Central Tendency

Central Tendency

- We are in section 2.5 of the textbook
- *Central tendency* describes the ”average” or ”typical” value of a data set, the center of the data.
- The *mode* is the most common (frequent) data value.
- The *median* is the middle of the data when sorted in order.
  - With an even number of data values, the median is halfway between the middle two points.
- The *mean* is calculated by adding the values and dividing by the number of values.
  Notation: sample mean $\bar{x} = \frac{\sum x}{N}$
  - N the size of a sample or data set (number of points)
  - The (random) variable x
  - Add values $\sum$
  - Mean of variable x is $\bar{x}$ (sample mean).

  Population mean μ
- *Outliers* are data points that are much larger or much smaller than the others.
  - Median is not very sensitive to outliers (unchanged if the same number of data points at the top and bottom are thrown out)
  - Mean is very sensitive to outliers (a very large outlier will increase the value of the mean; small will decrease the mean)
- Which measure of central tendency should you use?
  - Nominal data - use mode only
  - Ordinal data - mode, median, or mean based on discipline standards
  - Quantitative data with outliers - median
  - Mean is the most commonly used center measure

1. The university is studying the majors and grades of a collection of students. For each of the following variables, identify the measures of central tendency that apply.

   (a) The GPA of each student.

   (b) The major of each student.

   (c) The number of hours each student is taking this semester.

   (d) The high school each student attended.

2. Compute the mean and median of the following data set.

   8 12 15 16 20

3. Compute the mean and median of the following data set.

   40 70 74 80 90 100

4. We have a data set with 30 data points of ratio data, which are not all the same. Suppose that the highest value increases by 10.

   (a) Does the median change? If so, how much? If not, why not?

   (b) Does the mean change? If so, how much? If not, why not?

5. We have a data set with 30 data points of ratio data, which are not all the same. Suppose that the highest value increases by 12 and the lowest value decreases by 12.

   (a) Does the median change? If so, how much? If not, why not?

   (b) Does the mean change? If so, how much? If not, why not?