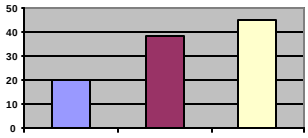
Data Display Glossary

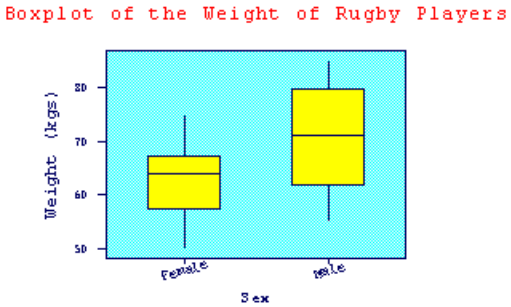
To learn more about these terms, visit the glossaries at these web sites:

HyperStat http://davidmlane.com/hyperstat/index.html
SurfStat Australia http://www.anu.edu.au/nceph/surfstat/surfstat-home/surfstat.html
Stat Soft, Inc. The Electronic Textbook http://www.statsoft.com/textbook/stathome.html
American Society for Quality http://www.asq.org/abtquality/glossary.html

Bar graph. A type of chart in which columns represent frequency or percentage distributions. Each bar is a separate column in the display. See example below.



Box plot. Also known as a Box-and-Whisker plot. A type of graph that depicts the relationship between observed and expected values. Expected values are represented by a line with end marks (to denote upper and lower limits) and a box with a line across the middle (most often denotes the median). The distance between an end mark and the top or bottom of the box is one quartile or ¼ of the distribution. Values that fall outside the upper limit are plotted on the graph. A box plot is often used to identify if a distribution is skewed and if outliers or unusual data values are present. See example below from the SurfStat Australia web site glossary.



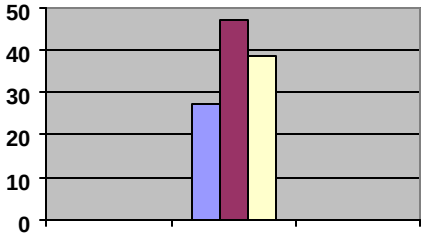
Concentration diagram. A picture is drawn that represents an item to be inspected or a process to be reviewed. Flaws, errors or major problems are marked on the picture where they occur. See example below.

```
xxx
    xx  x                                  xx
xxx
 x
```

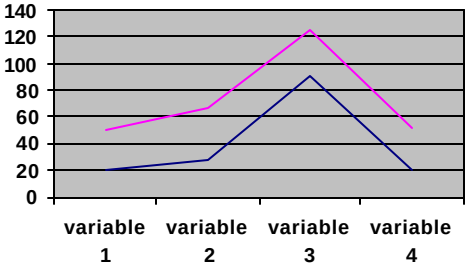
Recurring Grafitti in Visitor Area

Dot plot. A graph that is similar to a bar chart or histogram. Dots are used in place of bars and rectangles. The graph is used most often to identify outliers or to determine gaps in data.

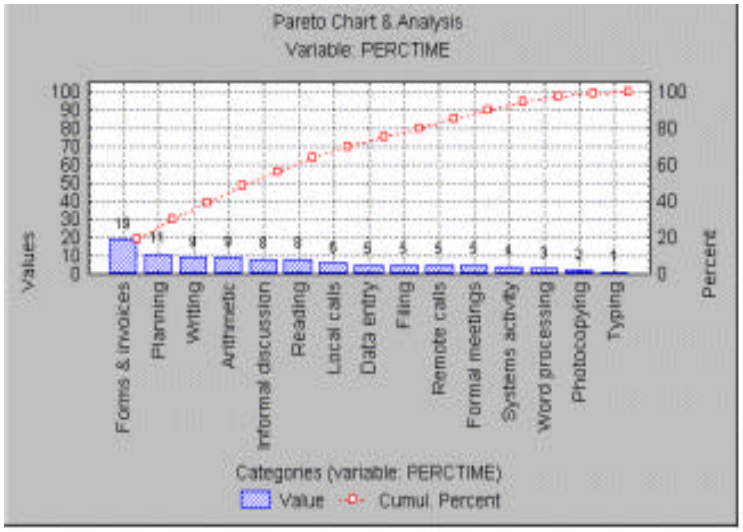
Histogram. A type of bar graph in which columns represent frequency or percentage distribution. The columns are adjacent to one another. See example below.



Line graph. A type of graph that displays values of the dependent variable (Y axis) for each category of an independent variable (X axis). See example below.

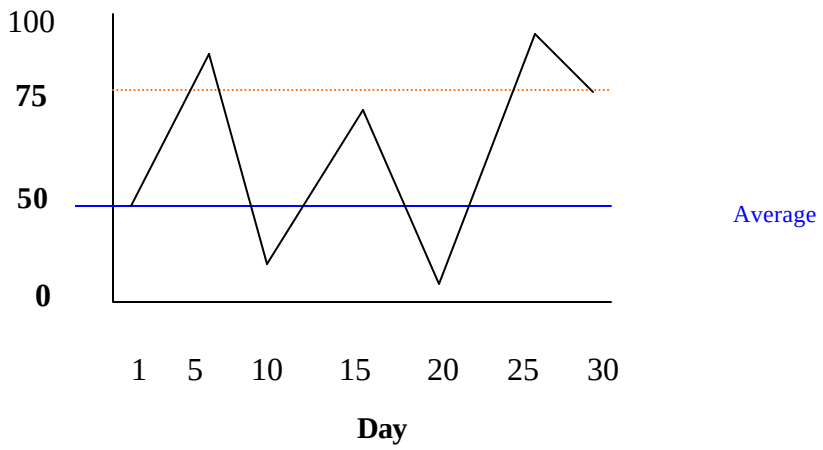


Pareto chart. A type of bar graph in which columns are arranged from highest to lowest frequency. Taller bars represent the relative frequency of more significant problems and shorter bars represent the relative frequency of less significant problems. The name is derived from the Pareto principle which states that 80% of the trouble comes from 20% of the problems. See example below from Stat Soft, Inc., The Electronic Textbook on-line chapter on Quality Control.

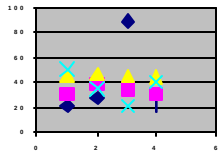


Run chart. A type of line graph used to monitor the performance of a process. Data values are plotted over a given period of time. See example below.

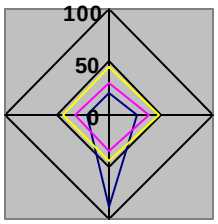
Percentage of Incomplete Intake Forms



Scatter plot. A statistical diagram that shows the relation between two variables. Data points are plotted onto a graph to display the spread of scores. See example below.



Spider chart. Also known as a "radar" chart. A diagram used to identify gaps in performance. Data on a critical performance measure for two or more agencies are plotted and connected on the diagram. Different colors or types of lines are used to differentiate between agencies. See example below.



Stem and leaf diagram. A type of chart in which values are separated into two components, a stem and a leaf. The stem is a multiple of the value. When the multiple is subtracted from the original value, the remaining portion becomes the leaf. For example, a value of 97 would have a stem of 9 (10x9) and leaf of 7. The diagram can be used to condense large amounts of data to identify the most and/or least commonly occuring value. See example below.

Stem	leaves
9	07
8	24
7	
6	01
5	105

Multiply stems by 10. Multiply leaves by 1 if stem is positive, -1 if negative.

Data: 97, 85, 88, 89, 82, 61, 59, 59, 59, 59, 59, 58, 59, 59, 59, 59, 59, 57

References:

American Society for Quality. (2000). Glossary [On-line]. Available http://www.asq.org

HyperStat Online. (2000). Statistical control chart tutorial [On-line]. Available http://davidmlane.com/hyperstat/index.html

HyperStat Online. (2000). Statistics Glossary [On-line]. Available http://davidmlane.com/hyperstat/index.html

Johns, M. L. (1997). Information management for health professionals. Albany, NY: Delmar.

Keehley, P., Medlin, S., MacBride, S., & Longmire, L. (1997). Benchmarking for best practices in the public sector. San Francisco: Jossey-Bass.

Office of Quality Services. (1995). QStp tool kit for quality. Columbus, OH: Author.

Shwartz, M., Ash, A. S., & Iezzoni, L. I. (1997). Comparing outcomes across providers. In L. I. Iezzoni (Ed.), Risk adjustment for measuring healthcare outcomes (2nd ed.). Chicago, IL: Health Administration Press.

Stat Soft, Inc. (1999). The electronic statistics textbook [On-line]. Available http://www.statsoft.com/textbook/stathome.html

SurfStat Australia. (2000). Statistics glossary [On-line]. Available http://www.anu.edu.au/nceph/surfstat/surfstat-home/surfstat.html (July, 2000).