

Perform a Three-Way ANOVA in R

Authored by
Mohammed looti

November 1, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *Perform a Three-Way ANOVA in R*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=7636>

The **Three-Way ANOVA** (Analysis of Variance) is a robust and sophisticated statistical technique employed when researchers need to assess the simultaneous influence of three distinct independent categorical variables (often referred to as factors) on a single continuous dependent variable. Unlike simpler tests, this method goes beyond merely comparing group means; it rigorously determines if there is a **statistically significant** difference among the averages of groups defined by the combined levels of these three factors. Crucially, the three-way ANOVA is essential for identifying potential interaction effects, where the effect of one factor on the outcome variable changes depending on the level of one or both other factors.

This expert guide offers a detailed, step-by-step walkthrough demonstrating the execution, analysis, and interpretation of a three-way ANOVA using the powerful and open-source statistical programming environment, **R**. We focus on providing clear code and detailed explanations to ensure accurate and reliable results.

Conceptualizing the Three-Way ANOVA Example in R

To illustrate the practical application of this test, we will use a hypothetical scenario involving athletic training research. Imagine a sports scientist aiming to maximize the improvement in vertical jump height among college basketball players. The primary focus is to compare the efficacy of two distinct training methods: Program 1 versus Program 2. The resulting change in jumping height, measured in inches, serves as our continuous dependent variable.

However, the researcher recognizes that training effectiveness might not be universal. They hypothesize that two inherent characteristics of the athletes--their **gender** (Male or Female) and their collegiate competitive level, or **division** (Division I or Division II)--may also significantly influence the outcome. Consequently, we have three categorical independent factors: Training Program (two levels), Gender (two levels), and Division (two levels).

The overarching objective of performing this three-way ANOVA is twofold: first, to assess the individual impact (main effects) of the training program, gender, and division on jump height improvement; and second, and perhaps more importantly, to determine if these factors interact. For instance, does Program 1 work best, but only for Division I male athletes? An interaction analysis is necessary to answer such nuanced questions.

Step 1: Structuring and Preparing the Data in R

The foundation of any robust statistical analysis is correctly organized data. For R, this often means constructing a well-defined **data frame**. In our example, we simulate a dataset containing 40 observations, capturing the measurements for the jump height improvement along with the levels of the three categorical factors: program, gender, and division.

The code snippet below is designed to generate this synthetic dataset, named `df`. It ensures that we have balanced groups (equal sample sizes across all factor combinations), a critical requirement for traditional ANOVA models. Following the creation of the data frame, we use the `head()` function to inspect the initial rows, confirming that the structure and variable types are correctly defined before proceeding to the analytical phase.

#create dataset

```
df <- data.frame(program=rep(c(1, 2), each=20),
gender=rep(c('M', 'F'), each=10, times=2),
division=rep(c(1, 2), each=5, times=4),
height=c(7, 7, 8, 8, 7, 6, 6, 5, 6, 5,
5, 5, 4, 5, 4, 3, 3, 4, 3, 3,
6, 6, 5, 4, 5, 4, 5, 4, 4, 3,
2, 2, 1, 4, 4, 2, 1, 1, 2, 1))
```

```
#view first six rows of dataset
```

```
head(df)
```

```
program gender division height
```

```
1 1 M 1 7
```

```
2 1 M 1 7
```

```
3 1 M 1 8
```

```
4 1 M 1 8
```

```
5 1 M 1 7
```

```
6 1 M 2 6
```

Step 2: Preliminary Data Exploration and Descriptive Statistics

Prior to executing any inferential statistical model, it is a crucial best practice to explore the data using descriptive statistics. This preliminary step allows the researcher to gain an intuitive understanding of the means and variances across different groups, potentially revealing obvious trends or disparities that the formal test will later confirm or refute. Descriptive analysis provides context, helping to frame the ultimate interpretation of the ANOVA output.

To calculate the mean jumping height for every possible combination of our three factors (resulting in $2 \times 2 \times 2 = 8$ distinct groups), we leverage the functionalities of the highly efficient R package, [dplyr](#), which is part of the tidyverse ecosystem. We group the data sequentially by program, gender, and division, and then use the `summarize()` function to calculate the mean of the `height` variable for each subgroup.

library(dplyr)

```
#calculate mean jumping height increase grouped by program, gender, and division
df %>%
group_by(program, gender, division) %>%
summarize(mean_height = mean(height))

# A tibble: 8 x 4
# Groups: program, gender
program gender division mean_height

1 1 F 1 4.6
2 1 F 2 3.2
3 1 M 1 7.4
4 1 M 2 5.6
5 2 F 1 2.6
6 2 F 2 1.4
7 2 M 1 5.2
8 2 M 2 4
```

This comprehensive output displays the mean improvement for all eight unique group combinations. For instance, the results suggest that Division I male athletes using Program 1 achieved the highest average improvement (7.4 inches), while Division II female athletes using Program 2 showed the lowest average increase (1.4 inches). While these descriptive statistics strongly suggest meaningful differences, only the formal ANOVA test can determine if these observed differences are statistically reliable and unlikely to be due merely to random chance.

Step 3 & 4: Model Execution and Comprehensive Interpretation

The core of the analysis involves running the ANOVA model in R using the built-in `aov()` function. The syntax used to specify the model is critical, especially when dealing with multiple factors and interactions. We define the model using the formula `height ~ program * gender * division`. The asterisk (*) notation in R is a powerful shorthand in modeling; it instructs the system to test not only for the main effect of each factor individually but also for every possible two-way interaction (e.g., program x gender, program x division) and the single three-way interaction (program x gender x division).

Once the model object (named `model`) is created, the `summary()` function generates the standard ANOVA table. This table is the primary output for interpretation, containing crucial metrics like Sum of Squares, F-values, Degrees of Freedom, and, most importantly, the p-values for each effect

being tested.

#perform three-way ANOVA

```
model <- aov(height ~ program * gender * division, data=df)
```

```
#view summary of three-way ANOVA
```

```
summary(model)
```

```
Df Sum Sq Mean Sq F value Pr(>F)
program 1 36.1 36.10 65.636 2.98e-09 ***
gender 1 67.6 67.60 122.909 1.71e-12 ***
division 1 19.6 19.60 35.636 1.19e-06 ***
program:gender 1 0.0 0.00 0.000 1.000
program:division 1 0.4 0.40 0.727 0.400
gender:division 1 0.1 0.10 0.182 0.673
program:gender:division 1 0.1 0.10 0.182 0.673
Residuals 32 17.6 0.55
```

```
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The interpretation hinges on the column labeled **Pr(>F)**, which presents the **p-value** for each effect. We compare these p-values against a predetermined significance level (alpha, typically 0.05). If the p-value falls below this threshold, the effect is deemed statistically significant, meaning we reject the null hypothesis that the means are equal.

Analyzing Interaction and Main Effects

The first step in interpreting a multifactor ANOVA is always to examine the highest-order interaction term--in this case, the three-way interaction (`program:gender:division`). If this term were significant ($p < 0.05$), it would mean the interpretation of the main effects is largely irrelevant, as the effect of one factor is entirely conditional on the specific combination of the other two.

In our results, the three-way interaction term yielded a p-value of 0.673, which is far greater than 0.05, indicating non-significance. We then proceed to examine the two-way interactions (`program:gender`, `program:division`, `gender:division`). All two-way interactions also resulted in high p-values (1.000, 0.400, and 0.673, respectively). The lack of statistically significant interactions is a crucial finding: it tells us that the impact of the training program on jump height does not depend on the athlete's gender or division level. The factors behave additively.

Since we found no significant interactions, we can confidently proceed to interpret the main effects:

Training Program: The p-value (2.98e-09) is extremely low, indicating a highly significant main effect. The choice of training program independently affects jump height improvement.

Gender: The p-value (1.71e-12) is also highly significant. Gender independently influences the amount of jump height increase.

Division: The p-value (1.19e-06) is highly significant. The competitive division level independently affects the improvement in jump height.

The highly significant main effects confirm that program, gender, and division all contribute unique variance to the model. However, the ANOVA table alone does not tell us which level within each factor performed better (e.g., was Program 1 better than Program 2?). To determine the direction and magnitude of these significant main effects, we must calculate the marginal means for each factor separately.

Step 5: Calculating Marginal Means for Clarity

To fully interpret the significant main effects, we use `dplyr` once more to calculate the average jump height increase across the levels of each factor, pooling the data across the other two factors. These marginal means provide the clearest picture of the independent effect of each variable.

library(dplyr)

```
#find mean jumping increase by program
df %>%
  group_by(program) %>%
  summarize(mean_height = mean(height))
```

```
# A tibble: 2 x 2
  program mean_height
```

```
1 1 5.2
2 2 3.3
```

```
#find mean jumping increase by gender
df %>%
  group_by(gender) %>%
  summarize(mean_height = mean(height))
```

```
# A tibble: 2 x 2
  gender mean_height
```

```
1 F 2.95
2 M 5.55
```

```
#find mean jumping increase by division
df %>%
group_by(division) %>%
summarize(mean_height = mean(height))

# A tibble: 2 x 2
  division mean_height
  <dbl>     <dbl>
1 1 4.95
2 2 3.55
```

By translating these quantitative outputs into practical findings, we arrive at clear and actionable conclusions for the sports researcher:

Training Program Effect: Program 1 is statistically superior, resulting in a significantly higher average jump height increase (**5.2 inches**) compared to Program 2 (**3.3 inches**).

Gender Effect: Male athletes showed a statistically greater mean increase (**5.55 inches**) than female athletes (**2.95 inches**), suggesting inherent biological or training adaptation differences based on gender.

Division Effect: Division I players achieved a statistically higher mean increase (**4.95 inches**) than Division II players (**3.55 inches**), potentially reflecting differences in baseline conditioning or overall athleticism.

Conclusion of the Three-Way ANOVA Analysis

In conclusion, the three-way ANOVA successfully evaluated the complex relationship between training program, gender, division, and jump height improvement. The analysis unequivocally confirmed that all three factors independently and significantly predict changes in athletic performance.

The most critical finding for the researcher is the absence of any significant interaction effects. This simplifies the practical recommendation: because the best training program (Program 1) is consistently superior regardless of the athlete's gender or division, the researcher can recommend Program 1 universally. The non-significant interaction implies that the effect of the program is additive to the effects of gender and division, rather than conditional upon them.

Further Statistical Resources

For those seeking to deepen their understanding of variance analysis and apply other models in R, the following tutorials provide guidance on related statistical techniques: