The Visual Display of Quantitative Data

Edward R. Tufte

Overall Keys to Success

- **Show the data**
- Induce viewer to think about substance, not form
- Avoid distortion
- Place many numbers in a small space
- Encourage the eye to make comparisons
- Have several layers of data
- Integrate statistical and verbal descriptions

Principles of Graphical Excellence

- Well-designed presentation of interesting data (substance, statistics, design)
- Complex ideas communicated with clarity, precision, efficiency
- Greatest number of ideas in the shortest time with the least ink in the smallest space
- Nearly always multivariate
- Tells the truth about the data

Graphical Integrity

- Representation should be physically proportional to data represented
- Thorough labeling w/ write-ups
- Should show data variation, not design variation
- Information-carrying dimensions should be less than or equal to dimensions of data
- Don't quote data out of context

The Lie Factor (LF):

$$LF = \frac{Size\ of\ effect\ shown\ in\ graph}{Size\ of\ effect\ in\ data} \tag{1}$$

Sources of Integrity and Sophistication

"No one can write decently who is distrustful of the reader's intelligence, or whose attitude is patronizing."

E.B. White

Data Ink and Graphical Redesign

"... the capital importance of criticism in the work of creation itself. Probably, indeed, the larger part of the labor of an author composing their work is critical labor; the labor of sifting, combining, constructing, expunging, correcting, testing: this frightful toil is as much critical as creative."
T.S. Eliot (in "The Function of Criticism")

- Above all, show the data
- Maximize the data-ink ratio
- Erase non-data ink
- Erase redundant data-ink
- Revise and edit

The Data-ink Ratio (DiR):

$$DiR = \frac{Data\ Ink}{Total\ Ink\ for\ Graphic} \tag{2}$$

Data Density and Small Multiples

"For non-data-ink, less is more. For data-ink, less is a bore."
Edward Tufte

Well-designed small multiples are:

- Inevitably comparitive
- Shrunken, high-density graphics
- Almost all data ink
- Deftly multivariate
- (Usually) based on a large matrix
- Efficient in interpretation

Data Density (DD):

$$DD = \frac{Number\ of\ entries\ in\ matrix}{Area\ of\ data\ graphic} \tag{3}$$

Miscellaneous

"Seeing is forgetting the name of the thing one sees."
Paul Valéry

- Forgo chartjunk, including moire vibration, the background (heavy) grid, and the Duck
- Various ideas for removing unnecessary axes, etc.
- Mobilize every graphical element (perhaps > 1 time) to show the data