

## Lesson 3.1 Extension - February 16, 2021

---

### Review

- Review of StatCrunch features and the necessary parameters for graphs
  - Recall how to graph categorical data, using stem-and-leaf plots and other types of graphs
  - Recall that you can create graphs by clicking through the **Graph** tab on StatCrunch
  - You can "restrict" a specific dataset so as to only graph data that is relevant to the question that you desire to answer
  - There are several "Select a Graph" questions in the textbook which you can practice on and improve your skills in this regard
- 

### Symmetric Distributions

- Standard Deviation tells us about the variance of the dataset, or how spread out the data is
  - Mean is the average of the dataset
  - Mathematically, the mean refers to the sum of all the numbers, divided by the total number of terms
  - The mean could be imagined as the fulcrum of a see-saw, being dynamic based on the relative weights of the objects on either side
  - Outliers have a very noticeable effect on the mean of a specific dataset
  - In StatCrunch Histograms, you can choose to display the mean in the graph.
- 

### Conventions

- The mean of a specific dataset is either represented as  $\bar{x}$  or as  $\mu$
- The standard deviation of a specific dataset is represented as  $S$
- The formal mean formula is:

$$\bar{x} = \frac{\sum x}{n}$$

- The sample size is usually referred to as  $n$

- Standard Deviation is calculated as follows:

$$S = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

---

## Symmetric Distribution Rules

- If the data is unimodal and symmetric, then the following rules apply:
- Approximately 68% of all data points are within 1 standard deviation from the mean value
- Approximately 95% are within 2 standard deviations from the mean value
- Approximately 99.7% (pretty much all of the values) are within 3 standard deviations from the mean value

---

## Standard Deviation

- Standard Deviations give context to the data set and to any given number that we may have from the dataset
- Knowing the mean can only tell you whether you are lower or higher than the average, but not how that delta compares with other people's difference from the mean
- Given a list of test scores and your test score, you can find what a good score is by using standard deviation
- The standard deviation helps you understand a certain number in the context of a dataset