name: ramya sri.v

reg.no:192124039

Day 3 Lab Manual

UNIVARIATE ANALYSIS IN R - MEASURES OF CENTRAL TENDENCY

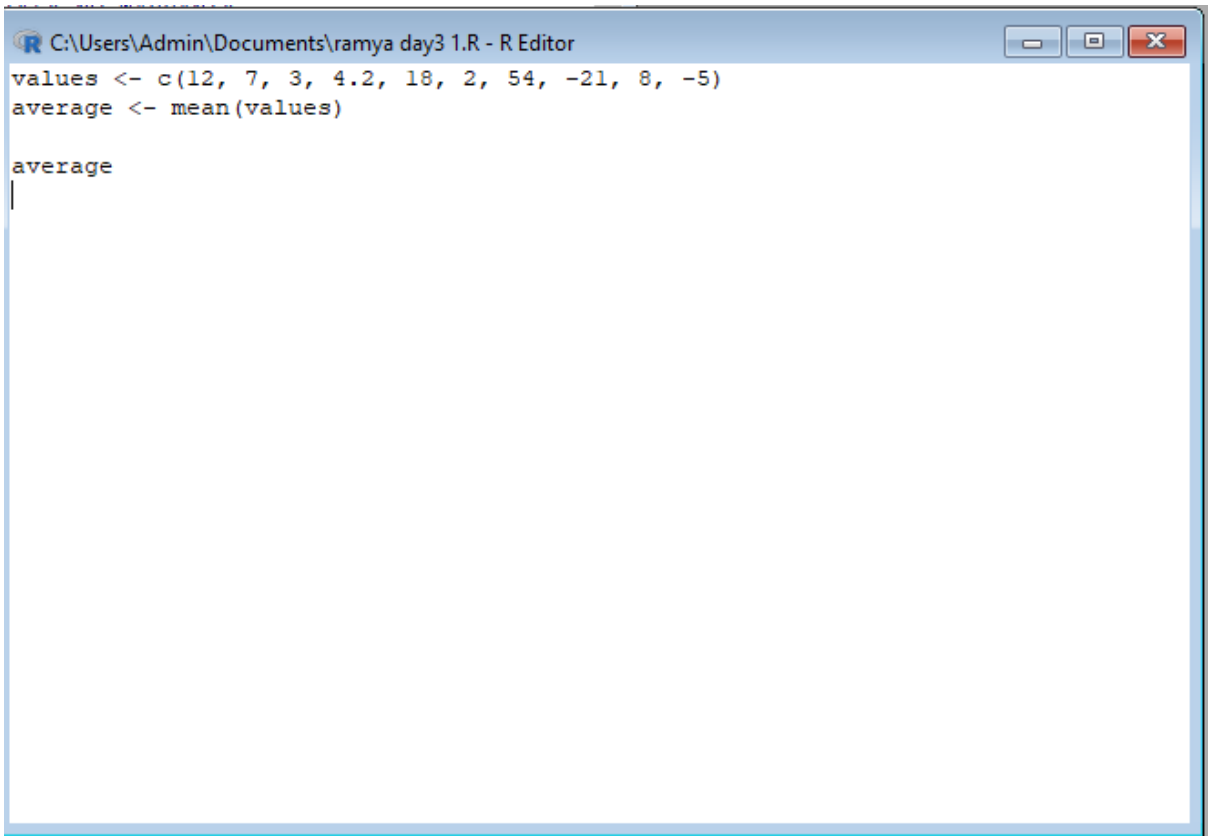
Exercise:

I. ARITHMETIC MEAN

a) Write suitable R code to compute the average of the following values.

12,7,3,4.2,18,2,54,-21,8,-5

program:



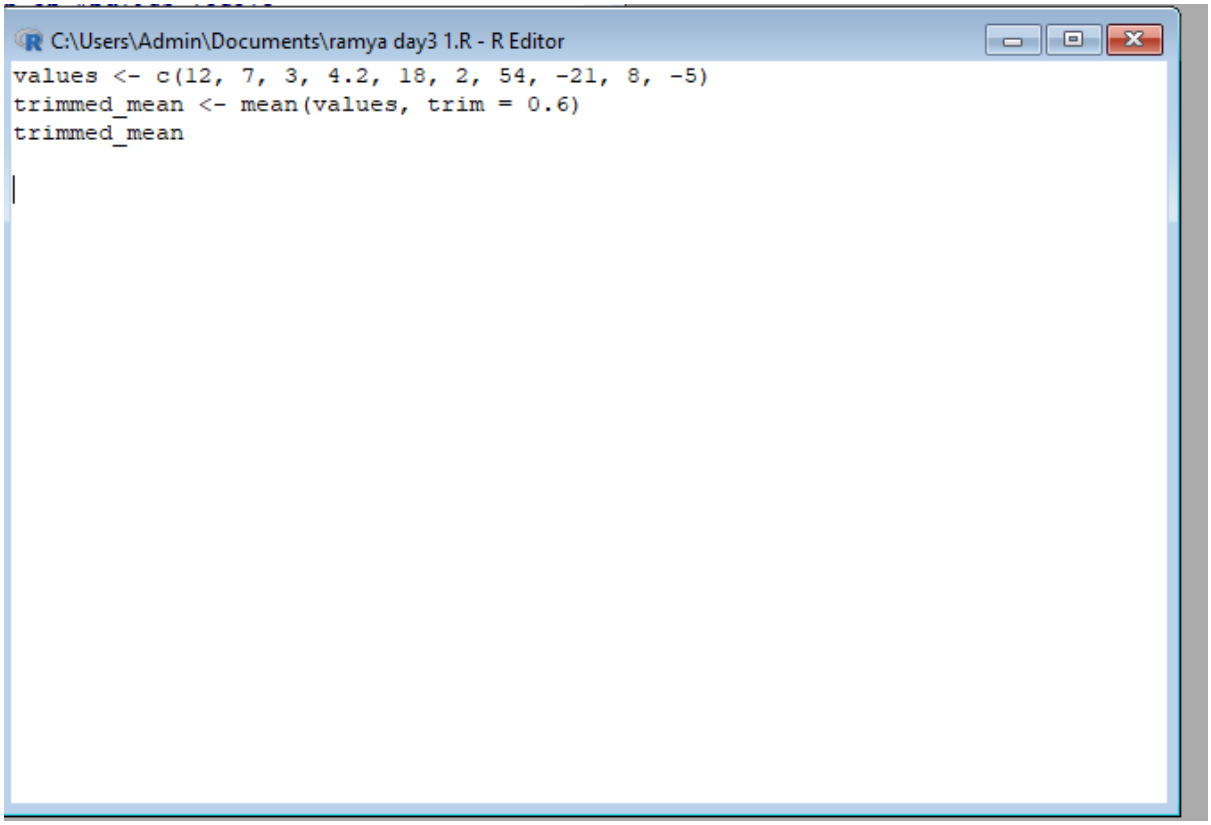
> values <- c(12, 7, 3, 4.2, 18, 2, 54, -21, 8, -5)

```
> average <- mean(values)
>
> average
[1] 8.22
>
```

b) Compute the mean after applying the trim option and removing 3 values from each end.

program:

```
> values <- c(12, 7, 3, 4.2, 18, 2, 54, -21, 8, -5)
> trimmed_mean <- mean(values, trim = 0.6)
> trimmed_mean
[1] 5.6
```



c) Compute the mean of the following vector .

(12,7,3,4.2,18,2,54,-21,8,-5,NA)

#If there are missing values, then the mean function returns NA.

Find mean dropping NA values.

#To drop the missing values from the calculation use na.rm = TRUE

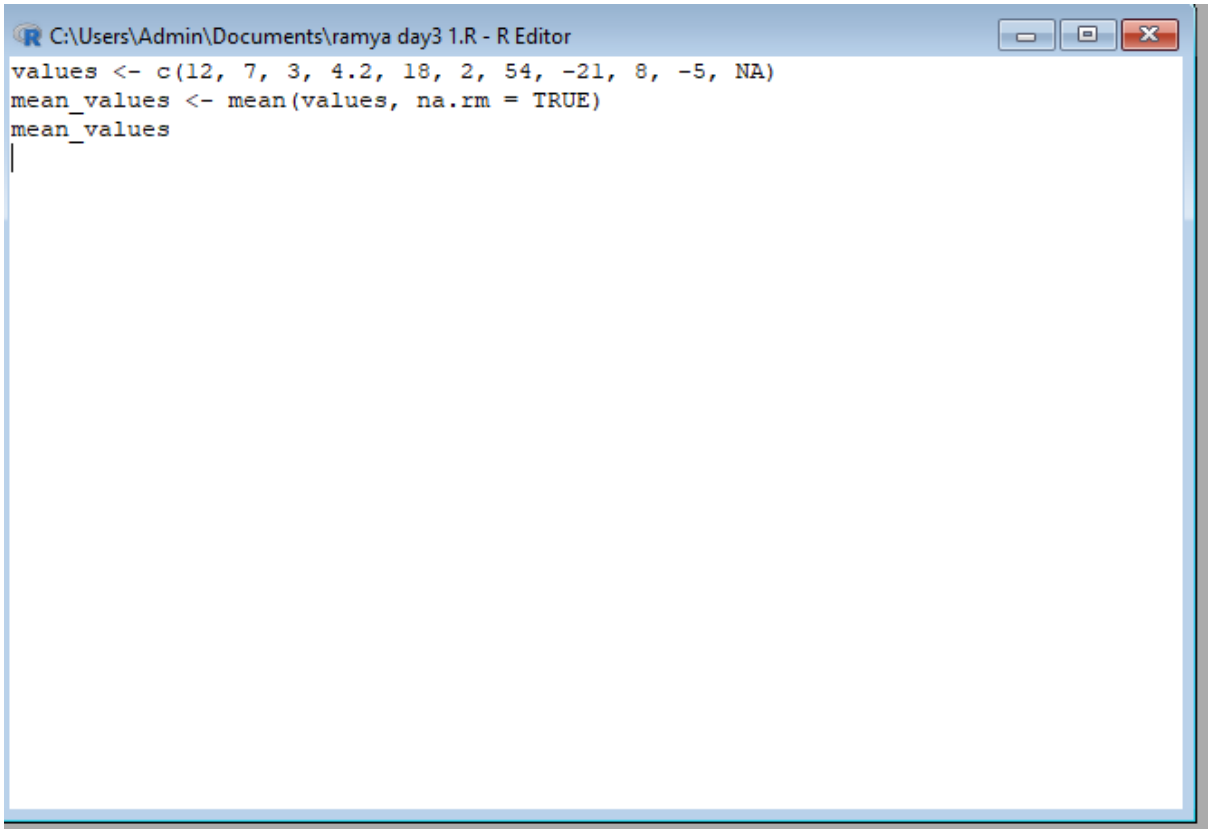
program:

> values <- c(12, 7, 3, 4.2, 18, 2, 54, -21, 8, -5, NA)

> mean_values <- mean(values, na.rm = TRUE)

> mean_values

[1] 8.22



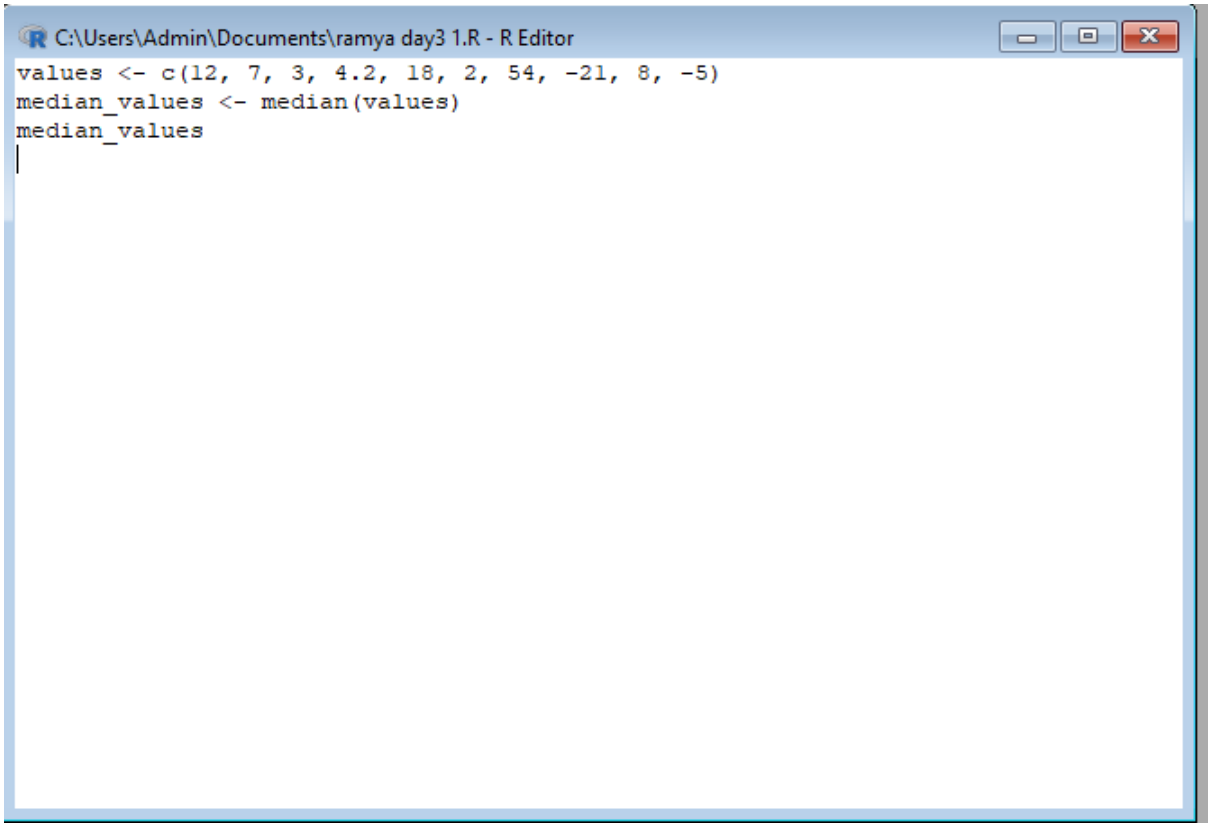
II.MEDIAN

Write suitable R code to compute the median of the following values.

12,7,3,4.2,18,2,54,-21,8,-5

program:

```
> values <- c(12, 7, 3, 4.2, 18, 2, 54, -21, 8, -5)
> median_values <- median(values)
> median_values
[1] 5.6
```



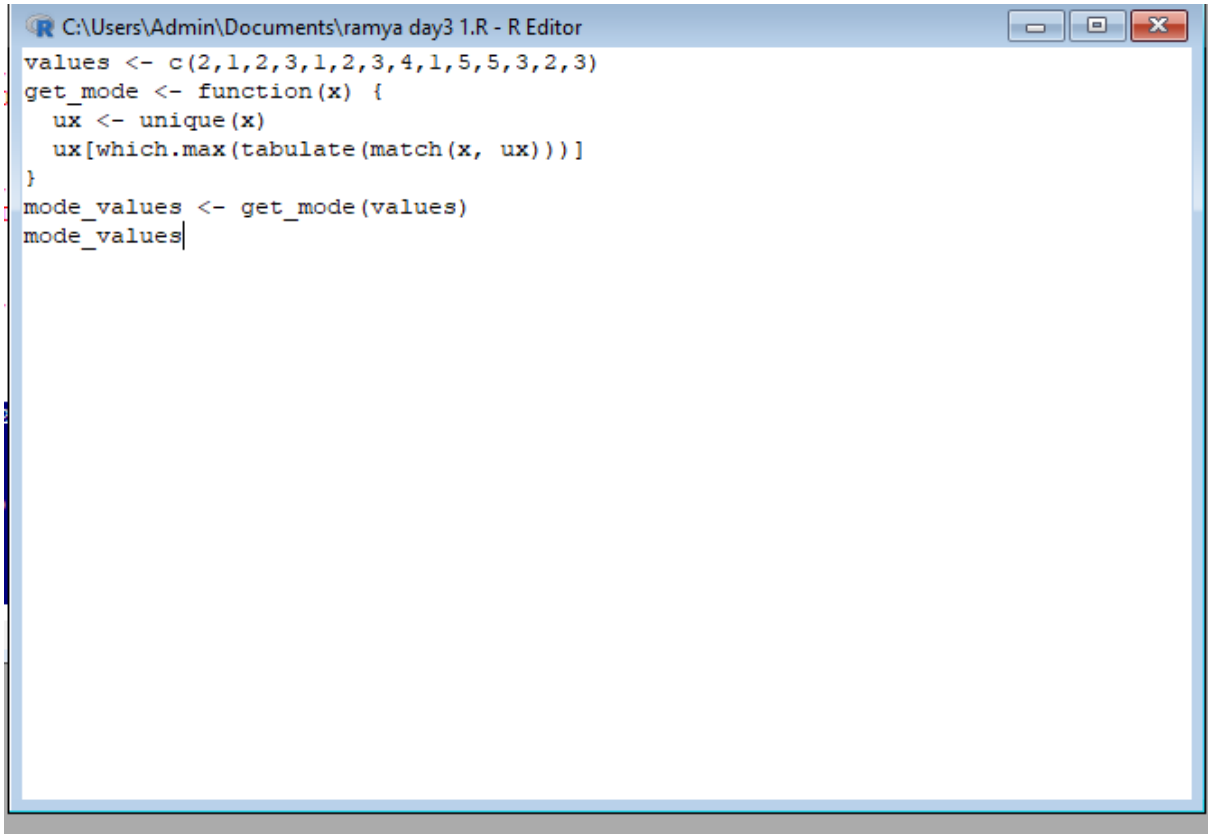
III. MODE

Calculate the mode for the following numeric as well as character data set in R.

(2,1,2,3,1,2,3,4,1,5,5,3,2,3) , ("o","it","the","it","it")

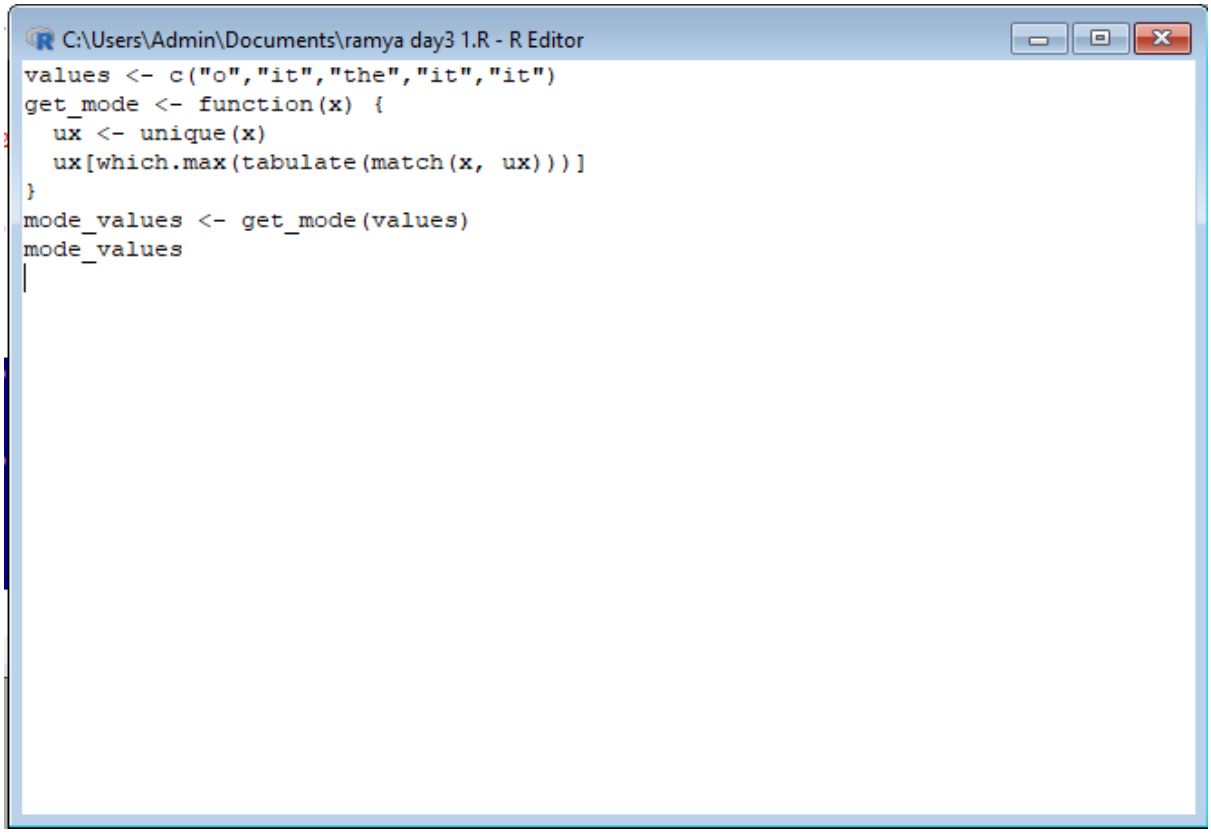
program for numeric values:

```
> values <- c(2,1,2,3,1,2,3,4,1,5,5,3,2,3)
> get_mode <- function(x) {
+   ux <- unique(x)
+   ux[which.max(tabulate(match(x, ux)))]
+ }
> mode_values <- get_mode(values)
> mode_values
[1] 2
```



program for characters:

```
> values <- c("o","it","the","it","it")
> get_mode <- function(x) {
+   ux <- unique(x)
+   ux[which.max(tabulate(match(x, ux)))]
+ }
> mode_values <- get_mode(values)
> mode_values
[1] "it"
>
```



UNIVARIATE ANALYSIS IN R - MEASURES OF DISPERSION

Exercise: 4

Download mpg dataset which contains Fuel economy data from 1999 and 2008 for 38 popular models of car from the URL given below.

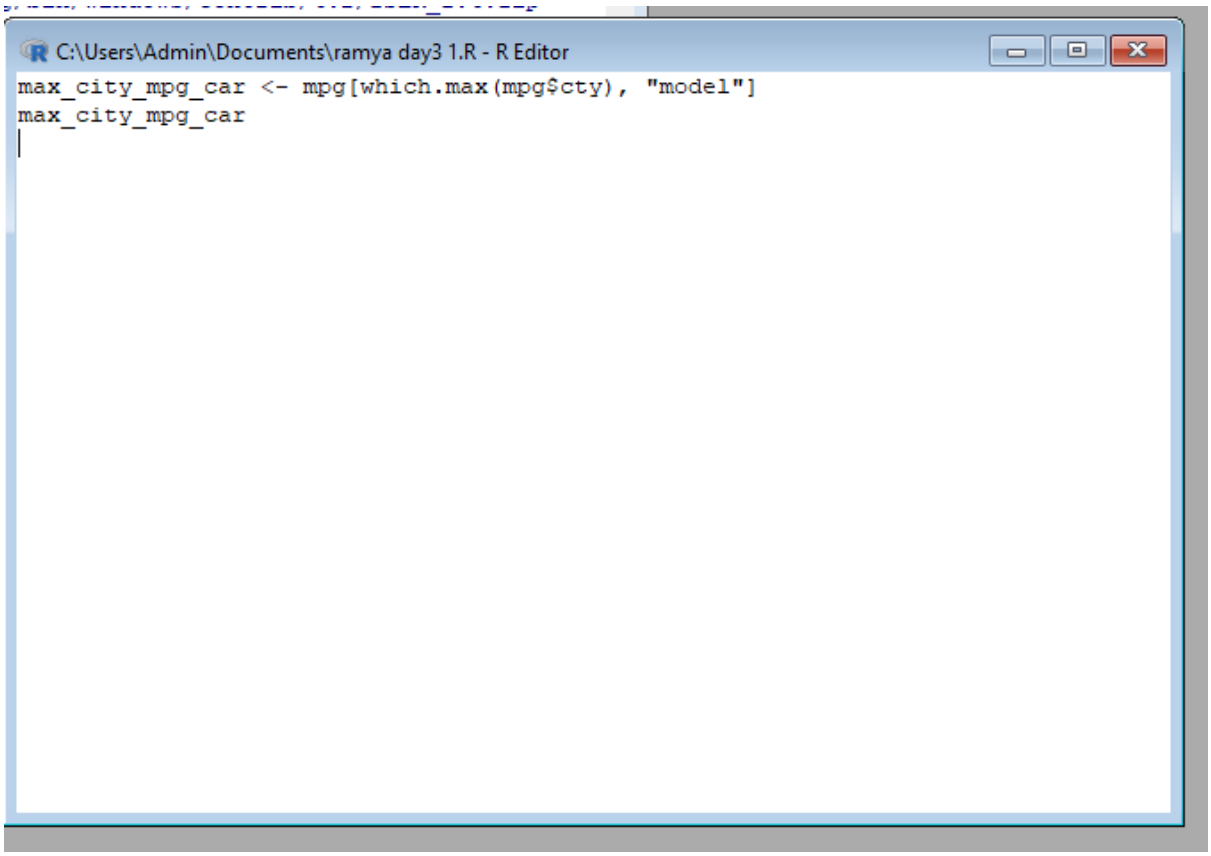
https://vincentarelbundock.github.io/Rdatasets/datasets.html

Answer the following queries

i) Find the car which gives maximum city miles per gallon

```
max_city_mpg_car <- mpg[which.max(mpg$cty), "model"]
```

max_city_mpg_car

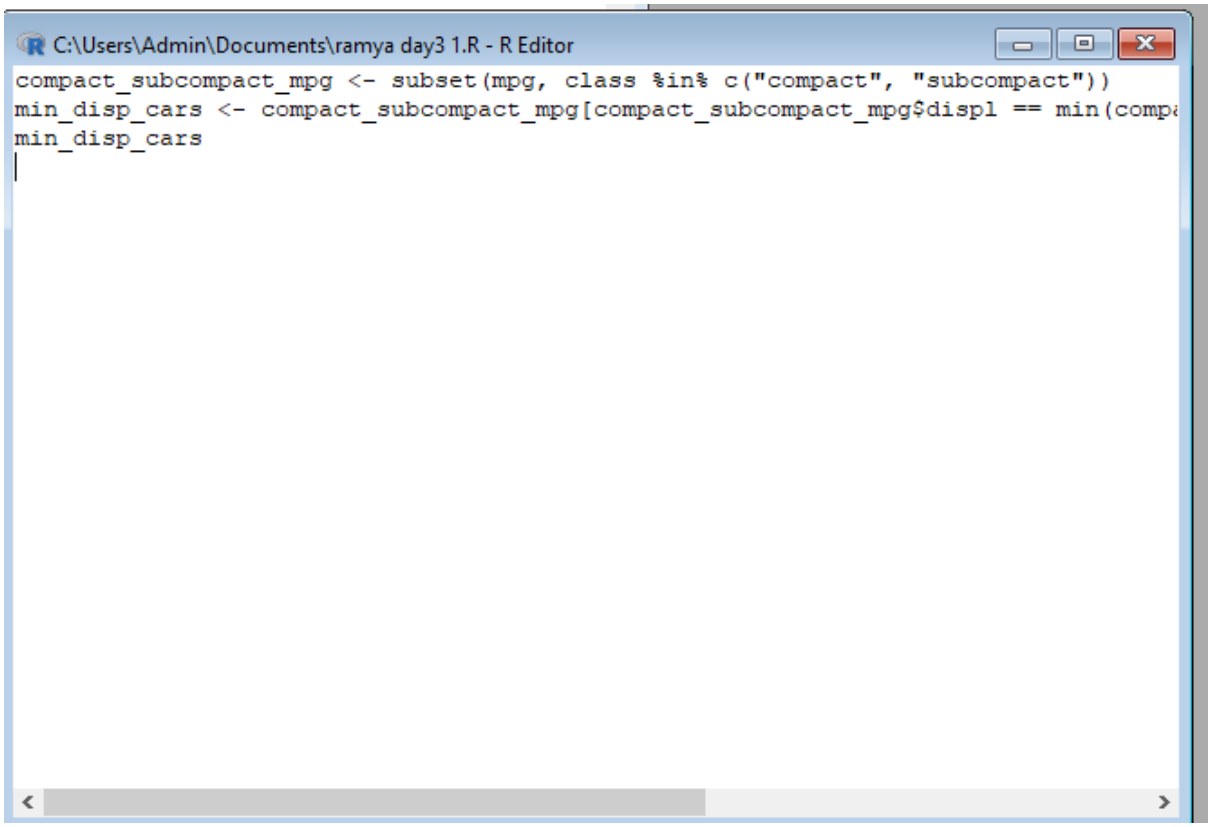


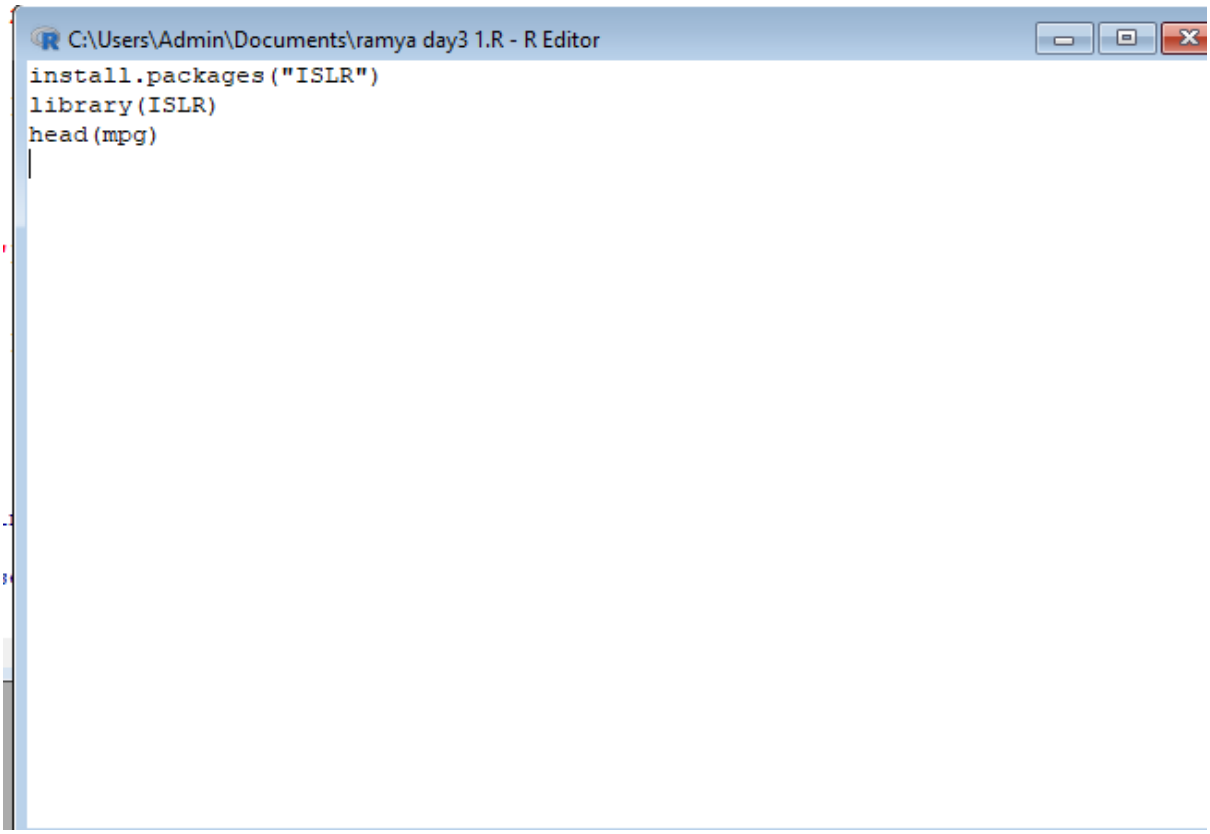
ii) Find the cars which gives minimum disp in compact and subcompact class

compact_subcompact_mpg <- subset(mpg, class %in% c("compact", "subcompact"))

min_disp_cars <- compact_subcompact_mpg[compact_subcompact_mpg$displ == min(compact_subcompact_mpg$displ), c("manufacturer", "model")]

min_disp_cars


C:\Users\Admin\Documents\ramya day3 1.R - R Editor
compact_subcompact_mpg <- subset(mpg, class %in% c("compact", "subcompact"))
min_disp_cars <- compact_subcompact_mpg[compact_subcompact_mpg$displ == min(compa
min_disp_cars


C:\Users\Admin\Documents\ramya day3 1.R - R Editor
install.packages("ISLR")
library(ISLR)
head(mpg)

install.packages("ISLR")

library(ISLR)

head(mpg)

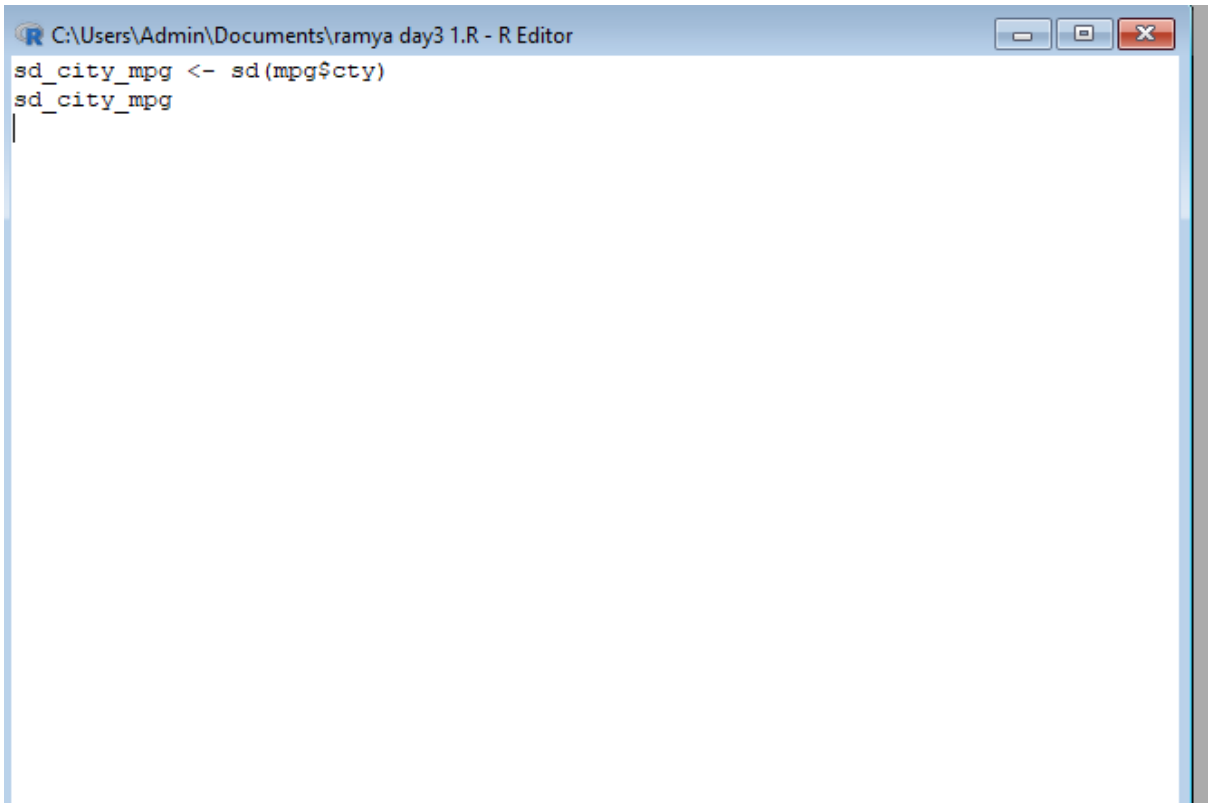
Exercise: 5

Use the same dataset as used in Exercise 4 and perform the following queries

i) Find the standard deviation of city milles per gallon

program:

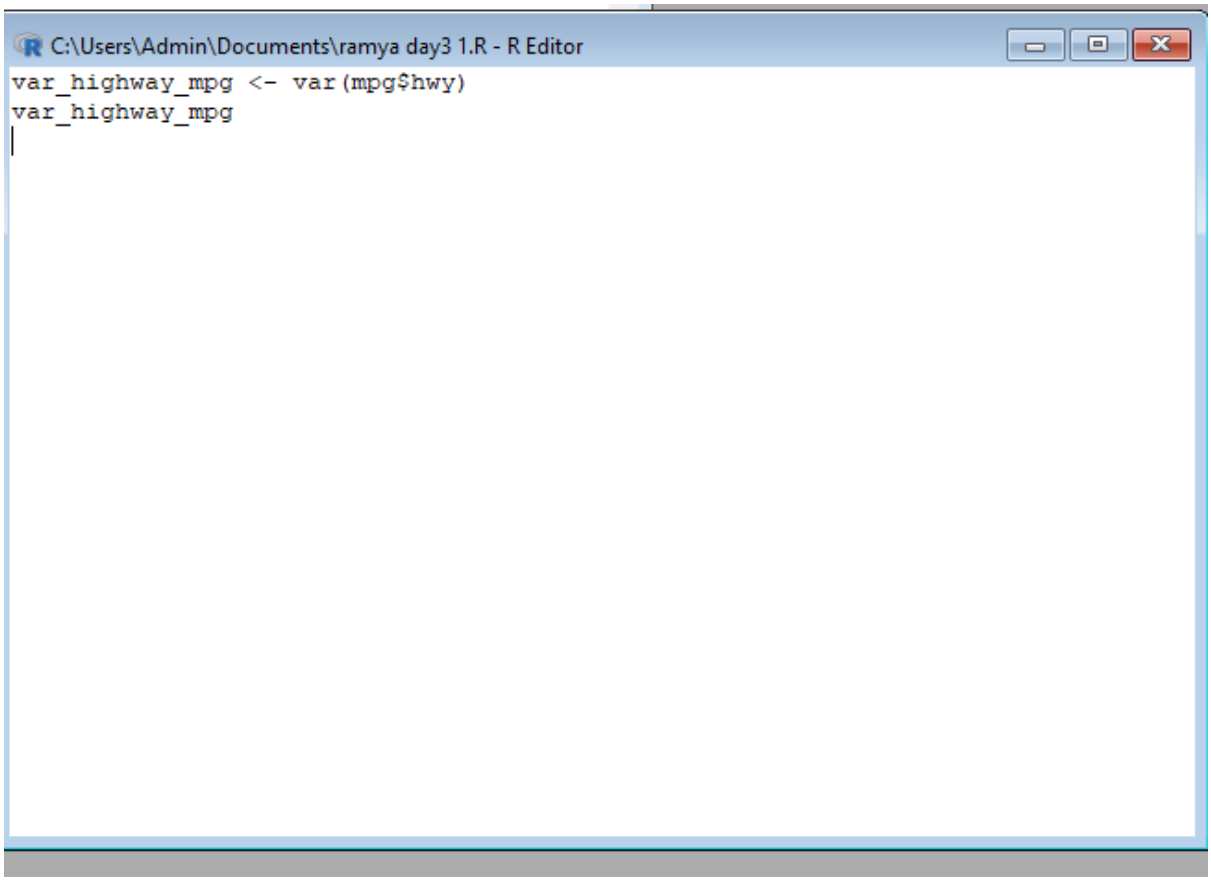
```
sd_city_mpg <- sd(mpg$cty)
sd_city_mpg
```



ii) Find the variance of highway milles per gallon

pprogram:

```
var_highway_mpg <- var(mpg$hwy)
var_highway_mpg
```



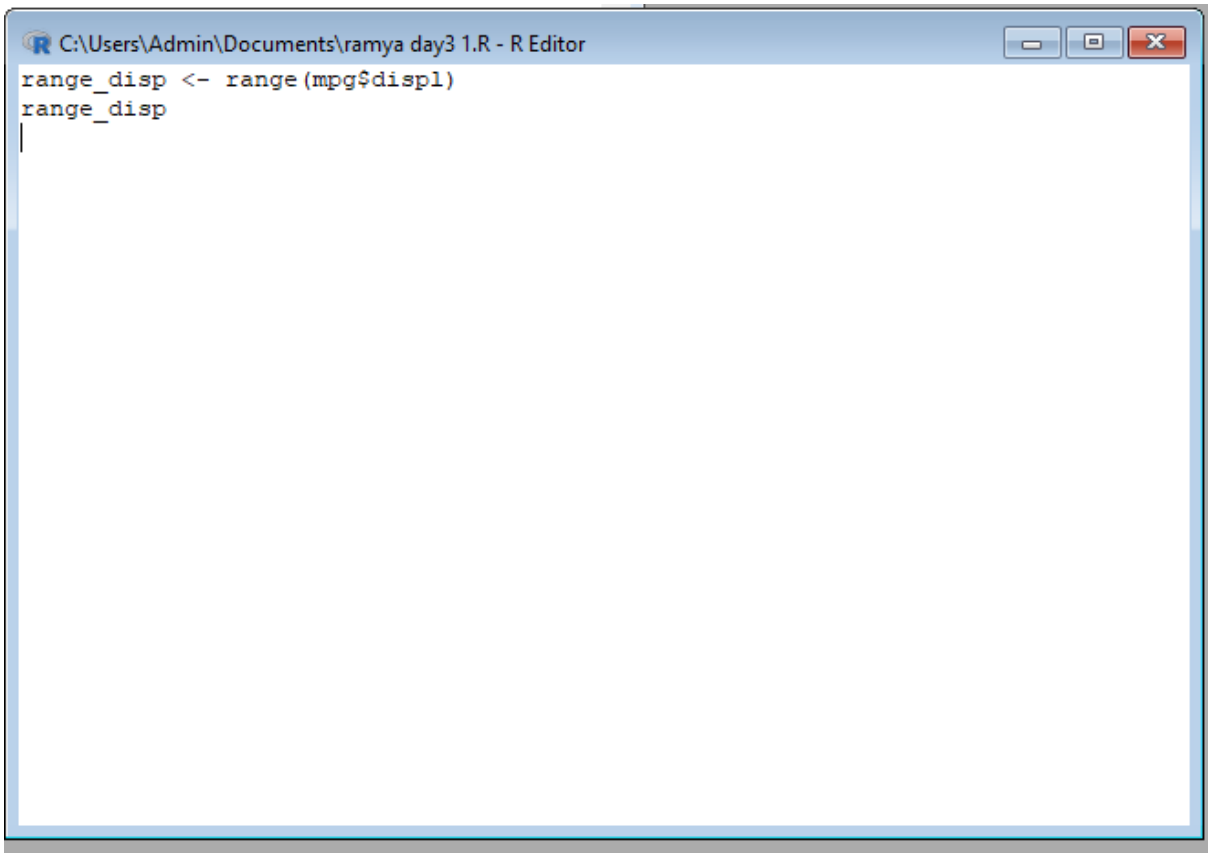
Exercise 6

Use the same dataset and perform the following queries

i) Find the range of the disp in the data set mpg

program:

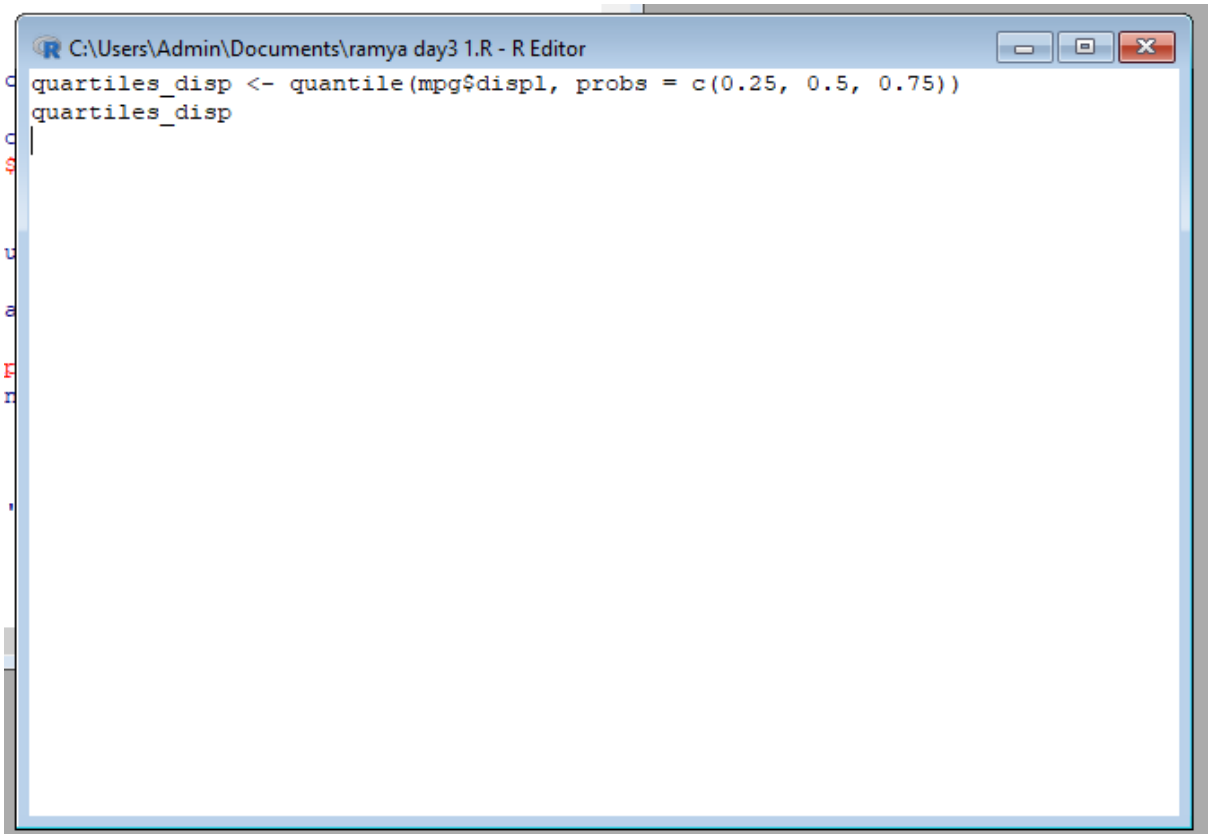
```
range_disp <- range(mpg$displ)
range_disp
```



ii) Find the Quartile of the disp in the data set mpg

program:

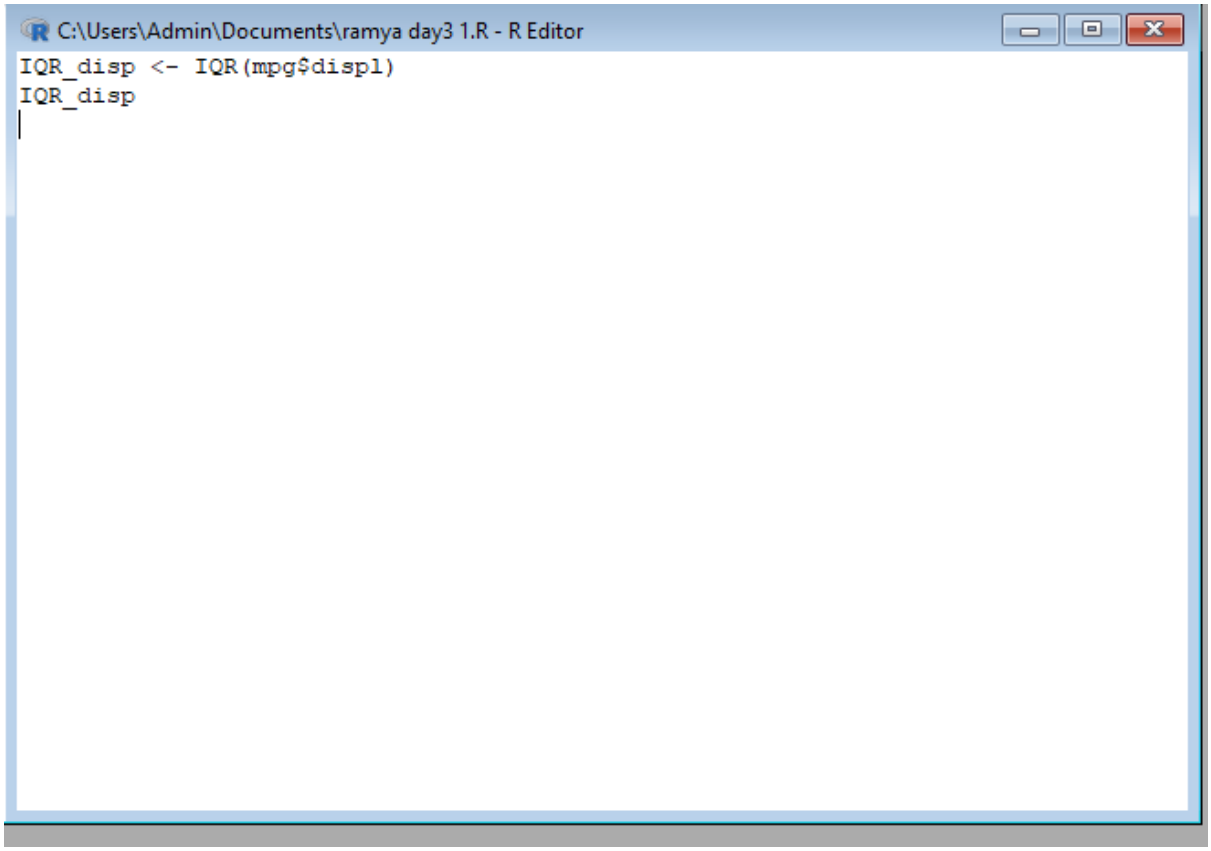
```
quartiles_disp <- quantile(mpg$displ, probs = c(0.25, 0.5, 0.75))
quartiles_disp
```



iii) Find the IQR of the disp column in the data set mpg

program:

```
IQR_disp <- IQR(mpg$displ)
IQR_disp
```



Exercise 7

#Install Library

library(e1071)

a. Find the skewness of city miles per mileage in the data set mpg ?

   Use qplot function and display the graph for the city miles per mileage column

b. Find the kurtosis of city miles per mileage in the data set mpg

program:

```
install.packages("e1071")
library(e1071)
skewness(mpg$cty)
library(ggplot2)
qplot(data = mpg, x = cty)
kurtosis(mpg$cty)
```

C:\Users\Admin\Documents\ramya day3 1.R - R Editor

```
install.packages("e1071")
library(e1071)
skewness(mpg$cty)
library(ggplot2)
qplot(data = mpg, x = cty)
kurtosis(mpg$cty)
```