List 2

Exercise 1. Two samples are given

(1) 5, 4, 3, 3, 5
(2) 1, 4, 6, 0, 8, 5.

Calculate the mean, variance and standard deviation for these samples. Draw the boxplots.

Exercise 2. The focal lengths in network cameras were tested and the following results (in mm) were obtained:

2.45, 3.6, 2.8, 2.45, 3.0, 2.8, 2.6, 2.8.

Determine the sample statistics (position and dispersion measures) and interpret the results.

Exercise 3. Create two 20-dimensional random vectors **r1** and **r2** generated from binomial distribution with parameters n=10 and p=0.3 and:

a) determine basic statistical characteristics of **r1** and **r2**: mean, variance, standard deviation, minimum, maximum, quartiles

b) compute and interpret variability of generated data

c) determine and interpret quantiles of order 0.1, 0.3, 0.6, and 0.9

d) repeat (a)-(c) to a vector consisting of **r1** and **r2**.

Exercise 4. Load the file strawberries.csv and:

a) verify the type of data; look to the data and pay attention for the lengths of variables and missing data

b) determine basic statistical characteristics for data labelled by „crop 2010”

c) draw probabilistic histograms for both years; automatize the process using „for” procedure; declare titles of histograms recalling data labels; arrange all the graphs in a single window

d) compare the data from consecutive years with the use of boxplots.

Exercise 5. The battery life [in hours] of the remote-controlled robot was measured. The following results were obtained

6, 7, 10, 5, 10, 12, 11, 7, 7, 9, 10, 8, 16, 11.

Plot the probabilistic histogram for the above data set. Plot a blue-colored histogram of the columns. There should be four columns. The first should contain values up to 8h, the second from 8h to 10h, the third from 10h to 14h, and the last up to 16 hours.

Determine the point and interval series for the given data.

Exercise 6. The weight of the drone was tested [in g] and the following results were obtained:

249, 254, 233, 284, 254, 271, 239, 282, 305, 276.

Plot the histogram for the above data in three different columns, broken down as desired in red.

Exercise 7. Create two vectors consisting of 10 elements each. Create boxplots for these vectors separately and in one plot area.

Exercise 8. The total capacity of the wind farm achieves dynamic growth thanks to the development of technology. Made 7 measurements of the nominal power of a wind farm (feature Y, in MW) depending on the propeller diameter (feature X, in m). The following results were obtained:

x_i	140	140	150	152	154	164	167
y_i	6	5.5	6	6.2	6.7	8	9.5

Determine the covariance between the propeller diameter and the nominal power of the power plant. Determine and interpret the correlation coefficient between the propeller diameter and the nominal power of the power plant.

Exercise 9. Load ranking.csv file. Check the data type and calculate the mean and standard deviation for the data from the first two columns. Create histograms for the next two.

Exercise 10. The weight of some two workpieces located in the car engine was measured [in g]. The following values were obtained:

workpiece I: 12.3; 15.7; 22.0; 13.8; 16.6; 12.1; 15.4; 16.2; 19.1; 17.6; 15.8

workpiece II: 11.3; 13.7; 12.0; 12.6; 15.1; 11.9; 14.7; 15.1; 16.2; 12.4; 14.7.

a) Plot the boxplot for the measured weight values of workpiece I.
b) Plot the boxplot for the measured weight values of workpiece II.
c) Plot the boxplot for the measured weight values of workpieces I and II on one graph.
d) Plot a histogram for the measured values of workpiece I.
e) Plot a histogram for the measured values of workpiece II.
f) Add appropriate titles and a description of the x axis to the histograms.
g) Reduce the number of columns to five for two histograms.