

# Learn How to Perform a One-Way ANOVA Test in Python

Authored by  
**Mohammed loot**

November 8, 2025

## RECOMMENDED CITATION

Mohammed loot (2025). *Learn How to Perform a One-Way ANOVA Test in Python*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=12691>

The [Analysis of Variance \(ANOVA\)](#) stands as a cornerstone statistical methodology used extensively for comparing the central tendencies, or means, of multiple distinct groups. Specifically, the One-Way ANOVA is a robust hypothesis test designed to evaluate whether there is a [statistically significant difference](#) among the average values derived from three or more independent samples, all stemming from a single categorical factor. This method offers a powerful alternative to performing multiple pairwise t-tests, which would increase the likelihood of committing a Type I error (false positive). By analyzing the total variation observed in the data and partitioning it into systematic (between-group) and random (within-group) components, ANOVA provides a comprehensive assessment of the factor's overall effect.

This comprehensive guide provides an in-depth, practical walkthrough detailing the implementation, execution, and interpretation of the [One-Way ANOVA](#) within the powerful Python scientific ecosystem. We will primarily leverage the optimized statistical functions available within the [SciPy library](#), a package essential for scientific computing and advanced data analysis in Python. Mastering this particular technique is fundamental for data scientists, researchers, and analysts who frequently handle experimental data and observational studies that require the simultaneous comparison of multiple conditions or treatments.

## Understanding the Theoretical Foundation

The necessity for employing ANOVA becomes clear when a researcher aims to assess the influence of an independent variable--or factor--that possesses three or more separate levels or categories. Consider a scenario where an agricultural scientist is studying the effect of three different fertilizer compositions (levels A, B, and C) on the resulting crop yield. Here, the factor is the "fertilizer type," and the quantitative outcome being measured is the dependent variable, "crop yield." The ANOVA test evaluates whether the differences in yield observed across the different fertilizer groups are substantially larger than the natural, inherent variability that occurs among crops treated with the same fertilizer.

Before any statistical comparison can be made, it is imperative to verify that the collected data adheres to the core assumptions underlying the One-Way ANOVA. These critical assumptions include the **\*\*independence of observations\*\*** (scores in one group must not influence scores in another), the **\*\*normal distribution\*\*** of the dependent variable within each group, and the **\*\*homogeneity of variances\*\***, which postulates that the variance (or spread) of scores is roughly equivalent across all treatment groups. Violations of these assumptions, especially the homogeneity requirement, can severely compromise the accuracy and reliability of the resulting statistical output, specifically the [F-statistic](#) and its associated [p-value](#). Although ANOVA is known to be relatively robust against minor deviations from normality, severe breaches might necessitate the use of specialized non-parametric statistical alternatives.

The mechanism of the ANOVA test involves calculating the F-ratio, which is the ratio of the variance explained by the differences between the sample means (Mean Square Between, MSB) to the variance unexplained, or error variance, within the samples (Mean Square Within, MSW). If the null hypothesis--which asserts that all population means are equal--is true, this ratio should ideally be close to 1. Conversely, a significantly larger F-statistic suggests that the differences observed between the group averages are too large to be merely attributed to random chance or sampling error. A sufficiently large F-ratio, combined with a small p-value, provides the necessary statistical evidence to reject the null hypothesis, thereby concluding that at least one of the treatments caused a statistically significant change in the measured outcome.

## Setting up the Educational Psychology Experiment

To demonstrate the practical application of the [One-Way ANOVA](#) in a controlled setting, we will examine a typical experimental scenario in educational psychology. Imagine a researcher enrolling 30 students to assess the efficacy of three distinct study methods. These students are randomly assigned to one of three groups: Technique A, Technique B, or Technique C. This crucial step of random assignment ensures that any pre-existing differences among students are evenly distributed across the groups, minimizing the risk of confounding variables influencing the final results. After a three-week preparation period utilizing their assigned technique, all 30 students take an identical standardized course exam, providing a quantifiable score for performance.

The primary analytical objective is to determine whether the average exam scores across the three groups are statistically equivalent. If the study techniques genuinely have varying levels of effectiveness, we expect the population means of the scores to differ significantly. Conversely, if the techniques are equally effective, or if the observed differences in the sample means are simply random fluctuations inherent to student performance, then we should conclude that the population means are equal. The subsequent steps guide us through using Python to formally test this research question against the collected sample data, moving from raw data preparation to final statistical inference.

## Data Organization and Structure in Python

The initial and most fundamental step in conducting this statistical test is organizing the raw experimental data accurately. When utilizing the standard `f_oneway` function provided by SciPy, the data must be structured such that the quantitative scores for each independent group are stored in separate sequences. These sequences are typically represented as Python lists or NumPy arrays. This specific requirement allows the statistical function to directly compare the variance within each group against the variance observed across the groups. It is crucial to ensure that every score is correctly assigned to the technique used by the corresponding student to maintain the integrity of the experimental design.

For our example, we have collected 10 exam scores for each of the three study techniques (Group 1, Group 2, and Group 3), resulting in a total sample size of 30 observations. We will represent these specific exam scores using standard Python lists, clearly labeling each list according to the study technique it represents. While the Pandas library is often used for complex data manipulation, for a simple comparison of means using SciPy's specialized function, defining three separate arrays is the most efficient and straightforward approach.

**Step 1: Enter the data.** We now define and populate the Python variables that hold the quantitative dependent variable scores (exam performance), categorized by the levels of the categorical independent variable (study technique).

```
#enter exam scores for each group
```

```
group1 =
```

```
group2 =
```

```
group3 =
```

## Executing the SciPy `f_oneway` Test

Once the data is correctly structured into separate arrays, the next phase involves invoking the statistical test itself. The [SciPy library](#) provides the highly specialized `f_oneway` function, which is precisely engineered for performing the One-Way ANOVA. This function accepts an arbitrary number of array arguments--each representing one of the independent group samples--and efficiently returns the two crucial metrics needed for hypothesis testing: the calculated F-statistic and the corresponding probability value (p-value).

To use this function, we must first explicitly import it from the `scipy.stats` module. This modular import strategy is characteristic of robust Python statistical packages, ensuring that the programming environment is loaded only with the necessary statistical tools. Once imported, we pass our three data arrays--`group1`, `group2`, and `group3`--as sequential positional arguments to the function. The `f_oneway` function then handles all the complex internal computations, which include calculating the degrees of freedom, determining the sums of squares, and ultimately deriving the final F-ratio.

**Step 2: Perform the one-way ANOVA.** We now apply the `f_oneway` function from the [SciPy library](#) to execute the One-Way ANOVA test on our three independent samples. The resulting output tuple immediately provides the necessary statistical metrics required for proceeding with hypothesis testing.

```
from scipy.stats import f_oneway
```

```
#perform one-way ANOVA
```

```
f_oneway(group1, group2, group3)  
  
(statistic=2.3575, pvalue=0.1138)
```

The execution of the function yields a tuple containing the F-statistic (2.3575) and the p-value (0.1138). The [F-statistic](#) serves to quantify the magnitude of the difference observed between the group means relative to the inherent variability existing within those groups. Conversely, the [p-value](#) represents the calculated probability of observing an F-statistic as extreme as, or more extreme than, the one computed, assuming that the [null hypothesis](#) is true. These two values form the sole basis upon which we make our formal statistical decision regarding the comparative effectiveness of the study techniques.

## Interpreting Results and Drawing Conclusions

The concluding step of the ANOVA process requires interpreting the statistical output in direct relation to the initial research question. Every [hypothesis](#) test begins with the formal establishment of two competing statements: the null hypothesis ( $H_0$ ) and the alternative hypothesis ( $H_1$ ). These hypotheses provide the critical framework for objectively determining whether the differences observed in the sample data are genuine effects or merely the result of random chance.

For a One-Way ANOVA, the null and alternative [hypotheses](#) are explicitly defined as follows:

**$H_0$  (Null Hypothesis):**  $\mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$ . This asserts that all population means associated with the study techniques are equal, meaning the techniques have no differential effect on exam scores.

**$H_1$  (Alternative Hypothesis):** At least one population mean is different from the rest. This asserts that at least one study technique leads to a significantly different average exam score compared to the others.

From the Python execution, we derived an F-statistic of **2.3575** and the corresponding p-value of **0.1138**. To render a statistical decision, we must compare this p-value against a predetermined threshold known as the significance level, or alpha ( $\alpha$ ), which is almost universally set at 0.05 (representing a 5% risk tolerance for Type I error). The established rule for decision-making is straightforward: if the calculated p-value is less than  $\alpha$  (0.05), we reject the null hypothesis; if the p-value is greater than  $\alpha$ , we fail to reject the null hypothesis.

In this specific case, the calculated [p-value](#) (0.1138) is clearly larger than the standard significance level of 0.05. Consequently, we must fail to reject the [null hypothesis](#) ( $H_0$ ). This is a vital statistical conclusion: it indicates that the variations observed in the average exam scores among the three study groups are most likely attributable to natural, random sampling variability among the students, rather than being caused by the distinct study techniques themselves. Based on this

sample evidence, we do not possess sufficient statistical proof to conclude that there is a [statistically significant difference](#) in exam performance among the three studying techniques when evaluated at the 95% confidence level.

### **Next Steps: Post-Hoc Analysis**

To further solidify your expertise in ANOVA and related statistical testing in Python, we strongly recommend reviewing the official documentation for the [SciPy stats module](#). While the One-Way ANOVA tells us whether an overall difference exists among the group means, it does not specify which particular pairs of groups are significantly different from one another. If our calculated F-statistic had been significant (i.e.,  $p < 0.05$ ), the necessary subsequent analytical step would involve performing post-hoc tests, such as Tukey's Honestly Significant Difference (HSD) test. These follow-up tests are essential for pinpointing the exact location of the statistical difference.