R Kurs Übungen / R Course Exercises

1. Übungen zum Erstellen, Indizieren und Berechnungen mit Vektoren in R Exercises on creation, indexing and calculations with vectors in R

Exercise 1.

Create a vector containing the numbers 3,9,7,6,4,1.

Then calculate the mean of the elements in the vector and the square root of the different elements.

Exercise 2:

Create a vector containig the numbers 1, 10 and 100 with data type of integer.
Use two different methods. Test the data type by using an R function call.

Exercise 3:

Create a vector with the elements "Easy", "work", "with", "R".
What's its data type?

Exercise 4:

Create a vector that contains the odd numbers between 1 and 100.
Try to find the simplest way to do so.

Exercise 5:

Create a vector x containig the numbers from 1 to 5 and a vector y with the numbers 7 to 20.
Combine both vectors to get vector z. Delete x and y from memory.
Sort z in decreasing order.

Exercise 6:

Calculate the mean of elemnts number 2, 7 and 9 of vector z (Ex. 5).

2. Übungen zu Matrizen

Exercise 1a:

Create this matrix as M1 in R:

```
     [,1] [,2]
[1,]    1    6
[2,]    2    7
[3,]    3    8
[4,]    4    9
[5,]    5   10
```

Exercise 1b:

Do the same as M2, but row-wise.

Exercise 1c:

Get the element in row 2, column 3 from that matrix (1b, M2). (Should be number 6.)

Exercise 1d:

Subtract 2 from every element of the second column. Calculate the mean of that new vector.

Exercise 1e:

Create a matrix M3 from M2 where the numbers in column 2 are replace by the numbers in column 1 each subtracted by 2.

Exercise 1f:

Add a sixth row to M3 containig the elements 12 and 14. And another row with 20 and 25.

Exercise 2:

Create a matrix with two rows and two columns containing numbers 1 and 4, row-wise. row names should be "a1" and "a2", column names "b1" and "b2.

Exercise 3:

Create the following matrix using the shortest way you can think of.

```
      [,1] [,2] [,3] [,4] [,5]
 [1,]  100   80   60   40   20
 [2,]   98   78   58   38   18
 [3,]   96   76   56   36   16
 [4,]   94   74   54   34   14
 [5,]   92   72   52   32   12
 [6,]   90   70   50   30   10
 [7,]   88   68   48   28    8
 [8,]   86   66   46   26    6
 [9,]   84   64   44   24    4
[10,]   82   62   42   22    2
```

3. Exercises with Dataframes

Exercise 3a:

Create the following data frame as "data".

	Plant_No	fertile_Seeds	aborted_Seeds
1	At1	377	10
2	At2	879	30
3	At3	216	41
4	At4	93	71
5	At5	98	22
6	At6	103	6

Exercise 3b:

Print out the statistical summary for "data"

Exercise 3c:

We forgot some data. Add the column "Counted_silics".
Use 6, 18, 4, 3, 2 and 2 silics accordingly.

Exercise 3d:

Sort the data by increasing number of counted silics.
Use the function order().

Exercise 2:

Open the iris dataset as dataframe "d" for the following exercises.

Exercise 3:

View the top of the dataset with head().
Now let's access all rows holding data for the species "virginica".

One possibility is to use the function which(). Which() will generate a numeric vector whose elements correspond to TRUE in a corresponding logical vector.
First create that numeric vector and a new dataframe "subset" in a second step.

Exercise 4:

Create a new dataframe "subset2" from "subset" that contains all virginica species with a Petal width below 2.

4. Exercises with Loops

Exercise 1:

Use the function ifelse() to print out "Nope" if an element of a vector containig the numbers from 10 to 20 is not 16 and "Hooray" if it is.

Exercise 2:

Use repeat() and break to implement a count-down from 9 to 0.

Exercise 3:

Create a loop that multiplies each of the numbers from 1 to 10 with 3 and print them out.

5. Übungen zu besonderen Funktionen

Exercise 1:

Benutzen Sie die grep() Funktion, um aus folgendem Vektor alle Elemente auszulesen, in denen RKURS steht:

```
x <- c("KURSFREITAGHEUTE", "HEUTERKURSTOLL", "RKUSISTEIN",
"ALLERKURSESINDAUSGEBUCHT", "ALLESKURSIVODER")
```

Exercise 2:

Please generate the following matix: x <- matrix(c(1:100), nrow = 10). Use the function appy() to read out the smallest number of each column..

6. Exercises for graphical presentations

Exercise 1:

Generate a scatterplot of the values of log(1) to log(100) on the x-achses against the nubmers from 1 to 100 on the y-achses..

Exercise 2:

Use a histogram to show the distribution of the petal lenghts of the species "setosa" of the Iris dataset.

Exercise 3:

Use a box and whiskers plot to show the lenght of the petals of the two species "setosa" and "virginica".

Exercise 4:

Plot the mean values of the lengths of length of petals of "setosa" and of "virginica" using a barplot (barplot()). Add the standard deviations of the means of the petal length as error bars. To do so you will need the function barplot2() of the package gplots. Install and load gplots and check the help for the function barplot2().

7. Excercises concerning distributions and hypothesis tests

Exercise 1:

Test if the mean values of the petal length of "setosa" and "virginica" are significantly different.

Exercise 2:

Based on the Iris dataset we would like to ask if the sepal width is significantly higher in the species "virginica" as compared to "setosa". To be on the safe side we test vice versa if the sepal width of "virginica" is significantly smaller as comared to "setosa"
Perform the test statistics and visualize the differences grafically.

Exercise 3:

In the laboratory you are checking mutant Arabidopsis lines with a defect during seed seed development. By opening the siliques you have counted all seeds that developed normally (norm), as well the the onces that show abort (abort). Of these lines 4 siliques each jof 3 plants have been counted and the data were safed in the file seedabort.txt.

Read in the data into R and inspect them. Analyse if the underlying defect is gamtophytic or zygotic. In case of a gametophytic defect you expect the ratio of norm : abort to be 1 : 1.
In case of a sporophytic effect you expect the ratio to be 3 : 1. Please test both possibilities using an appropriate test statistics.

8. Exercise concerning linear regression

Exercise 1:

Here we use an example dataset from (https://de.wikibooks.org/wiki/GNU_R:_Anwendungsbeispiele).

We generate the dataset using the following syntax:

```
> x <- c(1, 3, 6, 11, 12, 15, 19, 23, 28, 33, 35, 39, 47, 60, 66, 73)
> y <- c(3180, 2960, 3220, 3270, 3350, 3410, 3700, 3830, 4090, 4310, 4360, 4520, 4650, 5310, 5490, 5540)
> bsp5 <- data.frame(x,y)
> colnames(bsp5) <- c("Lebenstag", "Gewicht")
```

Erstellen Sie einen Scatterplot inklusive Regressionslinie.

9. Exercise concerning distributions

Exercise 1a:

Randomly draw five numbers between 1 and 6 for 1000 times and calculate the mean values. Draw a histogram of the mean values.

#Exercise *1b*:

Calculate the mean value of the mean and the standard deviation.

Exercise 1c:

Plot the normal distribution in the same graph as the histogram.

Exercise 2:

Plot the curve for the t-distribution with 4 degrees of freedom in the limits between ± 5 degrees of freedom. Mark the intervall containing 95 % of all values.