assignment 3

Oyedayo Oyelowo

2 November 2017

Exercise 1

- **Search information about function read.table with help-function. What are the data types of the following parameters and how do they affect to the read.table function.**

  a) ***header:*** could be TRUE or FALSE. It is a logical value which shows if the file has variables names in the first line. If set to FALSE, the headings would be omitted. If not included, the it determines the value from the format of the file.

  b) ***dec***: this is the character used in the file for decimal points. It is a character string that contains just one single-byte character.

  c) ***nrows***: this is an integer data type which specifies the maximum number of rows to read in. Negative and other invalid values are ignored.

  d) ***skip***: this is an integer data type and it specifies the number of lines of the data file to skip before beginning to read data.

  e) ***fill***: it is a logical data type. If TRUE then in case the rows have unequal length, blank fields are implicitly added.

Excercise 2

- Your research data has been corrupted. The first rows have changed to random letters and some. values are also missing from other rows. However, you try to save what's left because you desperately need to know the mean height of the measured trees.

  a) Modify puudata_rikki.xls (found in Moodle) data into text file and import it into R with read.table function. Set the parameters so that the first two rows are left unread (the rows contain random letters) and only the first 20 rows are imported into R.

  b) Define the mean height of the trees with mean-function. Modify parameters if necessary.

```
rm(list=ls())
setwd("C:/Users/oyeda/Desktop/R_COURSE/assignment3")
#?read.table
#Modify puudata_rikki.xls (found in Moodle) data into text file and
#import it into R with read.table function.
#Set the parameters so that the first two rows are left unread (the rows
#contain random letters) and only the first 20 rows are imported into R.
data <- read.table("puudata_rikki.txt", nrows = 20, header = T, skip = 2, sep = ";")
```

- calculate the mean height of the trees

```
mean.height<-mean(data$PITUUS)
mean.height
```

```
## [1] 188.3
```

Exercise 3

- Download file puudata.xls from Moodle and transform it into text file. Define the correlation between diameter and height with cor function

  a) for spruces (PUULAJI = 2) in the first crown layer (LATVKERROS = 1)

b) for birches in the first crown layer (PUULAJI = 3).

c) For which of the previous species the diameter explains greater amount of variation in tree height?

```
#read the data "puudata" in my working directory
data2<- read.table("puudata.txt", header=T, sep = "\t")
#data2
#summary(data2)

#Define the correlation between diameter and height with cor function
#a) for spruces (PUULAJI = 2) in the first crown layer (LATVKERROS = 1)
spruceLayer1 <- subset(data2, data2$PUULAJI==2 & data2$LATVKERROS==1)
#cor.test(spruceLayer1$LPM, spruceLayer1$PITUUS, method = c("pearson", "kendall",
"spearman"))

#correlation between diameter and the height for birches in the
#first crown layer (PUULAJI = 3).
cor(spruceLayer1$LPM, spruceLayer1$PITUUS,
    method = c("pearson", "kendall", "spearman"))

## [1] 0.923803

#b) for birches in the first crown layer (PUULAJI = 3).
birchLayer1<-subset(data2, data2$PUULAJI==3 & data2$LATVKERROS==1)
cor(birchLayer1$LPM, birchLayer1$PITUUS,
    method = "pearson")

## [1] 0.7943942
```

c) For which of the previous species the diameter explains greater amount of variation in tree height?

- ***Answer: the diameter of spruce explains greater amount of variation in the treee height.***

Exercise 4

-- Using command function, construct your own function that - a) calculates the product of three inserted parameters - b) prints letter " R" x times when x is functions parameter

```
#a) function that calculates the product of three inserted parameters
prod3<- function(x,y,z){
  pr<-x*y*z
  return(pr)
}
prod3(74,11,22)  #test function

## [1] 17908

#b) function  prints letter " R" x times when x is functions parameter
printR<-function(x){
  rprint<- rep("R", x)
  return(rprint)
}
printR(21)   #test function.

##  [1] "R" "R" "R" "R" "R" "R" "R" "R" "R" "R" "R" "R" "R" "R" "R" "R" "R"
## [18] "R" "R" "R" "R"
```