PRACTICAL 2

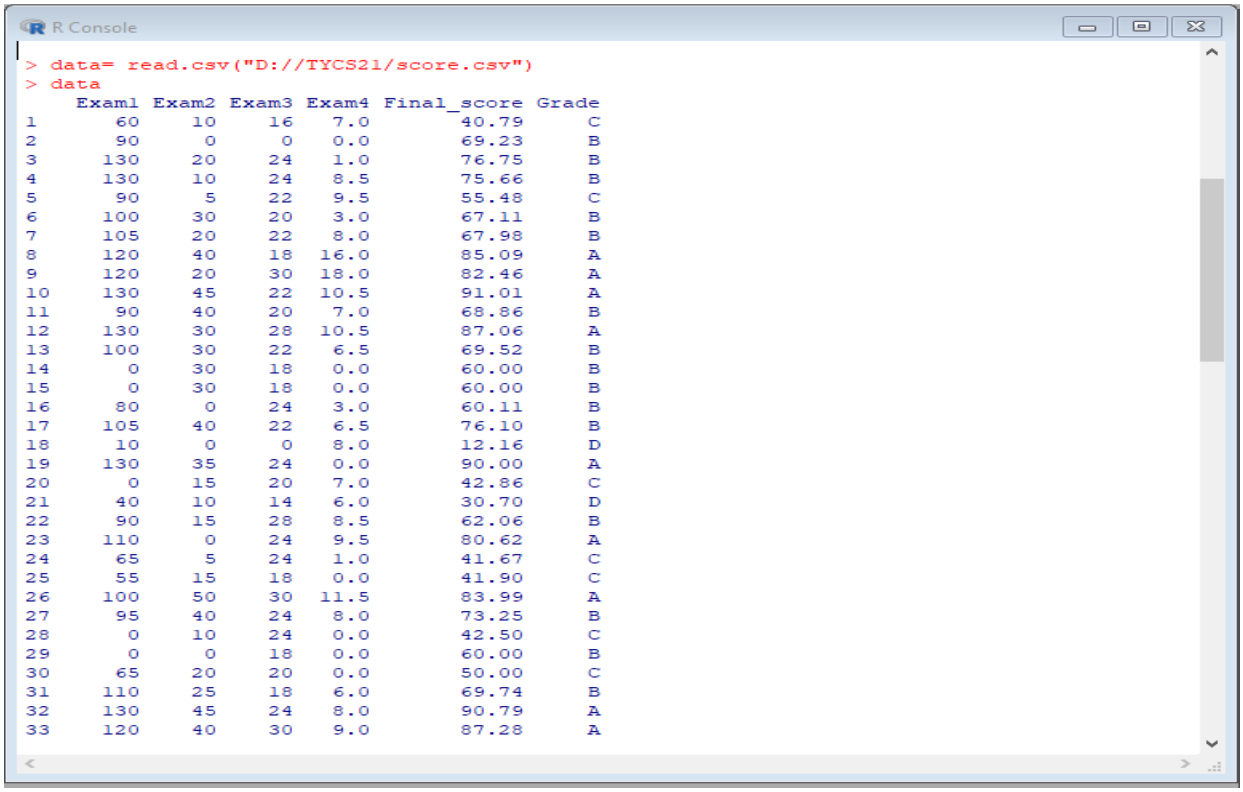
Aim: Simple /Multiple Linear Regressions.

#IMPOER DATASET:

Command:

>data=read.csv ("D://tycs/score.csv")

>data



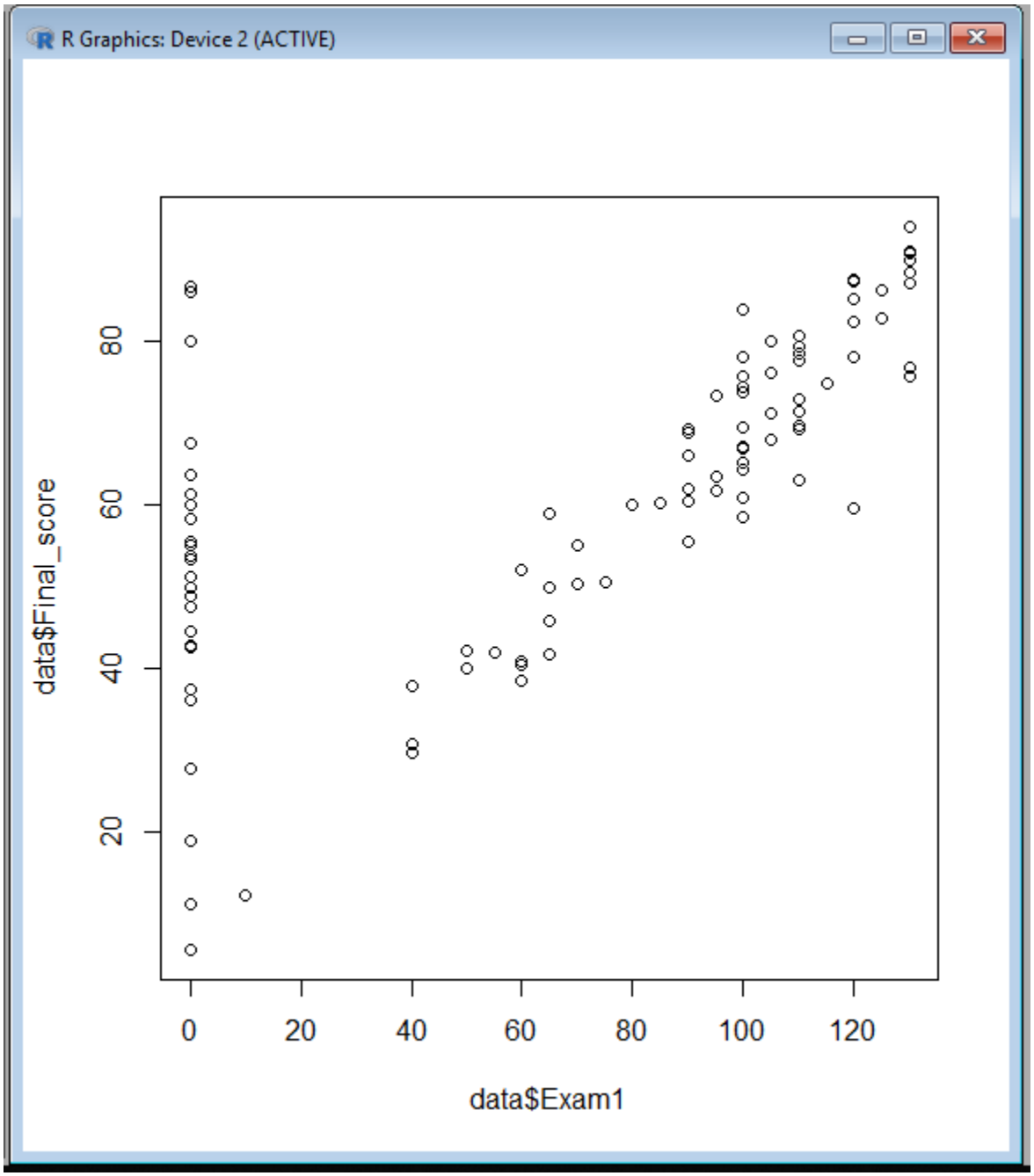
```
R Console
> data= read.csv("D://TYCS21/score.csv")
> data
   Exam1 Exam2 Exam3 Exam4 Final_score Grade
1     60    10    16   7.0       40.79     C
2     90     0     0   0.0       69.23     B
3    130    20    24   1.0       76.75     B
4    130    10    24   8.5       75.66     B
5     90     5    22   9.5       55.48     C
6    100    30    20   3.0       67.11     B
7    105    20    22   8.0       67.98     B
8    120    40    18  16.0       85.09     A
9    120    20    30  18.0       82.46     A
10   130    45    22  10.5       91.01     A
11    90    40    20   7.0       68.86     B
12   130    30    28  10.5       87.06     A
13   100    30    22   6.5       69.52     B
14     0    30    18   0.0       60.00     B
15     0    30    18   0.0       60.00     B
16    80     0    24   3.0       60.11     B
17   105    40    22   6.5       76.10     B
18    10     0     0   8.0       12.16     D
19   130    35    24   0.0       90.00     A
20     0    15    20   7.0       42.86     C
21    40    10    14   6.0       30.70     D
22    90    15    28   8.5       62.06     B
23   110     0    24   9.5       80.62     A
24    65     5    24   1.0       41.67     C
25    55    15    18   0.0       41.90     C
26   100    50    30  11.5       83.99     A
27    95    40    24   8.0       73.25     B
28     0    10    24   0.0       42.50     C
29     0     0    18   0.0       60.00     B
30    65    20    20   0.0       50.00     C
31   110    25    18   6.0       69.74     B
32   130    45    24   8.0       90.79     A
33   120    40    30   9.0       87.28     A
```

#PLOT THE DATASET:

COMMAND:

>plot(x=data$Exam1,y=data$Final_score)

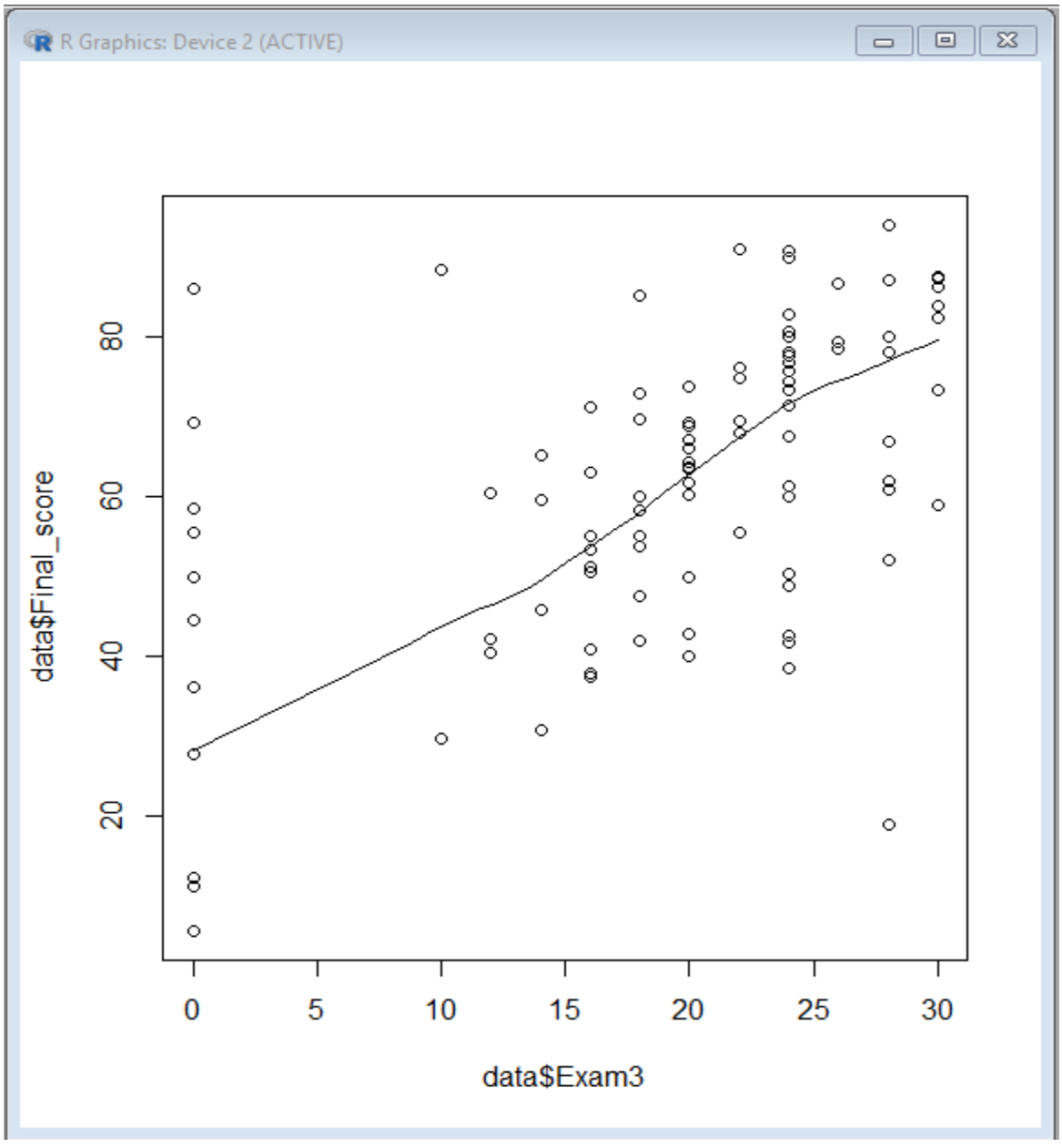
```
> plot(x=data$Exam1,y=data$Final_score)
>
```



#PLOT THE SCATTER DIAGRAM:

>scatter.smooth(x=data$Exam3,y=data$Final_score)

```
> scatter.smooth(x=data$Exam3,y=data$Final_score)
>
```




```
> cor(data$Exam3,data$Final_score)
[1] 0.6046352
>
```

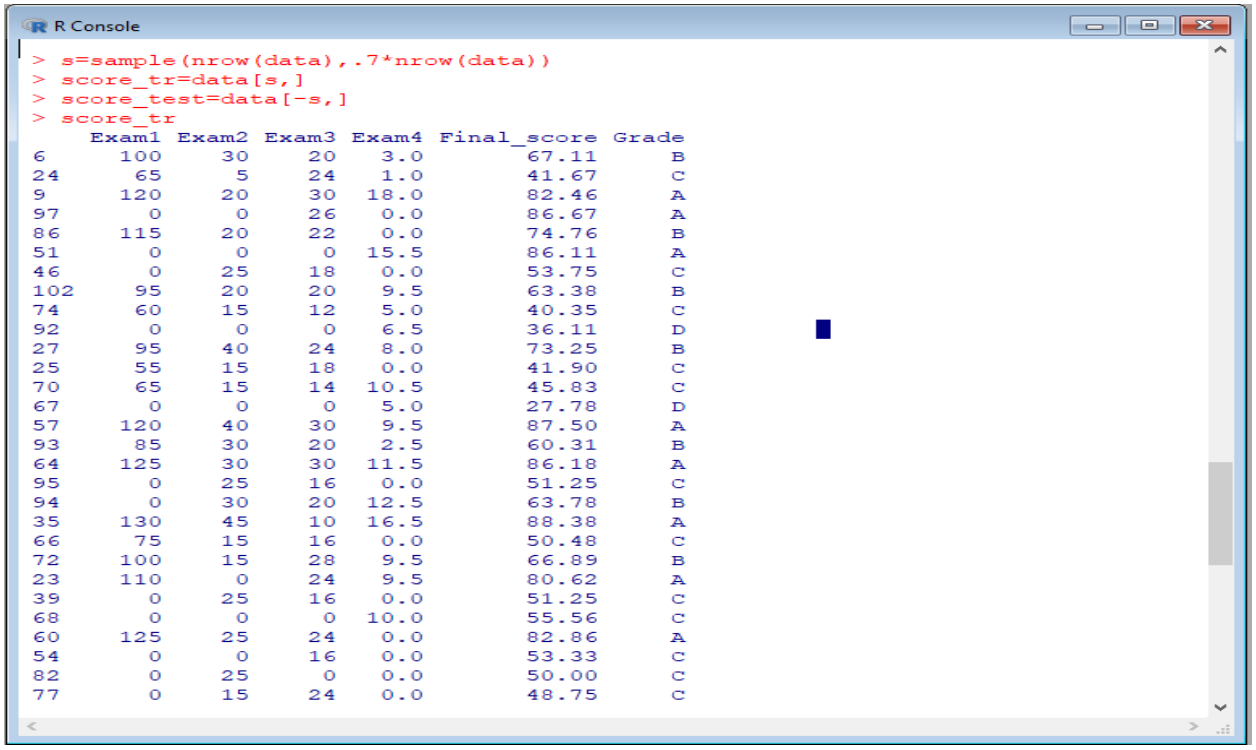
#PARTITIONING THE DATABASE INTO TRAINING AND TESTING SET

>s=sample(nrow(data),.7*nrow(data))

>score_tr=data[s,]

>score_test=[-s,]

Score_tr



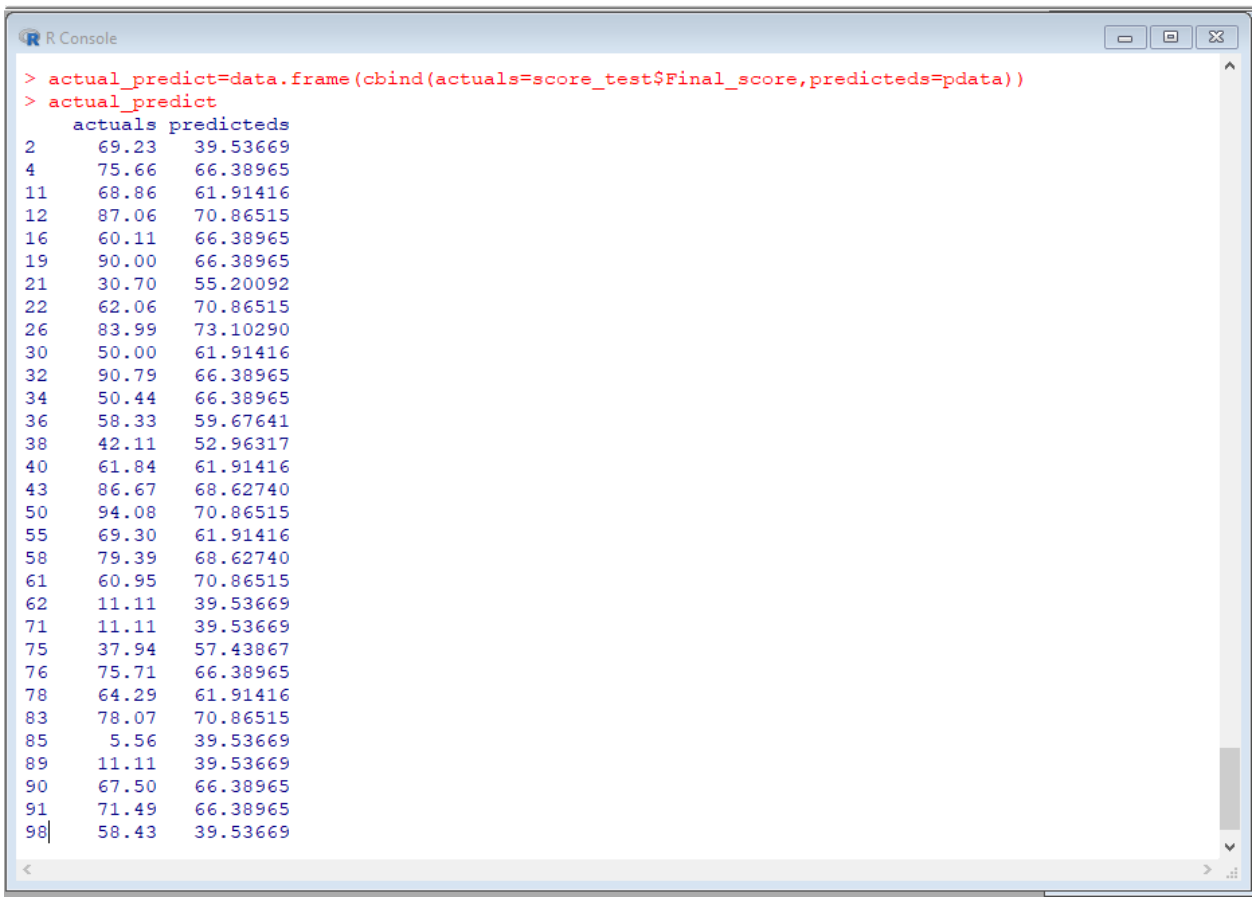
R Console

```
> score_test
    Exam1 Exam2 Exam3 Exam4 Final_score Grade
2      90     0     0   0.0       69.23     B
4     130    10    24   8.5       75.66     B
11     90    40    20   7.0       68.86     B
12    130    30    28  10.5       87.06     A
16     80     0    24   3.0       60.11     B
19    130    35    24   0.0       90.00     A
21     40    10    14   6.0       30.70     D
22     90    15    28   8.5       62.06     B
26    100    50    30  11.5       83.99     A
30     65    20    20   0.0       50.00     C
32    130    45    24   8.0       90.79     A
34     70    20    24   1.0       50.44     C
36      0     0    18  10.0       58.33     C
38     50    30    12   4.0       42.11     C
40     95    20    20   6.0       61.84     B
43      0     0    26   0.0       86.67     A
50    130    40    28  16.5       94.08     A
55    110    25    20   3.0       69.30     B
58    110    35    26  10.0       79.39     B
61    100     0    28   0.0       60.95     B
62      0     0     0   2.0       11.11     D
71      0     0     0   2.0       11.11     D
75     40    20    16  10.5       37.94     D
76    100    35    24   0.0       75.71     B
78    100    15    20   0.0       64.29     B
83    120    20    28  10.0       78.07     B
85      0     0     0   1.0        5.56     D
89      0     0     0   2.0       11.11     D
90      0    30    24   0.0       67.50     B
91    110    25    24   4.0       71.49     B
98    100     0     0   4.0       58.43     C
101   105    30    16  11.5       71.27     B
```

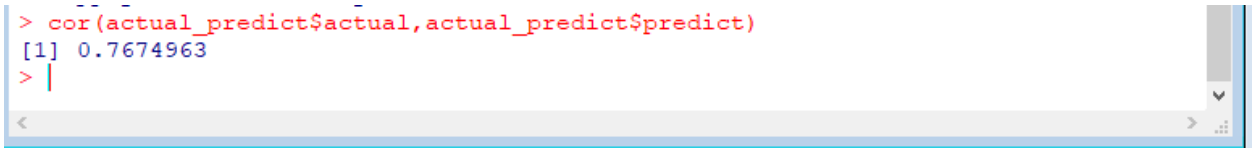
```
> linmon=lm(Final_score~Exam3,data=score_tr)
> print(linmod)
Error in print(linmod) : object 'linmod' not found
> print(linmon)

Call:
lm(formula = Final_score ~ Exam3, data = score_tr)

Coefficients:
(Intercept)        Exam3
     39.537        1.119

> 
```

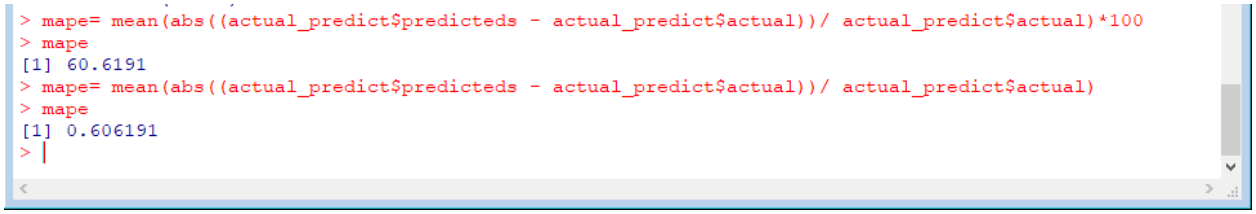
```
> pdata=predict(linmon,score_test)
> summary(linmon)

Call:
lm(formula = Final_score ~ Exam3, data = score_tr)

Residuals:
    Min      1Q  Median      3Q     Max
-52.005  -9.967   1.666  10.500  46.573

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  39.5367     4.8090   8.221 7.15e-12 ***
Exam3         1.1189     0.2362   4.737 1.10e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 16.42 on 70 degrees of freedom
Multiple R-squared:  0.2427,    Adjusted R-squared:  0.2319
F-statistic: 22.44 on 1 and 70 DF,  p-value: 1.101e-05

>
```

#CREATING A MODEL



#PREDICTING THE OUTPUT ON TEST DATASET





Conclusion : Successfully completed the simple and multiple linear regression