

Background



Lecture 2

Data vs Information

- Data is *not* information
- Data without context is useless
- Data + Interpretation -> Information / Knowledge
but
- Junk + Fancy Presentation -> Fancy Junk

Learn to see through glitz to the underlying information

Visualisation Goundwork

- We already have experience of visualisation in the form of *graphs*
- Graphical information displays are so common we don't think about them
- Skill with basic graphs is a foundation for more advanced techniques

Graphing Tips

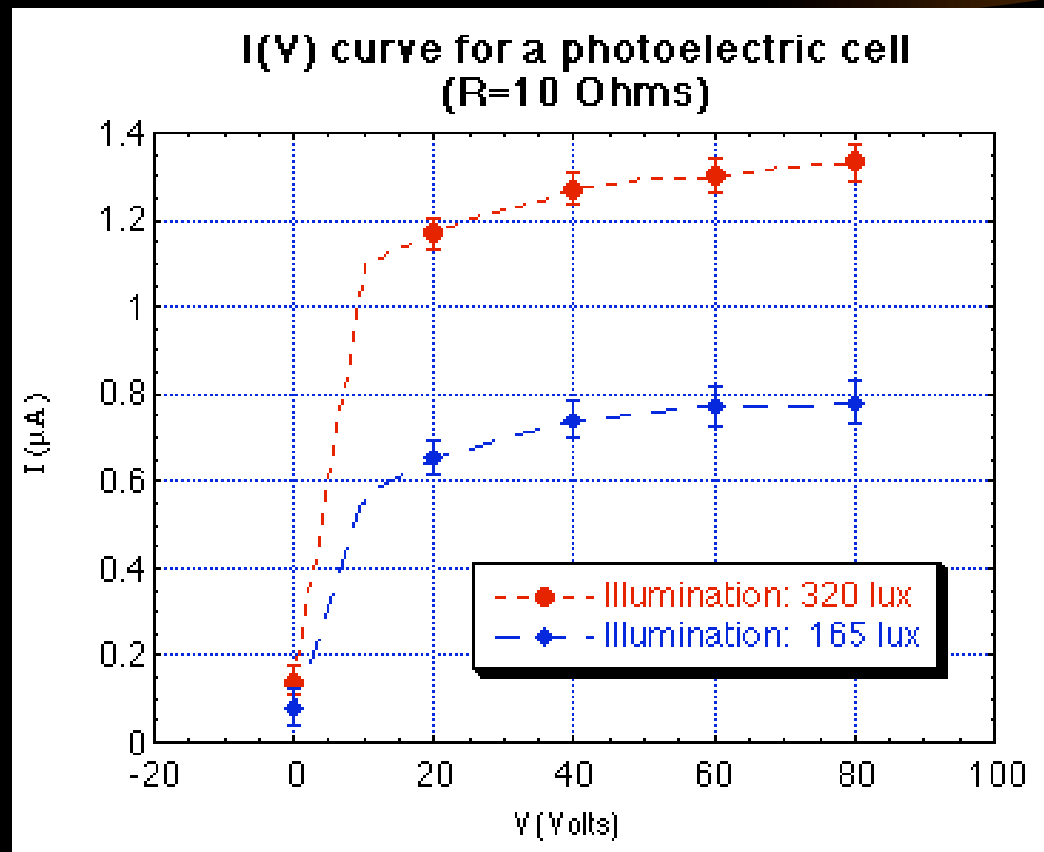
- A graph without context is meaningless
 - Add a Title / Caption
 - Label axes (text, numbers and units)
- Choose an appropriate graph type and scale
- Use a suitable origin
- Avoid visual ambiguity
- Avoid “chart junk”

Common Graph Types

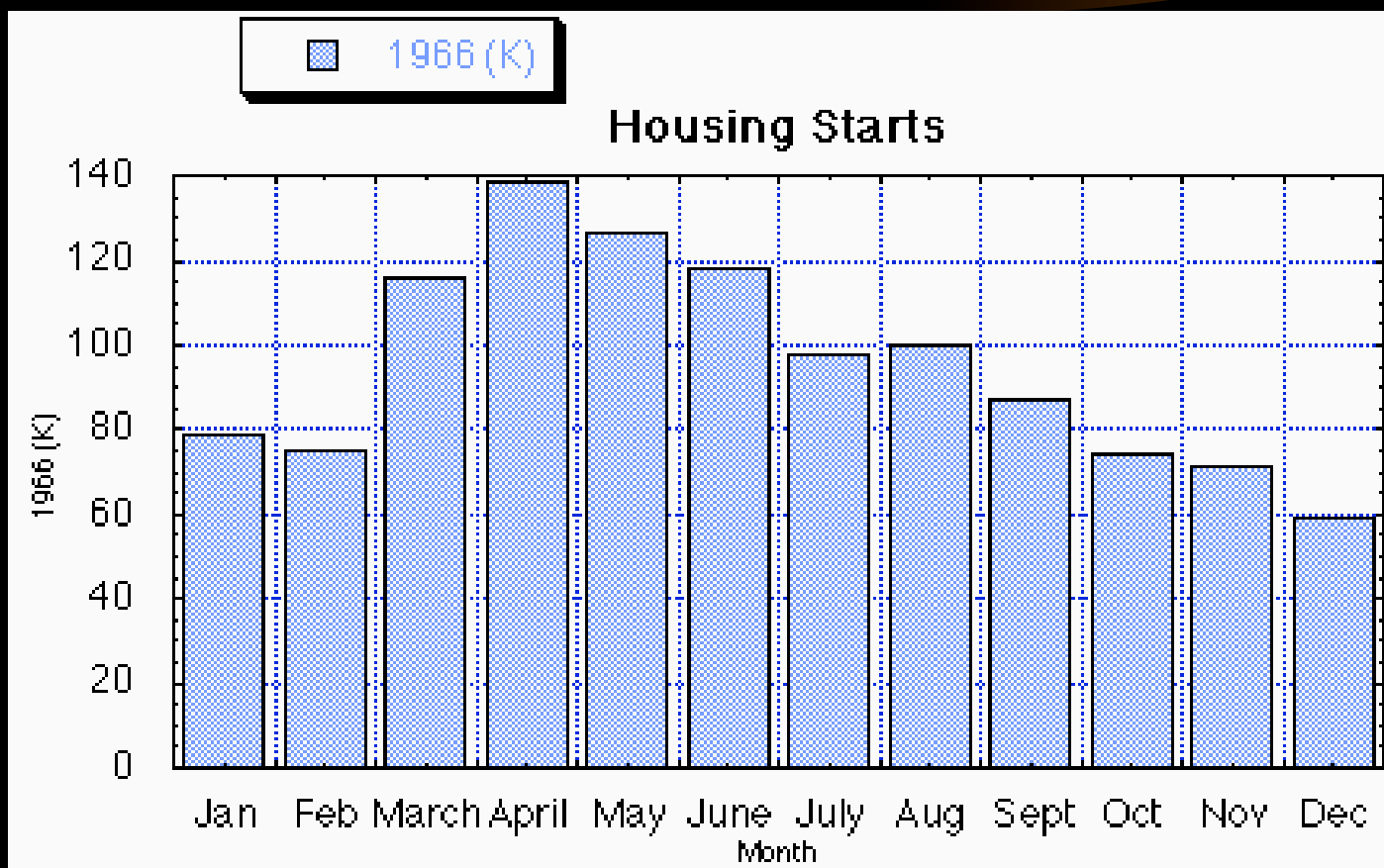


Line Graph

Not forgetting the error bars...

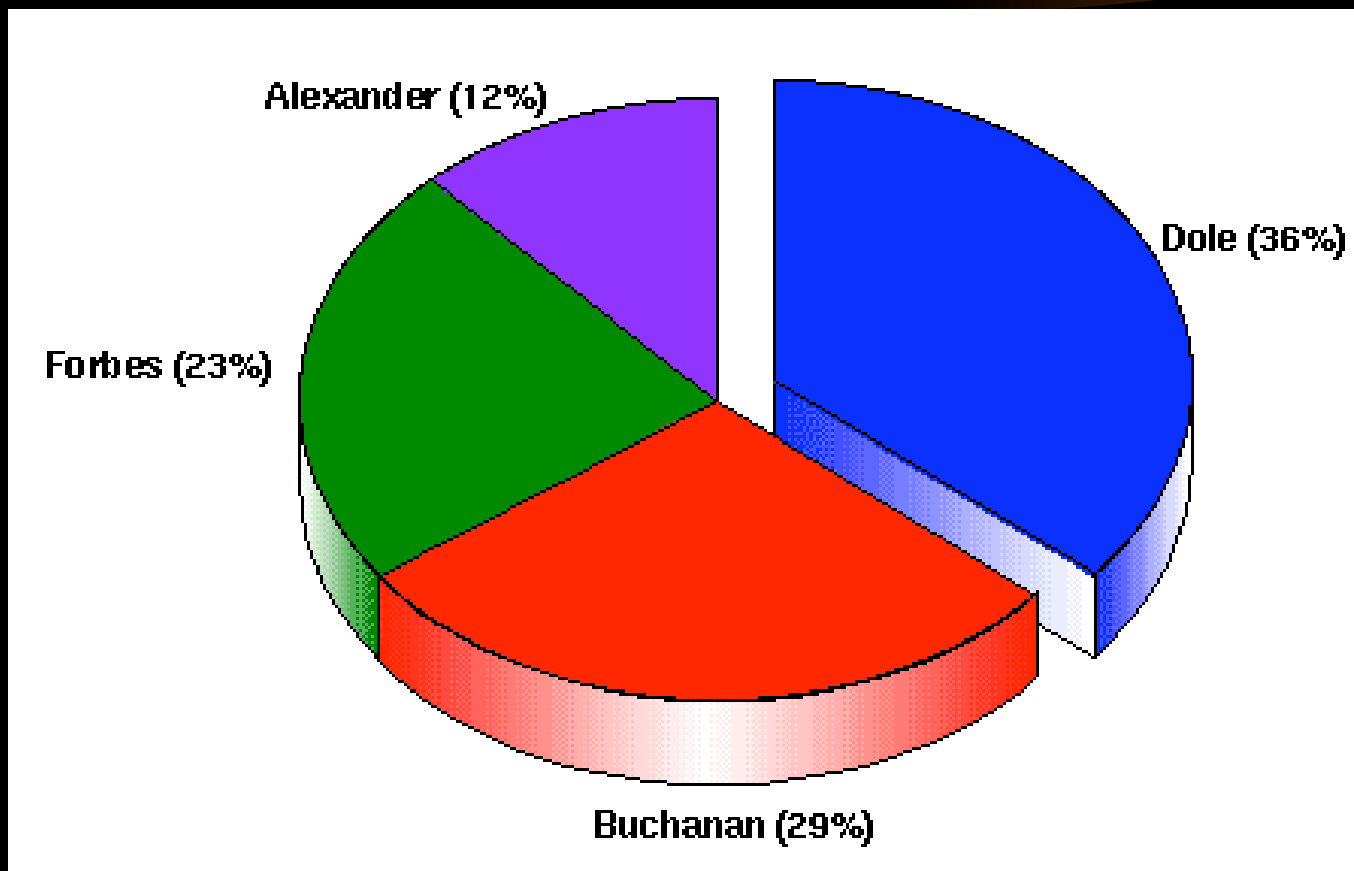


Bar Chart



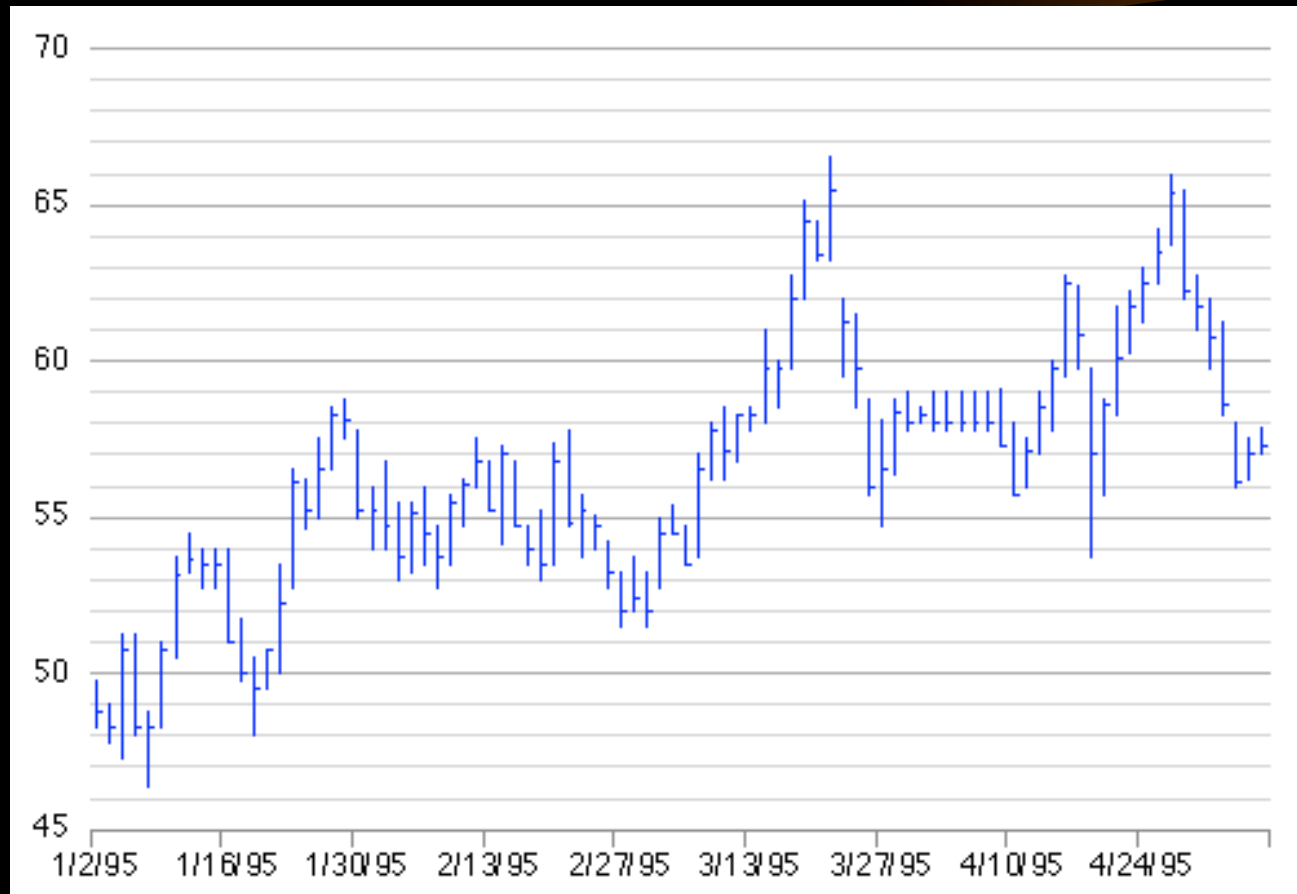
Pie Chart

Pie Charts compare values to the whole



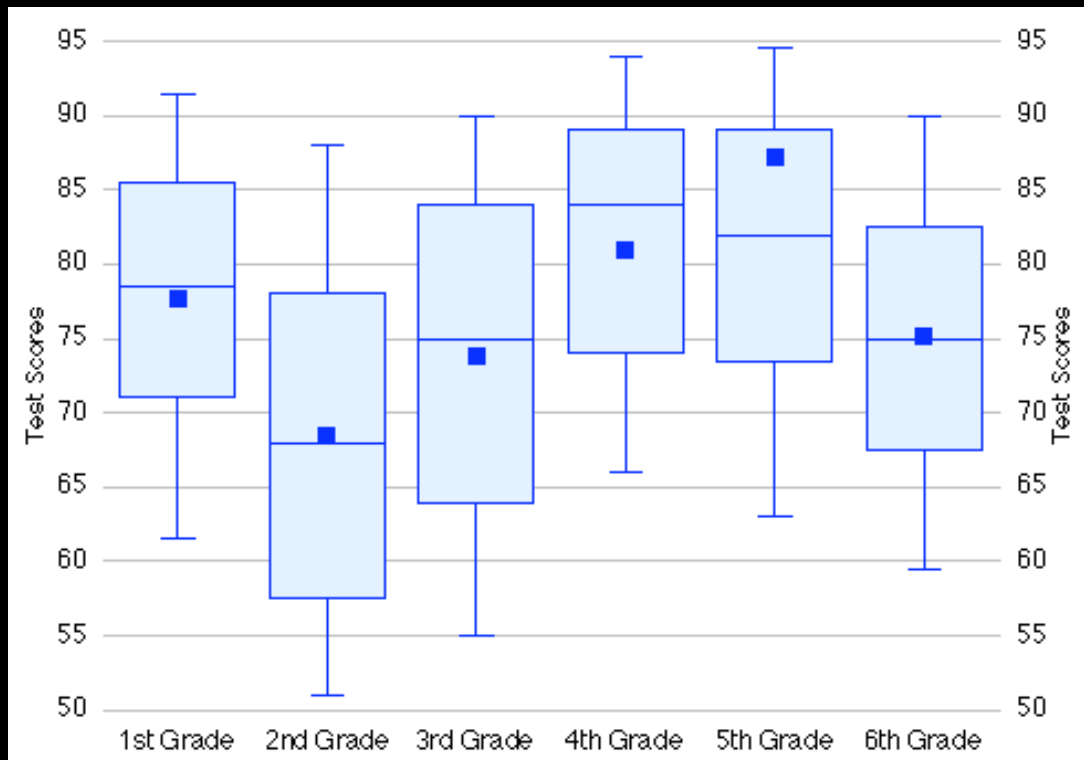
High-Low Graph

Share prices: Highest, Lowest and Closing Values

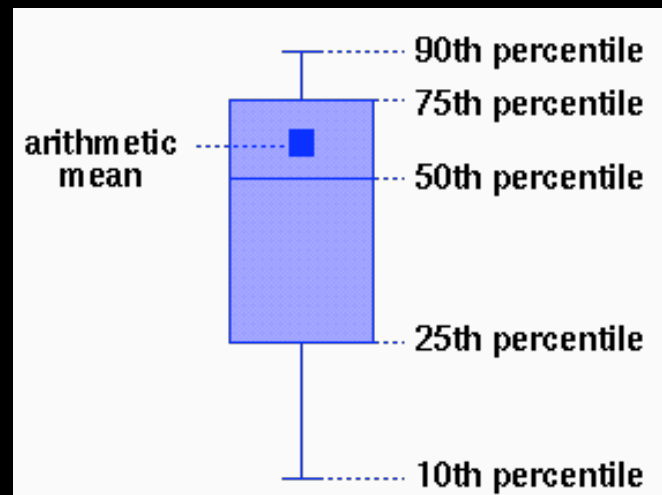


Box and Whisker Charts

Statistical data in a neatly presented form

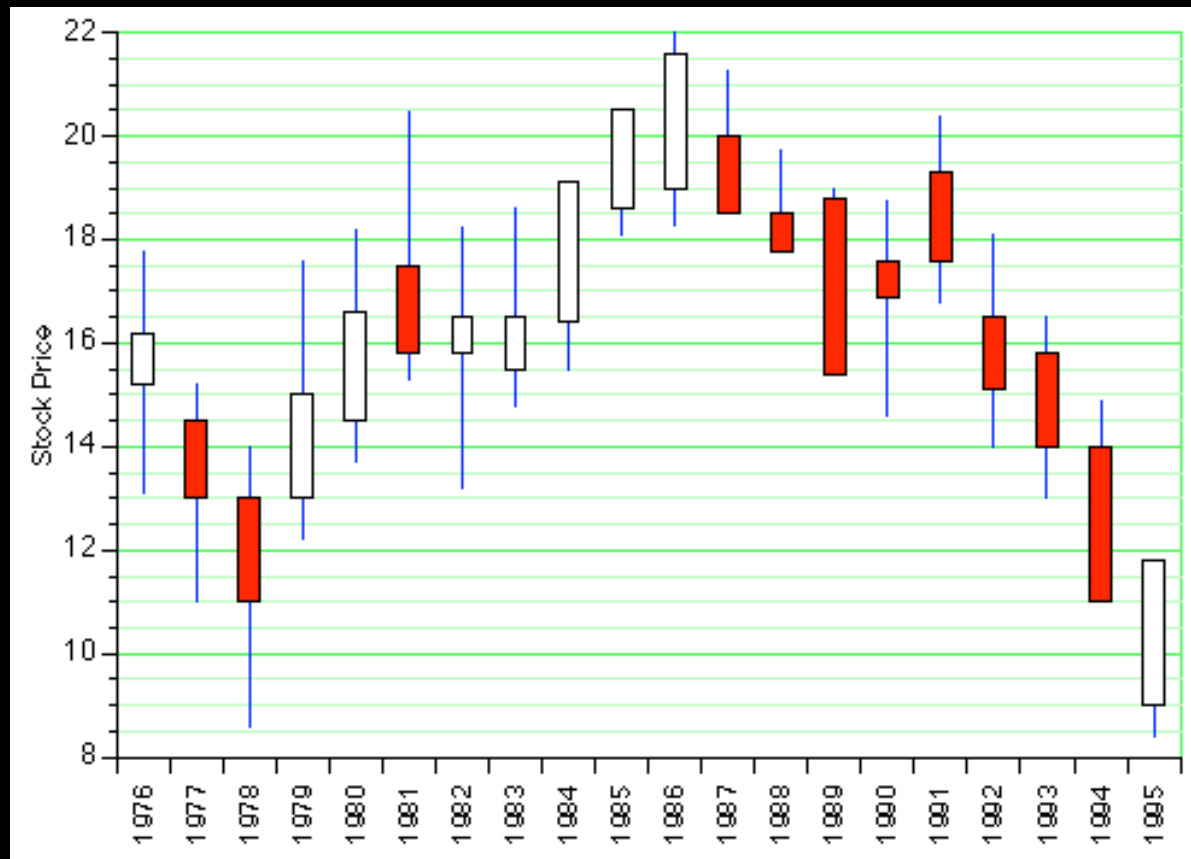


Key



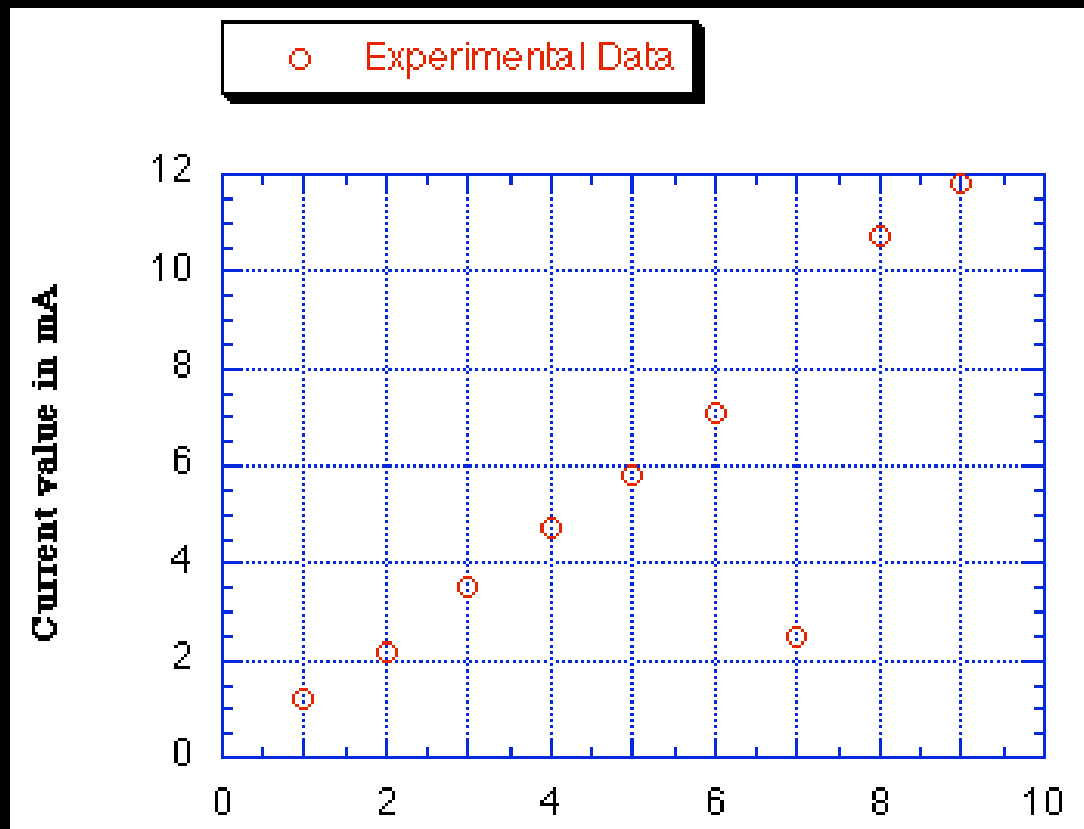
Box and Whisker Chart...

Variable changing over time



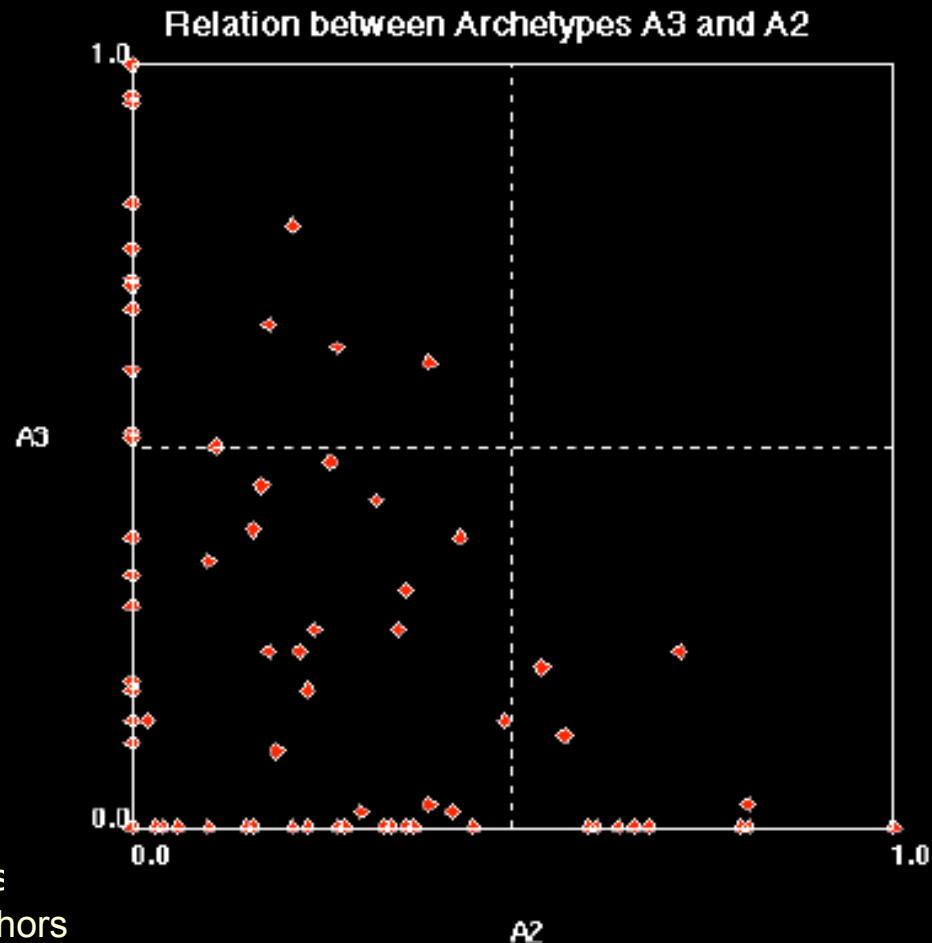
Scatter Plots

1D Scatter Plot - indicates values at test points.



Scatter Plots

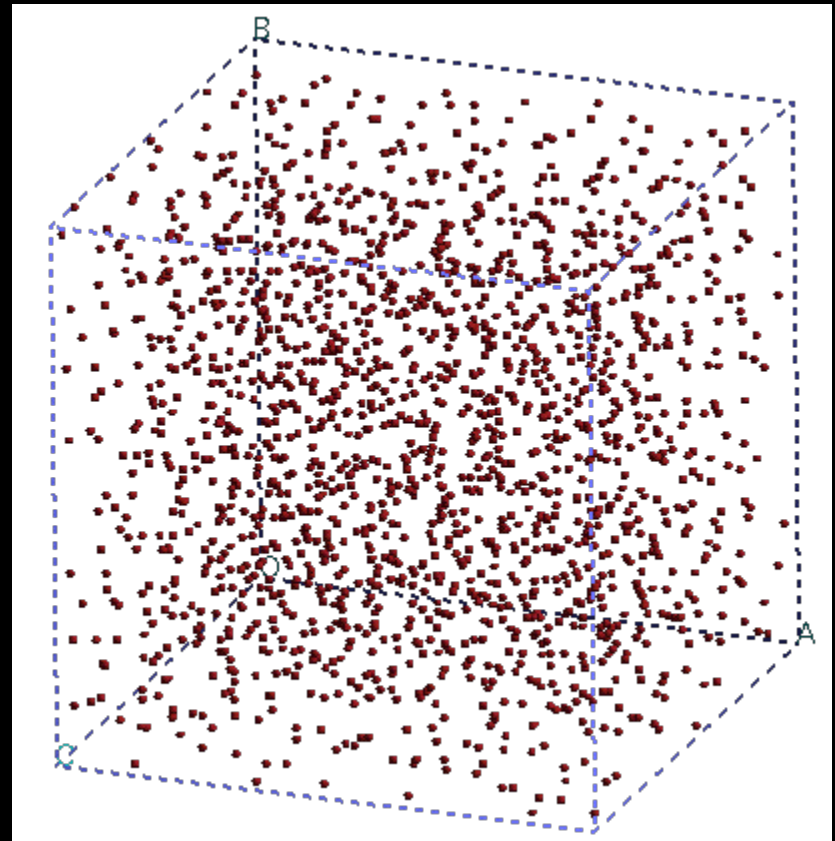
2D scatter plot - relationship between variables



Scatter Plots

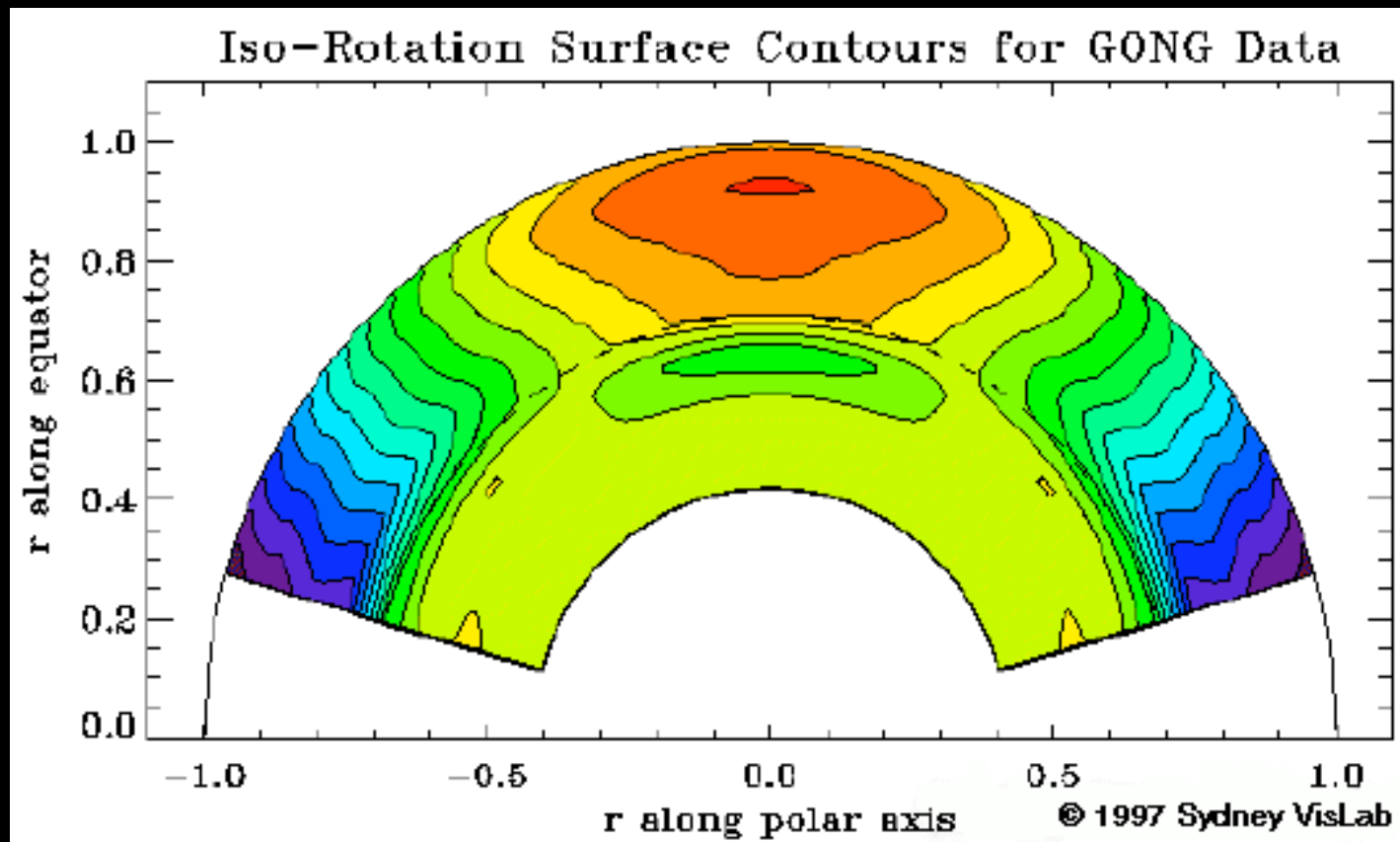
3D scatter plot - relationship between 3 variables

Note the results are unclear on a 2D screen unless the viewpoint can be rotated



Contour Maps

Contour maps contain lines of Iso-*








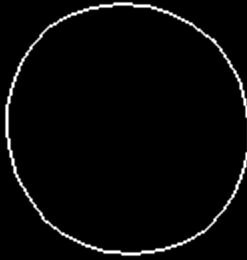





Glyph Charts

- Allows many variables to be presented in a clear manner.
- Each variable is assigned to a glyph or glyph element. The shape/colour/position of that element corresponds to the data values.

Glyph Charts

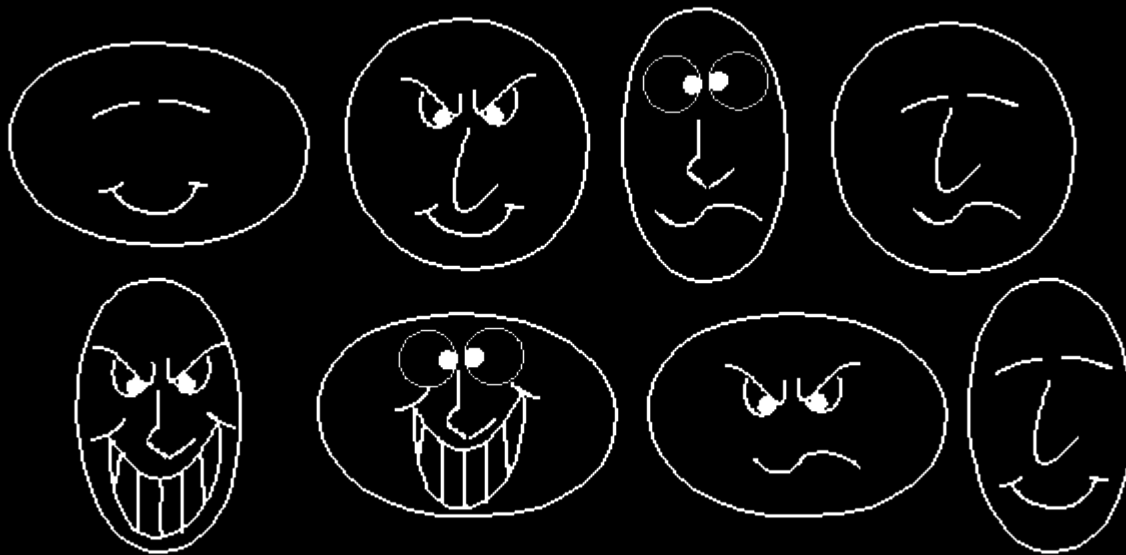
Example glyph symbols

Values of each attribute			
	0	1	2
m=1			
m=2			
m=3			
m=4			

4 Attributes

Glyph Charts

Adding the glyphs (attributes) together...



Graphing Packages

- Graphing Packages
 - CricketGraph PC/Mac
 - Kaleidagraph PC/Mac
 - Deltagraph PC/Mac
 - Xgraph Unix
 - GNUplot Unix
- Packages with graphing ability
 - Excel PC/Mac
 - Mathematica PC/Max/Unix
 - Matlab PC/Mac/Unix
 - AVS (PC)/Unix



Great Graphing Crimes

Not quite Australia's most wanted but...

- Presentation or manipulation
- The art of persuasion
- What you leave out is just as important
- Three types of lies



Rules for Displaying Data Badly

How *NOT* to do it.



Graphing Crimes...

Rule 1 - Show as little data as possible

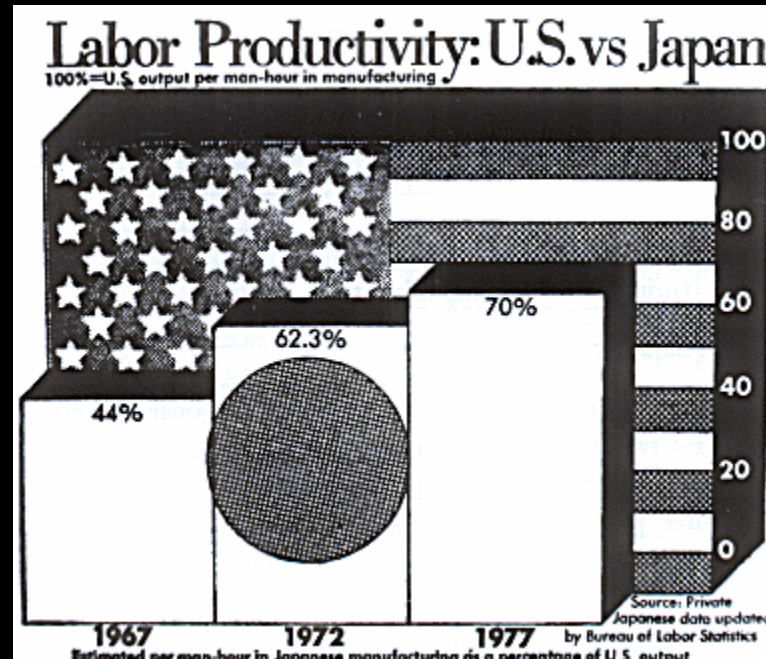
Tufte [*The Visual Display of Quantitative Information*, Graphics Press]
developed several measures of the information content in a display.

- **Data density index (ddi)** - the “number of numbers per square inch”
Measured graph ddi's have ranged from around 0.1 to over 300.
On the whole a larger ddi is preferable, as long as clarity is maintained.
- **Data / ink ratio** - "amount of ink to graph the data divided by the total amount of ink in the graph"
The lower the ratio, the worse the graph!
Often caused by “chart junk”



Graphing Crimes...

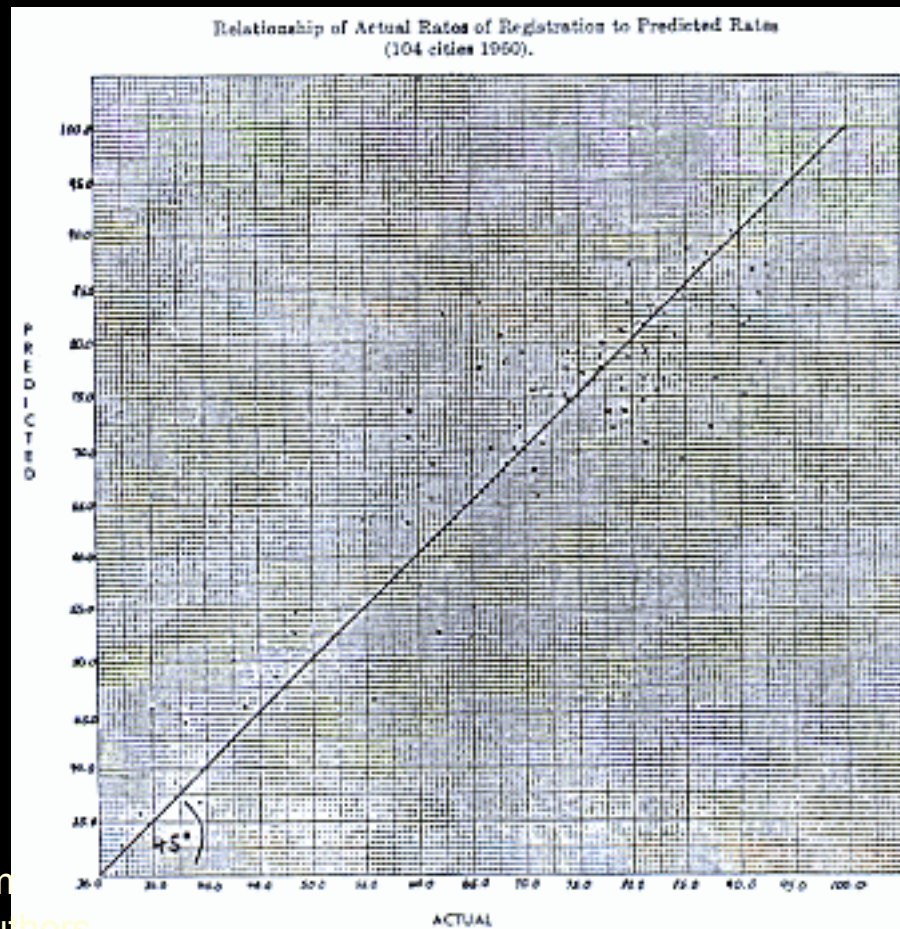
This graph is a bad example - it has a low ddi of 0.3 and contains chart junk.





Graphing Crimes...

Rule 2 - Hide the data that is shown



Lovely grid

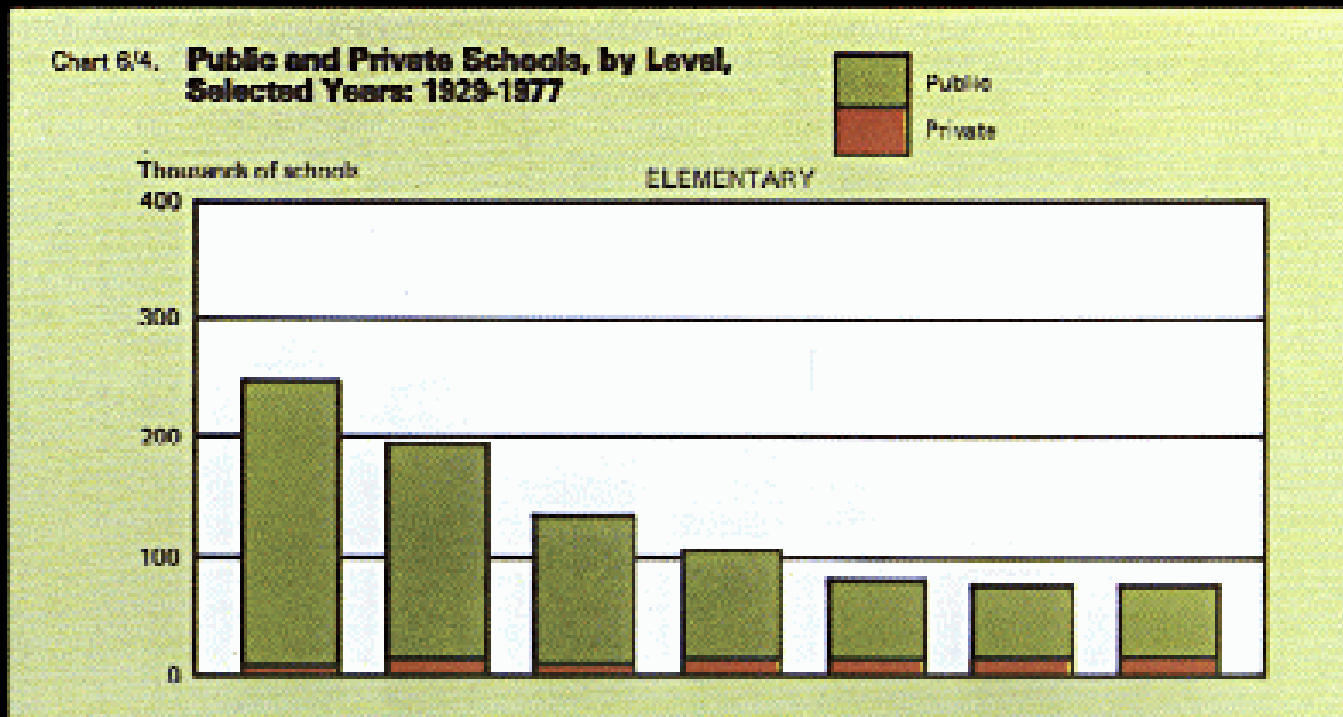
Pity about the data



Graphing Crimes...

Rule 2 - Hide the data that is shown

Private schools failing to expand !? No, they've doubled!



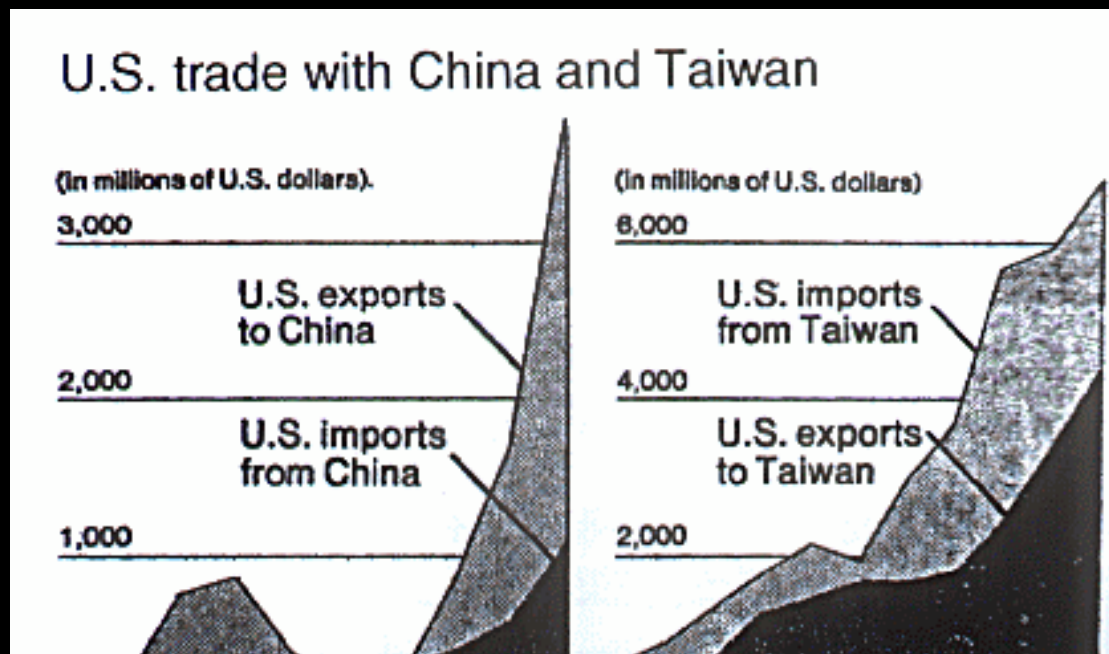


Graphing Crimes...

Rule 3 - Ignore the visual metaphor

Eg: Variable relationship between length and value in bar charts

Below: Different scales and shading vs direction changed between graphs





Graphing Crimes...

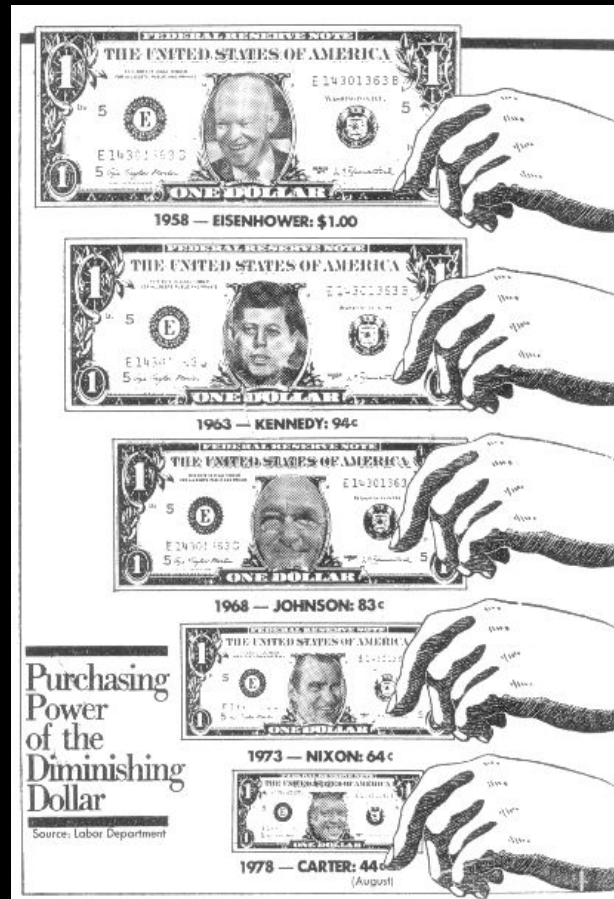
Rule 4 - Inappropriate visual metaphor

Eg: Using area when length is appropriate

Buying power of US dollar between 1958 and 1978.

\$1

\$0.44



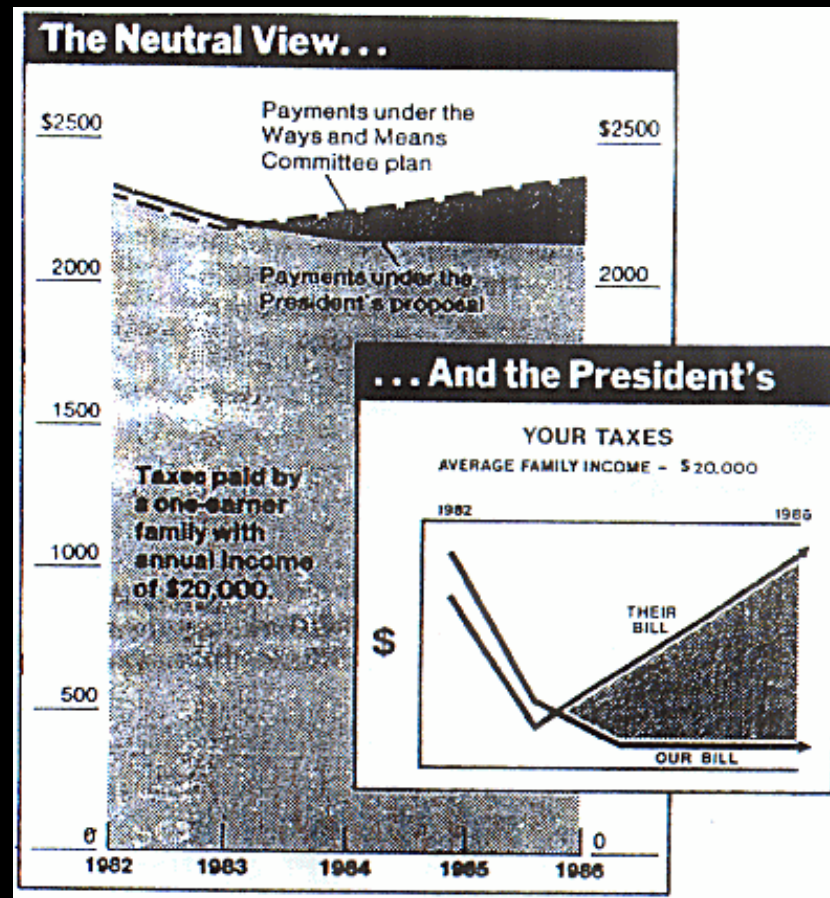
Using area makes the perceived difference much greater than it is.



Graphing Crimes...

Rule 5 - Display data out of context

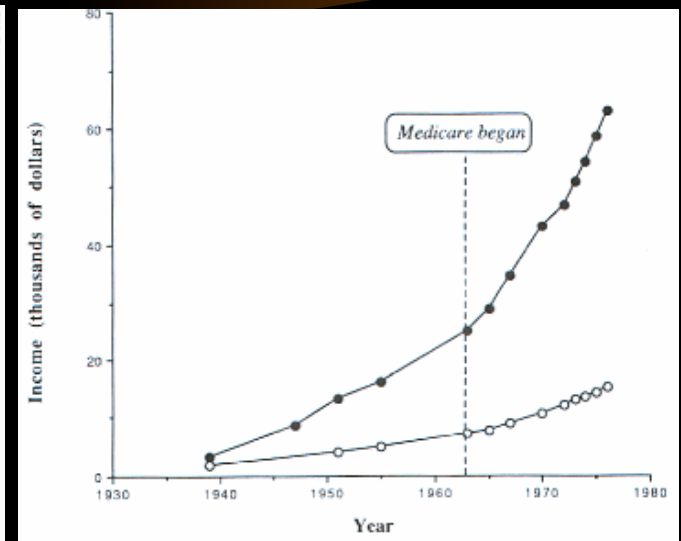
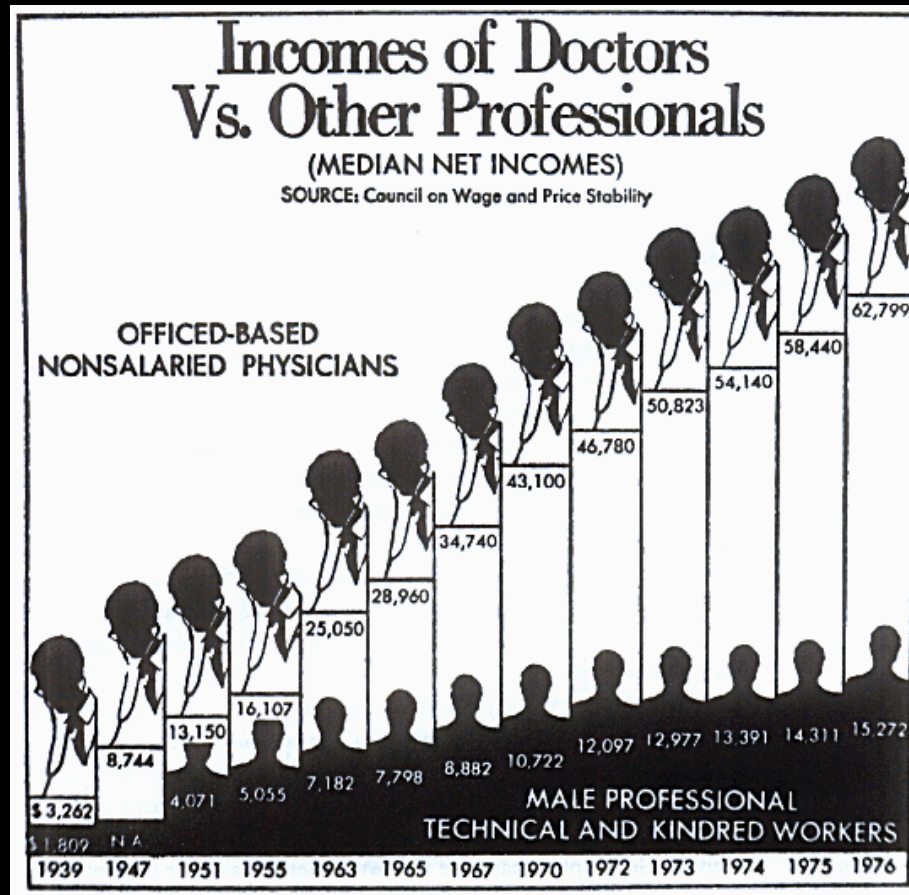
Proportions are
artificially
exaggerated





Graphing Crimes...

Rule 6 - Change scales in mid-axis



Left: Doctors claim proportional wage rises.

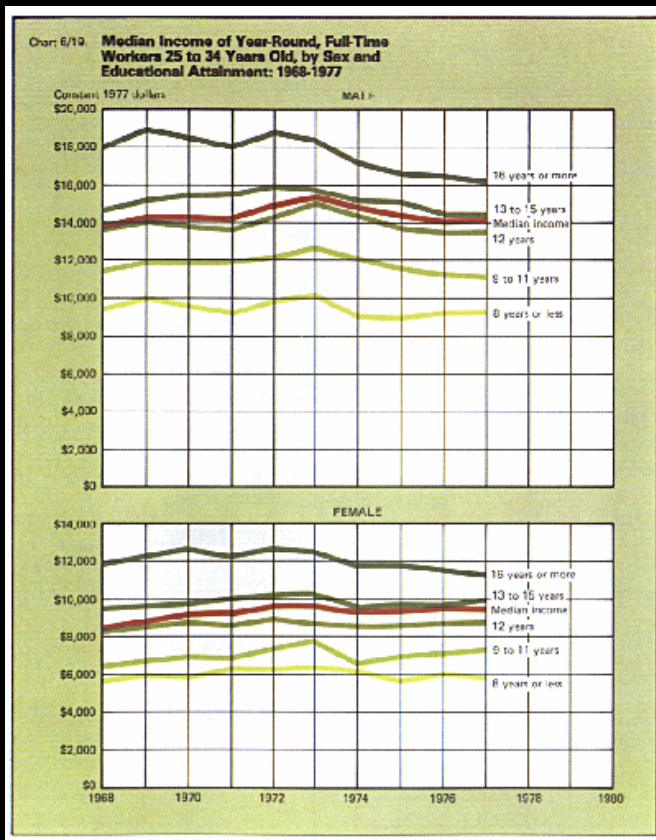
Above: The true story



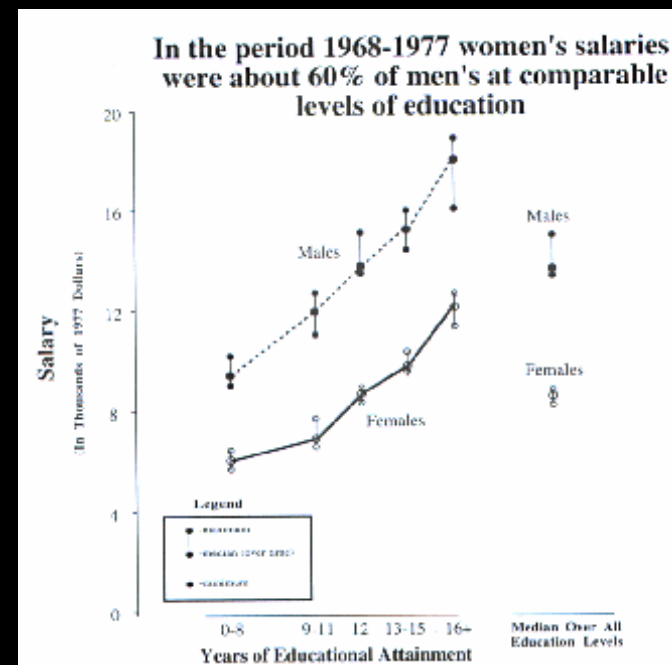
Graphing Crimes...

Rule 7 - Emphasize the trivial

Bad



Better



What's the real point ?
Wage variation with age
or male vs female wages ?

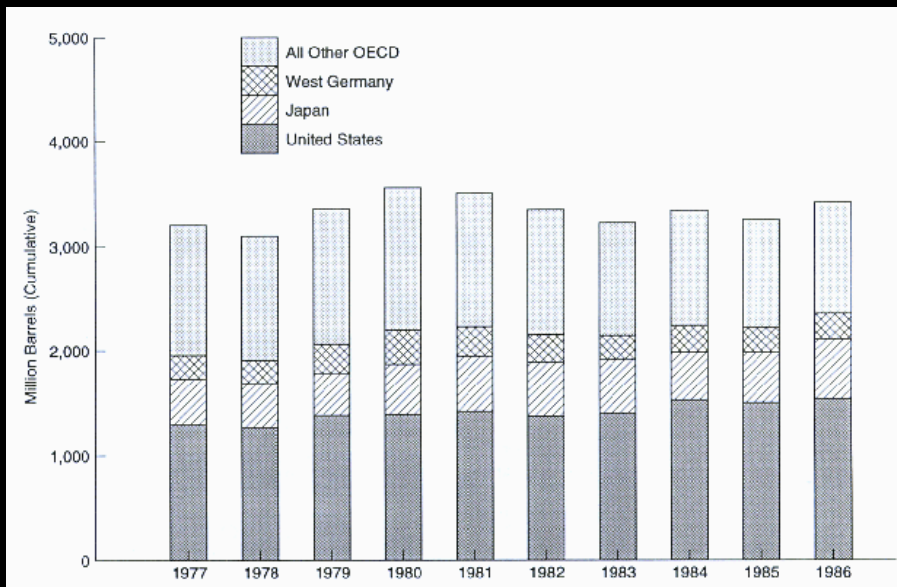


Graphing Crimes...

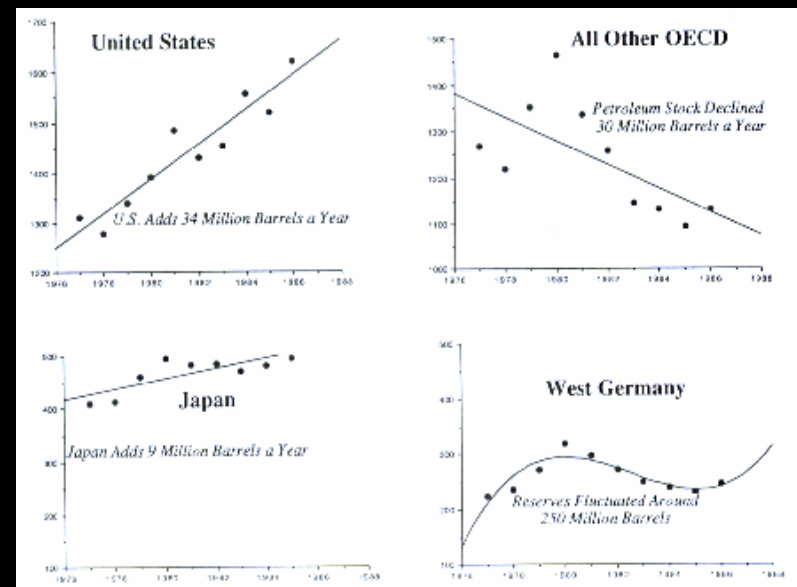
Rule 8 - Vary the baseline

How do Other OECD countries compare ?

Bad



Better





Graphing Crimes...

Rule 9 - Arrange data to obscure the point

Eg: Ordering bars in bar graphs alphabetically instead of numerically



Graphing Crimes...

Rule 10 - Label illegibly, incompletely, incorrectly, ambiguously

- Careless labeling
- Lying by omission or ambiguity



Graphing Crimes...

Rule 11 - Add unnecessary complexity

- Adding unnecessary digits visually complicates tables
- Creating 3D bar charts when 2D will suffice can add confusion



Graphing Crimes...

Rule 12 - Change previously successful formats

- If a good method was used in the past - change it!



Graphing Crimes

Rules for displaying data badly

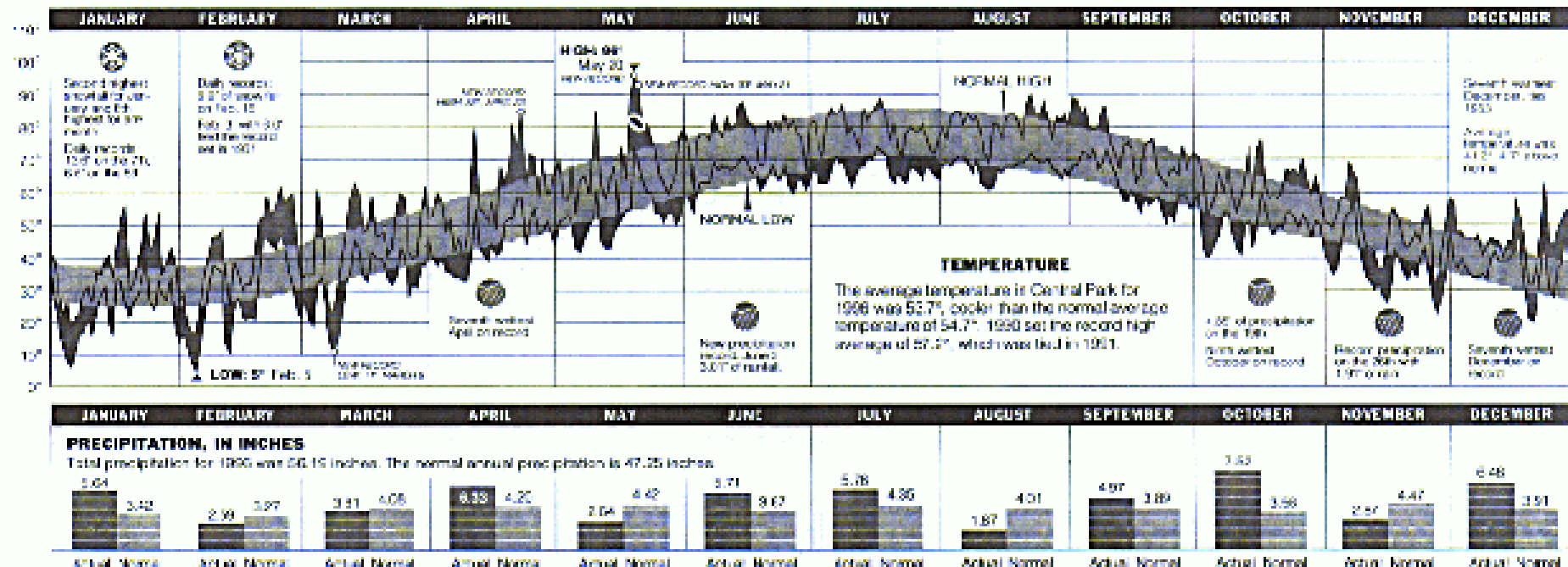
- 1) Show as little data as possible
- 2) Hide the data that is shown
- 3) Ignore the visual metaphor
- 4) Inappropriate visual metaphor
- 5) Display data out of context
- 6) Change scales in mid-axis
- 7) Emphasize the trivial
- 8) Vary the baseline
- 9) Arrange data to obscure the point
- 10) Label illegibly, incompletely, incorrectly, ambiguously
- 11) Add unnecessary complexity
- 12) Change previously successful formats



Good Graphing Examples

This graph is informative - it has a relatively high ddi of 54

New York City's Weather for 1996

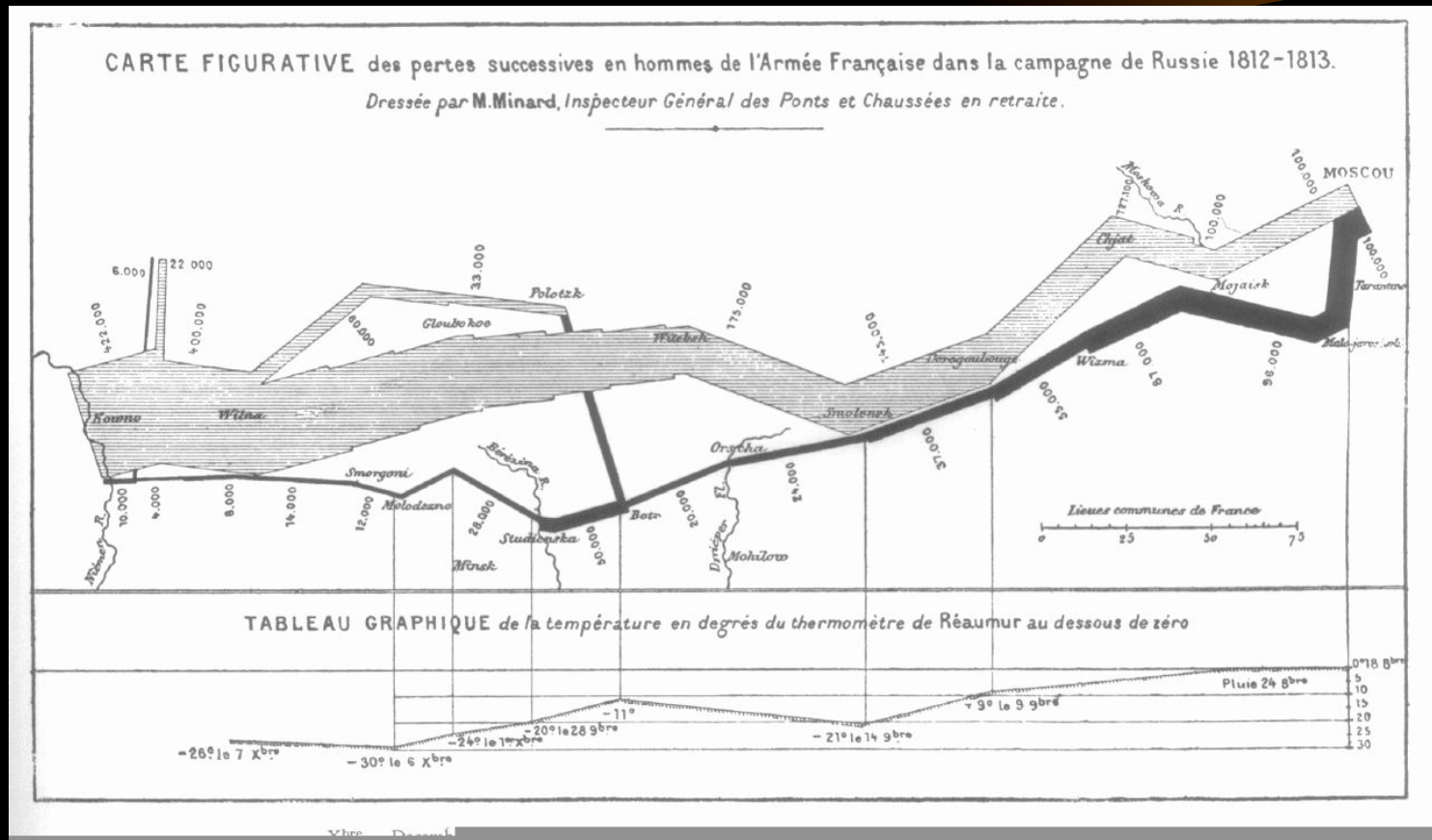


Source: National Weather Service



Good Graphing Examples

Minard's 1861 graph of Napoleon's march into Russia. 6 Variables shown.



End

Lecture 2