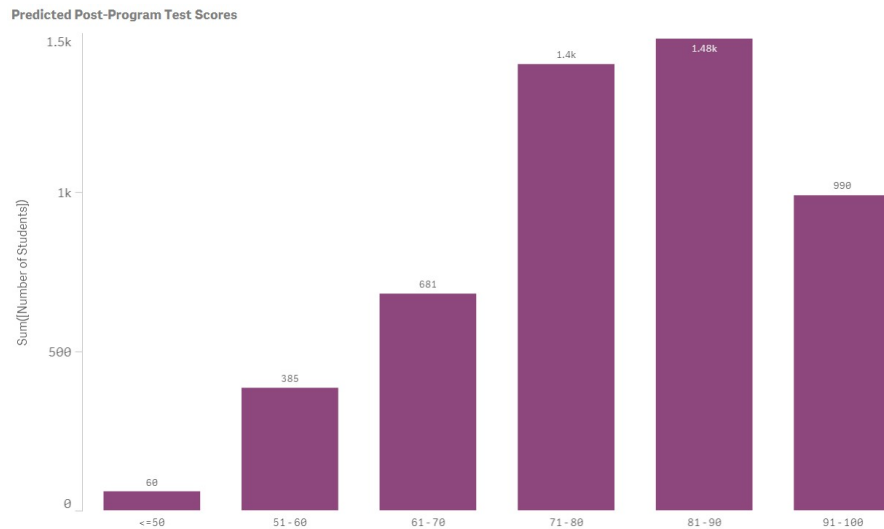


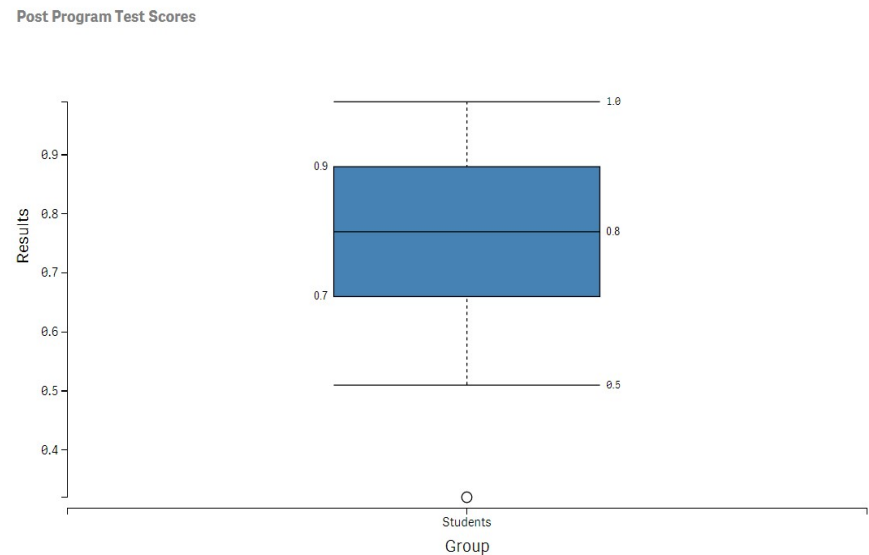
Distributions

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



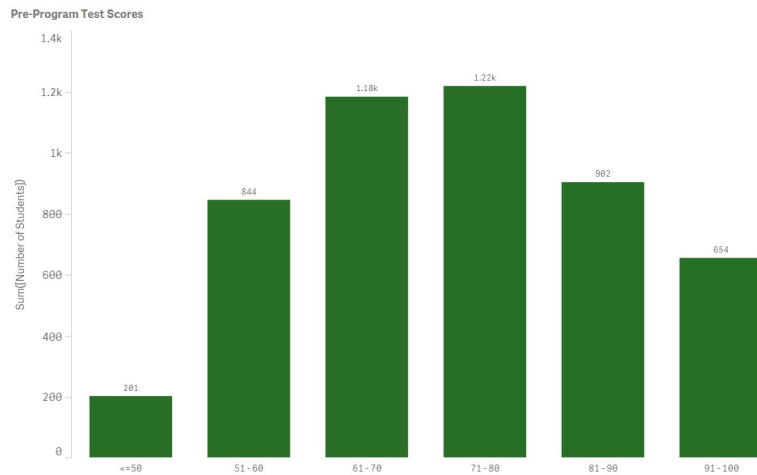
Example Box Plot

Types

Two common types of distributions are continuous and discrete

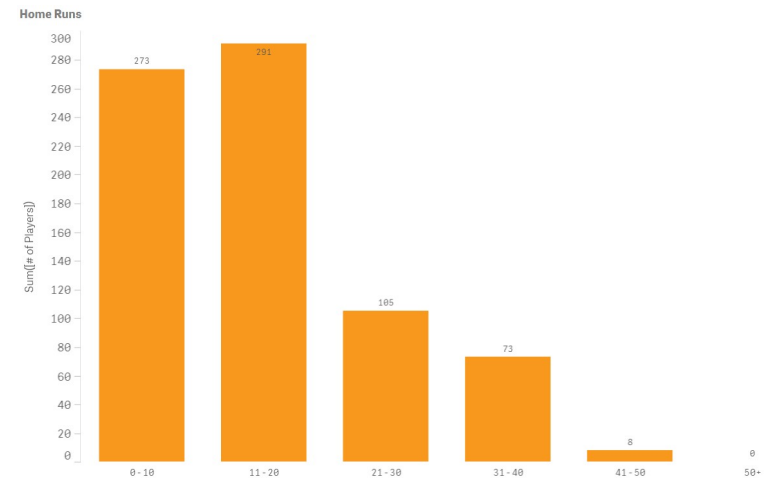
Continuous

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



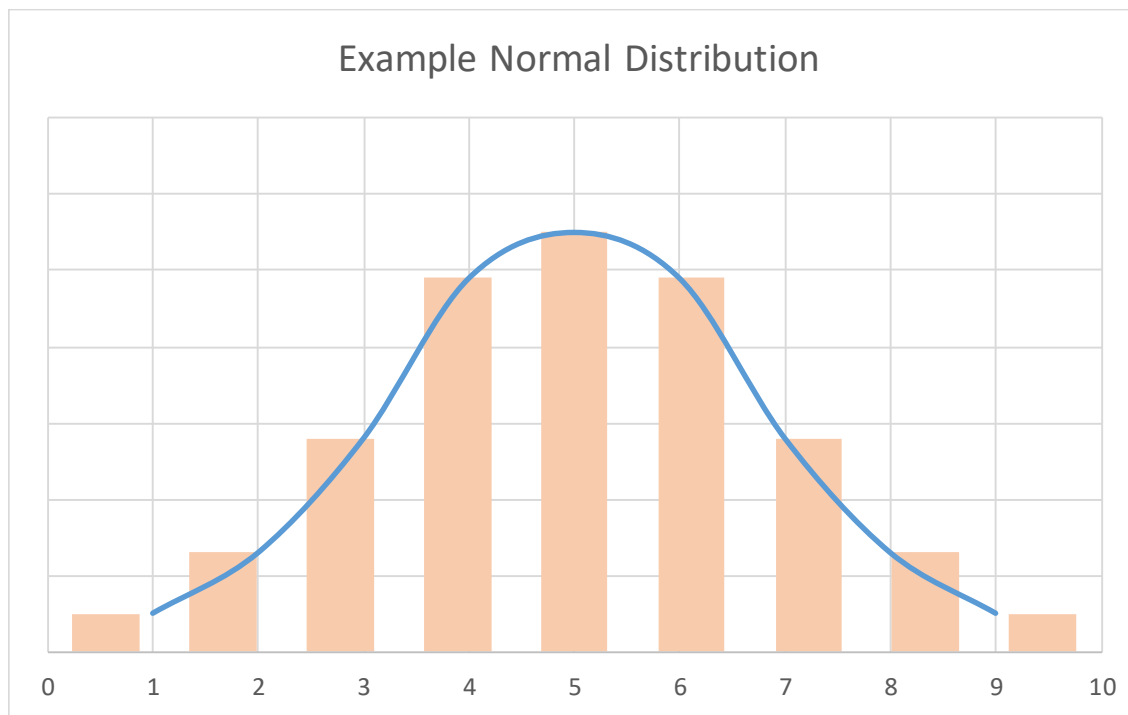
Discrete

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



Normal Distribution

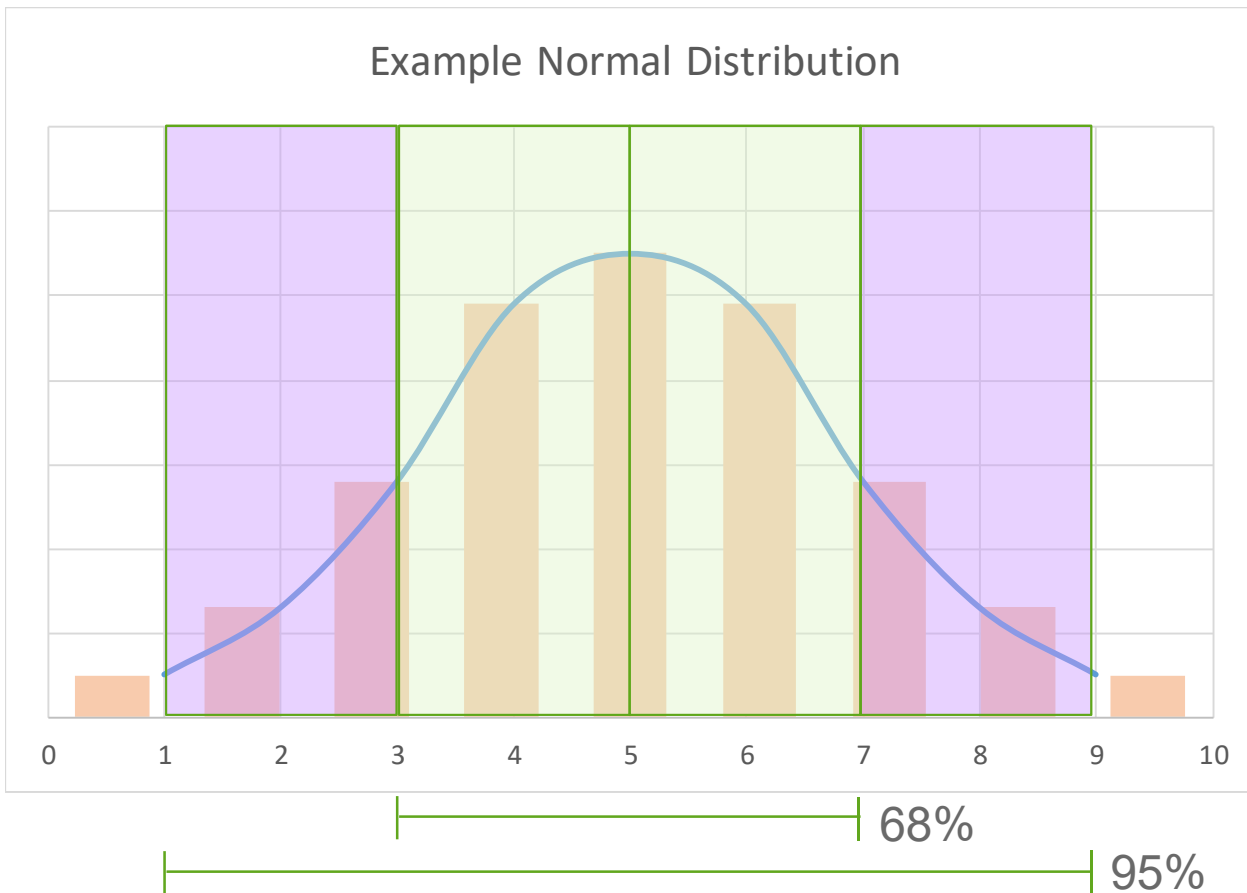
Bell shaped curve. Half the results fall above the mean and half fall below. The mean = the median = the mode.



Standard Deviation

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

Example Normal Distribution



σ = Standard Deviation

μ = Mean

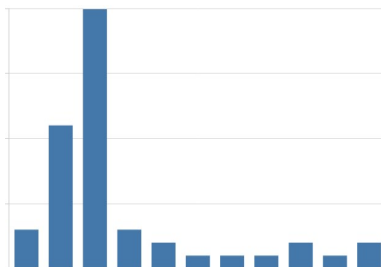
- $\pm 1\sigma = 68\%$
- $\pm 2\sigma = 95\%$
- $\pm 3\sigma = 99.7\%$

Characteristics

Skewness, Bimodal, Multimodal

Skewness

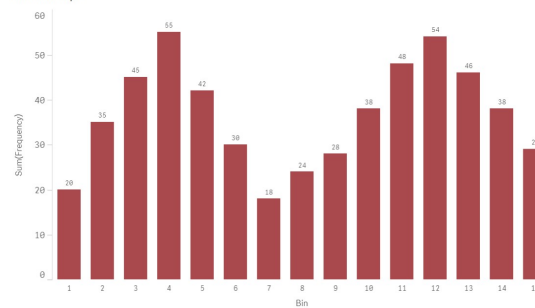
Right-skewed



Measure of asymmetry in a distribution

Bimodal

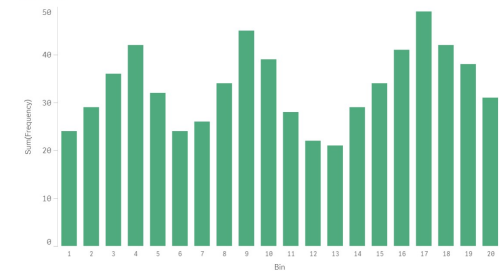
Bimodal Example



2 Peaks

Multimodal

Multimodal



2 or more peaks