Name: Kasturi Basu

Roll: 10020017900013

Paper Code: STSPC0R05P

Paper: Lab-1

Q1: Inserting Dataset on "TREES"

In [1]:
```
trees
```

A data.frame: 31 × 3

Girth <dbl>	Height <dbl>	Volume <dbl>
8.3	70	10.3
8.6	65	10.3
8.8	63	10.2
10.5	72	16.4
10.7	81	18.8
10.8	83	19.7
11.0	66	15.6
11.0	75	18.2
11.1	80	22.6
11.2	75	19.9
11.3	79	24.2
11.4	76	21.0
11.4	76	21.4
11.7	69	21.3
12.0	75	19.1
12.9	74	22.2
12.9	85	33.8
13.3	86	27.4
13.7	71	25.7
13.8	64	24.9
14.0	78	34.5
14.2	80	31.7
14.5	74	36.3
16.0	72	38.3
16.3	77	42.6
17.3	81	55.4
17.5	82	55.7
17.9	80	58.3
18.0	80	51.5
18.0	80	51.0
20.6	87	77.0

Finding Mean of each column of "TREES"

In [2]:
```
mean(trees$Volume)
```

30.1709677419355

In [3]:
```
mean(trees$Girth)
```

13.2483870967742

In [4]:
```
mean(trees$Height)
```

76

Finding Median of each column of "TREES"

In [5]:
```
median(trees$Volume)
```

24.2

In [6]:
```
median(trees$Girth)
```

12.9

In [8]:
```
median(trees$Height)
```

76

Finding Standard Deviation(sd) of each column of "TREES"

In [11]:
```
sd(trees$Volume)
```

16.4378464434647

In [12]:
```
sd(trees$Girth)
```

3.13813861683875

In [13]:
```
sd(trees$Height)
```

6.3718129288296

Finding Regression Equation of Volume on Girth(regrsn_eq_v_on_g)

In [14]:
```
dataset= trees
```

In [15]:
```
regrsn_eq_v_on_g = lm(Volume~Girth, data=dataset)
regrsn_eq_v_on_g
```

```
Call:
lm(formula = Volume ~ Girth, data = dataset)

Coefficients:
(Intercept)        Girth
    -36.943        5.066
```

This result shows the Intercept and the Beta coefficient for the Girth variable. The estimated regression line of volume on girth can be written as: Volume= -36.943+5.066*Girth

Finding Regression equation of Volume on Height(regrsn_eq_v_on_h)

In [16]:
```
regrsn_eq_v_on_h = lm(Volume~Height, data= dataset)
regrsn_eq_v_on_h
```

```
Call:
lm(formula = Volume ~ Height, data = dataset)

Coefficients:
(Intercept)       Height
    -87.124        1.543
```

This result shows the Intercept and the Beta coefficient for the Height variable. The estimated regression line of Volume on Height can be written as: Volume= -87.124+1.543*Height

Summary

In [22]:
```
summary(regrsn_eq_v_on_g)
```

```
Call:
lm(formula = Volume ~ Girth, data = dataset)

Residuals:
   Min     1Q Median     3Q    Max
-8.065 -3.107  0.152  3.495  9.587

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -36.9435     3.3651  -10.98 7.62e-12 ***
Girth         5.0659     0.2474   20.48  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.252 on 29 degrees of freedom
Multiple R-squared:  0.9353,    Adjusted R-squared:  0.9331
F-statistic: 419.4 on 1 and 29 DF,  p-value: < 2.2e-16
```

In [23]:
```
summary(regrsn_eq_v_on_h)
```

```
Call:
lm(formula = Volume ~ Height, data = dataset)

Residuals:
    Min      1Q  Median      3Q     Max
-21.274  -9.894  -2.894  12.068  29.852

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -87.1236    29.2731  -2.976 0.005835 **
Height        1.5433     0.3839   4.021 0.000378 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13.4 on 29 degrees of freedom
Multiple R-squared:  0.3579,    Adjusted R-squared:  0.3358
F-statistic: 16.16 on 1 and 29 DF,  p-value: 0.0003784
```

From the first summary, i.e,Volume on Girth,std.error of intercept and girth are 3.3651 and 0.2474 respectively and from the second summary, i.e, Volume on Height, std.error of intercept and 29.2731 and 0.3839 respectively.

So, we can say Volume on Girth prediction is less error prone than prediction of Volume on Height. So, Girth is more worthwhile than Height for explaining Volume.