Exercise 2

POINTS.SYD is a real archaeological dataset containing information on 160 ground slate points from Maritime Archaic sites in the Northeast (based on Rutledge 1980). The following variables have been recorded for each point:

SITE - This is the name of the site at which the point was found.

FACETS - Points can either be faceted or unfaceted. The former is indicated by a plus, the latter is indicated by a minus. Faceted points have distinctly beveled edges, resulting in a blade that is hexagonal in cross-section. Unfaceted points lack beveling, and thus have a blade that is biconvex in cross-section. It is generally believed that faceted points are chronologically later than unfaceted ones.

LENGTH - This is the total length of the point in cm. Obviously, this variable could only be measured on complete points.

ANGLE - This is a measure of the bluntness of the point's tip. It is equivalent to the angle in degrees between the longitudinal axis of the point and the line that is tangent to the blade edge at the tip. Except in cases of extensive resharpening, heavily utilized points tend to have blunter tips than less utilized points.

CONTEXT - This indicates the kind of provenience in which the point was found. A value of "B" means that the point was found as a grave offering in a burial feature; a value of "H" indicates that the point was found in habitation debris or midden.

Your assignment is to do the following:

1. Use a bar chart or pie diagram to show the proportion of all faceted points associated with each kind of context (i.e., burial vs. habitation). Do the same for unfaceted points and compare the two diagrams.

2. Compute descriptive statistics (mean, median, standard deviation, hinge spread) for the numeric measurements on all points, then do the same for faceted and unfaceted points separately. Explore the differences between faceted and unfaceted points using boxplots and stem-and-leaf diagrams. What's going on?

3. Using all the data on points, plot several histograms of length, experimenting with the class intervals (boundaries and widths) to see what such variation does to the shape of the histogram. Next, stratify these points by context and plot histograms of length for each context separately. How do you interpret your results? If you had to define two categories of points based on length, where would you set the boundary between the categories and why? (If the patterning in the frequency distribution of length for all points is not as clear as you would like, try looking at the frequency distributions for faceted and unfaceted points separately.

4. Is there evidence that points found in burials exhibit different amounts of utilization than those found in habitation contexts? (Answer this question for all points combined, then for faceted and unfaceted points separately.) Do points of different lengths differ in terms of their degree of utilization? Be sure to support your answers with appropriate charts and descriptive statistics. (Boxplots may be especially useful.) Interpret your results.

Reference: