



Components. A BOX PLOT identifies the **MEDIAN**, (50th percentile), the lower and upper quartiles (**25th and 75th PERCENTILES**), and the **RANGE** (extreme spread of the data). The edges of the box demarcate the 25th and 75th percentiles, and so represent the middle 50 percent (**INTERQUARTILE RANGE**) of the parameter values for the data subset. The line inside the box is the **MEDIAN**. The lines, or whiskers, extend outward from the box through the range of data, excluding outliers. Two outliers are defined, based on their distance from the nearest edge of the box, relative to the range of the box. **OUTSIDE VALUES** lie 1.5 to 3 interquartile ranges away from the nearest box edge, and **FAR-OUT VALUES** lie three or more interquartile ranges away from the nearest box edge. The **NOTCH** represents the approximate 95 percent confidence interval around the median.

Interpretation. If notches from different subsets of data overlap completely, one can conclude with 95% confidence that the groups have been sampled from a common population. If notches do not overlap at all, one can conclude (with 95% confidence) that the groups represent different populations. Cases of partial overlap require explicit tests (e.g., t-Test, ANOVA, Mann-Whitney, or Kruskal-Wallis) to specify significance of differences among groups.

Generic Box Plot

Figure A-6. Description of Box Plots