Stem-and-leaf Plot

- Choose a spot to split the data into two pieces. The first part of each number is the stem, the remaining piece is the leaf.
- Make a column of all the possible stems from the lowest that occurs in your data to the highest. Include all values, even if some do not occur in your data.
- Draw a vertical line next to the column of stems.
- Go through the dataset and attach the leaf for each observation to the right of the vertical line on the row for its stem. If the leaves are more than one digit, leave a space between leaves on any branch.
- After all the leaves are attached, it may be helpful to rewrite the plot with the leaves on each branch ordered from smallest to largest.
- If the plot looks too compact you may adjust the number of branches per stem from 1 to 2, 5, or 10.
- If the plot looks too spread out, split the data one digit further to the left, and redraw the plot with the new stems.
- If needed, you may use a 0 stem for numbers that would only have leaves, or use add a ".0" to whole numbers to be used as leaves.

Dotplot

- Draw a numberline that covers the range of the data.
- Divide the numberline up evenly.
- Place a dot above the numberline for each observation.
- If a particular value appears multiple times, stack the dots on top of each other. Keep the spacing even, so that two "stacks" that have three dots are the same height.
- NOTE: This method works best with discrete data.

Frequency Table

1. Choose the number of categories to use. The following table is a general guide that we will use. If n is between 16 and 31 we will generally use 5 categories, and between 32 and 63, we will use 6. Larger than that we would not do by hand, and smaller will not be interesting.
2. Find the range, that is, the largest observation minus the smallest.
3. Divide the range by the number of classes, and truncate the answer to the same accuracy as the data. Then add one to the last digit of this number.
4. Start the first category 1/2 unit below the first observation.
5. Complete the categories by adding the number found in step three to the number from step four to get the right boundary of the first category. Continue adding to complete the necessary number of categories.

6. Make a table with headings for the boundaries, the frequency, the relative frequency, the cumulative frequency, and the cumulative relative frequency.
   - **Frequency:** Simply the count of the number of observations in the category.
   - **Relative Frequency:** The frequency divided by the total number of observations, that is the proportion of the data that is in the category. This may also be thought of as the size of the category relative to the whole data set.
   - **Cumulative Frequency:** The count of the amount of data in this categories, and any previous category on the table. This is essentially accumulating the observations as we move from smaller values to bigger ones.
   - **Cumulative Relative Frequency:** The cumulative frequency divided by the sample size, that is the proportion of the data that has been seen so far. This may also be calculated by adding up the values in the relative frequency column, but this is more likely to be affected by rounding error.

Histogram

- A Frequency Table (with at least the columns for frequency and relative frequency) must be constructed to be able to construct a histogram.
- Draw a set of axes, and indicate the category boundaries along the x-axis. The y-axis is generally used to indicate the relative frequency.
- Bars are drawn for each category from one boundary to the other, with height determined by the relative frequency.

Boxplot

1. Find the median and both quartiles.
2. Find the intraquartile range.
3. Find $L_1 = 1.5 * \text{iqr}$ and $L_2 = 3 * \text{iqr}$.
4. Find the following:

$$\begin{array}{lll} \text{Inner Fences:} & f_1 = x_{.25} - L_1 & f_3 = x_{.75} + L_1 \\ \text{Outer Fences:} & F_1 = x_{.25} - L_2 & F_3 = x_{.75} + L_2 \end{array}$$

5. a_1 is the smallest value in the dataset greater than or equal to f_1, and a_3 is the largest value less than or equal to f_3.
6. Draw a numberline covering the range of the data, and locate a_1, $x_{.25}$, $x_.5$, $x_{.75}$, and a_3.
7. Draw a box from $x_{.25}$ to $x_{.75}$ with a line in the center at the median.
8. Draw "whiskers" from each side to the adjacent value.
9. Points between the inner and outer fences are indicated by a closed circle. $\bullet$
10. Points past the outer fence are indicated by an open circle. $\circ$