Using R software

Exercises: statements

1. Using R as a simple calculator

Exercise n°1:

The body mass index (BMI) is a measure of the amount of body fat in the body based on two parameters: weight, expressed in kilograms, and height, expressed in metres. It is calculated according to the following formula:

$$BMI = \frac{weight}{(height)^2}$$

a). Calculate your BMI using, directly, your weight and height (in the correct units).
Note: in R, a power is calculated using the symbol ^ (example: 2^2 will be written as 2^2 on the command line).

b). Keep a memory of your weight value as **`myWeight`** and your height value as **`myHeight`**. Calculate your BMI from **`myWeight`** and **`myHeight`**

Exercise n°2:

You want to start mountain biking and have a budget of 300 euros to buy a bike, a helmet and a polar jacket.

Store the amount of your budget in the variable **`myBudget`**.

a). You have found a mountain bike for 260 euros, a helmet for 30 euros and a polar jacket for 60 euros. Can you make this purchase?

b). The seller offers you a free polar jacket if you buy the mountain bike and helmet right away. Can you make this purchase?

c). You see that the sale will start next week with a 30% discount across the whole shop. What will be the total amount of the discount? What is the best offer?

d). After the purchase, how much money do you have left? Change the value of the variable **`myBudget`** if there was a purchase.

Exercise n°3:

The validation of a TU depends on the results of the continuous assessment (40% of the final mark) and the mark in the examination (60% of the final mark).

Here are the scores of 8 students:

- Continuous assessment: 11, 13, 15.5, 12, 8, 9, 13, 16
- Exam : 8.5, 14, 15, 10, 12, 13, 14, 17

a). Store these values in the variables **`ccont`** et **`exam`**.
Note: in R, to create a vector containing, for example, the values 1, 2, 3, you must write the command: `c(1, 2, 3)`.

b). What is each student's final TU score?

c). What is the score for the continuous assessment, the examination and the TU for student number 6?

d). Calculate the average of the continuous assessment, the examination and the TU for all 8 students. What is the highest score in the TU? The lowest score?

2. Using R software for descriptive analysis

Exercise n°4:

After importing the file "fallers.xlsx" on your computer, read the data file from this study into the R software.

a). What is the nature of the variables studied?
Note: use the class command **`str`** `(str(name_of_your_data.frame))` to understand what are the particularities of these variables.

b). If necessary, recode the variables to the appropriate format.

c). Analyse the **`sex`** variable (descriptive statistical parameters and at least one graphical representation).

d). Analyse the variable **`age`** (descriptive statistical parameters and at least one graphical representation).

e). Make the histogram of **`age`** by choosing the following intervals: [65, 75],]75, 85],]85, 100].
Note: use the command **`?hist`** to understand how to construct these intervals.

Exercise n°5:

With the same file "fallers" as before.

We want to give a representation of the age of individuals according to whether or not they have had a fall.

a). Describe the **`age`** variable in terms of the status of the **`fall`** variable.

b). Produce the histogram of the age of the patients for each level of the **`fall`** variable.
Note 1: you must first construct the 2 vectors **`ageYes`** (age of jumpers) and **`ageNo`** (age of non-fallers) in the following way:

```
ageYes <- age[fall=="Yes"]
```

Note 2: you can use the following commands to display the 2 histograms on the same figure:

```
old.par <- par(no.readonly = TRUE)
par(mfrow = c(1, 2))
hist(ageYes)
hist(ageNo)
par(mfrow = c(1, 1))
par(old.par)
```

Note 3: find out how to use the **`breaks`** and **`freq`** commands in the **`hist`** function to visually compare the 2 histograms.

c). Give a box-plot representation of age as a function of the level of the variable **`fall`**.

Exercise n°6:

With the same file "fallers" as before.

The question is whether falling is associated with lifestyle.

a). Provide the contingency table of the absolute frequencies of the mode of life (**`wayolife`**) in relation to the notion of falling (**`fall`**).

b). Provide the relative frequencies of fallers and non-fallers according to their lifestyle (**`wayolife`**).

Exercise n°7:

With the same file "fallers" as before.

The question is whether the number of drugs taken (**`nbdrugs`**) is associated with the age of the patients (**`age`**).

a). Plot the drug consumption as a function of patient age.

b). Provide a statistical result describing this association.