Assignment - 2

Session 7 – Basic Statistics

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

Ans:

> str(mtcars)

'data.frame': 32 obs. of 11 variables:

$ 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 ...

> # change the categorical variables to factor

> library(dplyr)

> 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))

> str(mtcars1)

'data.frame': 32 obs. of 11 variables:

$ 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)

'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

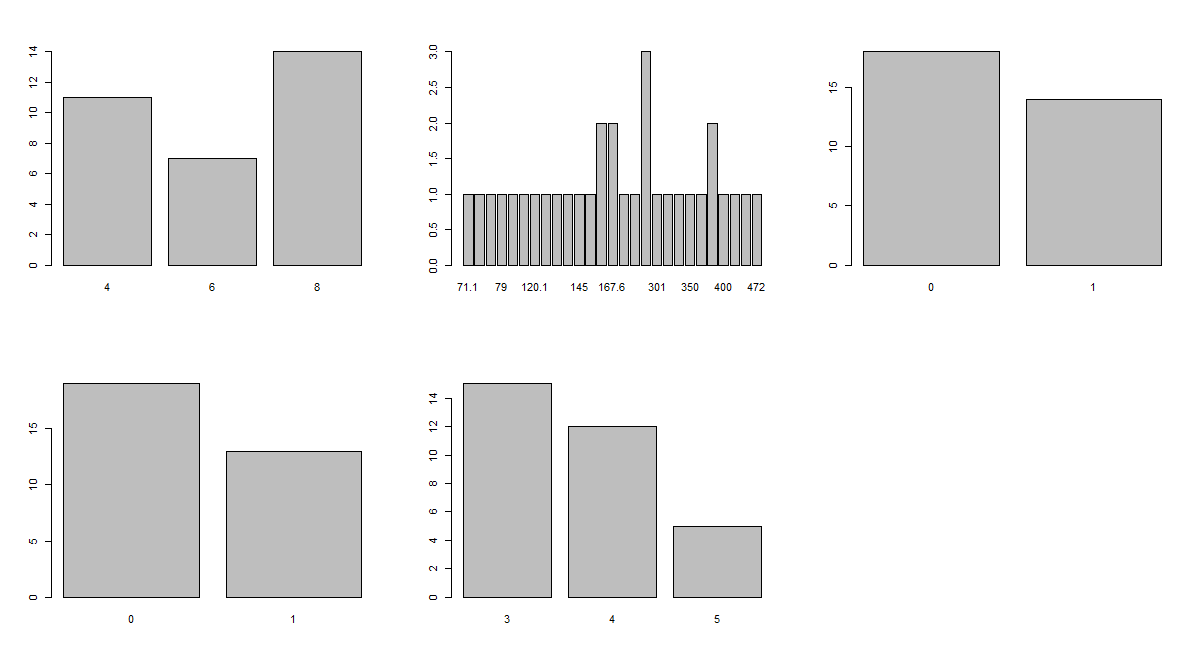
$gear

[,1]

[1,] 0.7

[2,] 1.9

[3,] 3.1



1. Create a **scatterplot** matrix by gear types in **mtcars** dataset.

Ans:

> plot(mtcars$gear , mtcars$mpg, xlab = 'Number of gears', ylab = 'mpg', main = 'Number of gears Vs mpg')



1. Write a program to create a **plot density** by class variable.

Ans:

> d <- density(mtcars$mpg)

> plot(d)

