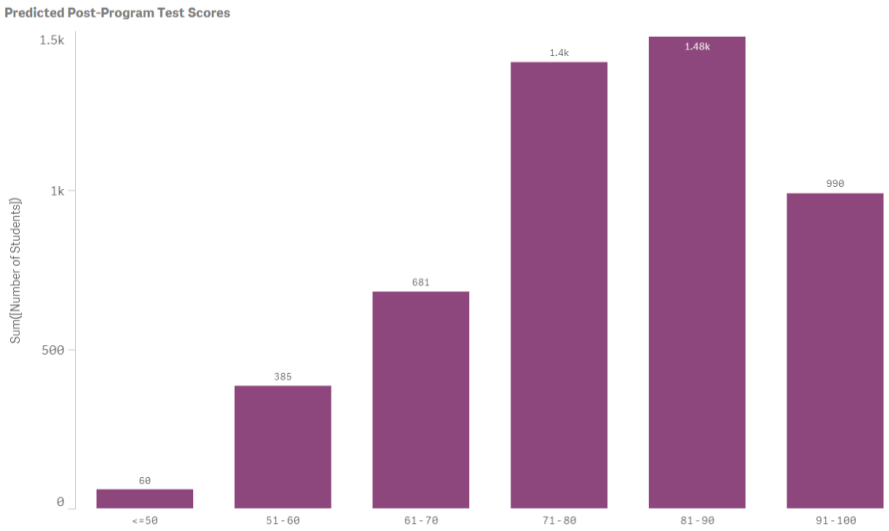
Understanding Distributions

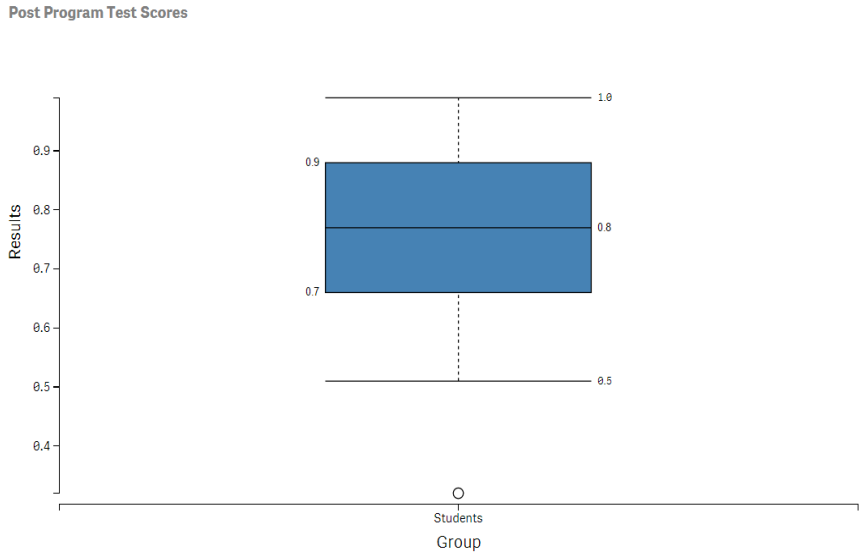
Distributions

An orientation of data points, broken down by their observed or predicted frequency of occurrence.

Visualizations - distributions are commonly visualized through histograms and box plots



Example Histogram



Example Box Plot

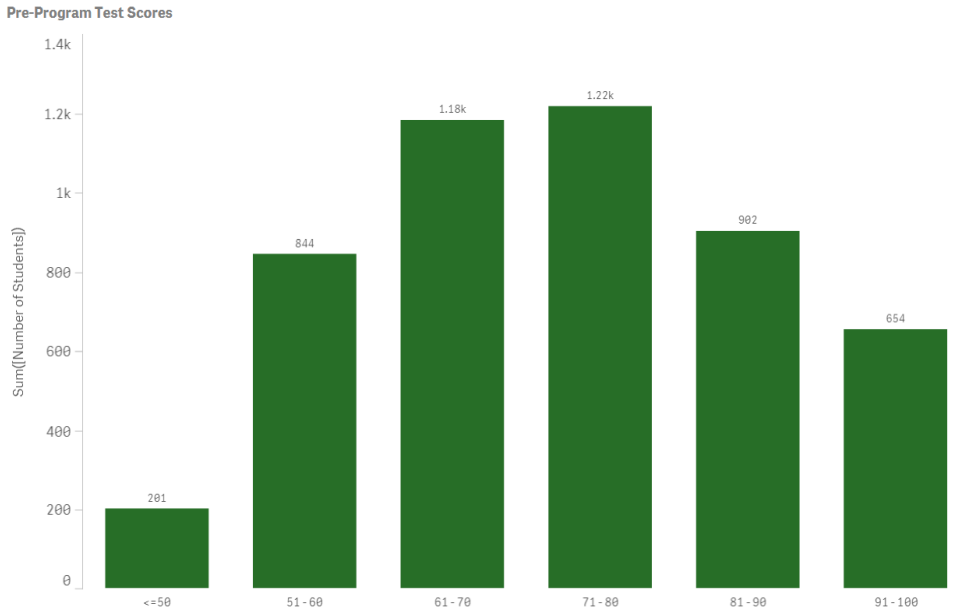
Understanding Distributions

Types

Two common types of distributions are continuous and discrete.

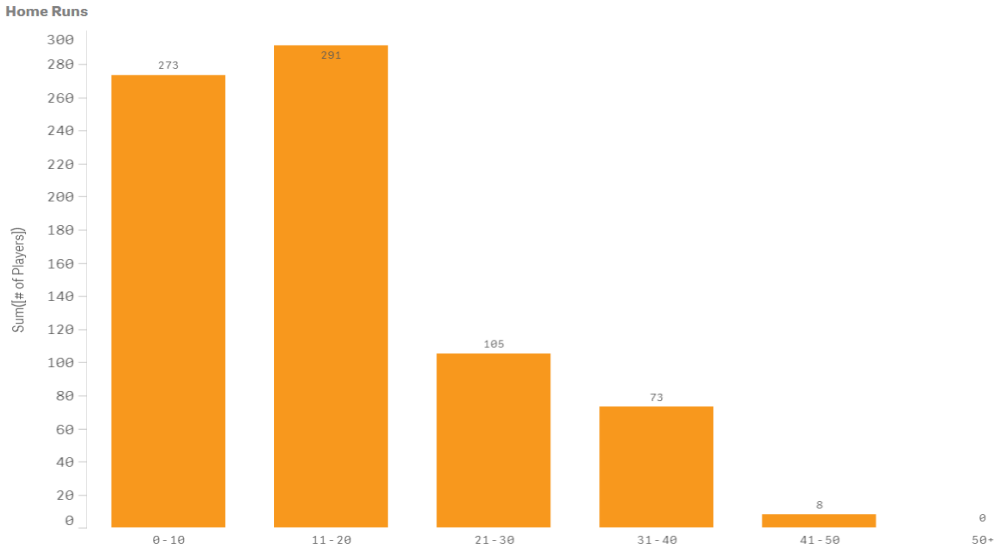
Continuous

Data can take on an infinite amount of possibilities within the range



Discrete

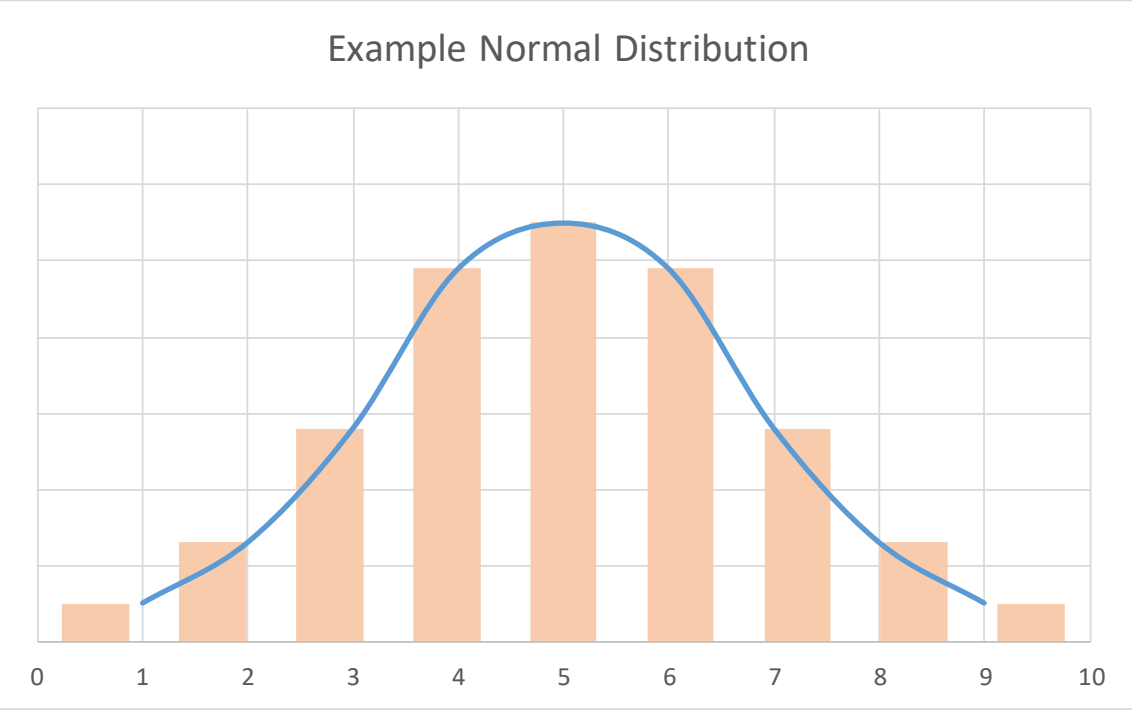
Where the potential results are finite or countably infinite number of potential results, based off a set of discrete variables



Understanding Distributions

Normal Distribution

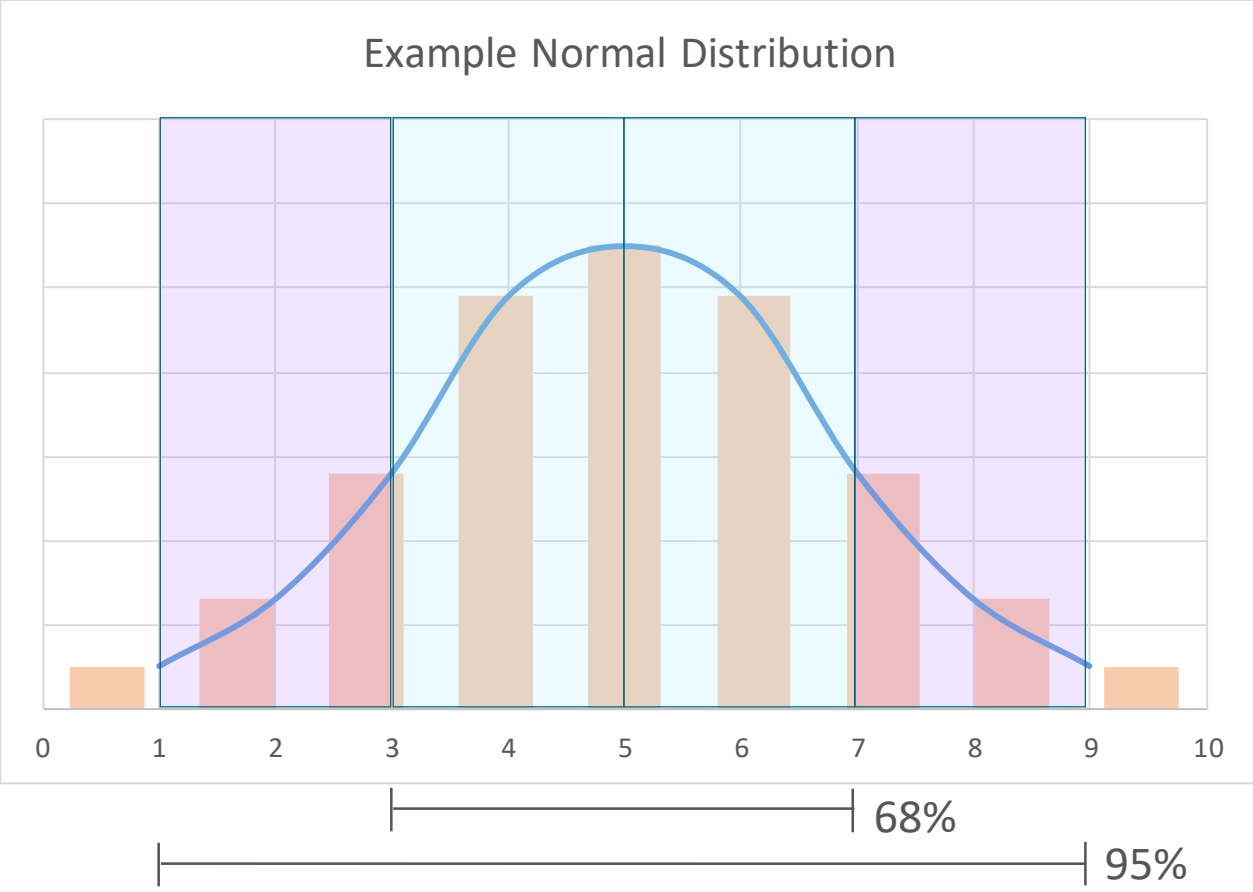
It has a bell shaped curve. Half the results fall above the mean and half fall below. The mean equals the median which equals the mode.



Understanding Distributions

Standard Deviation

A common tool in analytics that measures the dispersion of a data population.



σ = *Standard Deviation*

μ = *Mean*

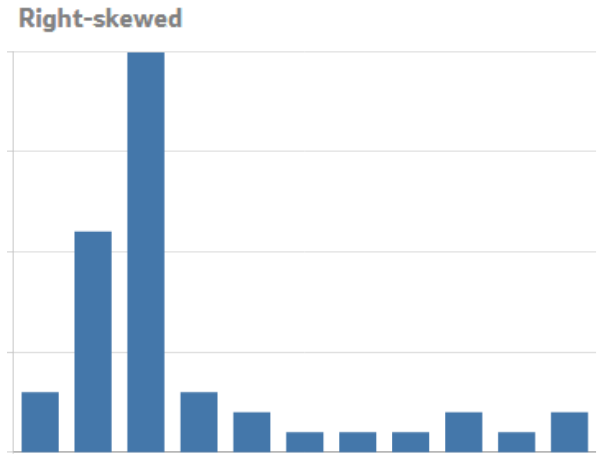
- +/- 1σ = *68.27%*
- +/- 2σ = *95.45%*
- +/- 3σ = *99.7%*

Understanding Distributions

Characteristics

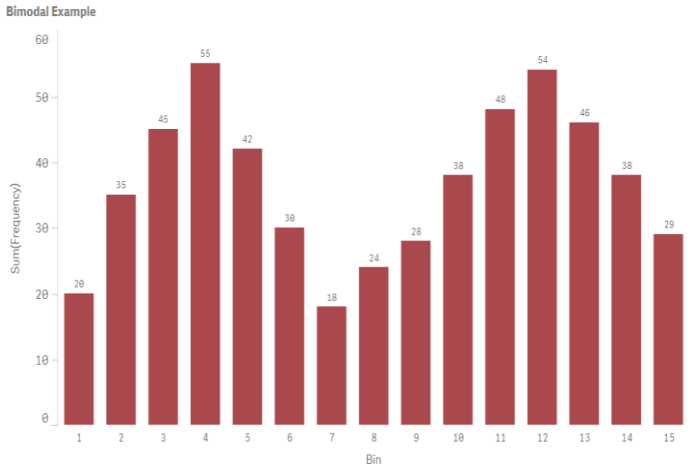
Three characteristics are skewness, bimodal, and multimodal.

Skewness



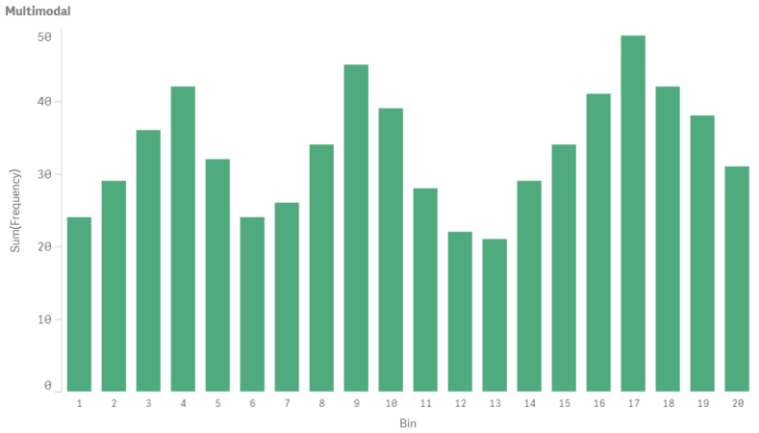
Measure of asymmetry in a distribution

Bimodal



2 Peaks

Multimodal



2 or more peaks