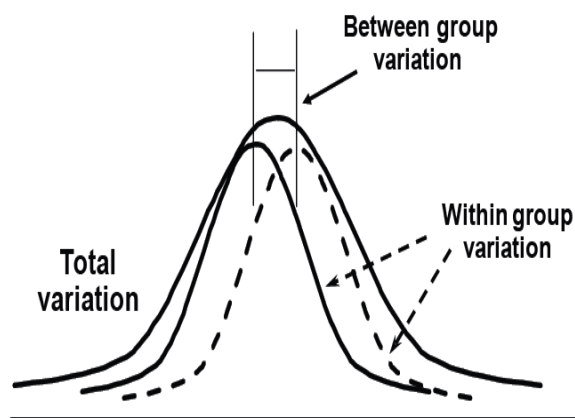


# Analysis of Variance (ANOVA) Documentation

## Analysis of Variance Definition :-

- Analysis of variance (ANOVA) is an analysis tool used in statistics that splits an observed aggregate variability found inside a data set into two parts:
  - ✓ Systematic factors
  - ✓ Random factors.
- The systematic factors have a statistical influence on the given data set, while the random factors do not. Analysts use the ANOVA test to determine the influence that independent variables have on the dependent variable in a regression study.
- The t- and z-test methods developed in the 20th century were used for statistical analysis until 1918, when Ronald Fisher created the analysis of variance method.
- ANOVA is also called the Fisher analysis of variance, and it is the extension of the t- and z-tests.
- The term became well-known in 1925, after appearing in Fisher's book, "Statistical Methods for Research Workers.
- It was employed in experimental psychology and later expanded to subjects that were more complex.

## Diagrammatic Representation of ANOVA :-



# Analysis of Variance (ANOVA) Documentation

## Formula for ANOVA :-

$$F = \text{MSE}/\text{MST}$$

Where

- ✓ F = ANOVA coefficient
- ✓ MST = Mean sum of squares due to treatment
- ✓ MSE = Mean sum of squares due to error

## What Does the Analysis of Variance Reveal :-

- ✚ The ANOVA test is the initial step in analyzing factors that affect a given data set. Once the test is finished, an analyst performs additional testing on the methodical factors that measurably contribute to the data set's inconsistency.
- ✚ The analyst utilizes the ANOVA test results in an f-test to generate additional data that aligns with the proposed regression models.
- ✚ The ANOVA test allows a comparison of more than two groups at the same time to determine whether a relationship exists between them.
- ✚ The result of the ANOVA formula, the F statistic (also called the F-ratio), allows for the analysis of multiple groups of data to determine the variability between samples and within samples.
- ✚ If no real difference exists between the tested groups, which is called the null hypothesis, the result of the ANOVA's F-ratio statistic will be close to 1.
- ✚ The distribution of all possible values of the F statistic is the F-distribution.
- ✚ This is actually a group of distribution functions, with two characteristic numbers, called the numerator degrees of freedom and the denominator degrees of freedom.

## Example of how to use ANOVA :-

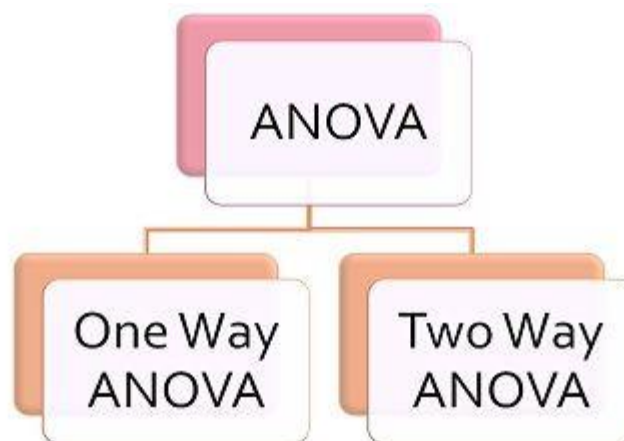
- A researcher might, for example, test students from multiple colleges to see if students from one of the colleges consistently outperform students from the other colleges.

# Analysis of Variance (ANOVA) Documentation

- In a business application, an R&D researcher might test two different processes of creating a product to see if one process is better than the other in terms of cost efficiency.
- The type of ANOVA test used depends on a number of factors. It is applied when data needs to be experimental.
- Analysis of variance is employed if there is no access to statistical software resulting in computing ANOVA by hand.
- It is simple to use and best suited for small samples.
- With many experimental designs, the sample sizes have to be the same for the various factor level combinations.
- ANOVA is helpful for testing three or more variables. It is similar to multiple two-sample t-tests.
- However, it results in fewer type I errors and is appropriate for a range of issues.
- ANOVA groups differences by comparing the means of each group and includes spreading out the variance into diverse sources.
- It is employed with subjects, test groups, between groups and within groups.

## Types of ANOVA :-

- ✚ ANOVA can be classified into two types they are listed below.



- ❖ One-Way ANOVA
- ❖ Two-Way ANOVA

# Analysis of Variance (ANOVA) Documentation

## One-Way ANOVA :-

- ✓ One-way or two-way refers to the number of independent variables in your analysis of variance test.
- ✓ A one-way ANOVA evaluates the impact of a sole factor on a sole response variable.
- ✓ It determines whether all the samples are the same.
- ✓ The one-way ANOVA is used to determine whether there are any statistically significant differences between the means of three or more independent (unrelated) groups.

## Formula For One-Way ANOVA :-

$$MS_b = \frac{SS_b}{K - 1}$$

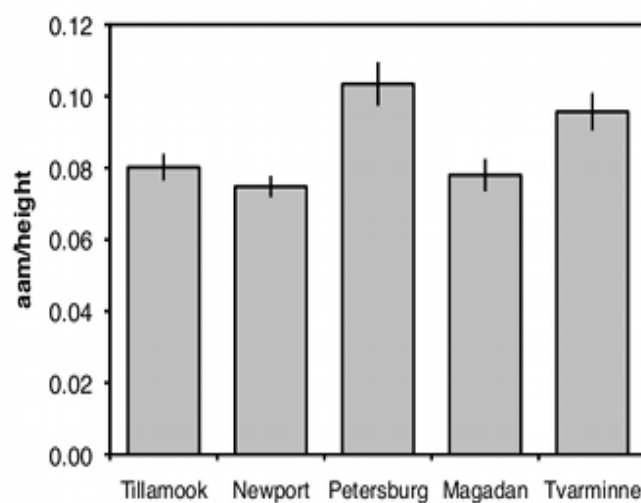
$MS_b$  = Mean Squares between groups

$SS_b$  = Sum of Squares between groups

$N$  = Number of cases combined across all groups

$K$  = Number of groups

## Diagram for One-Way ANOVA :-



# Analysis of Variance (ANOVA) Documentation

## Two-Way ANOVA :-

- ✓ A two-way ANOVA is an extension of the one-way ANOVA.
- ✓ With a one-way, you have one independent variable affecting a dependent variable.
- ✓ With a two-way ANOVA, there are two independents.
- ✓ For example, a two-way ANOVA allows a company to compare worker productivity based on two independent variables, such as salary and skill set.
- ✓ It is utilized to observe the interaction between the two factors and tests the effect of two factors at the same time.

## Formula For Two-Way ANOVA :-

$$SS = \sum (x - \bar{x})^2$$
$$MS = \frac{\sum (x - \bar{x})^2}{df}$$

$\bar{x}$  = measurement  
 $\bar{x}$  = group mean  
 $df$  = degrees of freedom

## Diagram for Two-Way ANOVA :-

