Statistics with R – Intermediate Level

Practice

Note: If you did not do it already, please download the CSV data files and extract them on your hard drive. You can find the download link in the section *Course Materials*.

Section 2

Mean Difference

Exercise #1

In the *students.csv* data frame, check whether there is a difference in average score between male and female students. (In the variable *gender*, the male students are coded with 1 and the female students with 0).

Exercise #2

In the *bankloan.csv* data set, determine whether there is a difference in average balance between people who have a loan (*loan*="yes") and people who do not (*loan*="no").

Exercise #3

In the data set *gym.csv* , determine whether there is a difference between the average score awarded by the judges from Romania and United States.

Exercise #4

In the data file *incomedata.csv*, determine whether the subjects' income is influenced by their education level, using the one-way analysis of variance.

Exercise #5

In the data file *incomedata.csv*, determine whether the subjects' income is influenced by their education level and age, using the two-way analysis of variance.

Exercise #6

In the data file *incomedata.csv*, determine whether the subjects' income is influenced by their education level, age and gender, using the three-way analysis of variance.

Exercise #7

In the data frame *vehicles.csv*, determine whether the variables price and sale are influenced by the vehicle type (automobile or truck), using the multivariate analysis of variance.

Exercise #8

In the data set *satisfaction.csv*, check whether there is a difference between the median satisfaction of the two types of travellers (business and pleasure) using the Mann-Whitney test.

Exercise #9

In the data set *teachers.csv*, check whether there is a difference between the median grades of the two teachers using the Wilcoxon test.

Exercise #10
In the data set *satisfaction.csv*, check whether there is a difference between the median satisfaction of the customer groups by price importance using the Kruskal-Wallis test.