

# **SOURCEBOOK**

## **R Articles**

### **Data Analysis**

**Abstract:** This chapter provides step-by-step written instructions and screenshots for obtaining basic statistical output using R. Simple examples for most undergraduate-level between-subjects and within-subjects research designs are provided.

**Keywords:** R, directions for use

**Original:** March 2021

**Updated:** May 2025

**This document is part of an online statistics Sourcebook.**

A browser-friendly viewing platform for this Sourcebook is available:

<https://cwendorf.github.io/Sourcebook>

## Table of Contents for This Chapter

Frequencies .....	3
Descriptives .....	4
Transformations and Standardized Scores .....	5
Correlations .....	6
Regression .....	7
Confidence Interval for a Mean .....	8
One Sample t Test.....	9
Paired Samples t Test.....	10
Independent Samples t Test .....	11
OneWay ANOVA.....	12
Post Hoc Comparisons .....	13
Repeated Measures ANOVA .....	14
Factorial ANOVA.....	15

# Frequencies

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Obtaining Frequencies

Get the frequency distribution for the variable.

```
FrequencyTable <- table(Outcome)
FrequencyTable
prop.table(FrequencyTable)
```

## Obtaining Summary Statistics

Get the percentiles for the variable.

```
length(Outcome)
summary(Outcome)
```

# Descriptives

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Obtaining Frequencies

Get the frequency distribution for the variable.

```
FrequencyTable <- table(Outcome)
FrequencyTable
prop.table(FrequencyTable)
```

## Obtaining Descriptive Statistics

Get the sample size, mean, variance, and standard deviation for the variable.

```
length(Outcome)
mean(Outcome)
var(Outcome)
sd(Outcome)
```

# Transformations and Standardized Scores

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Computing Transformations

Use a formula to calculate a new vector with the transformed scores.

```
trOutcome <- Outcome + 1
```

Create and display the data set in a frame.

```
data.frame(Outcome, trOutcome)
```

## Computing Standardized Scores

Create a new variable vector containing the standardized scores.

```
zOutcome <- scale(Outcome)
```

Create and display the data set in a frame.

```
data.frame(Outcome, trOutcome, zOutcome)
```

# Correlations

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Obtaining Descriptive Statistics

Get the sample sizes, means, and standard deviations for the variables.

```
lapply(CorrelationData, function(x) c(length(x), mean(x), sd(x)))
```

Get the covariance and correlation matrices for the variables.

```
cov(Outcome1, Outcome2)  
cor(Outcome1, Outcome2)
```

Get the correlation matrix for the variables.

```
(CorrelationData) |> describeCorrelations()
```

## Obtaining Inferential Statistics

Get the correlation, its test for statistical significance, and its confidence interval.

```
cor.test(Outcome1, Outcome2)
```

You can change the width of the confidence interval if desired.

```
cor.test(Outcome1, Outcome2, conf.level=.99)
```

# Regression

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Obtaining Descriptive Statistics

Get the sample sizes, means, and standard deviations for the variables.

```
lapply(CorrelationData, function(x) c(length(x), mean(x), sd(x)))
```

Get the covariance matrix for the variables.

```
cov(Outcome1, Outcome2)
```

Get the correlation matrix for the variables.

```
(CorrelationData) |> describeCorrelations()
```

## Obtaining Inferential Statistics

Get the proportion of variance accounted for by the model and the model coefficients, along with statistical significance for each of them.

```
model <- lm(Outcome2 ~ Outcome1)  
summary(model)
```

# Confidence Interval for a Mean

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Obtaining Descriptive Statistics

Get the sample size, mean, and standard deviation for the variable.

```
length(Outcome)
mean(Outcome)
sd(Outcome)
```

## Obtaining Inferential Statistics

Get the mean and its confidence interval.

```
t.test(Outcome)$conf.int
```

Also, you can change the width of the confidence interval if desired.

```
t.test(Outcome, conf.level=.99)$conf.int
```



# One Sample t Test

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Obtaining Descriptive Statistics

Get the sample size, mean, and standard deviation for the variable.

```
c(length(Outcome), mean(Outcome), sd(Outcome))
```

## Obtaining Inferential Statistics

Test the mean difference for statistical significance and get its confidence interval.

```
t.test(Outcome, mu=7)
```

Also, you can change the width of the confidence interval if desired.

```
t.test(Outcome, mu=7, conf.level=.99)
```

# Paired Samples t Test

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Obtaining Descriptive Statistics

Get the sample sizes, means, and standard deviations for the variables.

```
lapply(PairedData, function(x) c(length(x), mean(x), sd(x)))
```

## Obtaining Inferential Statistics

Test the mean difference for statistical significance and its confidence interval.

```
t.test(Outcome1, Outcome2, paired=TRUE)
```

Also, you can change the width of the confidence interval if desired.

```
t.test(Outcome2, Outcome1, paired=TRUE, conf.level=.99)
```

# Independent Samples t Test

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Obtaining Descriptive Statistics

Get the samples size, mean, and standard deviation for the dependent variable for each of the levels.

```
tapply(Outcome, Factor, function(x) c(length(x), mean(x), sd(x)))
```

## Obtaining Inferential Statistics

Test the mean difference for statistical significance and get its confidence interval.

```
t.test(Outcome~Factor, var.equal=T)
```

Also, you can change the width of the confidence interval if desired.

```
t.test(Outcome~Factor, var.equal=T, conf.level=0.99)
```

# OneWay ANOVA

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Obtaining Descriptive Statistics

Get the grand mean for the variable and the mean and standard deviation for the dependent variable for each of the levels.

```
mean(Outcome)
tapply(Outcome, Factor, function(x) c(length(x), mean(x), sd(x)))
```

## Obtaining Inferential Statistics

Get the analysis of variance source table with test of statistical significance.

```
Results <- aov(Outcome~Factor)
summary(Results)
```

# Post Hoc Comparisons

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Obtaining Descriptive Statistics

Get the grand mean for the variable and the mean and standard deviation for the dependent variable for each of the levels.

```
mean(Outcome)
tapply(Outcome, Factor, function(x) c(length(x), mean(x), sd(x))))
```

## Obtaining Inferential Statistics

Test each pairwise comparison for statistical significance.

```
Results <- aov(Outcome~Factor)
TukeyHSD(Results)
```

# Repeated Measures ANOVA

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Obtaining Descriptive Statistics

Get the sample sizes, means, and standard deviations for the variables.

```
lapply(RepeatedData, function(x) c(length(x), mean(x), sd(x)))
```

## Obtaining Inferential Statistics

Get the ANOVA source table with tests of statistical significance.

```
Results=aov(Outcome~factor(Factor)+Error(factor(Subject)))  
summary(Results)
```

# Factorial ANOVA

Prior to the steps below, enter the data as appropriate for the analyses (described elsewhere). As always, the following commands should be typed directly in the R console window.

## Obtaining Descriptive Statistics

Get the mean, sample size, and standard deviation for the dependent variable for each of the levels. Also, get the means for the main effects.

```
Results <- aov(Outcome~FactorA*FactorB)
model.tables(Results,"means")
tapply(Outcome, list(FactorA,FactorB), length)
tapply(Outcome, list(FactorA,FactorB), sd)
```

## Obtaining Inferential Statistics

Get the analysis of variance source table and a test of statistical significance.

```
summary(Results)
```