There are three basic approaches to generating forecasts: regression-based methods,

And heuristic smoothing methods. and general time series models.

Top of Form

Bottom of Form

# Some common formulas and calculations used in Descriptive Statistics

## Sum of Squares

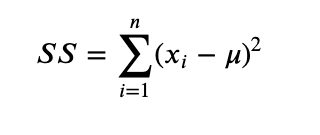
The sum of squares refers to the aggregated total of each data point's deviation from the mean, squared.

This calculation is a fundamental part of various statistical analyses, serving to quantify the variance within a dataset.

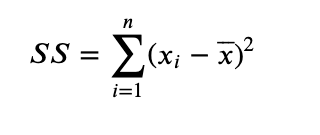
Top of Form

Bottom of Form

For a Population



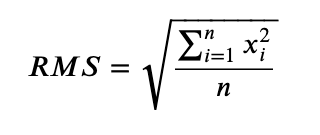
For a Sample



## 

## Root Mean Square

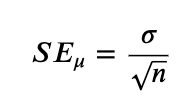
The root mean square (RMS) is a statistical measure that calculates the magnitude of a set of numbers. It is found by taking the square root of the average of the squares of the values in the set. This metric is especially useful in contexts where both positive and negative values in the dataset are treated equally, and it tends to give a higher value than the average due to the squaring of the values.



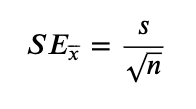
## Standard Error of the Mean

The standard error of the mean (SEM) is derived by dividing the standard deviation of the dataset by the square root of the number of observations (*n*). This metric indicates how much the sample mean is expected to vary from the true population mean.

For a Population



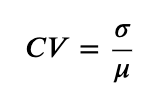
For a Sample



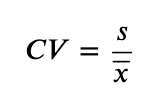
## Coefficient of Variation

The coefficient of variation (CV) measures the relative spread of data points around the mean, expressed as a ratio of the standard deviation to the mean. It's a useful statistic for comparing the degree of variability from one data series to another, even if the means are drastically different. The CV is calculated by dividing the standard deviation by the mean. This measure is particularly helpful in assessing the risk or variability in different contexts, such as finance and scientific research, where understanding relative dispersion is crucial.

For a Population



For a Sample



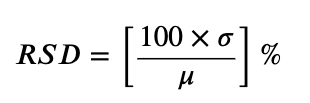
## Relative Standard Deviation

The relative standard deviation (RSD) quantifies the variation in a data set relative to its mean, presented as a percentage. It's computed by multiplying the standard deviation by 100 and then dividing by the mean. This statistic is valuable for comparing the variability of datasets with different units or means, providing a normalized measure of dispersion.

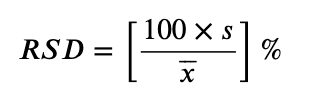
Top of Form

Bottom of Form

For a Population



For a Sample



## Frequency

Frequency measures how often each value appears in a dataset, essential for determining the mode, the value that occurs most frequently. This statistical concept helps in understanding the distribution and concentration of data points.

# These factors can significantly affect your central tendency measures.

# For example, a steadily increasing trend could make the mean over the entire series less representative of any specific point in time.

# Outliers: Especially with mean calculations, consider the impact of outliers.

# If your time series data includes extreme values, they could skew the mean.

# As can be seen in the described statistics, column Arrivals Air got close to 0