Summarizing Numerical Data

1. In the space below, outline a procedure to take a data set like the following and return a single statistic that fulfills the desiderata:

a. The statistic will increase in magnitude as variability increases
b. The statistic will decrease in magnitude as the variability decreases
c. The statistic is zero when there is no variability

You can express your procedure as a written recipe or using mathematical notation.

$$6 \quad 7 \quad 7 \quad 7 \quad 8 \quad 8 \quad 9 \quad 9 \quad 10 \quad 11 \quad 11$$

2. The small data set above was measured in meters, but what would have happened if it had been measured in decimeters (10 decimeters to a meter)? Provide reasoning for would happen to the measures of center - mean, median, mode - if it had instead been measured in decimeters. Repeat the exercise for three measures of spread: range, standard deviation, and IQR. Which measures remain the same after a multiplicative change in units?

3. Repeat exercise two, but consider a circumstance where the data set was recorded using a faulty measuring device that added .1 to each observation (instead of multiplying by 10). Which measures of center and spread remain the same after an additive change?

4. Sketch your best sense of the distribution of the following variable(s). For each, please:

    i. Use a form of statistical graphic that emphasizes the important elements of the distribution.
    ii. Label the axes and provide plausible values for the tick marks.
    iii. Describe in words the shape of the distribution.
    iv. State which measure of center and spread would be most appropriate and approximate their values.

Make a note of any assumptions you're making in interpreting these variable names.

Number of body piercings among Stat 20 students

Scores on an easy quiz among Stat 20 students

Distance from campus and *Year at Cal (first, second, etc)* among all Cal students