Good Graphics?

- Content, context, construction (and design)
  - information is content
  - interpretation needs context
  - good graphics are well constructed
  - design can make good graphics better
- Consistency
  - of scales
  - of formats
  - of style

From the R Graph Gallery

- addictedtor.free.fr/graphiques/RGraphGallery.php?graph= 39, 49, 54, 55, 81, 82, 93, 137

- What is the message?
- What was the aim?
- Is there other information to be found?
- Are other data relevant?
- What might be improved?
- What alternatives are there?

(Some of) Tufte's Principles

- Above all else show the data
- Graphical excellence is nearly always multivariate
- The lie factor should be 1
  - lie factor = (size of effect shown)/ (actual size of effect)
- Maximise the data-ink ratio
  - data-ink ratio = data-ink/total ink
  - erase non-data-ink
  - erase redundant data-ink

Cleveland's Comparison Ranks

Rank	*Code*
1	Positions along a common scale
2	Positions along identical, nonaligned scales
3	Lengths
4	Angles
4–10	Slopes
6	Areas
7	Volumes
8	Densities
9	Colour saturations
10	Colour hues

How are data represented?

- Values by
  - position
  - length
  - area, volume
- Counts by
  - length
  - area
- Grouping by
  - colour, shading, shape

Scales

- Continuous
  - Use standard values (e.g. powers of 10)
  - Use meaningful scales for the variables in question
  - Choose origin carefully
  - Include minimum and maximum (and a little more)
  - Use common scales for comparison and consistency
- Categorical
  - Consider different orderings of categories
  - Consider sorting categories by various criteria
  - Consider combining categories

Text

- Text is used with graphics in several different ways
  - Title
  - Caption
  - Axis labelling
  - Legend
  - Annotation
  - Accompanying narrative
- All should be consistent in what they say
- The text should not obscure the graphic

Formatting

- Size of plots
- Aspect ratio
- Axes
- Tickmarks
- Gridlines
- Points/symbols/glyphs
- Colour
  - colour scales
  - shading

Statistical models / guidelines

- Gridlines at specific values (e.g. y=0, x=100)
- Boundary lines (e.g. y=x)
- Outliers / extreme outliers on boxplots
- Density estimates on histograms, scatterplots
- Regression lines, smoothers on scatterplots (+ intervals)
- Confidence intervals on highlighted proportions
- Residuals from models on mosaicplots
- …

InterOcular Trauma Test (IOTT) on Phelps dominance!

August 10th, 2008 abhi Go to comments Leave a comment

A picture is worth a thousand words!

From NYTimes