Guidelines for Plots

Tagliatela College of Engineering

Graphical displays of data should be guided by the principles found in the works of Edward Tufte, a noted expert on the display of quantitative information. These are shown in Table 1.

Table 1. Principles of Analytical Design – from *Beautiful Evidence*[1] by Edward Tuft

1. Show comparisons, contrast and differences
2. Show causality, mechanism, explanation, systematic structure
3. Show multivariate data; that is show more than one or two variables
4. Completely integrate words, numbers, images, diagrams
5. Thoroughly describe the evidence. Provide a detailed title, data sources, scales, units, etc.
6. Content counts most of all. Analytical presentations ultimately stand or fall depending on the quality, relevance and integrity of the evidence.

A common pitfall in preparing graphical displays is to focus on format, color, etc. rather than on the data. Formatting should help bring out the message the author wants to deliver by showing data. Formatting must never distract from the message. Note that the principles of data display discussed here are general and apply regardless of the software package being used to create the plots.

A. Structure of the Plot:

In engineering work, the most commonly use type of plot is an "XY Plot" or "Scatter Diagram". This type of plot shows the relationship of one or more dependent variables (on the vertical axis) to an independent variable (on the horizontal axis).

Below are typical components of an XY plot.

- *Plot Title:* A brief description that identifies the information in the plot concisely, but is more informative than the titles of the axes. A sub-title may be added to further clarify the situation depicted in the plot.
- *Axis Labels and Numerical Value Labels*: Descriptive text that clearly explains the independent and dependent variables, including units. The numerical values listed on the axis are aligned with the major gridlines. These values should be shown with the minimum number of digits needed to distinguish the values (usually 1 or 2 significant figures).
- *Axis Scale:* The range for each variable should be selected to make full use of the plot area. Major divisions mark the value of the major gridlines and should be selected to make it easy for the viewer to read the numerical value of a point. Major divisions should divide the scale into 4 to 10 units. Selection of the scale and divisions should result in "round" numbers at each major gridline. Minor divisions are optional and are used to provide minor gridlines which enhance the ability to read values at a finer level of detail. There are no numbers aligned with the minor gridlines. Minor divisions should divide the major divisions so that

[1] Tufte, Edward R., *Beautiful Evidence*, Graphics Press LLC, Cheshire, CT, 2006; www.edwardtufte.com

the viewer can mentally assign numerical values. Usually the "missing" numbers should be exact values that require only 1 additional digit compared to the visible major divisions.

- *Legend:* If more than one set of data is shown, a legend is needed to explain the differences. The shape of plot symbols, type of line or color are used to distinguish among the data sets.
- *Figure Number:* A unique number that identifies each plot in a document. This number is needed if the plot is included in a document discussing the data.

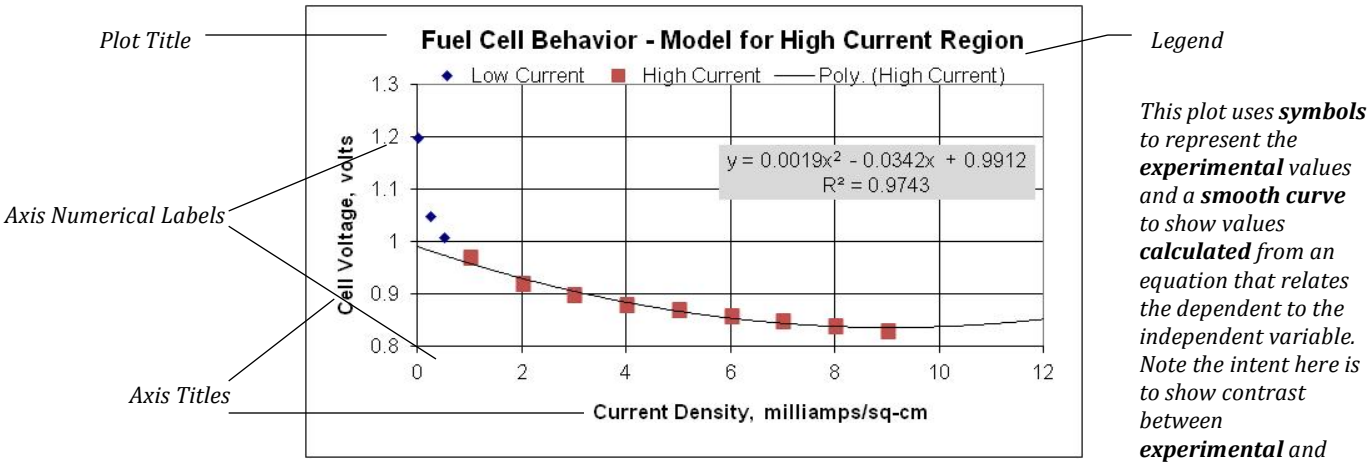


Figure 1. Example XY Plot

B. Typical Components of a Plot

Component	Specification
Figure Number	Figures used in a document should be numbered.
Figure Title	• Characterize the information concisely, augmenting axis titles. • **Do not** simply repeat the axis labels (for example, Plot of Power vs. Time).
Axis Labels	• Indicate clearly what is plotted on each axis, with units. • Select numerical labels for ease of estimating values.
Legends	• Add a legend for multiple data sets to describe each set. • Place legends so they do not obscure data nor waste space. • Change the Excel default placement that reduces the plot size and wastes area that should be used for data display (see Figure 2).
Plot symbols and lines *Experimental* *Calculated*	Add symbols and lines that help readers understand the meaning of the data and, in particular, guide the readers to observe relationships between variables, scatter in the data or contrasts among different data sets. • Show **experimental** values with **symbols** without connecting lines. • Show **calculated** values with a **line or smooth curve** through the plotted values, no symbols.

Component	Specification
Trend lines and predicted values	• A trend line is a plot of **calculated** values from a relationship between the plotted (dependent vs. independent) variables, often determined by regression or theory. A trend line may be helpful to illustrate a pattern in the data or a contrast between theoretical and actual results. • Add trend lines or curves to plots of data only if they serve a specific purpose in the analysis (don't always add a trend line). • Accompany trend lines by an information box showing the equation used and other relevant information.
Predicted values contrasted to experimental	• Show **Experimental** data as symbols (no lines), **calculated** data as curves (no symbols).

C. Issues & Advice in Plot Formatting

Issue	Advice
Numerical Scale	• Adjust the minimum and maximum values to fit the data set. • Make effective use of the plot area for showing the data, avoid having a significant region with not data because of the selected max or minimum. • A minimum of zero is often NOT the best choice, as it may leave a large portion of the plot devoid of data, wasting valuable space. In rare cases this empty space may be intended by the author to make a point, but usually it should be avoided. • Usually place the vertical axis at left edge of plot, horizontal axis at bottom, unless there is a need to place an axis inside the data area (for example, complex plane plot).
Major and Minor Divisions	• Major divisions set major gridline positions and align with numerical labels. • Select major divisions for ease of estimating values of plotted points, but avoid packing too densely. • Select major divisions to generate numerical labels that need only 1 or 2 digits to express the exact value (0.0, 0.2, 0.4, not 0.0, 0.1333, 0.267). • Minor divisions, if used, fill in between the number labels and help the reader assess the values of horizontal and vertical position of plotted data.
Scale Number Format	• Avoid trailing zeroes in the numerical labels (3 not 3.00, 3.2 not 3.20)) • Avoid showing values with many digits – use scientific notation for very large or very small numbers, or show as a "divided" value (for example, Investment in $millions, with value shown as 3.5, rather than $3,500,000).

Issue	Advice
Grid Lines	• Quantitative axes should have gridlines to help readers estimate values. • Major grid lines are needed for both axes on XY Plots. • Excel default shows horizontal axes only; **add vertical gridlines.** • Minor gridlines help estimate values between the major gridlines, but add to clutter. Use minor gridlines *only* as necessary to help readers understand the data.

D. Plot Examples

This plot was created using the same data as the plot example on page 2, but using the default settings in Excel. Note the following:

- Vertical scale starting at zero wastes half the plot area;
- Lack of meaningful title, no axis titles;
- Unnecessary legend (only one data set);
- Legend placement reduces data space considerably;
- Scale numbers vertical — crowded, horizontal — too few;
- Points connected with line segments;
- Missing vertical gridline.

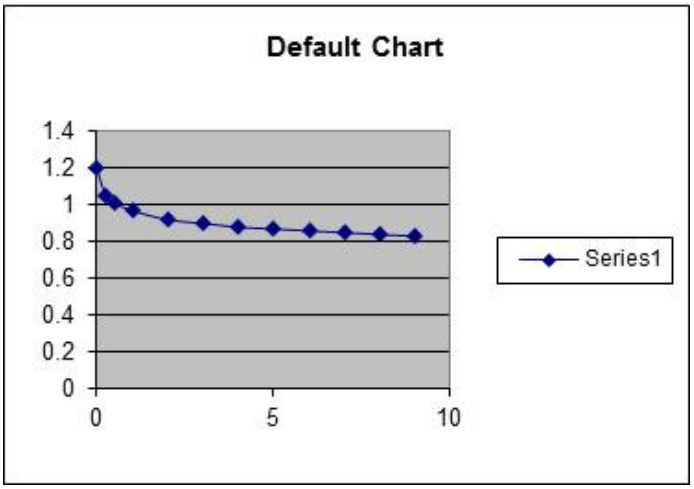
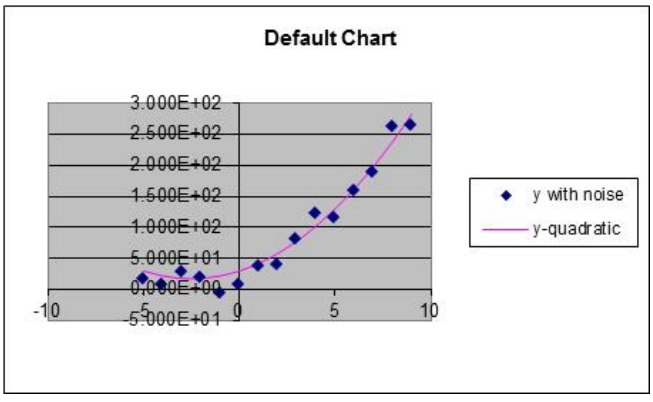


Figure 2. Poor Example of an XY Plot

Plots to left and right use the same data. Note the poor choices in the default plot (left) for:

- Placement of axes.
- Number scale and format.
- Legend placement.
- Main title, missing axis titles.



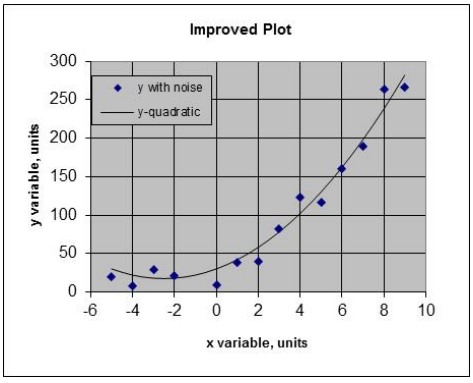


Figure 3. Default Plot Comparison to Improved Plot