



Mahavir Education Trust's
**SHAH & ANCHOR KUTCHHI ENGINEERING
 COLLEGE**
 Chembur, Mumbai - 400 088
UG Program in Cyber Security

Experiment Number: 11					
Date of Performance:					
Date of Submission:					
Program Execution/ formation/ correction/ ethical practices (07)	Documentation (02)	Timely Submission (03)	Viva Answer to sample questions (03)	Experiment Total (15)	Sign



Mahavir Education Trust's
**SHAH & ANCHOR KUTCHHI ENGINEERING
COLLEGE**
Chembur, Mumbai - 400 088
UG Program in Cyber Security

Experiment 11

Aim: Implement Linear regression using R tool.

Lab outcome: CSL 503.2: Implement data mining algorithms like classification.

Problem Statement: To implement Linear regression.

Theory:

Linear Regression

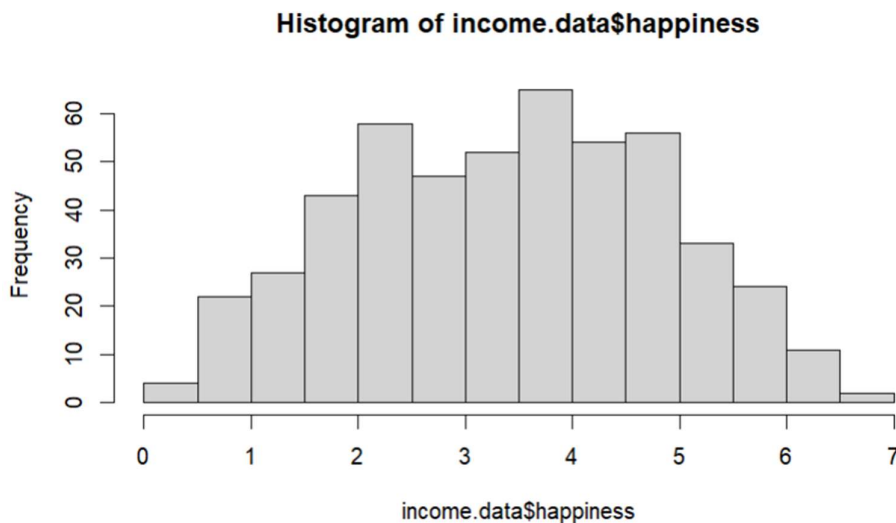
Linear Regression is one of the most widely used regression techniques to model the relationship between two variables. It uses a linear relationship to model the regression line. There are 2 variables used in the linear relationship equation i.e., predictor variable and response variable.

Program Listing and

Output: Simple regression

```
summary(income.data)
> summary(income.data)
      X      income      happiness
Min.  : 1.0   Min.  :1.506   Min.   :0.266
1st Qu.:125.2 1st Qu.:3.006   1st Qu.:2.266
Median :249.5 Median :4.424   Median :3.473
Mean   :249.5 Mean   :4.467   Mean   :3.393
3rd Qu.:373.8 3rd Qu.:5.992   3rd Qu.:4.503
Max.   :498.0 Max.   :7.482   Max.   :6.863
>
```

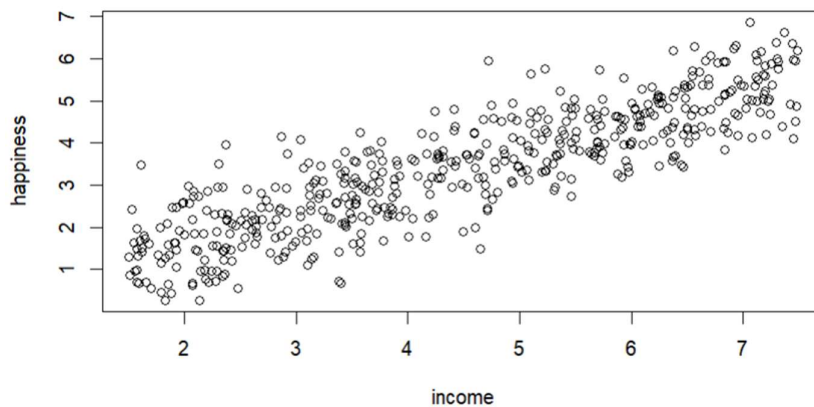
```
hist(income.data$happiness)
```





Mahavir Education Trust's
**SHAH & ANCHOR KUTCHHI ENGINEERING
COLLEGE**
Chembur, Mumbai - 400 088
UG Program in Cyber Security

```
plot(happiness ~ income, data = income.data)
```



Multiple Regression

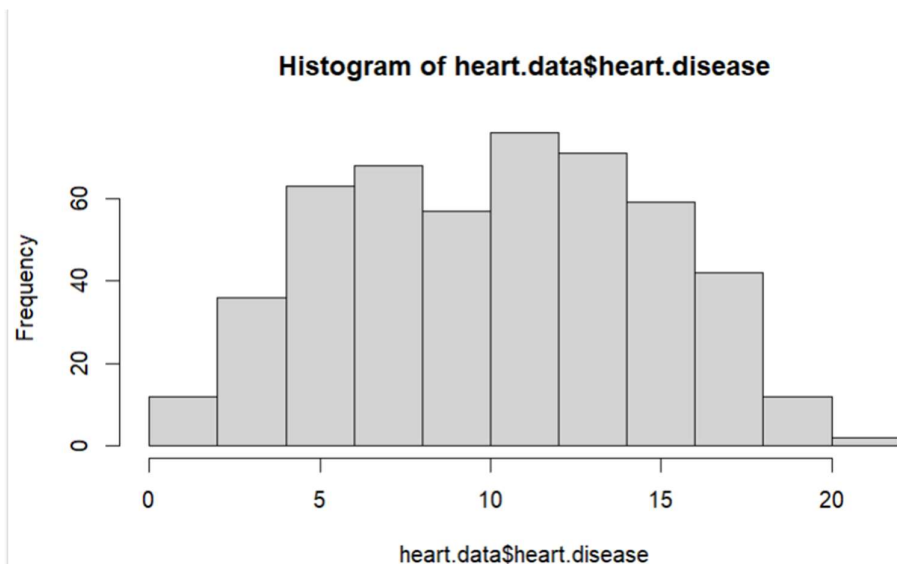
```
summary(heart.data)
```

```
> summary(heart.data)
```

X	biking	smoking	heart.disease
Min. : 1.0	Min. : 1.119	Min. : 0.5259	Min. : 0.5519
1st Qu.:125.2	1st Qu.:20.205	1st Qu.: 8.2798	1st Qu.: 6.5137
Median :249.5	Median :35.824	Median :15.8146	Median :10.3853
Mean :249.5	Mean :37.788	Mean :15.4350	Mean :10.1745
3rd Qu.:373.8	3rd Qu.:57.853	3rd Qu.:22.5689	3rd Qu.:13.7240
Max. :498.0	Max. :74.907	Max. :29.9467	Max. :20.4535

```
>
```

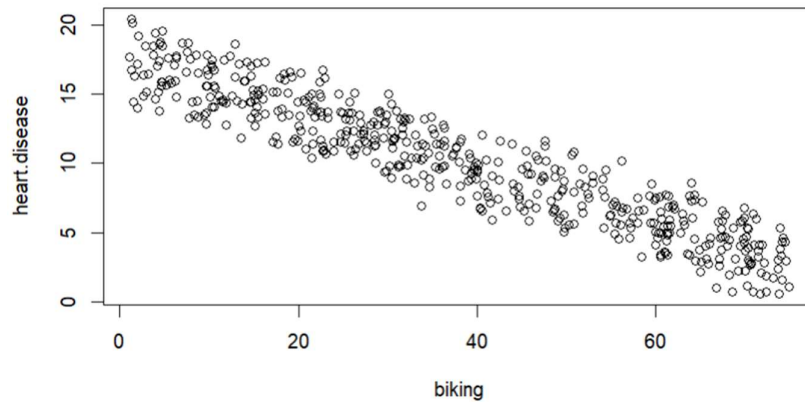
```
hist(heart.data$heart.disease)
```



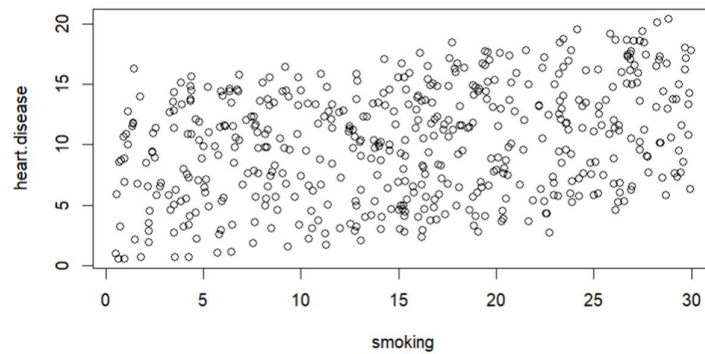


Mahavir Education Trust's
**SHAH & ANCHOR KUTCHHI ENGINEERING
COLLEGE**
Chembur, Mumbai - 400 088
UG Program in Cyber Security

```
plot(heart.disease ~ biking, data=heart.data)
```



```
plot(heart.disease ~ smoking, data=heart.data)
```



```
income.happiness.lm <- lm(happiness ~ income, data  
=income.data)
```

```
summary(income.happiness.lm)
```



Mahavir Education Trust's
SHAH & ANCHOR KUTCHHI ENGINEERING COLLEGE
Chembur, Mumbai - 400 088
UG Program in Cyber Security

```
> income.happiness.lm <- lm(happiness ~ income, data =income.data)
> summary(income.happiness.lm)
```

Call:

```
lm(formula = happiness ~ income, data = income.data)
```

Residuals:

Min	1Q	Median	3Q	Max
-2.02479	-0.48526	0.04078	0.45898	2.37805

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.20427	0.08884	2.299	0.0219 *
income	0.71383	0.01854	38.505	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7181 on 496 degrees of freedom

Multiple R-squared: 0.7493, Adjusted R-squared: 0.7488

F-statistic: 1483 on 1 and 496 DF, p-value: < 2.2e-16

Conclusion:

The above result shows that there is a significant positive relationship between income and happiness (p -value < 0.001), with a 0.713-unit (+/- 0.01) increase in happiness for every unit increase in income.