TMC 204
Statistical Data Analysis with R
Unit 5
Graphical Analysis in R Part 3
Scatter Plot

Presented By : Aditya Joshi
Asst. Professor
Department of Computer Application
Graphic Era Deemed to be University
02-05-2020

Scatter Plot :

The Scatter plot is used especially to show the relationship between two variables

A scatter plot can be created using the function **plot(x, y).** The function **lm()** will be used to fit linear models between y and x. A **regression line** will be added on the plot using the function **abline(),** which takes the output of **lm()** as an argument. You can also add a smoothing line using the function **loess().**

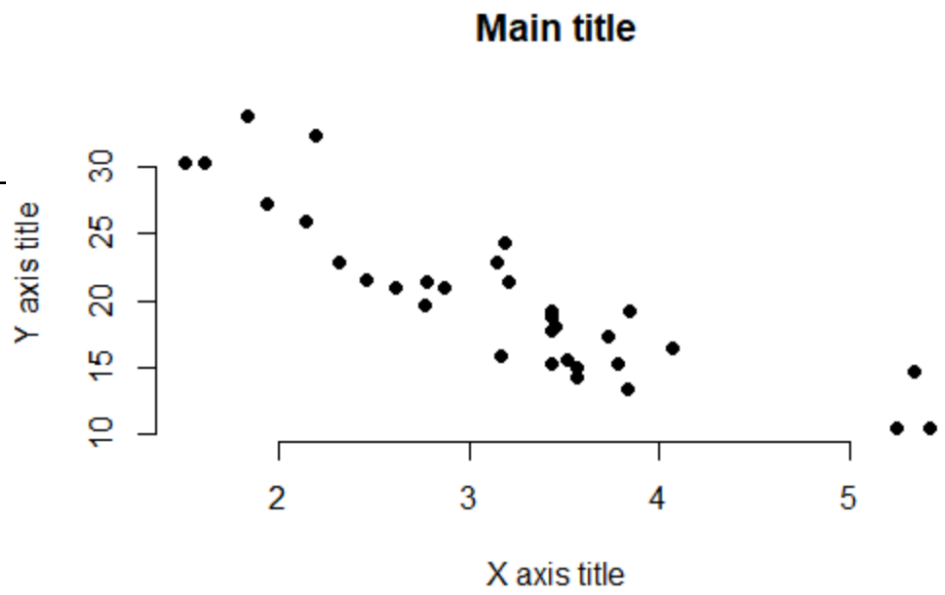
Preliminary tasks

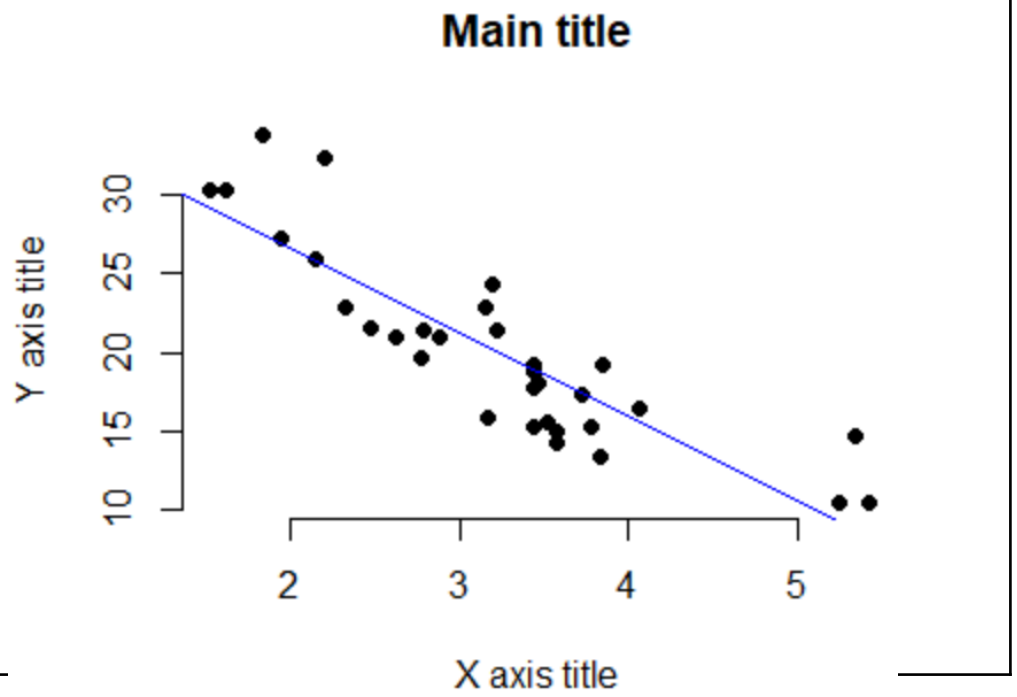
- Launch RStudio
- Prepare your data : prepare your data and save it in an external .txt tab or .csv files
- Import your data into R

Here, we'll use the R built-in **mtcars** data set.

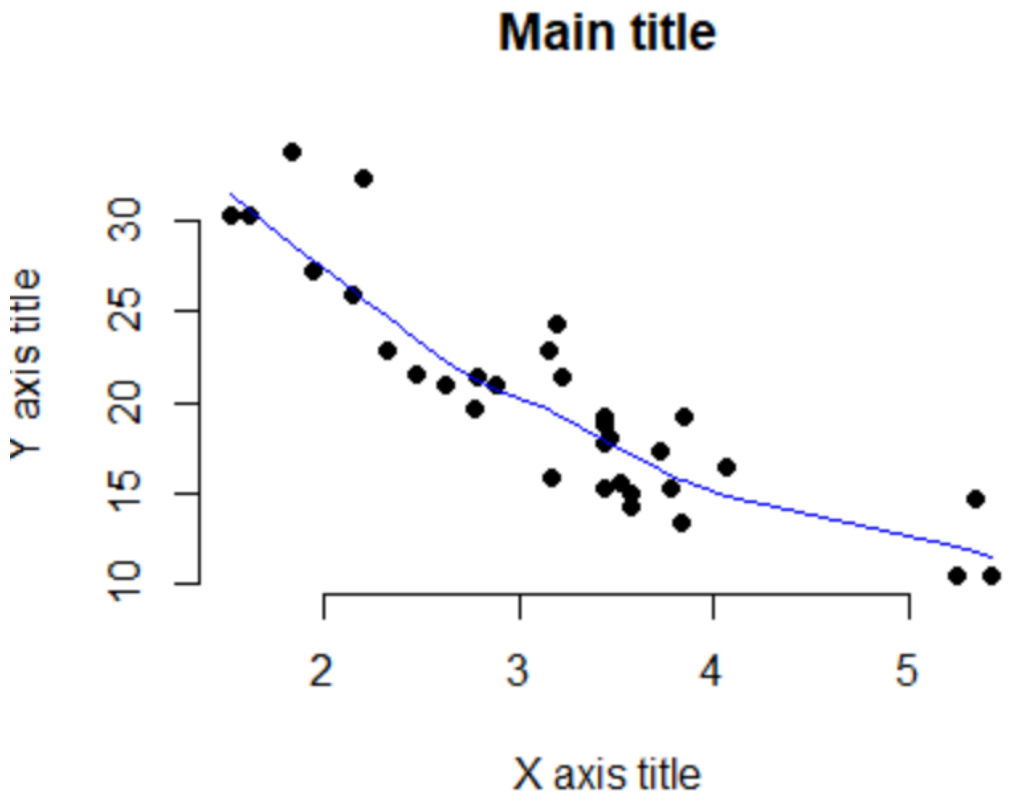
R basic scatter plot: plot()

```
x <- mtcars$wt
y <- mtcars$mpg
# Plot with main and axis titles
# Change point shape (pch = 19) and remove
frame.
plot(x, y, main = "Main title",
     xlab = "X axis title", ylab = "Y axis title",
     pch = 19, frame = FALSE)
# Add regression line
plot(x, y, main = "Main title",
     xlab = "X axis title", ylab = "Y axis title",
     pch = 19, frame = FALSE)
abline(lm(y ~ x, data = mtcars), col = "blue")
```





```
# Add loess fit
plot(x, y, main = "Main title",
     xlab = "X axis title", ylab = "Y axis title",
     pch = 19, frame = FALSE)
lines(lowess(x, y), col = "blue")
```

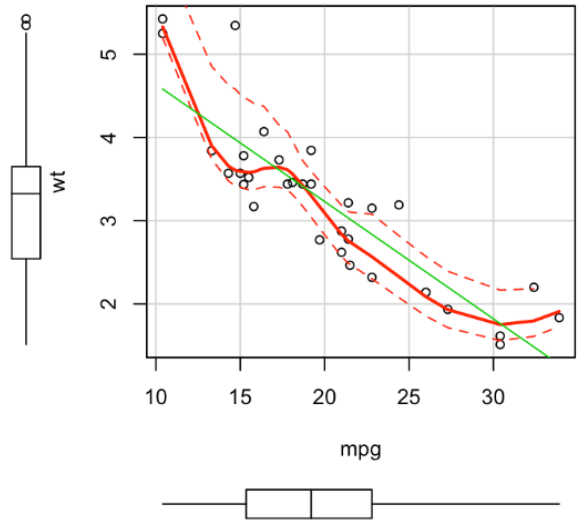


Enhanced scatter plots: car::scatterplot()

The function scatterplot() [in car package] makes enhanced scatter plots, with box plots in the margins, a non-parametric regression smooth, smoothed conditional spread, outlier identification, and a regression line

Install car package:

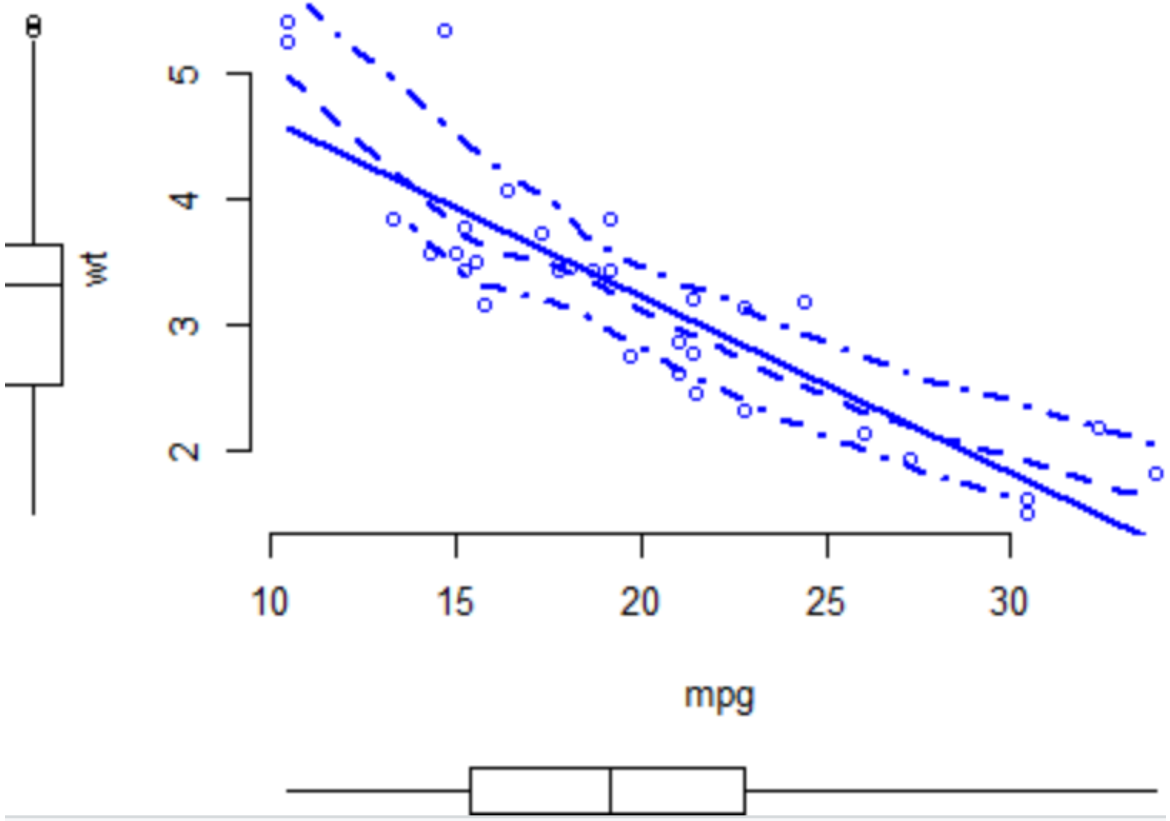
```
library("car")
scatterplot(wt ~ mpg, data = mtcars)
```



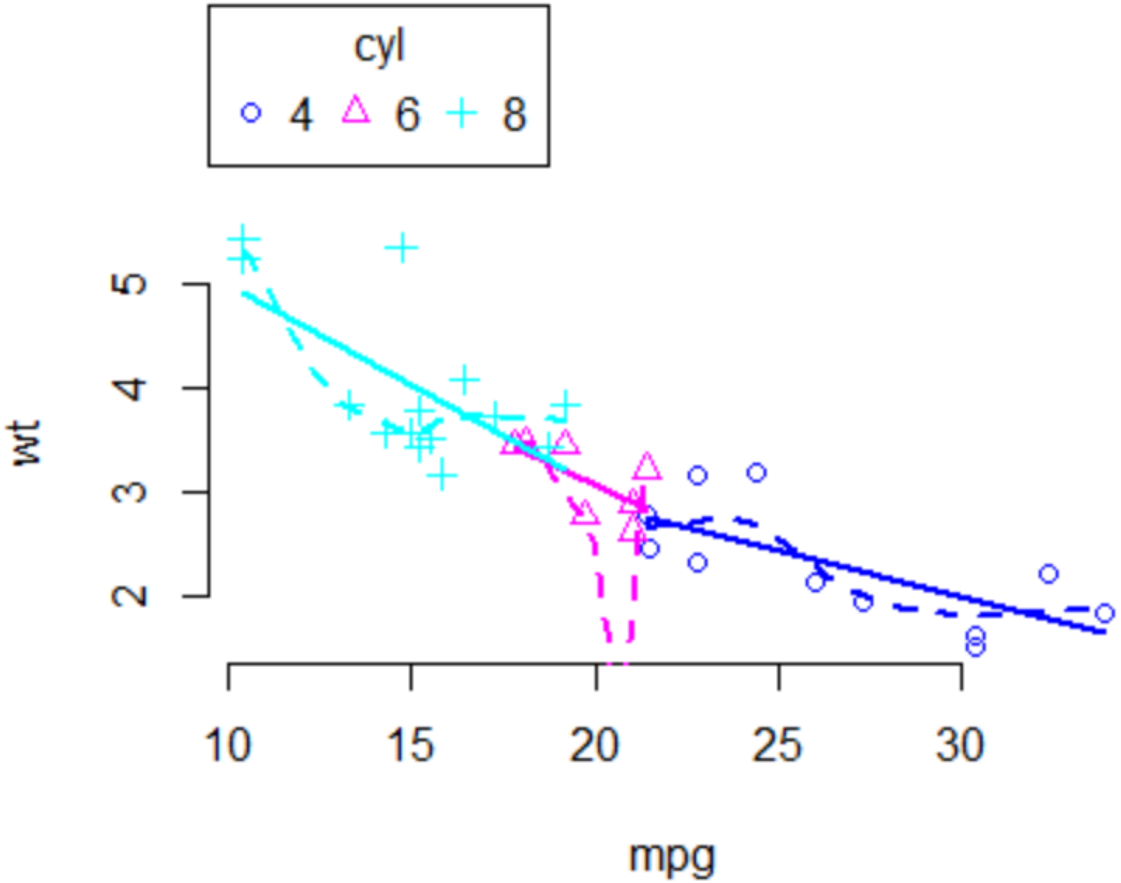
plot contains:

the points
the regression line (in green)
the smoothed conditional spread (in red dashed line)
the non-parametric regression smooth (solid line, red)

```
# Suppress the smoother and frame
scatterplot(wt ~ mpg, data = mtcars,
        smoother = FALSE, grid = FALSE, frame = FALSE)
```




```
# Scatter plot by groups ("cyl")
scatterplot(wt ~ mpg | cyl, data = mtcars,
        smoother = FALSE, grid = FALSE, frame = FALSE)
```



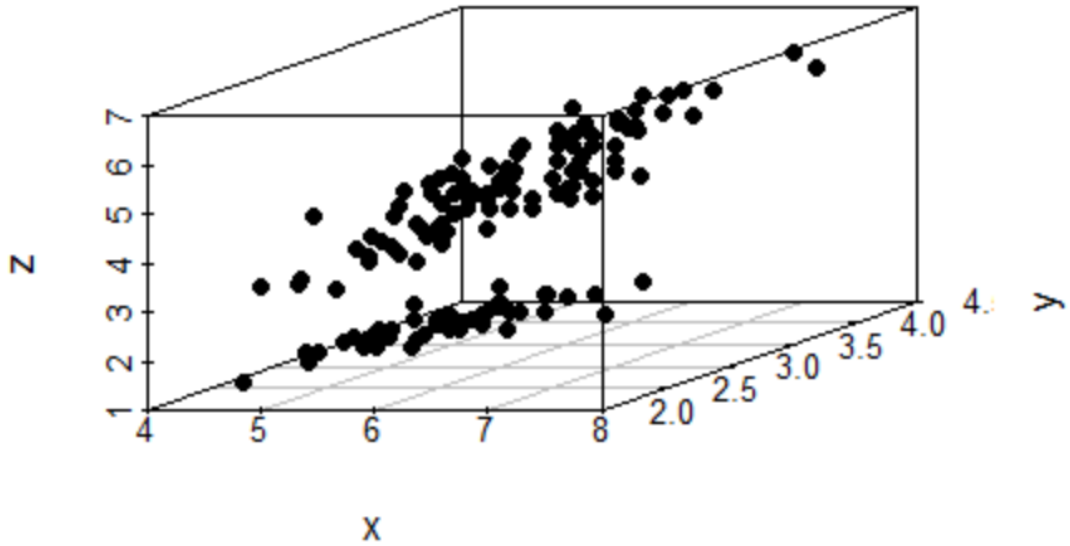
3D scatter plots

To plot a 3D scatterplot the function **scatterplot3D** [in **scatterplot3d** package can be used].

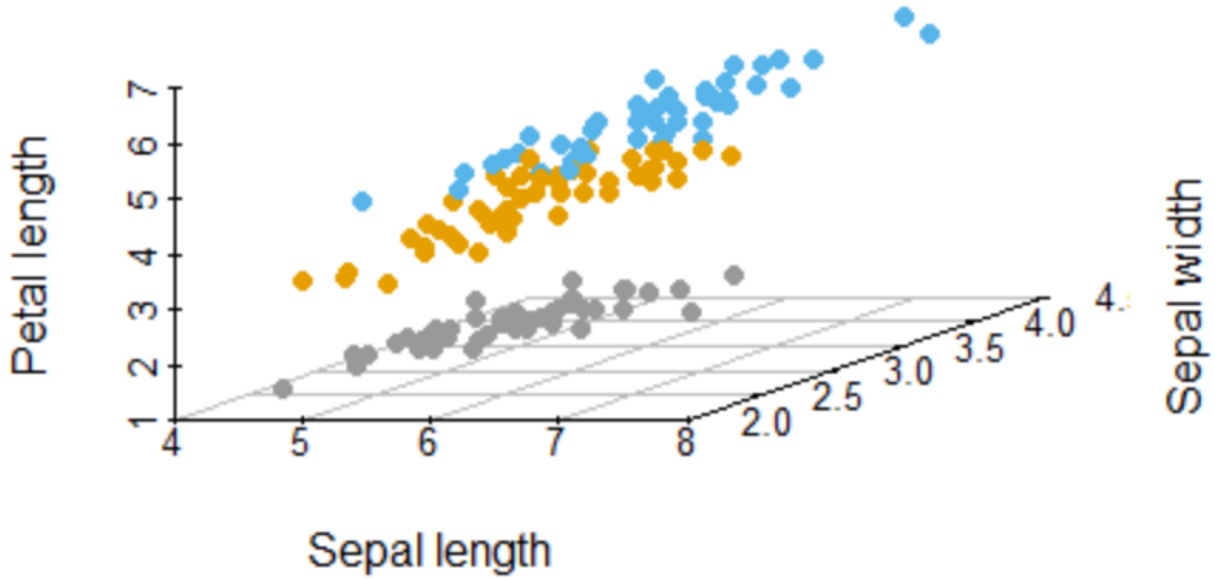
The following R code plots a 3D scatter plot using *iris* data set.

```
head(iris)
> install.packages("scatterplot3d")
# Prepare the data set
x <- iris$Sepal.Length
y <- iris$Sepal.Width
z <- iris$Petal.Length
grps <- as.factor(iris$Species)
# Plot
library(scatterplot3d)
scatterplot3d(x, y, z, pch = 16)
```

```
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1          5.1         3.5          1.4         0.2  setosa
2          4.9         3.0          1.4         0.2  setosa
3          4.7         3.2          1.3         0.2  setosa
4          4.6         3.1          1.5         0.2  setosa
5          5.0         3.6          1.4         0.2  setosa
6          5.4         3.9          1.7         0.4  setosa
There were 50 or more warnings (use warnings() to see the first 50)
```



```
# Change color by groups
# add grids and remove the box around the plot
# Change axis labels: xlab, ylab and zlab
colors <- c("#999999", "#E69F00", "#56B4E9")
scatterplot3d(x, y, z, pch = 16, color = colors[grps],
              grid = TRUE, box = FALSE, xlab = "Sepal length",
              ylab = "Sepal width", zlab = "Petal length")
```



Summary

Create a scatter plot:

Using R base function:

```
with(mtcars, plot(wt, mpg, frame = FALSE))
```

Using car package:

```
car::scatterplot(wt ~ mpg, data = mtcars,
                 smoother = FALSE, grid = FALSE)
```

3D scatter plot:

```
library(scatterplot3d)
with(iris,
    scatterplot3d(x = Sepal.Length, y = Sepal.Width,
                  z = Petal.Length, pch = 16,
                  grid = TRUE, box = FALSE)
)
```