QUESTION: 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

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

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

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

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")

1.Code:

```
data<-c(12,7,4.2,2,18,2,54,-21,8,-5)

result <- mean(data) result


trimmed_mean<-print(mean(data,trim=0.30))


data1<-c(12,7,3,4.2,2,18,-21,8,-5,NA)


mean(data1,na.rm = TRUE)


median(data)


data2<-c(2,1,2,3,1,2,3,4,1,5,5,3,2,3)

mode<-names(table(data2))[which.max(table(data2))]

mode
```

o/p:

```
C:\Users\ADMIN\AppData\Local\Temp\RtmpQBAMMY\downloaded_
> data<-c(12,7,4.2,2,18,2,54,-21,8,-5)
> result <- mean(data)
> result
[1] 8.12
>
> trimmed_mean<-print(mean(data,trim=0.30))
[1] 5.3
>
> data1<-c(12,7,3,4.2,2,18,-21,8,-5,NA)
>
> mean(data1,na.rm = TRUE)
[1] 3.133333
>
> median(data)
[1] 5.6
>
> data2<-c(2,1,2,3,1,2,3,4,1,5,5,3,2,3)
> mode<-names(table(data2))[which.max(table(data2))]
> mode
[1] "2"
>
```

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

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

ii) Find the variance of highway milles per gallon
Exercise 6 Use the same dataset and perform the following queries

i) Find the range of the disp in the data set mpg
ii) Find the Quartile of the disp in the data set mpg
iii) Find the IQR of the disp column in the data set mpg

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 columnb. Find the kurtosis of city miles per mileage in the data set mpg Use qplot function and display the graph for the city miles per mileage column

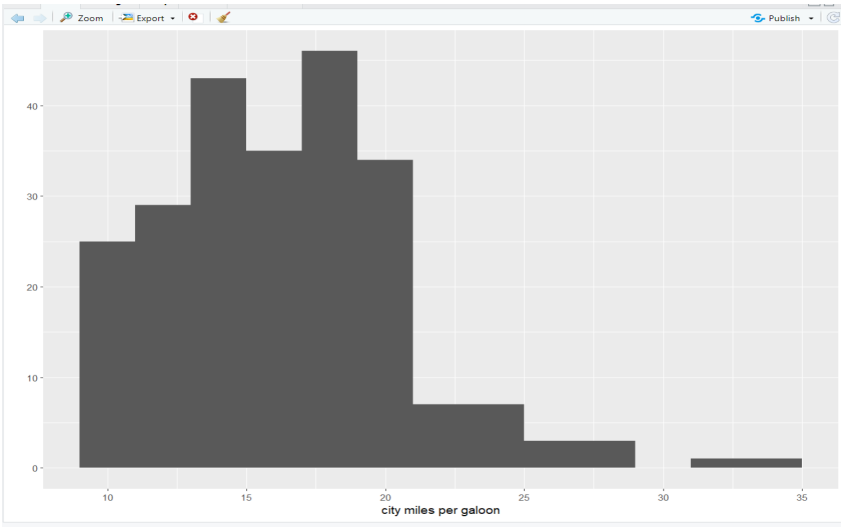
2.Code:

```
library("tidyverse")
mpg<-as.data.frame(ggplot2::mpg) mpg%>%
  filter(cty==max(cty))
mpg%>%
  filter(class %in% c("compact","subcompact")) %>%
group_by(class) %>%   filter(disp==min(disp)) %>%
select(manufacturer,model,class,disp)
install.packages("ggplot2")
library(ggplot2)
data(mpg)
install.packages("e1071")
library(e1071)
sd(mpg$cty)
var(mpg$hwy)
range(mpg$displ)
quantile(mpg$displ)
IQR(mpg$displ)
qplot(mpg$cty,goemn="histogram",binwidth=2,xlab=
"city miles per galoon")
skewness(mpg$cty)
kurtosis(mpg$cty)
```

op:

```
>
> mpg%>%
+   filter(cty==max(cty))
  manufacturer       model displ year cyl      trans drv cty hwy fl      class
1   volkswagen new beetle   1.9 1999   4 manual(m5)   f  35  44  d subcompact
> mpg%>%
+   filter(class %in% c("compact","subcompact")) %>%
+   group_by(class) %>%
+   filter(disp==min(disp)) %>%
+   select(manufacturer,model,class,disp)
```

```
> install.packages("ggplot2")
Error in install.packages : Updating loaded packages
> library(ggplot2)
> data(mpg)
> install.packages("e1071")
Error in install.packages : Updating loaded packages
> library(e1071)
> sd(mpg$cty)
[1] 4.255946
> var(mpg$hwy)
[1] 35.45778
> range(mpg$displ)
[1] 1.6 7.0
> quantile(mpg$displ)
  0%  25%  50%  75% 100%
 1.6  2.4  3.3  4.6  7.0
> IQR(mpg$displ)
[1] 2.2
> qplot(mpg$cty,goemn="histogram",binwidth=2,xlab="city miles per galoon")
```



QUESTION:

Reference Status Gender TestNewOrFollowUp 1 KRXH Accepted Female Test1 New 2 KRPT Accepted Male Test1 New 3 FHRA Rejected Male Test2 New 4 CZKK Accepted Female Test3 New 5 CQTN Rejected Female Test1 New 6 PZXW Accepted Female Test4 Follow-up 7 SZRZ Rejected Male Test4 New 8 RMZE Rejected Female Test2 New 9 STNX Accepted Female Test3 New 10 TMDW Accepted Female Test1 New

i) Load the dataset and Create a data frame and name it as dataframe1

ii) Load the function for crosstab

Code:

```
dataset<-data <- matrix(c("KRXH", "Accepted", "Female", "Test1", "New",
                          "KRPT", "Accepted", "Male", "Test1", "New",
                          "FHRA", "Rejected", "Male", "Test2", "New",
```

"CZKK", "Accepted", "Female", "Test3", "New",

"CQTN", "Rejected", "Female", "Test1", "New",

"PZXW", "Accepted", "Female", "Test4", "Follow-up",

"SZRZ", "Rejected", "Male", "Test4", "New",

"RMZE", "Rejected", "Female", "Test2", "New",

"STNX", "Accepted", "Female", "Test3", "New",

"TMDW", "Accepted", "Female", "Test1", "New"), ncol=5, byrow=TRUE)

dataframe1 <- data.frame(Reference=data[,1], Status=data[,2], Gender=data[,3], TestNewOrFollowUp=data[,5])

dataframe1

xtabs(~TestNewOrFollowUp+Status,data=dataframe1)

output:

```
> dataframe1 <- data.frame(Reference=data[,1], Status=data[,2], Gender=data[,3], TestNewOrFollowUp=data[,5])
> dataframe1
   Reference   Status Gender TestNewOrFollowUp
1       KRXH Accepted Female              New
2       KRPT Accepted   Male              New
3       FHRA Rejected   Male              New
4       CZKK Accepted Female              New
5       CQTN Rejected Female              New
6       PZXW Accepted Female        Follow-up
7       SZRZ Rejected   Male              New
8       RMZE Rejected Female              New
9       STNX Accepted Female              New
10      TMDW Accepted Female              New
```

```
> xtabs(~TestNewOrFollowUp+Status,data=dataframe1)
                Status
TestNewOrFollowUp Accepted Rejected
        Follow-up        1        0
        New              5        4
>
```

Write a program for creating a pie-chart in R using the input vector(21,62,10,53). Provide labels for the chart as ‘London’, ‘New York’, ‘Singapore’, ‘Mumbai’. Add a title to the chart as ‘city pie-chart’ and add a legend at the top right corner of the chart.

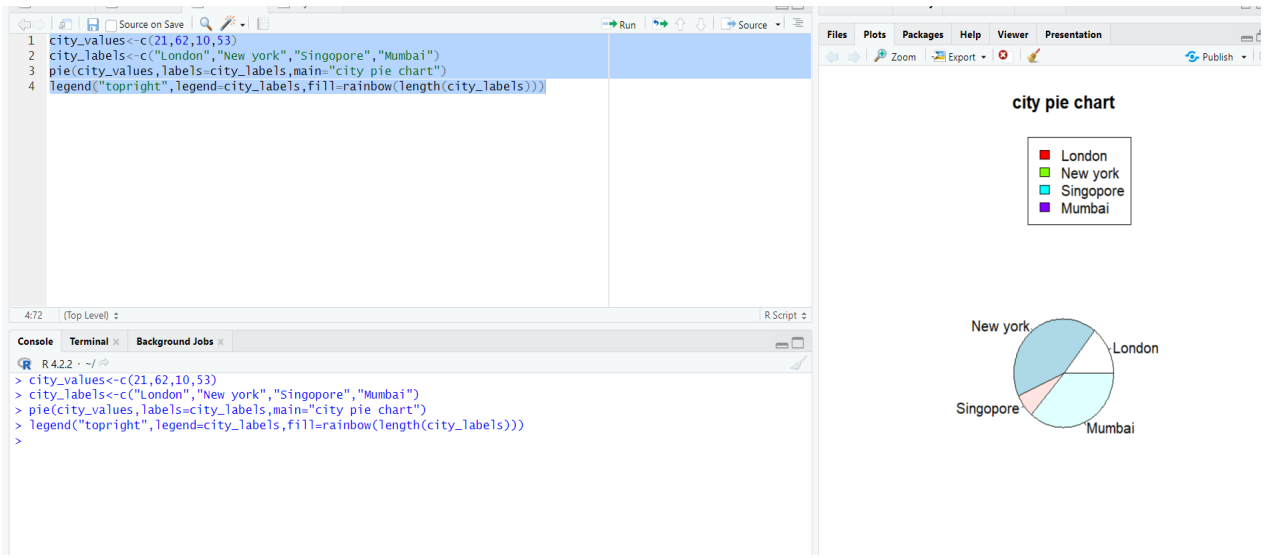
Code & output:

city_values<-c(21,62,10,53)

city_labels<-c("London","New york","Singopore","Mumbai")

pie(city_values,labels=city_labels,main="city pie chart")

legend("topright",legend=city_labels,fill=rainbow(length(city_labels)))

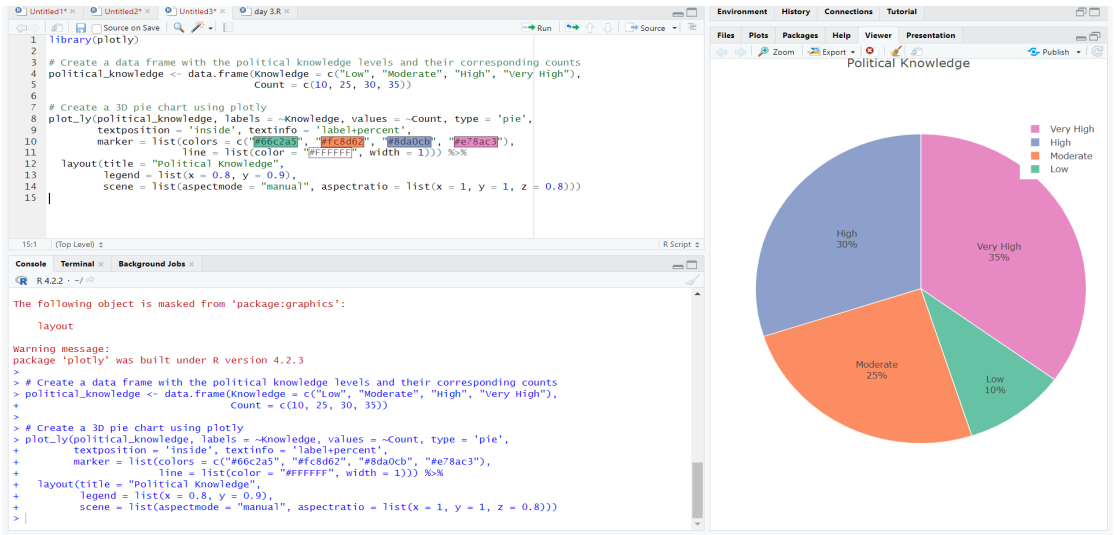


Create a 3D Pie Chart for the dataset “political Knowledge” with suitable labels,colours and a legend at the top right corner of the chart.

```
library(plotly)

# Create a data frame with the political knowledge levels and their corresponding counts
political_knowledge <- data.frame(Knowledge = c("Low", "Moderate", "High", "Very High"),
                                  Count = c(10, 25, 30, 35))

# Create a 3D pie chart using plotly
plot_ly(political_knowledge, labels = ~Knowledge, values = ~Count, type = 'pie',
        textposition = 'inside', textinfo = 'label+percent',
        marker = list(colors = c("#66c2a5", "#fc8d62", "#8da0cb", "#e78ac3"),
                      line = list(color = "#FFFFFF", width = 1))) %>%
  layout(title = "Political Knowledge",
         legend = list(x = 0.8, y = 0.9),
         scene = list(aspectmode = "manual", aspectratio = list(x = 1, y = 1, z = 0.8)))
```

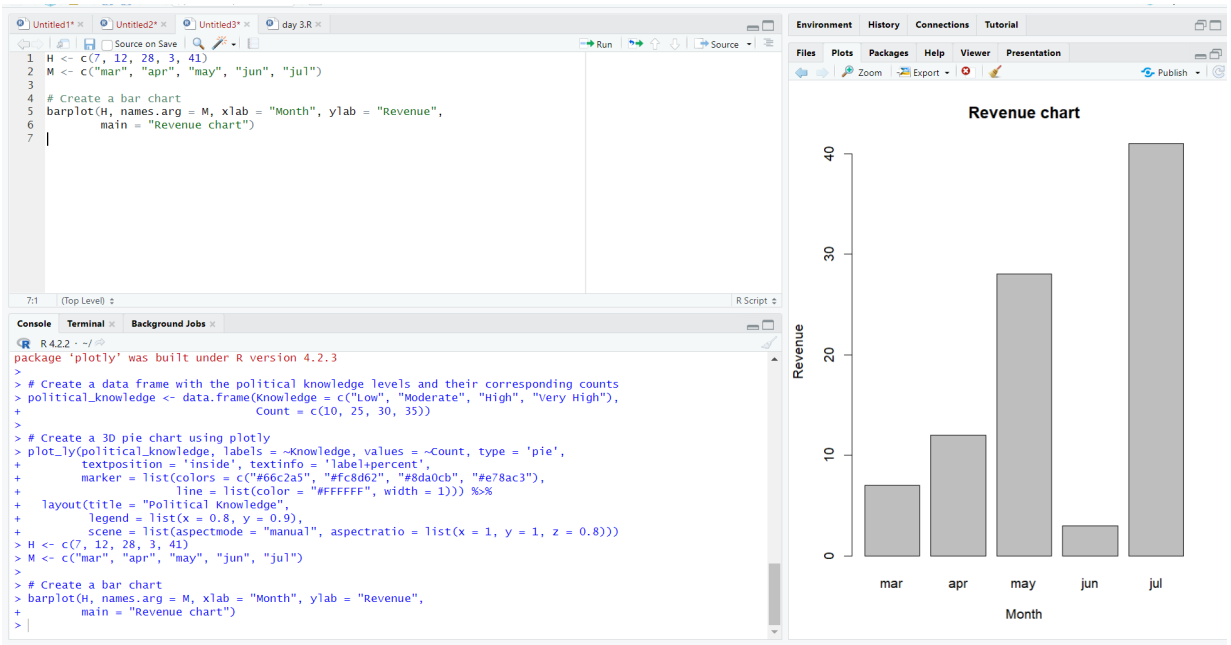


Write a program for creating a bar chart using the vectors H=c(7,12,28,3,41) and M=c(“mar”, “apr”, “may”, “jun”, “jul”). Add a title to the chart as “Revenue chart”.

Code:

```
H <- c(7, 12, 28, 3, 41)
M <- c("mar", "apr", "may", "jun", "jul")

# Create a bar chart
barplot(H, names.arg = M, xlab = "Month", ylab =
"Revenue",
        main = "Revenue chart")
```



Make a histogram for the “AirPassengers“dataset, start at 100 on the x-axis, and from values 200 to 700, make the bins 200 wide

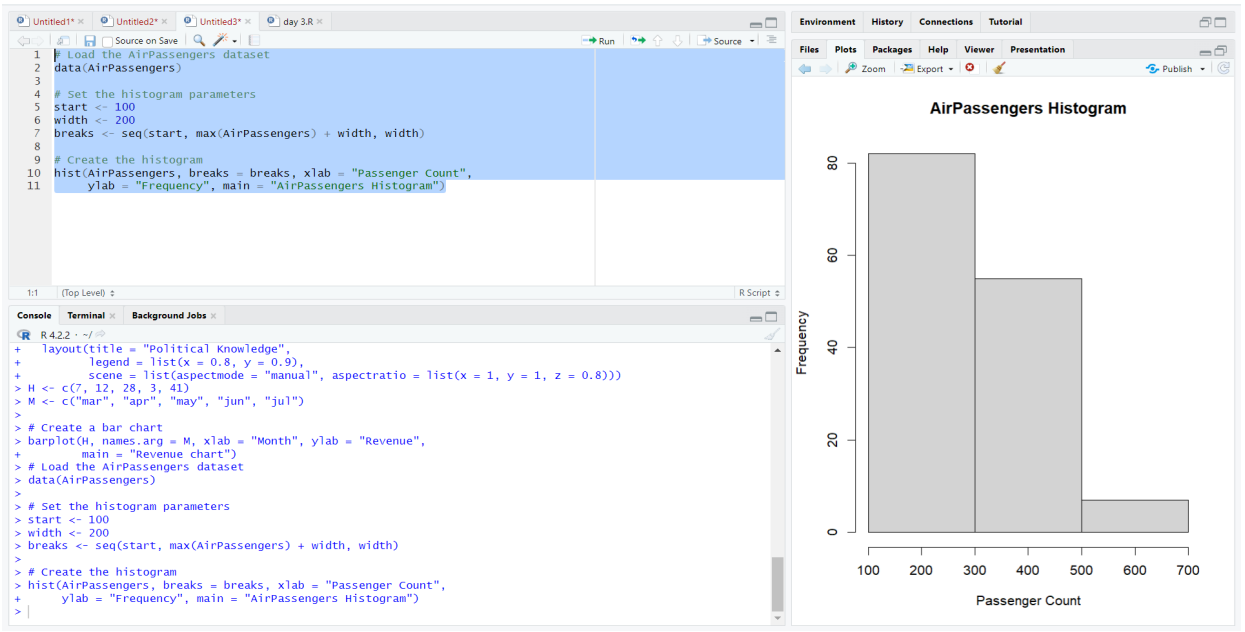
Code:

```
# Load the AirPassengers dataset
```

```
data(AirPassengers)

# Set the histogram parameters
start <- 100
width <- 200
breaks <- seq(start, max(AirPassengers) + width,
width)

# Create the histogram
hist(AirPassengers, breaks = breaks, xlab =
"Passenger Count",
     ylab = "Frequency", main = "AirPassengers
Histogram")
```



Create a Boxplot graph for the relation between "mpg"(miles per galloon) and "cyl"(number of Cylinders) for the dataset "mtcars" available in R Environment.

```
# Load the mtcars dataset
data(mtcars)

# Create a boxplot of mpg vs cyl
boxplot(mpg ~ cyl, data = mtcars, main = "Miles per
Gallon vs Number of Cylinders",
        xlab = "Number of Cylinders", ylab = "Miles per
Gallon", col = "lightblue")
```

```r
# Load the mtcars dataset
data(mtcars)

# Create a boxplot of mpg vs cyl
boxplot(mpg ~ cyl, data = mtcars, main = "Miles per Gallon vs Number of Cylinders",
        xlab = "Number of Cylinders", ylab = "Miles per Gallon", col = "lightblue")
```

```
> # Create a bar chart
> barplot(H, names.arg = M, xlab = "Month", ylab = "Revenue",
+         main = "Revenue chart")
> # Load the AirPassengers dataset
> data(AirPassengers)
>
> # Set the histogram parameters
> start <- 100
> width <- 200
> breaks <- seq(start, max(AirPassengers) + width, width)
>
> # Create the histogram
> hist(AirPassengers, breaks = breaks, xlab = "Passenger Count",
+      ylab = "Frequency", main = "AirPassengers Histogram")
> # Load the mtcars dataset
> data(mtcars)
>
> # Create a boxplot of mpg vs cyl
> boxplot(mpg ~ cyl, data = mtcars, main = "Miles per Gallon vs Number of Cylinders",
+         xlab = "Number of Cylinders", ylab = "Miles per Gallon", col = "lightblue")
>
```