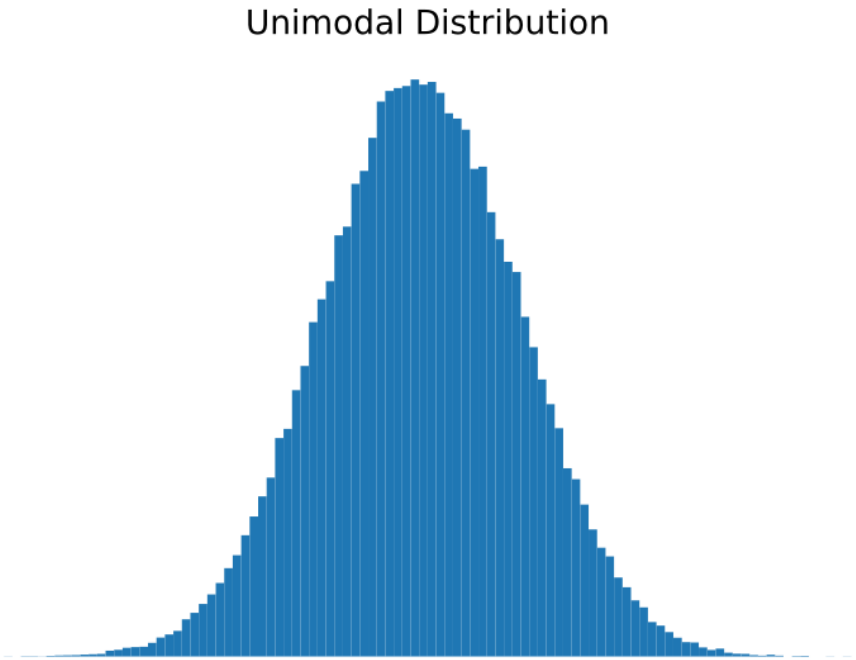
Describe a Histogram

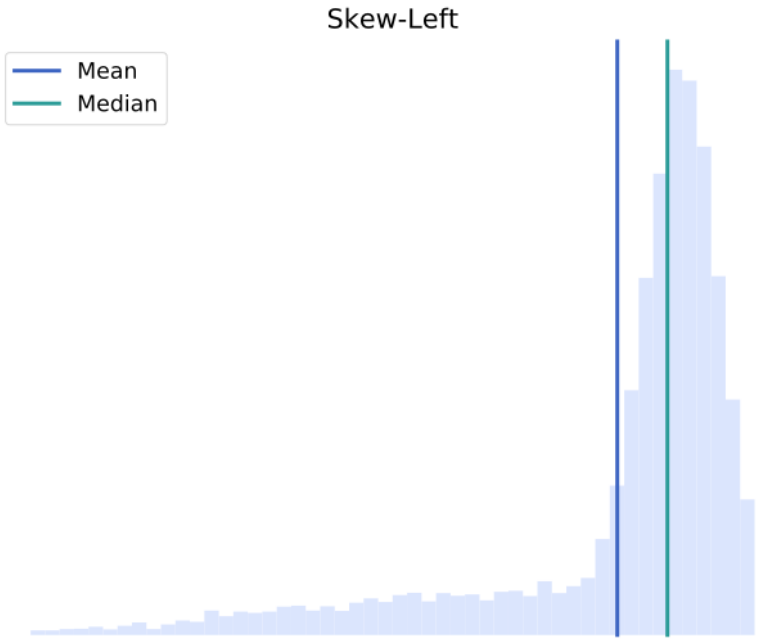
Unimodal Distribution

Modality describes the number of peaks in a dataset. A *unimodal* distribution in a histogram means there is one distinct peak indicating the most frequent value in a histogram.



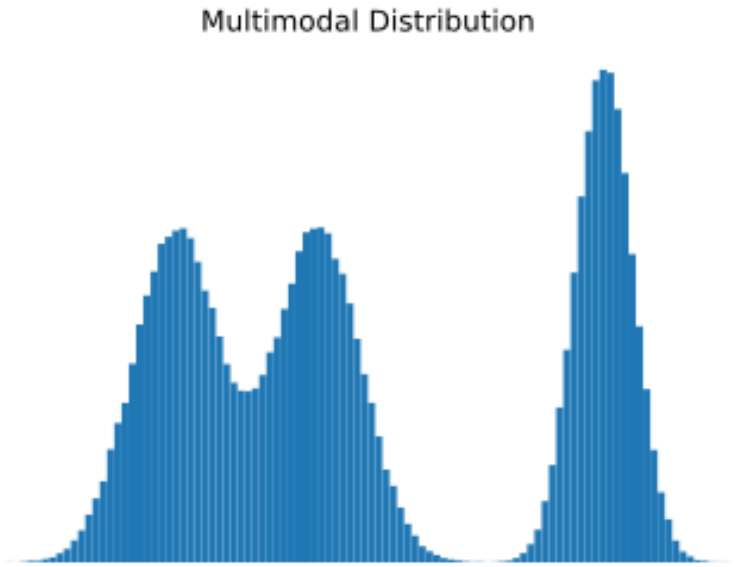
Left-Skewed Dataset

A left-skewed dataset has a long left tail with one prominent peak to the right. The median of this dataset is greater than the mean of this dataset.



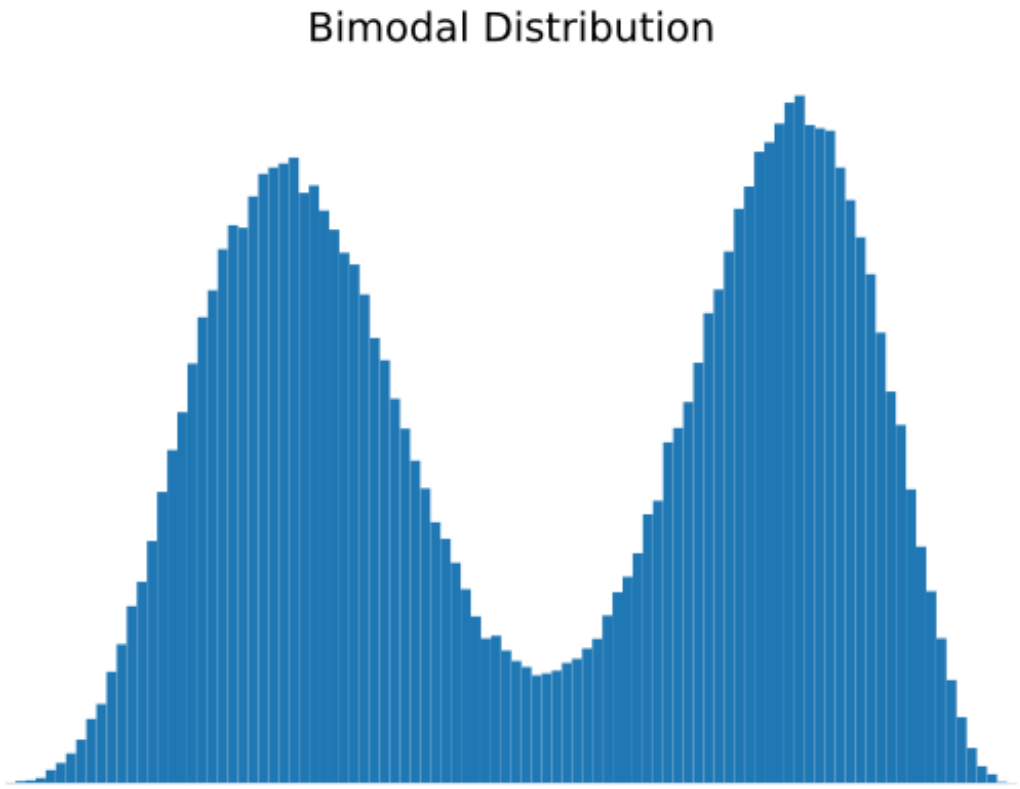
Multimodal Dataset

If a histogram has more than two peaks, then the dataset is referred to as *multimodal*.



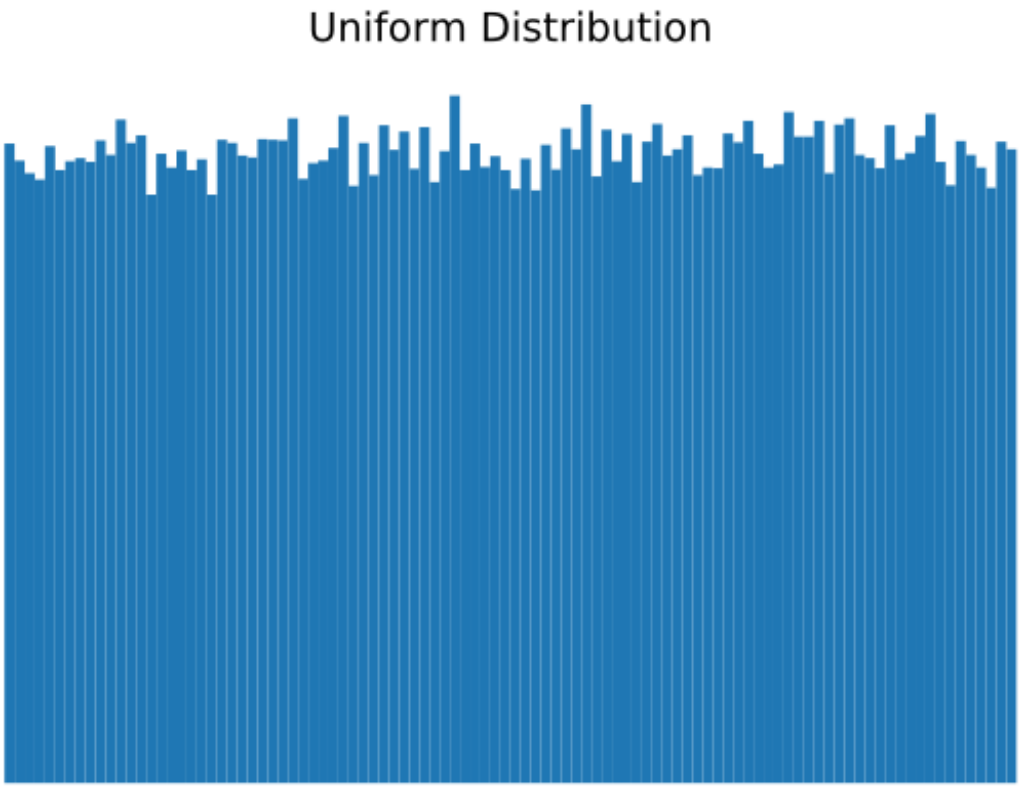
Bimodal Dataset

A bimodal dataset has two distinct peaks. This typically happens when the dataset contains two different populations.



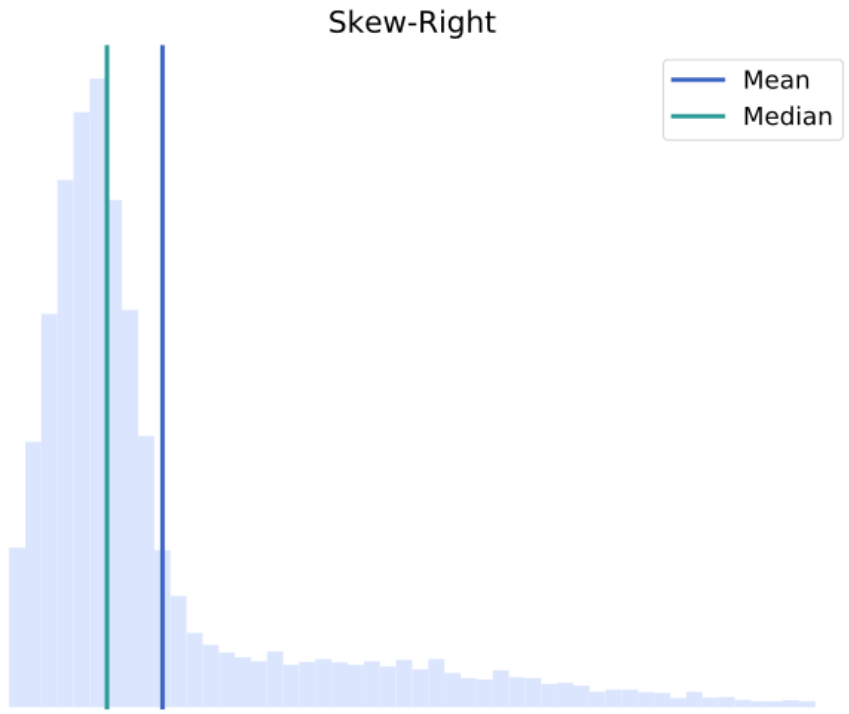
Uniform Dataset

A *uniform* dataset does not have any distinct peaks.
As seen in the histogram below, uniform datasets have approximately the same number of values in each group represented by a bar - there is no obvious clustering.



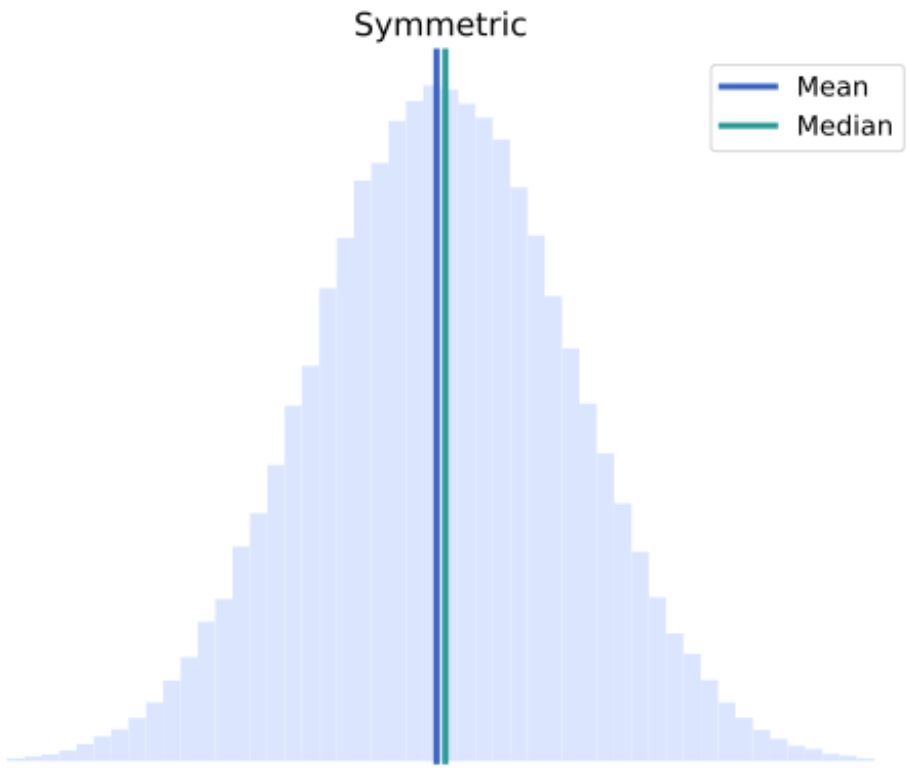
Right-skewed Dataset

In a histogram, if the prominent peak lies to the left with the tail extending to the right, then it is called a *right-skewed* dataset. In this case, the median is less than the mean of the dataset.



Symmetric Distribution in Histogram

In a histogram, the distribution of the data is symmetric if it has one prominent peak and equal tails to the left and the right. The Median and the Mean of a symmetric dataset are similar.



Dataset Outliers

An *outlier* is a data point that differs significantly from the rest of the values in a dataset.

For example, in the dataset `[1, 2, 3, 4, 100]` the value `100` is an outlier because it lies a large distance from the rest of the data.

Spread of a Dataset

The spread of a dataset is the dispersion from the dataset's center. The descriptive statistics that describe the spread are range, variance and standard deviation.

For example, for the dataset `[1, 4, 7, 10]`, the *range* of the dataset would be the maximum value of the set - the minimum value of the set, or `10 - 1 = 9`.

Peak of Unimodal Distribution

The center of a dataset is the peak of a unimodal distribution. The statistics that describe the center of a dataset are the mean and median.