**ICS-321**

**Data Warehousing and Data Mining**

**LAB-7**

**2020BCS0082**

**ADITYA SANJAY SHITALE**

1. Find Sum, Mean and Product of a Vector in R

Case 1: vec = c(1.1, 2, 3.0, 4.2 ) Case 2: vec = c(1.1,NA, 2, 3.0,NA )

## 1. Find Sum, Mean and Product of a Vector in R

      vec = c(1.1, 2, 3.0, 4.2)

      vec

      sum\_vec = sum(vec)

      sum\_vec

      mean\_vec = mean(vec)

      mean\_vec

      product\_vec = prod(vec)

      product\_vec

      vec1 = c(1.1,NA, 2, 3.0,NA)

      vec1

      sum\_vec1 = sum(vec1, na.rm = TRUE)

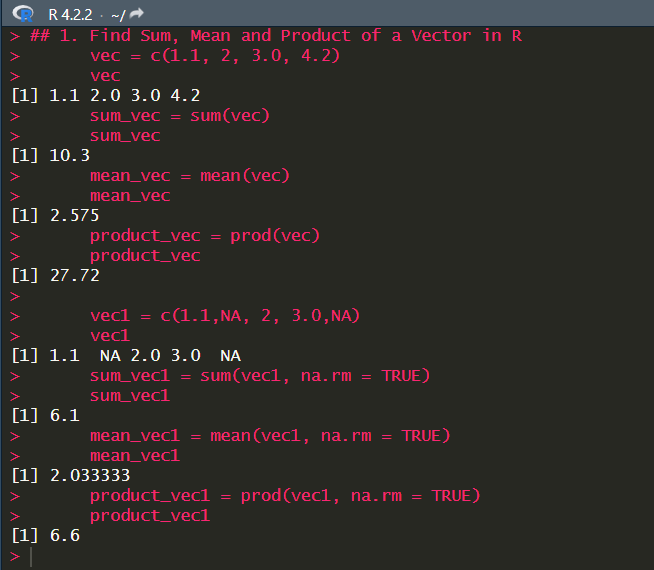
      sum\_vec1

      mean\_vec1 = mean(vec1, na.rm = TRUE)

      mean\_vec1

      product\_vec1 = prod(vec1, na.rm = TRUE)

      product\_vec1

****

2. Count Number of Occurrences of Certain Character in String in R

Case 1: str = c("$I%Love!R programming %","cs^e%portal", "le%.5%rty.in","join2022%")

Find the count of character “%” and “r” in the string.

## 2. Count Number of Occurrences of Certain Character in String in R

## Case 1: str = c("$I%Love!R programming %","cs^e%portal", "le%.5%rty.in","join2022%")

## Find the count of character “%” and “r” in the string.

      library(stringr)

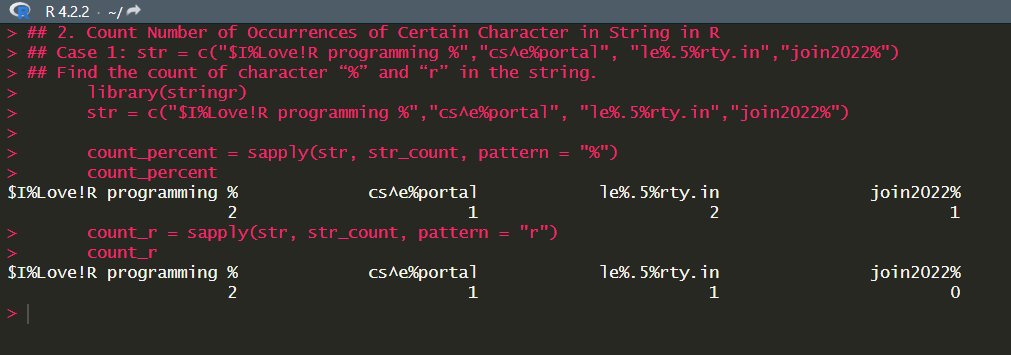
      str = c("$I%Love!R programming %","cs^e%portal", "le%.5%rty.in","join2022%")

      count\_percent = sapply(str, str\_count, pattern = "%")

      count\_percent

      count\_r = sapply(str, str\_count, pattern = "r")

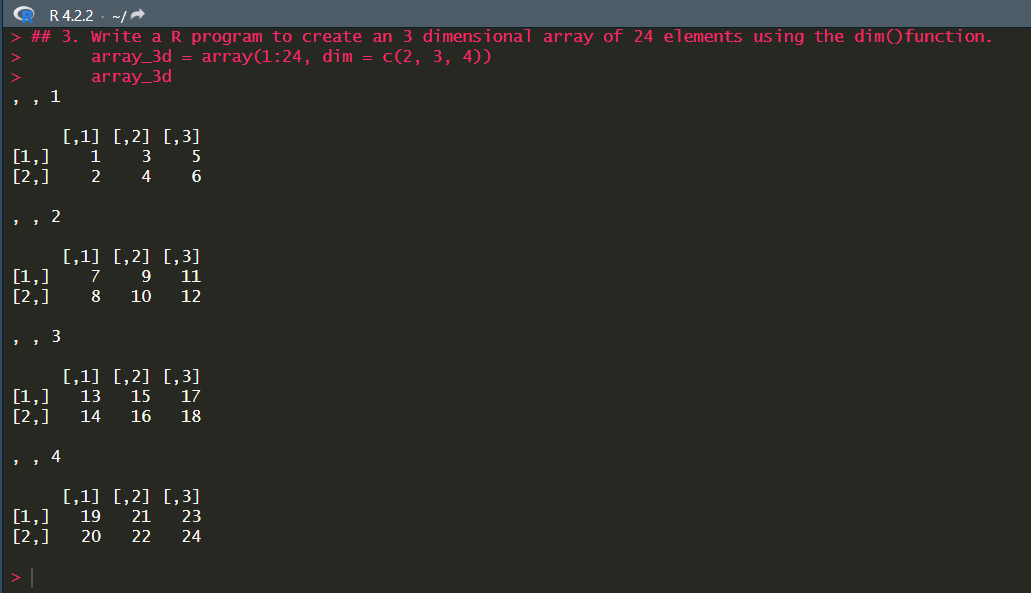
      count\_r

****

3. Write a R program to create an 3 dimensional array of 24 elements using the dim() function.

      array\_3d = array(1:24, dim = c(2, 3, 4))

      array\_3d

****

4. Write a R program to print the numbers from 1 to 100 and print "Buy" for multiples of 4, print "Now" for multiples of 5, and print "BuyNow" for multiples of both.

 for (i in 1:100) {

        if (i %% 4 == 0 && i %% 5 == 0) {

          print("BuyNow")

        } else if (i %% 4 == 0) {

          print("Buy")

        } else if (i %% 5 == 0) {

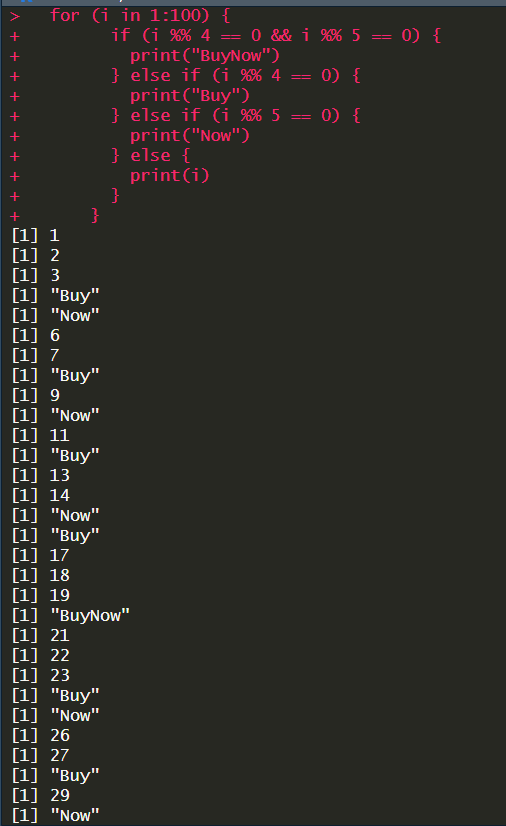
          print("Now")

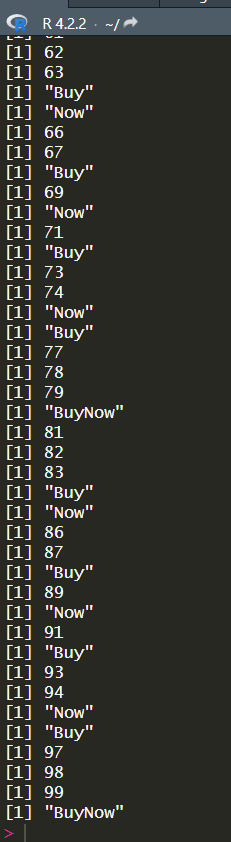
        } else {

          print(i)

        }

      }

****

****

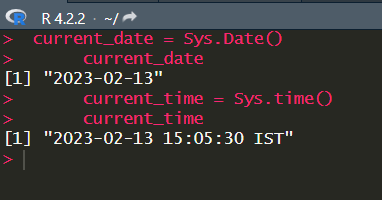
5. Write a R program to create the system's current date with and without time.

   current\_date = Sys.Date()

      current\_date

      current\_time = Sys.time()

      current\_time

****

6. Task will be based on mtcars dataset. This dataset consists of information regarding car characteristics like weight, fuel consumption, transmission type etc. Solve thefollowing queries in R using Select, Filter, Mutate, Summarise and Arrange functions.

  mtcars=read.csv("mtcars.csv")

      mtcars

      ##1. Select random n rows.

      library(dplyr)

      n = 10

      mtcars %>% sample\_n(n)

      ##2. Select random fraction of rows

      fraction = 0.1

      mtcars %>% sample\_frac(fraction)

      ##3.  Select cyl, hp, and wt columns from the mtcars dataset.

      mtcars %>% select(cyl, hp, wt)

      ##4. Filter the data by taking only cars that have an automatic transmission.

      mtcars %>% filter(am == 0)

      ##5. Retrieve cars with horsepower larger than 225.

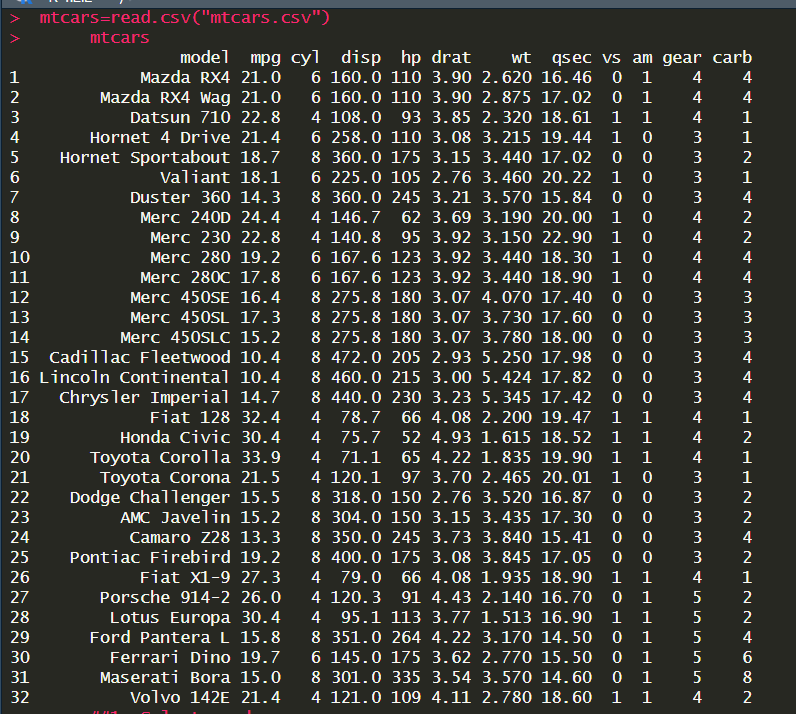
      mtcars %>% filter(hp > 225)

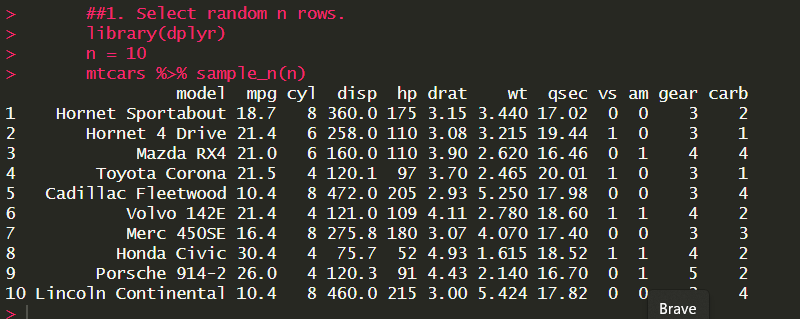
      ##6. Calculate the mean & median for the variable mpg.

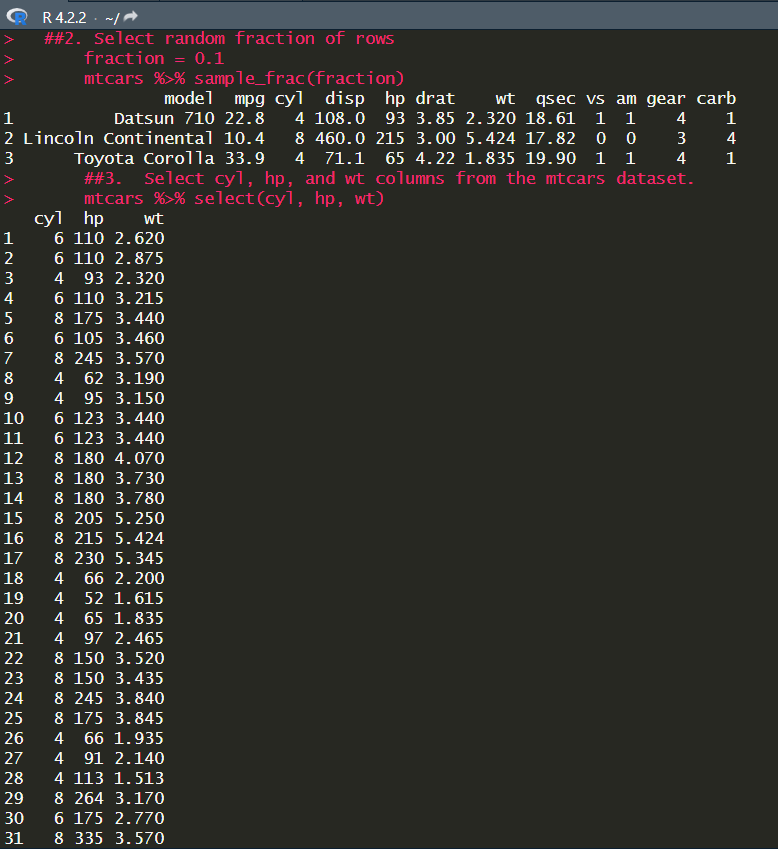
      mtcars %>% summarize(mean\_mpg = mean(mpg), median\_mpg = median(mpg))

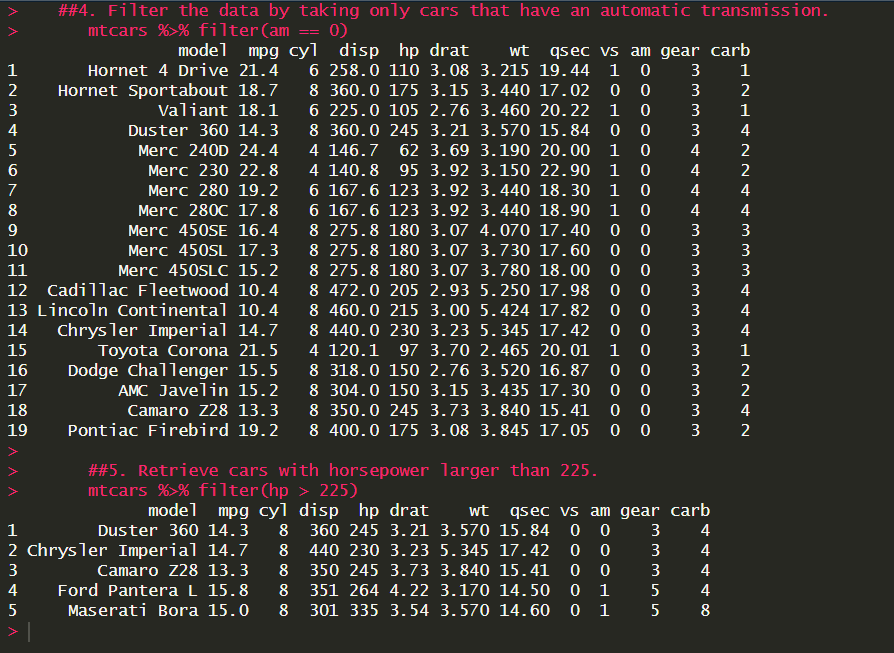
      ##7. Sort the variable hp in descending order.

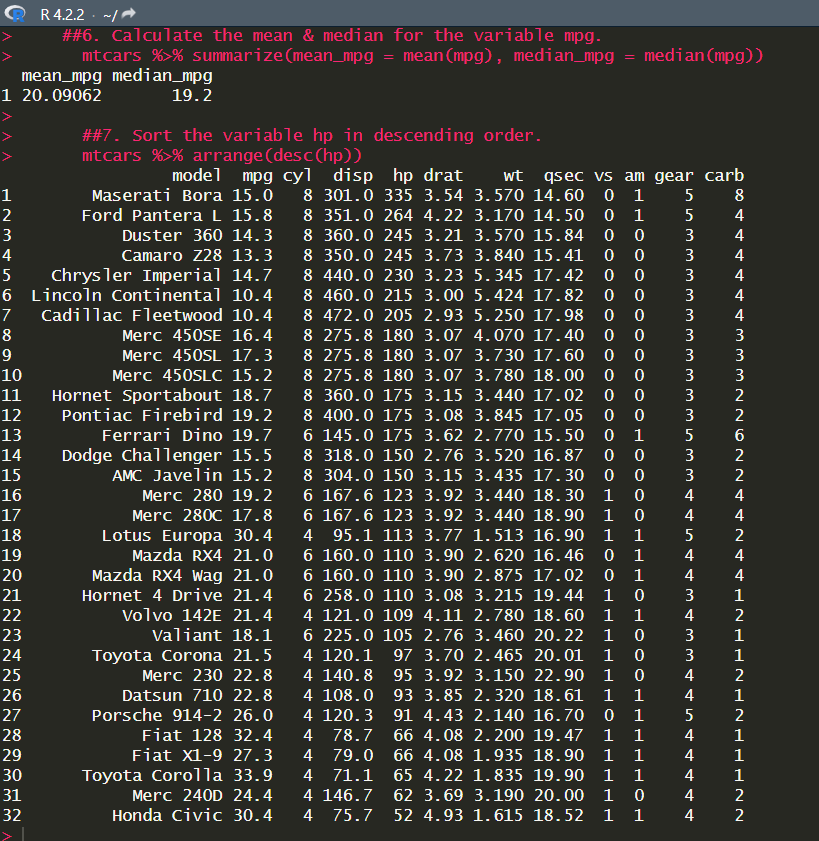
      mtcars %>% arrange(desc(hp))

****

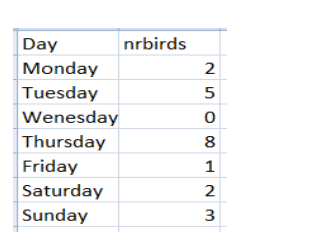
****

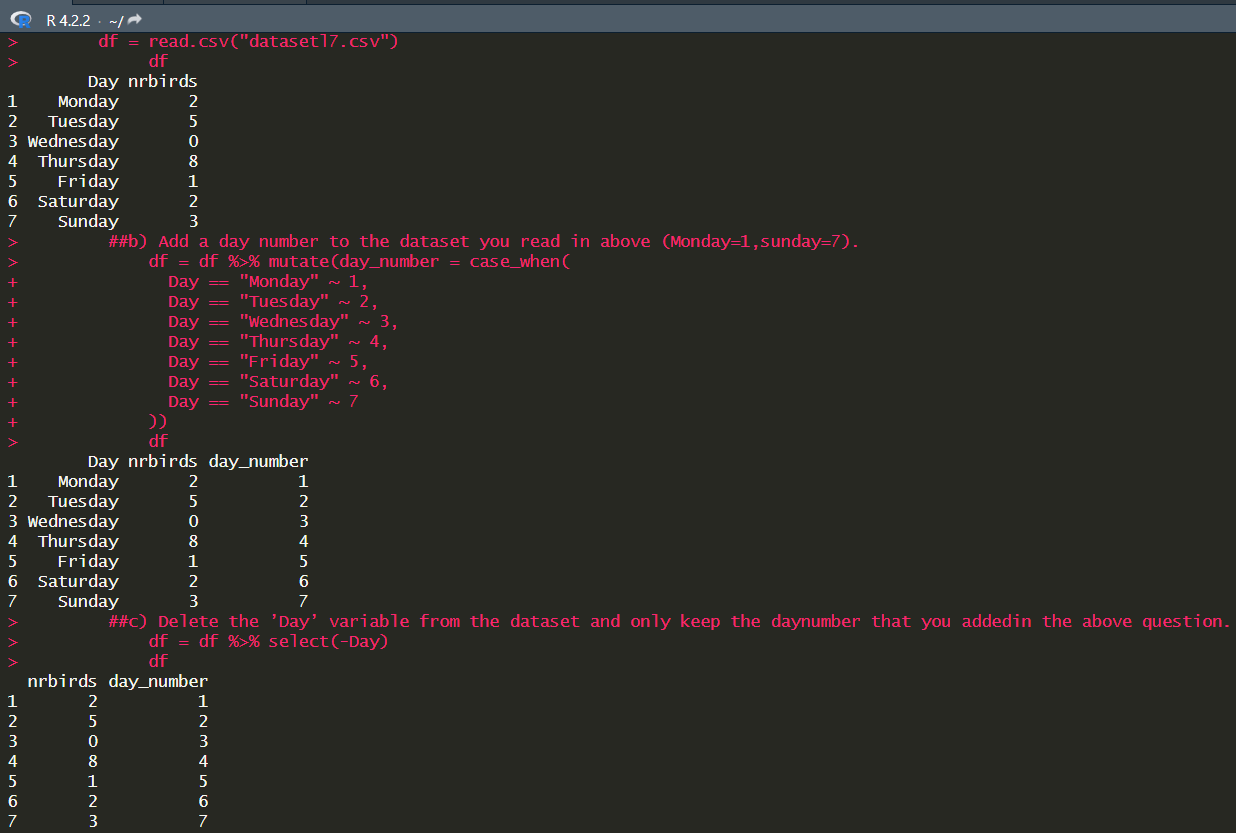
****

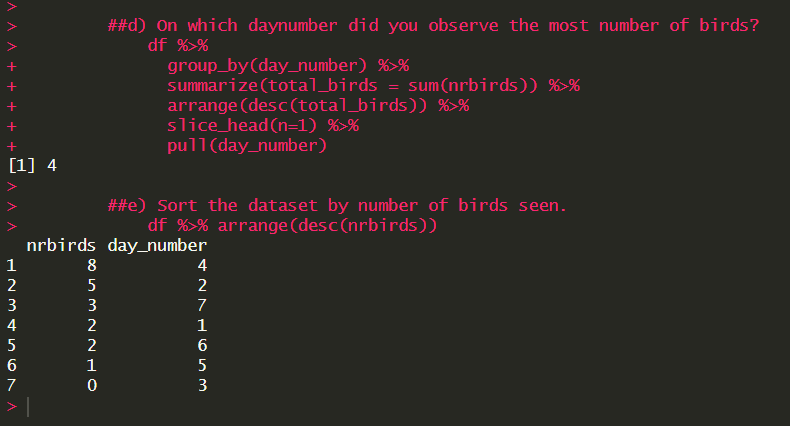
****

****

7. Answer the following with respect to the given dataset. Write the corresponding R code for all the questions.

****

****

****