

How to Perform a Two-Way ANOVA in R: A Step-by-Step Guide

Authored by
stats writer

March 3, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Perform a Two-Way ANOVA in R: A Step-by-Step Guide*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=133653>

A two-way ANOVA (analysis of variance) is a statistical test used to analyze the differences between two or more groups across two or more independent variables. In order to conduct a two-way ANOVA in R, there are a few steps that need to be followed. First, the data must be organized in a specific format, with each group and variable in a separate column. Next, the "aov" function is used to fit the ANOVA model to the data. Finally, the "summary" function is used to obtain the results of the ANOVA, including the F-statistic, p-value, and any significant differences between groups. Additional steps may be necessary depending on the specific data and research question being investigated. It is important to have a clear understanding of the data and research question in order to correctly conduct and interpret a two-way ANOVA in R.

Conduct a Two-Way ANOVA in R

A ("analysis of variance") is used to determine whether or not there is a statistically significant difference between the means of three or more independent groups that have been split on two factors.

This tutorial explains how to perform a two-way ANOVA in R.

Example: Two-Way ANOVA in R

Suppose we want to determine if exercise intensity and gender impact weight loss. In this case, the two factors we're studying are *exercise* and *gender* and the response variable is *weight loss*, measured in pounds.

We can conduct a two-way ANOVA to determine if exercise and gender impact weight loss and to

determine if there is an interaction between exercise and gender on weight loss.

We recruit 30 men and 30 women to participate in an experiment in which we randomly assign 10 of each to follow a program of either no exercise, light exercise, or intense exercise for one month.

The following code creates the data frame we'll be working with:

```
#make this example reproducible  
set.seed(10)  
  
#create data frame  
data <- data.frame(gender = rep(c("Male", "Female"),  
each = 30),  
exercise = rep(c("None", "Light", "Intense"), each = 10,  
times = 2),  
weight_loss = c(runif(10, -3, 3), runif(10, 0, 5), runif(10, 5,  
9),  
runif(10, -4, 2), runif(10, 0, 3), runif(10, 3, 8)))  
  
#view first six rows of data frame  
head(data)
```

```
# gender exercise weight_loss
```

```
#1 Male None 0.04486922
```

```
#2 Male None -1.15938896
```

```
#3 Male None -0.43855400
```

```
#4 Male None 1.15861249
```

```
#5 Male None -2.48918419
```

```
#6 Male None -1.64738030
```

```
#see how many participants are in each group
```

```
table(data$gender, data$exercise)
```

```
# Intense Light None
```

```
# Female 10 10 10
```

```
# Male 10 10 10
```

Exploring the Data

Before we even fit the two-way ANOVA model, we can gain a better understanding of the data by finding the mean and standard deviation of weight loss for each of the six treatment groups using the `dplyr` package:

```
#load dplyr package
```

```
library(dplyr)
```

```
#find mean and standard deviation of weight loss for
```

each treatment group

data %>%

group_by(gender, exercise) %>%

**summarise(mean = mean(weight_loss),
sd = sd(weight_loss))**

A tibble: 6 x 4

Groups: gender

gender exercise mean sd

#

#1 Female Intense 5.31 1.02

#2 Female Light 0.920 0.835

#3 Female None -0.501 1.77

#4 Male Intense 7.37 0.928

#5 Male Light 2.13 1.22

#6 Male None -0.698 1.12

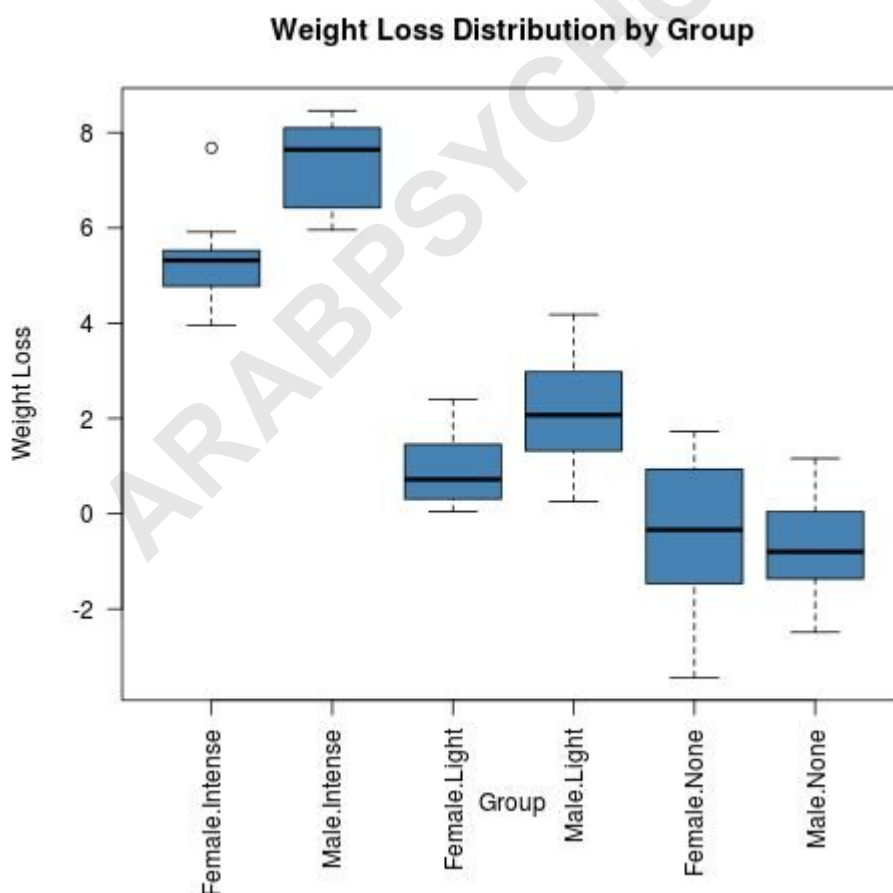
We can also create a for each of the six treatment groups to visualize the distribution of weight loss for each group:

#set margins so that axis labels on boxplot don't get cut off

par(mar=c(8, 4.1, 4.1, 2.1))

```
#create boxplots
```

```
boxplot(weight_loss ~ gender:exercise,  
data = data,  
main = "Weight Loss Distribution by Group",  
xlab = "Group",  
ylab = "Weight Loss",  
col = "steelblue",  
border = "black",  
las = 2 #make x-axis labels perpendicular  
)
```



Right away we can see that the two groups who participated in *intense* exercise appear to have greater weight loss values. We can also see that males tend to have higher weight loss values for the *intense* and *light* exercise groups compared to females.

Next, we'll fit the two-way ANOVA model to our data to see if these visual differences are actually statistically significant.

Fitting the Two-Way ANOVA Model

```
aov(response_variable ~ predictor_variable1 *  
predictor_variable2, data = dataset)
```

Note that the `*` between the two predictor variables indicates that we also want to test for an interaction effect between the two predictor variables.

In our example, we can use the following code to fit the two-way ANOVA model, using *weight_loss* as the response variable and *gender* and *exercise* as our two predictor variables.

We can then use the `summary()` function to view the output of our model:

```
#fit the two-way ANOVA model
```

```
model <- aov(weight_loss ~ gender * exercise, data =  
data)
```

```
#view the model output
```

```
summary(model)
```

```
# Df Sum Sq Mean Sq F value Pr(>F)
```

```
#gender 1 15.8 15.80 11.197 0.0015 **
```

```
#exercise 2 505.6 252.78 179.087 <2e-16 ***
```

```
#gender:exercise 2 13.0 6.51 4.615 0.0141 *
```

```
#Residuals 54 76.2 1.41
```

```
#---
```

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

From the model output we can see that *gender*, *exercise*, and the interaction between the two variables are all statistically significant at the .05 significance level.

Checking the Model Assumptions

Before we go any further, we should check to see that the assumptions of our model are met so that the our results from the model are reliable. In particular, a two-

way ANOVA assumes:

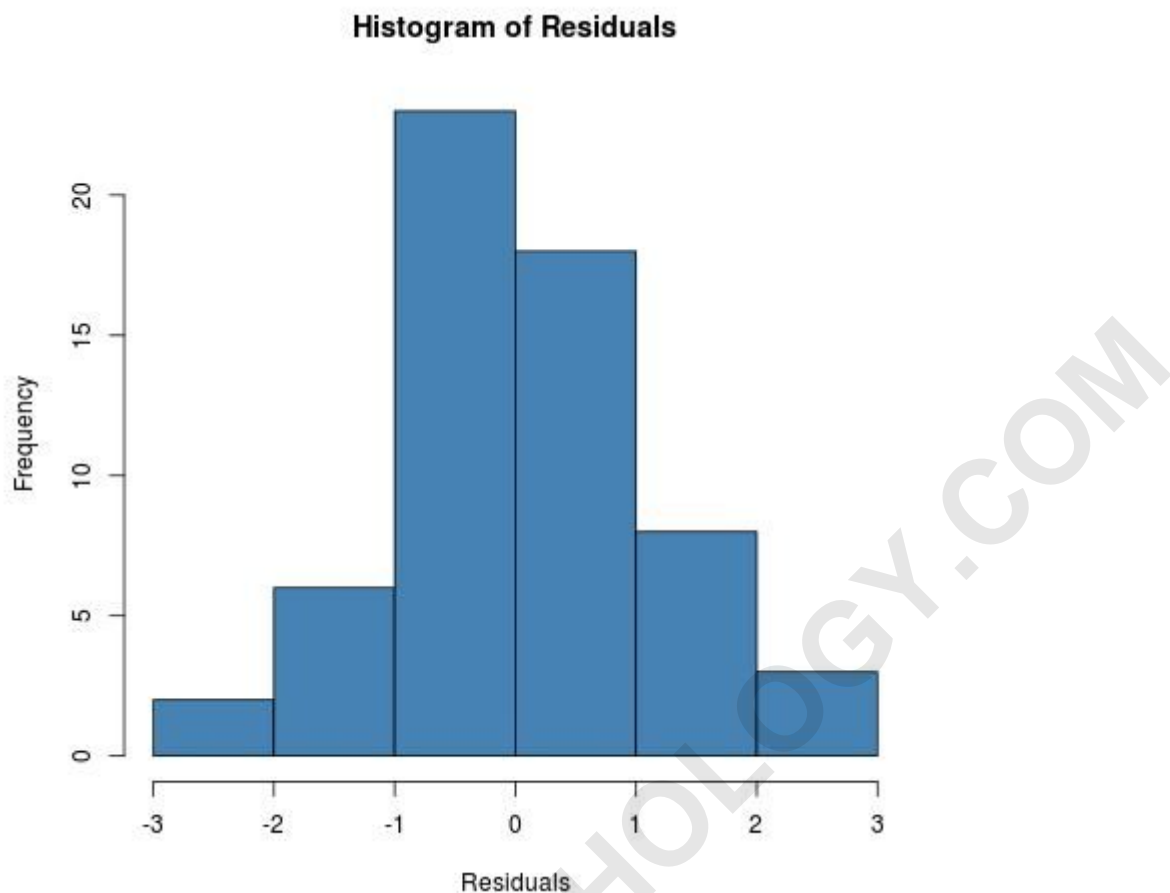
1. Independence - the observations in each group need to be independent of each other. Since we used a randomized design, this assumption should be met so we don't need to worry too much about this.

2. Normality - the dependent variable should be approximately normally distributed for each combination of the groups of the two factors.

One way to check this assumption is to create a histogram of the model residuals. If the residuals are roughly normally distributed, this assumption should be met.

#define model residuals

```
resid <- model$residuals#create histogram of residualshist(resid, main = Histogram of Residuals", xlab = "Residuals", col = "steelblue")
```



The residuals are roughly normally distributed, so we can assume the normality assumption is met.

3. Equal Variance - the variances for each group are equal or approximately equal.

One way to check this assumption is to conduct a Levene's Test for equality of variances using the `car` package:

```
#load car package
```

```
library(car)
```

```
#conduct Levene's Test for equality of variances
```

```
leveneTest(weight_loss ~ gender * exercise, data = data)
```

```
#Levene's Test for Homogeneity of Variance (center = median)
```

```
# Df F value Pr(>F)
```

```
#group 5 1.8547 0.1177
```

```
# 54
```

Since the p-value of the test is greater than our significance level of 0.05, we can assume that our assumption of equality of variances among groups is met.

Analyzing Treatment Differences

Once we have verified that the model assumptions are met, we can then conduct a to determine exactly which treatment groups differ from one another.

For our post hoc test, we will use the function `TukeyHSD()` to conduct Tukey's Test for multiple comparisons:

```
#perform Tukey's Test for multiple comparisons  
TukeyHSD(model, conf.level=.95)  
  
# Tukey multiple comparisons of means  
# 95% family-wise confidence level  
#  
#Fit: aov(formula = weight_loss ~ gender * exercise,  
data = data)  
#  
#$gender  
# diff lwr upr p adj  
#Male-Female 1.026456 0.4114451 1.641467 0.0014967  
#  
#$exercise  
# diff lwr upr p adj  
#Light-Intense -4.813064 -5.718493 -3.907635 0.0e+00  
#None-Intense -6.938966 -7.844395 -6.033537 0.0e+00  
#None-Light -2.125902 -3.031331 -1.220473 1.8e-06  
#  
#$`gender:exercise`  
# diff lwr upr p adj  
#Male:Intense-Female:Intense 2.0628297 0.4930588  
3.63260067 0.0036746  
#Female:Light-Female:Intense -4.3883563 -5.9581272
```

-2.81858535 0.0000000
#Male:Light-Female:Intense -3.1749419 -4.7447128
-1.60517092 0.0000027
#Female:None-Female:Intense -5.8091131 -7.3788841
-4.23934219 0.0000000
#Male:None-Female:Intense -6.0059891 -7.5757600
-4.43621813 0.0000000
#Female:Light-Male:Intense -6.4511860 -8.0209570
-4.88141508 0.0000000
#Male:Light-Male:Intense -5.2377716 -6.8075425
-3.66800066 0.0000000
#Female:None-Male:Intense -7.8719429 -9.4417138
-6.30217192 0.0000000
#Male:None-Male:Intense -8.0688188 -9.6385897
-6.49904786 0.0000000
#Male:Light-Female:Light 1.2134144 -0.3563565
2.78318536 0.2185439
#Female:None-Female:Light -1.4207568 -2.9905278
0.14901410 0.0974193
#Male:None-Female:Light -1.6176328 -3.1874037
-0.04786184 0.0398106
#Female:None-Male:Light -2.6341713 -4.2039422
-1.06440032 0.0001050
#Male:None-Male:Light -2.8310472 -4.4008181

```
-1.26127627 0.0000284  
#Male:None-Female:None -0.1968759 -1.7666469  
1.37289500 0.9990364
```

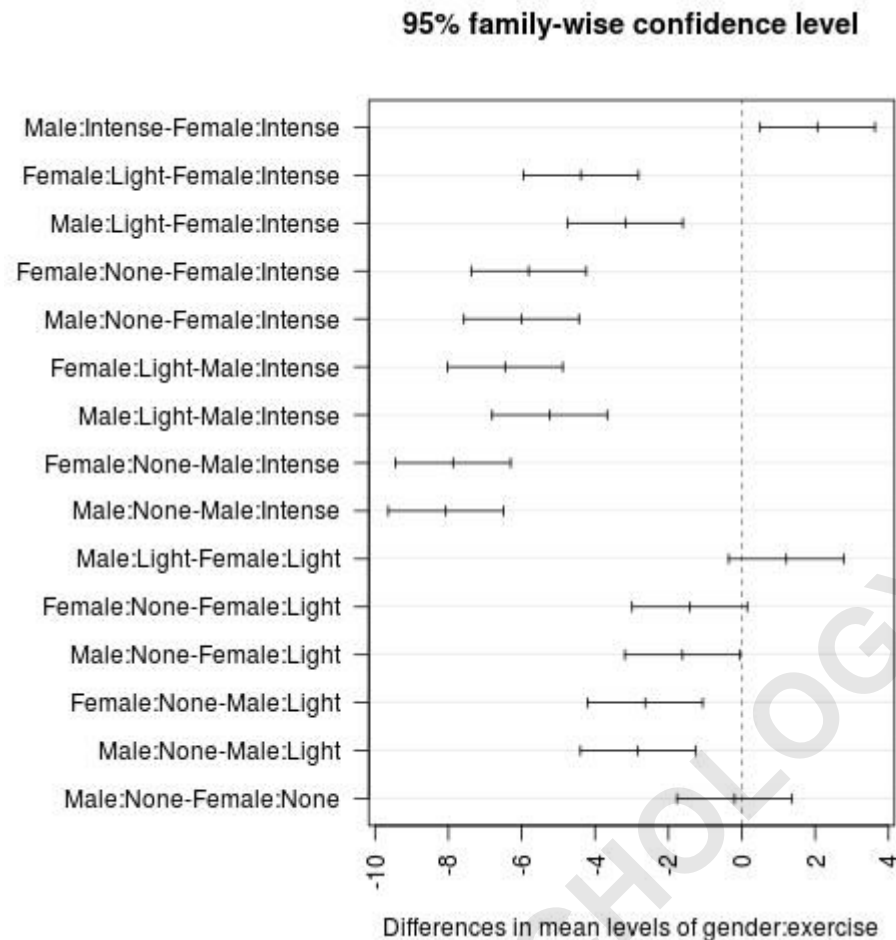
The p-value indicates whether or not there is a statistically significant difference between each group.

For example, in the last row above we see that the male group with no exercise did not experience a statistically significant difference in weight loss compared to the female group with no exercise (p-value: 0.990364).

We can also visualize the 95% confidence intervals that result from the Tukey Test by using the `plot()` function in R:

```
#set axis margins so labels don't get cut off  
par(mar=c(4.1, 13, 4.1, 2.1))
```

```
#create confidence interval for each comparison  
plot(TukeyHSD(model, conf.level=.95), las = 2)
```



Reporting the Results of the Two-Way ANOVA

Lastly, we can report the results of the two-way ANOVA in such a way that summarizes the findings:

A two-way ANOVA was conducted to examine the effects of gender (*male, female*) and exercise regimen (*none, light, intense*) on weight loss (*measure in lbs*). There was a statistically significant interaction between the effects of gender and exercise on weight loss ($F(2, 54) = 4.615, p = 0.0141$). Tukey's HSD post hoc tests

were carried out.

For males, an *intense* exercise regimen lead to significantly higher weight loss compared to both a *light* regimen ($p < .0001$) and *no exercise* regimen ($p < .0001$). In addition for males, a *light* regimen lead to significantly higher weight loss compared to *no exercise* regimen ($p < .0001$).

For females, an *intense* exercise regimen lead to significantly higher weight loss compared to both a *light* regimen ($p < .0001$) and *no exercise* regimen ($p < .0001$).

Normality checks and Levene's test were conducted to verify that the ANOVA assumptions were met.