MAT 2009 Lab

A Lab Report

FALL SEM 2021-22

MAT2009

(Computational Statistics)

Submitted by

Gelle Hruthesh Reddy(20BCB7031)

VIT-AP UNIVERSITY

AMARAVATI

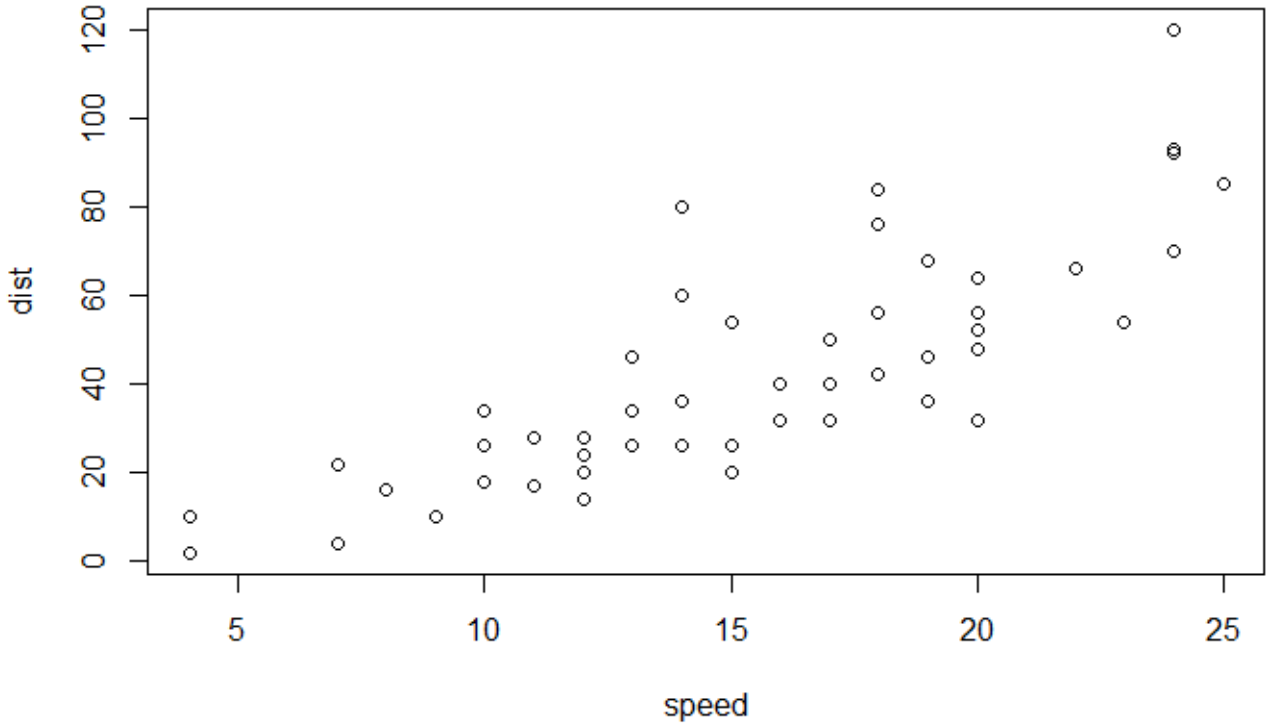
ANDHRA PRADESH, INDIA

2021

1) Learning about plotting in graph:

```{r}
plot(cars)
```

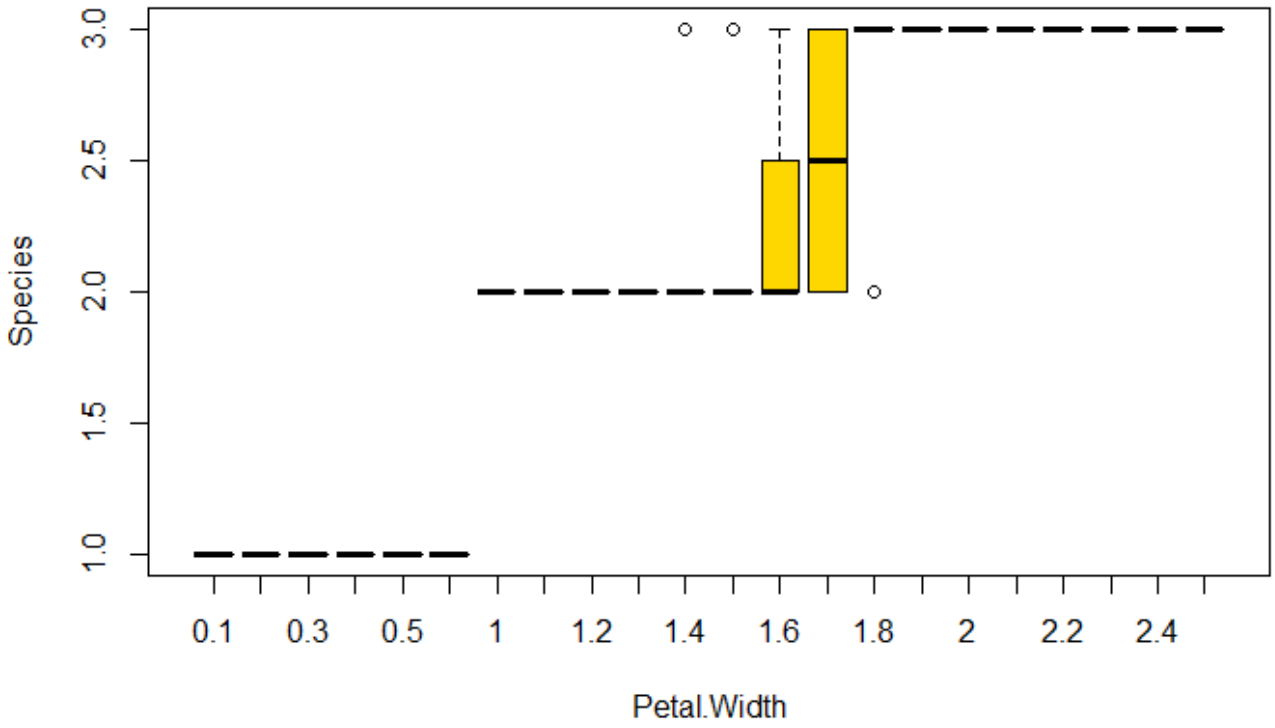
Output:



2) Using Rstudio to learn about Box plot:

```{r}
boxplot(Species~Petal.Width, data=iris, col="gold")
```

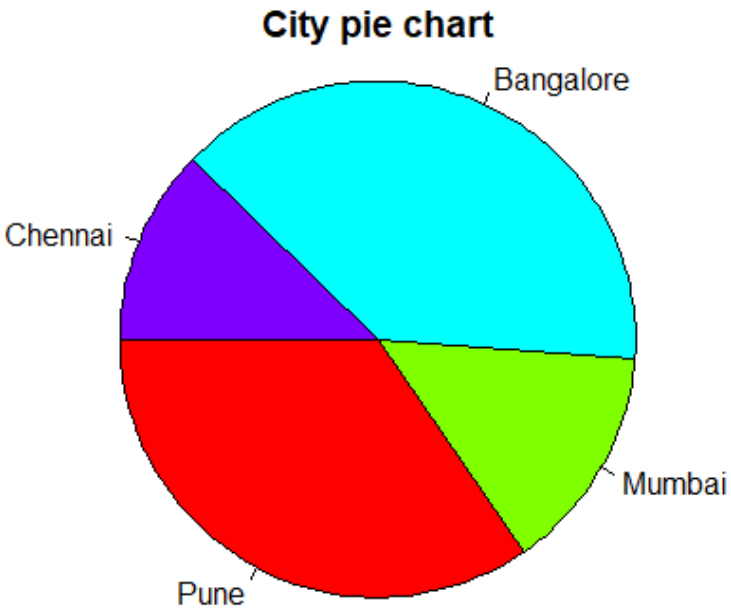
Output:



3) Using Pie chart to show information:

```{r}
x<- c(34.6, 14.2, 38.9 ,12.3)
names(x)<-c("Pune", "Mumbai", "Bangalore", "Chennai")
pie(x, labels=names(x),col=rainbow(4),
main="City pie chart", radius=-1,col.main="black")
```

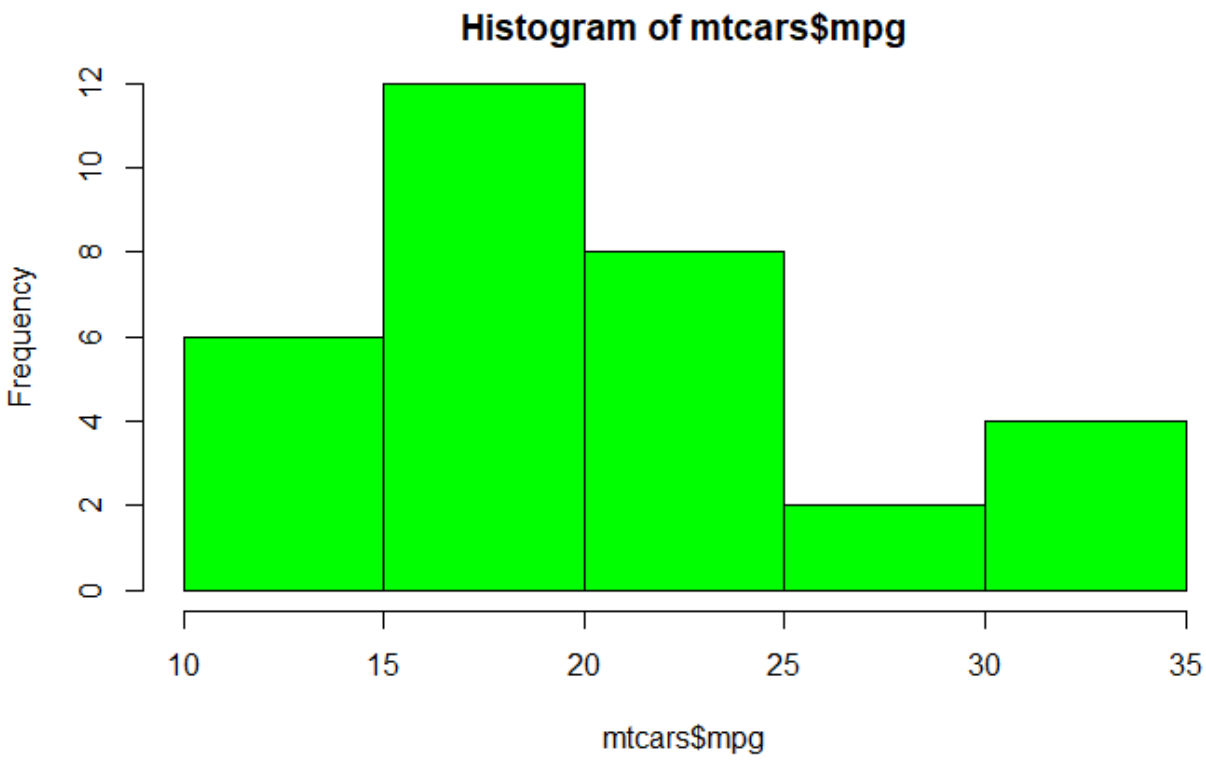
Output:

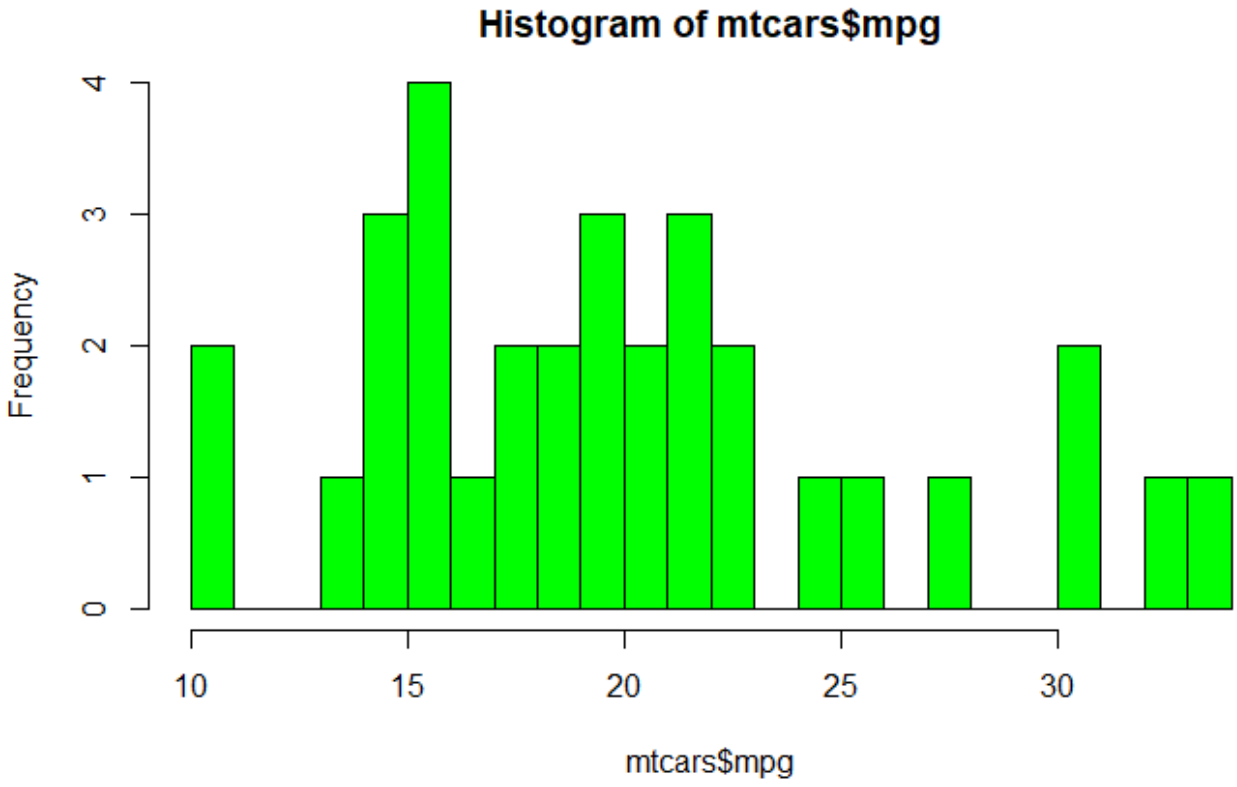


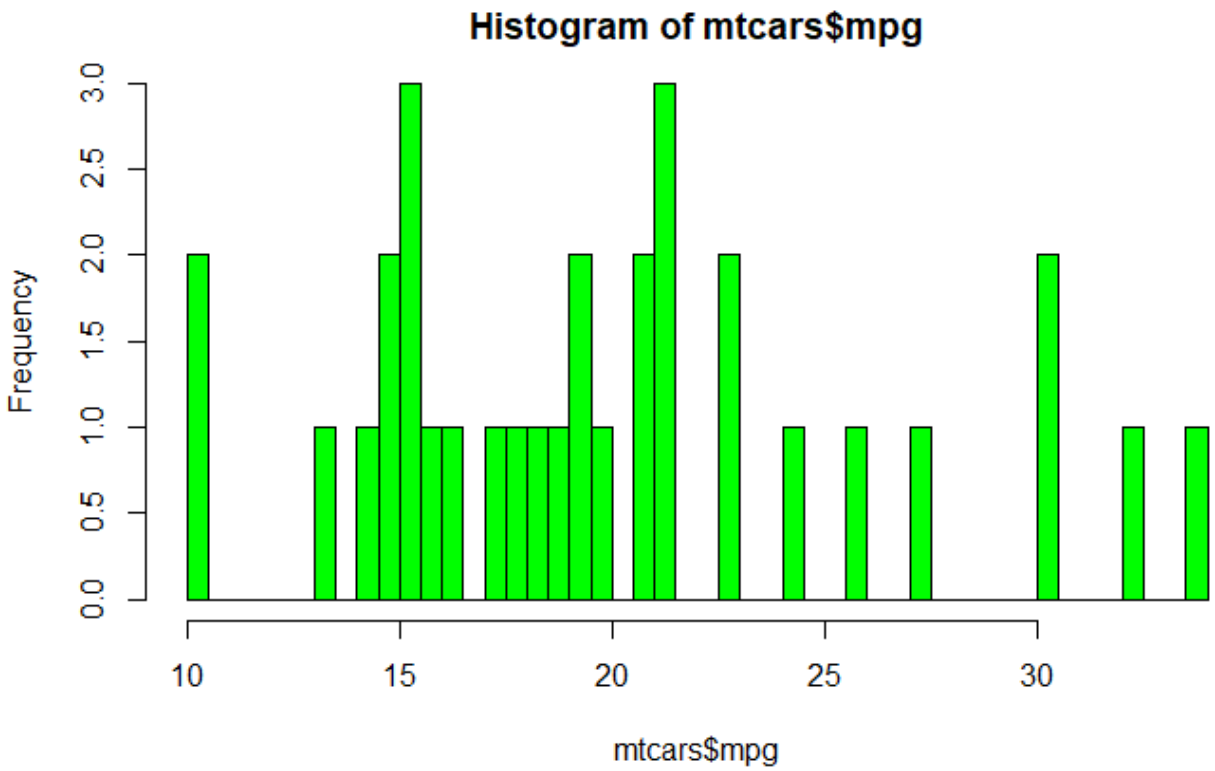
4) Using Histogram to show the information:

```{r}
hist(mtcars$mpg, col = "green")
hist(mtcars$mpg, col = "green", breaks = 25)
hist(mtcars$mpg, col = "green", breaks = 50)
```

Output:


Histogram of mtcars$mpg
Frequency
0
2
4
6
8
10
12
10
15
20
25
30
35
mtcars$mpg


Histogram of mtcars$mpg
Frequency
0
1
2
3
4
10
15
20
25
30
mtcars$mpg



5) Creating a CSV file to store data and read it to display:

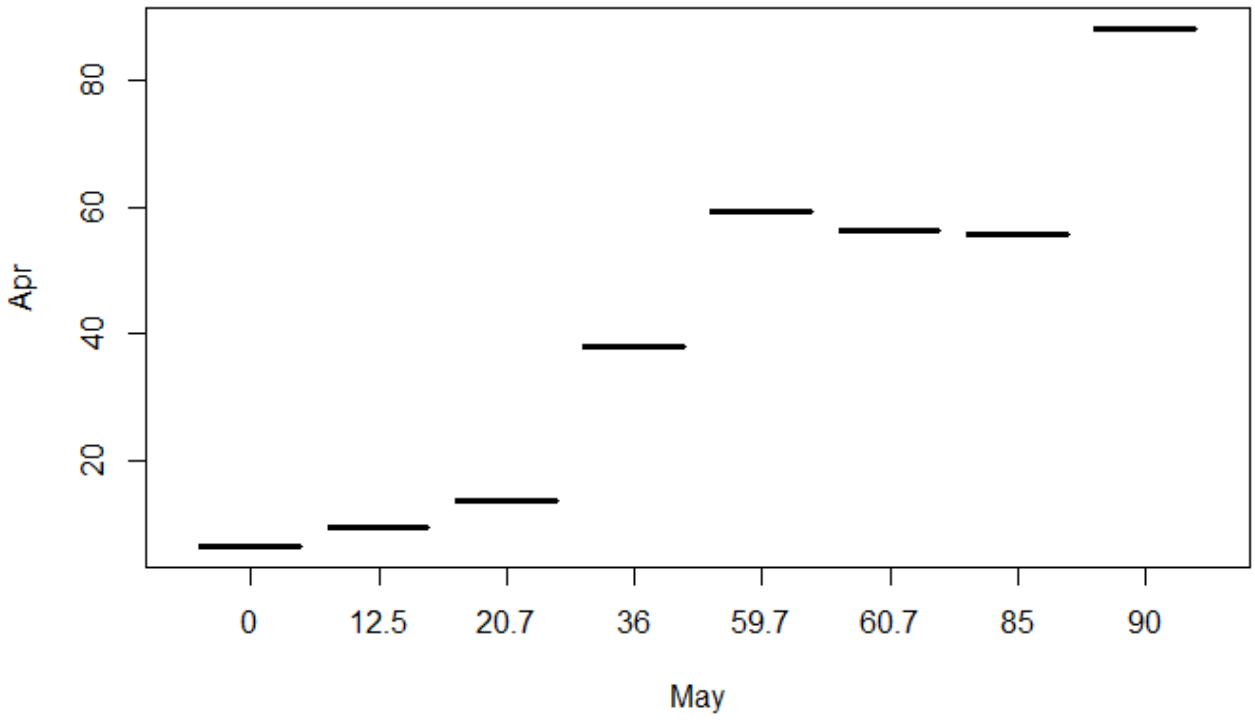
```{r}
Apr<-c(55.7, 88.3, 56.4, 59.3, 38.0, 13.7, 6.5, 9.4)
May<-c(85.0, 90.0, 60.7, 59.7, 36.0, 20.7, 0.0, 12.5)
Jun<-c(83.5, 100.0, 50.7, 39.6, 31.7, 26.7, 0.0, 12.9)
Jul<-c(94.8, 85.4, 66.3, 78.4, 22.2, 18.8, 8.6, 3.6)
Aug<-c(66.3, 103.0, 59.8, 29.3, 39.0, 12.3, 3.2, 6.4)
Sep<-c(75.9, 91.2, 23.5, 27.1, 28.0, 8.2, 17.8, 11.8)
Oct<-c(75.5, 65.7, 23.2, 46.6, 25.0, 24.1, 23.7, 14.3)
df<-data.frame(Apr, May, Jun, Jul, Aug, Sep, Oct)
print(df)
boxplot(Apr~May,data=df, col="red")
write.csv(df,"C:\\Users\\Hruthesh\\Documents\\Lab_1.csv",row.names = FALSE)
read.csv("C:\\Users\\Hruthesh\\Documents\\Lab_1.csv")
```

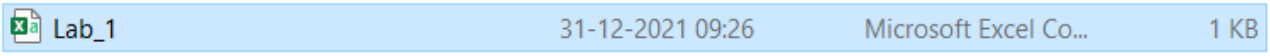
Output:

data.frame
8 x 7

data.frame
8 x 7

Apr <dbl>	May <dbl>	Jun <dbl>	Jul <dbl>	Aug <dbl>	Sep <dbl>	Oct <dbl>
55.7	85.0	83.5	94.8	66.3	75.9	75.5
88.3	90.0	100.0	85.4	103.0	91.2	65.7
56.4	60.7	50.7	66.3	59.8	23.5	23.2
59.3	59.7	39.6	78.4	29.3	27.1	46.6
38.0	36.0	31.7	22.2	39.0	28.0	25.0
13.7	20.7	26.7	18.8	12.3	8.2	24.1
6.5	0.0	0.0	8.6	3.2	17.8	23.7
9.4	12.5	12.9	3.6	6.4	11.8	14.3





The above is the csv file we saved by taking the above information and we used box plot to show in the graph.

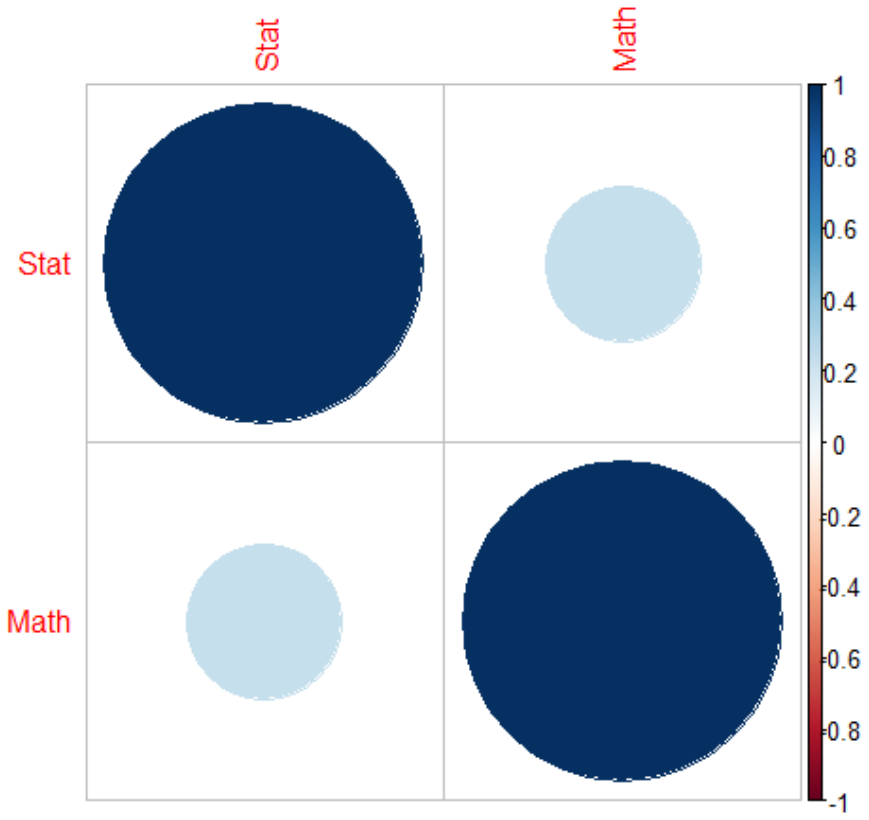
6) Using correlogram to plot information using different methods:

```{r}

library(corrplot)

Stat<-c(70,92,80,74,65,83)

Math<-c(74,84,63,87,78,90)
```

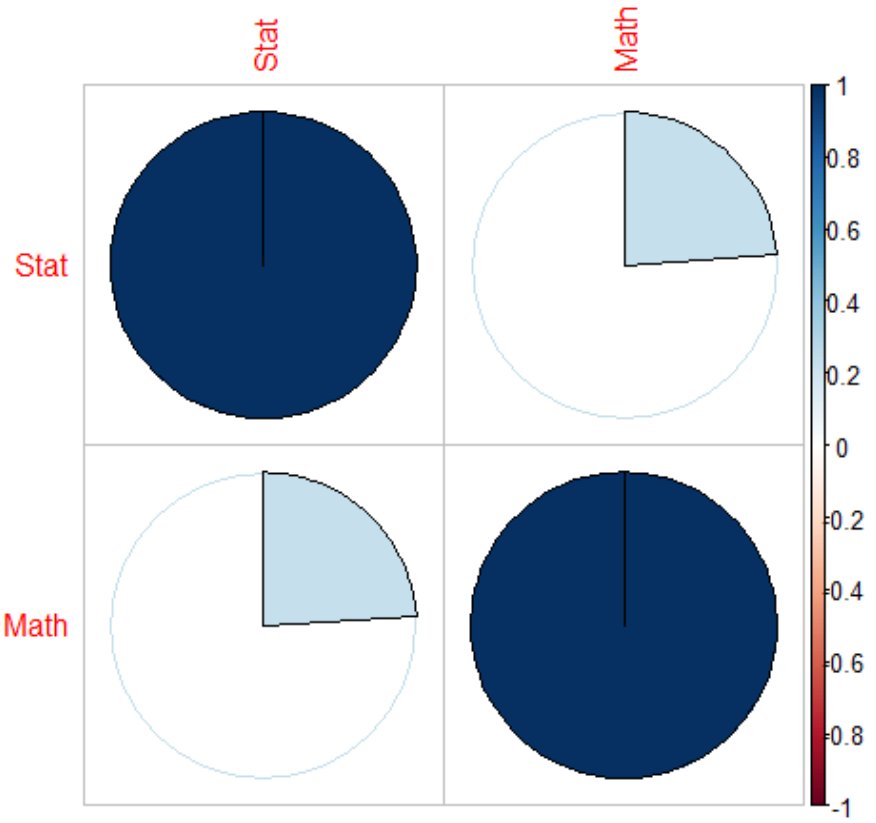
```
df<-data.frame(Stat, Math)
M<-cor(df)
corrplot(M, method="circle")
corrplot(M, method="pie")
corrplot(M, method="color")
corrplot(M, method="number")
```

Output:

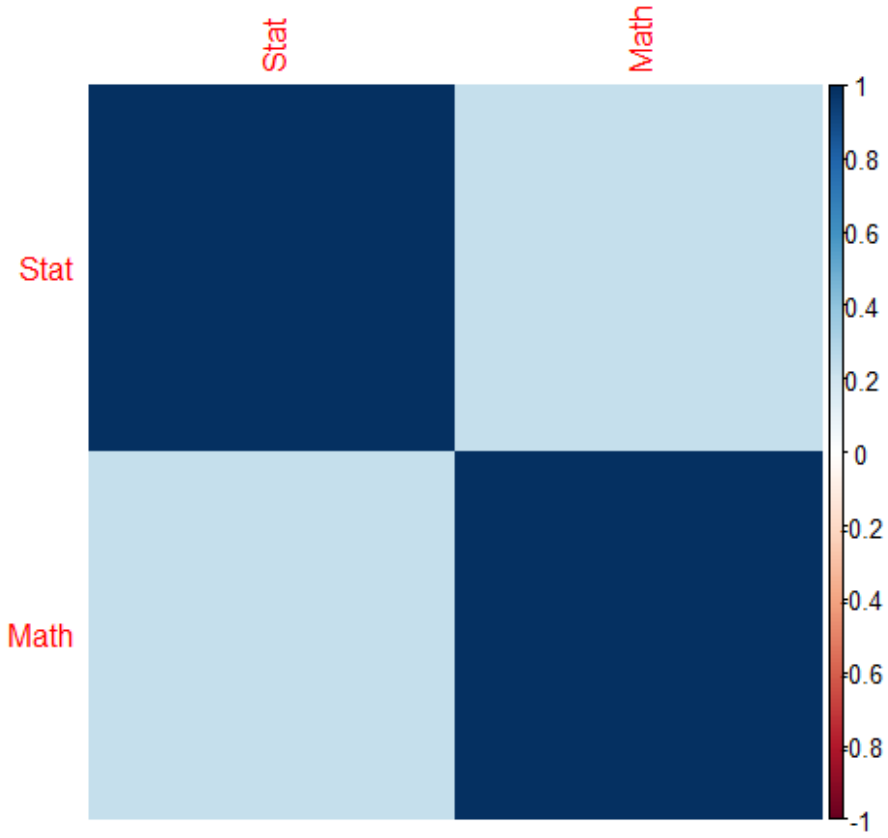
Circle Method:



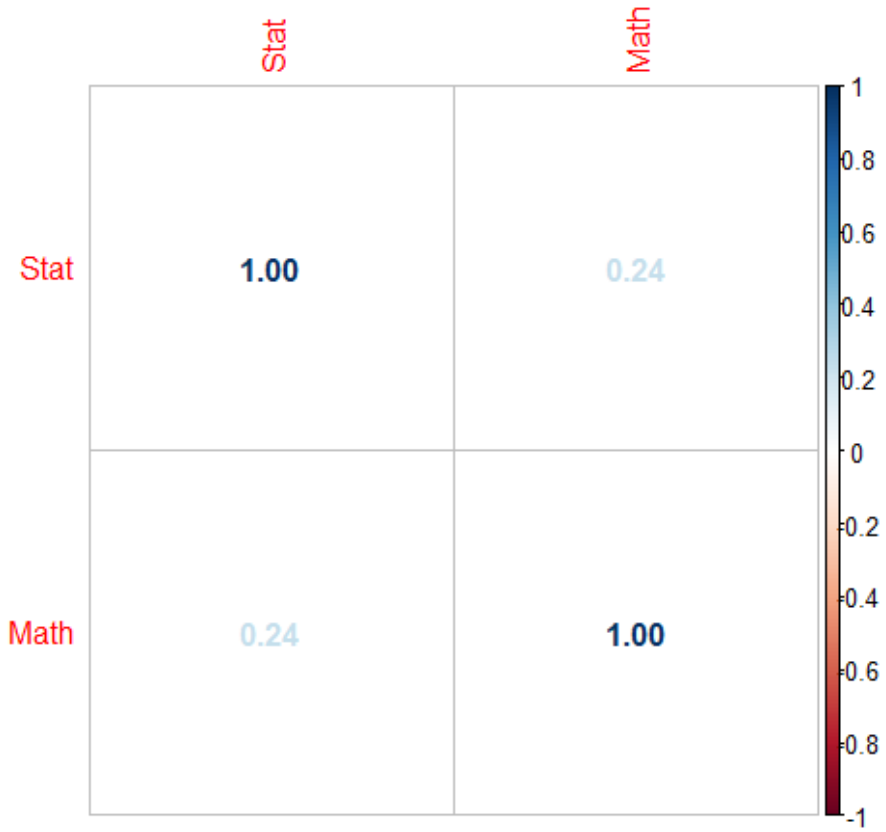
Pie Method:



Colour Method:



Number Method:



7) Calculating correlations using different methods:

```{r}
x <- c(1, 3, 5, 10)
y <- c(2, 4, 6, 20)
print(cor(x, y))
print(cor(x, y, method = "pearson"))
print(cor(x, y, method = "kendall"))
print(cor(x, y, method = "spearman"))

```

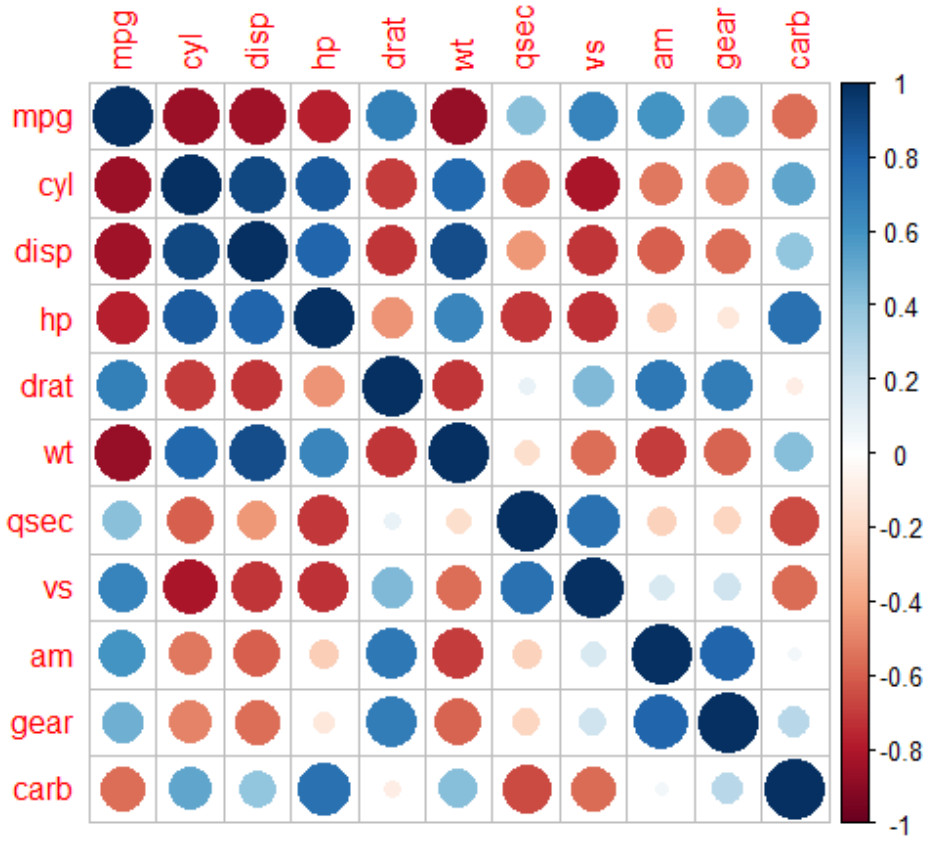
Output:

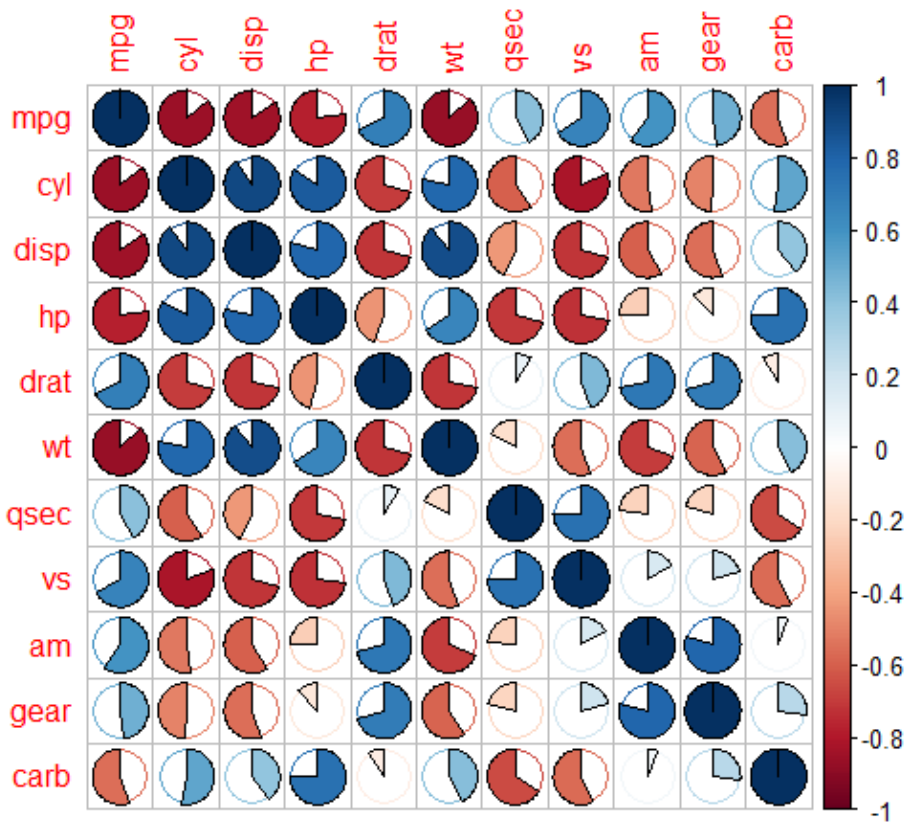
```
[1] 0.9724702
[1] 0.9724702
[1] 1
[1] 1
```

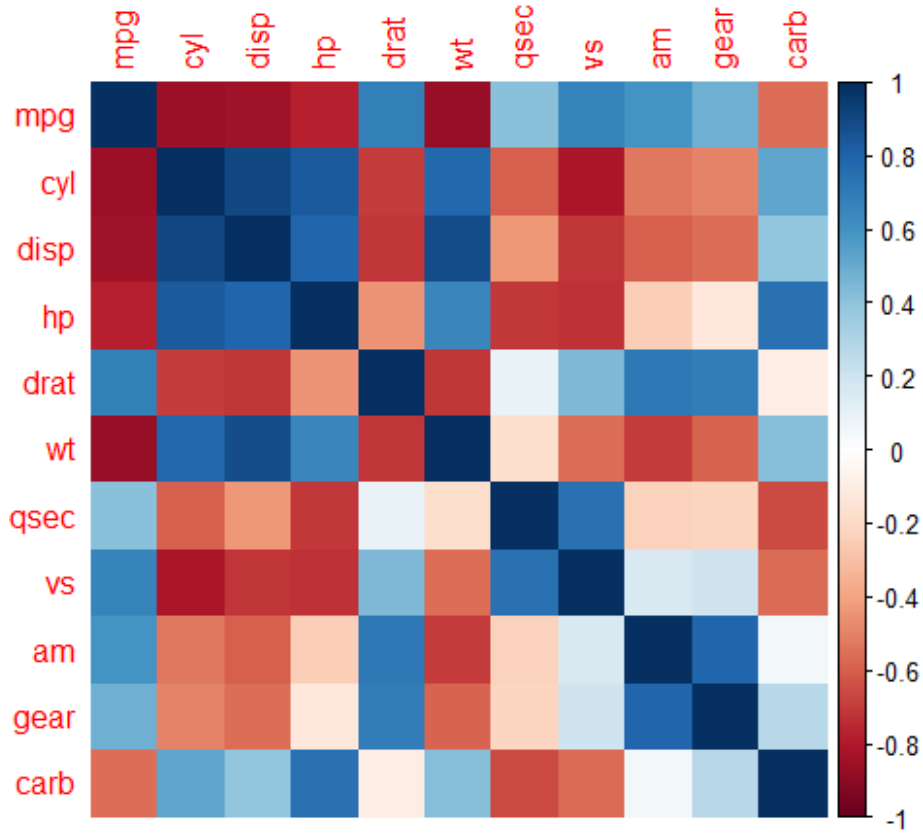
8) Correlation matrix using correlogram:

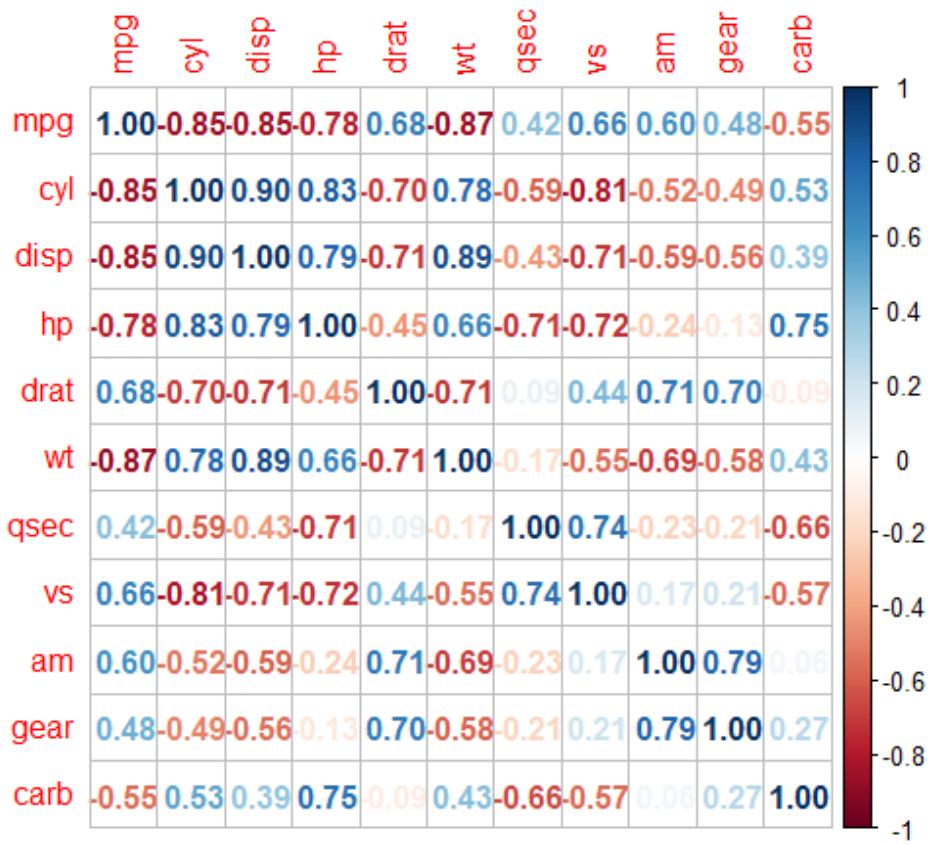
```{r}
library(corrplot)

head(mtcars)
M<-cor(mtcars)
head(round(M,2))

corrplot(M, method="circle")
corrplot(M, method="pie")
corrplot(M, method="color")
corrplot(M, method="number")
```

Output:




mpg
cyl
disp
hp
drat
wt
qsec
vs
am
gear
carb
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1


mpg
cyl
disp
hp
drat
wt
qsec
vs
am
gear
carb
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

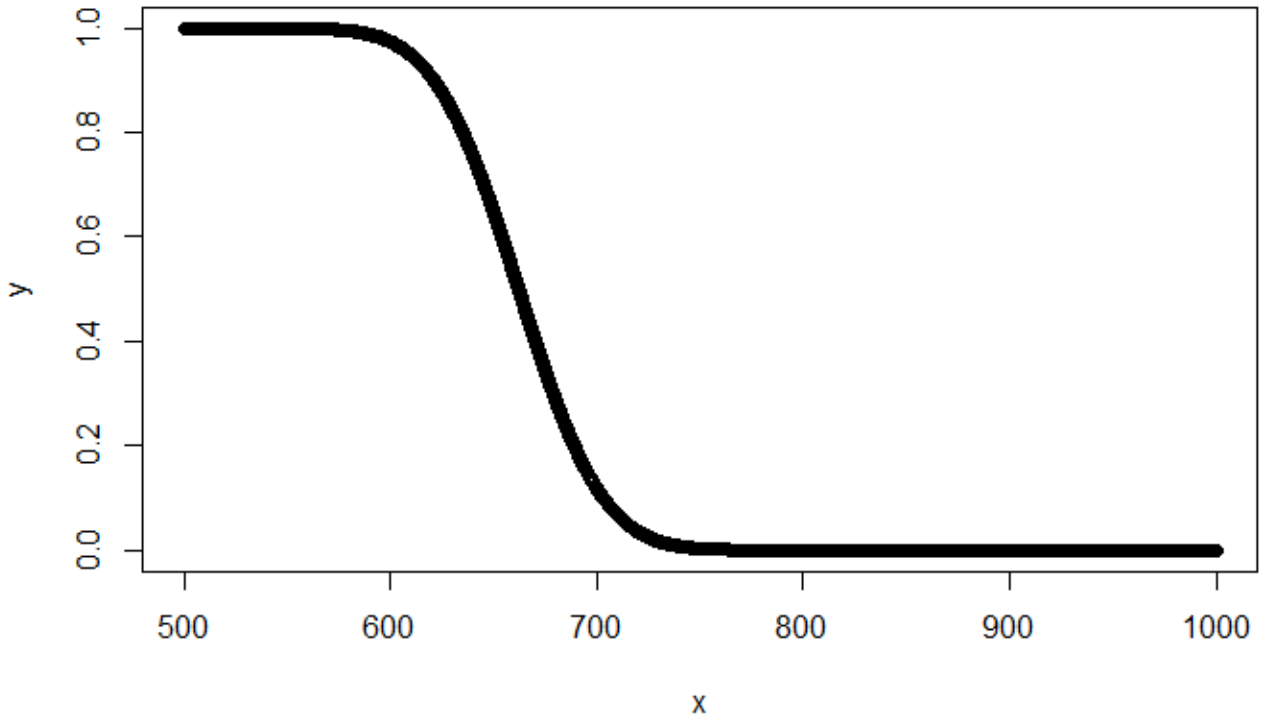


9) Plotting values with given values and cumulative density function of normal distribution:

i)

```{r}
x<-seq(500,1000,by=0.1)
y<-pnorm(x,mean=662,sd=32,lower.tail=FALSE)
plot(x,y)
```

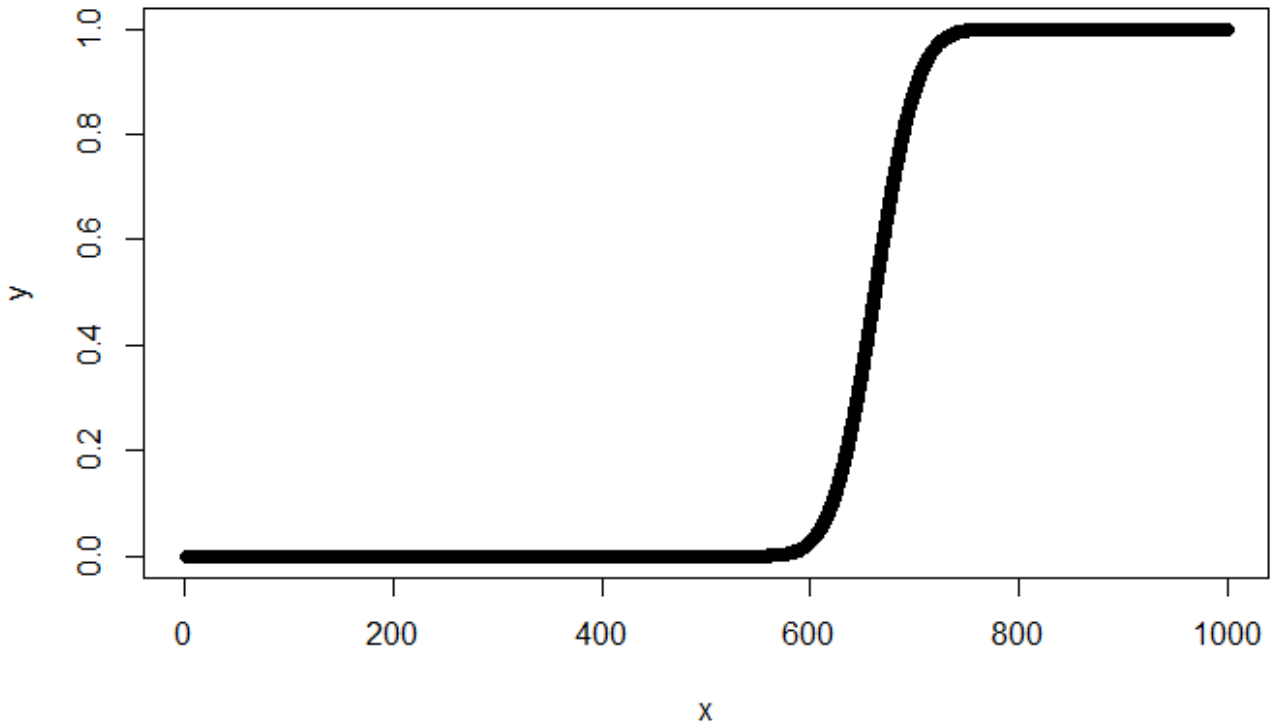
Output:



ii)

```{r}
x<-seq(1,1000,by=0.1)
y<-pnorm(x,mean=662,sd=32,lower.tail=TRUE)
plot(x,y)
```

Output:



10) Generating Multivariate normal distribution

```{r}
library("MASS")
mu=c(2,3)
sigma=matrix(c(9,6,6,16),2,2) #positive definite 2x2 matrix

variables=mvrnorm(1000,mu,sigma)

head(variables)
tail(variables)
```

Output:

```
          [,1]      [,2]
[1,] -3.544528 -5.572673
[2,]  3.483154  3.256930
[3,]  1.785419 -5.483503
```

```
[4,]  3.890355  9.793126
[5,]  4.089184  9.867275
[6,]  4.672605  4.422452
            [,1]      [,2]
 [995,] -2.5922118 -5.389564
 [996,]  2.0716408  2.093541
 [997,]  2.7428303  5.254111
 [998,]  0.6236855 -1.309172
 [999,]  0.5013315 -2.765284
[1000,] -0.8429362  3.746520
```

11) Program to illustrate linear regression:

```{r}
x <- c(153, 169, 140, 186, 128,
          136, 178, 163, 152, 133)

y <- c(64, 81, 58, 91, 47, 57,
          75, 72, 62, 49)

model <- lm(y~x)

print(model)

df <- data.frame(x = 182)
res <- predict(model, df)
cat("\nPredicted value of a person
                              with height = 182")
print(res)

```

Output:

Call:

lm(formula = y ~ x)

Coefficients:

(Intercept) x

-39.7137 0.6847

Predicted value of a person

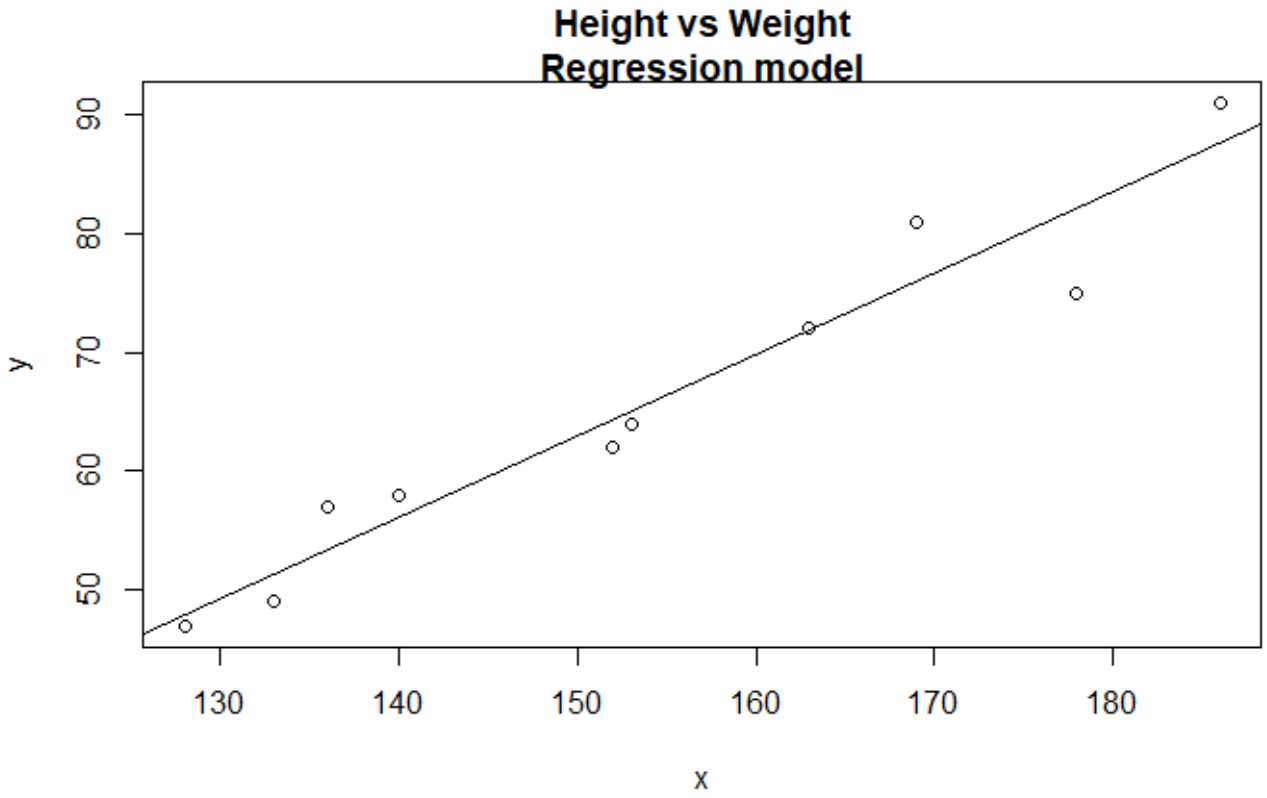
with height = 182 1

84.9098

Plot the values of the regression model:

```{r}

# Plot

plot(x, y, main = "Height vs Weight

Regression model")

abline(lm(y~x))

```

Output:



12) Program to illustrate multiple linear regression:

```{r}
input <- airquality[1:50,
                    c("Ozone", "Wind", "Temp")]

model <- lm(Ozone~Wind + Temp,
                        data = input)

cat("Regression model:\n")
print(model)

png(file = "multipleRegGFG.png")

plot(model)
dev.off()
```

Output:

Regression model:

Call:

lm(formula = Ozone ~ Wind + Temp, data = input)

Coefficients:

(Intercept)	Wind	Temp
-58.239	-0.739	1.329

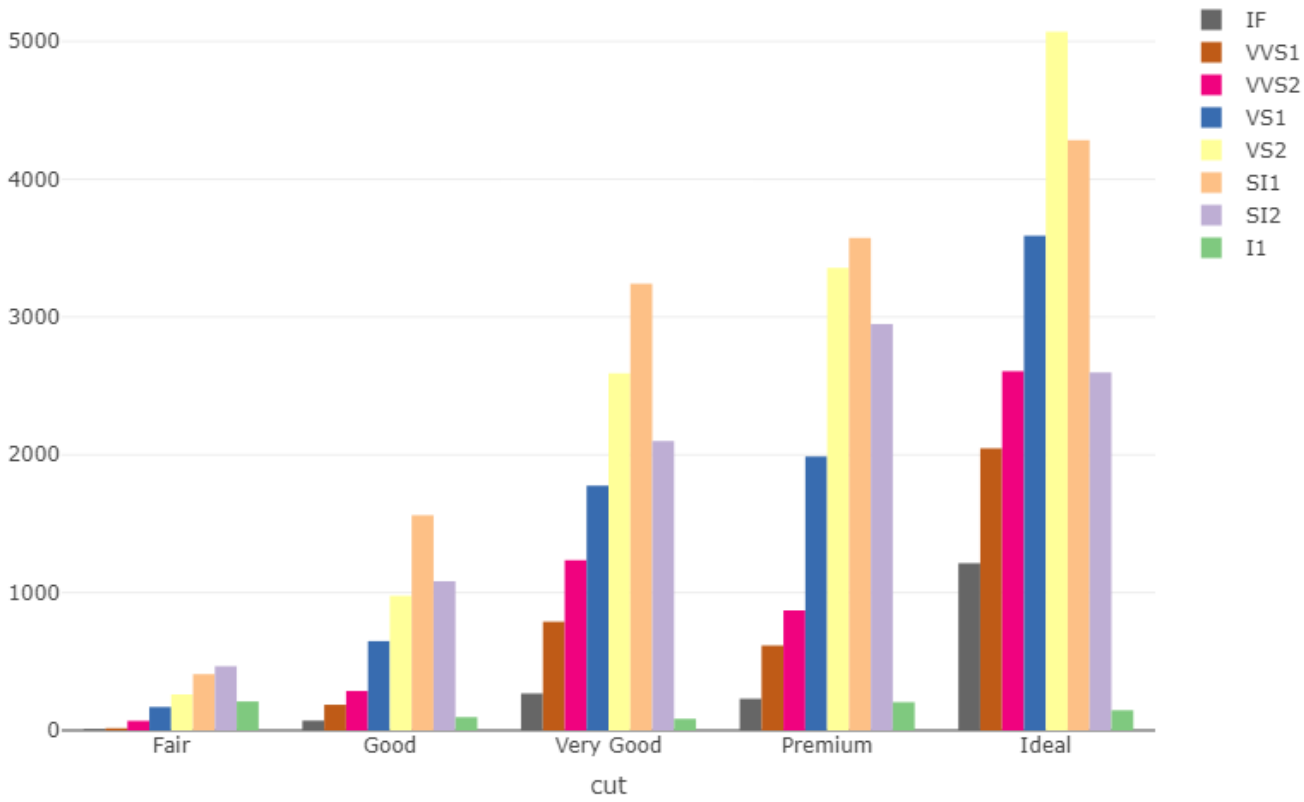
null device

1

13) Program to illustrate the visualization:

```{r}
plot_ly(diamonds, x = ~cut, color = ~clarity, colors = "Accent")
```

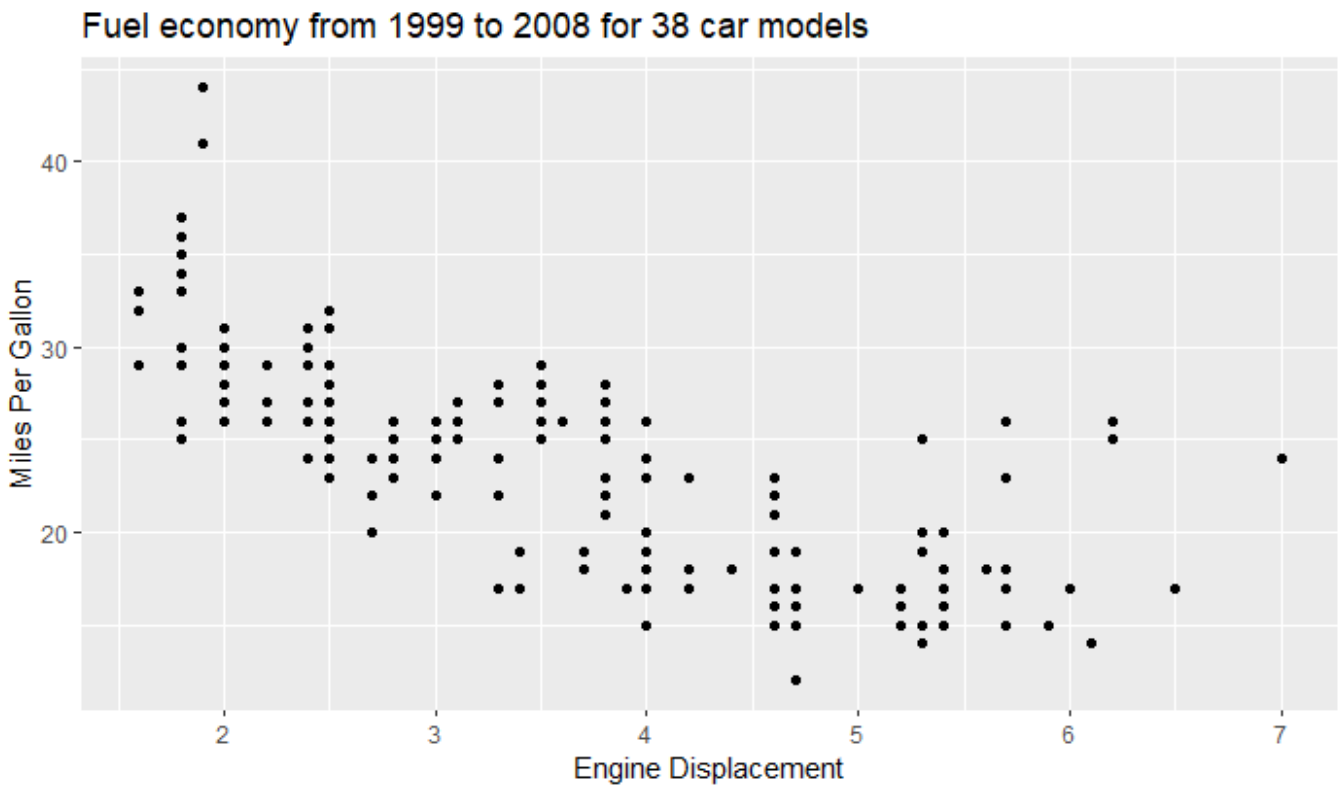
Output:



14) Using ggplot to visualize the data:

```{r}
library(ggforce)
ggplot(mpg, aes(displ, hwy)) +
geom_point() +
geom_mark_hull(aes(filter = model == "corvette", label = model)) +
labs(
title = "Fuel economy from 1999 to 2008 for 38 car models",
caption = "Source: https://fueleconomy.gov/",
x = "Engine Displacement",
y = "Miles Per Gallon")
```

Output:


Fuel economy from 1999 to 2008 for 38 car models
Miles Per Gallon
40
30
20
2
3
4
5
6
7
Engine Displacement