Laboratory Exercise Week 2

Ryan Estes | Section 006

8/30/2017

*Directions*:

* Write your R code inside the code chunks after each question.
* Write your answer comments after the # sign.
* To generate the word document output, click the button Knit and wait for the word document to appear.
* RStudio will prompt you (only once) to install the knitr package.
* Submit your completed laboratory exercise using Blackboard's Turnitin feature. Your Turnitin upload link is found on your Blackboard Course shell under the Laboratory folder.

1. You will work with a matrix whose entries are all prime numbers below 30.
   1. Create this matrix using the function matrix() with five rows. Save the matrix as P.
   2. Extract the second and third row out of P.
   3. Extract the entry in the fourth row and first column of `P.
   4. Generate the transpose of the matrix using the function t(). What is the new dimension of this matrix.

## Code chunk

# Insert your code for this question after this line  
#i  
P <- matrix(c(2,3,5,7,11,13,17,19,23,29), nrow = 5)  
   
#ii  
row2 <- P[2, ]  
row3 <- P[3, ]  
  
#iii  
row4col1 <- P[4,1]  
  
#iv  
transpose <- t(P)  
## The new dimension is a 5 column matrix instread of 5 rows.

1. Create and manipulate a data frame.
   1. Create a data.frame named my.trees that has the following columns:
      * Girth = c(8.3, 8.6, 8.8, 10.5, 10.7, 10.8, 11.0)
      * Height = c(70, 65, 63, 72, 81, 83, 66)
      * Volume = c(10.3, 10.3, 10.2, 16.4, 18.8, 19.7, 15.6)
   2. Extract the fifth observational unit.
   3. Extract the Girth column referring to it by name.
   4. Print out a data frame of all the observations except for the last observation.

## Code chunk

# Insert your code for this question after this line  
#i  
my.trees <- data.frame(  
 Girth = c(8.3, 8.6, 8.8, 10.5, 10.7, 10.8, 11.0),  
 Height = c(70, 65, 63, 72, 81, 83, 66),  
 Volume = c(10.3, 10.3, 10.2, 16.4, 18.8, 19.7, 15.6)  
)  
  
#ii  
my.trees[5,]

## Girth Height Volume  
## 5 10.7 81 18.8

#iii  
my.trees[, "Girth"]

## [1] 8.3 8.6 8.8 10.5 10.7 10.8 11.0

#iv  
print(my.trees[1:6,])

## Girth Height Volume  
## 1 8.3 70 10.3  
## 2 8.6 65 10.3  
## 3 8.8 63 10.2  
## 4 10.5 72 16.4  
## 5 10.7 81 18.8  
## 6 10.8 83 19.7

1. The RailTrail dataset within the mosaic package includes the counts of crossings of a rail trail in Northampton, Massachusetts for 90 days in 2005. City officials are interested in understanding usage of the trail network, and how it changes as a function of temperature and day of the week.
   1. Check the structure of theRailTrail.
   2. How many variables and observations are in the data set?
   3. Which variables are integer type?
   4. Display the first 4 rows of the data.

## Code chunk

# Insert your code for this question after this line  
#install.packages("mosaic") #needed for RailTrail  
  
#i  
RailTrail <- mosaicData::RailTrail  
str(RailTrail)

## 'data.frame': 90 obs. of 10 variables:  
## $ hightemp : int 83 73 74 95 44 69 66 66 80 79 ...  
## $ lowtemp : int 50 49 52 61 52 54 39 38 55 45 ...  
## $ avgtemp : num 66.5 61 63 78 48 61.5 52.5 52 67.5 62 ...  
## $ spring : int 0 0 1 0 1 1 1 1 0 0 ...  
## $ summer : int 1 1 0 1 0 0 0 0 1 1 ...  
## $ fall : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ cloudcover: num 7.6 6.3 7.5 2.6 10 ...  
## $ precip : num 0 0.29 0.32 0 0.14 ...  
## $ volume : int 501 419 397 385 200 375 417 629 533 547 ...  
## $ weekday : Factor w/ 2 levels "0","1": 2 2 2 1 2 2 2 1 1 2 ...

#ii  
#RailTrail contains 90 observations and 10 variables.  
  
#iii  
#Variables of type integer:  
#hightemp  
#lowtemp  
#spring  
#summer  
#fall  
#volume  
  
#iv  
head(RailTrail, 4)

## hightemp lowtemp avgtemp spring summer fall cloudcover precip volume  
## 1 83 50 66.5 0 1 0 7.6 0.00 501  
## 2 73 49 61.0 0 1 0 6.3 0.29 419  
## 3 74 52 63.0 1 0 0 7.5 0.32 397  
## 4 95 61 78.0 0 1 0 2.6 0.00 385  
## weekday  
## 1 1  
## 2 1  
## 3 1  
## 4 0

# last R code line

1. You will work with the passenger Titanic data found on this [link](https://goo.gl/NHb1Pg).
   1. Read this data directly from the web using its URL.
   2. Read this data by saving it first into your working directory. Check your working directory using getwd() or change its location using the steps covered in the lesson.
   3. How many passengers and what variables are in the data set?
   4. Display the first 5 rows of the data.

## Code chunk

# Insert your code for this question after this line  
#i  
TitanicPass.data <- read.csv("https://goo.gl/NHb1Pg", header = TRUE)  
  
#ii  
currentdir <- getwd()  
LocalTitanicPass.data <- "/localtitanicdata.csv"  
write.table(TitanicPass.data, file=paste(currentdir, LocalTitanicPass.data, sep = ""))  
  
#iii  
#passengers = 1316  
#variables include:  
#X  
#class  
#age  
#sex  
#survived  
  
#iv  
head(TitanicPass.data, 5)

## X class age sex survived  
## 1 1 1st class adults man yes  
## 2 2 1st class adults man yes  
## 3 3 1st class adults man yes  
## 4 4 1st class adults man yes  
## 5 5 1st class adults man yes