Class #3 - Monday February 8
Measures of Central Location

Textbook readings:

- Ross, Sec 3.2: Sample Means & Sec 3.3: Sample Medians
- Phillips, Chapter 3

Definitions & Spreadsheet Functions

- for frequency tables: `=frequency(data,class_endpoints)`
  - in Excel: select output cells, enter formula, press Control+Shift+Enter
- Sigma (Σ) notation for sums: for a sample of n data points whose values are $x_1, x_2, \ldots x_n$

$$\Sigma_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \ldots x_n$$

  - `=sum(data)`
- sample mean ("x-bar"):

$$\bar{x} = \frac{\Sigma_{i=1}^{n} x_i}{n}$$

  - `=average(data)`
- sample median: value m such that half of the data points in the sample are smaller than m (and hence half are larger than m)
  - `=median(data)`
- sample percentile: the p-th percentile is the value such that $p\%$ of the data points are smaller than that value
  - Note that the 50th percentile is the same as the median
  - Typically used when the number of data points is much larger than 100; often used in health care (e.g., height and weight) and education (e.g., SAT scores)
  - `=percentile(data, 0.95)` would output the 95th percentile

Homework #1 (due Wednesday February 10): see `https://openlab.citytech.cuny.edu/mat1372-statistics-spring2016-ganguli/assignments/`

- For each exercise, enter the data into a spreadsheet
- Choose appropriate class intervals and construct a frequency table using `=frequency`
- Sort the data using `=sort(data)` and check that your frequency table is correct
- Produce a frequency histogram; either sketch it by hand, or (preferably) use the spreadsheet to create it
- Hand in printouts of your spreadsheets and your written answers to additional questions in the textbook