R for Psychology Research

Week 1 - Exercises

Marcus Lindskog

Vectors

Write a R script...

1. ... to create four vectors, all of length 7. Create one vector each of data type *logical*, *character*, *integer*, and *numeric*. Name your vectors vec_1 - vec_4.
2. ... to check the data type of your vectors.
3. ... that creates two vectors of integers type and length 3 and adds them.
4. ... that creates two vectors of numeric type and length 5 and multiplies them.
5. ... that creates two vectors of numeric type and length 6 and divides them.
6. ... that creates a vector of numeric type and length 7 and find the Sum, Mean and Product of that vector.
7. ... to find the minimum and the maximum of a the vector you created in 6.
8. ... that creates a vector of type integer and tests whether it contains a specified element.
9. ... that creates a vector of type logical and length 4 and access the last value in it.
10. ... that creates two vectors of type numeric and length 6. Assign to a new vector every element in the second vector that is larger than the corresponding element in the first.

Matrices

Write a R script...

11. ... that creates a 4 (row) by 4 (column) matrix from the numbers 15 through 30 where the elements are entered column wise.
12. ... that creates a 4 by 4 matrix from the numbers 15 through 30 where the elements are entered row wise.
13. ... that creates a 3 by 2 matrix from the numbers `1, 26, 24, 68, 32, 99` with row names `Row 1, Row 2, Row 3`, and column names `Col 1, Col 2`.
14. ... to access the element at the 2nd column and the 3rd row in the matrix you created in `13.`
15. ... that to creates a 2x3 matrix from the numbers 9 through 14 and add, subtract, multiply and divide the matrix by 9.
16. ... that creates a 3 x 7 matrix from the numbers 1 through 21 and extracts the sub-matrix whose entries are > 7.

Data frames - Examination Exercises

The solution to the exercises in this section should be handed in as a part of the examination. Your solution should be contained in a single R-script that is emailed to `marcus.lindskog@psyk.uu.se`. Your code should be well commented and easy to follow. Answers to any questions below should be written as a comment in the R-script after the code that produces the answer.

There are several built in data sets in R. For this exercise we will use the `bfi` data set from the `psych` package.

1. Install the `psych` package with `install.packages('psych')`
2. Load the `psych` packages with `library(psych)`
3. Assign the `bfi` data set to your own data frame called `my_data` with `my_data <- bfi`.
4. Explore the data set with `str()` and `summary()`. How many variables and entries does the data set contain?
5. Which variable has the most missing values?
6. Create a new data frame called `demographics` that only contains the variables `gender`, `education`, and `age` from `my_data`.
7. What is the mean age of the participants?
8. The `runif(n)` function let's you create n uniform random numbers in the range $[0, 1]$. Add a variable, `select` to `demographics` using this function.
9. Create a new data frame, `demographics_subset` which contains the subset of `demographics` where `select >= .5`.
10. Add a new variable, `sum_of_rows`, to `my_data` that is the sum of five variables beginning with `A`. Hint 1: R has a built in function called `rowSums`.
11. The five variables A1 - A5 contain missing data `NA`. What argument, and how, should you specify in `rowSums` to not include them in the calculations?
12. Add a new variable, `sum_of_rows_no_na`, to `my_data` that is the sum of five variables beginning with `A`, when `rowSums` does not include `NAs`.