Data Collection and Frequency Distributions

- Trouble with Design
- Sampling Types
- Frequency Distributions

Frequency Distributions

- **Frequency Distribution:** A table that shows the number of values in each category

Age	Frequency
15-24	19
25-34	8
35-44	5

Example Frequency Distributions

- **Example:** How many miles from here is your birthplace?
  - Sacramento = 100
  - San Francisco = 180
  - Los Angeles = 480
  - Seattle = 750
  - New York = 2,750
  - Acapulco = 1,867

Frequency Distribution Attributes

- **Lower Class Limits:** The smallest numbers in each class.

- **Upper Class Limits:** The largest numbers in each class.

- **Class Boundaries:** The numbers separating the classes (the midpoint of the upper class boundary and the next lower class boundary.

- **Class Midpoints:** The midpoint of the lower and upper boundary in each class.

- **Class Width:** The difference between two consecutive lower class boundaries.

$$\text{Width} = \frac{\text{Maximum} - \text{Minimum}}{\text{Number of Classes}} \uparrow$$

Relative and Cumulative

- **Relative Frequency Distribution:** Instead frequency, use

- **Cumulative Frequency Distribution:** Instead of frequency use the number at or below that class.

Normal Distribution

- The data is symmetric.
- The frequencies start small, get larger and then end small.

Interval	Freq.
0-9	3
10-19	9
20-29	20
30-39	45
40-49	21
50-59	8
60-69	3

Interval	Freq.
0-9	3
10-19	9
20-29	20
30-39	45
40-49	65
50-59	89
60-69	140

Approximately Normal   Not Normal