assignment 2

Oyedayo Oyelowo

2 November 2017

## Exercise 1

* 1. Create folder "assignment2" and set it as your working directory through R-console
  2. Download file puudata\_20.xls from Moodle and open it in MS Excel. Column name LPM stands for diameter, ELAVALARAJA stands for the lower limit of living canopy and LATVUSLEV stands for width of the canopy.
  3. Save the data to your working directory as a text file and name it treedata.txt: save as other formats text tab limited (tab as a column delimiter).
  4. Read the file in R and name it properly.
  5. Check that the data appears right in R (columns, decimal separators, column names)
  6. Find out what are the mean diameter and mean height of the trees in the data.

#clear memory list  
rm(list=ls())  
#set working directory  
setwd("C:/Users/oyeda/Desktop/R\_COURSE/assignment2")  
#get working directory  
getwd()

## [1] "C:/Users/oyeda/Desktop/R\_COURSE/assignment2"

#read the first table  
data<-read.table("puudata\_20.txt", header = TRUE, sep="\t")  
summary(data) #summary of the data

## TUNNISTE KOEALA PUUNRO SUUNTA   
## Min. : 1.00 Min. :764.0 Min. : 1.00 Min. : 13.0   
## 1st Qu.: 5.75 1st Qu.:764.0 1st Qu.: 3.00 1st Qu.: 36.0   
## Median :10.50 Median :865.0 Median : 5.50 Median :100.5   
## Mean :10.50 Mean :834.7 Mean : 6.30 Mean :137.2   
## 3rd Qu.:15.25 3rd Qu.:865.0 3rd Qu.: 9.25 3rd Qu.:259.2   
## Max. :20.00 Max. :865.0 Max. :14.00 Max. :355.0   
##   
## ETAISYYS PUULAJI LATVKERROS LPM   
## Min. :261.0 Min. :2.0 Min. :1.00 Min. : 5.50   
## 1st Qu.:570.8 1st Qu.:2.0 1st Qu.:1.00 1st Qu.:13.40   
## Median :727.0 Median :2.0 Median :1.00 Median :26.60   
## Mean :712.0 Mean :2.8 Mean :1.15 Mean :22.35   
## 3rd Qu.:903.8 3rd Qu.:2.0 3rd Qu.:1.00 3rd Qu.:29.43   
## Max. :976.0 Max. :6.0 Max. :2.00 Max. :35.00   
##   
## PITUUS ELAVALARAJA LATVUSLEV X   
## Min. : 4.40 Min. : 1.400 Min. :1.00 Mode:logical   
## 1st Qu.:12.95 1st Qu.: 4.925 1st Qu.:1.80 NA's:20   
## Median :23.55 Median : 8.350 Median :2.50   
## Mean :18.83 Mean : 7.265 Mean :2.34   
## 3rd Qu.:24.45 3rd Qu.: 9.575 3rd Qu.:2.70   
## Max. :26.40 Max. :12.100 Max. :3.50   
##   
## X.1 X.2   
## Min. :22.35 Min. : 27.33   
## 1st Qu.:22.35 1st Qu.:443.03   
## Median :22.35 Median :593.66   
## Mean :22.35 Mean :533.33   
## 3rd Qu.:22.35 3rd Qu.:667.36   
## Max. :22.35 Max. :891.52   
## NA's :19 NA's :13

#calculate the mean diameter of the trees  
meanDiameter <- mean(data$LPM)   
meanDiameter

## [1] 22.35

#calulate the mean height of the trees  
meanHeight <- mean(data$PITUUS)   
meanHeight

## [1] 18.83

* Create a matrix from treedata columns LPM-, PITUUS- and PUULAJI.
* DBH is diameter, h is height and s is the species

dbh <- data$LPM  
h <- data$PITUUS  
s <- data$PUULAJI  
mat <- matrix(c(dbh, h, s), nrow=length(dbh))  
mat

## [,1] [,2] [,3]  
## [1,] 29.3 23.1 2  
## [2,] 33.7 25.5 2  
## [3,] 28.5 24.6 2  
## [4,] 29.8 24.7 2  
## [5,] 27.5 23.6 2  
## [6,] 19.3 19.9 2  
## [7,] 5.9 4.4 2  
## [8,] 27.5 23.6 2  
## [9,] 35.0 25.9 2  
## [10,] 25.7 23.5 2  
## [11,] 23.3 23.9 2  
## [12,] 29.3 24.4 2  
## [13,] 31.0 23.8 2  
## [14,] 32.5 26.4 2  
## [15,] 10.4 8.7 2  
## [16,] 13.1 10.1 2  
## [17,] 13.5 13.9 6  
## [18,] 5.7 6.5 6  
## [19,] 20.5 15.2 6  
## [20,] 5.5 4.9 6

#Calculate basal area (BA) for every tree and   
#add the results in the matrix as a new column. BA is basal area.  
data$BA <- with(data, (pi\*h^2)/4)  
  
#joining the BA column to the matrix created earlier  
mat <- cbind(mat, data$BA)  
  
#create column names  
colnames(mat) <- c("DBH", "H", "S", "BA")  
mat

## DBH H S BA  
## [1,] 29.3 23.1 2 419.09631  
## [2,] 33.7 25.5 2 510.70516  
## [3,] 28.5 24.6 2 475.29155  
## [4,] 29.8 24.7 2 479.16357  
## [5,] 27.5 23.6 2 437.43536  
## [6,] 19.3 19.9 2 311.02553  
## [7,] 5.9 4.4 2 15.20531  
## [8,] 27.5 23.6 2 437.43536  
## [9,] 35.0 25.9 2 526.85294  
## [10,] 25.7 23.5 2 433.73614  
## [11,] 23.3 23.9 2 448.62728  
## [12,] 29.3 24.4 2 467.59465  
## [13,] 31.0 23.8 2 444.88094  
## [14,] 32.5 26.4 2 547.39110  
## [15,] 10.4 8.7 2 59.44679  
## [16,] 13.1 10.1 2 80.11847  
## [17,] 13.5 13.9 6 151.74678  
## [18,] 5.7 6.5 6 33.18307  
## [19,] 20.5 15.2 6 181.45839  
## [20,] 5.5 4.9 6 18.85741

#save the data  
#?write.table  
write.table(mat, file = "treeData20.txt", sep="\t", col.names = TRUE, row.names = FALSE)

## Exercise 3

* 1. Download file puudata\_300.txt from Moodle, read it into R and name it.
  2. Cross tabulate the data by latvuskerros (crown layer) and puulaji (tree species) with function xtabs()
  3. Which is the most common tree species in crown layer 1? How many trees of this species can be found in the whole data set?
  4. Which tree species' relative portion in second crown layer is the highest (amount in 2.layer/amount in layer 1 and 2)? Consider only those species that are present in both crown layers.

#load the data "puudata\_300.txt" from the directory.  
data300 <- read.table("puudata\_300.txt", sep = "\t", header = TRUE)  
#?xtabs  
  
#s.tab<-table(data300$LATVKERROS, data300$PUULAJI)  
layerSpp <-xtabs(~LATVKERROS+PUULAJI, data300)  
layerSpp

## PUULAJI  
## LATVKERROS 1 2 3 4 5 6 8 11 12  
## 1 91 79 21 32 3 6 1 0 1  
## 2 0 48 3 11 0 3 0 1 0

* **Which is the most common tree species in crown layer 1?**
* answer: the most common tree species in the crown layer 1 is species 1
* **How many trees of this species can be found in the whole data set?**
* Answer: there are 91 of these trees in the whole data set.
  1. Which tree species' relative portion in second crown layeris the highest (amount in 2.layer/amount in layer 1 and 2)?
* divide the total in layer 2 by the sum of the total in layer 1 and 2. Consider only those species that are present in both crown layers.

highSpp2 <- layerSpp[2,] / (layerSpp[2,] + layerSpp[1,])  
highSpp2

## 1 2 3 4 5 6 8   
## 0.0000000 0.3779528 0.1250000 0.2558140 0.0000000 0.3333333 0.0000000   
## 11 12   
## 1.0000000 0.0000000

* answer: species 2 has the highest relative portion of crown layer2 in layers 1 and 2 crown layers, considering those that have both layers present.

## Exercise 4

* Create the following subsets from file puudata\_300.txt:
  1. Trees that are measured from plot (KOEALA) 865
  2. Trees that are measured from plots 865 and 490
  3. Those spruce trees (PUULAJI=2) that belong to second crown layer and are over 10 meters tall (the height is given in decimeters)
  4. Those trees in the first crown layer, whose diameter is over 150 mm and that are not pines or spruces (species 1 and 2)

#Create the following subsets from file puudata\_300.txt:  
#a) Trees that are measured from plot (KOEALA) 865  
treesA <- subset(data300, data300$KOEALA==865)  
head(treesA)

## TUNNISTE KOEALA PUUNRO SUUNTA ETAISYYS PUULAJI LATVKERROS LPM PITUUS  
## 1 1 865 1 16 750 2 1 293 231  
## 2 2 865 2 23 507 2 1 337 255  
## 3 3 865 3 81 901 2 1 285 246  
## 4 4 865 4 84 480 2 1 298 247  
## 5 5 865 5 117 912 2 1 275 236  
## 6 6 865 6 147 644 2 1 193 199  
## ELAVALARAJA LATVUSLEV  
## 1 72 24  
## 2 88 27  
## 3 107 25  
## 4 89 27  
## 5 95 25  
## 6 75 21

#b) Trees that are measured from plots 865 and 490  
treesB<- subset(data300, data300$KOEALA==865 | data300$KOEALA==490)  
summary(treesB)

## TUNNISTE KOEALA PUUNRO SUUNTA   
## Min. : 1.0 Min. :490.0 Min. : 1.00 Min. : 13.00   
## 1st Qu.:174.5 1st Qu.:490.0 1st Qu.: 7.75 1st Qu.: 83.25   
## Median :241.5 Median :490.0 Median :14.50 Median :164.50   
## Mean :188.2 Mean :583.8 Mean :18.43 Mean :175.23   
## 3rd Qu.:255.2 3rd Qu.:583.8 3rd Qu.:28.25 3rd Qu.:259.25   
## Max. :269.0 Max. :865.0 Max. :46.00 Max. :355.00   
## ETAISYYS PUULAJI LATVKERROS LPM   
## Min. : 156.0 Min. :1.000 Min. :1.000 Min. : 50.00   
## 1st Qu.: 510.0 1st Qu.:1.000 1st Qu.:1.000 1st Qu.: 95.25   
## Median : 741.5 Median :2.000 Median :1.000 Median :220.50   
## Mean : 712.7 Mean :1.839 Mean :1.268 Mean :200.30   
## 3rd Qu.: 890.5 3rd Qu.:2.000 3rd Qu.:2.000 3rd Qu.:275.00   
## Max. :1434.0 Max. :4.000 Max. :2.000 Max. :350.00   
## PITUUS ELAVALARAJA LATVUSLEV   
## Min. : 42.0 Min. : 0.00 Min. :10.00   
## 1st Qu.:105.5 1st Qu.: 65.75 1st Qu.:15.75   
## Median :195.5 Median :101.00 Median :19.50   
## Mean :170.1 Mean : 92.70 Mean :20.43   
## 3rd Qu.:232.2 3rd Qu.:137.50 3rd Qu.:25.25   
## Max. :264.0 Max. :169.00 Max. :36.00

#c) Those spruce trees (PUULAJI=2) that belong to second crown layer   
#and are over 10 meters tall (the height is given in decimeters).  
treesC <- subset(data300, data300$PUULAJI==2 & data300$LATVKERROS==2 & data300$PITUUS>100)  
head(treesC)

## TUNNISTE KOEALA PUUNRO SUUNTA ETAISYYS PUULAJI LATVKERROS LPM PITUUS  
## 16 16 764 2 29 970 2 2 131 101  
## 36 36 764 22 216 729 2 2 117 102  
## 293 293 505 24 180 256 2 2 139 104  
## ELAVALARAJA LATVUSLEV  
## 16 14 17  
## 36 19 18  
## 293 10 17

#d) Those trees in the first crown layer, whose diameter is over 150 mm and   
#that are not pines or spruces (species 1 and 2)  
treesD<- subset(data300, data300$LATVKERROS==1 & data300$LPM>150 & data300$PUULAJI!=1 & data300$PUULAJI!=2)  
head(treesD)

## TUNNISTE KOEALA PUUNRO SUUNTA ETAISYYS PUULAJI LATVKERROS LPM PITUUS  
## 19 19 764 5 37 726 6 1 205 152  
## 21 21 764 7 71 954 3 1 207 197  
## 27 27 764 13 139 835 3 1 290 231  
## 31 31 764 17 186 598 3 1 162 196  
## 34 34 764 20 197 300 3 1 274 247  
## 38 38 764 24 226 592 3 1 215 232  
## ELAVALARAJA LATVUSLEV  
## 19 52 35  
## 21 91 29  
## 27 79 40  
## 31 101 22  
## 34 81 40  
## 38 127 25