# Class 2

Sri Akhil Jonnalagadda
sjonnala@stedewards.edu

# Descriptive Statistics

❖ The idea is to explain the data at face value

❖ Summarize key metrics

❖ Describe what we have

❖ Present the data efficiently

# Summary Measures

* Where is the data centered?

  * IE where is the mean + standard deviation

  * Distribution?

  * Mode or Median

* Do we trust the data

  * Does it make sense

  * Outliers

  * What are we missing

* What's the range of outcomes?

  * Is it skewed?

# Spread Measures

❖ Variance is the average squared deviations from the mean

❖ Standard deviation is the square root of this number

❖ Covariance measures the direction of the relationship between two variables

# Probability distribution

❖ What is it?

❖ Z scores?

❖ Why do we care

# How to find skew

- ❖ Statistically:

    - ❖ The kurtosis coefficient is a summary measure that tells us whether the tails of the distribution are more or less extreme than the normal distribution.

    - ❖ Tails that are more extreme than the normal distribution are called

    $$\frac{n}{(n-1)(n-2)(n-3)} \sum \left( \frac{x_i - \bar{x}}{s} \right)^4$$

- ❖ Better to just plot it!

# Lets hop into it

❖ Get R going

❖