

**Variance:** variance measures the average squared deviation of each data point from the mean of the dataset.

**How** - The way you find variance is by calculating the mean of a data set, Subtracting the mean from each data point to find the deviation of each point, Square each deviation to eliminate negative values, and Calculate the average of these squared deviations to get the result as the variance.


**Why** - Variance is more commonly used in theoretical statistics, especially in the analysis of variance (ANOVA).

**Standard Deviation** - Standard deviation is the square root of the variance. It provides a measure of the spread of data points in the same units as the data itself, making it more interpretable.

**How** - Calculate the variance as described above and Take the square root of the variance to get the standard deviation.

**Why** - Standard Deviation is more commonly used in descriptive statistics to describe the spread of data.

**Formulas:**



	Population	Sample
Variance	$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$	$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$
Standard deviation	$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$	$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$

**N or n:**

- **N** is the total number of data points in the population.
- **n** is the total number of data points in the sample.

**xi:**

- **xi** represents each individual data point in the dataset.

**μ** (Population Mean) or (**Sample Mean**):

- **μ** is the mean (average) of the entire population.

**xi - μ or xi - sample mean (Deviation):**

- This represents the difference between each data point and the mean. It shows how much each data point deviates from the mean.

$(x_i - \mu)^2$  or  $(x_i - \text{sample mean})^2$  (Squared Deviation):

- Each deviation is squared to eliminate negative values, ensuring that all differences contribute positively to the total variance. Squaring also gives more weight to larger deviations, making the variance more sensitive to outliers.

(Sum of Squared Deviations):

- The squared deviations for all data points are summed to get a total measure of deviation in the dataset.

$1/N$  or  $1/n-1$  (Averaging the Squared Deviations):

- For the population variance, you divide the sum of squared deviations by  $N$ , the number of data points.
- For the sample variance, you divide the sum by  $n-1$  instead of  $n$ . This adjustment, called Bessel's correction, corrects the bias in the estimation of the population variance from a sample.