CSE – 4020 – Machine Learning

Name: Priyank Kumar

Reg. No: 16BCE2031

Slot: L59 + L60

Lab Assignment – 4

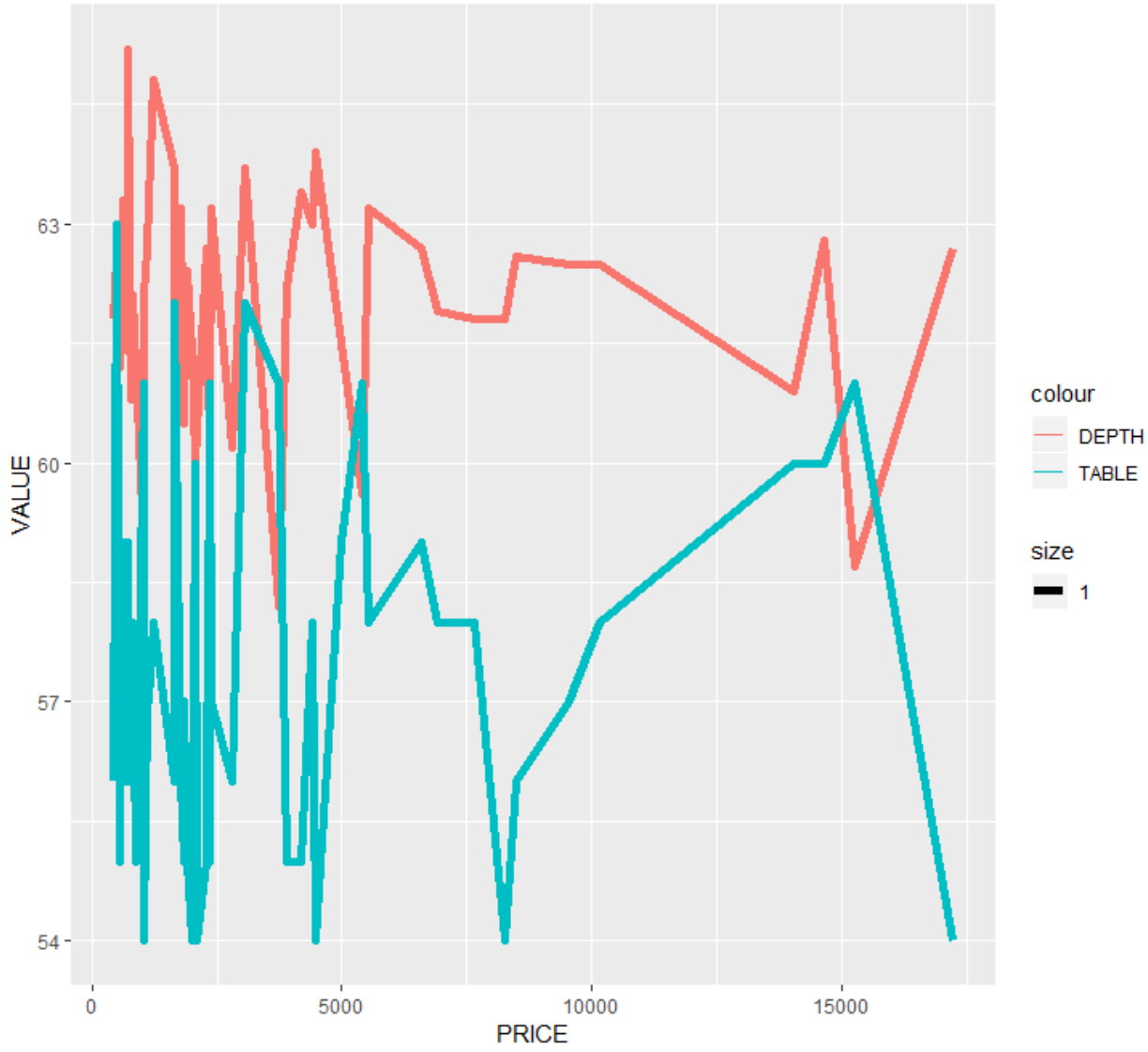
Dataset 1: Diamonds

Splitting the Dataset:

```
> library(ggplot2)
> data<-diamonds
> smp_size = floor(0.999*nrow(data))
> t_ind = sample(seq_len(nrow(data)),size = smp_size)
> train = data[t_ind,]
> test = data[-t_ind,]
```

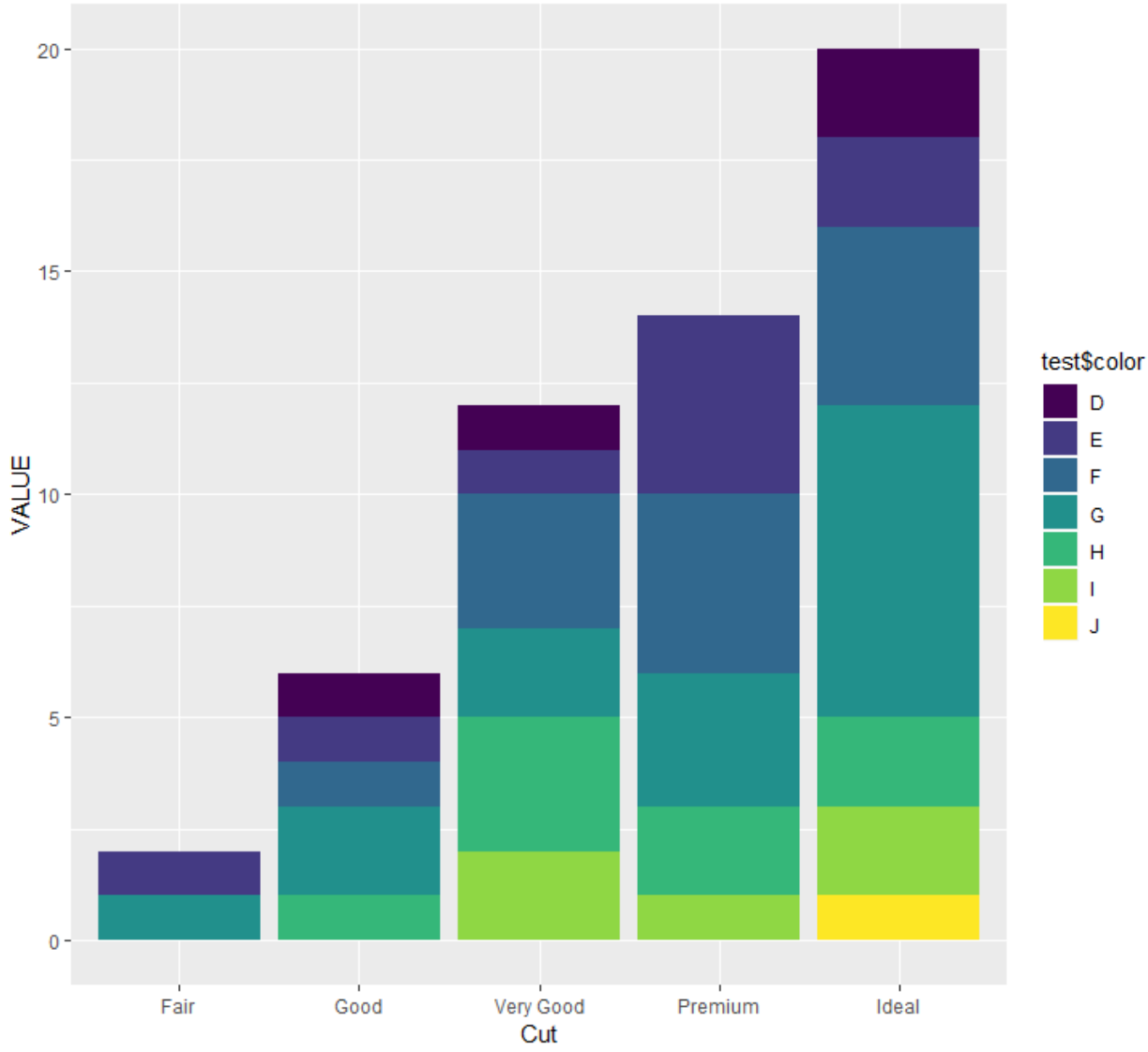
LINE GRAPH:

```
ggplot(test, aes(test$price, size="1"))+
  geom_line(aes(y=test$depth, colour="DEPTH"))+
  geom_line(aes(y=test$table, colour="TABLE"))+
  ylab("VALUE")+
  xlab("PRICE")
```



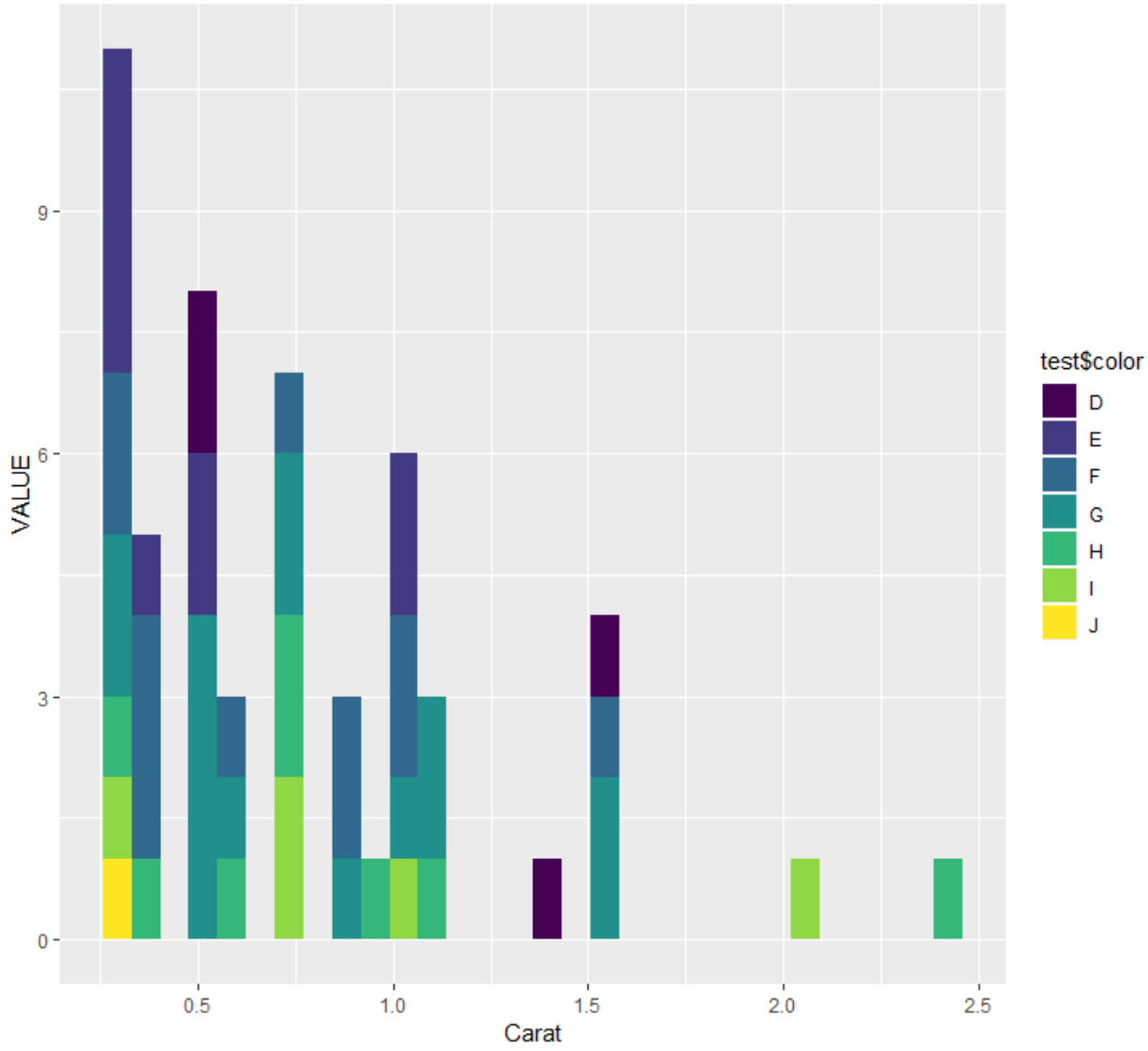
BAR GRAPH:

```
ggplot(test, aes(test$cut,fill=test$color))+
  geom_bar()+
  ylab("VALUE")+
  xlab("Cut")
```



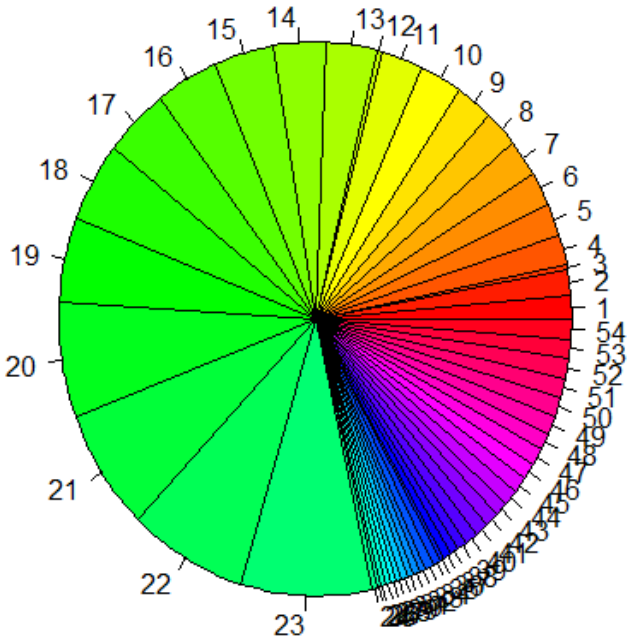
HISTOGRAM

```
ggplot(test, aes(test$carat,fill=test$color))+
  geom_histogram()+
  ylab("VALUE")+
  xlab("Carat")
```



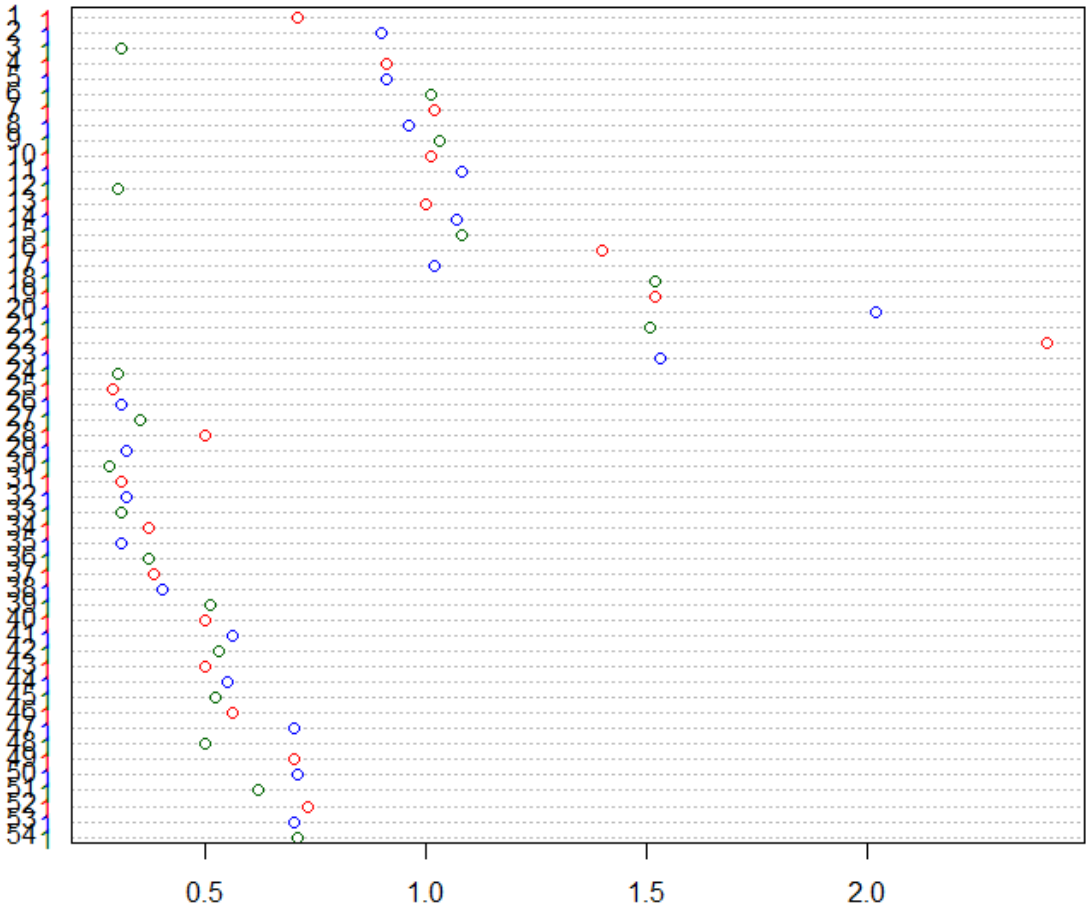
PIE CHART

```
pie(test$price,col=rainbow(length(test$price)))
```



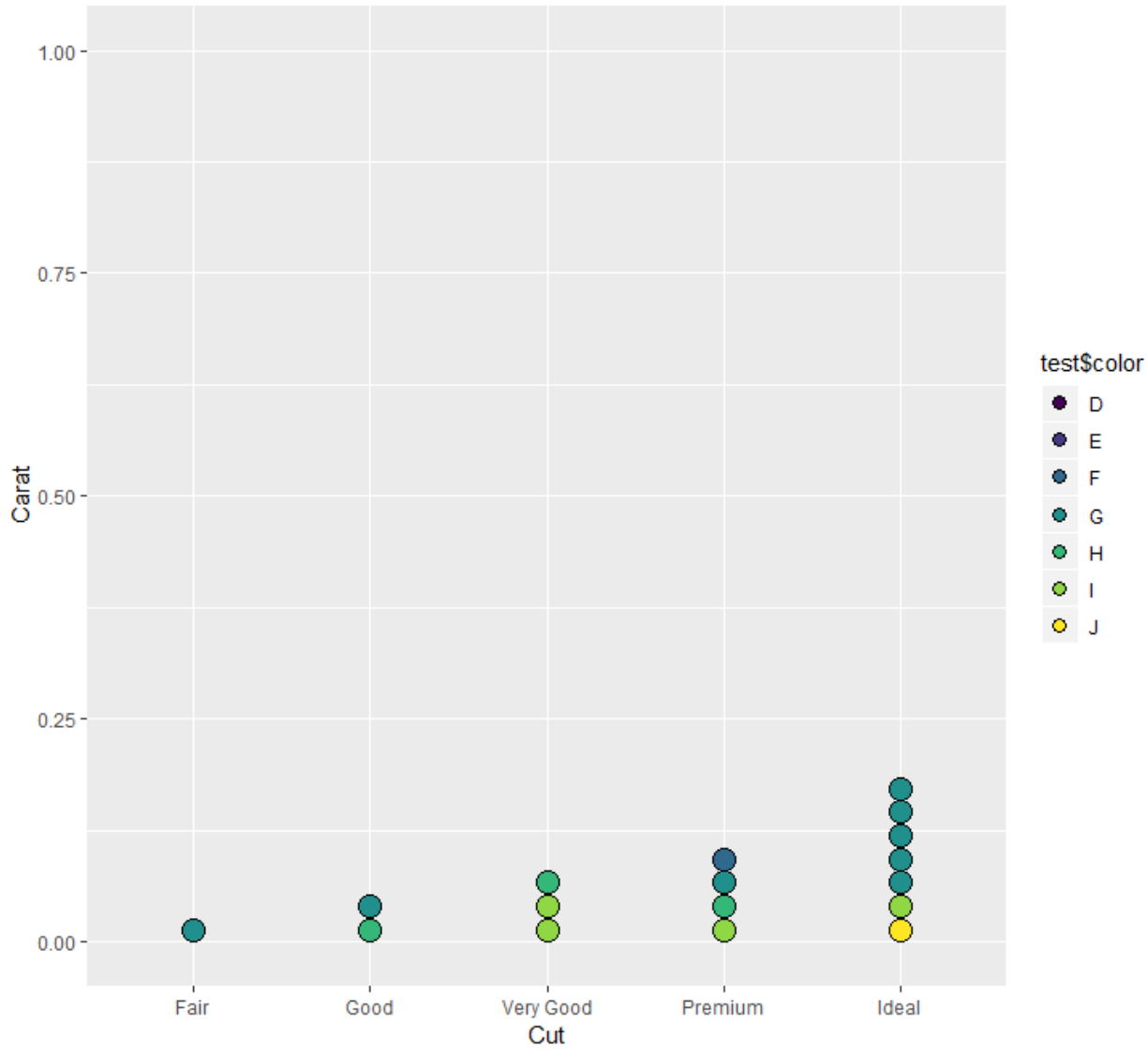
DOT CHART

```
dotchart(t(test$carat),col=c("red","blue","darkgreen"))
```



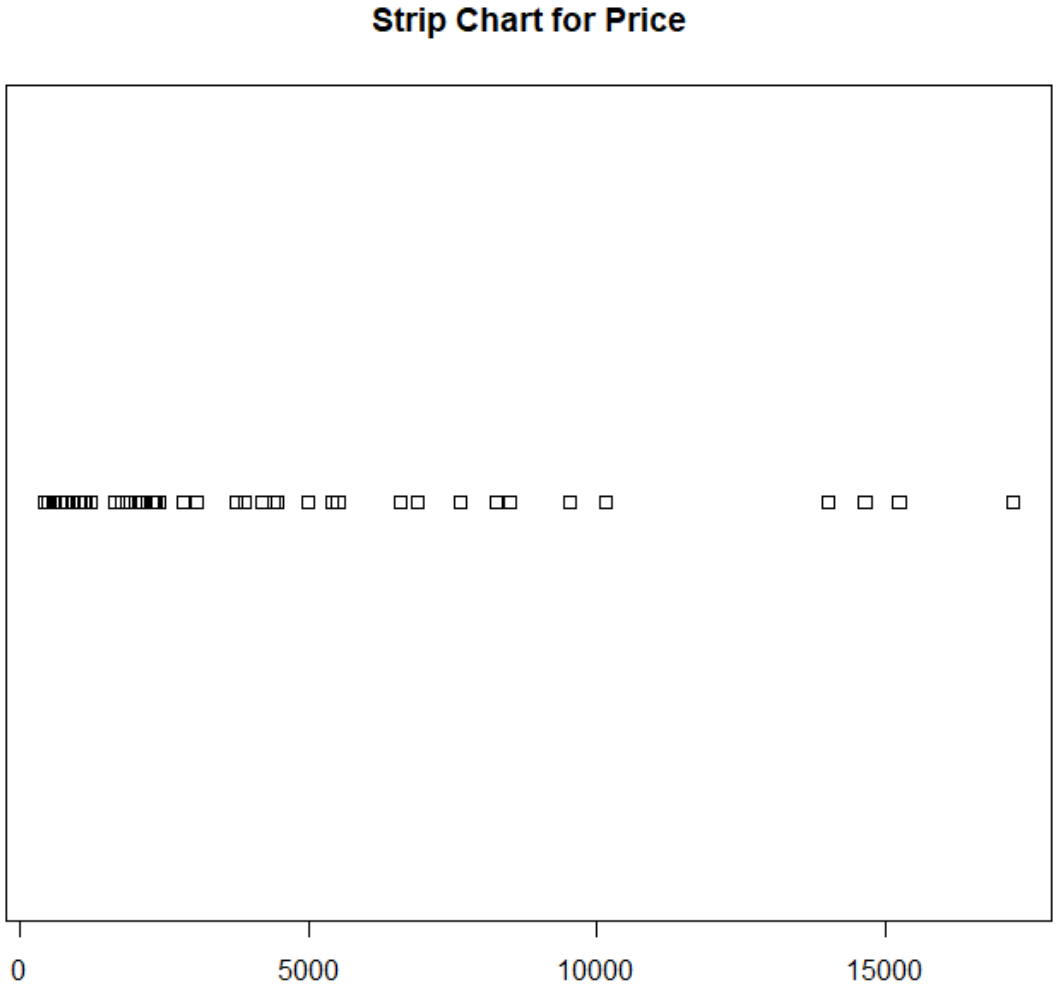
DOTPLOT

```
ggplot(test, aes(x=test$cut,y=test$carat,fill=test$color))+
  geom_dotplot()+
  ylab("Carat")+
  xlab("Cut")
```



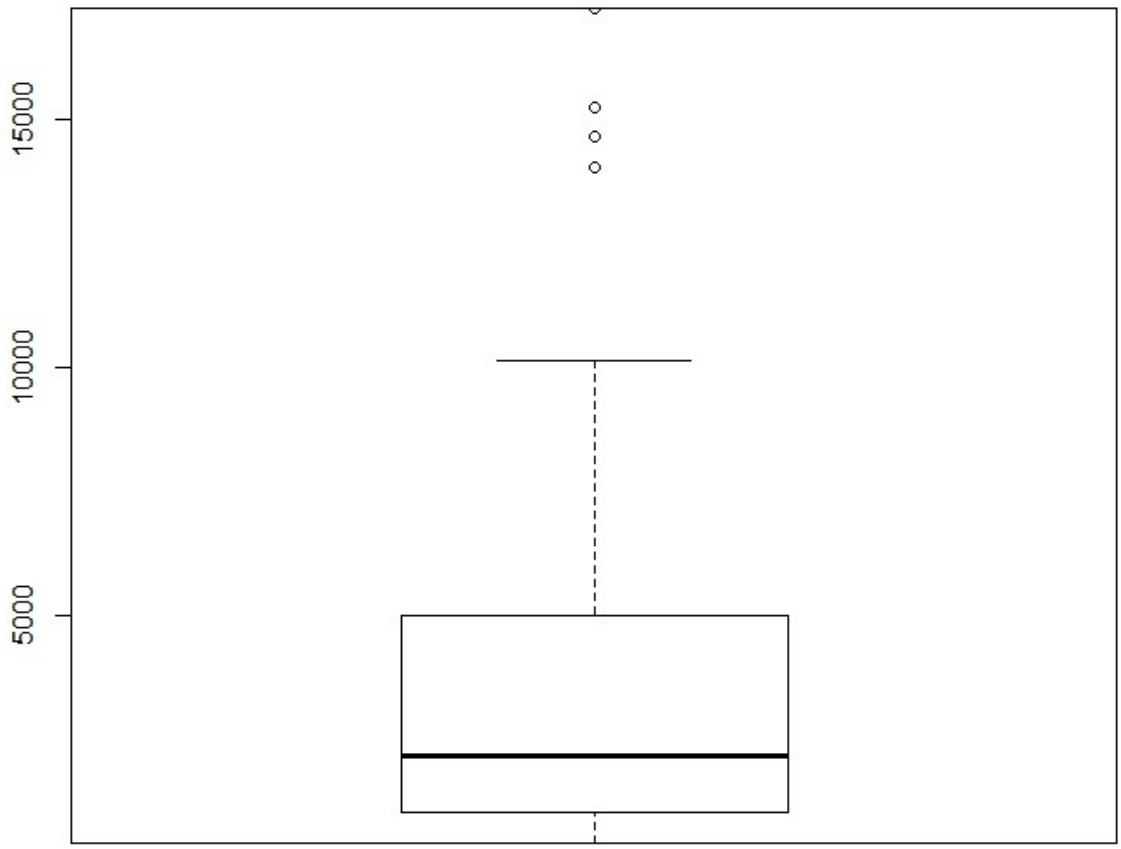
STRIP CHART

```
stripchart(test$price,method = "stack",main="Strip Chart for Price")
```



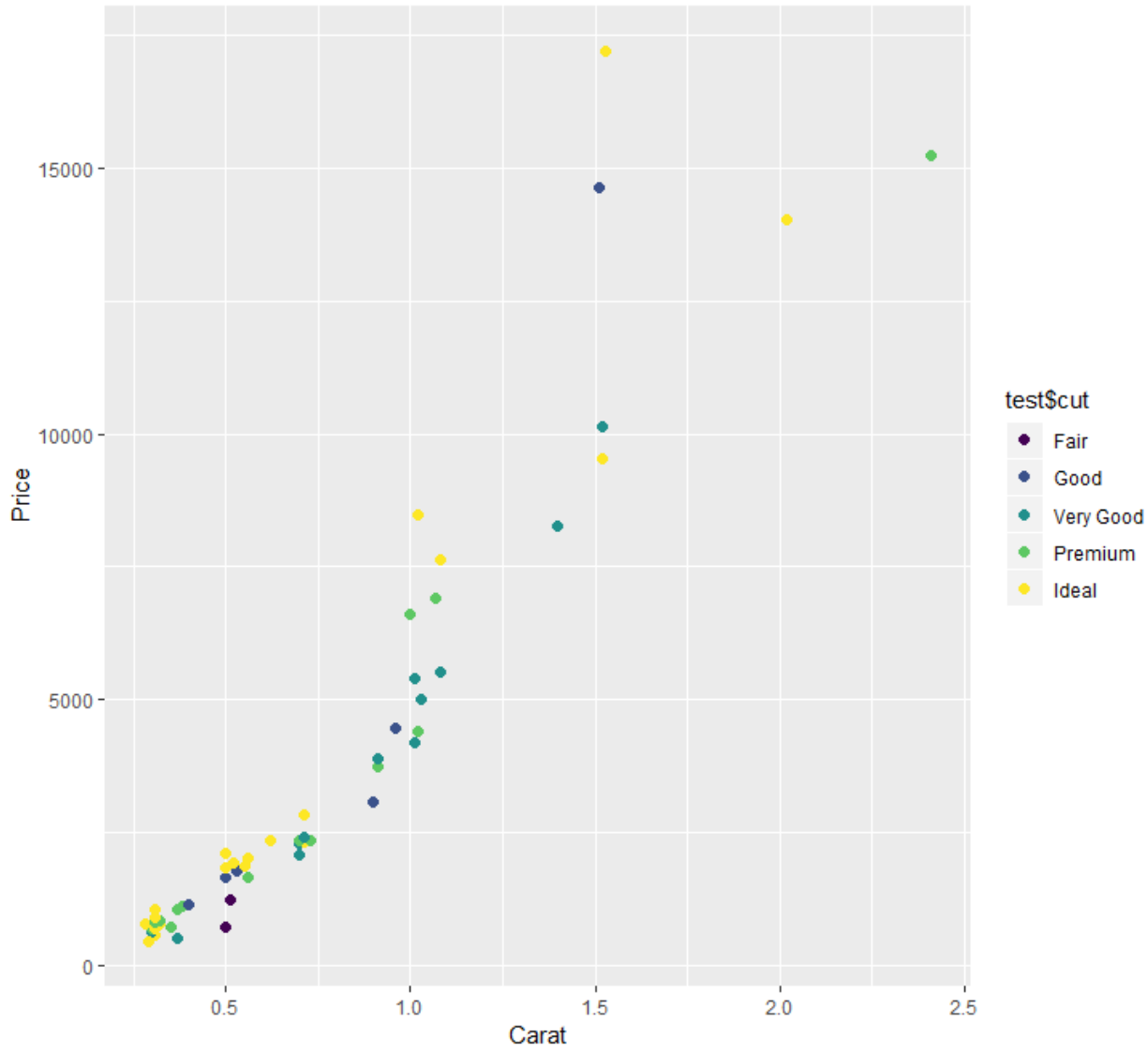
BOX PLOT

```
boxplot(test$price)
ggplot(test, aes(x=test$carat,y=test$price,fill=test$color))+
  geom_boxplot()+
  ylab("Carat")+
  xlab("Cut")
```



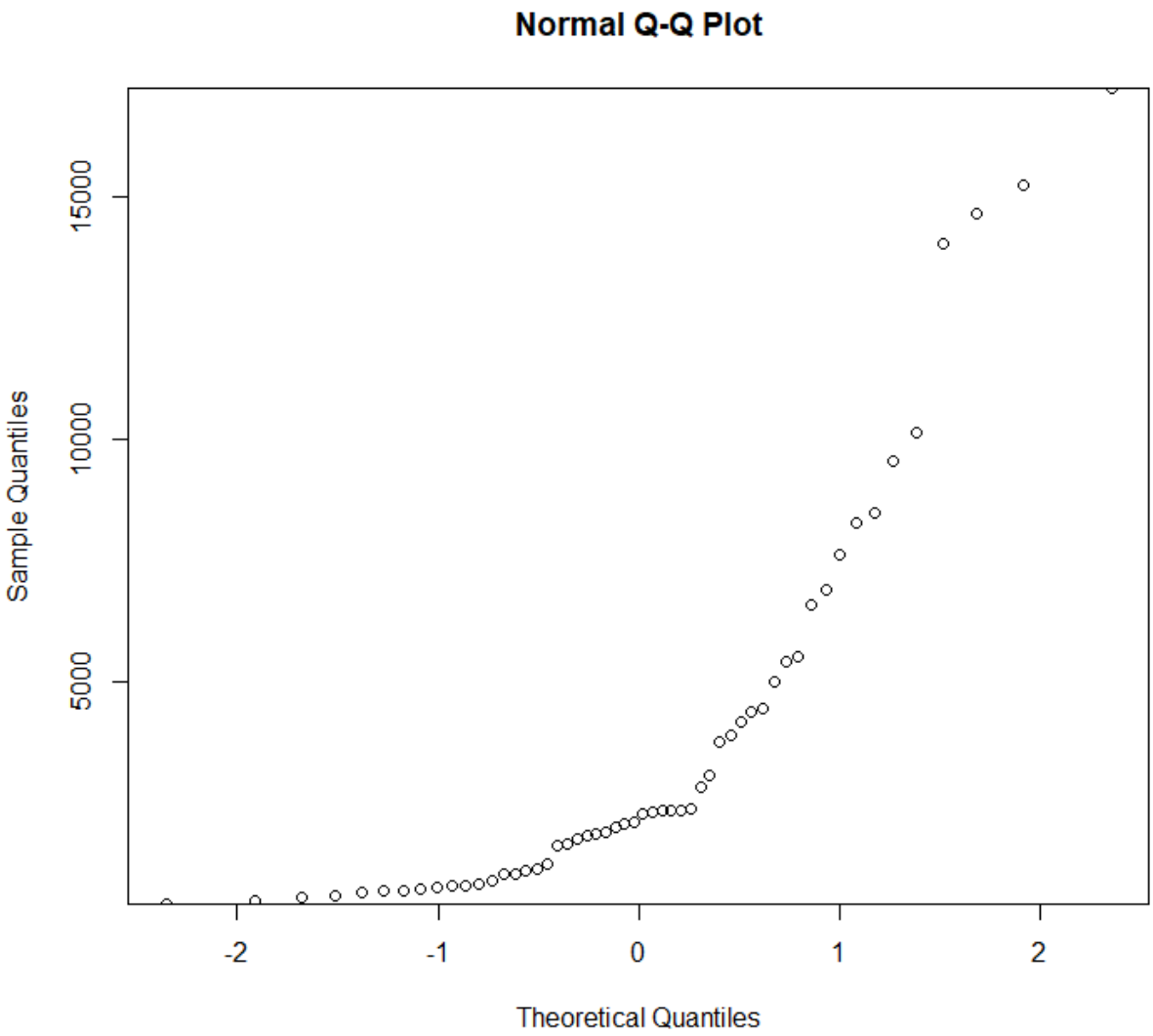
POINTS PLOT

```
ggplot(test, aes(x=test$carat,y=test$price,colour=test$cut))+
  geom_point(size=2.0)+
  ylab("Price")+
  xlab("Carat")
```



NORMAL QQ PLOT

```
qqnorm(test$price)
```

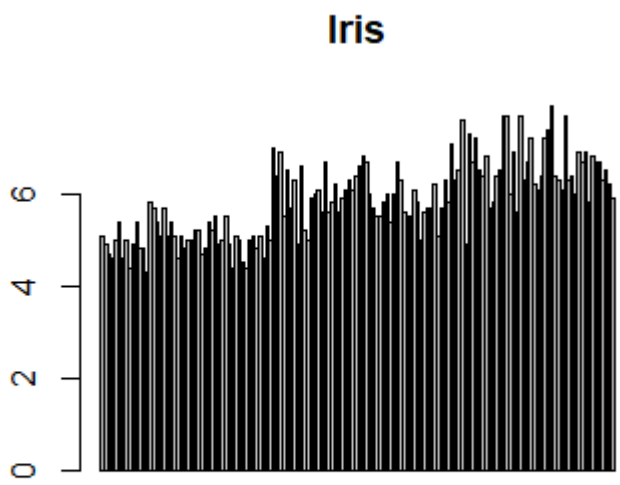


DATASET 2: IRIS

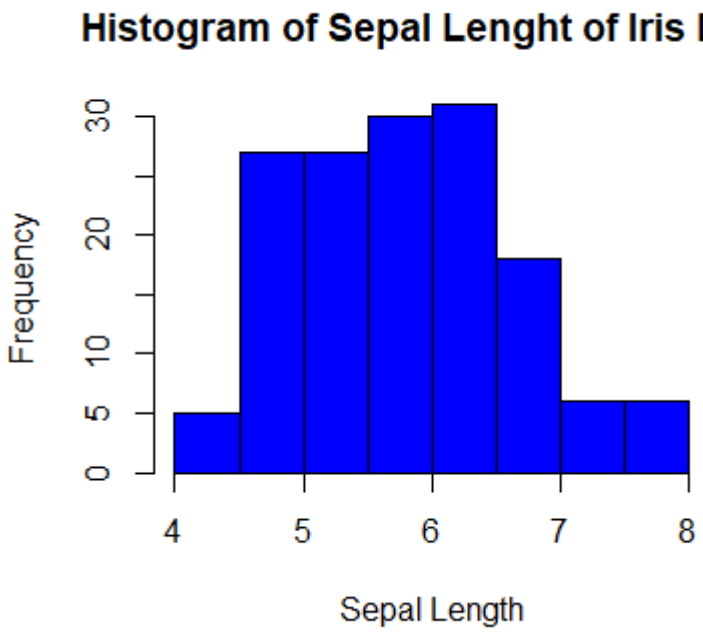
CODE:

```
head(iris)
plot(iris$Sepal.Width,type="o",col="yellow")
title(main="Iris Sepal length")
lines(iris$Sepal.Length,type="o",col="red")
a<-c(1,2,3,4,9)
b<-c(2,7,8,10)
plot(a,type="o",col="blue",ylim=c(0,12))
```

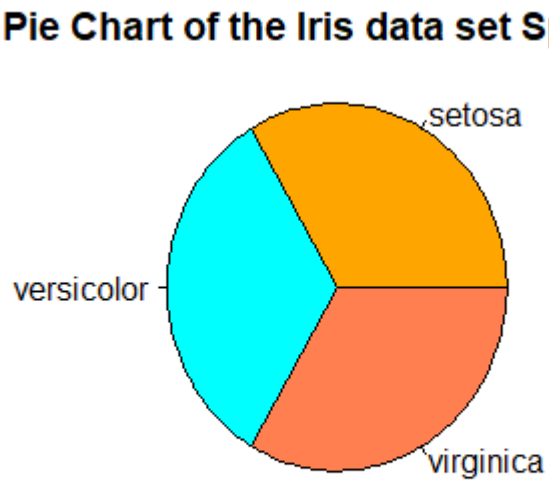
```
lines(b,type="o",pch=22,lty=2,col="red")

g_range<-range(0,iris$Sepal.Length,iris$Sepal.Width)

plot(iris$Sepal.Length,type="o",col="red",ylim=g_range,axes=FALSE,ann=FALSE)

axis(1,at=1:50)

axis(2,las=1,at=4*0:g_range[2])

box()



barplot(iris$Sepal.Length,main="Iris")

counts <- table(iris$Sepal.Length,iris$Sepal.Width)

barplot(counts, main="Iris Distribution by length and width",xlab="length",
col=c("darkblue","red"),legend = rownames(counts))
```




```
#histogram

hist(iris$Sepal.Length, col = "blue", xlab = "Sepal Length", main =

    "Histogram of Sepal Lenght of Iris Data")
```

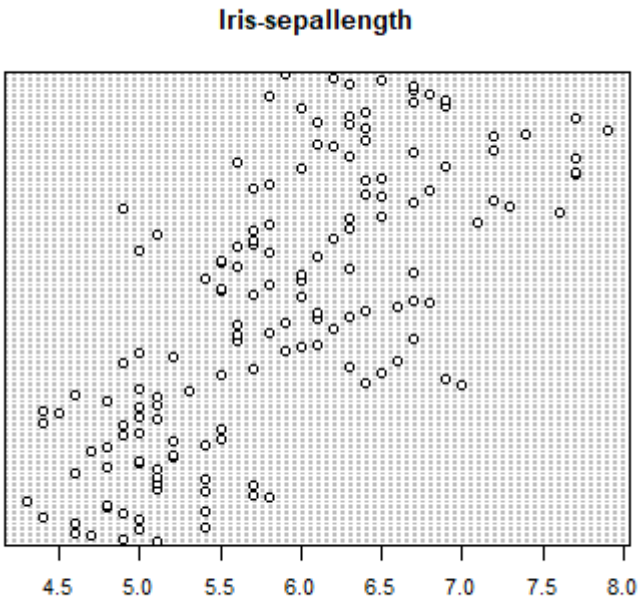



```
#pie
table(iris$Species)
pie(table(iris$Species), main = "Pie Chart of the Iris data set Species",
   col = c("orange", "cyan", "coral"), radius = 1)
```



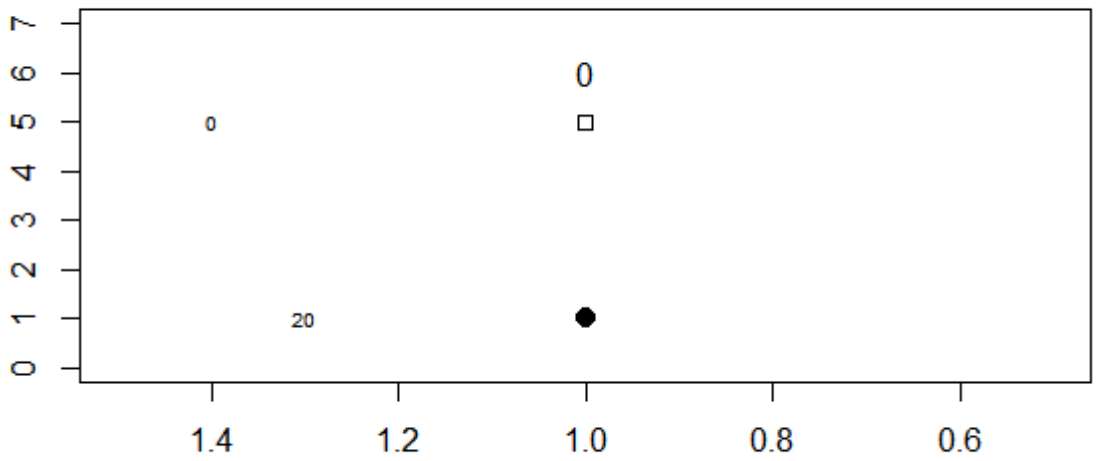
#dotchart

```
dotchart(iris$Sepal.Length,labels=row.names(iris$Sepal.Length),cex=.7,main="Iris-sepallength", )
```



#MISC

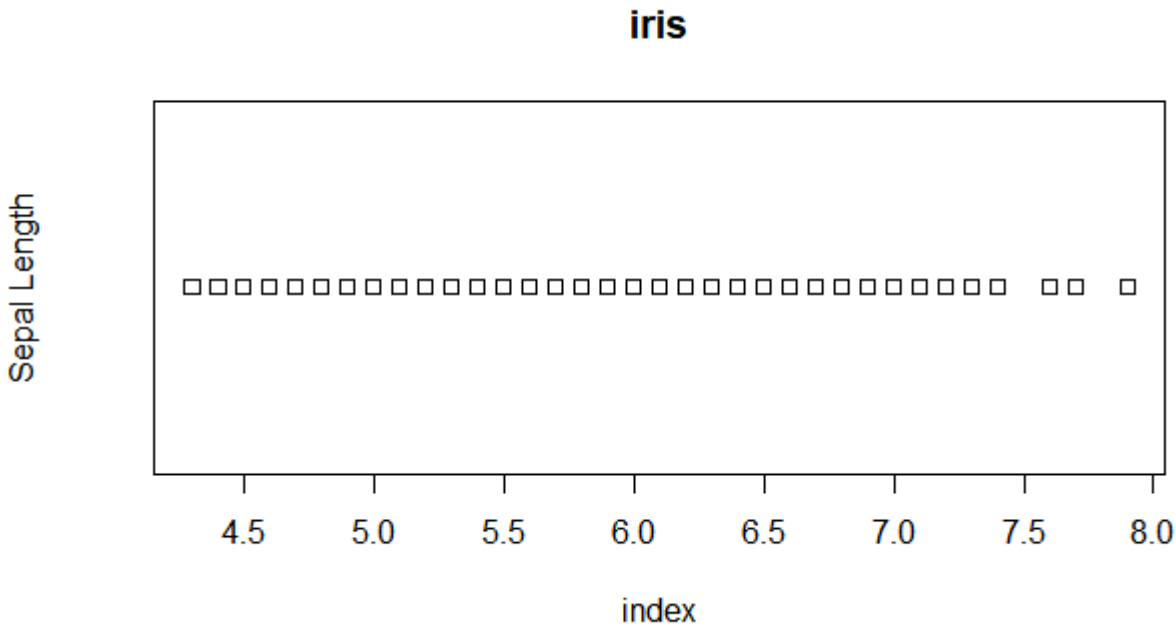
```
plot(1,1,xlim=c(1.5,0.5),ylim=c(0,7),type="n",ann="false")
text(1:5,rep(6,5),labels=c(0:4),cex=1:15,col=1:5)
points(1:5,rep(5,5),cex=1:5,col=1:5,pch=0:4)
text((1:5)+0.4,rep(5,5),cex=0.6,(0:4))
points((1:6)*0.8+0.2,rep(1,6),cex=2,pch=(20:25))
text((1:6)*0.8+0.5,rep(1,6),cex=0.6,(20:25))
```



#stripchart

stripchart(iris$Sepal.Length,xlab="index",ylab="Sepal Length")

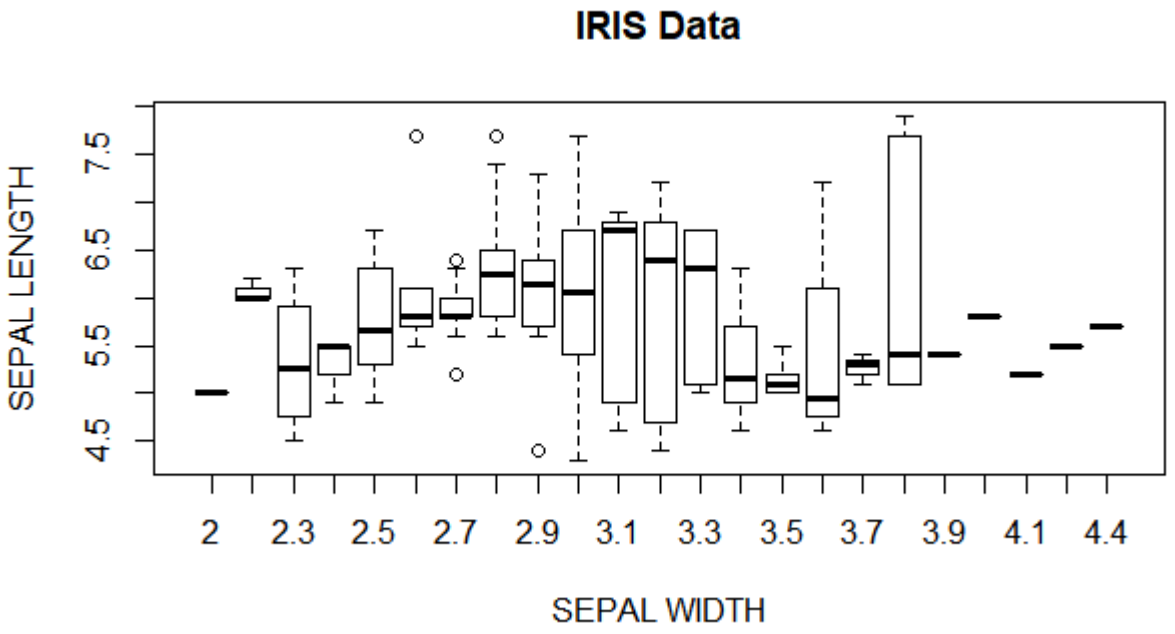
title("iris")



#boxplot

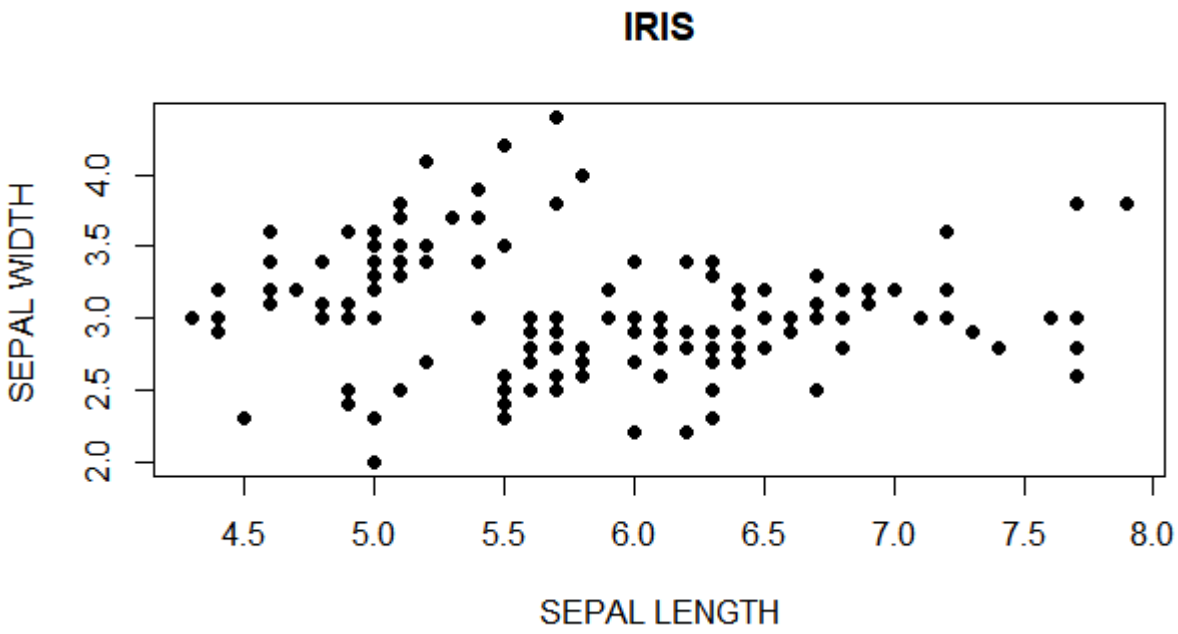
boxplot(iris$Sepal.Length~iris$Sepal.Width, main="IRIS Data",

xlab="SEPAL WIDTH", ylab="SEPAL LENGTH")



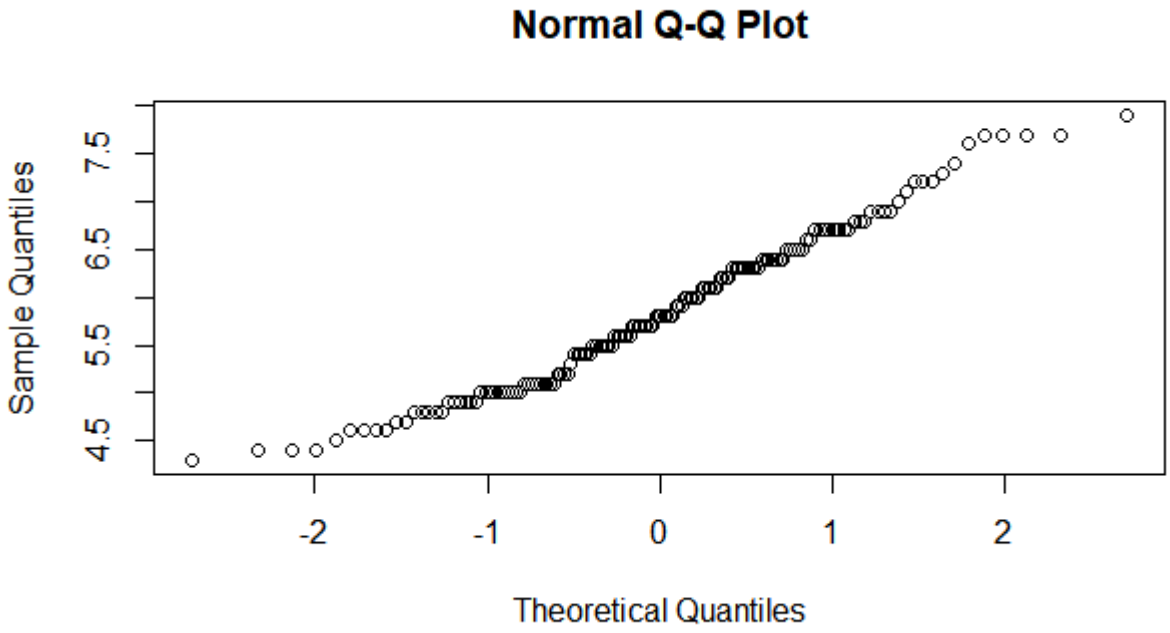
#SCATTEPLOT

```
attach(iris)
plot(Sepal.Length, Sepal.Width, main="IRIS",
    xlab="SEPAL LENGTH", ylab="SEPAL WIDTH ", pch=19)
```



#QQPLOT

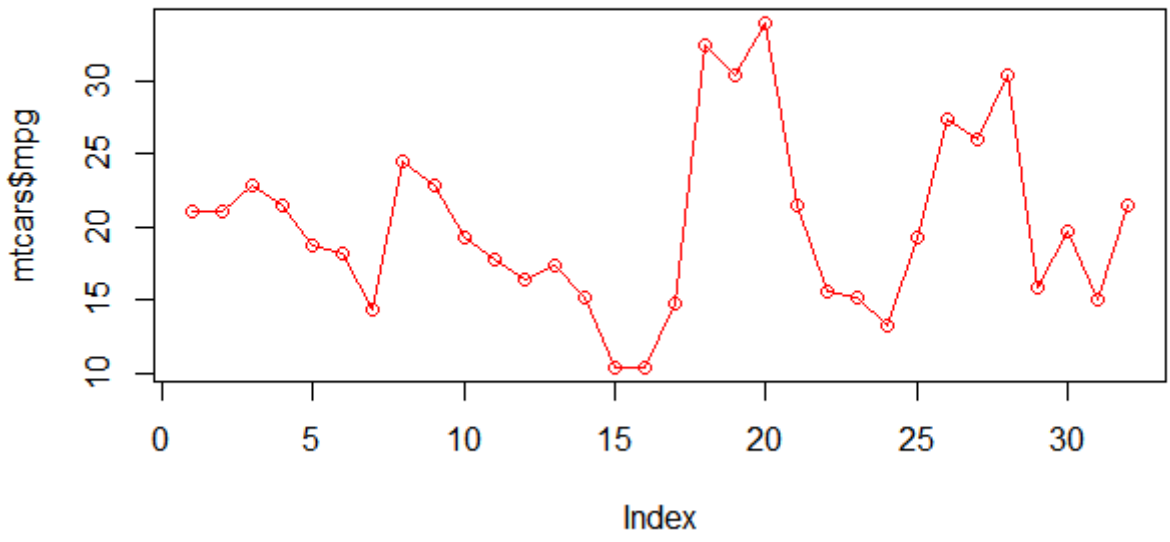
```
qqnorm(iris$Sepal.Length)
qqline(iris$Sepal.Width)
```



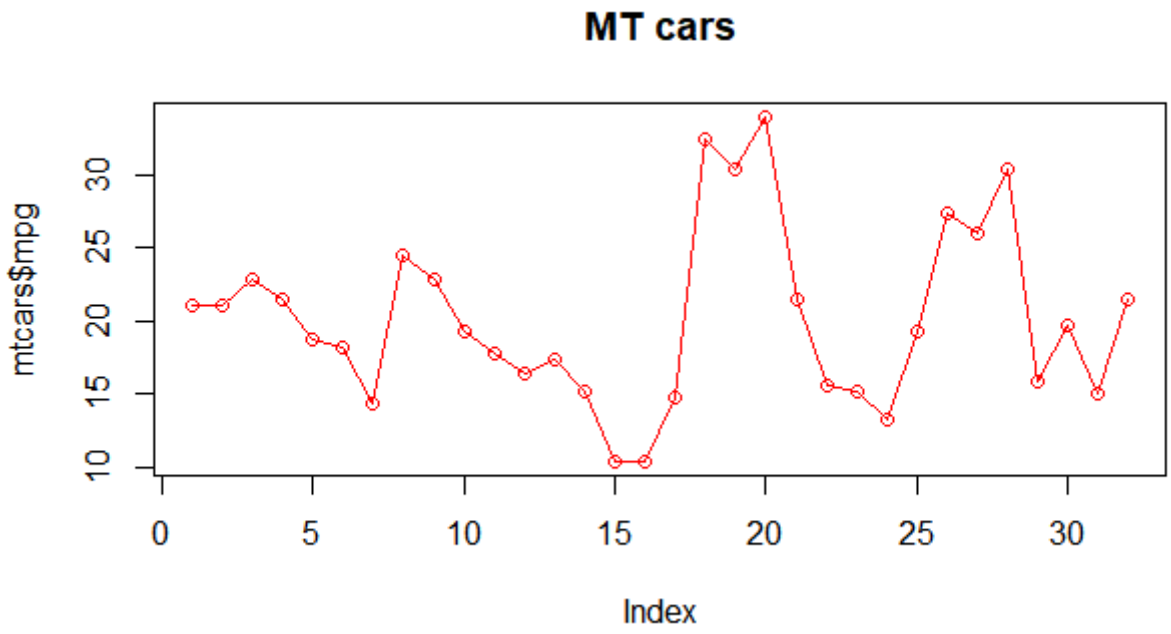
Dataset 3 - MTCARS

LINECHART

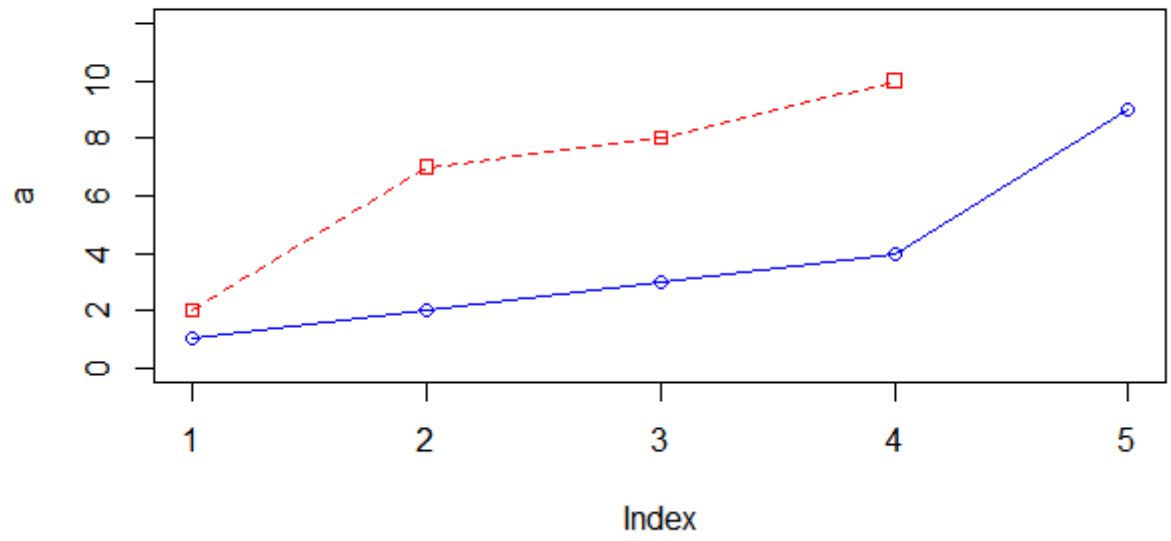
```
> plot(mtcars$mpg,type="o",col="red")
```



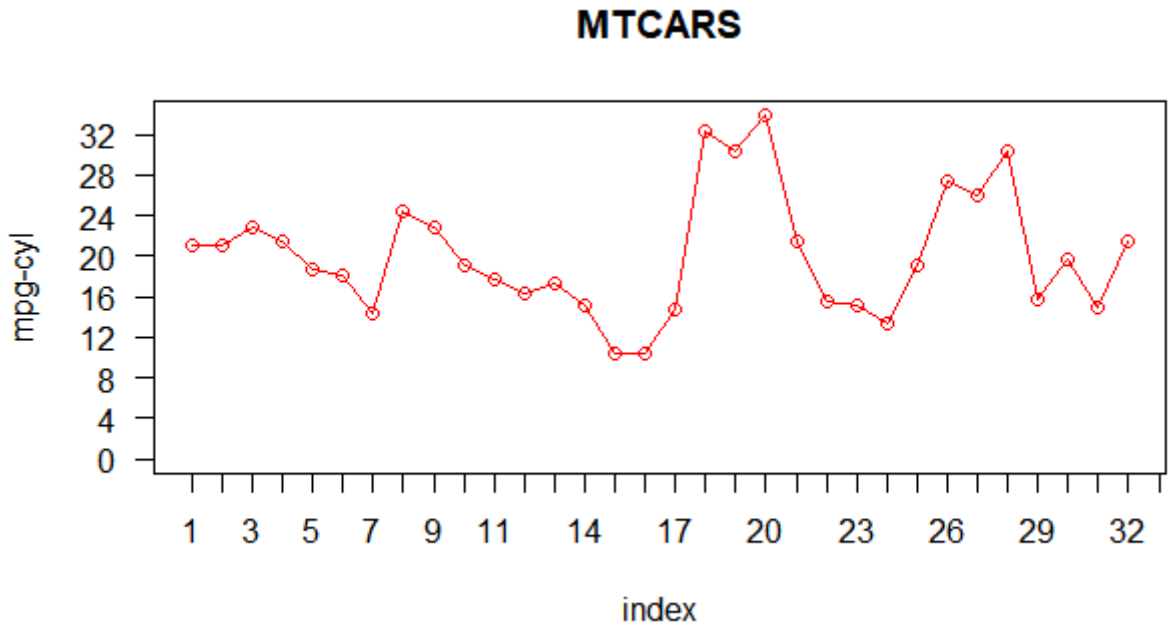
```
> head(mtcars)
                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
> plot(mtcars$mpg,type="o",col="red")
> title(main="MT cars")
```



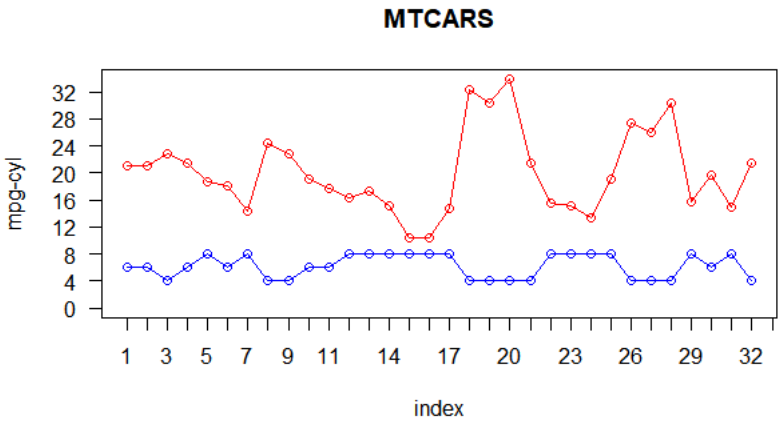

```
> plot(a,type="o",col="blue",ylim=c(0,12))
> lines(b,type="o",pch=22,lty=2,col="red")
```



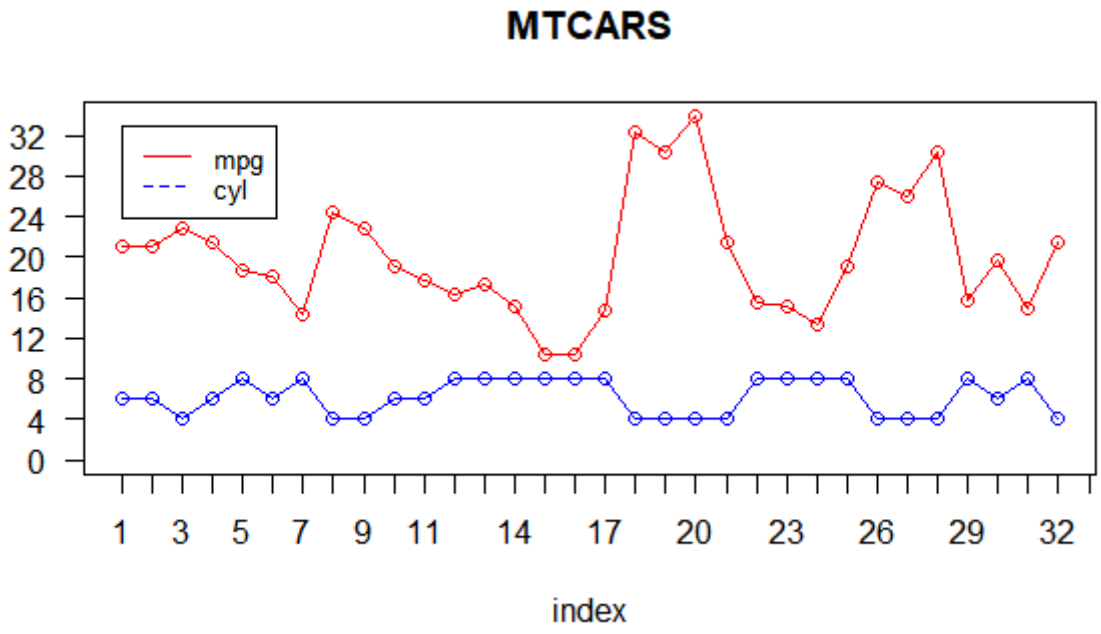

```
> g_range<-range(0,mtcars$mpg,mtcars$cyl)
> plot(mtcars$mpg,type="o",col="red",ylim=g_range,axes=FALSE,ann=FALSE)
> axis(1,at=1:50)
> axis(2,las=1,at=4*0:g_range[2])
> box()
> title(main="MTCARS",xlab="index",ylab="mpg-cyl")
```



```
> lines(mtcars$cyl,type="o",col="blue")
```

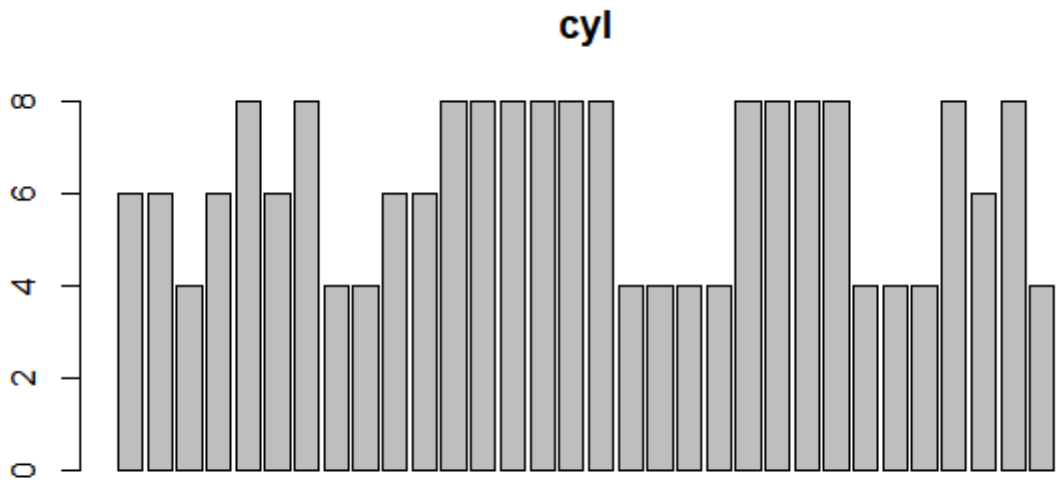



```
> plot(mtcars$mpg,type="o",col="red",ylim=g_range,axes=FALSE,ann=FALSE)
> axis(1,at=1:50)
> axis(2,las=1,at=4*0:g_range[2])
> box()
> title(main="MTCARS",xlab="index",ylab="")
> lines(mtcars$cyl,type="o",col="blue")
> lines(mtcars$disp,type="o",col="green")
> legend(1,33,legend=c("mpg","cyl"),col=c("red","blue"),lty=1:2,cex=0.8)
```

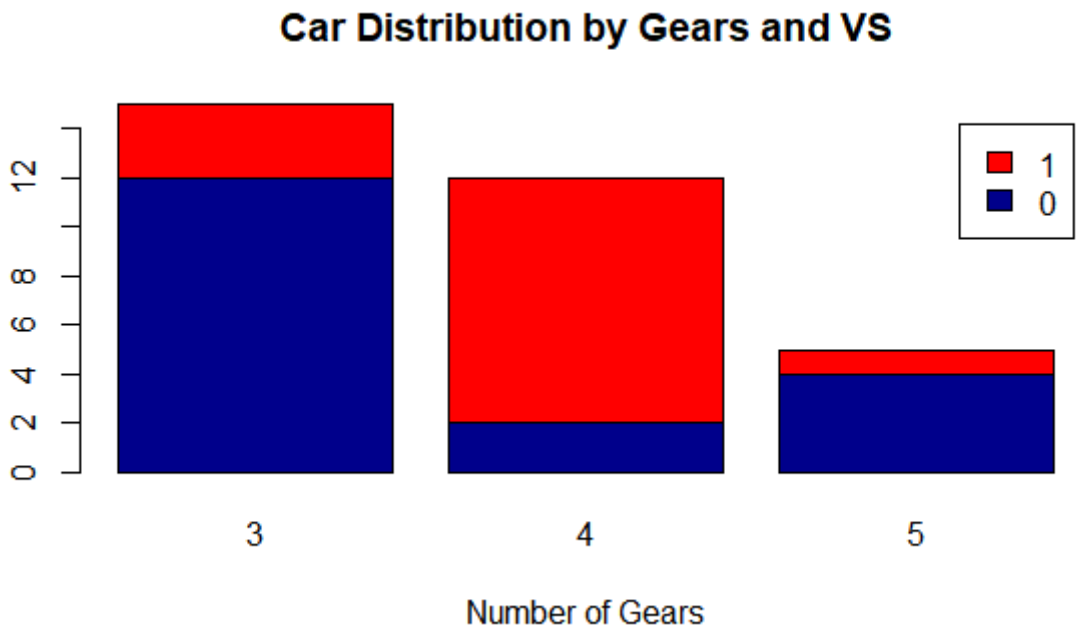


Barplot

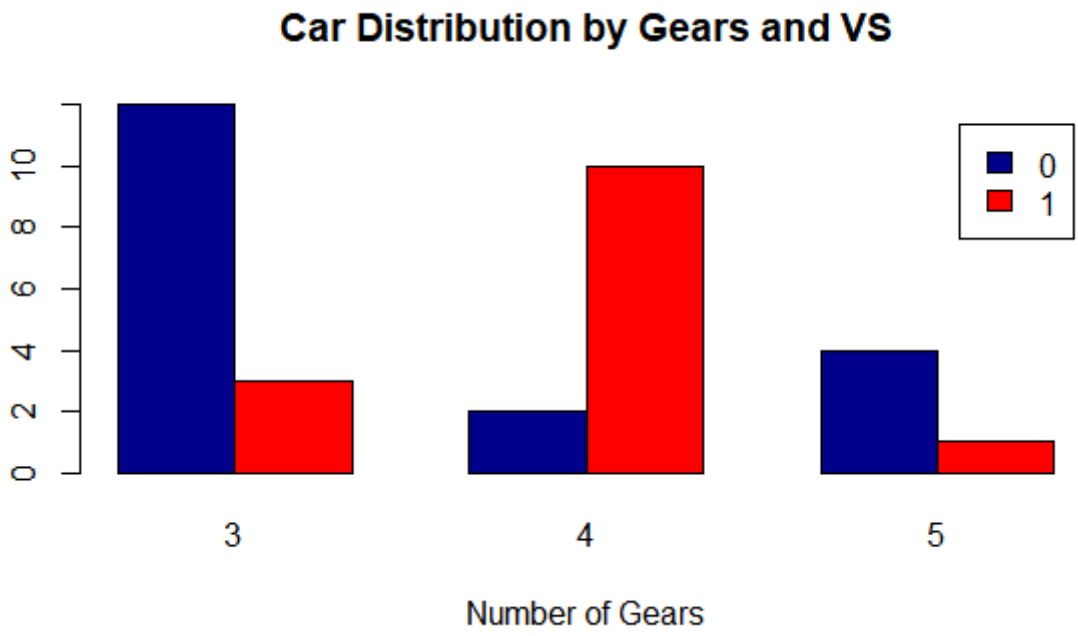
```
> barplot(mtcars$cyl,main="cyl")
```




```
> counts <- table(mtcars$vs, mtcars$gear)
> barplot(counts, main="Car Distribution by Gears and VS",xlab="Number of
Gears", col=c("darkblue","red"),legend = rownames(counts))
```

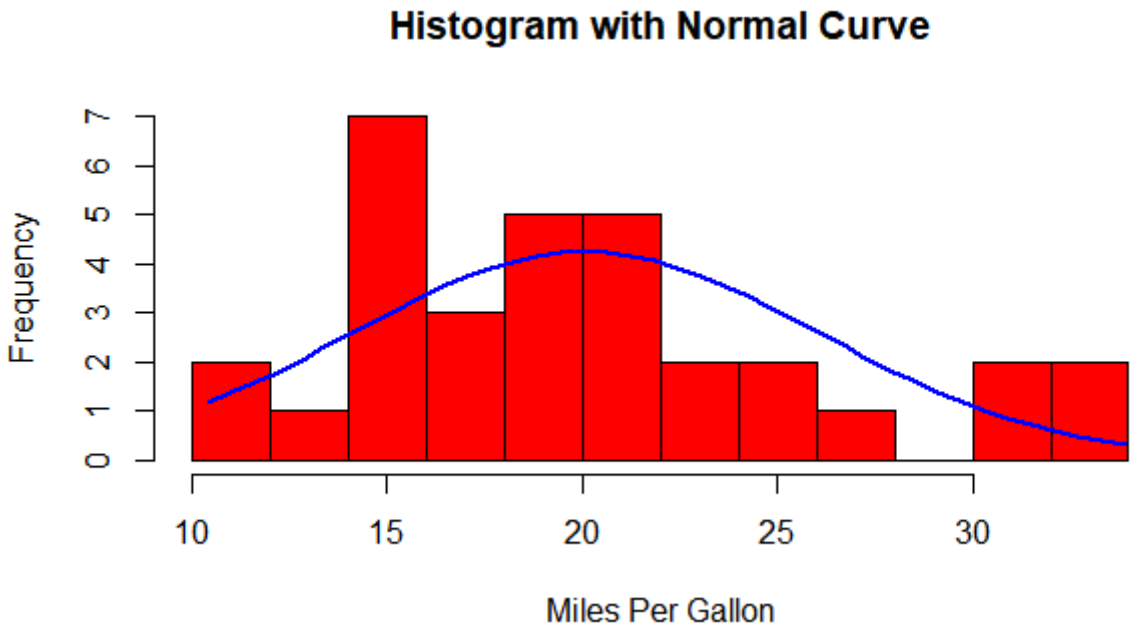



```
> counts <- table(mtcars$vs, mtcars$gear)
> barplot(counts, main="Car Distribution by Gears and VS",xlab="Number of
Gears", col=c("darkblue","red"),legend = rownames(counts), beside=TRUE)
```

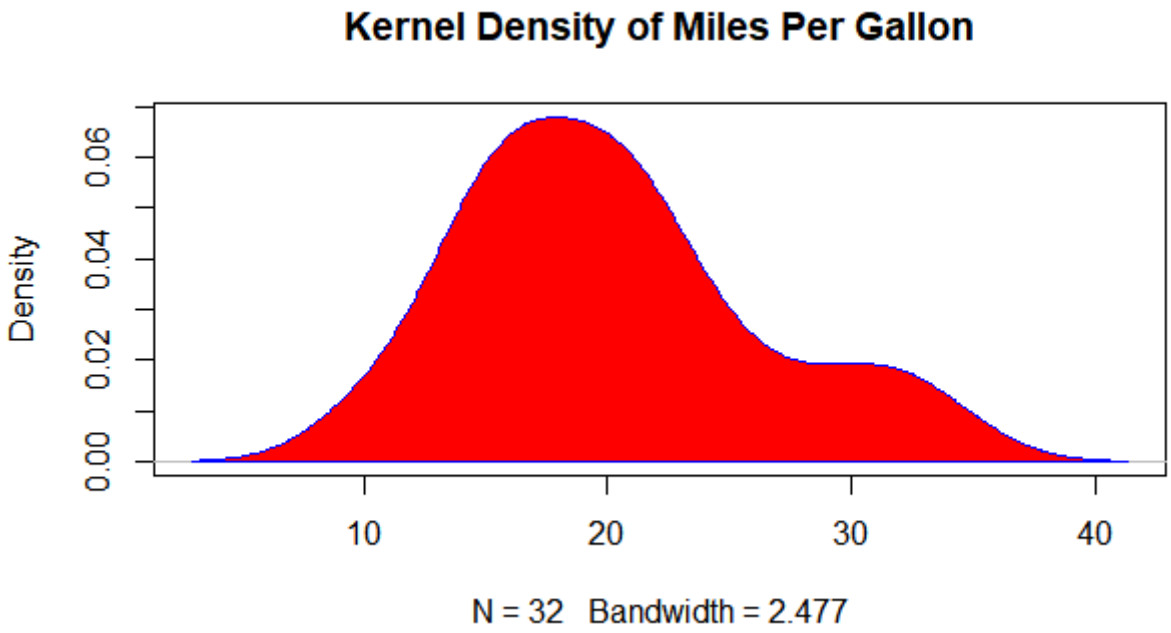


Histogram

```
x <- mtcars$mpg
> h<-hist(x, breaks=10, col="red", xlab="Miles Per Gallon",
main="Histogram with Normal Curve")
> xfit<-seq(min(x),max(x),length=40)
> yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
> yfit <- yfit*diff(h$mids[1:2])*length(x)
> lines(xfit, yfit, col="blue", lwd=2)
```

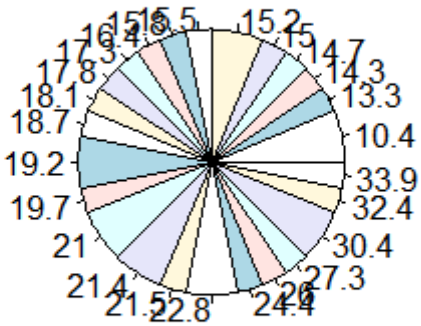


```
> d <- density(mtcars$mpg)
> plot(d, main="Kernel Density of Miles Per Gallon")
> polygon(d, col="red", border="blue")
```



Pie Chart

pie(table(mtcars$mpg))

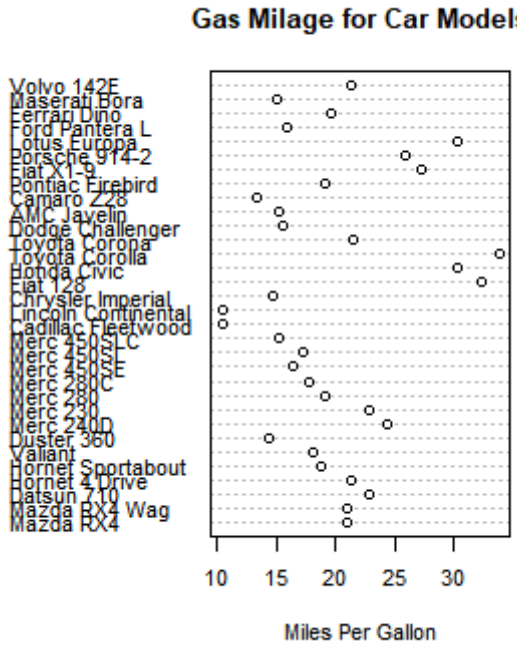



```
> cars<-c(1,2,3,4,5)
```

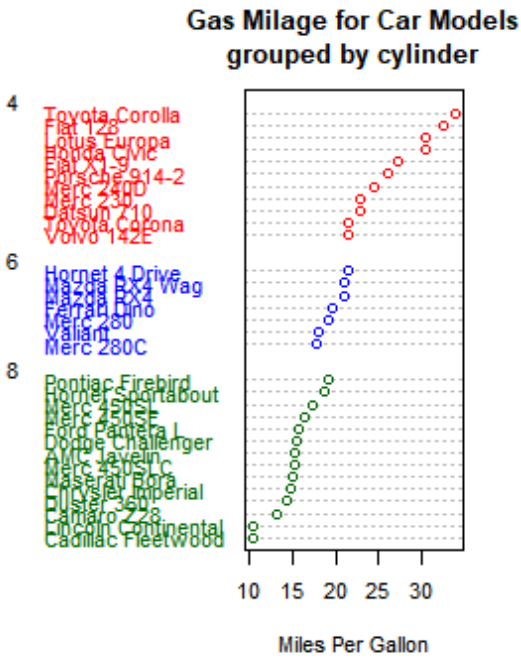
```
> pie(cars,main="cars",col=rainbow(length(cars)))
```

Dot Chart

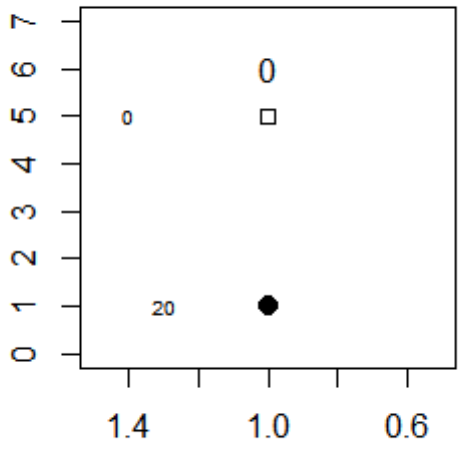
```
>
dotchart(mtcars$mpg,labels=row.names(mtcars),cex
=.7,main="Gas Milage for Car Models",
+          xlab="Miles Per Gallon")


>
```




```
> x <- mtcars[order(mtcars$mpg),] #
sort by mpg
> x$cyl <- factor(x$cyl) # it must be a
factor
> x$color[x$cyl==4] <- "red"
> x$color[x$cyl==6] <- "blue"
> x$color[x$cyl==8] <- "darkgreen"
>
dotchart(x$mpg,labels=row.names(x),cex=
.7,groups= x$cyl,
+           main="Gas Milage for Car
Models\ngrouped by cylinder",
+           xlab="Miles Per Gallon",
gcolor="black", color=x$color)


>
```

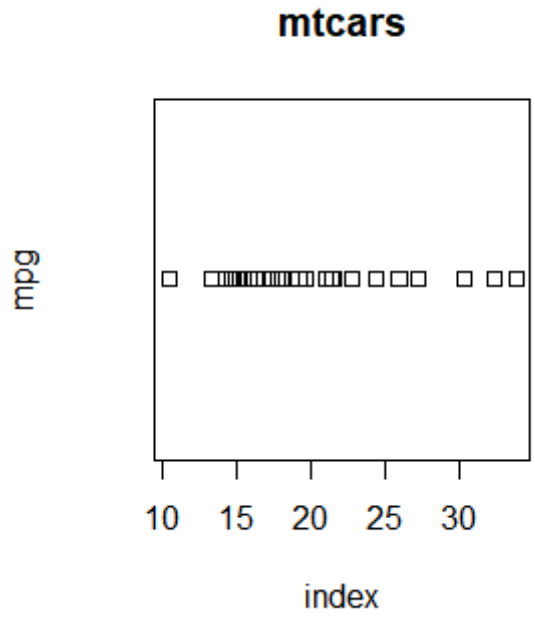



```
>
plot(1,1,xlim=c(1.5,0.5),ylim=c(0,7),ty
pe="n",ann="false")
>
text(1:5,rep(6,5),labels=c(0:4),cex=1:1
5,col=1:5)
>
points(1:5,rep(5,5),cex=1:5,col=1:5,pch
=0:4)
> text((1:5)+0.4,rep(5,5),cex=0.6,
(0:4))
>
points((1:6)*0.8+0.2,rep(1,6),cex=2,pch
=(20:25))
> text((1:6)*0.8+0.5,rep(1,6),cex=0.6,
(20:25)
+ text((1:6)*0.8+0.5,rep(1,6),cex=0.6,
(20:25)
Error: unexpected symbol in:
"text((1:6)*0.8+0.5,rep(1,6),cex=0.6,
(20:25)
text"
> )
Error: unexpected ')' in ")"
> text((1:6)*0.8+0.5,rep(1,6),cex=0.6,
(20:25))


>
```



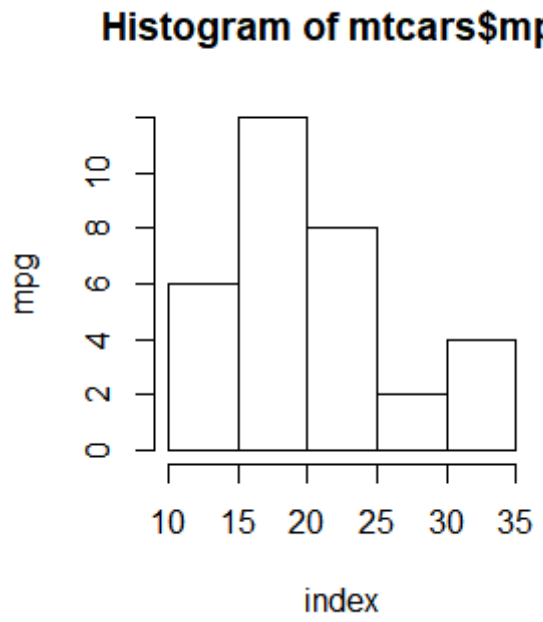
STRIPCHARTS

```
>
stripchart(mtcars$mpg,xlab="index",ylab=
"mpg")
> title("mtcars")
```



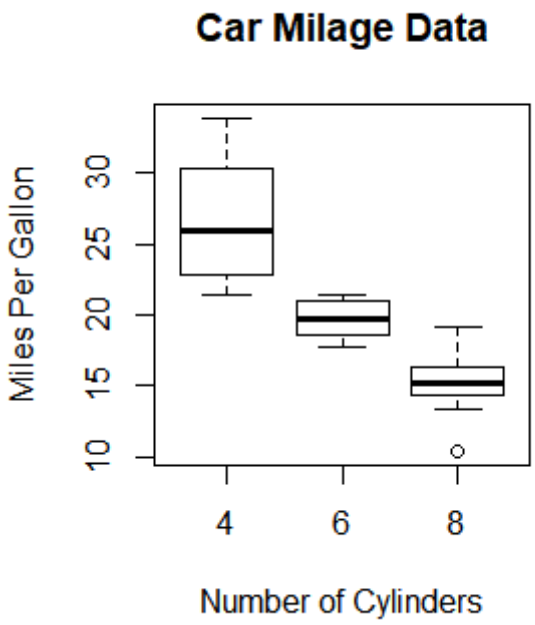
HISTOGRAM

```
hist(mtcars$mpg,xlab="index",ylab
="mpg")
```



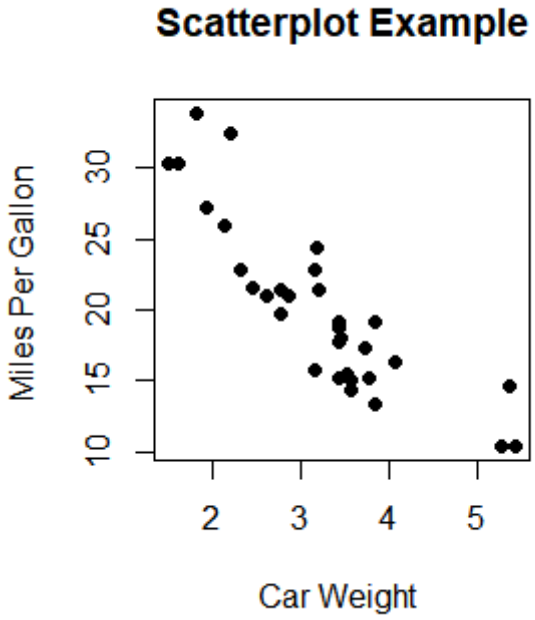
BOXPLOT

```
> boxplot(mpg~cyl,data=mtcars, main="Car
Milage Data",
+         xlab="Number of Cylinders",
ylab="Miles Per Gallon")


>
```



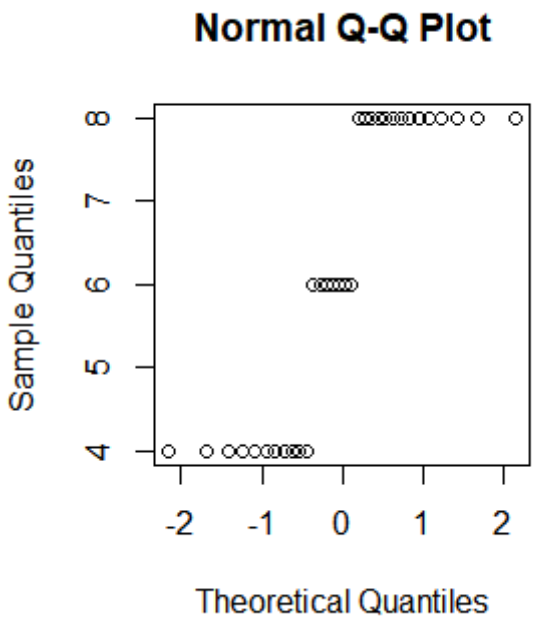
SCATTERPLOT

```
> plot(wt, mpg, main="Scatterplot Example",
+      xlab="Car Weight ", ylab="Miles Per Gallon ", pch=19)
```



Qqnorm

```
> qqnorm(mtcars$cyl)
```




```
> qqline(mtcars$mpg)
```