

# How to Perform an ANCOVA in R: A Step-by-Step Guide

Authored by  
**stats writer**

December 30, 2025

## RECOMMENDED CITATION

stats writer (2025). *How to Perform an ANCOVA in R: A Step-by-Step Guide*.  
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=110038>

The Analysis of Covariance (ANCOVA) is a powerful statistical technique that extends traditional analysis of variance (ANOVA) by incorporating one or more continuous variables, known as covariates. Its primary utility lies in comparing the means of two or more independent groups while statistically adjusting the response variable for the influence of these covariates. This adjustment is crucial for enhancing statistical power and reducing within-group error variance, allowing for a more precise estimation of treatment effects.

In the R programming environment, ANCOVA is typically implemented using the built-in `lm()` function, which is designed to fit a linear model. By properly specifying both the grouping factor (the independent variable) and the continuous covariate(s) within this function, researchers can effectively control for extraneous variation. The resulting model summary provides the necessary output to rigorously evaluate the impact of the primary treatment effect after accounting for the covariate's influence.

This comprehensive tutorial walks through a practical example demonstrating how to perform a robust ANCOVA in R, detailing every step from data preparation to final interpretation.

## ANCOVA Application Example: Evaluating Study Techniques in R

To illustrate the application of ANCOVA, we will analyze a scenario designed to test whether different studying techniques significantly impact student performance on an exam. By utilizing ANCOVA, we can isolate the effect of the technique while controlling for pre-existing differences in student academic ability.

This analysis focuses on three essential variables, defining their roles within the ANCOVA framework:

**Studying technique:** This serves as the primary **independent variable** (the factor of interest) which dictates group membership (Techniques A, B, or C).

**Student's current grade:** This is the **covariate**, a continuous variable used to adjust the outcome, accounting for initial differences in student proficiency.

**Exam score:** This represents the **response variable** (or dependent variable), the outcome we are measuring after the intervention period.

The underlying dataset comprises information from 90 students who were randomly assigned and evenly split into three groups of 30, corresponding to the three study techniques. The randomization helps ensure group parity before the intervention begins.

The dataset records the specific studying technique each student employed (A, B, or C), their current grade in the class at the start of the study (the covariate), and the final exam score they attained after utilizing the technique for one month:

### # Ensure reproducibility of the random dataset generation **set.seed(10)**

```
# Create the sample dataset structure
data <- data.frame(technique = rep(c("A", "B", "C"), each = 30),
  current_grade = runif(90, 65, 95),
  exam = c(runif(30, 80, 95), runif(30, 70, 95), runif(30, 70, 90)))
```

```
# Display the first six observations
head(data)
```

```
# technique current_grade exam
#1 A 80.22435 87.32759
#2 A 74.20306 90.67114
#3 A 77.80723 88.87902
#4 A 85.79306 87.75735
#5 A 67.55408 85.72442
#6 A 71.76310 92.52167
```

## Step 1: Preliminary Data Exploration

Before fitting any complex statistical model like ANCOVA, it is standard practice to thoroughly explore the dataset. This exploration helps in identifying the structure of the data, confirming variable types, and checking for potential extreme outliers that might disproportionately influence the model results.

We begin by generating a summary of the descriptive statistics for each variable within the dataset:

### **summary(data)**

```
# technique current_grade exam
# A:30 Min. :65.43 Min. :71.17
# B:30 1st Qu.:71.79 1st Qu.:77.27
# C:30 Median :77.84 Median :84.69
# Mean :78.15 Mean :83.38
# 3rd Qu.:83.65 3rd Qu.:89.22
# Max. :93.84 Max. :94.76
```

The summary output confirms that the categorical variable, `studying technique`, has 30 observations for each level (A, B, and C), ensuring balanced group sizes. Furthermore, the

statistics for the continuous variables--`current_grade` and `exam`--provide insights into their distributions. The current grades range from 65.43 to 93.84 with a mean of 78.15, while the final exam scores range from 71.17 to 94.76 with a mean of 83.38.

Next, we utilize the powerful `dplyr` package to calculate key group statistics, specifically focusing on the mean and standard deviation of both the covariate (current grades) and the response variable (exam scores) broken down by studying technique. This step is critical for evaluating whether the groups started at comparable levels regarding the covariate:

**# Load the specialized data manipulation package, dplyr**

**library(dplyr)**

```
data %>%
group_by(technique) %>%
summarise(mean_grade = mean(current_grade),
sd_grade = sd(current_grade),
mean_exam = mean(exam),
sd_exam = sd(exam))

# A tibble: 3 x 5
# technique mean_grade sd_grade mean_exam sd_exam
#1 A 79.0 7.00 88.5 3.88
#2 B 78.5 8.33 81.8 7.62
#3 C 76.9 8.24 79.9 5.71
```

Reviewing these calculated means and standard deviations, we observe that the average current grades and their variability (standard deviations) for students across the three techniques are relatively similar. This initial indication suggests that the randomization process was successful in creating comparable groups prior to the intervention, lending credibility to the subsequent ANCOVA adjustment.

## Visualizing Group Distributions with Boxplots

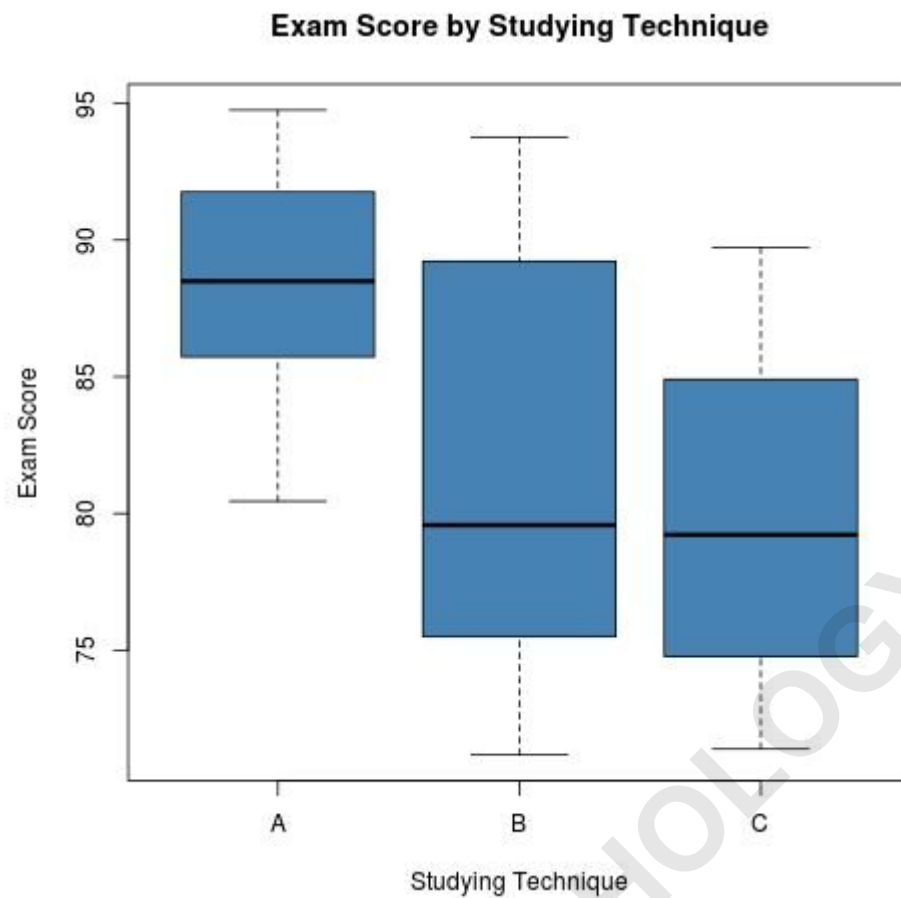
Data visualization provides a quick and intuitive understanding of group differences and data spread. We can use boxplots to graphically compare the distribution of the outcome variable (Exam Score) across the different studying techniques.

```
boxplot(exam ~ technique,
data = data,
main = "Exam Score by Studying Technique",
xlab = "Studying Technique",
```

```
ylab = "Exam Score",  
col = "steelblue",  
border = "black"  
)
```

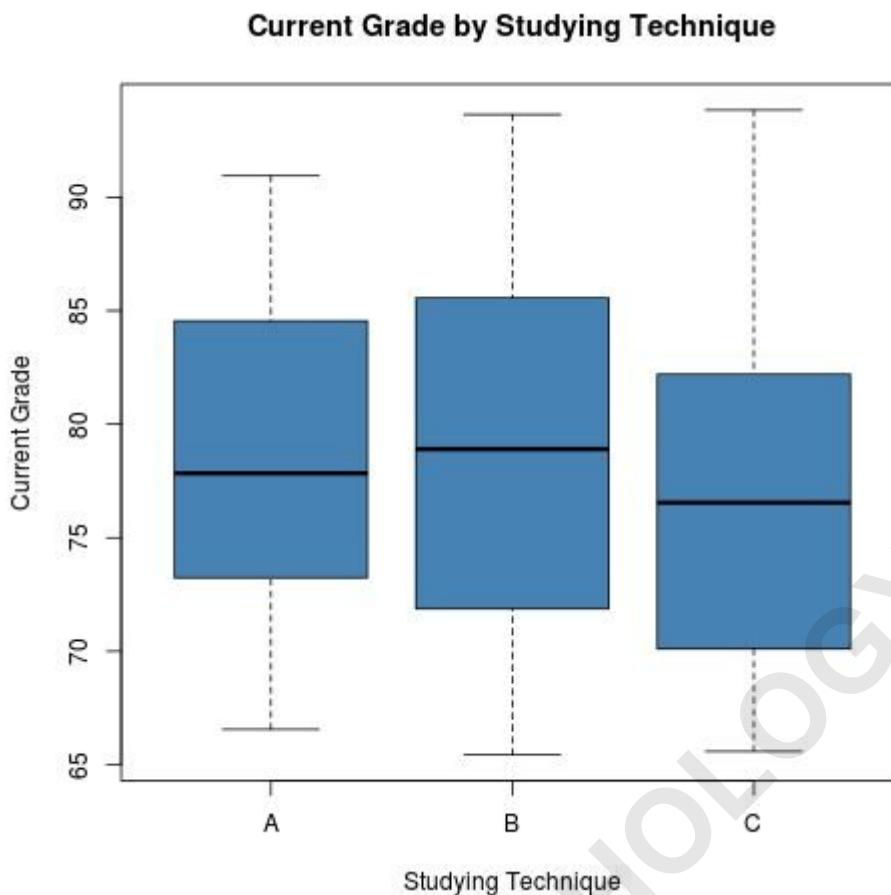
ARABPSYCHOLOGY.COM

ARABPSYCHOLOGY.COM



Similarly, visualizing the distribution of the covariate (Current Grade) across groups helps confirm that the covariate is spread relatively evenly across the treatment levels, which supports the independence assumption we will formally test next:

```
boxplot(current_grade ~ technique,  
data = data,  
main = "Current Grade by Studying Technique",  
xlab = "Studying Technique",  
ylab = "Current Grade",  
col = "steelblue",  
border = "black"  
)
```



## Step 2: Verifying Key Model Assumptions

For the ANCOVA model results to be statistically reliable and interpretable, several core assumptions must be satisfied. Having completed the initial data exploration, we must formally test two critical assumptions related to the relationship between the covariate and the treatment groups.

The primary assumptions we must verify are:

**Independence of the Covariate and Treatment:** It is essential that the covariate (Current Grade) and the treatment factor (Studying Technique) are independent of each other. If the treatment assignment is dependent on the covariate, the covariate might partially absorb the treatment effect, making the interpretation of the main factor difficult.

**Homogeneity of Variance:** This assumption requires that the variance of the response variable (Exam Score) is approximately equal across all levels of the independent variable (Studying Technique).

To test the independence assumption, we run a simple ANOVA with `current_grade` as the

dependent variable and `technique` as the predictor. If the studying technique does not significantly predict the current grade (i.e., the p-value is high), the assumption of independence is met:

#### # Fit ANOVA model to test independence

```
anova_model <- aov(current_grade ~ technique, data = data)
```

#### # View summary of the ANOVA model

```
summary(anova_model)
```

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

```
#technique 2 74 37.21 0.599 0.552
```

```
#Residuals 87 5406 62.14
```

The p-value associated with `technique` is 0.552, which is significantly greater than the standard significance level of 0.05. This result confirms that the covariate (Current Grade) and the treatment (Studying Technique) are indeed statistically independent, satisfying this key ANCOVA requirement.

Next, we test for homogeneity of variance across the groups using Levene's Test. This test evaluates whether the spread of exam scores is roughly equal across groups A, B, and C:

#### # Load the 'car' library, which contains Levene's Test

```
library(car)
```

```
# Conduct Levene's Test on the response variable
```

```
leveneTest(exam~technique, data = data)
```

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

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

```
#group 2 9.4324 0.0001961 ***
```

```
# 87
```

The output shows a highly significant p-value (0.0001961), which is far below 0.05. This result leads us to reject the null hypothesis of equal variances, indicating that the assumption of homogeneity of variance has been violated. While this violation suggests potential issues--which could ideally be addressed via data transformation or using robust alternatives--we will proceed with the standard ANCOVA model for demonstration, noting this limitation in the interpretation.

### Step 3: Fitting the ANCOVA Model

With the preliminary steps complete, we now fit the core ANCOVA model using the `aov()` function. The model structure specifies `exam` (response variable) as a function of `technique` (categorical

predictor) plus `current_grade` (continuous covariate).

For robust inference, especially in unbalanced designs (though ours is balanced, it is good practice), we use the `Anova()` function from the `car` package to specify Type III Sum of Squares. This method ensures that the effect of each predictor is tested after accounting for all other effects in the model, providing a more reliable interpretation of the unique contribution of the studying technique.

### # Load the 'car' library for Type III Sum of Squares

```
library(car)
```

```
# Fit the ANCOVA model using aov (which is equivalent to lm for this model type)
```

```
ancova_model <- aov(exam ~ technique + current_grade, data = data)
```

```
# View Type III summary of the model
```

```
Anova(ancova_model, type="III")
```

```
#Response: exam
```

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

```
 #(Intercept) 7161.2 1 201.4621 < 2.2e-16 ***
```

```
 #technique 1242.9 2 17.4830 4.255e-07 ***
```

```
 #current_grade 12.3 1 0.3467 0.5576
```

```
 #Residuals 3057.0 86
```

The results of the ANCOVA model provide clear evidence regarding the effects of our variables. Crucially, the p-value for the primary predictor, `technique`, is extremely small (4.255e-07). Since this value is much less than 0.05, we conclude that studying technique has a statistically **significant effect** on final exam scores, even after adjusting for the initial differences captured by the student's current grade. Conversely, the p-value for the covariate, `current_grade` (0.5576), indicates that after accounting for the study technique, the covariate itself is not a significant predictor in this specific model.

## Step 4: Conducting Post Hoc Analysis

The statistically significant F-test for `technique` tells us that at least one pair of study techniques differs significantly from the others. However, it does not specify which techniques are different. To pinpoint these specific differences, we must conduct post hoc tests (or multiple comparisons).

We will employ Tukey's Honest Significant Difference (HSD) test, which is highly suitable for pairwise comparisons following an ANOVA or ANCOVA, as it controls the family-wise error rate. We use the `glht()` function within the `multcomp` package in R to perform this procedure:

**# Load the 'multcomp' library for multiple comparisons****library(multcomp)**

# Re-fit the ANCOVA model (necessary for glht function structure)

ancova\_model &lt;- aov(exam ~ technique + current\_grade, data = data)

# Define the post hoc comparisons using Tukey contrasts

postHocs &lt;- glht(ancova\_model, linfct = mcp(technique = "Tukey"))

# View a summary of the post hoc comparisons, including p-values

summary(postHocs)

#Multiple Comparisons of Means: Tukey Contrasts

#

#Fit: aov(formula = exam ~ technique + current\_grade, data = data)

#

#Linear Hypotheses:

# Estimate Std. Error t value Pr(&gt;|t|)

#B - A == 0 -6.711 1.540 -4.358 0.000109 \*\*\*

#C - A == 0 -8.736 1.549 -5.640 &lt; 1e-04 \*\*\*

#C - B == 0 -2.025 1.545 -1.311 0.393089

# View the confidence intervals associated with the multiple comparisons

confint(postHocs)

# Simultaneous Confidence Intervals

#

#Multiple Comparisons of Means: Tukey Contrasts

#

#Fit: aov(formula = exam ~ technique + current\_grade, data = data)

#

#Quantile = 2.3845

#95% family-wise confidence level

#

#Linear Hypotheses:

# Estimate lwr upr

#B - A == 0 -6.7112 -10.3832 -3.0392

#C - A == 0 -8.7364 -12.4302 -5.0426

#C - B == 0 -2.0252 -5.7091 1.6588

The results from the Tukey HSD test clearly delineate the differences between the study

techniques:

First, there is a statistically significant difference ( $p = 0.000109$ ) in adjusted exam scores between studying technique **A** and technique **B**. The negative estimate (-6.711) indicates that technique A results in higher scores than technique B. Second, technique **A** also shows a highly significant difference ( $p < 1e-04$ ) compared to technique **C**, with technique A again yielding superior adjusted mean scores (Estimate = -8.736).

Finally, the comparison between technique **B** and technique **C** yields a p-value of 0.393089, which is far above the 0.05 threshold. This means there is **no statistically significant difference** in adjusted exam scores between techniques B and C. The corresponding confidence interval for B - C includes zero, confirming the lack of a significant difference between these two groups.

In summary, based on the ANCOVA and post hoc analysis, we confidently conclude that utilizing studying technique **A** leads to a statistically significantly greater adjusted exam score for students compared to techniques B and C. This finding remains robust even after controlling for the student's initial current grade in the class.