ANOVA Test

ANOVA (Analysis Of Variance):

* An ANOVA test is a type of statistical test used to determine if there is a statistically significant difference between two or more categorical groups by testing for differences of means using variance.
* Another Key part of ANOVA is that it splits the independent variable into 2 or more groups. For example, one or more groups might be expected to influences the dependent variable while the other group is used as a control group, and is not expected to influence the dependent variable.

## Assumptions of ANOVA

The assumptions of the ANOVA test are the same as the general assumptions for any parametric test:

* *An ANOVA can only be conducted if there is****no relationship between the subjects****in each sample. This means that subjects in the first group cannot also be in the second group (e.g. independent samples/between-groups).*
* *The different groups/levels must have****equal sample sizes****.*
* *An ANOVA can only be conducted if the dependent variable is****normally distributed****, so that the middle scores are most frequent and extreme scores are least frequent.*
* *Population variances must be equal (i.e. homoscedastic).****Homogeneity of variance****means that the deviation of scores (measured by the range or standard deviation for example) is similar between populations.*

## Types of ANOVA Tests:

There are different types of ANOVA tests. The two most common are a “One-Way” and a “Two-Way.”

The difference between these two types depends on the number of independent variables in your test.

### One-way ANOVA

* *A one-way ANOVA (analysis of variance) has one categorical independent variable (also known as a factor) and a normally distributed continuous (i.e., interval or ratio level) dependent variable.*
* *The independent variable divides cases into two or more mutually exclusive levels, categories, or groups.*
* *The one-way ANOVA test for differences in the means of the dependent variable is broken down by the levels of the independent variable.*
* *An example of a one-way ANOVA includes testing a therapeutic intervention (CBT, medication, placebo) on the incidence of depression in a clinical sample.*
* ***Note****: Both the One-Way ANOVA and the Independent Samples t-Test can compare the means for two groups. However, only the One-Way ANOVA can compare the means across three or more groups.*

. A [one-way ANOVA](https://www.statology.org/one-way-anova/) (“analysis of variance”) compares the means of three or more independent groups to determine if there is a statistically significant difference between the corresponding population means.

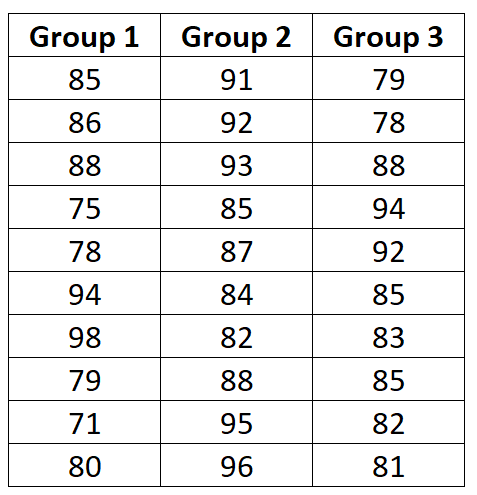
This tutorial explains how to perform a one-way ANOVA by hand.

### **Example: One-Way ANOVA by Hand**

Suppose we want to know whether or not three different exam prep programs lead to different mean scores on a certain exam. To test this, we recruit 30 students to participate in a study and split them into three groups.

The students in each group are randomly assigned to use one of the three exam prep programs for the next three weeks to prepare for an exam. At the end of the three weeks, all of the students take the same exam.

The exam scores for each group are shown below:



Use the following steps to perform a one-way ANOVA by hand to determine if the mean exam score is different between the three groups:

**Step 1: Calculate the group means and the overall mean.**

First, we will calculate the mean for all three groups along with the overall mean:

**Step 2: Calculate SSR.**

Next, we will calculate the regression sum of squares (SSR) using the following formula:

nΣ(Xj – X..)2

where:

* **n**: the sample size of group j
* **Σ**: a greek symbol that means “sum”
* **Xj**: the mean of group j
* **X..**: the overall mean

**Step 3: Calculate SSE.**

Next, we will calculate the error sum of squares (SSE) using the following formula:

Σ(Xij – Xj)2

where:

* **Σ**: a greek symbol that means “sum”
* **Xij**: the ith observation in group j
* **Xj**: the mean of group j

**Step 4: Calculate SST.**

SST = SSR + SSE

**Step 5: Fill in the ANOVA table.**

Here is how we calculated the various numbers in the table:

* **df treatment:**k-1 = 3-1 = 2
* **df error:**n-k = 30-3 = 27
* **df total:**n-1 = 30-1 = 29
* **MS treatment:**SST / df treatment = 192.2 / 2 = 96.1
* **MS error:**SSE / df error = 1100.6 / 27 = 40.8
* **F:**MS treatment / MS error = 96.1 / 40.8 = 2.358

***Note:****n = total observations, k = number of groups*

**Step 6: Interpret the results.**

The F test statistic for this one-way ANOVA is **2.358**. To determine if this is a statistically significant result, we must compare this to the F critical value found in the [F distribution table](https://www.statology.org/f-distribution-table/) with the following values:

* α (significance level) = 0.05
* DF1 (numerator degrees of freedom) = df treatment = 2
* DF2 (denominator degrees of freedom) = df error = 27

We find that the F critical value is **3.3541**.

Since the F test statistic in the ANOVA table is less than the F critical value in the F distribution table, we fail to reject the null hypothesis. This means we don’t have sufficient evidence to say that there is a statistically significant difference between the mean exam scores of the three groups.

### Two-way (factorial) ANOVA

* A two-way ANOVA (analysis of variance) has two or more categorical independent variables (also known as a factor), and a normally distributed continuous (i.e., interval or ratio level) dependent variable.
* The independent variables divide cases into two or more mutually exclusive levels, categories, or groups. A two-way ANOVA is also called a factorial ANOVA.
* An example of a factorial ANOVAs include testing the effects of social contact (high, medium, low), job status (employed, self-employed, unemployed, retired), and family history (no family history, some family history) on the incidence of depression in a population.