Practice Problems 2

Econometrics, Spring 2019
Belk College of Business, UNCC
Ercan Karadas

1)-13) As we discussed in class the paper **A (very) short introduction to R** gives a very nice overview of R in relatively short time. It is posted on Canvas. In this exercise, I want you read and complete all **ToDo** tasks there. There are 13 of them and that is why this makes the first 13 exercises on this problem set. As stated in the document "*Going through all text and exercises takes 1-2 hours*".

14) Store the values $-20, -15, -5, 8, 12, 9, 2, 23, 19$ in the R variable `x`.

a) Use the R command `sum` to verify that the sum of the values is 33.

b) Compute an average by using the R command `mean`?

c) Compute the average with using the R command `sum`?

d) Use R to sum the positive values in `x`.

e) Use the `which` command to get the average of the values ignoring the largest value.

f) Speculate about the values corresponding to the command `x[abs(x)>=8 & x<8]`. Verify your speculation running this R command.

15) Let `x = c(1, 8, 2, 6, 3, 8, 5, 5, 5, 5)`

a) Describe two different R commands for summing the values in `x` ignoring the value 2 stored in `x[3]` and the value 3 stored in `x[5]`.

b) Use two different R commands to sum all of the values not equal to 5.

c) Use a single R command to change all values equal to 8 to 7.

16) a) Create a 10×5 matrix M whose elements are random draws from a normal distribution with mean 5 and variance 2.

b) Create another matrix N of the same size which contains all zeros, except 5 NA and locations of these NAs are randomly determined (i.e. in the sense that each time you run your code their locations are expected to change).

c) Using M and N generate a random matrix from $N(5, 2)$ which contains 5 NA values that are arbitrarily located in the matrix.

d) Describe how the R function `is.na` can be used to eliminate the rows with missing values.

17) R has a built-in data set called `chickwts`, which is stored in a data frame with two columns. The first column contains the weight of chicks, and the second column indicates the type of feed they received, one of which is labeled `horsebean`. Use R to compute the average weight among chicks that were fed horsebean.

18) Create a vector named `my_vec` containing the integers 1 through 100 and then divide each element of `my_vec` by 3 and store the result as `my_vec2`. (Your answer should contain two lines of R commands)

19) Compute the average of the vector `my_vec` you created in part (a) without using built-in R function `mean`.

20) Create a vector named `my_vec3` containing the elements of `my_vec` that are between 20 and 35. (Your answer should contain a single line of R commands)

21) Briefly explain what would be the output of following R command: `mean(rnorm(1000))`

22) What would be the output of following R commands:

```
vect <- c("foo" = 11, "bar" = 2, "norf" = NA)
vect
```

23 Create a 2 by 3 matrix named `my_matrix` containing 6 random draws from the standard normal distribution. (Your answer should contain one line of R commands)

24) Make a script file which constructs two random normal vectors of length 10. Call these vectors `x1` and `x2`. Make a data frame called `T` with two columns (called `a` and `b`) containing respectively `x1` and `x1 + x2`.

25) What would be the output of following R commands:

```
x <- seq(from=1, to=10, by=1)
for (i in x){
  if (i<5 | i>8)
  {
    x[i] = 0.5*i
  }else{
    x[i] = 5*i
  }}
x
```

26) When the function `head` called on the data frame `Orange` it produces the following output:

```
> head(Orange, n=2)
  Tree  age circumference
1    1  118            30
2    1  484            58
```

Write the command that adds up Y (defined as `circumference`) and X (defined as square root of `age`). (Your answer should contain at most three lines of R commands)

27) Put all even integers from 30 to 89 in a vector named P and then in a matrix with 6 rows and 5 columns named Q.

28) What would be the output of following R commands:

```
d <- data.frame(a = c(1,2,4,8,16,32),
                g = c(2,4,8,16,32,64),
                x = c(3,6,12,24,48,96))
```

29) `Orange` is a data frame with two numeric variables `circumference` and `age`. What would be the output of following R commands (just describe what these commands do):

```
Orange$age[Orange$age <= 900] <- 0
Orange$age[Orange$age > 900] <- 1
```

30) Declare a function in R, named `my_function` which takes a vector, say `x`, as input and returns the sum of the elements in the vector and the mean of the values in the vector.

31) Using a `for` loop in R, write a script that produces the sum of the first $n = 40$ integers.

32) What would be the output of following R commands:

```
myFun1 <- function(x){
  n = length(x)
  for (i in 1:n)  {
    if (i<5 | i>90)
    {
      x[i] = 10*i
    }else{
      x[i] = 0.1*i
    }  }
  return(x)}
myFun1(1:10)
```