

Graphical design

Carl J. Schwarz
Simon Fraser University, 2006

excerpt from course materials online—reproduced with permission—
<http://www.math.sfu.ca/~cschwarz/Stat-301/Handouts/node1.html>

Note: A. Caelleigh made minor corrections to text (such as typos).

There are three kinds of lies, “lies, damned lies, and statistics”. This is never more true than when poor statistical graphs are drawn.

A picture is worth a thousand words, and a graph is worth a thousand numbers only if it is clear, concise, and correct.

Always graph your data—often a properly chosen graph will obviate the need for any further analysis. Don't lose sight of the purpose of the graph when you are drawing it.

Principles of good graphical design

Basic Principles. Some basic principles to follow when constructing graphs are:

- **High data-to-chart ration.** There should be a high data to chart ratio. This means that the data points should be clearly visible, form the heart of the graph, and should not be overwhelmed by axes, titles, reference lines, and chart junk. Grid lines should not be too dark, missing, or not relevant to the graph. Grid lines, if present, are best presented in a light grey screen so that they are visible, but not obtrusive.
- **Use appropriate graph.** Use the appropriate graph for the appropriate purpose. Most of the many graphs presented in Excel are POOR CHOICES! In particular, never use a pie chart!

Basic types of graphs

- **Trend graphs.** If you wish to emphasize the trend in a time series, a line chart (i.e. use a line to connect the data points to show the trend) is better than a series of side-by-side bars.
- **Relative size graphs.** Here side-by-side bar graphs are best, but all bars must be anchored at zero. All bars should be equal width, otherwise, readers of the graph will be confused by differences in area, rather than difference in lengths of the bars.
- **Composition graphs.** This is where pie-charts are often (badly) used. The trouble with pie-charts is that people are not well programmed to compare angles of pies. A better graph is a segmented bar-chart where the bar (that stretches from 0 to 100%) is segmented into pieces. Put the most important segments at the top or the bottom of the bar (so that they are anchored at 0% or 100%)—this enables most readers to accurately estimate the percentage of the bar used by the category.

- Make sure that the graph is complete. All axes must be labeled. There should be a title on the graph.
- Think about the overall presentation of the graph. The points on a plot should be spread over the area of the graph without being shoved into one corner. The axes scales should be appropriate. In some cases, a log-scale is a better representation of data that spans several orders of magnitude. Where is the 0 point on a graph? In particular, bar charts should always be anchored at zero. Use different plotting symbols or line-types to differentiate among groups on the graph. The independent variables is usually plotted on the X-axis; the dependent variable usually on the Y-axis.

The best graph is one that is self-explanatory!

Common Graphical Errors

There are many common errors that are made in poor graphs. Here are some of the most common errors:

- **Wrong graph type.** Think about what you want to present. Trends are best displayed using lines. Compositions best displayed using segmented-bar-charts.
- **Missing text.** All tick-marks and axes must be labeled. The graph needs a title.
- **Inconsistent scale.** The scale must be constant across the graph; don't change the increments between tick marks. Most people read increasing scales from left to right and from bottom to top. Comparative graphs must be plotted on the same axes to facilitate comparisons.
- **Misplaced zero point.** Most people assume that the zero point is at the bottom of the graph. This can give a very misleading impression of the amount of change present in a data series.
- **Poor chart effects.** Shading, 3-D effects, or ducks are often added to liven up a graph. In most cases they are useless since they distort the graph and add little new information to the story. 3-D effects are particularly poor as no information is being added; it is difficult to read the chart values; and often the graph is also tilted to make it even harder to read the graph.
- **Confusing of area and length.** If you make a picture twice as large, it looks as if it has four times the area!
- **No adjustment for inflation.** Dollar amounts must be adjusted for inflation. Otherwise, any comparison is misleading.
- **Too much precision.** We've all seen graphs reporting that the amount of money raised is \$13,456,234.32. Most people can't distinguish objects at a resolution better than one part in a hundred. Consequently, giving 10 significant digits is just silly. It would be far better to present this number as simply \$13 million (i.e., get rid of all the extra zeroes and use an appropriate scale).

Note: For this handout, the format was changed in minor ways to clarify the text.