# Data Presentation and Interpretation

## Plot the Data

Prof. J. Ramkumar
Department of ME & Design
Indian Institute of Technology Kanpur

# Plotting

- The purpose of plotting scientific data is to visualize variation or show relationships between variables, but not all data sets require a plot.

- If there are only one or two points, it is easy to examine the numbers directly, and little or nothing is gained by putting them on a graph.

- If there is no variation in the data, it is easy enough to see or state the fact without using a graph of any sort.
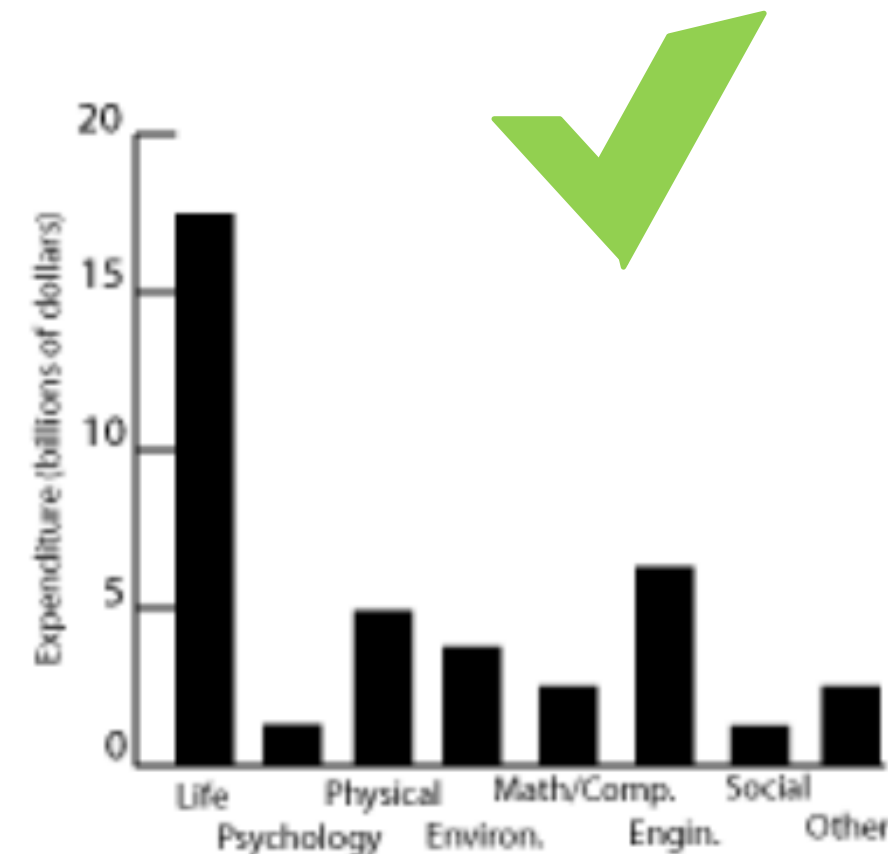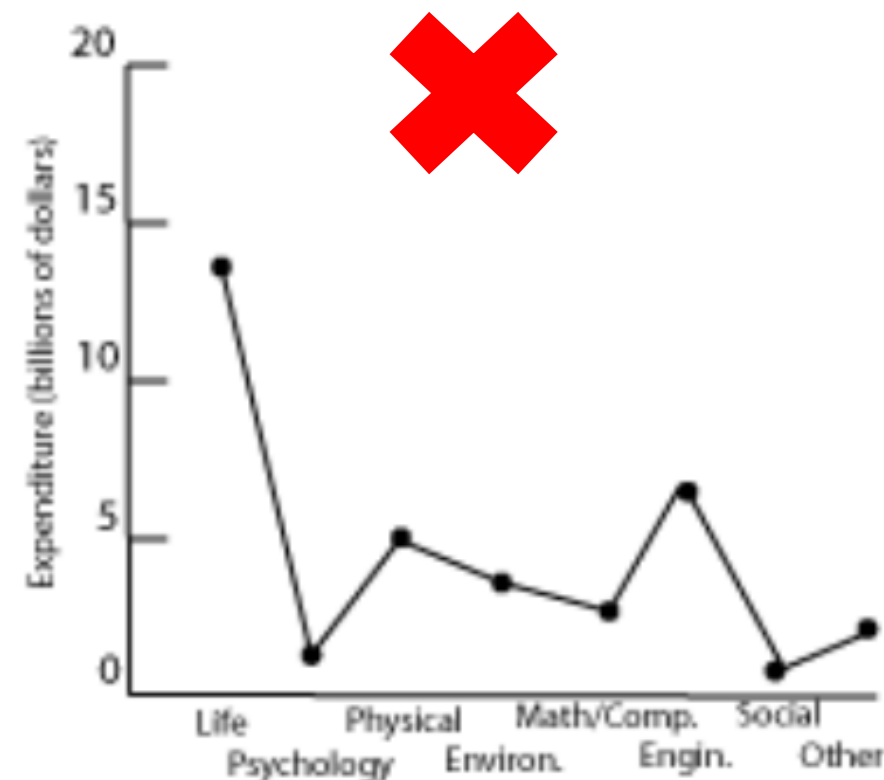


Research expenditures for various scientific fields.

Source: www.ruf.rice.edu

# Mistakes

There are a lot of mistakes in the graph:

1. Missing units – is it total consumption and deaths, or normalized for population?

2. Reversed axes – we suspect that smoking leads to cancer, not the converse. The independent or causal variable goes on the x-axis.

3. The jagged line connecting the points has no basis. The scatter of the data suggests large random effects, not real changes from point to point.

4. A caption that is not particularly helpful.
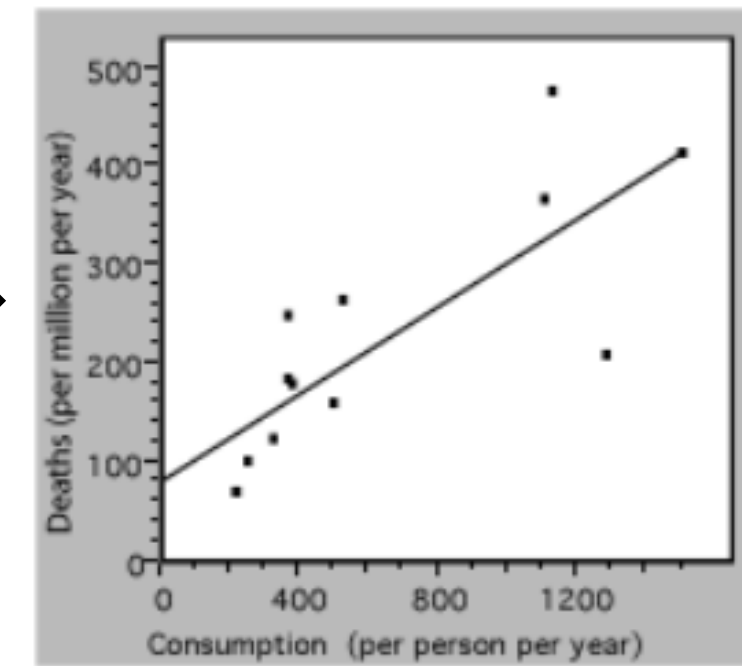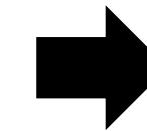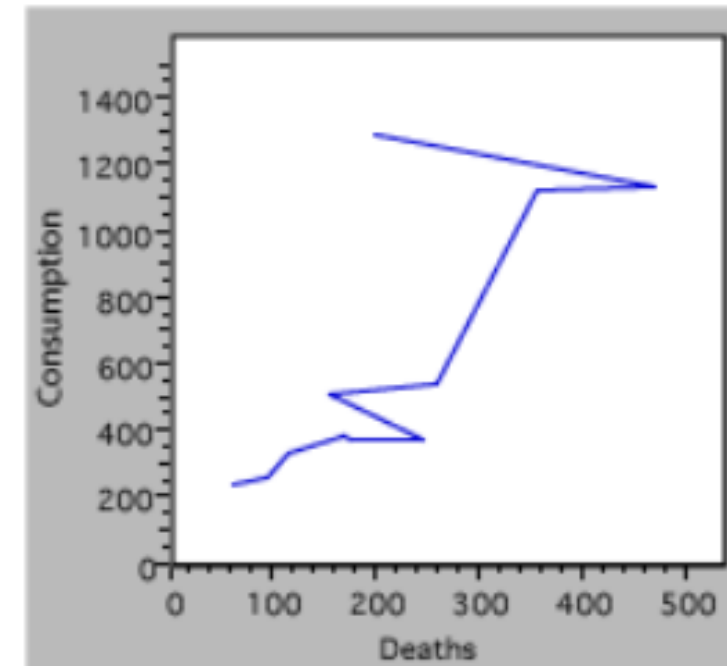


Source: www.ruf.rice.edu

# Misleading Scales

Important points in data plotting:

1.  Axes are not labeled with the quantity measured, nor are units identified.

2.  The axes are very unequal in length, for no visible reason.

3.  The vertical scale has too wide a range to display the range of the data.

4.  The horizontal scale is also too long, extending well beyond the data range.

5.  Grid lines add clutter but not information.

6.  A fitted straight line is shown, but the scales make it hard to tell if it is accurate.

7.  The fit extends far beyond the data, without justification.
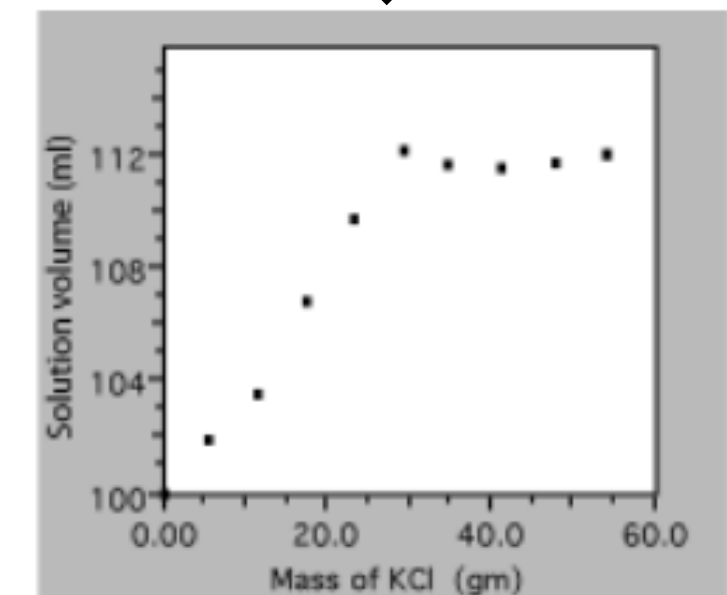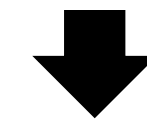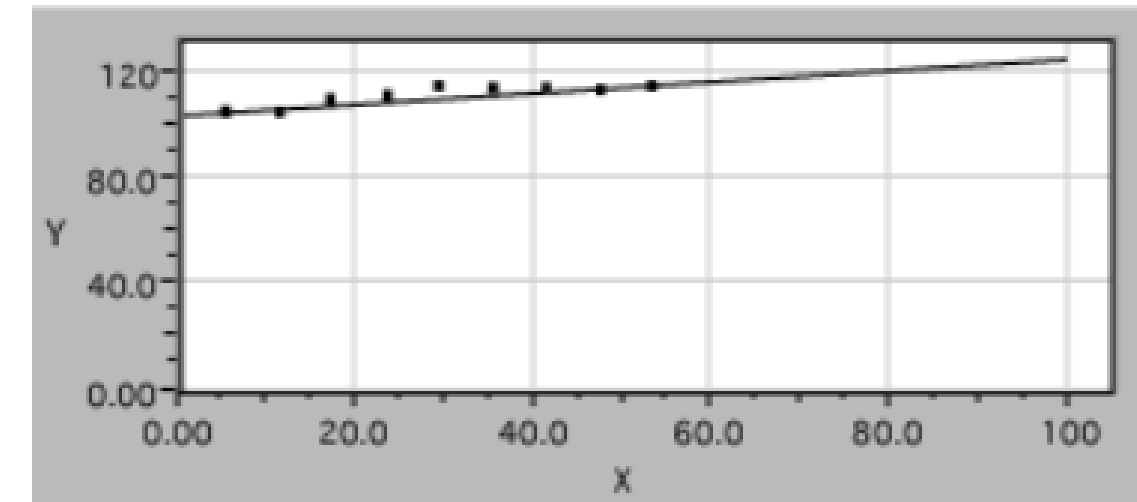
# Misleading Scales
## (Computer fits)

- Experimenters often use computer-generated best-fit lines to demonstrate agreement with some model or theory.

- Using the defaults in a poor fitting program might produce this result.

- One line consists of data, the other theory, but it is hard to follow either one (impossible with a monochrome version).

- Connecting the data dots is also incorrect because it implies that there are more data than actually present.



Source: www.ruf.rice.edu
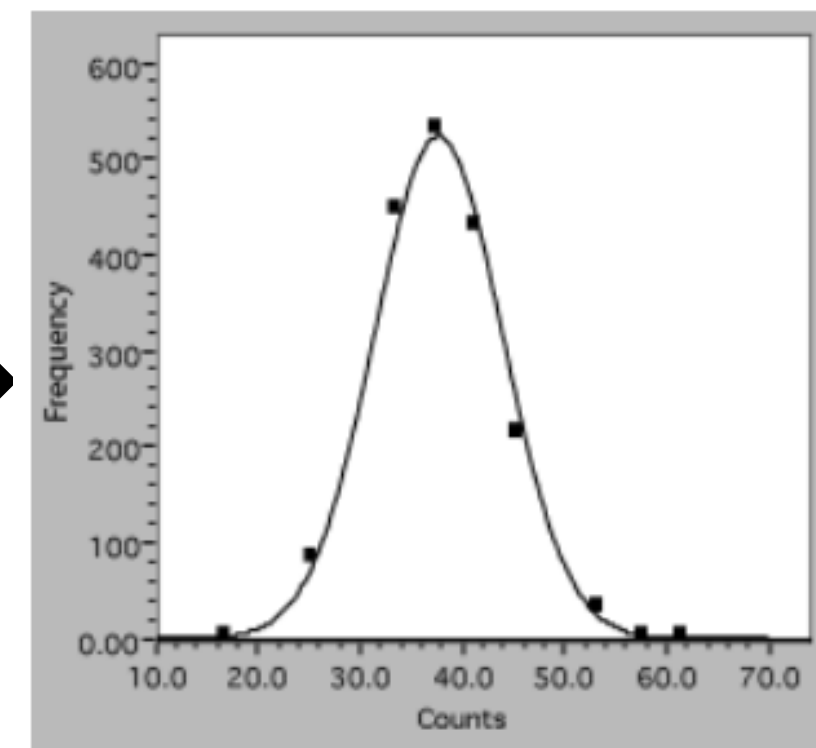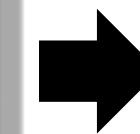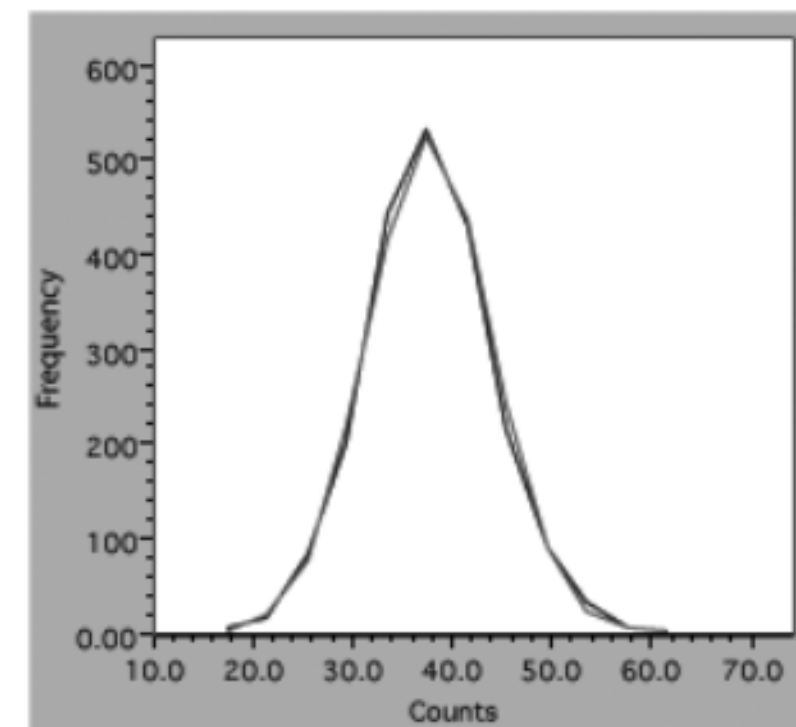
# Guiding experimentation

- Data plots are often a useful guide to experimentation.

- A plot will quickly show if parameters are varying as expected.

- These may indicate regimes where more or less data are needed.
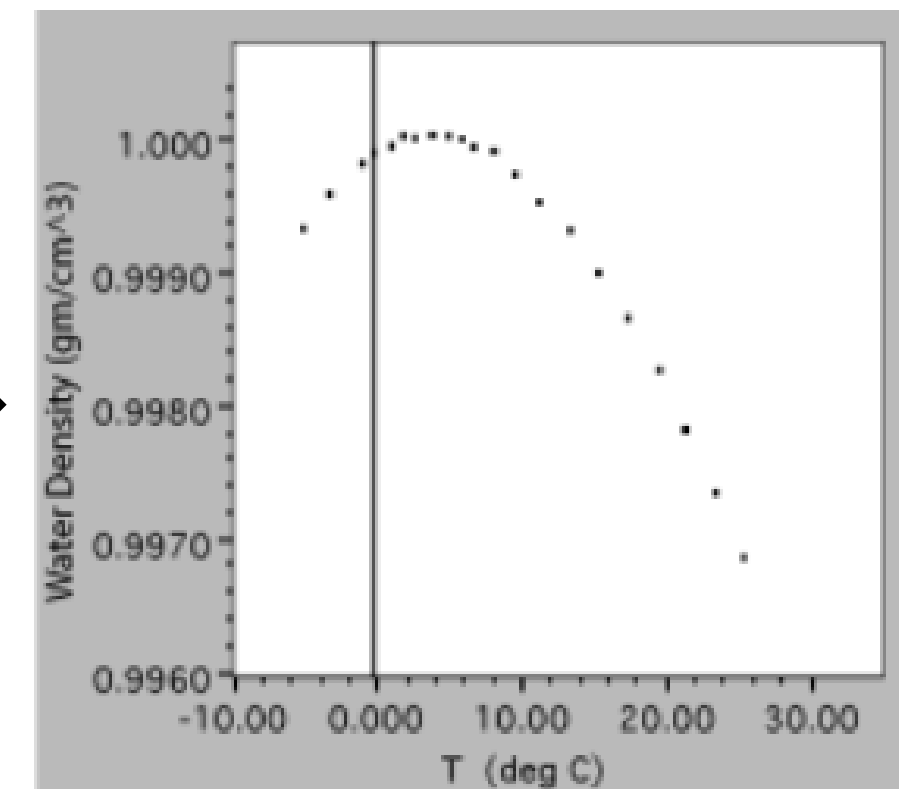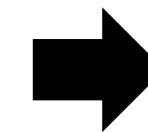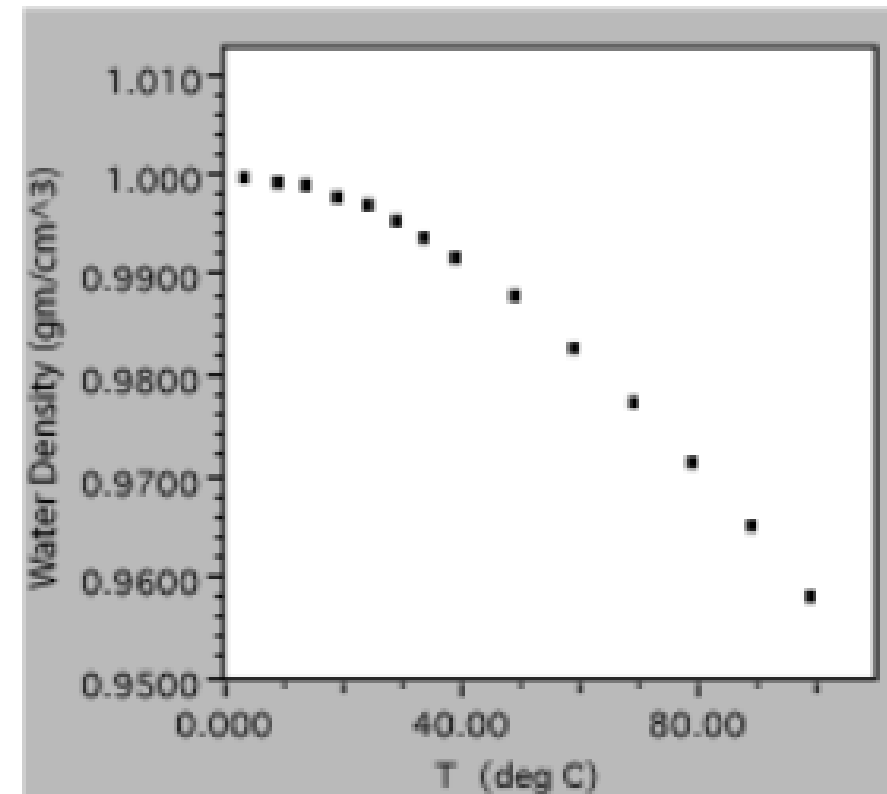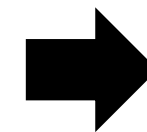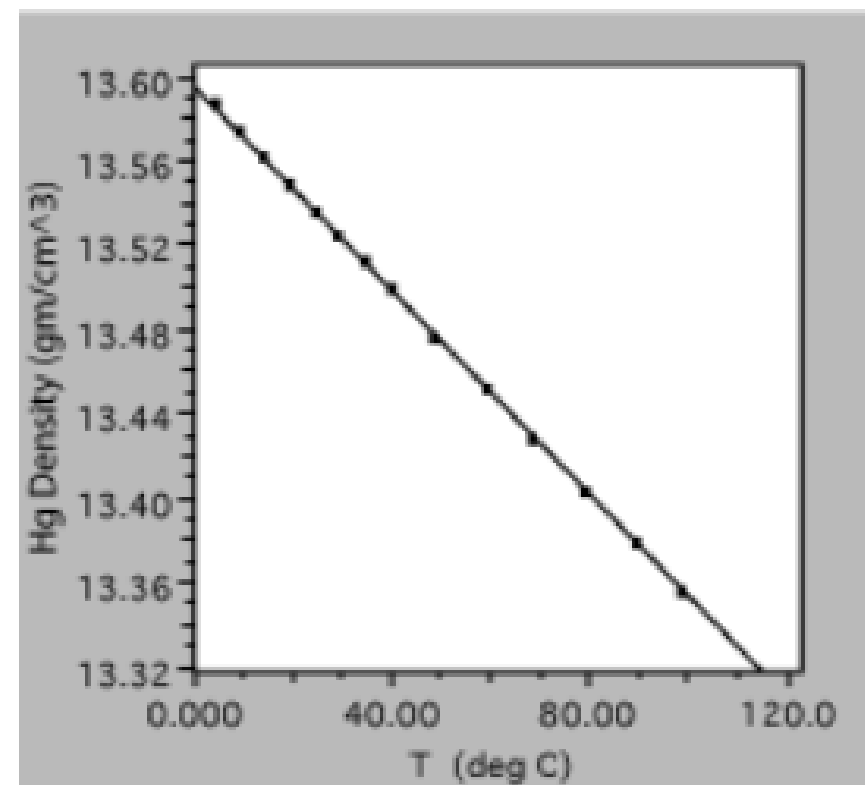
# Transformation of variables

- It is sometimes helpful to mathematically transform one or both of the variables before plotting.

- The technique can be used to linearize data to simplify model fitting, or to change the way data are distributed to clarify display. The exact procedure will depend on the situation.

1. Making a relationship linear

2. Changing the distribution

# Thank you