high quality graphical displays

lorenz gygax

October 17, 2001

The following notes are based on Edward R. Tufte (1999) The visual display of quantitative data. Graphics Press, Cheshire. 17th printing.

aims

- statistics is not boring and tedious, readers are sophisticated
- to communicate complex ideas with clarity, precision, and efficiency. Multivariate is more interesting. A question of substance matter, statistics, design.
- make large data sets coherent, put it in context, reveal the data at several levels of detail, serve a clear purpose, be closely integrated with the statistical and verbal description of a data set.

criteria of quality

- lie factor $= \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}} = 1$ (avoid distortions)
- maximise data-ink ratio $= \frac{\text{data-ink}}{\text{total ink used to print the graphic}}$ (within reason)
- maximise data density of a graphic $= \frac{\text{number of entries in data matrix}}{\text{area of data graphic}}$ (within reason)
- the number of information carrying dimensions should not exceed the number of dimensions in the data (e. g. cake-diagrams)
- forgo chartjunk including hatching (or equally spaced bars and lines), the grid (if necessary light it up) and the duck (fake perspective)

further hints

- double functioning axis labels: detailed information on distribution
- clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.
- use redundancy if it makes a pattern more visible or comparisons more simple (small multiples: multivariate, narrative, high-density)
- in time series (of money), deflated and standardised units of (monetary) measurements are nearly always better than nominal units
- use graphs wider than high (about 1.5:1), calm shading if horizon is suggested, consider cause (X) and effect (Y)
- use where adequate sentences, text-tables, tables, semi-graphics and graphics
- friendly data graphics: spell out words in clear type (e. g. upper-and-lowercase with serifs), words run from left to right, little messages explain the data, without elaborate shading or encoding, use colours that can be read by the colour-blind, place labels on the graphic itself, provoke curiosity.