Exercises 1

You should try to carry out each of the following exercises in as few steps as possible. Make use of the help available for each function if you are stuck by typing help(*name_of_function*).

R as calculator

1. Write down R commands that perform/calculate the following
   a. 10 + 54
   b. 20 divided by 60
   c. 2^{20}
   d. assign the value 1.5 to the variable x and the value 2.6 to the variable y
   e. calculate $(x / (y + 20/15))^{18}$

c and some other basic functions

2. Create the vector of numbers –77, 239, 16, 1, 8, 12, 71, 2007, 0, 10 and assign it to a variable k

3. Calculate the

   a. Mean – mean()
   b. Median – median()
   c. Standard deviation – sd()
   d. Maximum - max() and
   e. Minimum – min()

   of k

conditional indexing

4. Calculate
   a. the sum of the entries of k
   b. the sum of the entries in k which are greater than zero

5. Calculate the sum of the odd numbers of k using R (in R %% is the modulo operator. E.g. 10%%7 = 3; the sum() function adds the entries of a vector).

6. Add all the entries of k which are either even or negative.

7. Write an R command that will calculate the mean of the entries of k that are greater than 10 and less than 1000.

8. Write an R command to work out the standard deviation of the positive entries of k.

9. Create a new vector s by adding 101 to each of the elements of k. Work out the sum of the elements of s for which the corresponding element of k is greater than 10.

10. Use an appropriate R function to assign the 2^{nd} and 5^{th} elements of this vector to another vector called m
   a. Add m and k

b. Multiply m by k

using seq and rep to make vectors

11. Create a vector containing all the numbers between 111 and 767 (inclusive). Add all these numbers and multiply the result by the sum of all the numbers between 28 and 35 (inclusive).

12. Create a vector, k, that consists of every number divisible by 3 between 8 and 376 (inclusive). Sum the entries of *k* and divide the result by the sum of every number which is greater than 4 and less than 3428 and is divisible by 39.

13. Create a vector consisting of 11 copies of the vector k from question 12. Work out the average of the entries of this vector.

14. Create the vector r, consisting of 100 zeros, 100 ones and 169 twos (in that order). Calculate the sum of the entries of k from the previous question for which the corresponding element of r is an even number.

15. Write an R command to generate a new vector, a, consisting of the numbers 3.5, 7, 10.5, 14, 17.5 up to 3,500

lists and data frames

16. Create a list containing three vectors called a, b and c, where a is the even numbers between 1 and 21, b is the odd numbers between 2 and 10 and c is a vector containing the names of seven colours.

17. Create the data frame student_data with the vectors height = c (157, 201, 198, 180, 183, 167, 175) weight = c (55, 98, 78, 67, 70, 64, 70) age = c (27, 22, 21, 19, 22, 23, 34). Write an R command to multiply the weight of student three by her height and divide by her age. Write a command that will create a new vector called crude_bmi consisting of weights divided by heights. Work out the average value of crude_bmi. Add crude_bmi to the data frame.

18. Add a factor vector to the data frame in **8** that assigns students to the two groups “Male” and “Female” (assume that every second student is male and that the first student is female). The subset() function allows you to take a sub-portion of a data frame. Make a new data frame called male_data that includes only the data from male students (how to use subset(): subset(*dataframe, condition*))

Writing data to a file

19. Make a new directory called r_course (or whatever you like) on your desktop. Change directory (by clicking on this option under the ‘file’ menu) to the new folder. Use the write() function to write the vector (1,3,8,2,5,1,2,8) to a file called ‘test’. Open the file you created in a text editor and take a look. Looking at the help files for the function write(), if necessary, write the vector out to the file ‘test’ as a single column of numbers.

20. Use the scan file to read the data from the file called test into a new vector called k

21. This time use the scan() function, with the what=list() option, to read the numbers in the file 'test' into a list called my_list such that the numbers (1,8,5,2) get put into the list element x and the numbers (3,2,1,8) get put into the list element y (i.e. you want scan to read in numbers a pair at a time and put the first number in the pair into x and the second number into y in the list my_list).

22. Use the data.frame() function to make a dataframe out of the list and save it as df1. Display df1.

23. Use the write.table function to write the dataframe you created in **9** to a file on the F drive (whenever you save a file in this course you will save it to the F drive – all data gets wiped from the local machines every time they are restarted).

24. Use the read.table function to read the data back in from the file you saved it in and save it in another data frame called df.

25. Check that it's still a data frame using the attributes() function.

26. Load the dataset cars that comes with R using the command data(cars). You can find out what kind of data object this is by typing attributes(cars). You will see that it's a dataframe with two vectors – speed and dist. You can get more information about the dataset by typing help(cars). Work out what the average car speed was in the study (don't worry about units). What is the standard deviation of the speed? What was the mean and standard deviation of the stopping distance?