

Learn How to Conduct a Two-Way ANOVA in Python

Authored by
Mohammed looti

November 8, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *Learn How to Conduct a Two-Way ANOVA in Python*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=12690>

The Foundation of Two-Way Analysis of Variance (ANOVA)

The [Two-Way ANOVA](#), or Analysis of Variance, is an essential tool in [inferential statistics](#), designed specifically for analyzing experiments where two distinct categorical independent variables--known as **factors**--may influence a continuous dependent variable, often referred to as the [response variable](#). This method significantly advances beyond the simpler One-Way ANOVA by allowing researchers to simultaneously evaluate the impact of both factors and, critically, the way these two factors might interact with one another. The core utility of the Two-Way ANOVA lies in its ability to determine if a [statistically significant difference](#) exists among the means of groups that are defined by the unique combinations of the levels of these two factors.

A Two-Way ANOVA addresses three primary statistical questions simultaneously. First, we examine the **main effect** of the first factor, which assesses its influence on the response variable irrespective of the second factor's levels. Second, we quantify the **main effect** of the second factor, disregarding the first. The third, and often most revealing, objective is the investigation of the **interaction effect**. This interaction term reveals whether the effect of Factor A on the response variable changes depending on the specific level of Factor B. For example, if Factor A only produces a strong effect when Factor B is present at its highest setting, a significant interaction is detected. Discovering this interaction is vital because it uncovers complex, synergistic relationships that simple analyses focusing only on individual factors would invariably overlook.

To perform this sophisticated analysis using **Python**, researchers rely on specialized statistical packages. This tutorial centers on the use of the [statsmodels](#) library, a comprehensive framework built for fitting various statistical models, including those derived from the General Linear Model (GLM). By leveraging the intuitive R-like formula API within **statsmodels**, we can clearly define the relationship between our factors and the response variable, allowing the library to efficiently compute the necessary sums of squares, F-statistics, and [p-values](#) required to draw rigorous conclusions about the existence and magnitude of the main and interaction effects.

Defining the Experimental Design and Hypotheses

To solidify the concepts of the Two-Way ANOVA, let us consider a classic experimental setup. Imagine a botanist conducting research to understand how different environmental conditions affect the growth rate of a specific plant species. The botanist selects two primary categorical factors: **watering frequency** and **sunlight exposure**. The watering factor has two levels (daily or weekly), and the sunlight factor has three levels (low, medium, or high). The quantitative outcome, or [response variable](#), being measured is the height of each plant, recorded in inches, after an observational period of two months.

This design results in a total of six unique experimental groups (2 Water levels × 3 Sun levels = 6 groups). If the experiment involves 30 plants, each of these six groups contains 5 replicates. This

structured, balanced design is ideal for the Two-Way ANOVA, which is employed to test three distinct sets of formal null hypotheses (H_0) derived directly from the experimental setup. These hypotheses cover the main effect of each factor and their combined interaction:

H_0 (Water): There is no [statistically significant difference](#) in the mean plant height observed between plants watered daily versus those watered weekly (The Main Effect of Water is zero).

H_0 (Sun): There is no statistically significant difference in the mean plant height across the three levels of sunlight exposure (low, medium, high) (The Main Effect of Sun is zero).

H_0 (Interaction): There is no interaction effect; the effect of watering frequency on plant height does not depend on the specific level of sunlight exposure, meaning the effects are purely additive (The Interaction Effect of Water $\times$ Sun is zero).

The subsequent analysis uses the collected height data to calculate the F-statistics and associated [p-values](#) for each of these three effects. By comparing these p-values to a predetermined significance level (typically $\alpha = 0.05$), we can rigorously determine whether watering, sunlight, or the combination of the two, exerts a significant influence on the final height of the plants. This process allows the botanist to confidently reject or fail to reject the stated null hypotheses, transforming raw measurements into actionable scientific conclusions.

Step 1: Preparing the Data Structure in Python

A prerequisite for any robust statistical analysis in [Python](#) is the proper preparation and organization of the raw experimental data. We rely heavily on the **pandas** library, which is indispensable for structuring data into a manageable format. For statistical modeling, data must be structured in a 'long' format, where every row represents a single observation (a single plant), and columns are dedicated to the categorical factors and the continuous response variable.

Our **Pandas DataFrame** requires three key variables to proceed with the ANOVA:

water: The categorical factor indicating the watering regime ('daily' or 'weekly').

sun: The categorical factor defining the light exposure level ('low', 'med', or 'high').

height: The continuous, numerical measurement (in inches) that serves as our [response variable](#).

The following code snippet uses **NumPy** for efficient array generation (specifically `np.repeat` and `np.tile`) to accurately simulate the balanced experimental design involving 30 plants. This ensures that the data structure perfectly reflects the design described in the scenario. Once the **Pandas DataFrame** is constructed, viewing the initial rows confirms that the categorical factor levels are correctly aligned with the measured height outcomes, ensuring the data is ready for input into the **statsmodels** package.

```
import numpy as np
```

```
import pandas as pd
```

```
#create data
```

```
df = pd.DataFrame({'water': np.repeat(, 15),
```

```
'sun': np.tile(np.repeat(, 5), 2),
```

```
'height': })
```

```
#view first ten rows of data
```

```
df
```

```
water sun height
```

```
0 daily low 6
```

```
1 daily low 6
```

```
2 daily low 6
```

```
3 daily low 5
```

```
4 daily low 6
```

```
5 daily med 5
```

```
6 daily med 5
```

```
7 daily med 6
```

```
8 daily med 4
```

```
9 daily med 5
```

Step 2: Implementing and Fitting the Two-Way ANOVA Model

With the data correctly organized within a [Pandas DataFrame](#), the next essential step is to execute the statistical model. We utilize the [Ordinary Least Squares \(OLS\)](#) function, imported from `statsmodels.formula.api`, to define our ANOVA, treating it as a specialized case of the General Linear Model. The model definition relies on a powerful and concise R-like formula syntax, which simplifies the specification of complex factor relationships.

The core of the execution lies in the formula string: `'height ~ C(water) + C(sun) + C(water):C(sun)'`. This string directs the [statsmodels](#) OLS estimator to model the outcome variable, **height**. The crucial element is the use of the `C()` wrapper around `water` and `sun`; this explicitly informs **statsmodels** that these variables are categorical **factors**, ensuring they are correctly coded using indicator variables for the ANOVA calculation. The terms `C(water)` and `C(sun)` capture the independent **main effects** of the factors.

The term `C(water):C(sun)` is the defining characteristic of the Two-Way design. It represents the **interaction effect**, testing whether the unique combination of watering frequency and sunlight exposure produces a synergistic effect on height that cannot be explained simply by summing the

individual main effects. After fitting the model object using `.fit()`, we generate the final ANOVA table using `sm.stats.anova_lm(model, typ=2)`. We specifically request Type II Sums of Squares (`typ=2`) because it provides a conservative and highly reliable method for assessing main effects in factorial designs, especially when those designs might involve minor or anticipated imbalances, though it remains robust even in our balanced scenario.

```
import statsmodels.api as sm
```

```
from statsmodels.formula.api import ols
```

```
#perform two-way ANOVA
```

```
model = ols('height ~ C(water) + C(sun) + C(water):C(sun)', data=df).fit()
```

```
sm.stats.anova_lm(model, typ=2)
```

```
sum_sq df F PR(>F)
```

```
C(water) 8.533333 1.0 16.0000 0.000527
```

```
C(sun) 24.866667 2.0 23.3125 0.000002
```

```
C(water):C(sun) 2.466667 2.0 2.3125 0.120667
```

```
Residual 12.800000 24.0 NaN NaN
```

Step 3: Interpreting the Statistical Output

The ANOVA table produced by **statsmodels** contains all the necessary data points--Sum of Squares (`sum_sq`), Degrees of Freedom (`df`), the F-statistic (`F`), and the crucial [p-value](#) (`PR(>F)`)--to test the three hypotheses. Interpretation is straightforward: we compare the `PR(>F)` value for each term against our predetermined significance level, $\alpha = 0.05$. If the p-value is less than 0.05, we reject the corresponding null hypothesis, concluding that the effect is statistically significant.

A systematic review of the results reveals the following conclusions for the botanist's experiment:

Water (Main Effect): The p-value for the watering factor is 0.000527. Since this value is considerably less than 0.05, we confidently reject H_0 for water. This finding confirms that **watering frequency** has a strong, [statistically significant effect](#) on the mean plant height. The difference in height between the daily-watered group and the weekly-watered group is highly unlikely to be the result of random sampling error.

Sun (Main Effect): The p-value for the sunlight factor is 0.000002 (or 2×10^{-6}). This result is exceptionally significant ($p < 0.05$). We must therefore reject H_0 for sunlight exposure, concluding that **sunlight exposure** significantly influences the mean plant height. This means that at least one of the three sunlight levels (low, medium, or high) produces a significantly different mean height compared to the others.

Water × Sun (Interaction Effect): The p-value for the interaction term, `C(water):C(sun)`, is 0.120667. As this value is greater than the 0.05 threshold, we fail to reject the null hypothesis for the interaction. This is a critical finding: it suggests that the effect of watering frequency on plant height remains consistent, regardless of whether the plant receives low, medium, or high sunlight. In other words, there is no significant synergistic **interaction** between these two factors.

In summary, the statistical model demonstrates that both **watering frequency** and **sunlight exposure** are potent, independent predictors of plant height. The large F-statistics (16.00 and 23.31) corresponding to these factors indicate that the variance explained by the differences in group means is substantially greater than the unexplained residual variance. Furthermore, the low residual Sum of Squares (12.80) relative to the total variance confirms that the model is a strong fit, successfully accounting for a large proportion of the observed variability in plant height.

Conclusion and Mandatory Post-Hoc Analysis

The Two-Way ANOVA successfully established that our two independent **factors**--watering frequency and sunlight exposure--each exert a statistically significant and independent influence on plant growth. However, it is essential to remember that ANOVA is an omnibus test. While it signals that differences exist among group means, it does not specify the location of those differences. This limitation requires researchers to proceed to further analysis steps, particularly for factors with more than two levels.

For the water factor, which only has two levels (daily vs. weekly), the significant ANOVA result already confirms that these two means are statistically different. However, for the sunlight factor, which has three levels (low, medium, high), we only know that the means are not all equal. To determine exactly which pairs of sunlight levels (e.g., low vs. medium, low vs. high, or medium vs. high) are significantly different, we must conduct **Post-Hoc Tests**.

Standard procedures for these pairwise comparisons include tests like **Tukey's Honestly Significant Difference (HSD)** or the **Bonferroni correction**. These specialized tests are necessary because they adjust the [p-value](#) threshold to mitigate the risk of Type I errors (false positives) that accumulate when performing multiple comparisons across the same data set. The botanist would calculate the mean height for each sunlight level and then apply a Post-Hoc test to pinpoint the specific contrasts that are [statistically significant](#). The absence of a significant interaction effect greatly simplifies this final interpretation, allowing the botanist to discuss the effects of water and sun separately, transforming the general statistical finding into precise, actionable ecological knowledge.