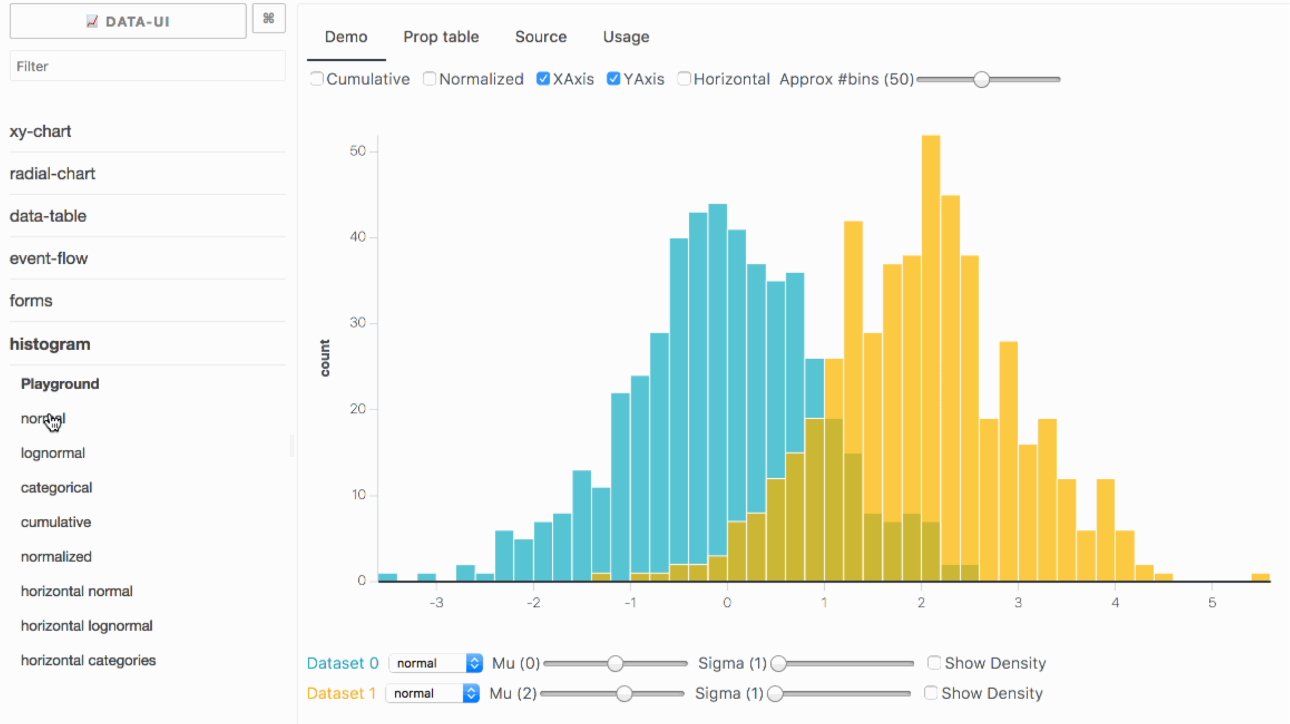


☐ Cumulative ☐ Normalized ☑ XAxis ☑ YAxis ☐ Horizontal Approx #bins (50)

Dataset 0 normal Mu (0) Sigma (1) ☐ Show Density

Dataset 1 normal Mu (2) Sigma (1) ☐ Show Density

EXPLORATORY DATA ANALYSIS

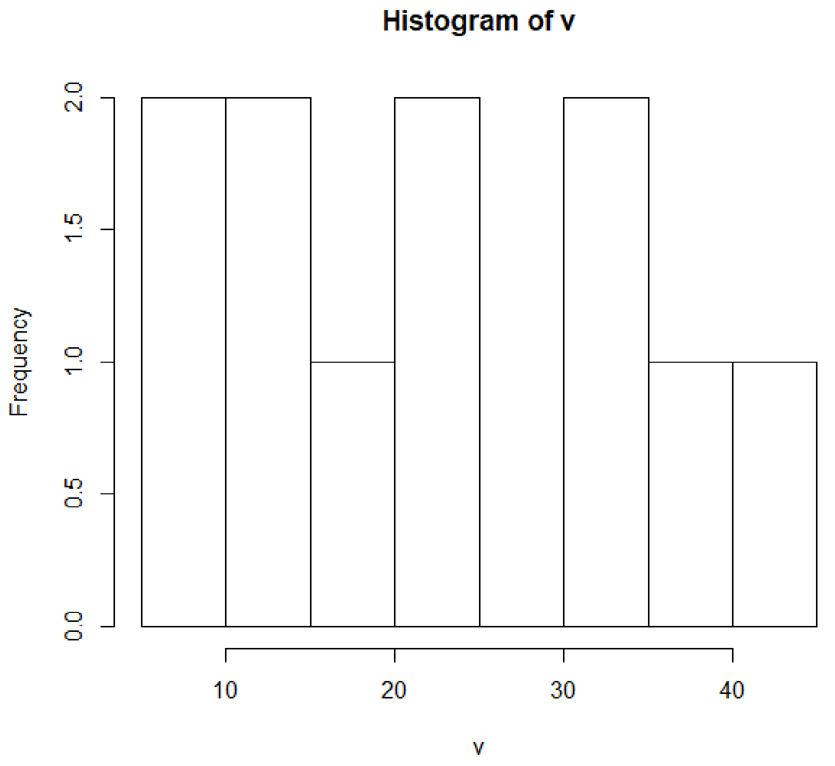
HISTOGRAM

- R creates histogram using hist() function.
- This function takes a vector as an input and uses some more parameters to plot histograms

Ex: Simple Histogram

x <- c(9,13,21,8,36,22,12,41,31,33,19)

To draw a simple histogram, the function is hist(x)



Note: If you don’t specify labels explicitly, it considers them by it’s own

HISTOGRAM

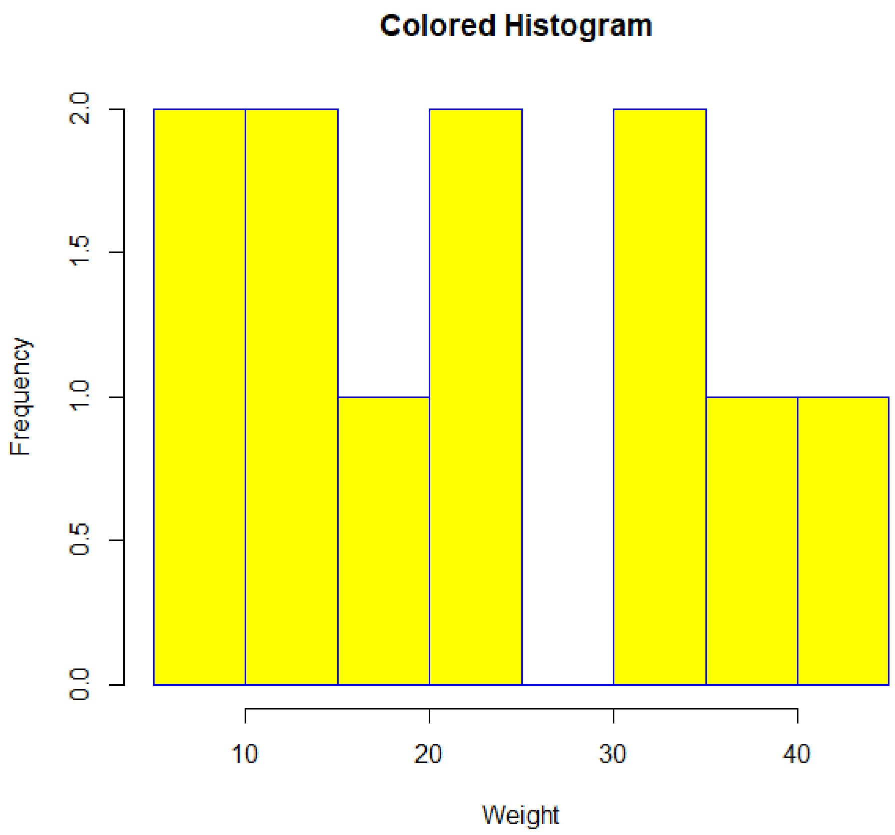
Ex2: Histogram with Parameters

```
x <-  c(9,13,21,8,36,22,12,41,31,33,19)
```

To draw a simple histogram, having

Label of x-axis = "Weight", color of bars = "yellow", border color of bars = "blue“, then function is

```
hist(x,
        xlab = "Weight",
        col = "yellow",
        border = "blue “,
         main = “Colored Histogram ")
```



HISTOGRAM

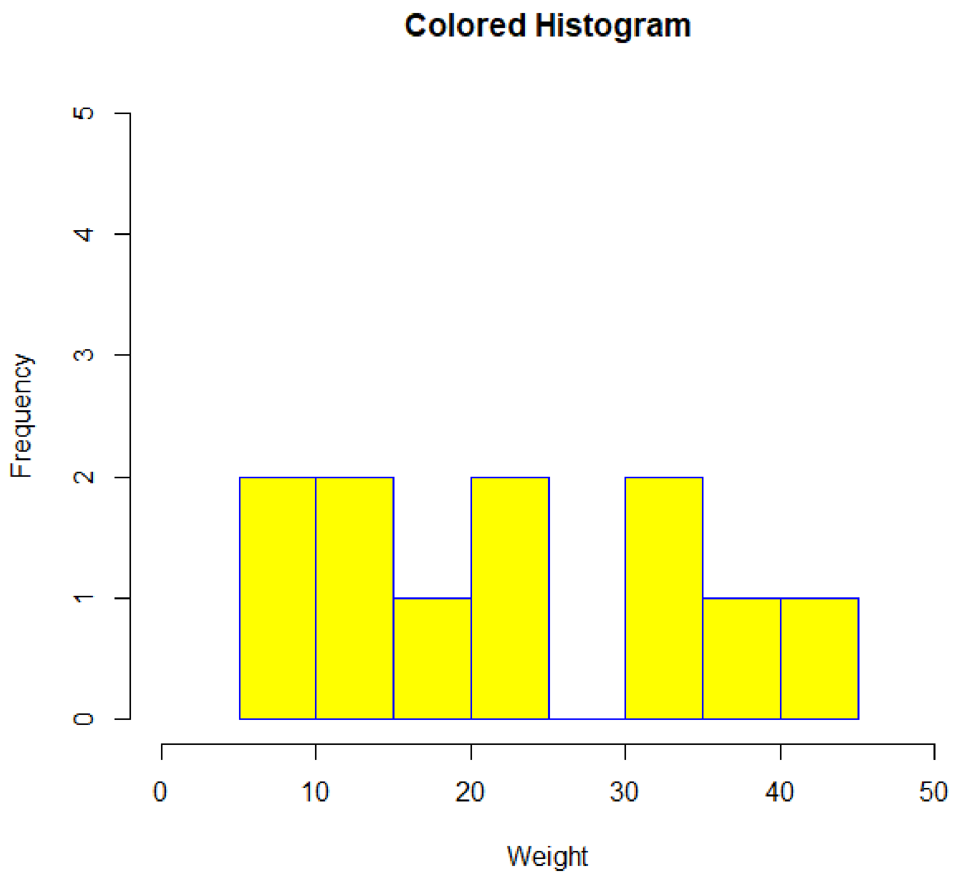
Ex2: Histogram with Parameters

x <- c(9,13,21,8,36,22,12,41,31,33,19)

To draw a simple histogram, having

Label of x-axis = "Weight", color of bars = "yellow", border color of bars = "blue“, limits of x-axis are 0 to 50, limits of y-axis are 0 to 5 then function is

```
hist(x,
        xlab = "Weight",
        col = "yellow",
        border = "blue “,
         main = “Colored Histogram “,
        xlim = c(0,50),
        ylim = c(0,5))
```



BAR CHART/GRAPH

- A bar chart represents data in rectangular bars
- length of the bar proportional to the value of the variable.
- R uses the function barplot() to create bar charts.
- R can draw both vertical and Horizontal bars in the bar chart.
- In bar chart each of the bars can be given different colors.

BAR CHART/GRAPH

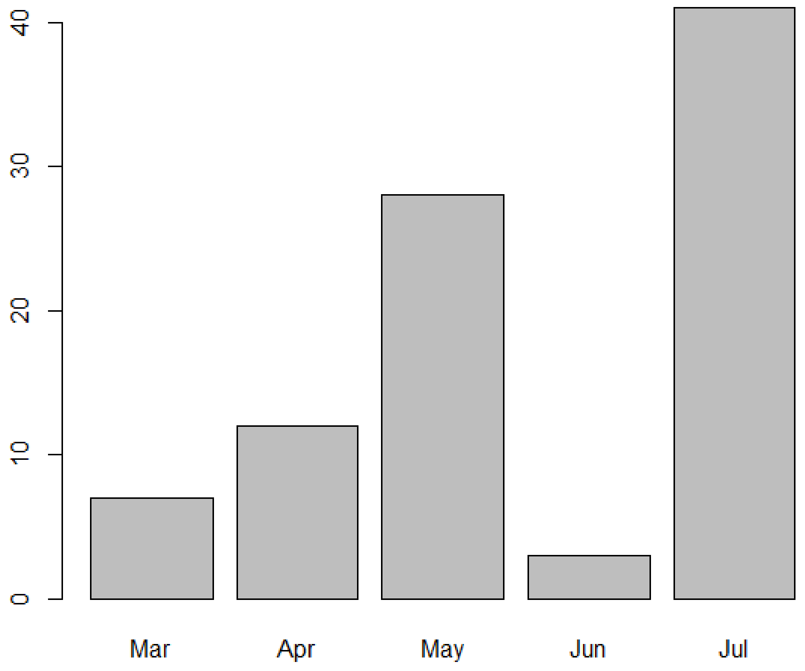
Ex: Simple Bar graph

x <- c("Mar","Apr","May","Jun","Jul")

y <- c(7,12,28,3,41)

To draw a simple barplot

barplot(names.arg = x, y)



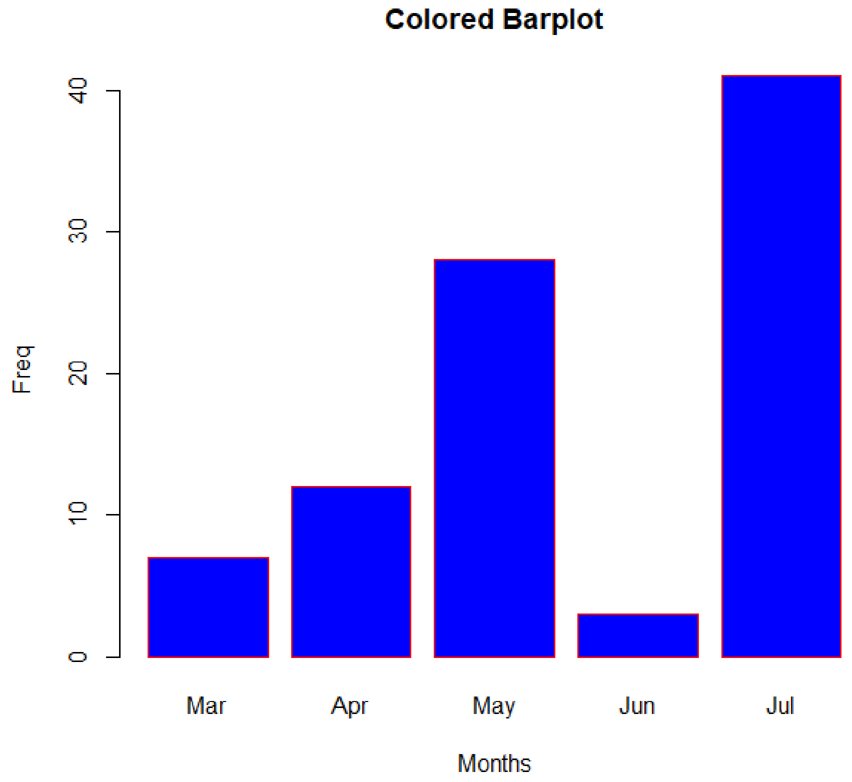
BAR CHART/GRAPH

Ex2: Bar graph with parameters

```
x <- c("Mar","Apr","May","Jun","Jul")
y <- c(7,12,28,3,41)
```

To draw a simple barplot

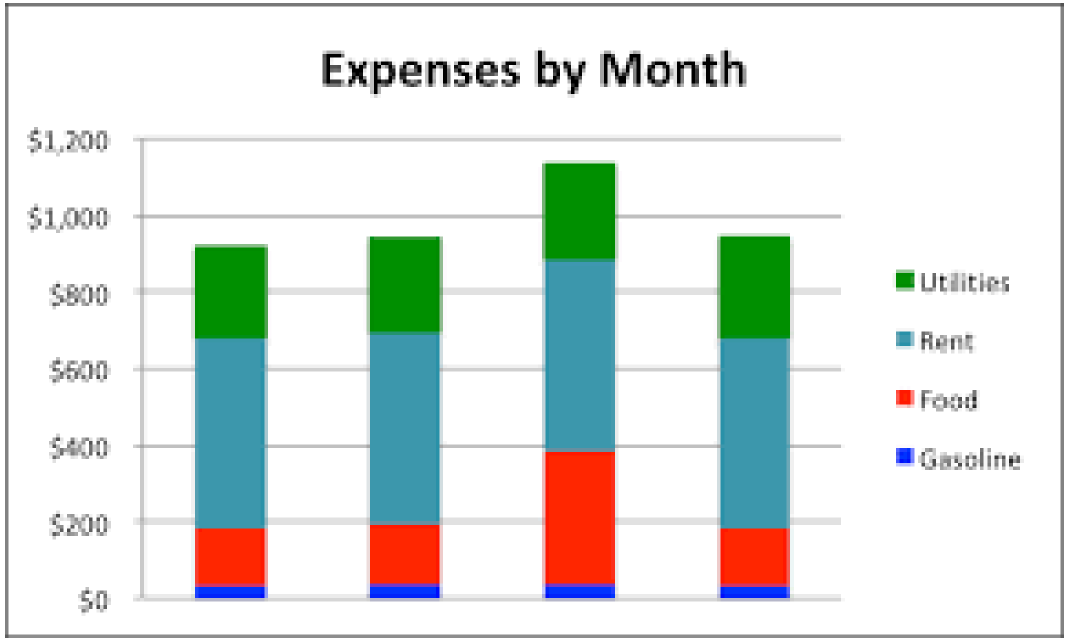
```
        barplot(names.arg = x,
        y,
        main = "Colored Barplot",
        xlab = "Months",
        ylab = "Freq",
        col = "blue",
        border = "red")
```



GROUP BAR CHART AND STACKED BAR CHART

- We can create bar chart with groups of bars and stacks in each bar by using a matrix as input values.

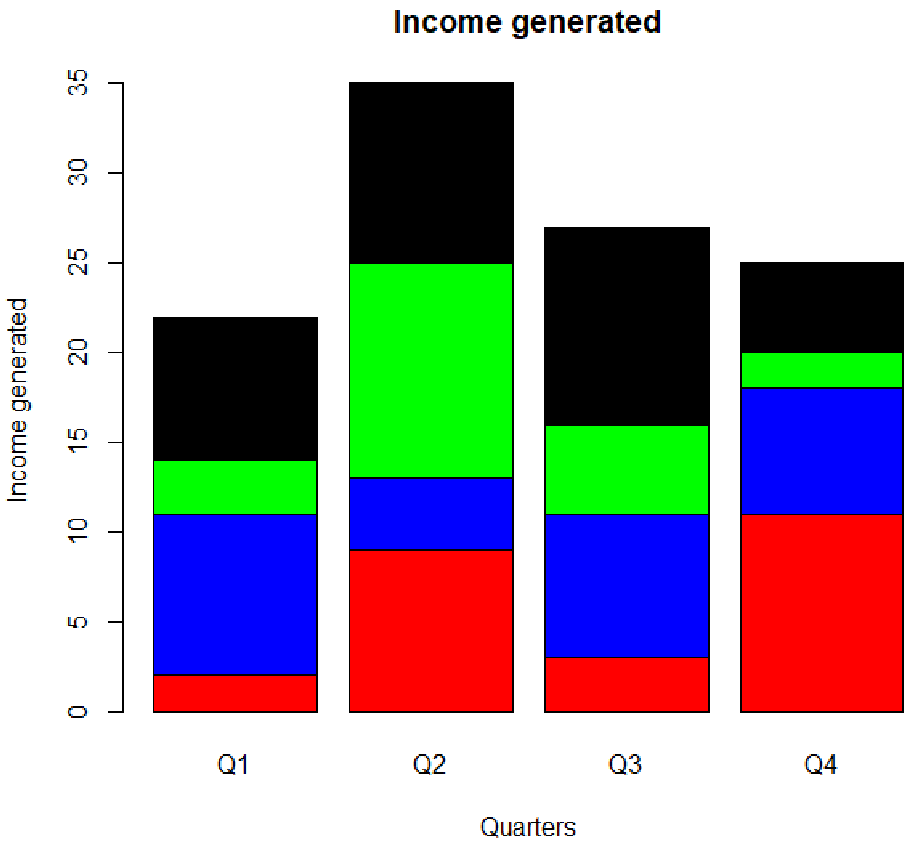
- More than two variables are represented as a matrix which is used to create the group bar chart and stacked bar chart.



GROUP BAR CHART AND STACKED BAR CHART

- Ex: Draw a stacked bar chart with the following data

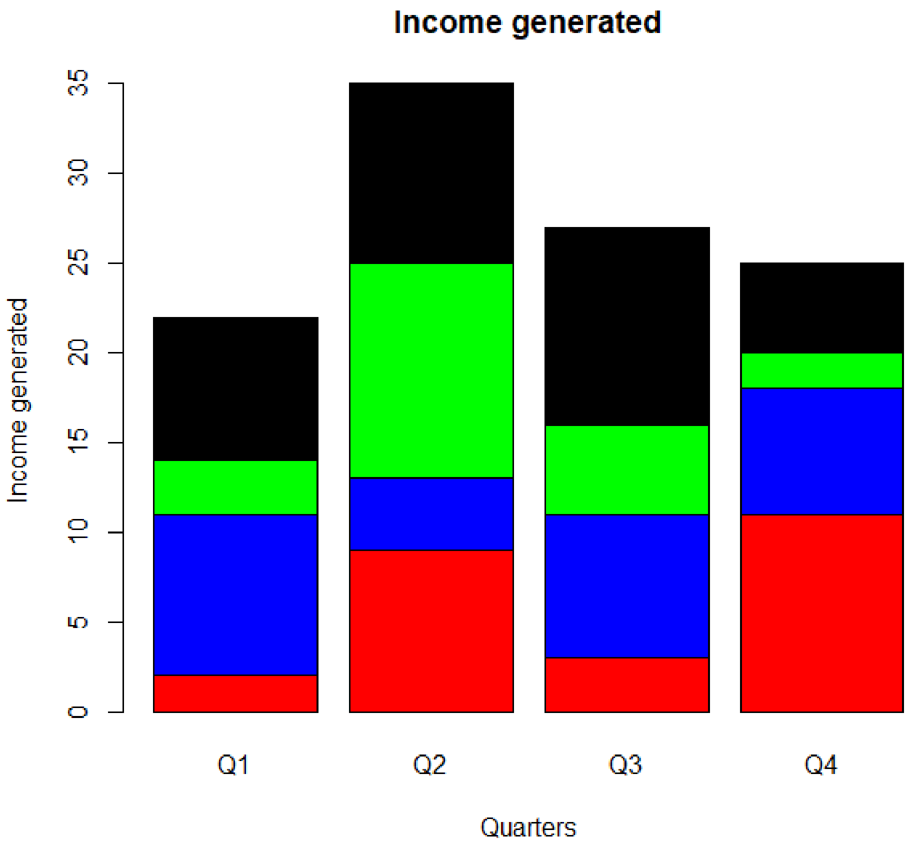
	Q1	Q2	Q3	Q4
state1	2	9	3	11
state2	9	4	8	7
state3	3	12	5	2
state4	8	10	11	5



GROUP BAR CHART AND STACKED BAR CHART

	Q1	Q2	Q3	Q4
state1	2	9	3	11
state2	9	4	8	7
state3	3	12	5	2
state4	8	10	11	5

```
> states <- c("state1", "state2", "state3", "state4")
> colors <- c("red","blue","green","black")
> quarters <- c("Q1","Q2","Q3","Q4")
>  Values <- matrix(c(2, 9, 3, 11, 9, 4, 8, 7, 3, 12, 5, 2, 8, 10, 11, 5), nrow = 4, ncol = 4, byrow = TRUE)
> barplot(Values, main = "Income generated", names.arg = quarters, xlab = "Quarters", ylab = "Income generated", col = colors);
```



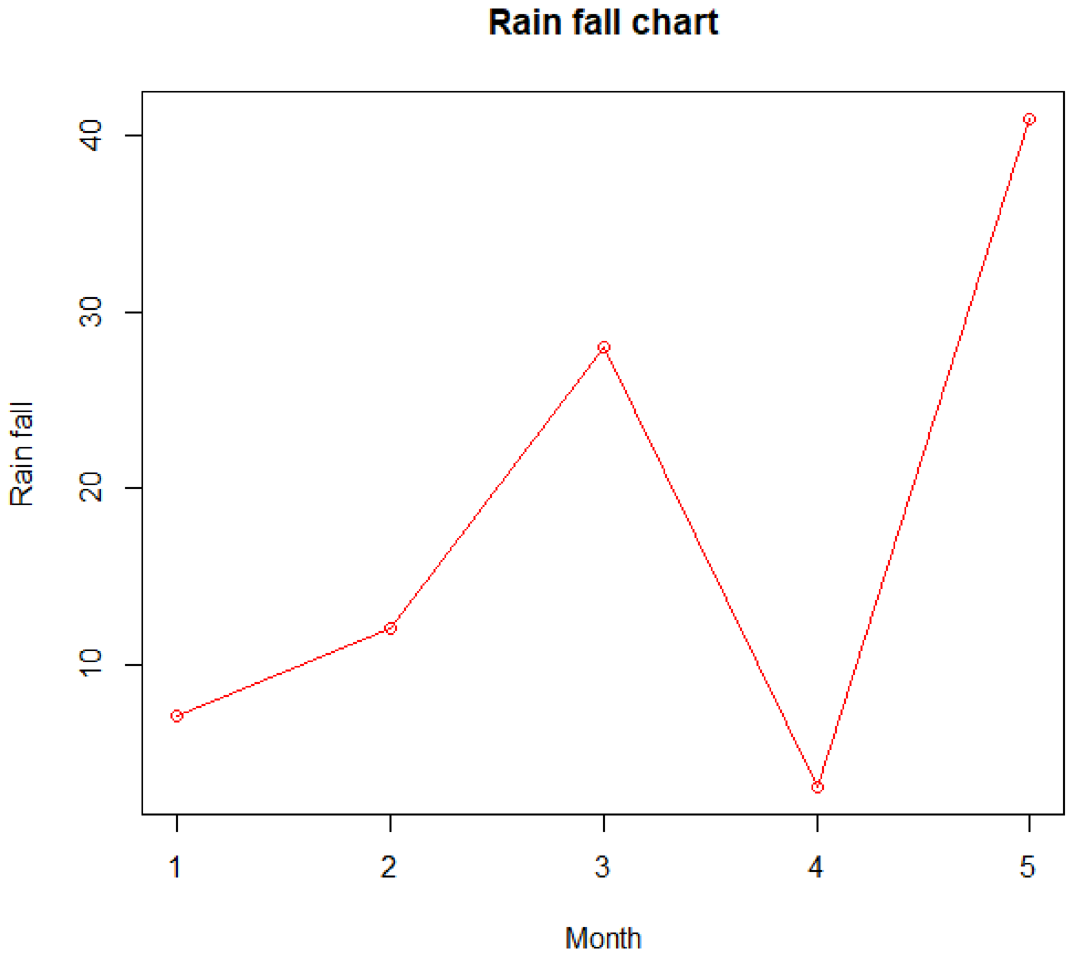
LINE GRAPH

- a graph that connects a series of points by drawing line segments between them.
- These points are ordered in one of their coordinate (usually the x-coordinate) value.
- Line charts are usually used in identifying the trends in data.
- The plot() function in R is used to create the line graph.

LINE GRAPH

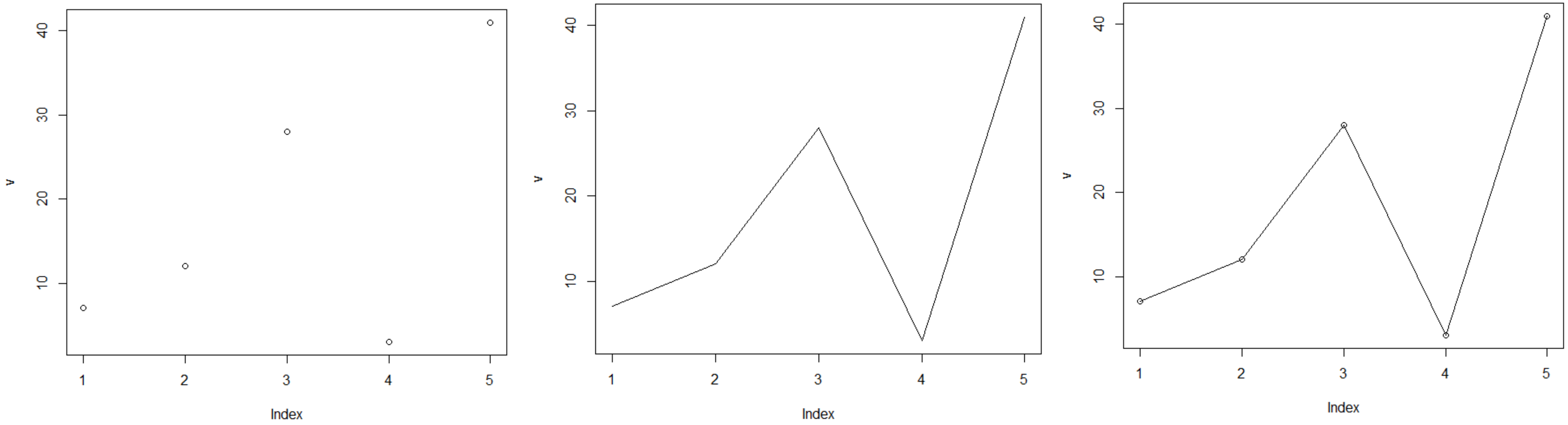
- Ex: Draw a line graph single line

```
x <- c(7,12,28,3,41)
plot(x,
     type = "o",
     col = "red",
     xlab = "Month",
     ylab = "Rain fall",
     main = "Rain fall chart")
```



LINE GRAPH

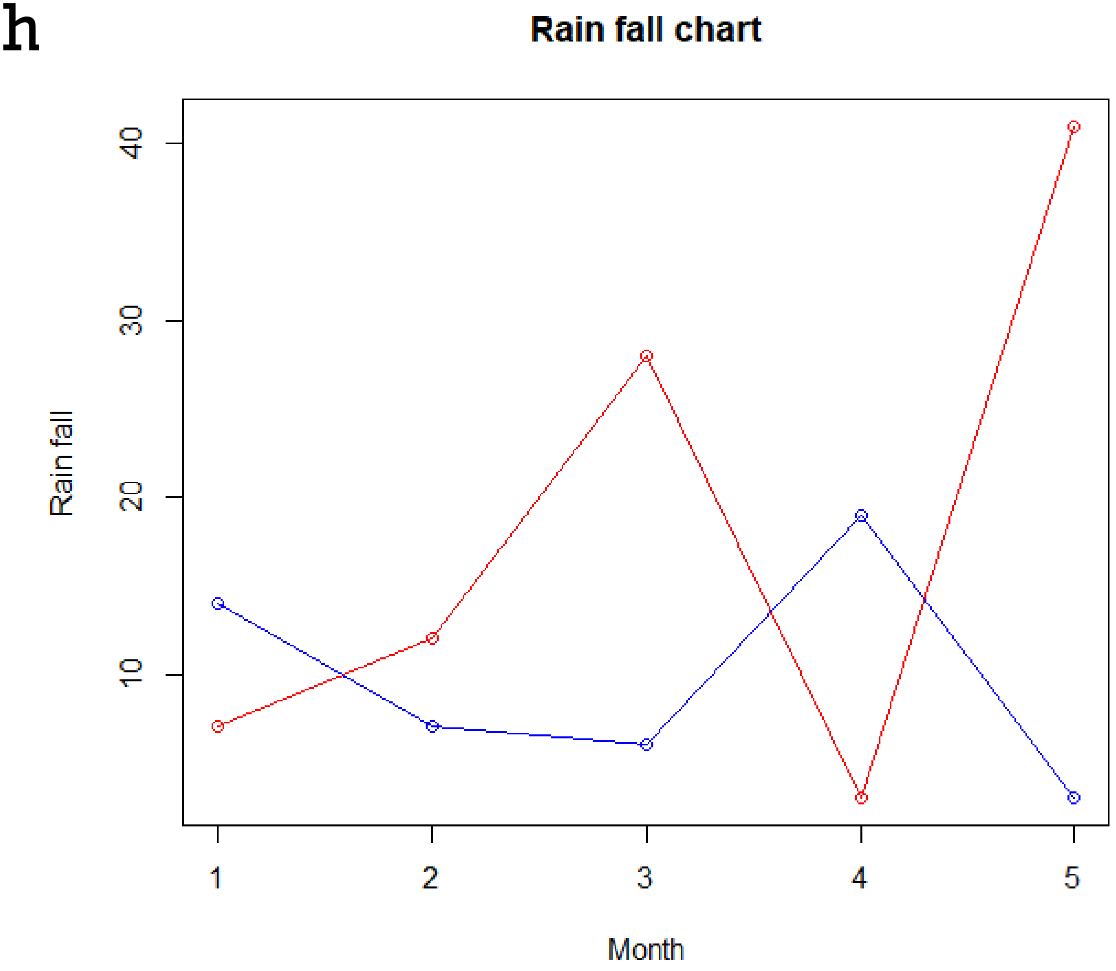
type takes the value "p" to draw only the points, "l" to draw only the lines and "o" to draw both points and lines



LINE GRAPH

- Ex: Draw a line graph with multiple lines

```
x1 <- c(7,12,28,3,41)
x2 <- c(14,7,6,19,3)
plot(x1,
     type = "o",
     col = "red",
     xlab = "Month",
     ylab = "Rain fall",
     main = "Rain fall chart")

lines(x2, type = "o", col = "blue")
```



BOX PLOT

- a measure of how well distributed is the data in a data set.

- It divides the data set into three quartiles.

- This graph represents the minimum, maximum, median, first quartile and third quartile in the data set.

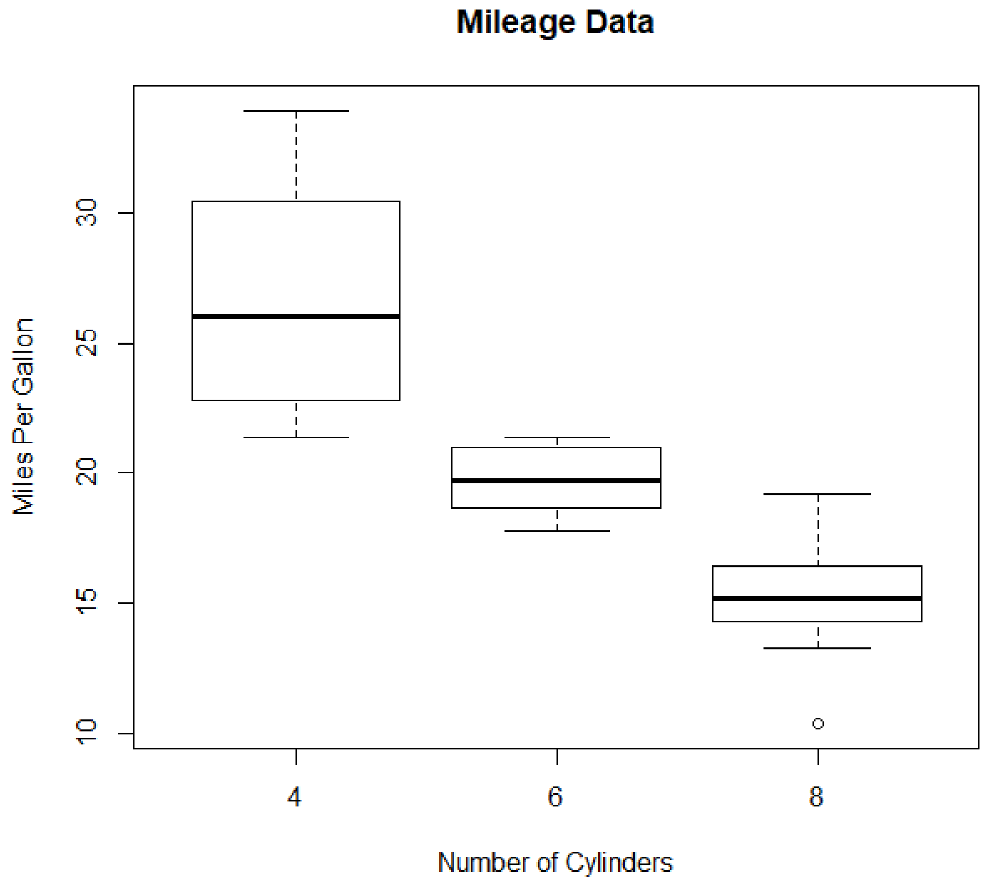
- It is also useful in comparing the distribution of data across data sets by drawing boxplots for each of them.

- Boxplots are created in R by using the boxplot() function.

BOX PLOT

- Ex: Simple Box plot

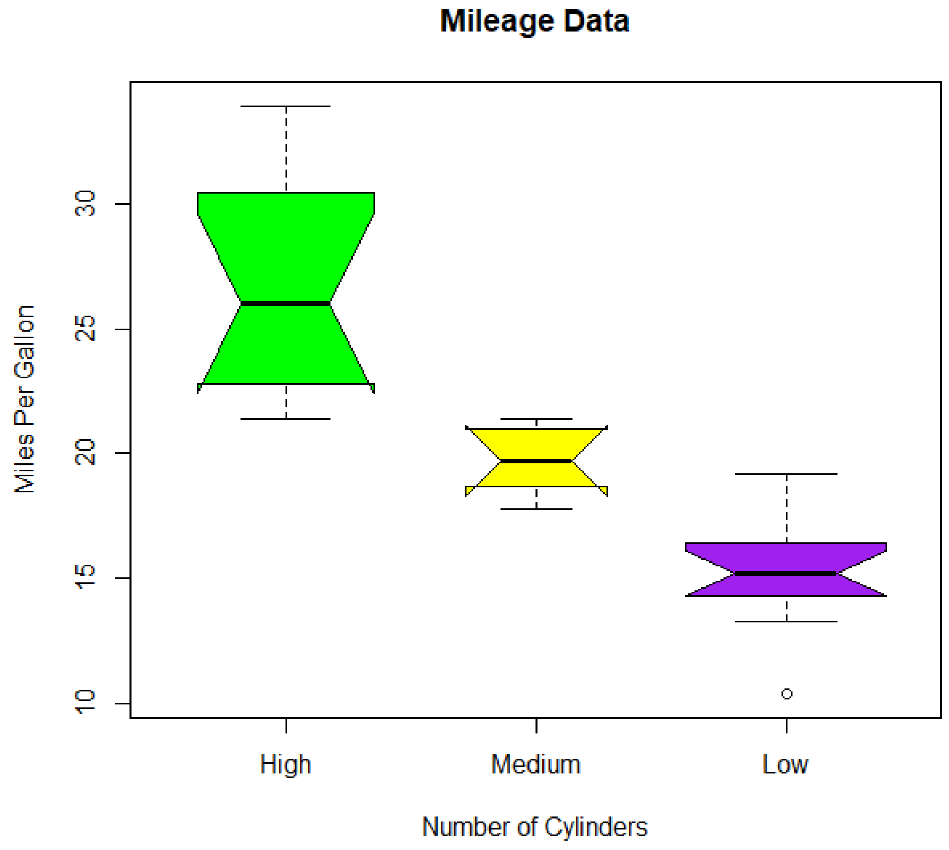
```
boxplot(mpg ~ cyl,
        data = mtcars,
        xlab = "Number of Cylinders",
        ylab = "Miles Per Gallon",
        main = "Mileage Data")
```



BOX PLOT

- Ex: Box plot with parametrs

```
boxplot(mpg ~ cyl, data = mtcars,
   xlab = "Number of Cylinders",
   ylab = "Miles Per Gallon",
   main = "Mileage Data",
   notch = TRUE,
   varwidth = TRUE,
   col = c("green", "yellow", "purple"),
   names = c("High", "Medium", "Low")
)
```



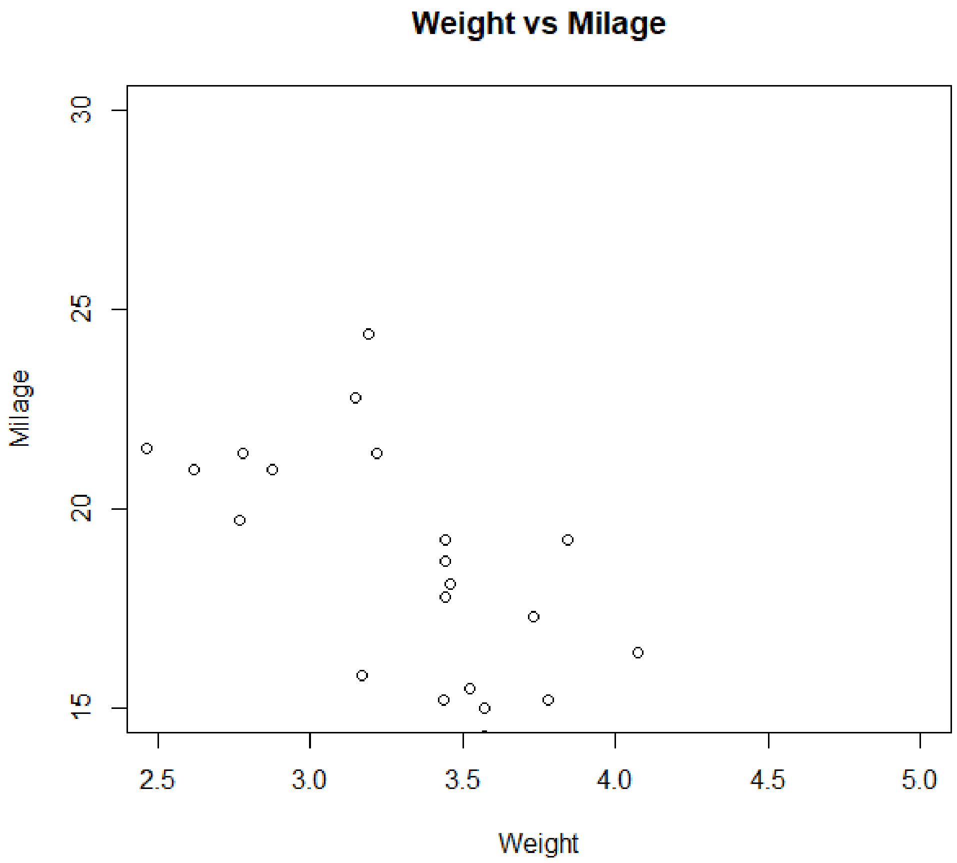
SCATTER PLOT

- Scatterplots show many points plotted in the Cartesian plane.
- Each point represents the values of two variables.
- One variable is chosen in the horizontal axis and another in the vertical axis.
- The simple scatterplot is created using the plot() function.

SCATTER PLOT

- Ex: Simple scattered plot

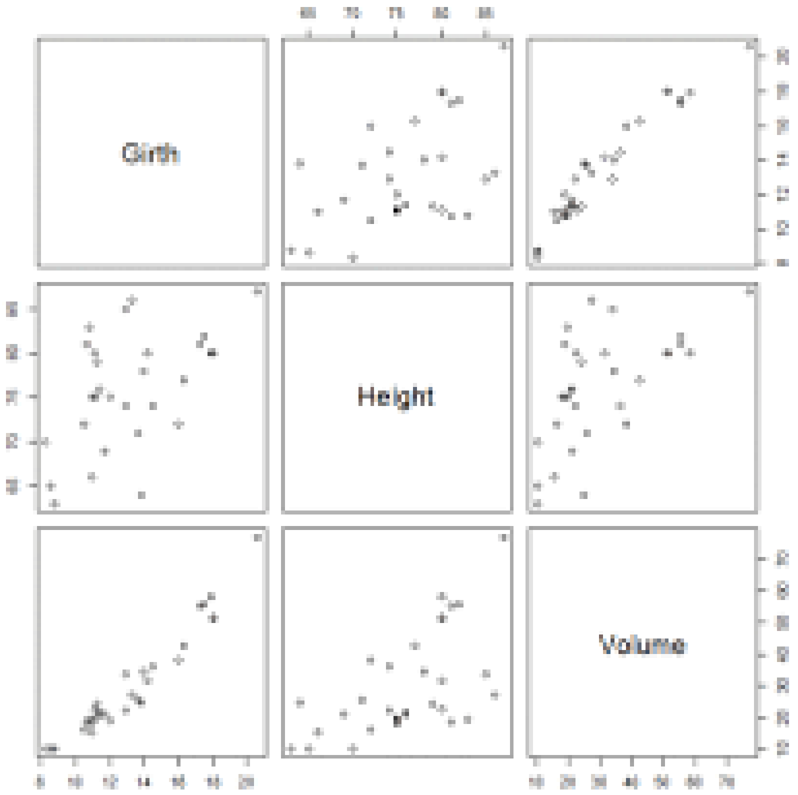
```
input <- mtcars[,c('wt','mpg’)]

plot(x = input$wt,y = input$mpg,
   xlab = "Weight",
   ylab = "Milage",
   xlim = c(2.5,5),
   ylim = c(15,30),
   main = "Weight vs Milage"
)
```



SCATTERPLOT MATRICES

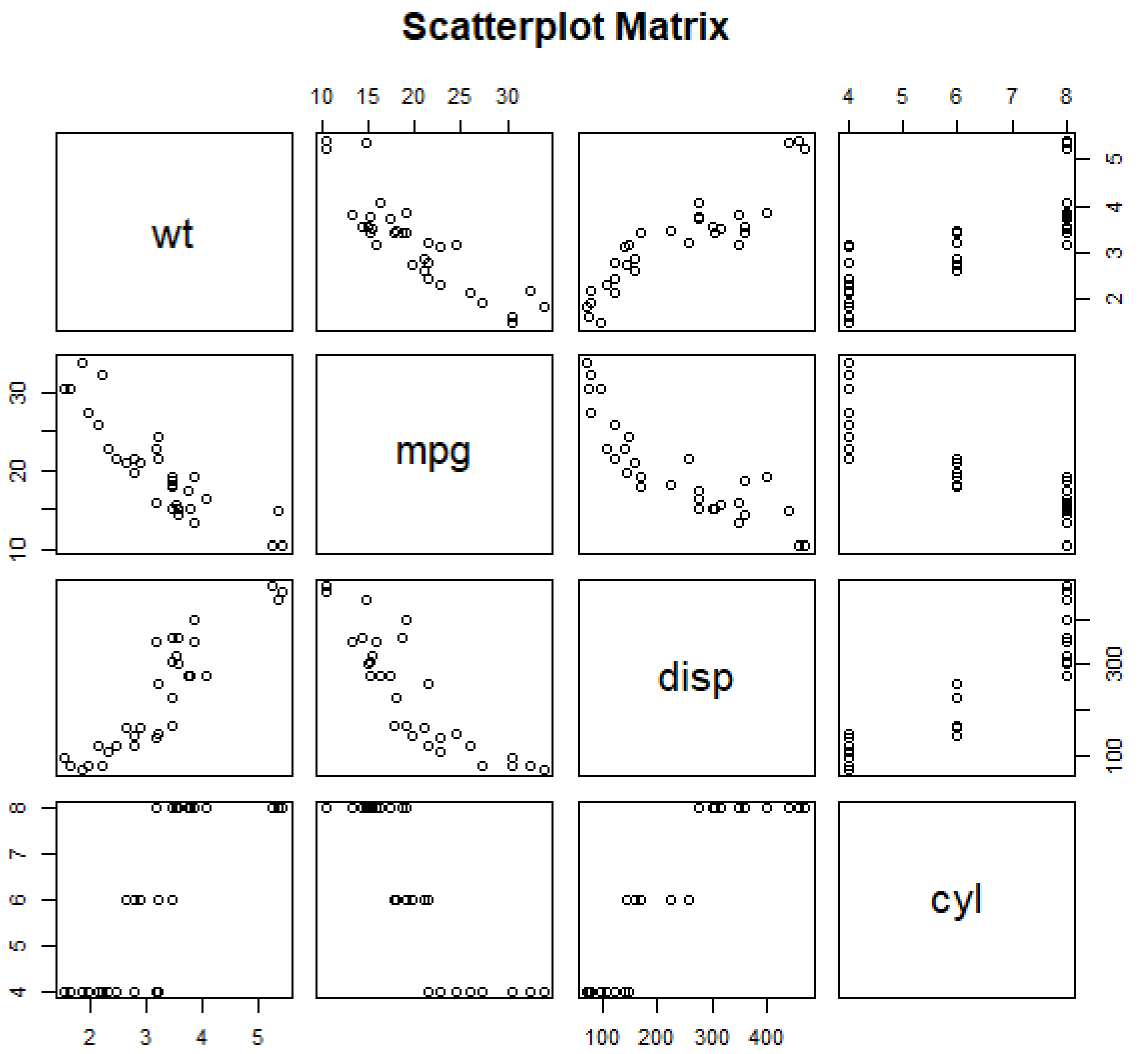
- The basic syntax for creating scatterplot matrices in R is –

```
pairs(formula, data)
```



SCATTERPLOT MATRICES

- Ex: pairs(~wt+mpg+disp+cyl,data = mtcars, main = "Scatterplot Matrix")



PIE CHART

- A pie-chart is a representation of values as slices of a circle with different colors.

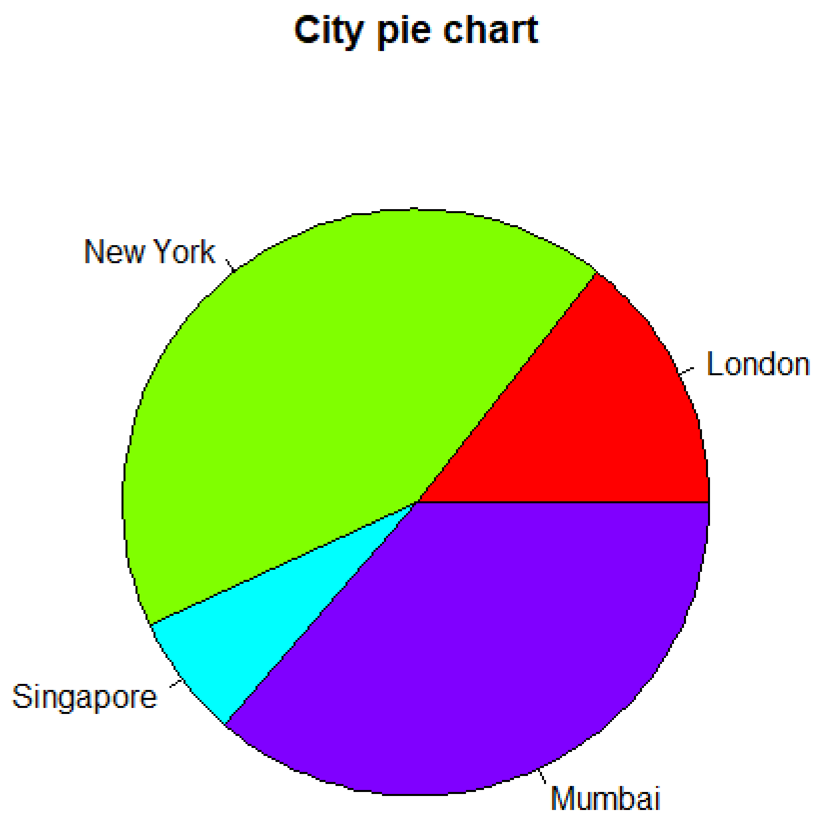
- The slices are labeled and the numbers corresponding to each slice is also represented in the chart.

- In R the pie chart is created using the pie() function which takes positive numbers as a vector input.

PIE CHART

- Ex:

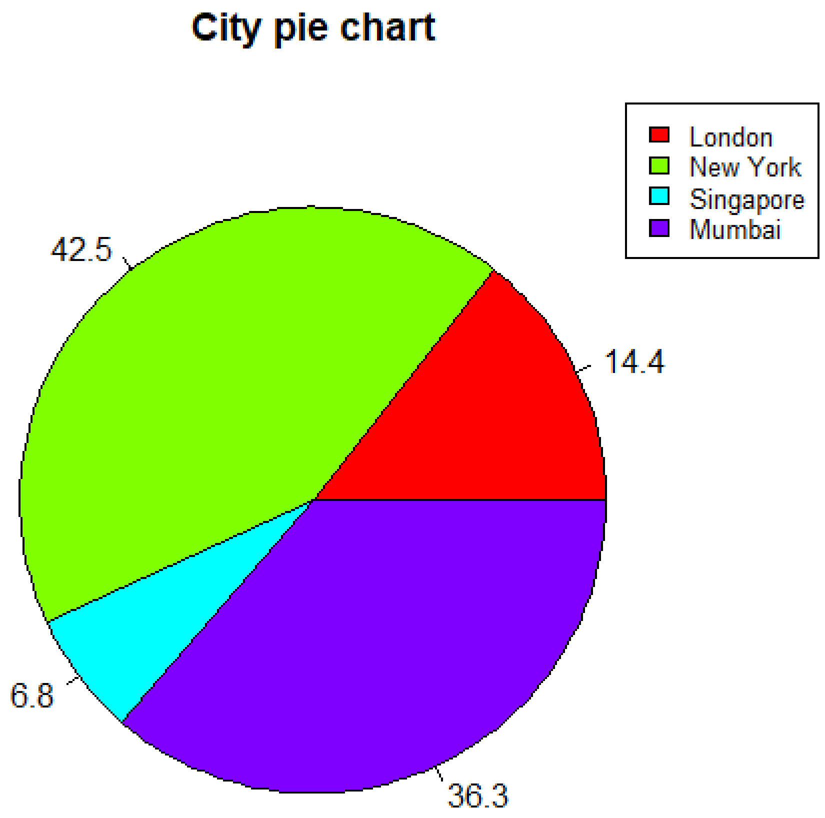
```
x <- c(21, 62, 10, 53)

labels <- c("London", "New York" "Singapore", "Mumbai")

pie(x,
labels,
main = "City pie chart",
col = rainbow(length(x)))
```



SLICE PERCENTAGES AND CHART LEGEND

- Ex:

```
x <- c(21, 62, 10,53)
labels <- c("London","New York",
"Singapore", "Mumbai")

piepercent<- round(100*x/sum(x), 1)

pie(x, labels = piepercent, main = "City
pie chart",col = rainbow(length(x)))

legend("topright", c("London","New
York","Singapore","Mumbai"), cex = 0.8,
 fill = rainbow(length(x)))
```


City pie chart
🟥 London
🟩 New York
🟦 Singapore
🟪 Mumbai