Association of Variables :: Univariate and Bivariate Scatter Plots

Association of Two Variables

- Number of hours of study affect the marks obtained in an examination.

- Electricity/power consumption increases when the weather temperature increases.

- Weight of infants and small children increases as their height increases under normal circumstances.

Association of Two Variables

The observations on both the variables are related to each other.

How to know the variables are related?

How to know the degree of relationship between the two variables?

Graphical procedures – Two dimensional plots, three dimensional plots etc.

Quantitative procedures – Correlation coefficients, contingency tables, Chi-square statistic, linear regression, nonlinear regression etc.

Association of Two Variables

How to judge or graphically summarize the association of two variables?
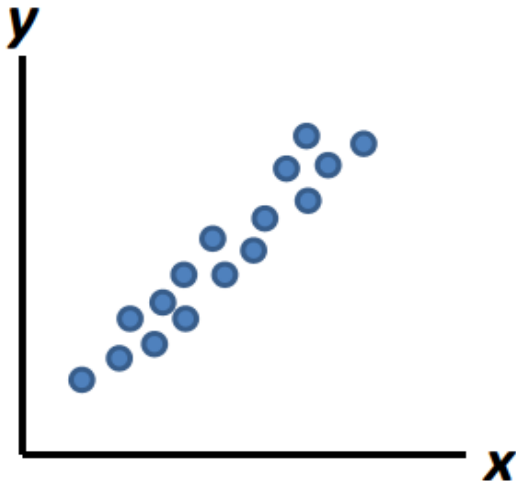
X, Y : Two variables

n pairs of observations are available as (x_1, y_1), (x_2, y_2),...,(x_n, y_n)
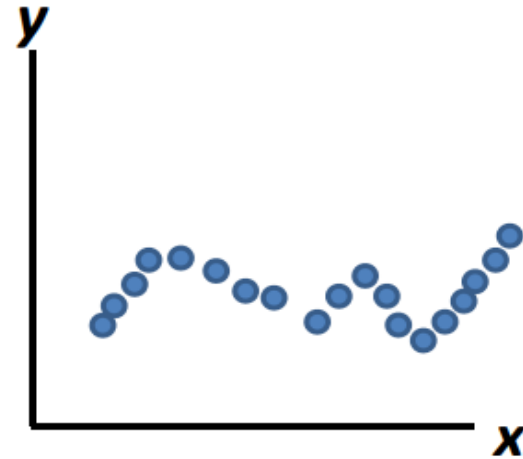
Scatter Plot

Plot the paired observations in a single graph, called as scatter plot.

Scatter plot reveals the nature and trend of possible relationship.

Relationships : Linear   or   nonlinear.

Linear relationship between *X* and *Y*

Nonlinear relationship between *X* and *Y*

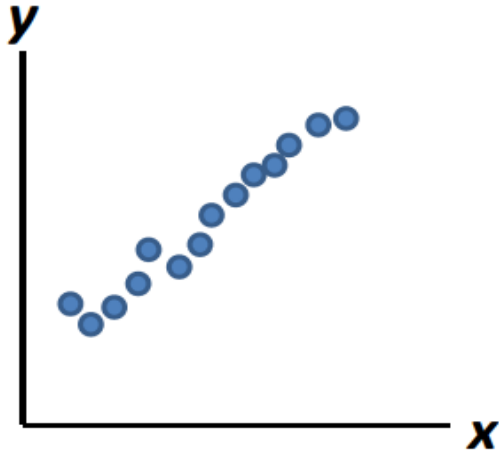# Scatter Plot

Strength and Trend of Relationships
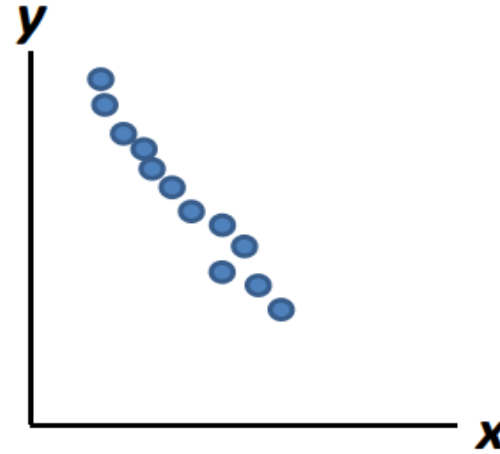

Fig. 1: Strong positive linear relationship

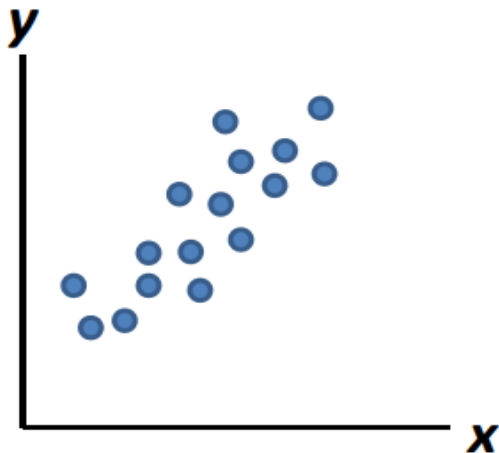Fig. 2: Strong negative linear relationship
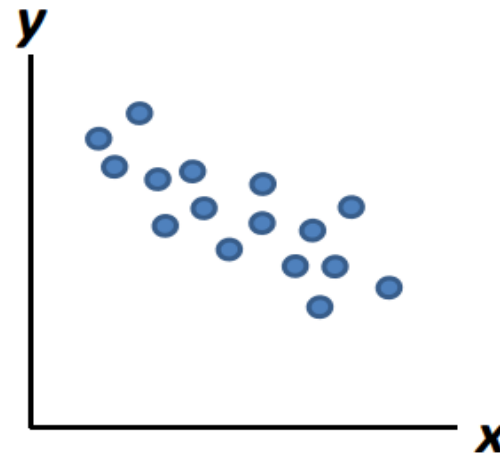
Fig. 3: Moderate positive linear relationship

Fig. 4: Moderate negative linear relationship

Scatter Plot

Strength and Trend of Relationships
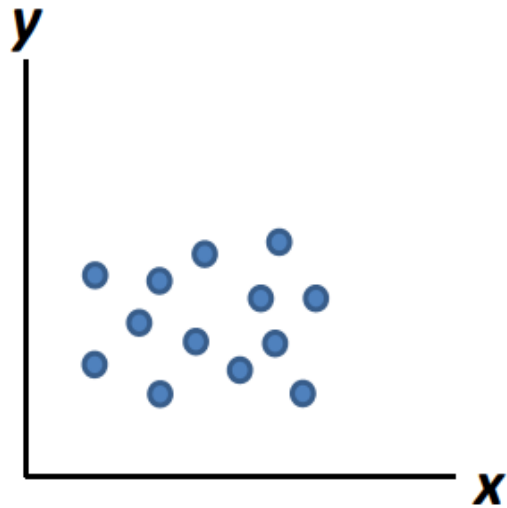

Fig. 5: No clear relationship

Univariate Scatter Plot

Plot command for one variable:

x: Data vector

plot(x)

Example

Height of 50 persons are recorded as follows:

166,125,130,142,147,159,159,147,165,156,149,164,137,166,135,142,

133,136,127,143,165,121,142,148,158,146,154,157,124,125,158,159,

164,143,154,152,141,164,131,152,152,161,143,143,139,131,125,145,

140,163

```
> height = c(166,125,130,142,147,159,159,147,
165,156,149,164,137,166,135,142,133,136,127,143,
165,121,142,148,158,146,154,157,124,125,158,159,
164,143,154,152,141,164,131,152,152,161,143,143,
139,131,125,145,140,163)
```
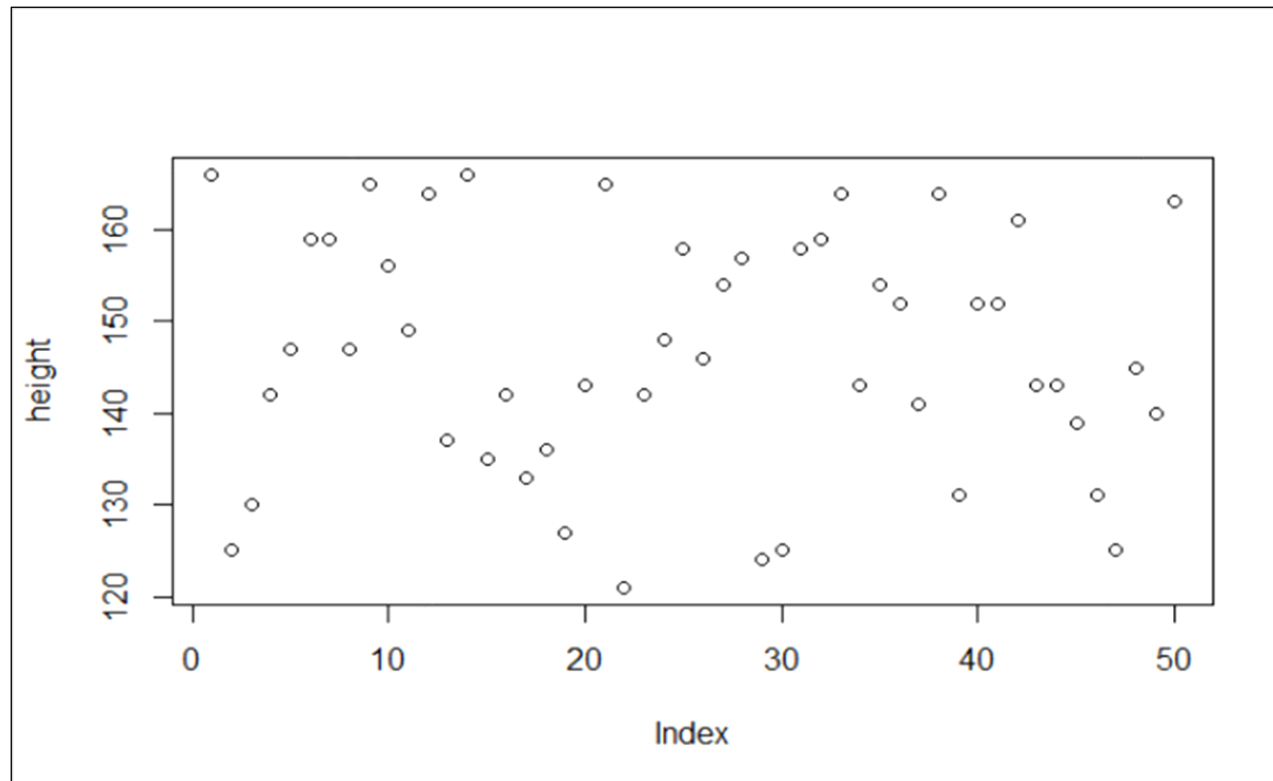
Univariate Scatter Plot

```
> height = c(166,125,130,142,147,159,159,147,165,156,149,164,137,166,135,142,13
3,136,127,143,165,121,142,148,158,146,154,157,124,125,158,159,164,143,154,152,14
1,164,131,152,152,161,143,143,139,131,125,145,140,163)
> height
 [1] 166 125 130 142 147 159 159 147 165 156 149 164 137 166 135 142 133 136
[19] 127 143 165 121 142 148 158 146 154 157 124 125 158 159 164 143 154 152
[37] 141 164 131 152 152 161 143 143 139 131 125 145 140 163
```

```
> plot(height)
```

Bivariate plots

Provide first hand visual information about the nature and degree of relationship between two variables.

Relationship can be linear or nonlinear.

We discuss several types of plots through examples.

Scatter Plot

Plot command:

`x, y:` **Two data vectors**

`plot(x, y)`

`plot(x, y, type)`

type	
"p" for points	"l" for lines
"b" for both	"c" for the lines part alone of "b"
"o" for both 'overplotted'	"s" for stair steps.
"h" for 'histogram' like (or 'high-density') vertical lines	

Scatter Plot

Plot command

`x, y:` **Two data vectors**

`plot(x, y)`

`plot(x, y, type)`

Get more details from help: `help("type")`

Other options:

`main`	an overall title for the plot.
`suba`	sub title for the plot.
`xlaba`	title for the x axis.
`ylaba`	title for the y axis.
`aspthe`	*y/x* aspect ratio.

Scatter Plot

Example

Data on marks obtained by 20 students out of 500 marks and the

number of hours they studied per week are recorded as follows:

We know from experience that marks obtained by students increase

as the number of hours increase.

Marks	337	316	327	340	374	330	352	353	370	380
Number of hours per week	23	25	26	27	30	26	29	32	33	34

Marks	384	398	413	428	430	438	439	479	460	450
Number of hours per week	35	38	39	42	43	44	45	46	44	41

Scatter Plot

Example

marks =
c(337,316,327,340,374,330,352,353,370,380,384,398,413,428,430,438,439,479,460,450)

hours =
c(23,25,26,27,30,26,29,32,33,34,35,38,39,42,43,44,45,46,44,41)

Scatter Plot

Example
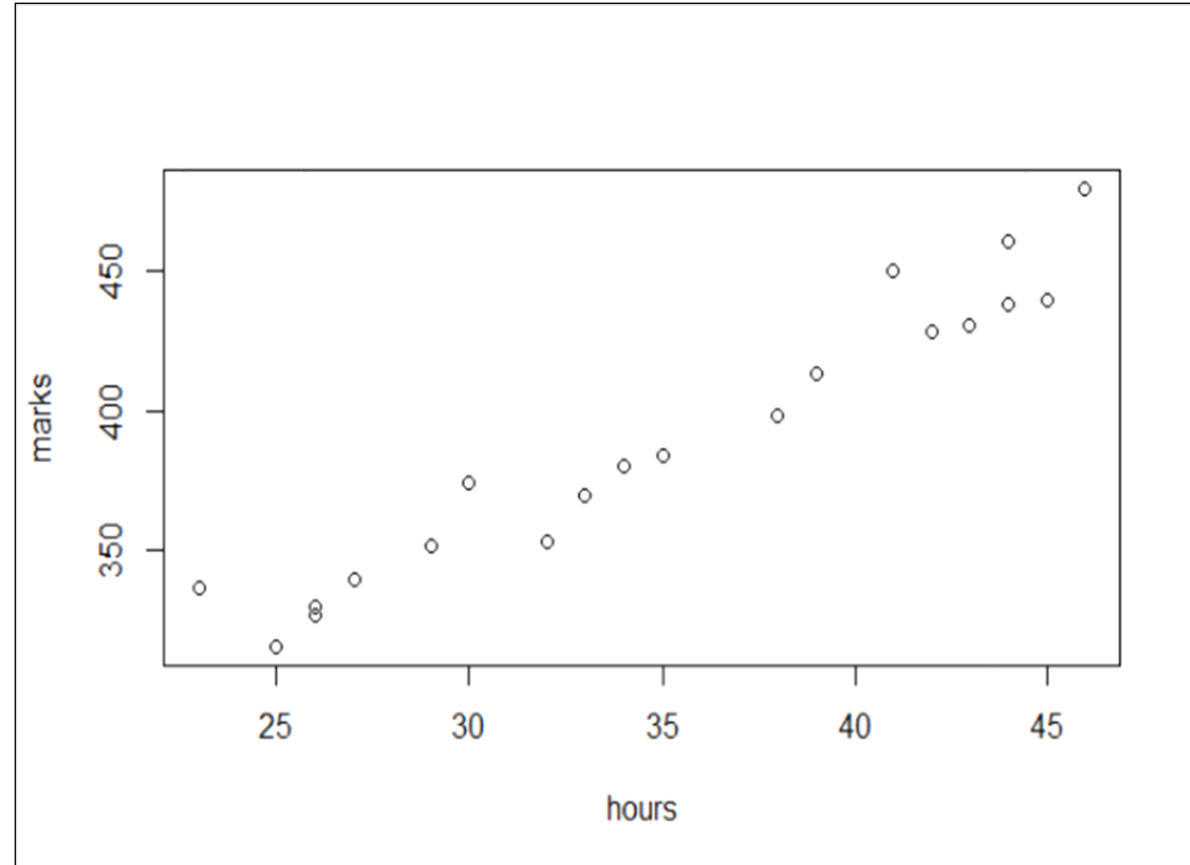
`plot` command:

`x, y:` Two data vectors

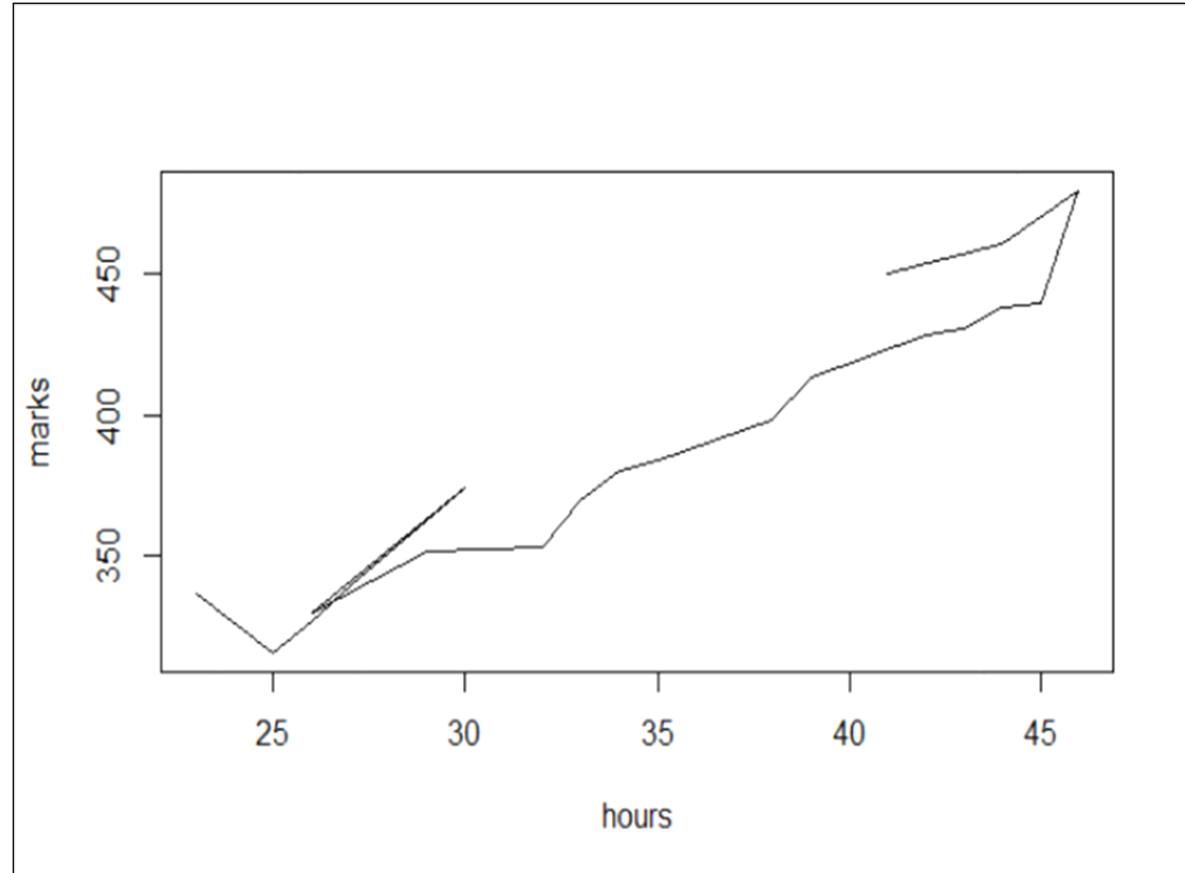Various type of plots are possible to draw.

`plot(x, y)`

`plot(hours, marks)`

Scatter Plot

Example

plot(hours, marks, "l")

"l" for lines,

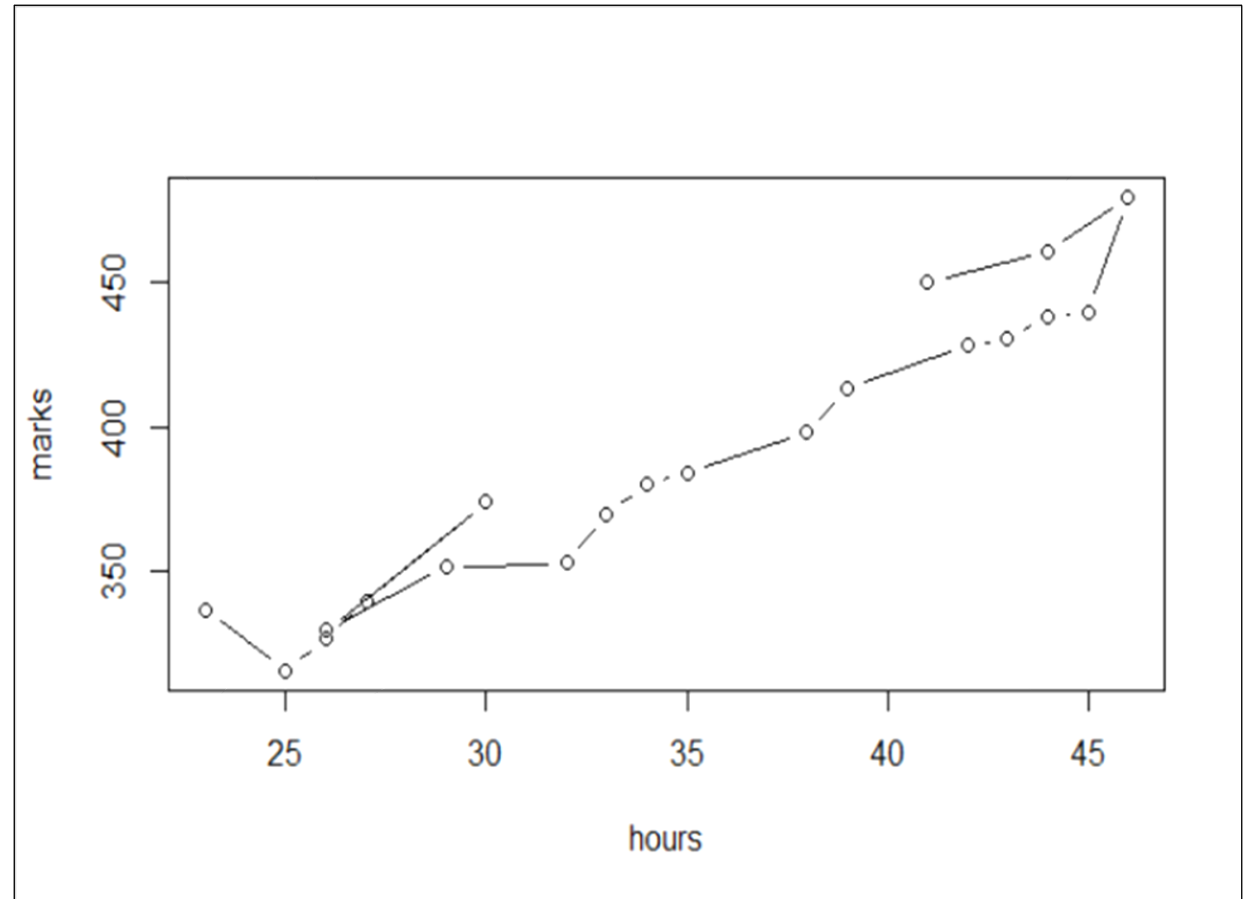# Scatter Plot

Example
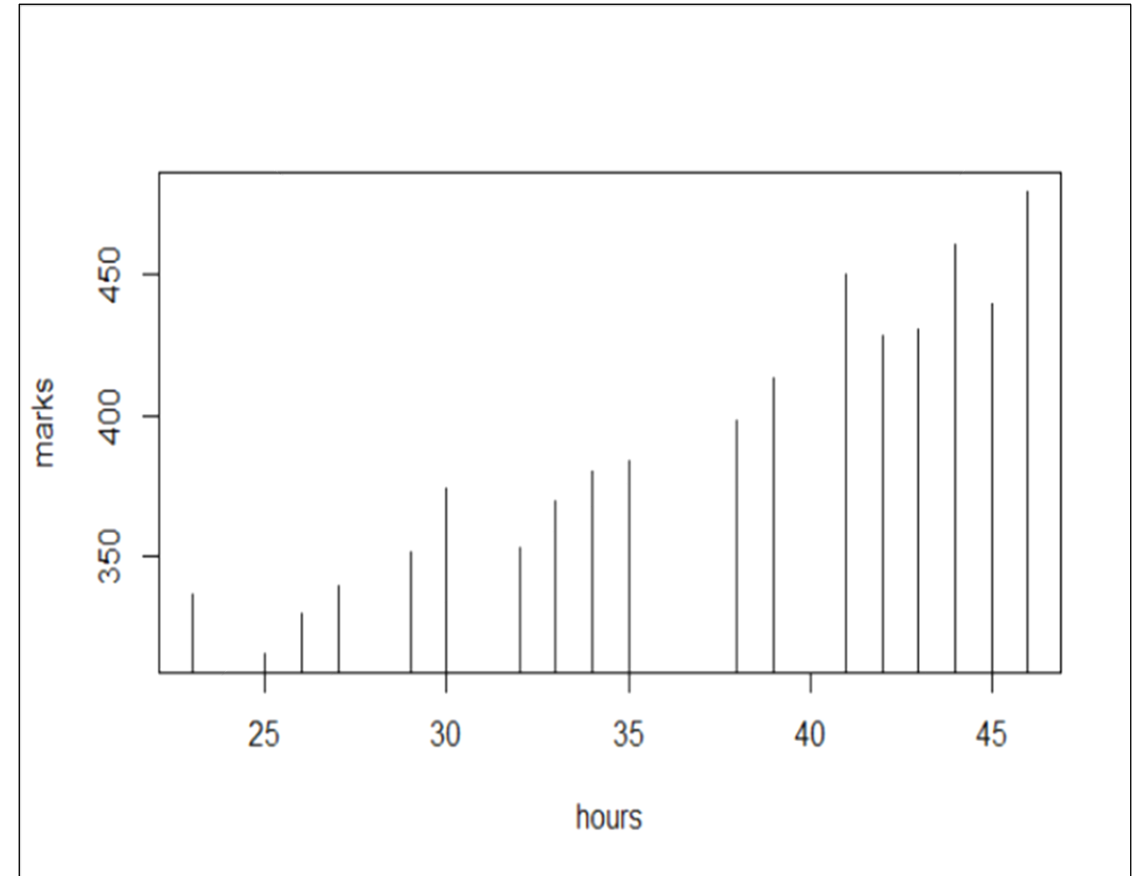
`plot(hours, marks, "b")`

"b" for both – line and point

Scatter Plot

Example

plot(hours, marks, "h")

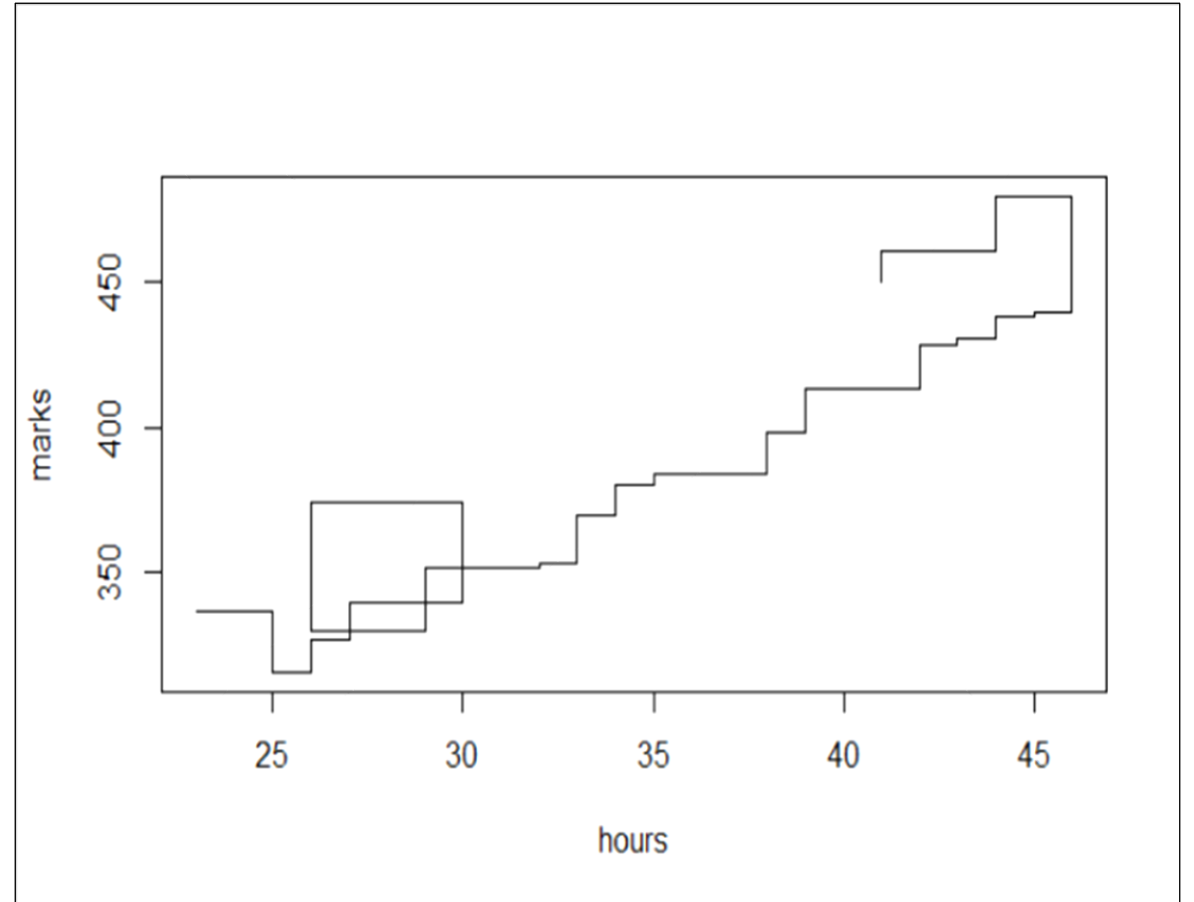"h" for 'histogram' like (or 'high-density') vertical lines

Scatter Plot

Example

plot(hours, marks, "s")

"s" for stair steps.

Scatter Plot

Example

```
plot(hours, marks, xlab="Number of weekly
hours", ylab="Marks obtained",  main="Marks
obtained versus  Number of hours per week")
```


Marks obtained versus Number of hours per week