Electronics and
Computer Science

UNIVERSITY OF Southampton

Intro to data visualisation: History

Lecture 3, 28.01.2016

Dr. Chris Phethean

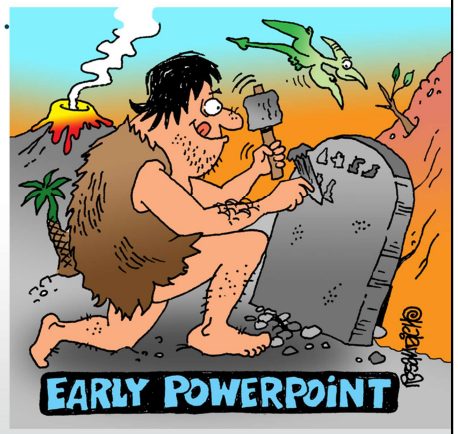
Data Visualisation (COMP6234)

Summary

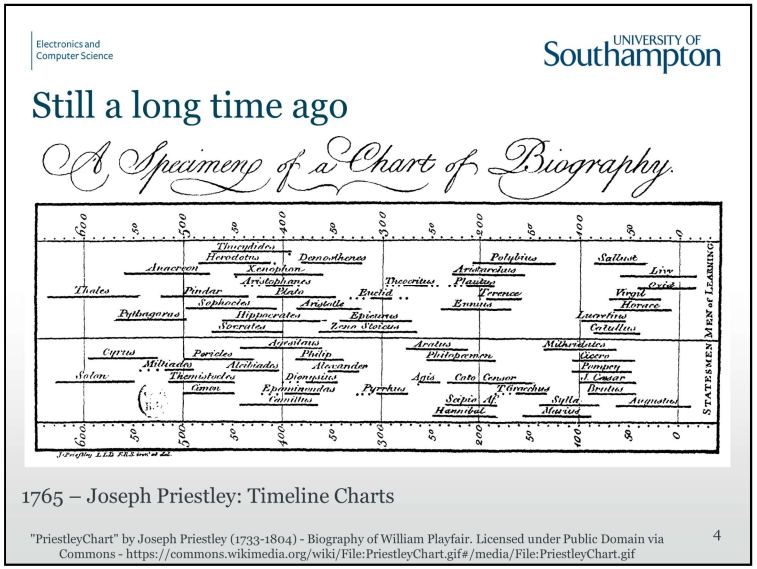
- Origin of graphs
- Visualisation forms and context
- Visualisations today

A long time ago…

- There were no charts
- No Venn diagrams
- No PowerPoint!



http://danscartoons.com/tombstone-cartoons/early-powerpoint/



A Chart of Biography

Births and deaths of “statesmen of learning”

One of first charts showing relative length and durations as a chart rather than a table,

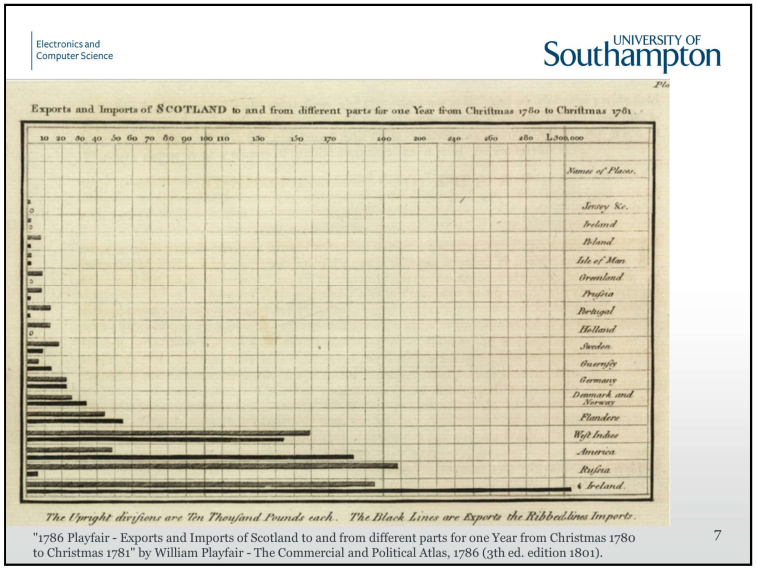
Still a long time ago

1786 - William Playfair publishes 'The Commercial and Political Atlas"

Still a long time ago

1786 - William Playfair publishes 'The Commercial and Political Atlas"

Any ideas about why this is important?



Licensed under Public Domain via Commons - https://commons.wikimedia.org/wiki/File:1786_Playfair_-_Exports_and_Imports_of_Scotland_to_and_from_different_parts_for_one_Year_from_Christmas_1780_to_Christmas_1781.jpg#/media/File:1786_Playfair_-_Exports_and_Imports_of_Scotland_to_and_from_different_parts_for_one_Year_from_Christmas_1780_to_Christmas_1781.jpg

Created bar charts

First ever bar chart, created because Mayfair didn't have enough data to make a time series. (He actually apologised for using it)

The First Bar Chart

- Before this, all charts were line graphs or surface charts (line graphs with shaded areas underneath)
- Line charts useful for showing time-series data
- Due to a lack of data, Playfair used a single year's worth of data and represented each country as a bar.
- It was the first graphic that did not locate data in space (e.g. co-ordinates) or time.

Read more: Beniger, James R., and Dorothy L. Robyn. 1978. "Quantitative Graphics in Statistics: A Brief History". *The American Statistician* 32 (1). [American Statistical Association, Taylor & Francis, Ltd.]: 1–11. doi:10.2307/2683467.

The First Bar Chart

This Chart is different from the others in principle, as it does not comprehend any portion of time, and it is much inferior in utility to those that do; for though it gives the extent of the different branches of trade, it does not compare the same branch of commerce with itself at different periods; nor does it imprint upon the mind that distinct idea, in doing which, the chief advantage of Charts consists: for as it wants the dimension that is formed by duration, there is no shape given to the quantities.

Read more: Beniger, James R., and Dorothy L. Robyn. 1978. "Quantitative Graphics in Statistics: A Brief History". *The American Statistician* 32 (1). [American Statistical Association, Taylor & Francis, Ltd.]: 1–11. doi:10.2307/2683467.

The First Bar Chart: An apology

This Chart is different from the others in principle, as it does not comprehend any portion of time, and it is much inferior in utility to those that do; for though it gives the extent of the different branches of trade, it does not compare the same branch of commerce with itself at different periods; nor does it imprint upon the mind that distinct idea, in doing which, the chief advantage of Charts consists: for as it wants the dimension that is formed by duration, there is no shape given to the quantities.

Read more: Beniger, James R., and Dorothy L. Robyn. 1978. "Quantitative Graphics in Statistics: A Brief History". *The American Statistician* 32 (1). [American Statistical Association, Taylor & Francis, Ltd.]: 1–11. doi:10.2307/2683467.

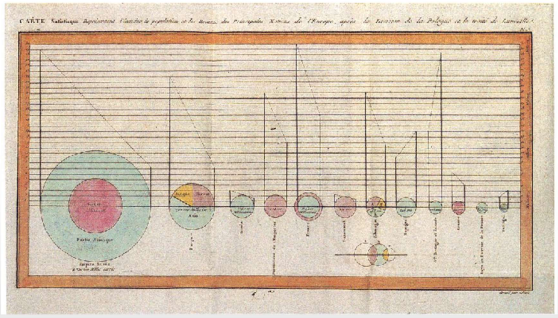
Bar Charts

- Arose from the *lack* of data for one example
- But popularity is now because of the need to explain *large amounts* of data.

Electronics and Computer Science

UNIVERSITY OF Southampton

Playfair Continues…



Pie Charts and Circle Charts soon follow

"Playfair piecharts" by William Playfair (1759–1823) - Scanned from the book "The Commercial and Political Atlas and Statistical Breviary", Cambridge University Press 2005

12

"Playfair piecharts" by William Playfair (1759–1823) - Scanned from the book "The Commercial and Political Atlas and Statistical Breviary", Cambridge University Press 2005.. Licensed under Public Domain via Commons - https://commons.wikimedia.org/wiki/File:Playfair_piecharts.jpg#/media/File:Playfair_piecharts.jpg

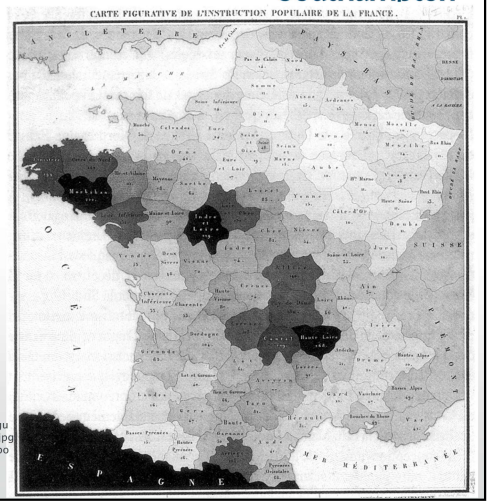
Electronics and Computer Science

UNIVERSITY OF Southampton

Choropleth Maps - 1826

- Charles Dupin
- Cartogram
- Regions coloured/shaded to illustrate a statistic
- E.g. Literacy rate.

https://commons.wikimedia.org/wiki/File:Carte_figurative_de_l%27instruction_populaire_de_la_France.jpg#/media/File:Carte_figurative_de_l%27instruction_populaire_de_la_France.jpg



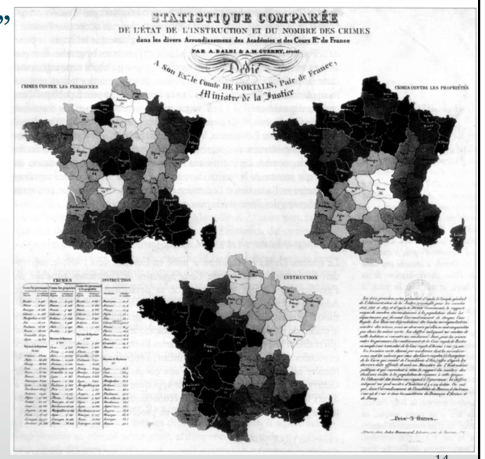
"Carte figurative de l'instruction populaire de la France" by Charles Dupin (1784-1873) - http://math.yorku.ca/SCS/Gallery/images/dupin1826-map_200.jpg. Licensed under Public Domain via Commons - https://commons.wikimedia.org/wiki/File:Carte_figurative_de_l%27instruction_populaire_de_la_France.jpg#/media/File:Carte_figurative_de_l%27instruction_populaire_de_la_France.jpg

Electronics and Computer Science

UNIVERSITY OF Southampton

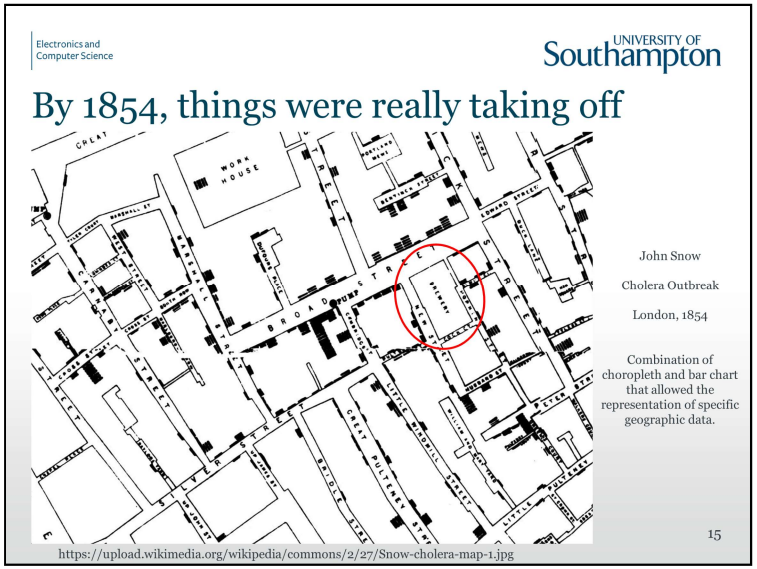
"Moral Statistics"

- André-Michel Guerry expanded the idea in 1833 with his 'Essay on moral statistics of France'
- Suicide and crime stats, with maps for almost every variable available
- Social and economic impact of the charts
- Led to the new field of 'moral statistics'



14

Data representing social pathology ('abnormality') of groups – crime, illiteracy, suicide, divorce etc.



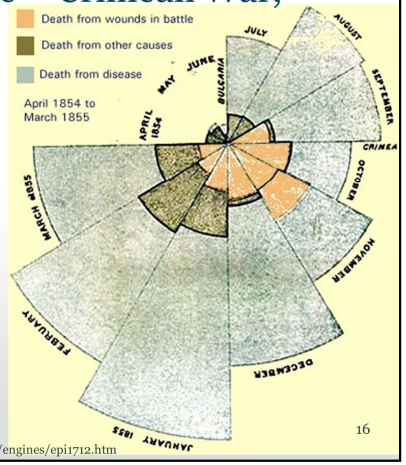
Use of visualisation to actually change something.
Everyone thinks that cholera was spread through air, used this to show it is water based
- Combines heat map with bar chart. Bar at each house to represent death tolls in that house.
Relates it to pump locations (brewery has its own water pump)

Electronics and Computer Science

UNIVERSITY OF Southampton

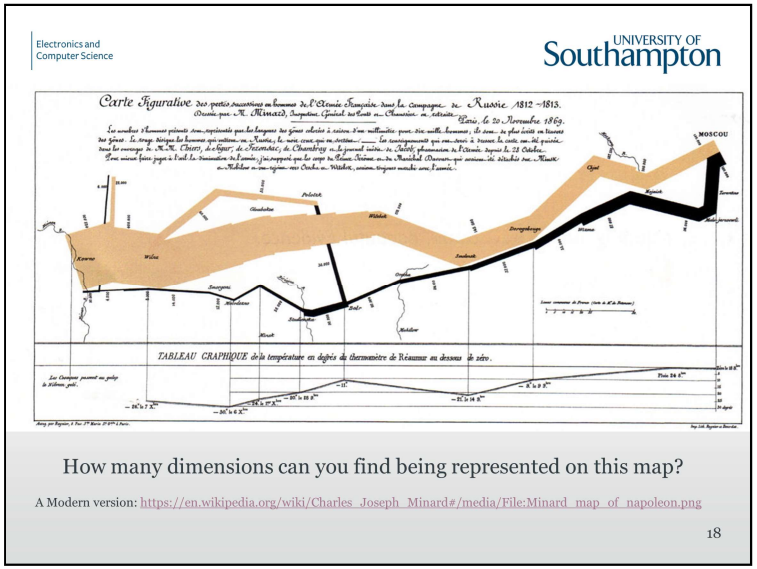
Florence Nightingale – Crimean War, 1858

- Invented the "polar area chart"
  - Based on Playfair's pie charts from 1801…
- Wanted to highlight how many deaths were from various causes
- And the importance of sanitary hospital conditions!



http://www.uh.edu/engines/epi1712.htm

Now things go really crazy…



Ask class to come up and look at the screen closely. Then write on the board how many different dimensions they found.

1) Size of the army
2) And 3) Path (lat / long) taken on map
3) As above
4) Direction army was travelling
5) Temperature
6) Dates

So, why were all these charts invented?

So, why were all these charts invented?

- Nightingale and Snow charts:
  - Improving hospital conditions
  - Improving understanding about disease

- Clear interface to the underlying data
- Tell a story that others can understand
  - Resulted in real, beneficial changes for humanity

So, why were all these charts invented?

- Main problem that they all solve
  - Raw data is space and time consuming
    - Graphs/charts are much smaller, and quicker to process

Lets think about some more modern things

- What do you think is the most widely recognised alphabet in the world?

Alphabet (according to Wikipedia)

An **alphabet** is a standard set of letters (basic written symbols or graphemes) which is used to write one or more languages based on the general principle that the letters represent phonemes (basic significant sounds) of the spoken language.

Does this not just sound like visualisation?

Who recognises these?

Devised in the 1970s for VHS tapes!

Lots of Data

- The popularity of the VHS symbols is due to the scale by which electronic and digital devices have spread
- These also help create lots more data!
  - We discussed last time how much data was created every minute
- But modern technology also allows new ways to show this data

Communication Types for Explaining Data

- Presentations
- Briefings
- Bulletins
- Dashboards
- Campaigns

Each of these can be used to tell a story in a different way, and to a different audience

Electronics and Computer Science

UNIVERSITY OF Southampton

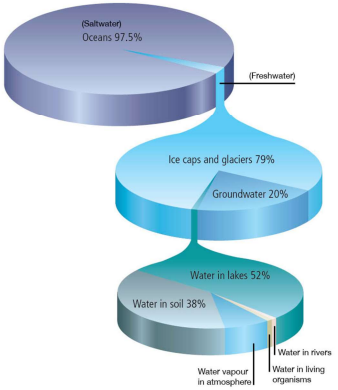
Communication Types for Explaining Data

- Presentations
  - Information broken down into slides
- Briefings
  - Info with short summaries, indexed and signposted
- Bulletins
  - Short, focusing on a clear story
- Dashboards
  - Simple ways to highlight key messages
- Campaigns
  - Little or prior knowledge assumed; desired action as clear as possible

27

Examples:
Bulletins - Emails
Campaigns - Posters

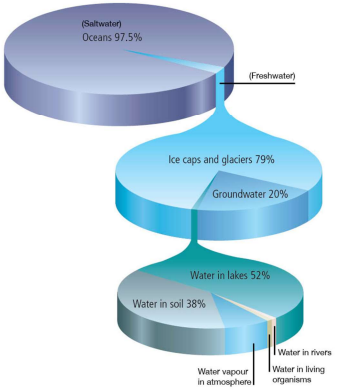
Not all visualisations are good…



Just because some things *can* be done, doesn't mean they should be done

https://thiswoo.wordpress.com/2008/10/21/i-give-you-statistical-pie-charts/

Not all visualisations are good…



Pros:

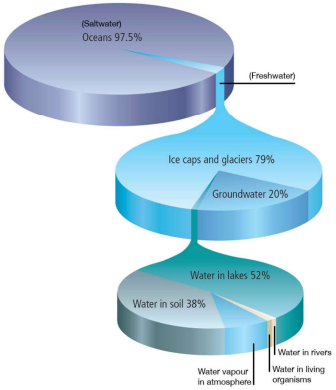
- Tells a story of how much (or little) water on the planet is consumable

Cons:

- 3D Pies are deceptive!
- How big is the % of freshwater?
- 3d effects make the freshwater slice look bigger than it actually is
- Colours in 2nd and 3rd tier are confusing (groundwater appears separate to freshwater)
- Sizing of each pie is misleading

https://thiswoo.wordpress.com/2008/10/21/i-give-you-statistical-pie-charts/
http://chandoo.org/wp/2012/11/09/pie-of-pie-of-pie-chart/

Not all visualisations are good…



The problem?

Adding "design" to graphics which obscures the story.

https://thiswoo.wordpress.com/2008/10/21/i-give-you-statistical-pie-charts/
http://chandoo.org/wp/2012/11/09/pie-of-pie-of-pie-chart/

- More and more techniques are becoming available for visualising and telling stories from data
- Not *all* of them are good.
- But they all stem from a need to communicate value and meaning from data

And not everyone has the skill to make a clear chart…

- https://www.youtube.com/watch?v=-rbyhj8uTT8

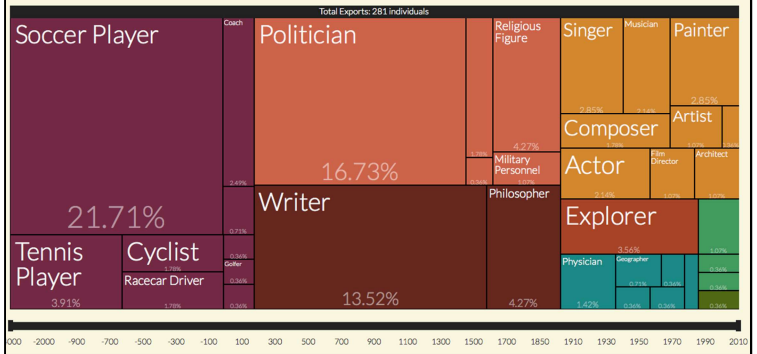
Electronics and Computer Science

UNIVERSITY OF Southampton

A modern chart: Treemaps

Who are the globally known people born within present day **Spain***?

[4000 B.C. – 2010]



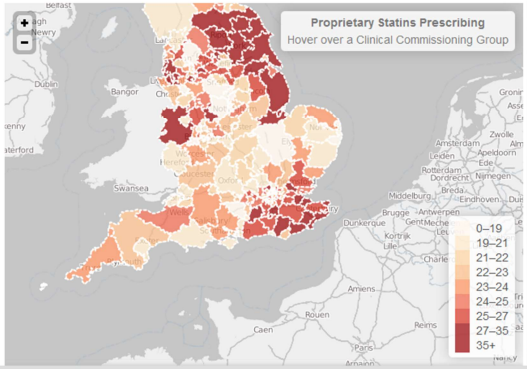
http://pantheon.media.mit.edu/treemap/country_exports/ES/all/-4000/2010/H15/pantheon

Electronics and Computer Science

UNIVERSITY OF Southampton

Same charts, Modern Technology

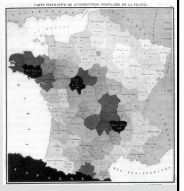
Percentage of proprietary statin prescribing by CCG Sep 2011 - May 2012



Open prescription data analysis

Where are branded drugs being prescribed be GPs?

How can the NHS save money?



http://www.prescribinganalytics.com

Modern visualisations

- http://youtu.be/ystkKXzt9Wk

Summary

- Visualisation is not new
- Ways of telling stories from data have evolved over hundreds of years
  - New techniques have been introduced to help convey meanings from increasingly complex datasets
  - The focus should always be on the story, and communicating this as effectively as possible
- Unnecessary design can cloud that story

Further reading

- Suda, Brian. *A Practical Guide to Designing with Data*. Mark Boulton Design Limited, 2010.
- Beniger, James R., and Dorothy L. Robyn. 1978. "Quantitative Graphics in Statistics: A Brief History". *The American Statistician* 32 (1). [American Statistical Association, Taylor & Francis, Ltd.]: 1–11. doi:10.2307/2683467.

Next Week…

- Monday
  - Sri will go over the steps involved in planning a visualisation
  - And how various types of data can be encoded visually
- Thursday
  - First tutorial session
    - Google Charts