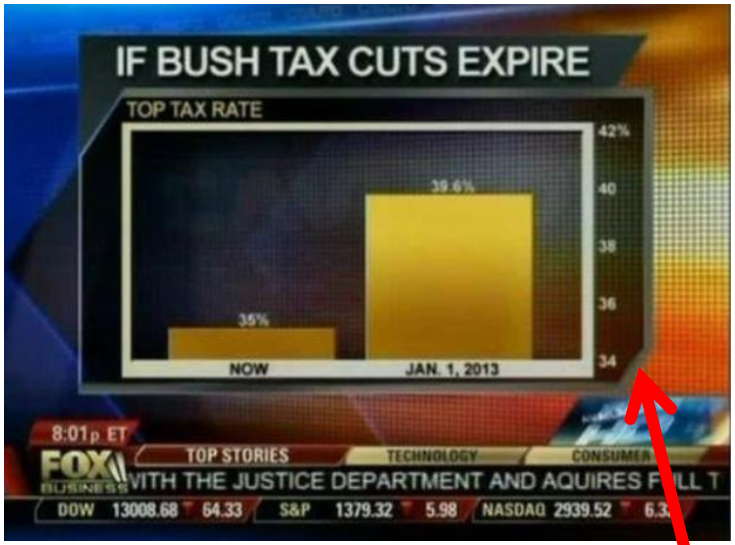
Visualizations

Computational Social Science Course

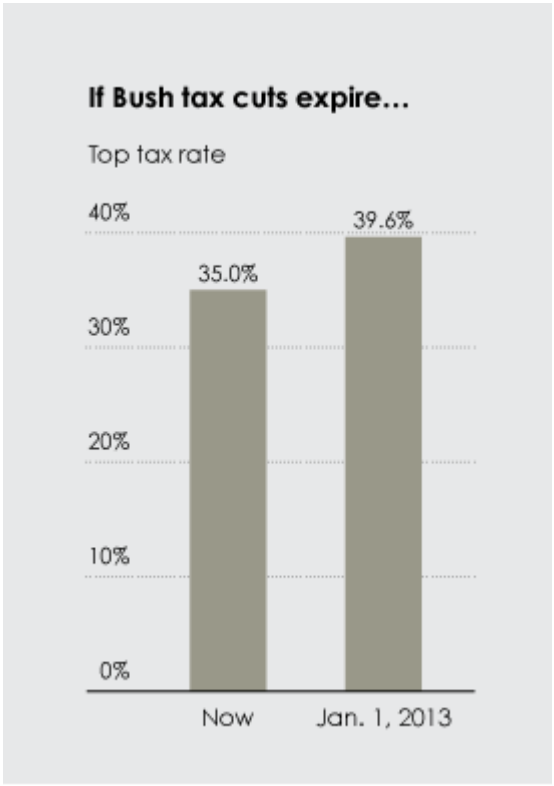
JProf. Dr. Claudia Wagner

Missleading Visualizations

- Avoid Distortion

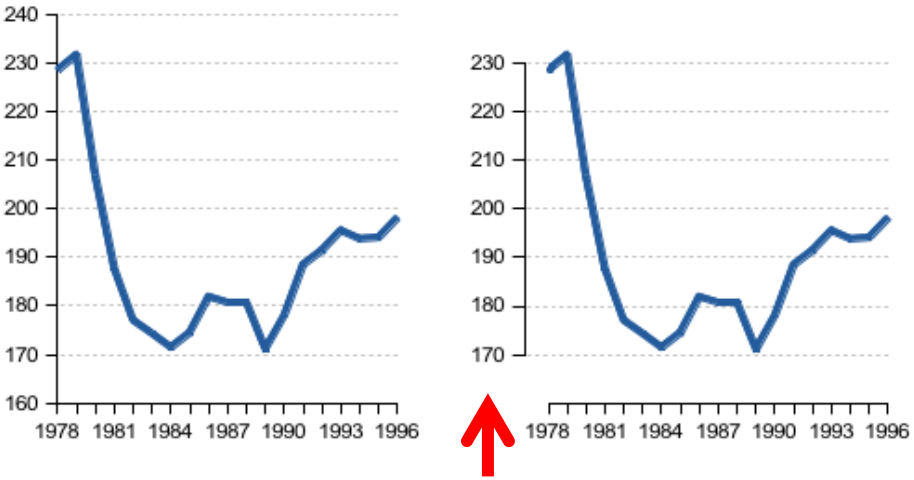


Axis does not start with zero



http://flowingdata.com/2012/08/06/fox-news-continues-charting-excellence/

- Avoid Chartjunk
  - anything that can be removed from a chart without changing its meaning is chartjunk

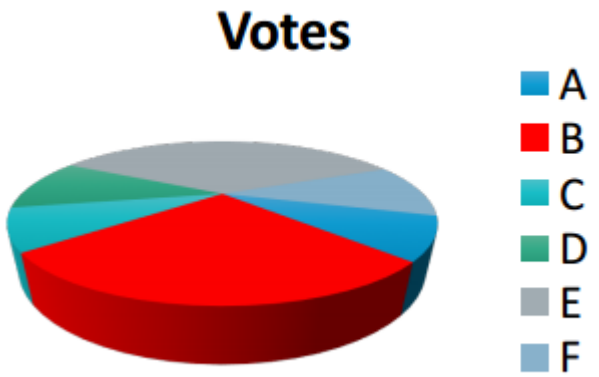


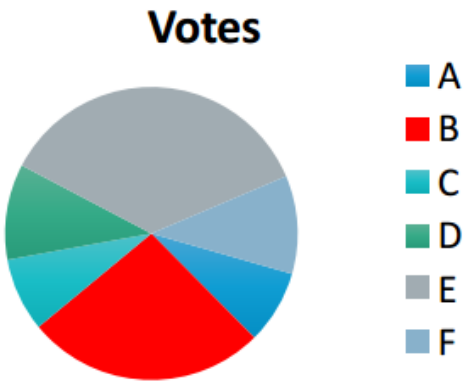
Gap helps to see that line does not start from zero

http://flowingdata.com/2012/08/06/fox-news-continues-charting-excellence/

Good Charts

- What did most people vote for?
- What did least people vote for?



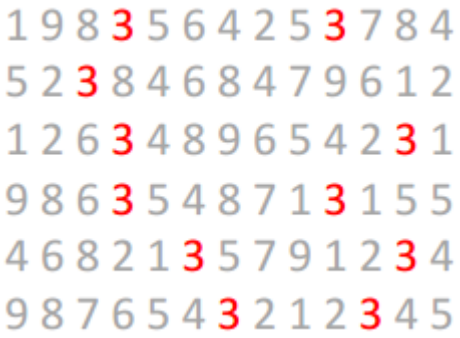


Party	Votes
A	60
B	195
C	60
D	78
E	265
F	80

- 3D charts are usually bad

How many 3‘s can you see?

1 9 8 3 5 6 4 2 5 3 7 8 4
5 2 3 8 4 6 8 4 7 9 6 1 2
1 2 6 3 4 8 9 6 5 4 2 3 1
9 8 6 3 5 4 8 7 1 3 1 5 5
4 6 8 2 1 3 5 7 9 1 2 3 4
9 8 7 6 5 4 3 2 1 2 3 4 5



Good visualizations convey information effectively

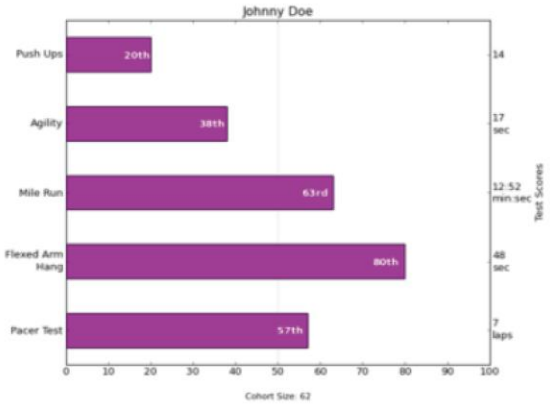
Source: Christ Phethean http://videolectures.net/eswc2016_phethean_understanding_communicating/

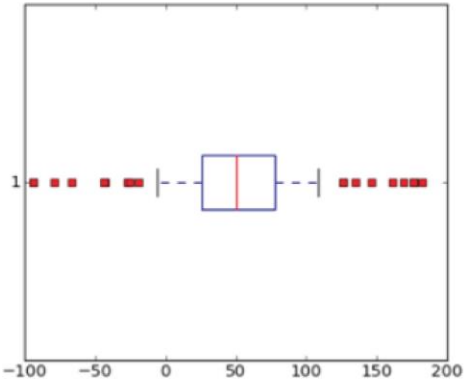
Univariate Data

1 Variable only!

Describe it: e,g, central tendency, dispersion,

Plot it: frequency distributions, bar graph, histogram, pie chart, line, graph, box-and-whisker plot

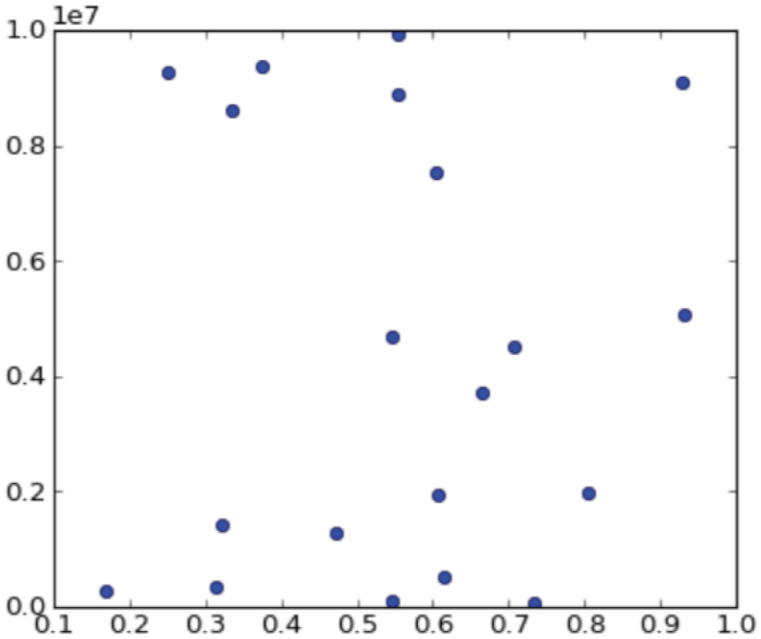




Based on slide from M. Agrawala

Bivariate Data

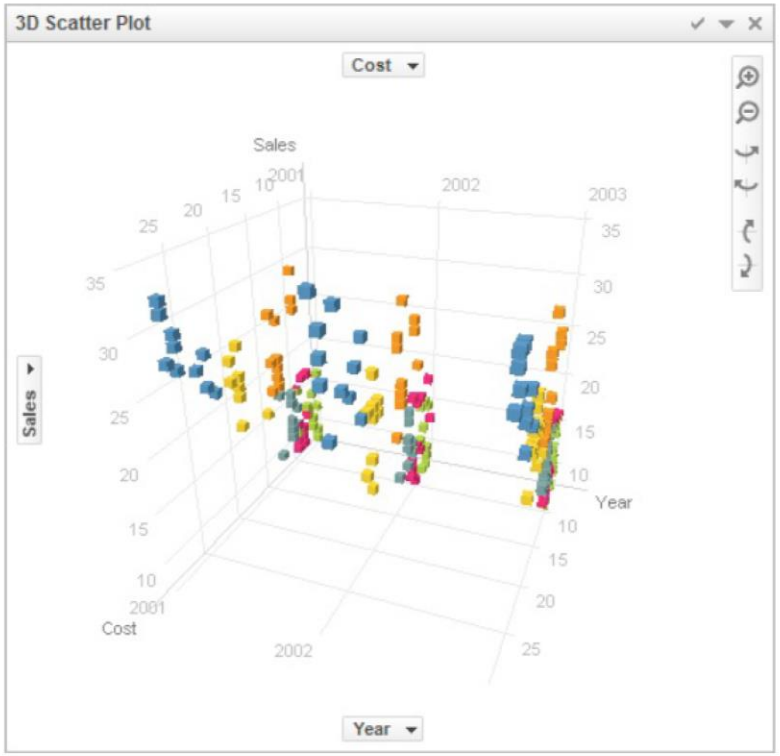
Scatterplot is common



Based on slide from M. Agrawala

Trivariate Data

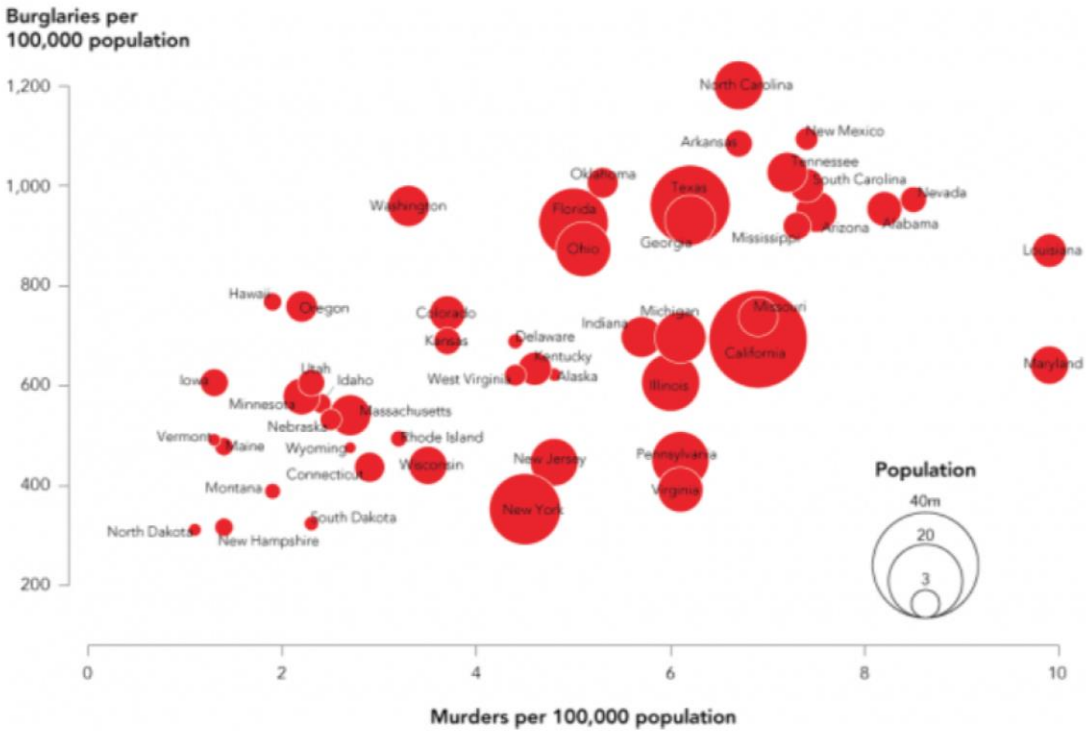
Do NOT use 3D scatterplots!



Based on slide from M.Agrawala

Trivariate Data

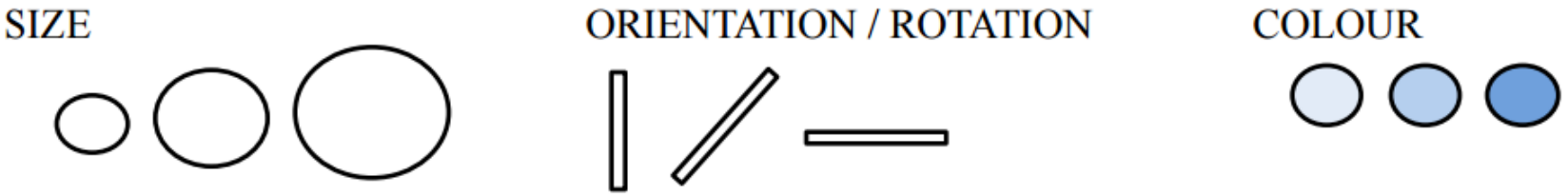
Map the third dimension to some other visual attribute



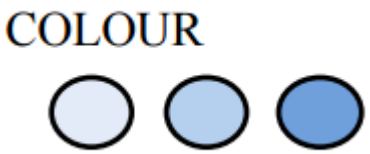
Based on slide from M. Agrawala

Types of Data

- Ordered Data

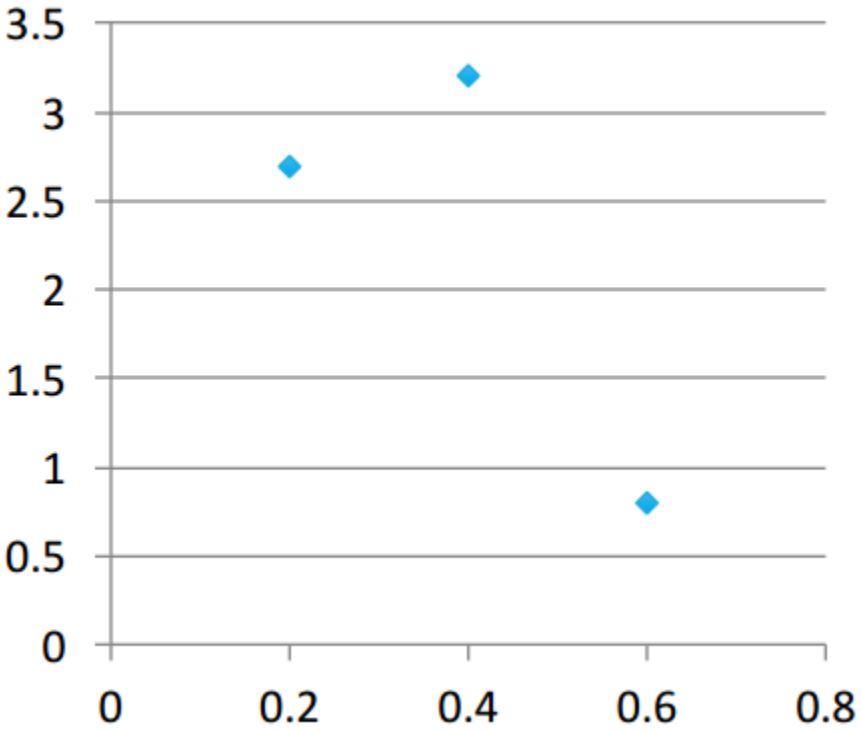


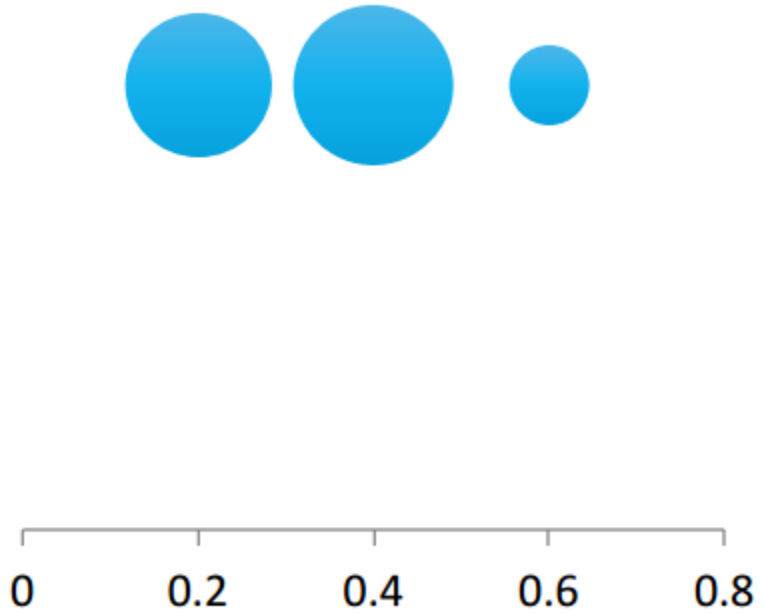
- Categorical Data



Source: Christ Phethean http://videolectures.net/eswc2016_phethean_understanding_communicating/

Y(0.2) < Y(0.3)





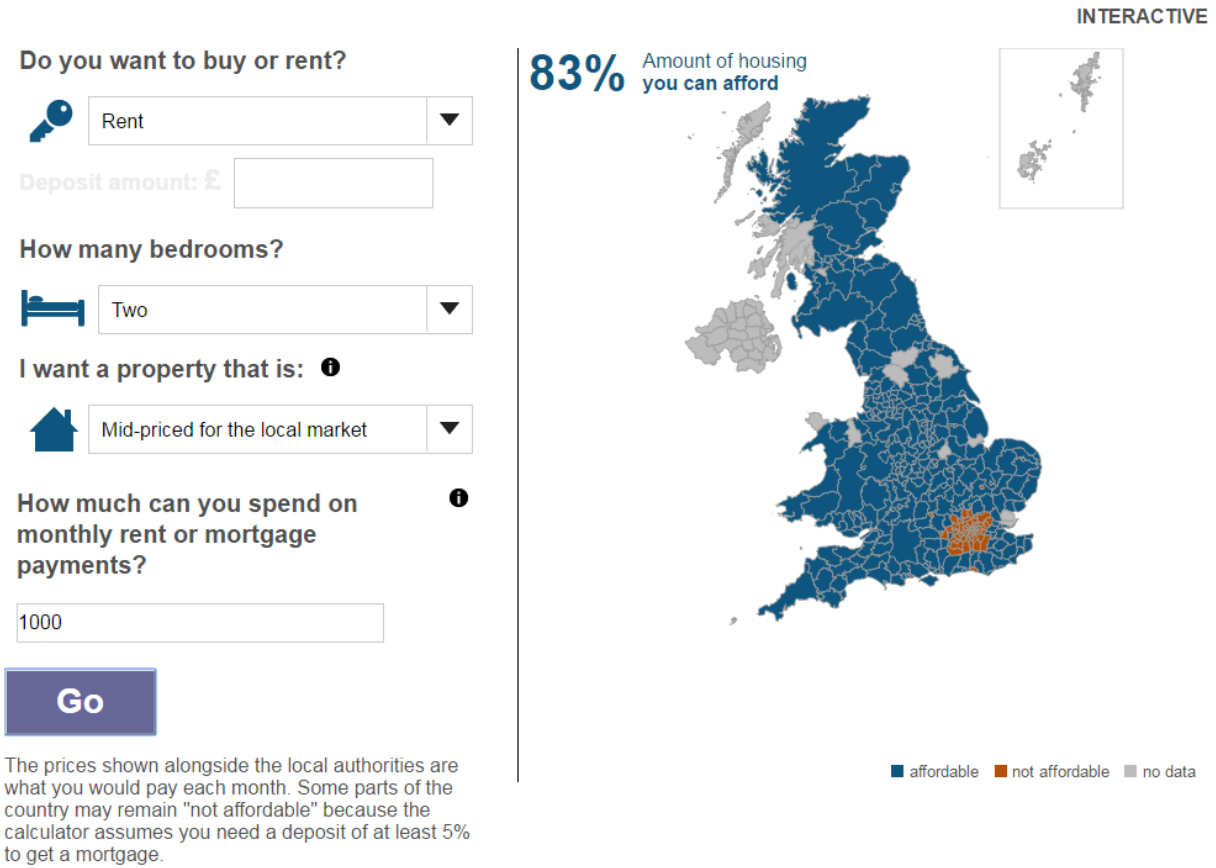
Source: Christ Phethean http://videolectures.net/eswc2016_phethean_understanding_communicating/

Good Charts

- Check the data (e.g. explain outlier points)
- Explain encodings (e.g. colors)
- Label Axes
  - Include units (i.e., percentage, cm)
- Include your sources (online that's often useful; people copy images and reshare them)
- Consider your audience and situation (how much details do they need; prior knowledge? What visualizations are they used to? Science versus journalism; how much time do they have?)

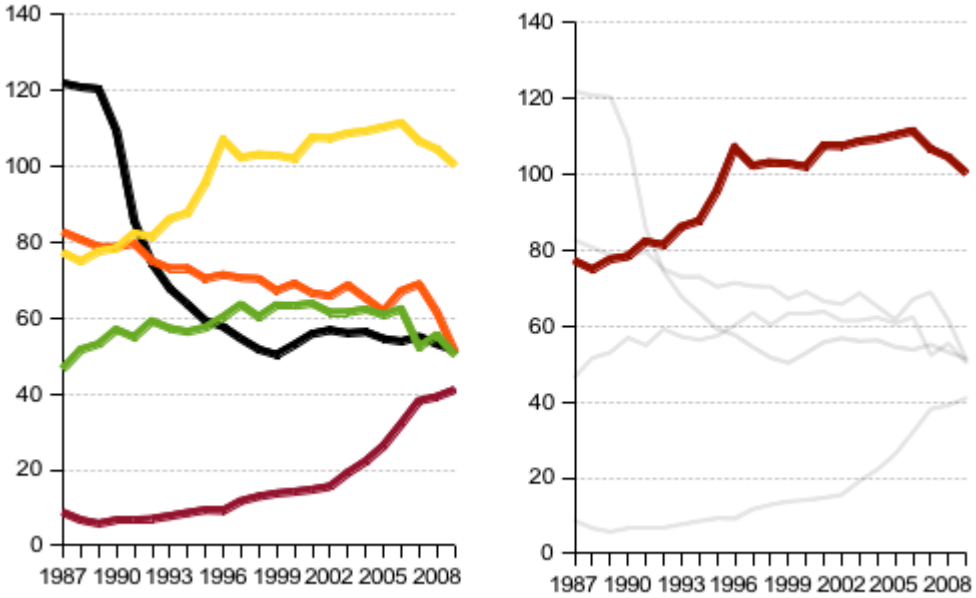
Explanatory versus Exploratory Charts

- Exploratory Charts:
  - E.g. interactive chart
    - http://www.bbc.com/news/business-23234033

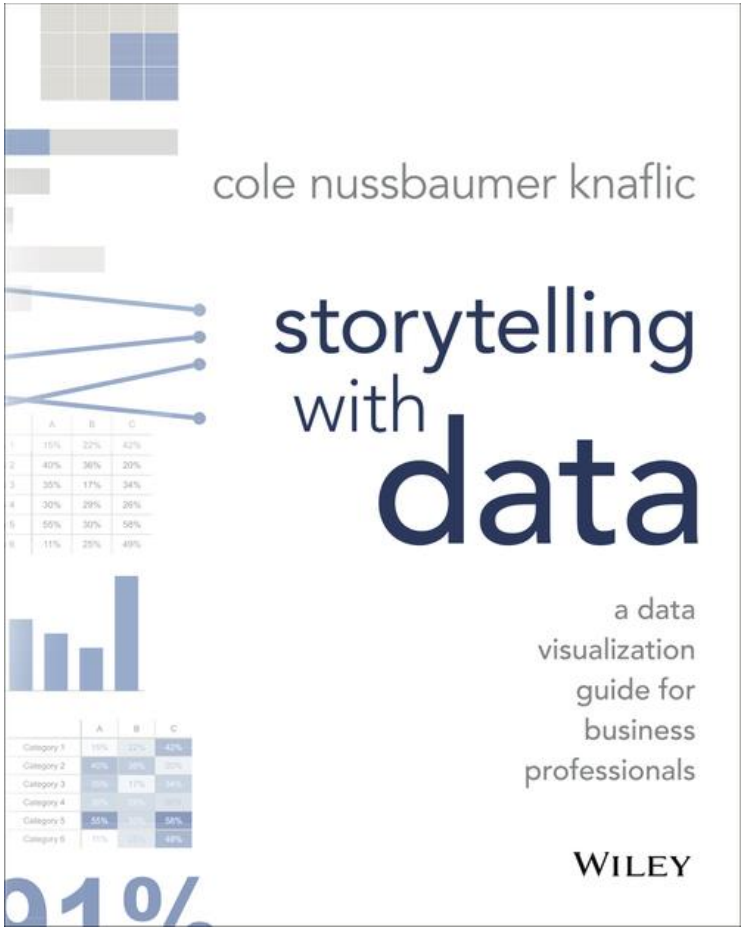


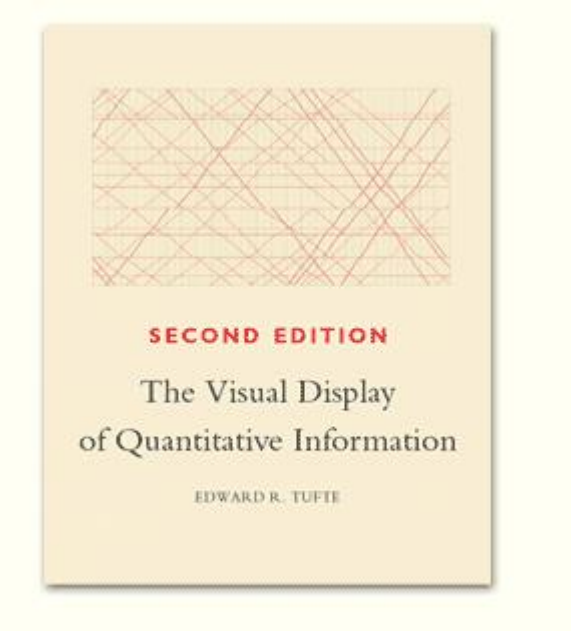
Explanatory versus Exploratory Charts

- Explanatory Chart:
  - What is your main point? What should people see
    - Highlight it!



Background Reading





Any further questions?

See you next week