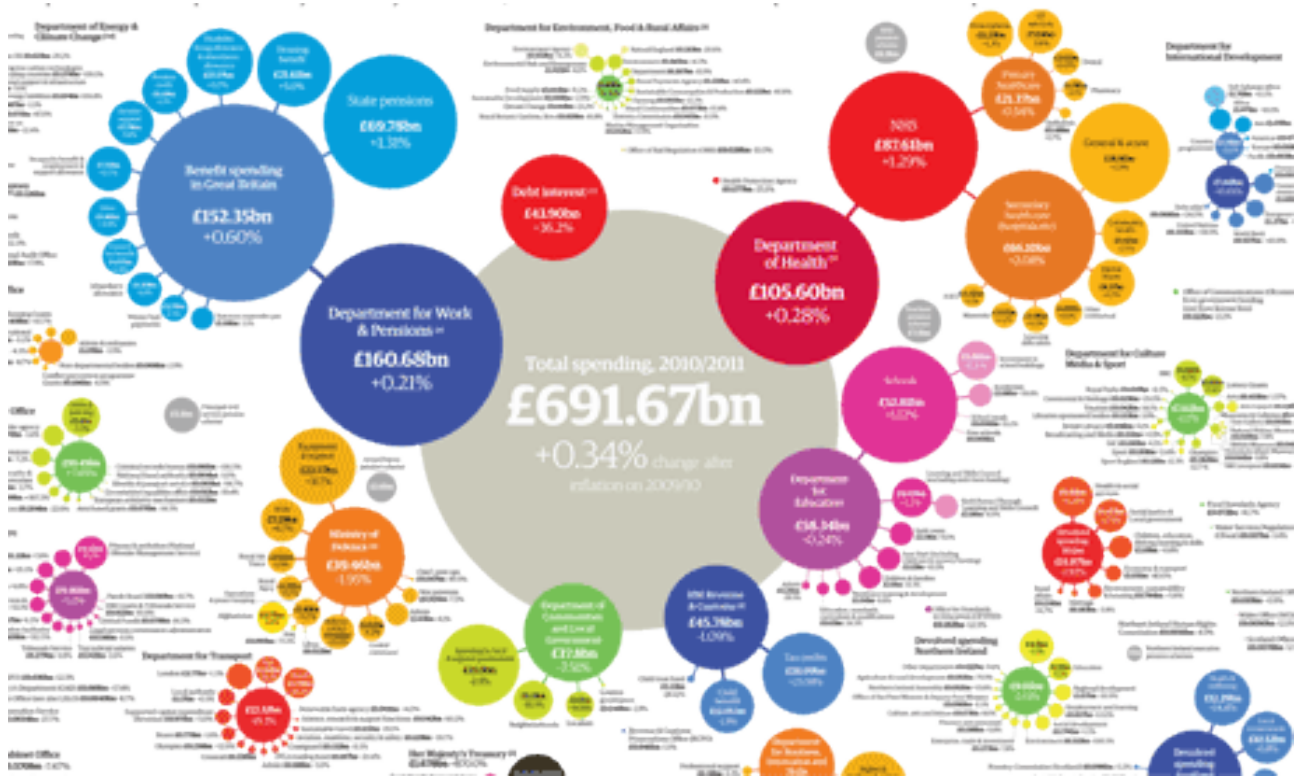


Figure 4.1: [Image extracted from [Government spending by department, 2010-11: get the data].]

Visualizations : Doctor Who villains - Oct 2011 update

Uploaded by: smrtogers
Description:

Created at: Oct 12 2011

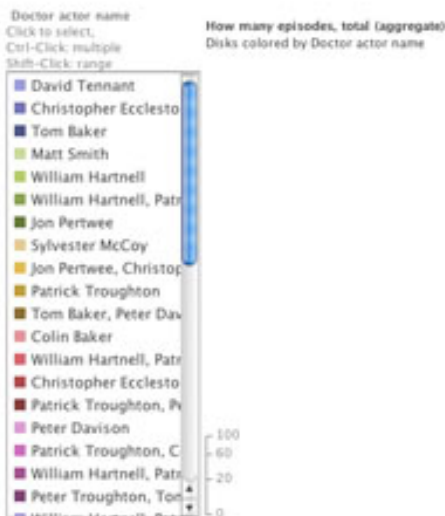


Figure 4.2: [Image extracted from [Data visualization DIY: Our Top Tools].]

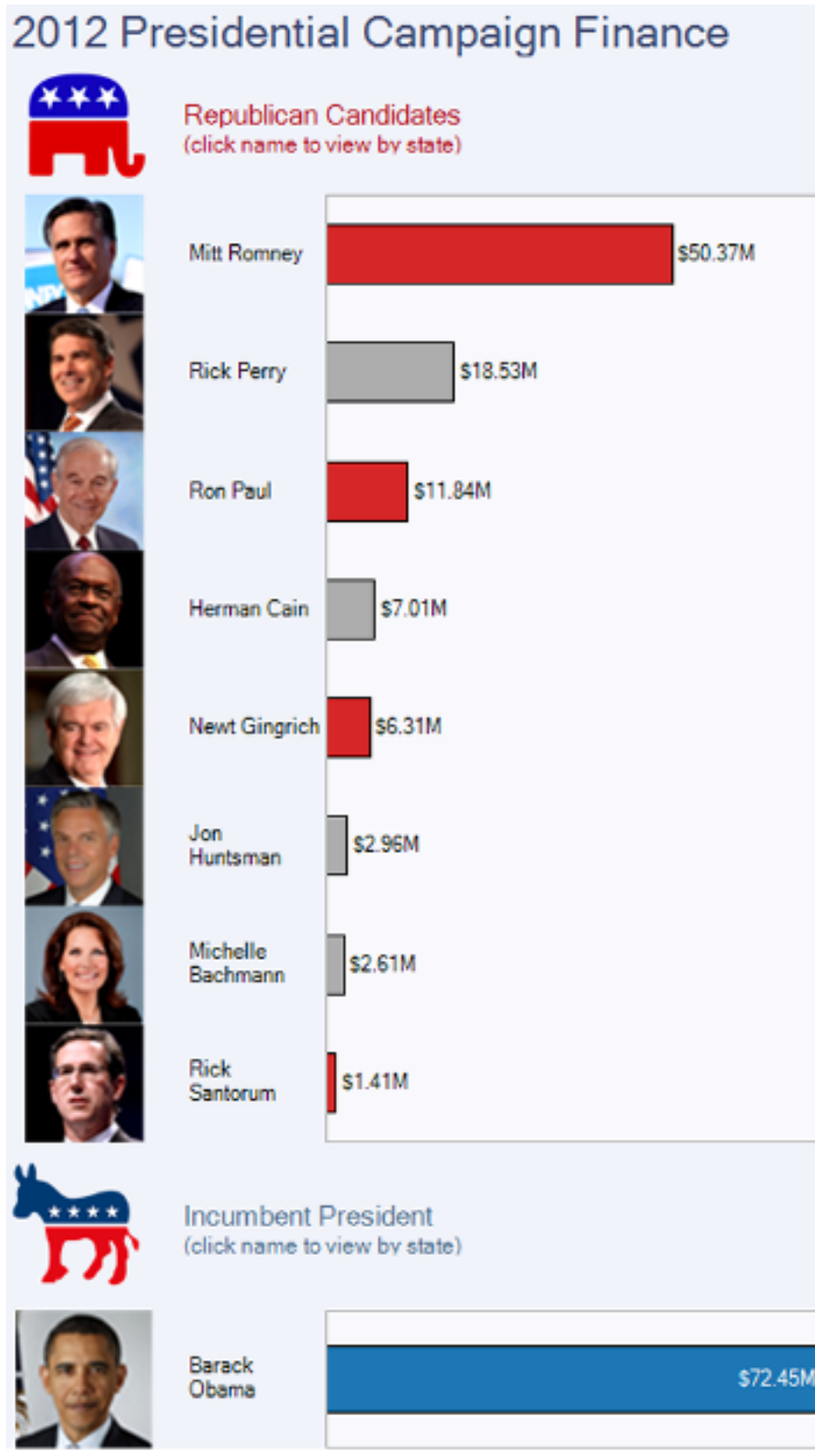


Figure 4.3: [Image extracted from [Data visualization DIY: Our Top Tools].]

Chapter 5

Concluding Remarks

In this survey, most of the purposes and advantages of data visualization in journalism had been presented. Different forms had been introduced, although they are not easy to separate as they all related and combined. In fact, all of these forms make a data and story more clear and easier to comprehend. For example, most of the stories contain a narrative and active form. Both of them help a viewer to understand even some of the most complex datasets.

During exploring examples from the various media, it had become obvious that most of them give much importance and a value to visualisation of the data. Online resources of newspapers are using a number of different tools like Adobe Illustrator, Google Fusion Tables, Datamarket etc to show the data to their readers.

The idea of visualization is a powerful and complements every dataset making it accessible to the viewer. However, without visualization the story consisting of data would be just a statistics report not a journalism story.

Bibliography

- Allen, Paddy. *German bomber to be raised from the sea - interactive*. May 3, 2013. <http://www.guardian.co.uk/uk/interactive/2013/may/03/german-bomber-dornier-goodwin-sands-interactive>.
- Card, Stuart K., Jock D.Mackinlay, and Ben Shneiderman. "Readings in Information Visualization: Using Vision to Think". *San Francisco: Morgan Kaufmann* (1999).
- DeFanti, Thomas A., Maxine D.Brown, and Bruce H. McCormick. "Visualization: Expanding Scientific and Engineering Research Opportunities". *IEEE Computer* (Aug. 22, 1989), pages 12–25.
- et.al., Stephanie Clark. *Growth of domestic counter terrorism organizations*. 2010. <http://projects.washingtonpost.com/top-secret-america/map>.
- GABRIEL DANCE ARON PILHOFER, ANDY LEHREN and JEFF DAMENS. *Faces of the Dead*. Aug. 17, 2012. http://www.nytimes.com/interactive/us/faces-of-the-dead.html?_r=0#/copes_gregory_t.
- Gapminder. *Wealth & Health of Nations**. 2008. <http://www.gapminder.org/world/#example=75>.
- Green, Marc. "Toward a Perceptual Science of Multi-dimensional Data Visualization: Bertin and Beyond". *ERGO/GERO Human Factors Science* (1998).
- Guardian Interactive team Rob Procter, Farida Vis and Alex Voss. *How riot rumours spread on Twitter*. Dec. 7, 2011. <http://www.guardian.co.uk/uk/interactive/2011/dec/07/london-riots-twitter/>.
- Interactive, Pitch. *Out of Sight, Out of Mind. A visualization of all documented drone strikes in Pakistan since 2004*. 2013. <http://drones.pitchinteractive.com/>.
- Oliver, Christine. *The world economy turns the corner*. Jan. 26, 2010. <http://www.guardian.co.uk/business/interactive/2010/jan/26/recession-gdp>.
- Pesanha, Sergio. *The Chances of Getting a Free Upgrade on a Flight*. June 9, 2012. http://www.nytimes.com/interactive/2012/06/10/travel/the-chances-of-getting-a-free-upgrade-on-a-flight.html?_r=0.
- Rogers, Simon. *Data visualization DIY: Our Top Tools*. http://datajournalismhandbook.org/1.0/en/delivering_data_7.html.
- Rogers, Simon. *Government spending by department, 2010-11: get the data*. <http://www.guardian.co.uk/news/datablog/2011/oct/26/government-spending-department-2010-11>.
- Rosenthal, Susken. *Football Drawings*. 1982. <http://www.visualcomplexity.com/vc/project.cfm?id=452>.
- Salcher, Maximilian. *Euripa-Atlas*. 2013. <http://www.sueddeutsche.de/app/wirtschaft/europaatlas/>.

- Segel, Edward and Jeffrey Heer. "Narrative Visualization: Telling Stories with Data". *IEEE Transactions on Visualization and Computer Graphics* 16.6 (2010), pages 1139–1148. doi:10.1109/TVCG.2010.179. <http://vis.stanford.edu/files/2010-Narrative-InfoVis.pdf>.
- Shan Carter, Amanda Cox and Joe Ward. *Paths to the Top of the Home Run Charts*. Aug. 8, 2007. http://www.nytimes.com/ref/sports/20070731%5C_BONDS%5C_GRAPHIC.html.
- Soble, Jonathan. *Toyota timeline: a company history*. Feb. 26, 2010. <http://www.ft.com/cms/s/0/1f8f077c-2301-11df-a25f-00144feab49a.html#axzz2TIlqfrtr>.
- The social component bar on New York Times article webpage*. http://www.nytimes.com/interactive/2012/06/10/travel/the-chances-of-getting-a-free-upgrade-on-a-flight.html?_r=1%5C&.
- Tufte, Edward. *The Life Cycle of a Japanese Beetle*. http://www.edwardtufte.com/tufte/advocate_1099.