William CLEVELAND's Elements of Graphing Data
Outline of principles
[prepared by Ryan Womack, rwomack@rutgers.edu, 1/2014]

- *CLEAR VISION*
  - *Make the data stand out. Avoid superfluity*
  - *Use visually prominent graphical elements to show the data.*
  - *Use a pair of scale lines for each variable. Make the data region the interior of the rectangle formed by the scale lines. Put tick marks outside of the data region.*
  - *Do not clutter the data region [legend outside, etc.].*
  - *Do not overdo the number of tick marks.*
  - *Use a reference line when there is an important value that must be seen across the entire graph, but do not let the line interfere with the data.*
  - *Do not allow data labels in the data region to interfere with the quantitative data or to clutter the graph.*
  - *Avoid utting notes, keys, and markers in the data region. put keys and markers just outside the data region and ut nots in the legend or in the text.*
  - *Overlapping ploutting sybols must be visually distinguishable.*
  - *Superposed data sets must be readily visually discriminated.*
  - *Visual clarity must be preserved under reduction and reproduction.*
- *CLEAR UNDERSTANDING*
  - *Put major conclusions into graphical form. Make legends comprehensive and informative.*
  - *Error bars should be clearly explained [error bars can be 1) sample s.d. of data; 2) estimate of s.d. of a quantity; 3. confidence interval of a quantity]*
  - *When logarithms of a variable are graphed, the scale label schould correspond to the tick mark labels (can show both original and log scale on opp. sides of graph)*
  - *Proofread graphs.*
- *SCALES*
  - *Choose the range of the tick marks to include or nearly include the range of the data.*
  - *Subject to the constraints that scales have, choose the scales so that the data fill up as much of the data region as possible.*
  - *It is sometimes helpful to use the pair of scale lines for a variable to show different scales.*
  - *Choose appropriate scales when graphs are compared [use panels when one scale won't work].*
  - *Do not insist that zero always be included on a scale showing magnitude [except for bar charts].*
  - *Use a logarithmic scale when it is important to understand percent change or multiplicative factors.*
  - *Showing data on a logarithmic scale can improve resolution.*
  - *Use a scale break only when necessary. If a break cannot be avoided, use a full scale break. Do not connect numerical values on two side of a break.*
- *GENERAL STRATEGY*
  - *A large amount of quantitative information can be packed into a small region.*
  - *Graphing data should be an iterative, experimental process.*
  - *Graph data two or more times when it is needed.*
  - *Many useful graphs require careful, detailed study.*

TECHNIQUES

- *logs, % change, residuals*
- *point graph [2d histogram], histogram, percentile graph [and with comparisons/reference line], box plot [Tukey]*
- *dot charts - best way to attach label to quantity 2-way dot chart {multiway} grouped dot chart*
- *overlap is dealt with by jitter, distinguishable symbols {+sunflowers}, taking log or other transformation*
- *box plots for high multiples*
- *visually distinguish curves and points [this has gotten easy by now]*

THREE OR MORE VARIABLES

- *Framed-rectangle graphs*
- *scatterplot matrices*
- *interaction/brushing*
- *3d wireframe or stereogram (points)*

PERCEPTION

- *pie v. dot chart*
- *distance and detection*
- *length in a stacked bar*
- *45 degree banking*
- *strive for clarity*