Problem Statement

1. Write a program to create barplots for all the categorical columns in mtcars.

Answer :

```
> library(readr)
> library(ggplot2)
Warning message:
package 'ggplot2' was built under R version 3.5.2
> library(ggplot2)
> mtcars <- read_csv('C:/Users/Vikram/Desktop/Acad/mtcars.csv')
Parsed with column specification:
cols(
  model = col_character(),
  mpg = col_double(),
  cyl = col_double(),
  disp = col_double(),
  hp = col_double(),
  drat = col_double(),
  wt = col_double(),
  qsec = col_double(),
  vs = col_double(),
  am = col_double(),
  gear = col_double(),
  carb = col_double()
)
> View(mtcars)
> str(mtcars)
Classes 'tbl_df', 'tbl' and 'data.frame':    32 obs. of  12 variables:
 $ model: chr  "Mazda RX4" "Mazda RX4 Wag" "Datsun 710" "Hornet 4 Drive" ...
 $ mpg  : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
 $ cyl  : num  6 6 4 6 8 6 8 4 4 6 ...
 $ disp : num  160 160 108 258 360 ...
 $ hp   : num  110 110 93 110 175 105 245 62 95 123 ...
 $ drat : num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
 $ wt   : num  2.62 2.88 2.32 3.21 3.44 ...
 $ qsec : num  16.5 17 18.6 19.4 17 ...
 $ vs   : num  0 0 1 1 0 1 0 1 1 1 ...
 $ am   : num  1 1 1 0 0 0 0 0 0 0 ...
 $ gear : num  4 4 4 3 3 3 3 4 4 4 ...
 $ carb : num  4 4 1 1 2 1 4 2 2 4 ...
 - attr(*, "spec")=
  .. cols(
  ..   model = col_character(),
  ..   mpg = col_double(),
  ..   cyl = col_double(),
  ..   disp = col_double(),
  ..   hp = col_double(),
  ..   drat = col_double(),
  ..   wt = col_double(),
  ..   qsec = col_double(),
  ..   vs = col_double(),
  ..   am = col_double(),
  ..   gear = col_double(),
  ..   carb = col_double()
  .. )
> library(dplyr)
> library("dplyr", lib.loc="~/R/win-library/3.5")
```

```
> mtcars1 <- mutate(mtcars,cyl = as.factor(cyl),disp = as.factor(disp),vs = as.factor(
vs),am = as.factor(am),gear = as.factor(gear),carb = as.factor(carb))
Warning message:
package 'bindrcpp' was built under R version 3.5.2
> str(mtcars1)
Classes 'tbl_df', 'tbl' and 'data.frame':    32 obs. of  12 variables:
 $ model: chr  "Mazda RX4" "Mazda RX4 Wag" "Datsun 710" "Hornet 4 Drive" ...
 $ mpg  : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
 $ cyl  : Factor w/ 3 levels "4","6","8": 2 2 1 2 3 2 3 1 1 2 ...
 $ disp : Factor w/ 27 levels "71.1","75.7",..: 13 13 6 16 23 15 23 12 10 14 ...
 $ hp   : num  110 110 93 110 175 105 245 62 95 123 ...
 $ drat : num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
 $ wt   : num  2.62 2.88 2.32 3.21 3.44 ...
 $ qsec : num  16.5 17 18.6 19.4 17 ...
 $ vs   : Factor w/ 2 levels "0","1": 1 1 2 2 1 2 1 2 2 2 ...
 $ am   : Factor w/ 2 levels "0","1": 2 2 2 1 1 1 1 1 1 1 ...
 $ gear : Factor w/ 3 levels "3","4","5": 2 2 2 1 1 1 1 2 2 2 ...
 $ carb : Factor w/ 6 levels "1","2","3","4",..: 4 4 1 1 2 1 4 2 2 4 ...
> is.fact <- sapply(mtcars1, is.factor)
> mtcars2 <- mtcars1[,is.fact]
> str(mtcars2)
Classes 'tbl_df', 'tbl' and 'data.frame':    32 obs. of  6 variables:
 $ cyl : Factor w/ 3 levels "4","6","8": 2 2 1 2 3 2 3 1 1 2 ...
 $ disp: Factor w/ 27 levels "71.1","75.7",..: 13 13 6 16 23 15 23 12 10 14 ...
 $ vs  : Factor w/ 2 levels "0","1": 1 1 2 2 1 2 1 2 2 2 ...
 $ am  : Factor w/ 2 levels "0","1": 2 2 2 1 1 1 1 1 1 1 ...
 $ gear: Factor w/ 3 levels "3","4","5": 2 2 2 1 1 1 1 2 2 2 ...
 $ carb: Factor w/ 6 levels "1","2","3","4",..: 4 4 1 1 2 1 4 2 2 4 ...
> par(mfrow= c(2,3))
> lapply(lapply(mtcars2[,1:5], table), barplot)
$`cyl`
     [,1]
[1,]  0.7
[2,]  1.9
[3,]  3.1

$disp
      [,1]
 [1,]  0.7
 [2,]  1.9
 [3,]  3.1
 [4,]  4.3
 [5,]  5.5
 [6,]  6.7
 [7,]  7.9
 [8,]  9.1
 [9,] 10.3
[10,] 11.5
[11,] 12.7
[12,] 13.9
[13,] 15.1
[14,] 16.3
[15,] 17.5
[16,] 18.7
[17,] 19.9
[18,] 21.1
[19,] 22.3
[20,] 23.5
[21,] 24.7
[22,] 25.9
[23,] 27.1
[24,] 28.3
[25,] 29.5
[26,] 30.7
```

```
[27,] 31.9

$vs
     [,1]
[1,]  0.7
[2,]  1.9

$am
     [,1]
[1,]  0.7
[2,]  1.9

$gear
     [,1]
[1,]  0.7
[2,]  1.9
[3,]  3.1
```

2. Create a scatterplot matrix by gear types in mtcars dataset.

Answer :

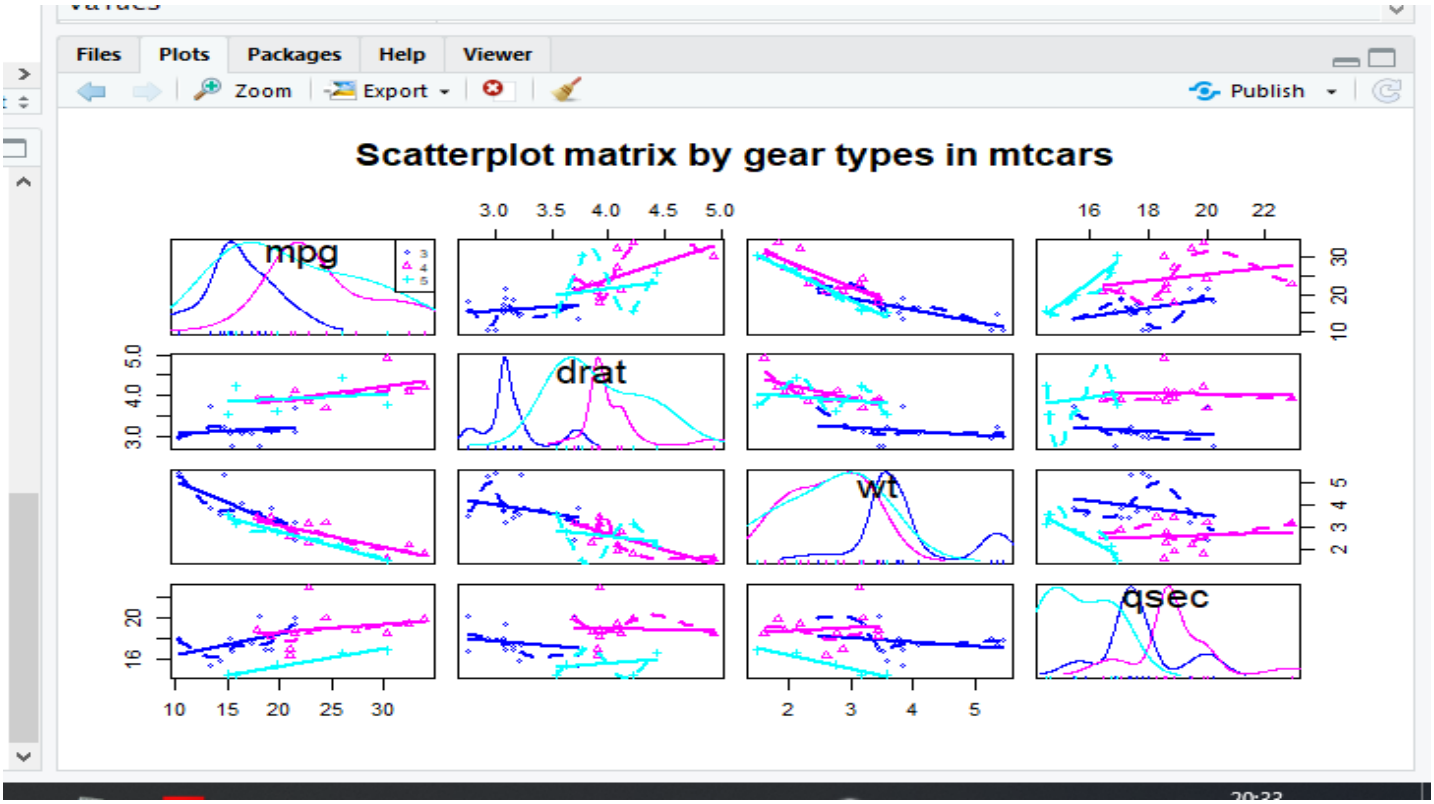
```
> library(car)
Loading required package: carData
```

```
Attaching package: ‘car’

The following object is masked from ‘package:dplyr’:

    recode

Warning messages:
1: package ‘car’ was built under R version 3.5.2
2: package ‘carData’ was built under R version 3.5.2
> str(mtcars)
Classes ‘tbl_df’, ‘tbl’ and 'data.frame':    32 obs. of  12 variables:
 $ model: chr  "Mazda RX4" "Mazda RX4 Wag" "Datsun 710" "Hornet 4 Drive" ...
 $ mpg  : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
 $ cyl  : num  6 6 4 6 8 6 8 4 4 6 ...
 $ disp : num  160 160 108 258 360 ...
 $ hp   : num  110 110 93 110 175 105 245 62 95 123 ...
 $ drat : num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
 $ wt   : num  2.62 2.88 2.32 3.21 3.44 ...
 $ qsec : num  16.5 17 18.6 19.4 17 ...
 $ vs   : num  0 0 1 1 0 1 0 1 1 1 ...
 $ am   : num  1 1 1 0 0 0 0 0 0 0 ...
 $ gear : num  4 4 4 3 3 3 3 4 4 4 ...
 $ carb : num  4 4 1 1 2 1 4 2 2 4 ...
 - attr(*, "spec")=
  .. cols(
  ..   model = col_character(),
  ..   mpg = col_double(),
  ..   cyl = col_double(),
  ..   disp = col_double(),
  ..   hp = col_double(),
  ..   drat = col_double(),
  ..   wt = col_double(),
  ..   qsec = col_double(),
  ..   vs = col_double(),
  ..   am = col_double(),
  ..   gear = col_double(),
  ..   carb = col_double()
  .. )
> scatterplotMatrix(~mpg+drat+wt+qsec|gear, data=mtcars,main="Scatterplot matrix by ge
ar types in mtcars")
```



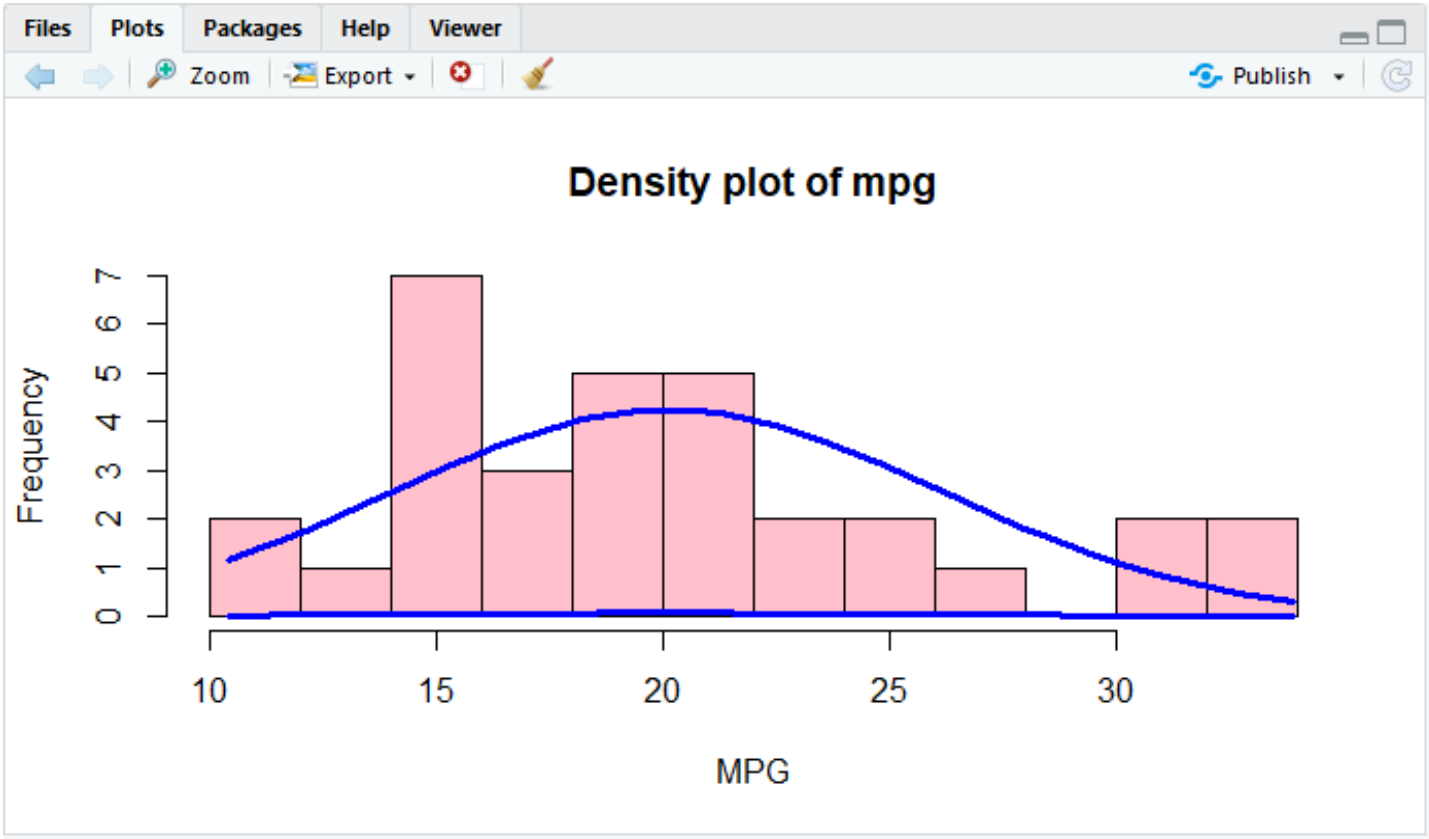
3. Write a program to create a plot density by class variable.

Answer :

```
> par(mfrow = c(1,1))
> x <- mtcars$mpg
> h <- hist(x, breaks = 10, col = "pink",xlab = "MPG",main = "Density plot of mpg"
+ h <- hist(x, breaks = 10, col = "pink",xlab = "MPG",main = "Density plot of mpg")
Error: unexpected symbol in:
"h <- hist(x, breaks = 10, col = "pink",xlab = "MPG",main = "Density plot of mpg"
h"
>

>

> h <- hist(x, breaks = 10, col = "pink",xlab = "MPG",main = "Density plot of mpg")
> xfit <- seq(min(x), max(x), length = 40)
> yfit <- dnorm(xfit, mean = mean(x), sd= sd(x))
> lines(xfit, yfit, col="Blue", lwd = 3)
> xfit <- seq(min(x), max(x), length = 40)
> yfit <- dnorm(xfit, mean = mean(x), sd= sd(x))
> yfit <- yfit*diff(h$mids[1:2]*length(x))
> lines(xfit, yfit, col="Blue", lwd = 3)
```


Files
Plots
Packages
Help
Viewer
Zoom
Export
Publish
Density plot of mpg
Frequency
0
1
2
3
4
5
6
7
10
15
20
25
30
MPG