

## Lab2\_1740256

Jeevan

19/06/2019

### **AIM-**

Fit a simple linear regression to the data that involving the following steps by using R.

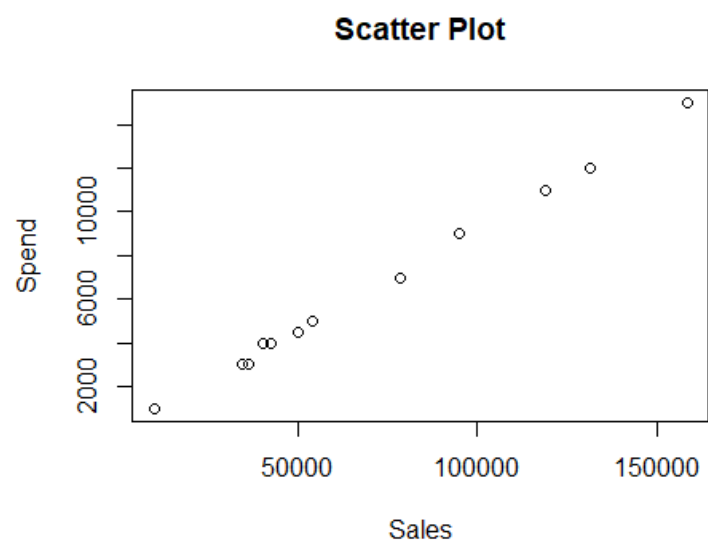
- a) Plot the scatter diagram for the data .
- b) Estimate the parameters of a simple linear regression model.
- c) Obtain the fitted values of the model.
- d) Test the significance of the model parameters and interpret the results.
- e) Obtain the coefficient of determination and interpret the results.
- f) Obtain the 95 percent and 99 percent confidence interval for the slope and intercept parameters.
- g) Check whether the assumptions of the residuals are satisfied or not by using a suitable statistical test.

### **ANALYSIS-**

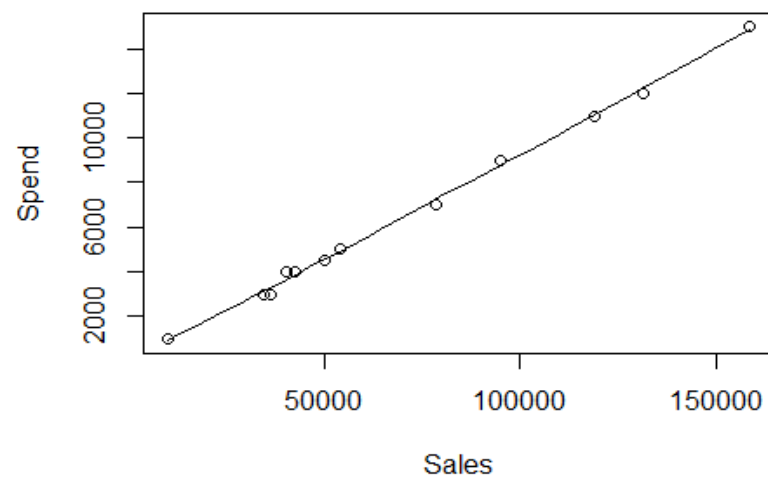
```
library(lmtest)

Spend_sales <- read.csv("C:/Users/Jeevan/Desktop/Christ
University/Statistics/Spend_sales.csv")
View(Spend_sales)
attach(Spend_sales)

plot(Sales,Spend,main = "Scatter Plot")
```



```
scatter.smooth(Sales,Spend)
```



```
cor(Sales,Spend)
```

```
## [1] 0.9988322
```

```
model = lm(Spend~Sales)
```

```
model
```

```
##
```

```
## Call:
```

```
## lm(formula = Spend ~ Sales)
```

```
##
```

```
## Coefficients:
```

```
## (Intercept)      Sales
## -114.67027      0.09392

summary(model)

##
## Call:
## lm(formula = Spend ~ Sales)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -293.22 -165.15  -20.82   188.67   312.02
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.147e+02  1.196e+02  -0.959    0.36
## Sales        9.392e-02  1.437e-03   65.378 1.71e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 217.5 on 10 degrees of freedom
## Multiple R-squared:  0.9977, Adjusted R-squared:  0.9974
## F-statistic: 4274 on 1 and 10 DF, p-value: 1.707e-14

fit=fitted.values(model)
fit

##           1           2           3           4           5           6
## 816.4799 3687.9803 4987.5895 4585.6001 3146.2338 3881.8369
##           7           8           9          10          11          12
## 8795.8754 11054.0602 14770.5836 12221.8957 7258.6417 3293.2230

res=resid(model)
sum(res)

## [1] 1.882938e-13

confint(model, level = 0.95)

##              2.5 %       97.5 %
## (Intercept) -381.20401860 151.86346955
## Sales        0.09072177   0.09712374

confint(model, level = 0.99)

##              0.5 %       99.5 %
## (Intercept) -493.78404056 264.44349151
## Sales        0.08936972   0.09847579

dwtest(Spend~Sales)
```

```
##
## Durbin-Watson test
##
## data: Spend ~ Sales
## DW = 1.1625, p-value = 0.03414
## alternative hypothesis: true autocorrelation is greater than 0
```

## **INTERPRETATION-**

The parameters for the simple linear regression model are -

**Coefficients:**

```
## (Intercept)      Sales
## -114.67027      0.09392
```

The fitted values of the model are -

```
##          1          2          3          4          5          6
## 816.4799 3687.9803 4987.5895 4585.6001 3146.2338 3881.8369
##          7          8          9         10         11         12
## 8795.8754 11054.0602 14770.5836 12221.8957 7258.6417 3293.2230
```

The significance of the model parameters -

1.71e-14 is less than 0.05 and is therefore highly significant and therefore we reject the null hypothesis.

The coefficient of determination is **0.9974** and this shows that the model is a good fit to the data.

The 95 percent and 99 percent confidence interval for the slope and intercept parameters are -

```
##          2.5 %          97.5 %
## (Intercept) -381.20401860 151.86346955
## Sales      0.09072177  0.09712374

##          0.5 %          99.5 %
## (Intercept) -493.78404056 264.44349151
## Sales      0.08936972  0.09847579
```

The assumptions of the residuals are satisfied by the Durbin Watson Test

```
## Durbin-Watson test
## data: Spend ~ Sales
## DW = 1.1625, p-value = 0.03414
## alternative hypothesis: true autocorrelation is greater than 0
```

-----