PRESENTING YOUR DATA GRAPHICALLY

OVERVIEW

This paper outlines best practices for the graphical presentation of data. It begins with a brief discussion of various aspects of giving a presentation. It then discusses how to decide which type of graphic to use. This decision is influenced by the type of data being shown and the purpose behind the presentation of the data. The paper then discusses design characteristics for graphics, including choices pertaining to font, labeling, axes, line design and the use of color. The paper closes with a bibliography, which includes several websites which include demonstrations of both good and bad graphic designs.

INTRODUCTION

Presentations generally contain three elements: a visual presentation (often done in PowerPoint), a handout and the verbal discussion that takes place during the presentation. Not all presentation situations are the same, and therefore not all presentations should be of the same style. Many presentations, unfortunately, do follow the same design. These presentations consist of a PowerPoint presentation, a handout, which is simply a printout of the PowerPoint presentation, and the discussion, in which the facilitator mostly reads the PowerPoint slides. This style of presentation has led many commentators to criticize PowerPoint. But PowerPoint is not to blame. The problem with this presentation style is in its design, not in the tool (in this case, PowerPoint) itself.

Generally speaking, the handouts for a presentation should be distinct from the visual presentation, and the verbal discussion is an additional distinct element. The visual presentation is an excellent place to showcase charts that highlight patterns and relationships. The handouts are a better place for tables of data, often the underlying data for the charts in the visual presentation. The visual presentation is an appropriate place for powerful visual images without much text. The text should be included in the discussion. A visual presentation does not substitute for the verbal discussion. Any particularly salient points may be reinforced in the handout, along with any supporting details that the audience may want to have available following the presentation.

GRAPHIC TYPE

The choice of graphic type is generally driven by two factors: (1) the type of data to be presented and the relationship to be demonstrated, and (2) the human mind's ability to perceive, interpret and analyze the visual representation of the data that it is seeing. Bar charts, line charts, scatter plots, box plots, tables, radar graphs, pie charts and stacked bar graphs are all commonly used graphics.

Perception

The human mind is very good at perceiving differences in vertical heights. It is less able to distinguish between varying widths. It is very poor at perceiving two-dimensional representations of three-dimensional objects. It is also very poor at judging the distance along a curve.

Bar charts

Bar charts enable the comparison of individual data values. These help the audience to focus on subsets within the bigger picture. These are most appropriate when comparing discrete categories across continuous data.

Good bar chart design includes the following elements:

- Set the vertical axis to zero at the origin.
    - A bar that is twice as high as another bar should represent a value that is twice as large. Using a non-zero-based axis can lead to exaggerated differences in the data and confuse the audience.
    - If a non-zero axis must be used, the vertical axis should be broken with a jagged line or two angled slashes to alert the audience to the use of a non-zero axis.
    - If a non-zero axis is used, it may be best to show the chart first with a zero axis and then present the "zoomed in" version with the non-zero axis.
- Limit the use of bar charts to a few bars or sets of bars.
    - A bar chart with too many bars becomes overwhelming to the eye.
    - Consider replacing it with multiple bar charts or with one or more line charts.
- Have related bar charts use the same scale.
    - If you are showing related data sets of different bar charts, make sure the vertical axis is the same on both charts to facilitate an accurate comparison between the two charts.

Line charts

Line charts are best for the presentation of trends. The most common line charts plot a continuous variable as a function of time. Line charts should generally not be used for categorical data. Do not overload a line chart with too many lines. A double bar graph may be better than two line charts.

If you are using a line chart to show a growth rate or multiple growth rates, plot the logarithmic values of the underlying data. This allows the mind to easily interpret a straight line as a constant growth rate. Remember that the mind has trouble judging distance along a curve. It is even more daunting to figure out if data is following a smooth exponential growth curve.

Scatter plots

Scatter plots are best used to show correlation between variables.

Tables

Tables should be used when the presenter needs to show precise numbers. When using a table, the categorical axis should be sorted by the numerical axis value, not by the categorical value. For example, when showing a list of countries and their populations, the list should be sorted by population, not alphabetically by country. This also applies to horizontal bar charts.

Pie charts

Pie charts should only be used to show a group of items that combine to represent 100% of a larger set. Even in that case, a pie chart should only be used if there are four or less segments. As a best practice, pie charts should not be used even if this case, as the following type of graphic generally works better.

Stacked bar graphs

A stacked bar graph is the preferred graphic to represent data elements that combine to represent 100% of a larger data set. The same qualifications apply to stacked bar graphs as to pie charts. They should be used for four or less segments, and only for displays of how 100% of a large data element is broken down into smaller segments.

Radar graphs and 3D representations

Rader graphs and 3D representations should never be used. Shading should also never be used. Do not let a desire to make the representation artsy or pretty interfere with communicating the data.

DESIGN CHARACTERISTICS

Good graphical design requires attention to the many finer points that combine to create the overall graphic. As a rule, good design seeks to maximize the data-ink ratio. That is, a good design is one that involves most of the ink being used to show the data, rather than superfluous decorative elements. The following guidelines seek to increase the data-ink ratio and promote good graphic design:

Gridlines

- Minimize the number of gridlines
- Make gridlines thin, black or gray, and unobtrusive
- Never use both a reference line and gridlines
- Avoid having data lie along a gridline

Font

- Use serif fonts, such as Garamond or Times, for printed documents (handouts)
- Use sans serif fonts, such as Arial or Calibri, for on screen presentations
- Avoid decorative specialty fonts

Labels, Legends and Titles

- Give the graphic a title only if it needs one
    - If it is the only graphic on a PowerPoint slide and the slide already has a good headline, then the graphic may not need a title of its own
    - Make sure the title communicates to the reader and is not simply a repetition of information presented elsewhere or easily inferred from the graphic
- Avoid repeating the same information in multiple places
    - For example, do not label the vertical axis as dollars, then label the data points as "$40" and include a title of "Dollars per Year"
- Avoid vertical labels on the horizontal axis
    - Ideally, the elements along the horizontal axis are far enough apart to allow for horizontal labeling
    - Other options include using a horizontal bar chart with the categories on the vertical axis
    - Some authors favor having the labels for the horizontal axis turned 45°, while others argue against this practice, but all agree this is superior to labels that are turned 90° from horizontal
- It is preferable to label the data lines to avoid needing to use a legend
- When a legend is required for categorical data, be sure to order the legend in the same order as the data in the graphic
- Order the data in a consistent and logical order
    - Years should progress from oldest to most recent
    - Data periods need to be consistent
        - Do not mix years and quarters in the same exhibit
        - Do not omit years in the middle of a series

Alignment and Number Formatting

- Left align dates and text
- Right align numbers and numeric column headers
- Put numbers to be compared into a column, not a row
    - Make sure all numbers are formatted similarly, so that the eye naturally perceives that the number which protrudes the farthest to the left is the largest number
    - This is the same rationale that favors putting the numbers into a column, not a row
- Use only a few digits when formatting a number

Axes

- Axis tick marks should be avoided whenever possible
- Axis tick marks that are included should point outward, away from the data

Lines

- Do not use line markers, such as boxes or crosses. Line graphs are generally used to show the overall trend, making the individual points less relevant.
- To highlight a comparison between two or three points on the line chart, use a bar chart to show that comparison.

Color

- Use color only to explain, highlight, or show data – not to make things look lively or pretty
- Avoid background color
- Do not use red for positive financial values
- Do not use multiple colors to represent the same type of data
- Objects in red, orange and yellow appear larger than objects in blue, violet and gray
- Avoid using both blue and green (for color-blind viewers)
- Avoid very high contrasting colors
- Use different colors for different series
- Use color gradations for trend within a series

An in-depth discussion of the use and impact of color on presentation is beyond the scope of this document. For those who are interested in learning more about how to effectively use color in presentations, www.colorbrewer.org is an excellent resource, as is the discussion in Carr and Harrington on pages 3-5.

Attribution

- Be sure to indicate the source of the data
- Include any appropriate qualifications about the data, but avoid undermining the data's credibility

POWERPOINT TIPS

Most visual presentations will involve PowerPoint. When putting a graphic into PowerPoint, be sure to use the headline feature in PowerPoint. Keep in mind that a strong headline guides the reader's thoughts and helps to guide the conversation. In most situations, this is the impact that the presenter desires. There are situations where this is not appropriate. In particular, the use of a strong headline can inhibit dialogue and the challenging of assumptions.

As was discussed above in the context of not overloading a single chart, also be careful not to put too much into a single slide. When using a slide to convey information to your audience, it should clearly deliver its message. Do not include too many points within a single slide.

If the purpose of a slide is to support a dialogue, then a very complicated slide can be appropriate and can support a very lengthy discussion of that single slide.

CLOSING COMMENTS

This document has presented a summary of best practices around the graphical presentation of data. Much of this content is drawn from experts in the field. Their work can be referenced directly by referring to the bibliography.

BIBLIOGRAPHY

Bray, Tim. “The Data-Ink Ratio Example.” *Tbray.org*. Tim Bray, 13 Mar. 2003. Web.

Carr, Rebeca, Mary Harrington. “Effective Communication Through Visual Design: Tables and Charts.” *Aaude.org*. Association of American Universities Data Exchange, 03 Feb, 2011. Web.

“Communicating Your Data Graphically.” *Enfovis.com*. Enfovis, Oct. 2010. Web.

Duke, Susan P. “What Makes A Good Graph?” *Phuse.eu*. PhUSE, n.d. Web.

Hunter, Matt. “Presenting Data in PowerPoint.” *Slideshare.net*. Slideshare.net, 22 June 2012. Web.

Lyons, Ray. “Best Practices in Graphical Data Presentation.” *Libraryassessment.org*. Association of Research Libraries, 25 Oct. 2010. Web.

Smith, Jerry A. “Five Graphical Perception Best Practices Every Data Scientist Should Know.” *Datascientistinsights.com*. Data Scientist Insights, 09 Feb. 2013. Web.

Tufte, Edward R. *The Visual Display of Quantitative Information: Second Edition*. Cheshire, CT: Graphics Press, 2001. Print.