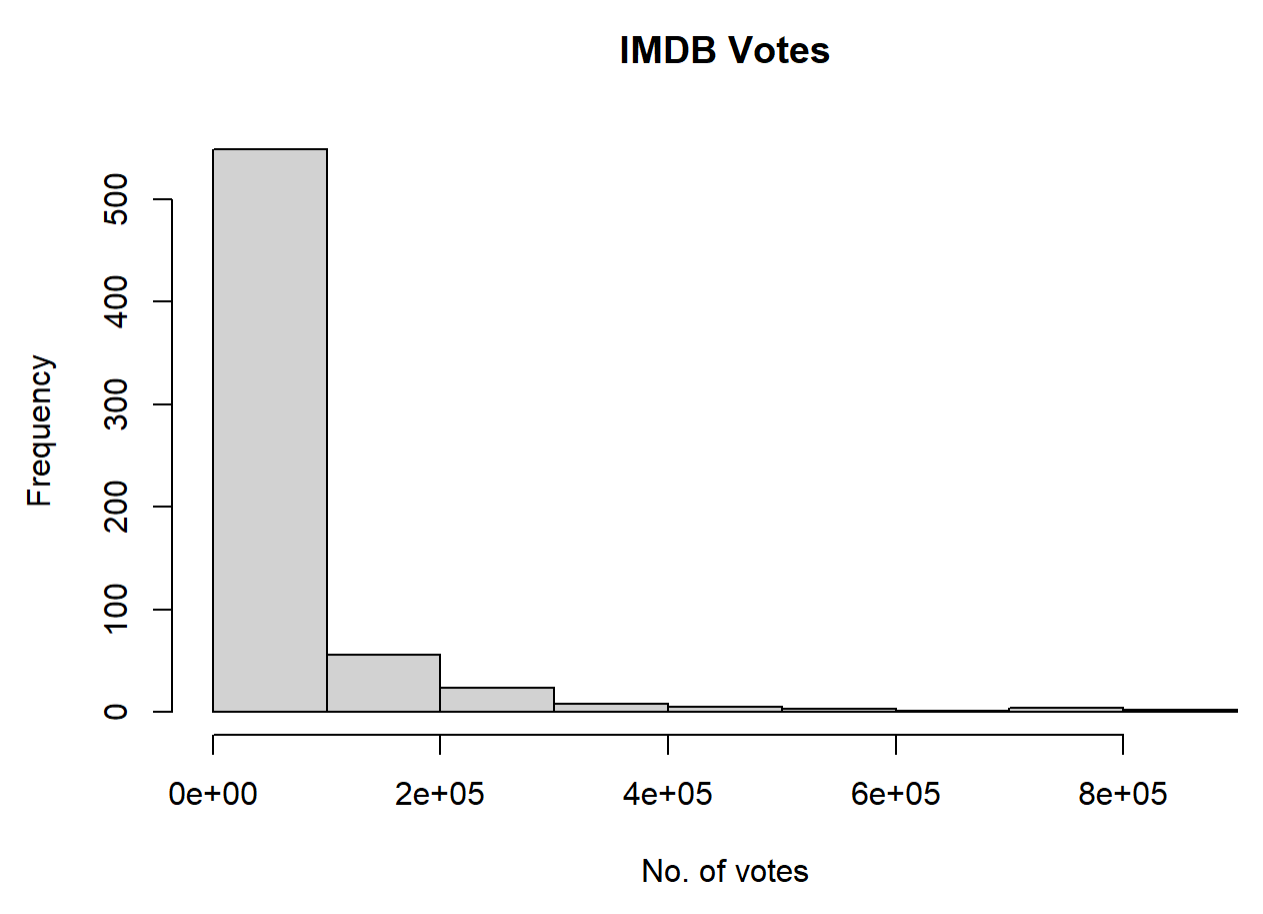
612303050 Deshmukh Mehmood Rehan's Assignment 2

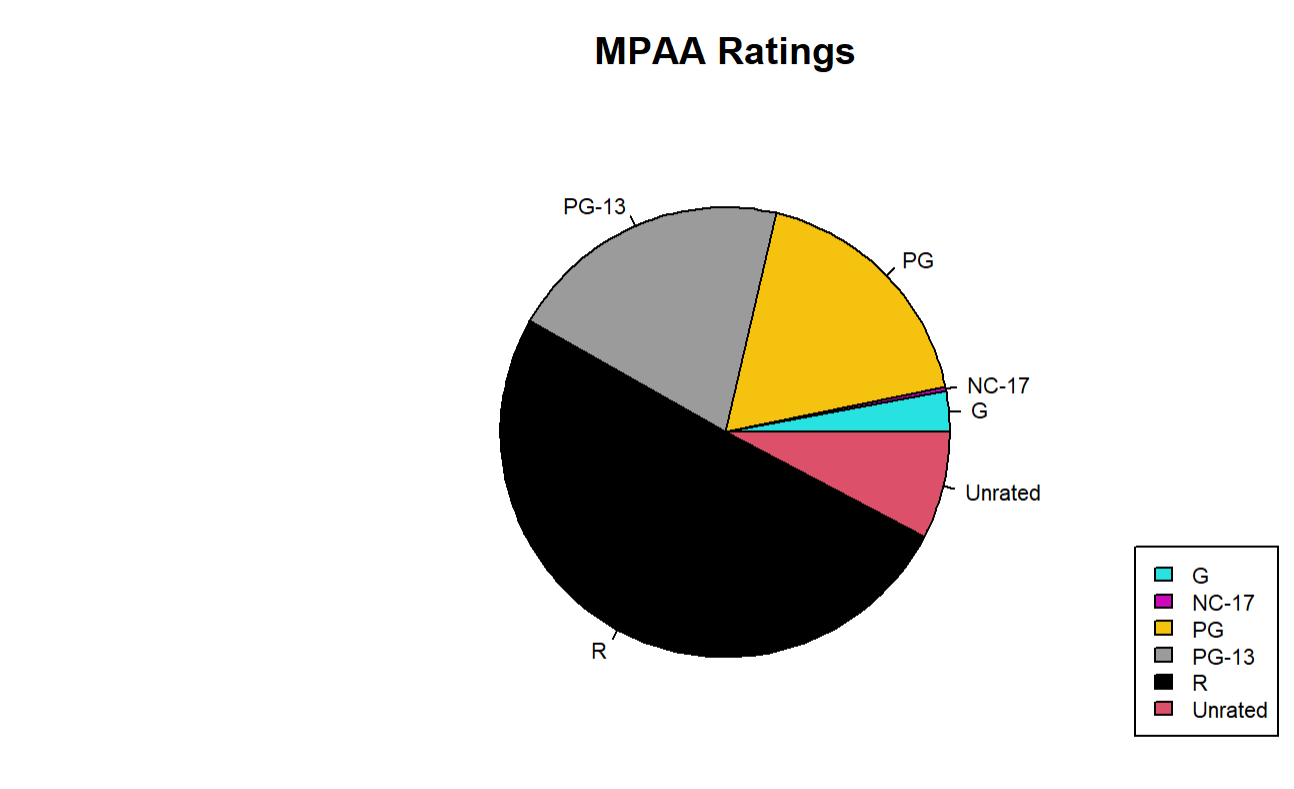
Q1. Use the moviesData. Create a histogram of the object named imdb_num_votes in this file

```
moviesData = read.csv("moviesData.csv")
hist(moviesData$imdb_num_votes, main = "IMDB Votes", xlab = "No. of votes")
```



Q2. Create a pie chart of the object mpaa_rating and save the plot.

```
mpaa_ratings = factor(moviesData$mpaa_rating)
mpaa_ratings_table = table(mpaa_ratings)
pie(mpaa_ratings_table, col = 5:11, main = "MPAA Ratings", cex = 0.7)
legend("bottomright", levels(mpaa_ratings), fill = 5:11, cex = 0.7)
```

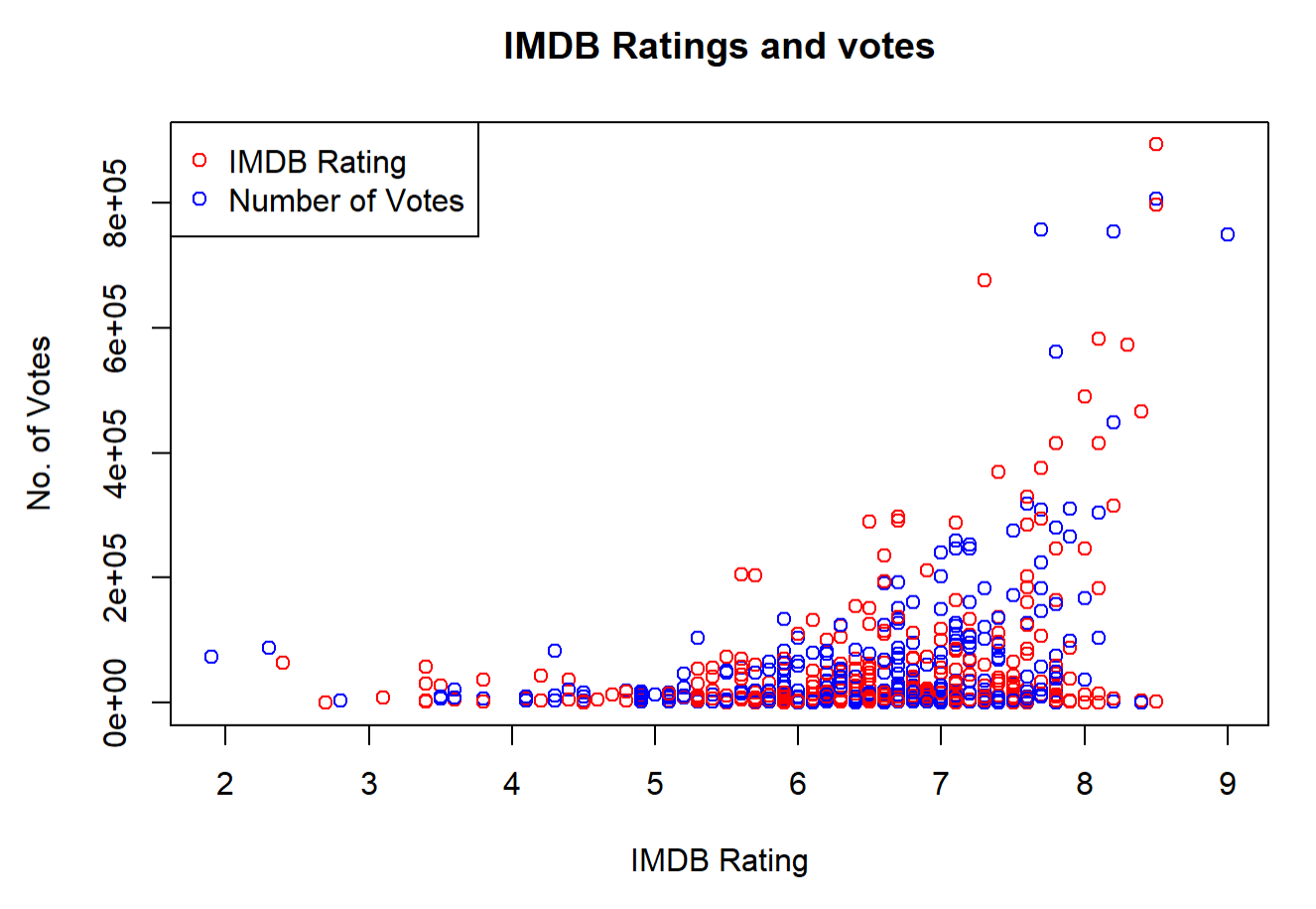


Q3. Read the file moviesData.csv Create a bar chart of critics_score for the first 10 movies.

```
moviesData = read.csv("moviesData.csv")
score_of_critics = (moviesData$critics_score[1:10])
barplot(score_of_critics, main = "Critics' Score", names.arg = 1:10, xlab = "Movie No.", ylab = "Score")
```

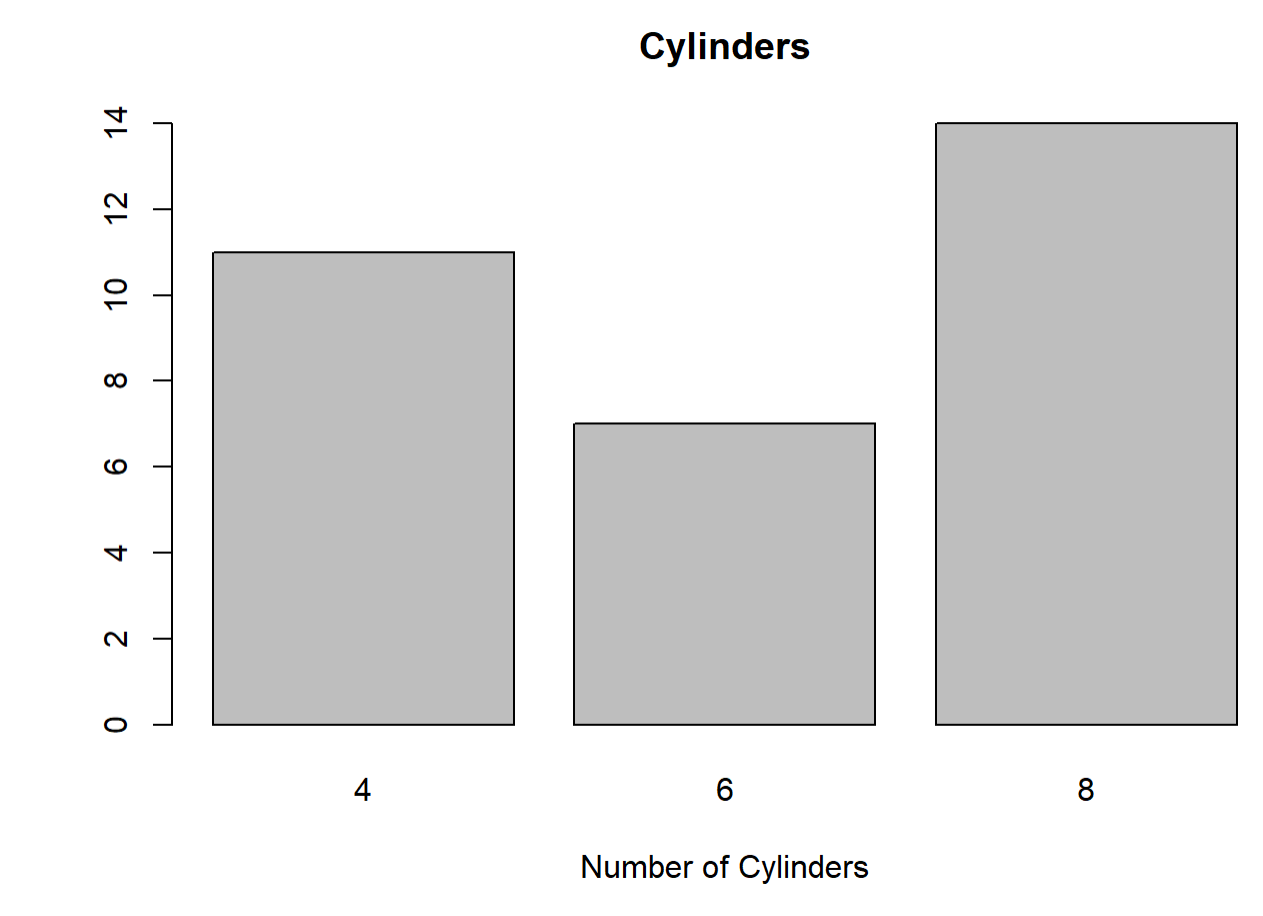
Q4. Create a scatter plot of imdb_rating and imdb_num_votes to see their relation. Also save the plot.

```
imdb_rating = moviesData$imdb_rating
imdb_num_votes = moviesData$imdb_num_votes
plot(imdb_rating, imdb_num_votes, col = c("red", "blue"), main = "IMDB Ratings and votes", xlab="IMDB Rating", yl
ab="No. of Votes")
legend("topleft", legend = c("IMDB Rating", "Number of Votes"), col = c("red", "blue"), pch = 1)
```



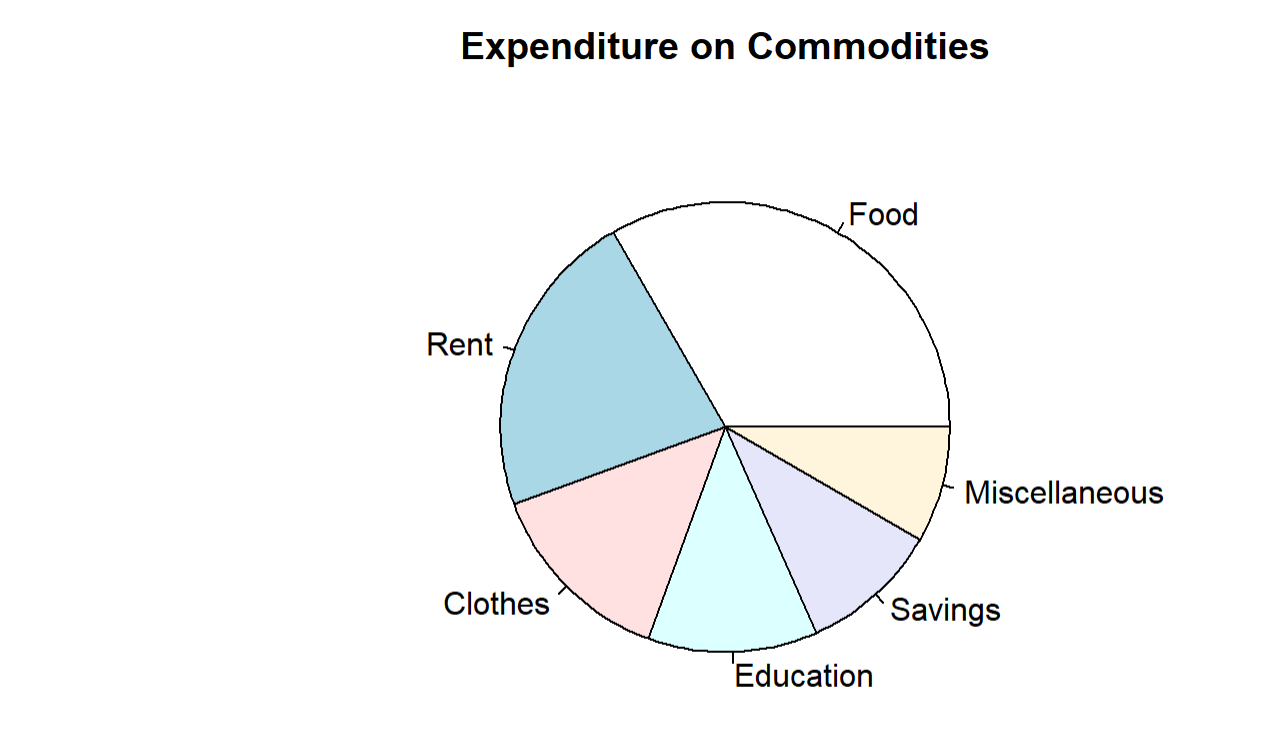
Q5. Using the mtcars dataset create a barplot using cyl variable

```
data(mtcars)
cylinder_table <- table(mtcars$cyl)
barplot(cylinder_table, main = "Cylinders",
        xlab = "Number of Cylinders")
```



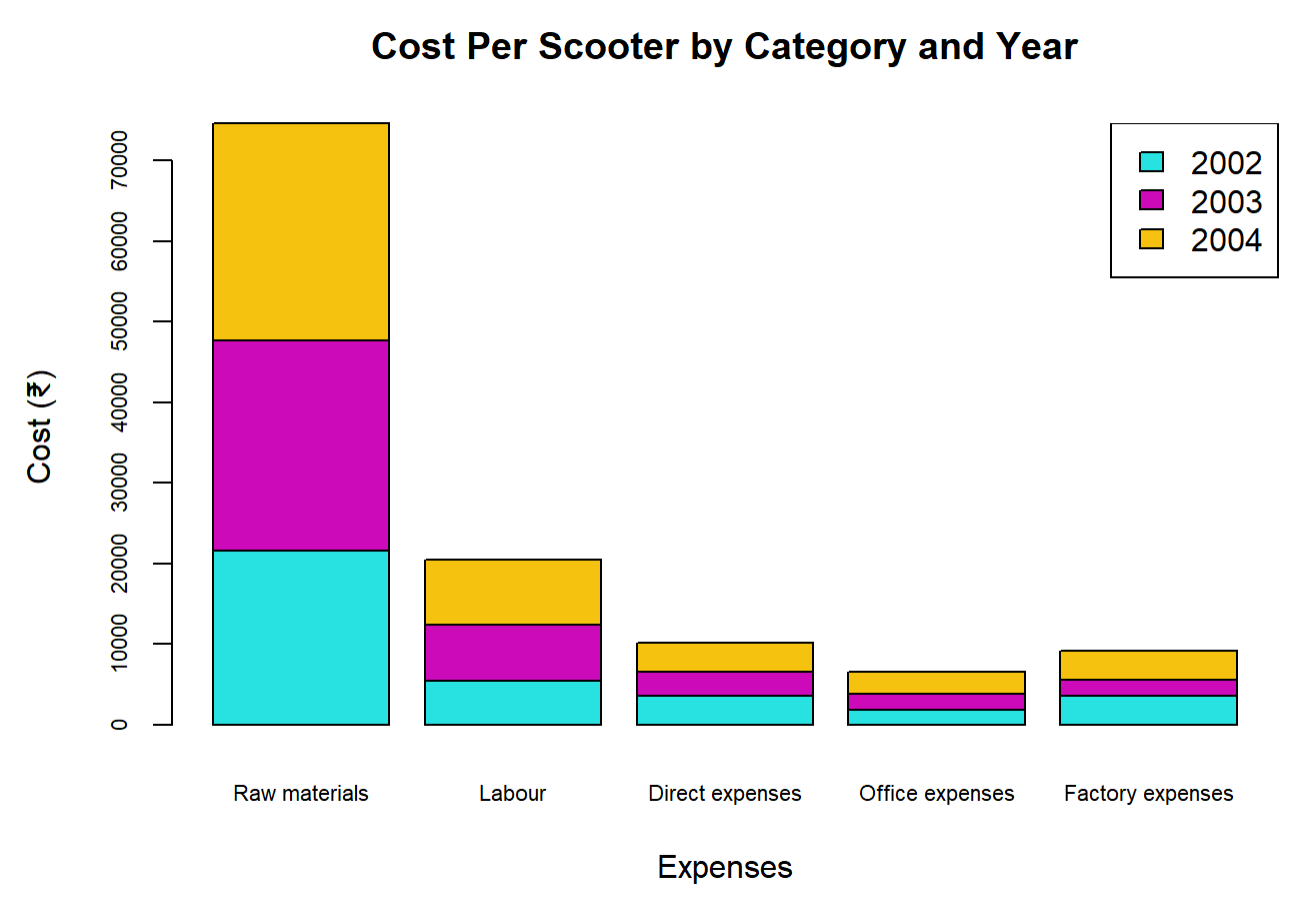
Q6. Represent the following data by pie diagram

```
commodity <- c("Food", "Rent", "Clothes", "Education", "Savings", "Miscellaneous")
expenditure <- c(300, 200, 125, 110, 90, 75)

pie(expenditure, labels = commodity, main = "Expenditure on Commodities")
```



Q7. Represent the following data by subdivided bar diagram

```
years <- c(2002, 2003, 2004)
raw_materials <- c(21600, 26000, 27000)
labour <- c(5400, 7000, 8100)
direct_expenses <- c(3600, 3000, 3500)
office_expenses <- c(1800, 2000, 2700)
factory_expenses <- c(3600, 2000, 3600)

scooter_row_names <- c("Raw materials", "Labour", "Direct expenses","Office expenses", "Factory expenses")

cost_df <- data.frame(raw_materials, labour, direct_expenses, office_expenses, factory_expenses)
names(cost_df) <- scooter_row_names
barplot(as.matrix(cost_df), col = 5:7, main = "Cost Per Scooter by Category and Year", ylab = "Cost (₹)",xlab = "E
xpenses" ,names.arg = scooter_row_names, cex.axis = 0.7, cex.names = 0.7)
legend("topright", legend = years, fill = 5:7)
```



Q8. The following frequency distribution table gives age distribution gives the age of drivers who were at fault in accidents during a 1-week period in a city Draw a histogram

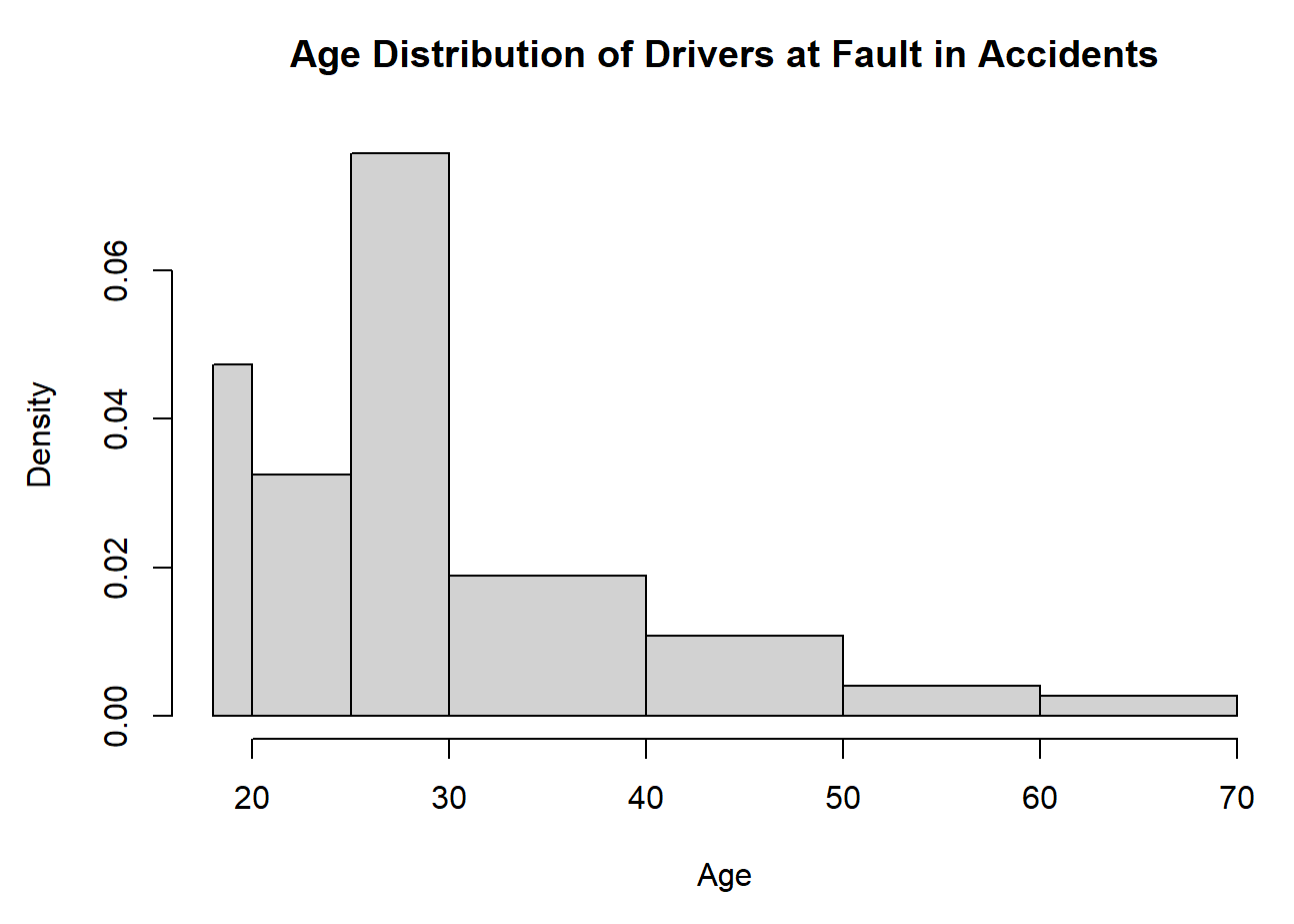
```
age_groups <- c("18-20", "20-25", "25-30", "30-40", "40-50", "50-60", "60-70")

lowerbound=c(18,20,25,30,40,50,60)
upperbound=c(20,25,30,40,50, 60, 70)
age_freq=c(7, 12, 28, 14, 8, 3, 2)
breaks=c(18,upperbound)
x=(lowerbound+upperbound)/2
x
```

```
## [1] 19.0 22.5 27.5 35.0 45.0 55.0 65.0
```

```
y=rep(x,age_freq)
y
```

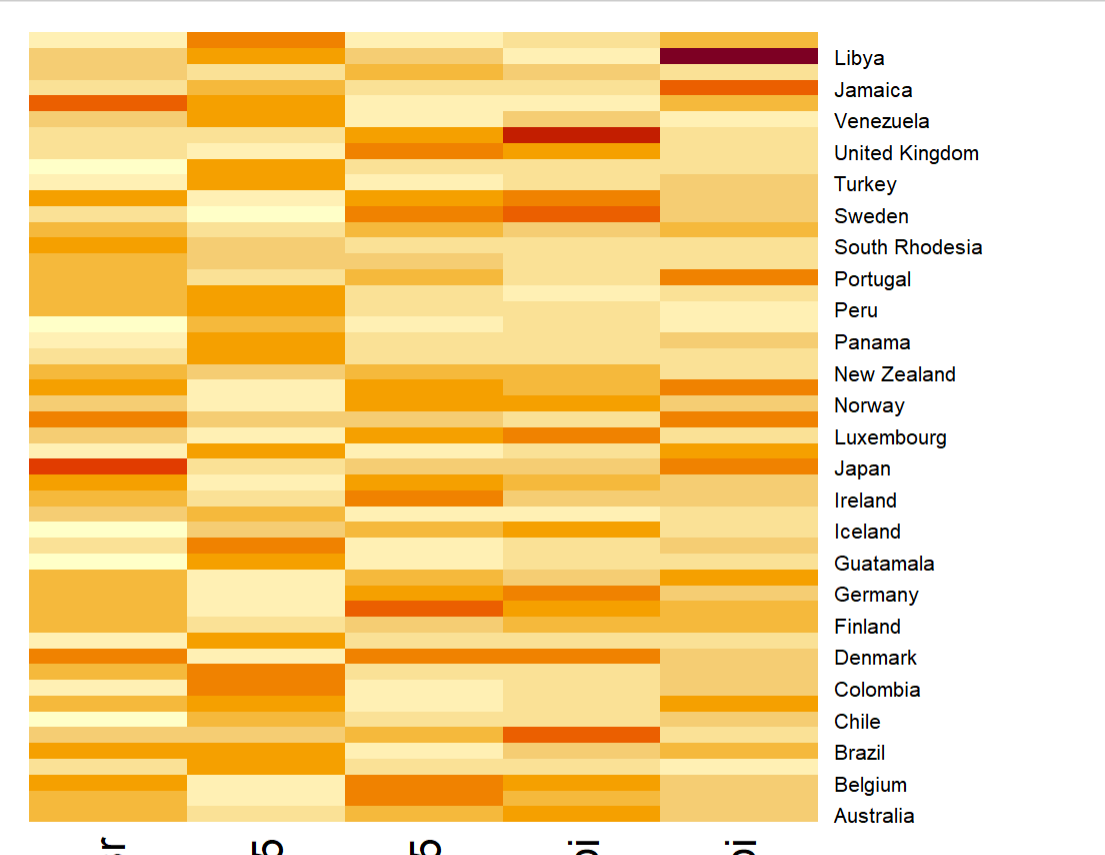
```
##  [1] 19.0 19.0 19.0 19.0 19.0 19.0 19.0 22.5 22.5 22.5 22.5 22.5 22.5 22.5 22.5
## [16] 22.5 22.5 22.5 22.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5
## [31] 27.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5 27.5
## [46] 27.5 27.5 35.0 35.0 35.0 35.0 35.0 35.0 35.0 35.0 35.0 35.0 35.0 35.0 35.0
## [61] 35.0 45.0 45.0 45.0 45.0 45.0 45.0 45.0 45.0 55.0 55.0 55.0 65.0 65.0
```

```
hist(y,breaks,main = "Age Distribution of Drivers at Fault in Accidents",  xlab = "Age")
```



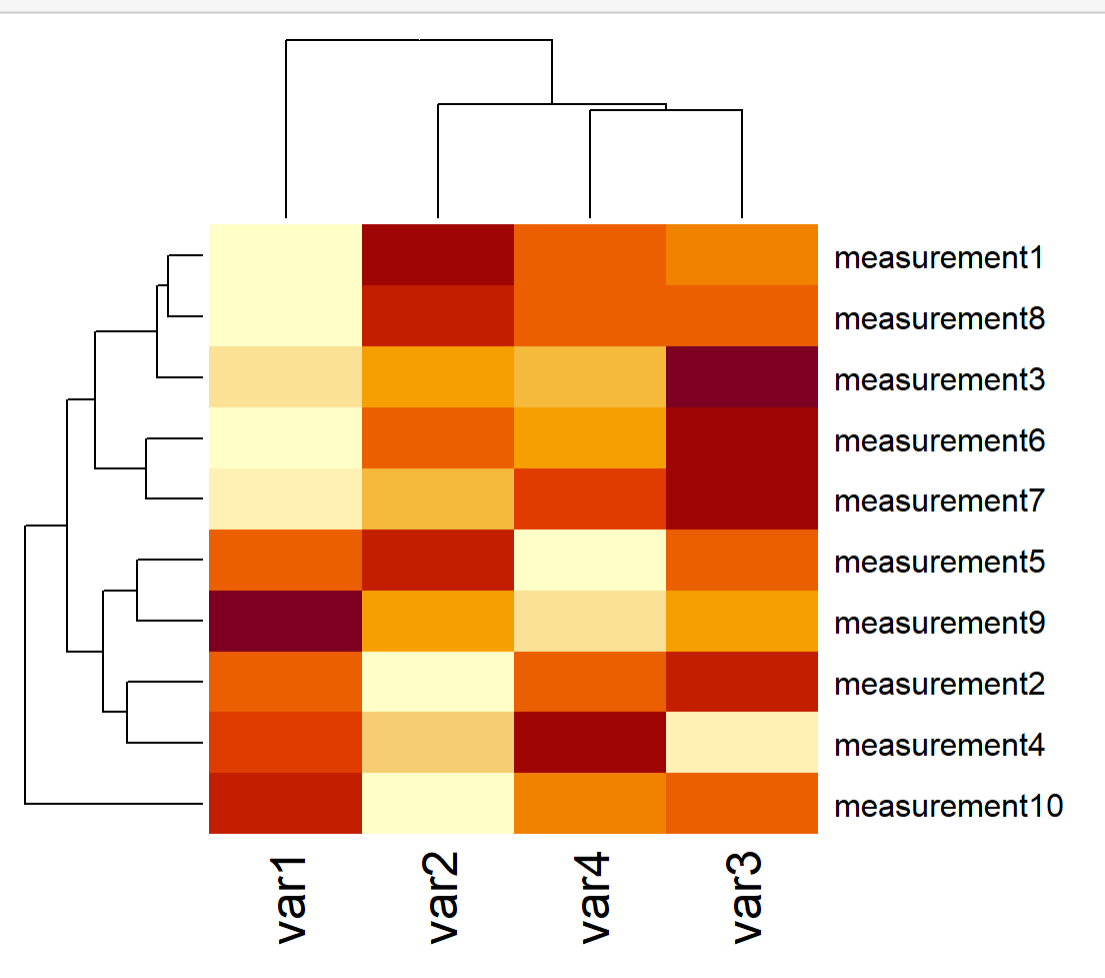
Q9.Create a heatmap using LifeCycleSavings dataset

```
data("LifeCycleSavings")
data=as.matrix(LifeCycleSavings)
heatmap(data, scale = "column", Colv = NA, Rowv = NA )
```

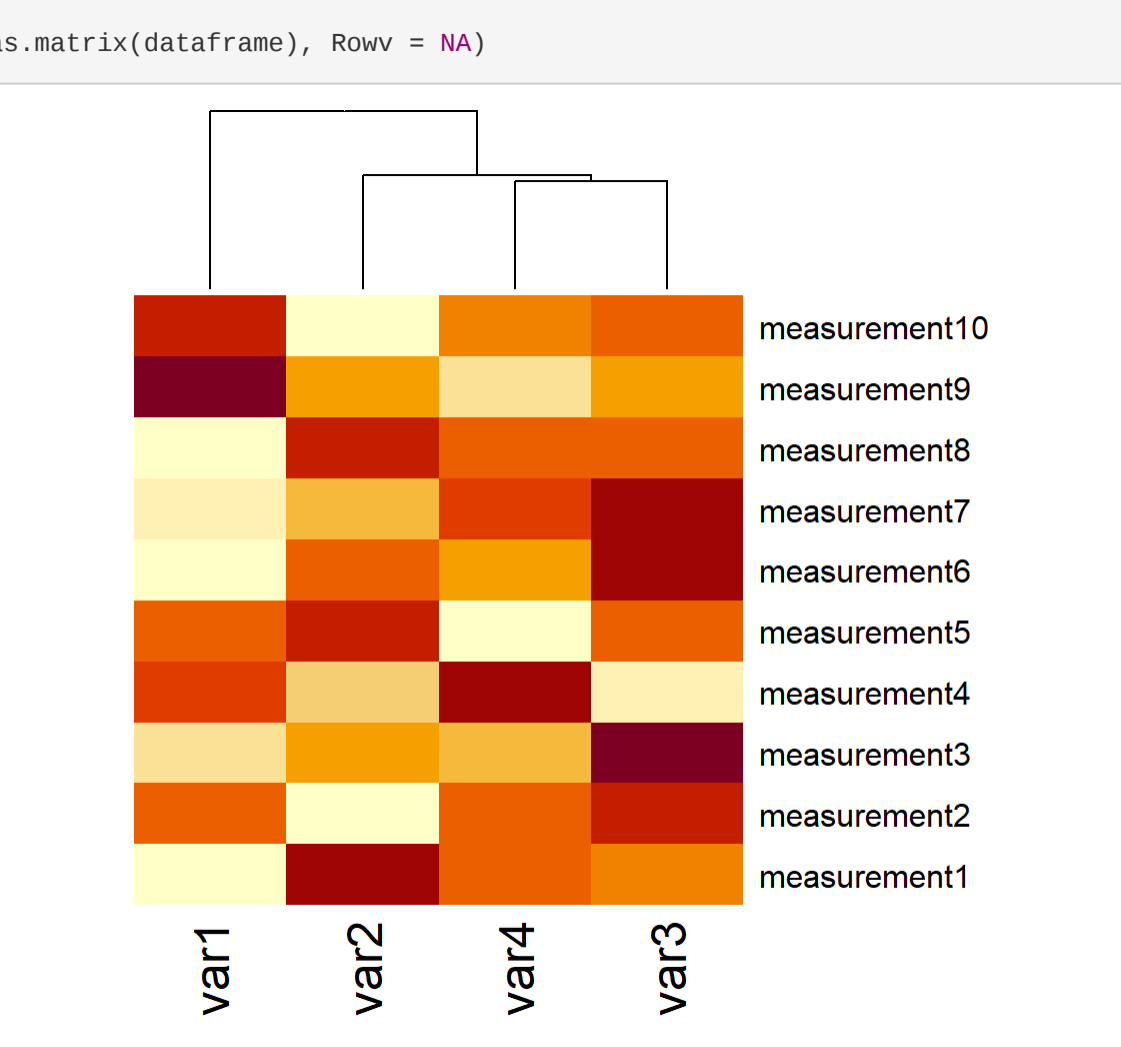

Libya
Jamaica
Venezuela
United Kingdom
Turkey
Sweden
South Rhodesia
Portugal
Peru
Panama
New Zealand
Norway
Luxembourg
Japan
Ireland
Iceland
Guatamala
Germany
Finland
Denmark
Colombia
Chile
Brazil
Belgium
Australia
sr
pop15
pop75
dpi
ddpi

Q10. Create a heatmap using following datasets.

```
var1 = c(0.094, 0.1138, 0.1893, -0.0102, 0.1587, -0.4558, -0.6241, -0.227, 0.7365, 0.9761)
var2 = c(0.668, -0.3847, 0.3303, -0.4259, 0.2948, 0.2244, -0.3119, 0.499, -0.0872, 0.4355)
var3 = c(0.4153, 0.2671, 0.5821, -0.5967, 0.153, 0.6619, 0.3642, 0.3067, -0.069, 0.8663)
var4 = c(0.4613, 0.1529, 0.2632, 0.18, -0.2208, 0.0457, 0.2003, 0.3289, -0.4252, 0.8107)
dataframe = data.frame(var1, var2, var3, var4)
rownames(dataframe) = c("measurement1", "measurement2", "measurement3", "measurement4", "measurement5", "measurem
ent6", "measurement7", "measurement8", "measurement9", "measurement10")
heatmap(as.matrix(dataframe))
```




```
#a) Write code to remove row dendogram

heatmap(as.matrix(dataframe), Rowv = NA)
```




```
#b) Write code to remove column dendograms

heatmap(as.matrix(dataframe), Colv = NA)
```