Measures of Center

Reading

- Section 2.5 (up to problem 2.29 in page 101)
- Section 2.6

Practice Problems

2.5 (Page 129) 43, 44, 46, 47
2.6 (Page 130) 49, 50, 52, 53, 57, 63

Notes

- There are two main measures of center, and another less used:

  Median The "middle" value. Half the values are below it, half above.
  Denoted by M.
  The median is stable to the effects of outliers.

  Mean Numerical average of all values.
  Denoted by $\bar{x}$.
  The mean has nice numerical properties, but it is affected by outliers.

  Trimmed-mean A mean computed after a percent of data from each end is removed.
  Resembles the mean, but is not much affected by outliers.

- NOTE: The book, and other resources, talk about the **mode**, referring to the "most frequent value". We will not use that term for this purpose in this class, but you should be aware of it.

- One important consideration is how these measures behave to changes in the data. This is the notion of "resistance":

  Resistant A measure is called *resistant*, if its value is not considerably affected by extreme changes to a few data values.
  In particular, outliers can be considered as values that were near others but their value somehow changed radically. So resistant measures are not affected much by outliers.

- Based on this description, the median is a resistant measure, but the mean is not. (THINK ABOUT IT!)