SOURCEBOOK

R

DATA ANALYSIS

Abstract: This chapter provides step-by-step instructions on how to obtain basic statistical output using R via written instructions. Simple examples for most undergraduate-level between-subjects and within-subjects research designs are provided.

Keywords: R, directions for use

Original: March 2021
Updated: January 2025

This document is part of an online statistics sourcebook.

A browser-friendly viewing platform for the sourcebook is available:
https://cwendorf.github.io/Sourcebook

All data, syntax, and output files are available:
https://github.com/cwendorf/Sourcebook

TABLE OF CONTENTS FOR THIS CHAPTER

Frequencies 3

Descriptives 4

Transformations and Standardized Scores 5

Correlations 6

Confidence Intervals 7

One Sample t Test 8

Paired Samples t Test 9

Independent Samples t Test 10

OneWay ANOVA 11

Post Hoc Comparisons 12

Repeated Measures ANOVA 13

Factorial ANOVA 14

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 mean and standard deviation for the variable.

```
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 a 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 means and standard deviations for the variables.

```
mean(Outcome1)
sd(Outcome1)
mean(Outcome2)
sd(Outcome2)
```

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)
```

Confidence Intervals

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 and standard deviation for the variable.

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

Obtaining Inferential Statistics

Get the mean and its confidence interval.

```
t.test(Outcome)$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 mean and standard deviation for the variable.

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

Obtaining Inferential Statistics

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

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

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 means and standard deviations for the variables.

```
mean(Outcome1)
sd(Outcome1)
mean(Outcome2)
sd(Outcome2)
```

Obtaining Inferential Statistics

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

```
t.test(Outcome1-Outcome2,mu=0)
```

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

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

```
by(Outcome,Factor,mean)
by(Outcome,Factor,sd)
```

Obtaining Inferential Statistics

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

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

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

```
by(Outcome,Factor,mean)
by(Outcome,Factor,sd)
```

```
Results <- aov(Outcome~Factor)
model.tables(Results,"means")
```

Obtaining Inferential Statistics

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

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

```
by(Outcome,Factor,mean)
by(Outcome,Factor,sd)
```

```
Results <- aov(Outcome~Factor)
model.tables(Results,"means")
```

Obtaining Inferential Statistics

Test each pairwise comparison for statistical significance.

```
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 means and standard deviations for the variables.

```
mean(Outcome1)
sd(Outcome1)
mean(Outcome2)
sd(Outcome2)
```

Change the data format for use with R procedures.

```
StackData=reshape(RepeatedData,varying=c("Outcome1","Outcome2"),v.names="
Outcome",timevar="Factor",idvar="Subject",direction="long")
attach(StackData)
StackData
```

```
mean(Outcome1)
Results=aov(Outcome~factor(Factor)+Error(factor(Subject)))
model.tables(Results,"means")
```

Obtaining Inferential Statistics

Get the ANOVA source table with tests of statistical significance.

```
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 and standard deviation of the dependent variable for each of the levels.

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

Obtaining Inferential Statistics

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

```
summary(Results)
```