

PRACTICAL 2

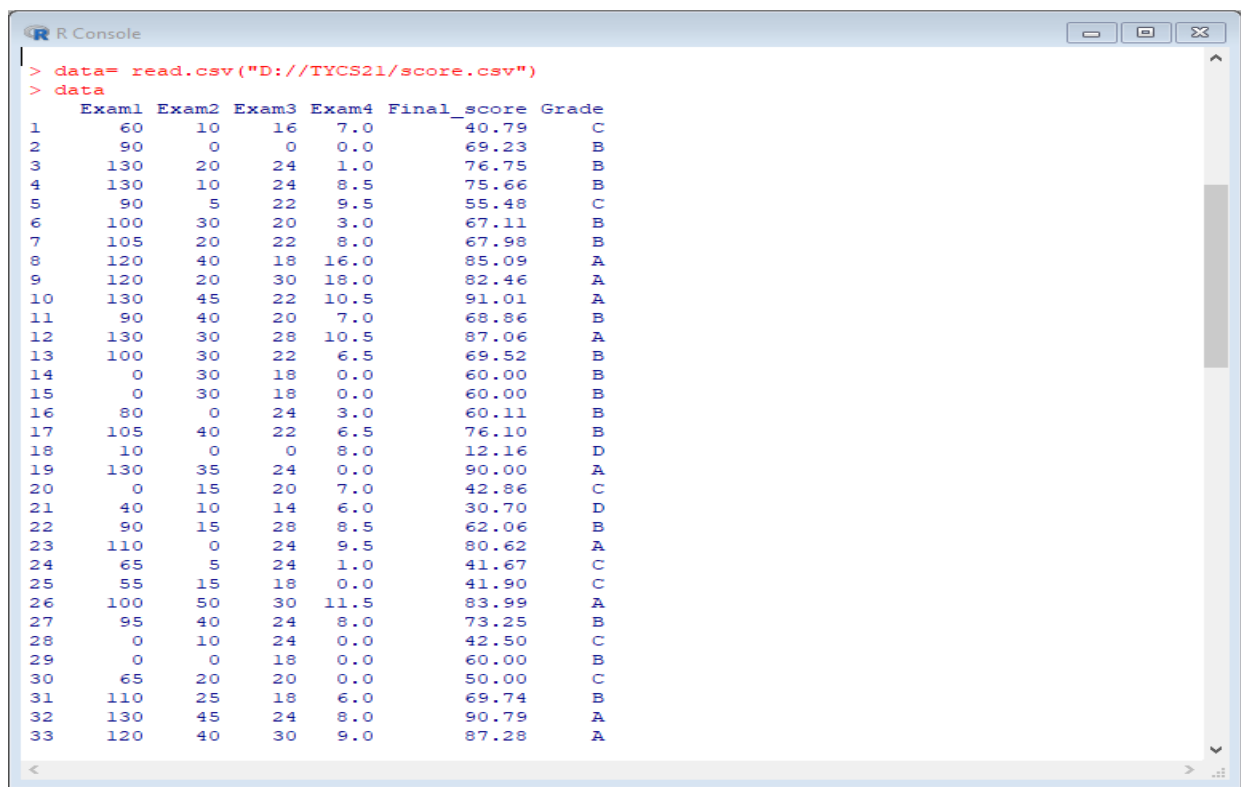
Aim: Simple /Multiple Linear Regressions.

#IMPOER DATASET:

Command:

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

```
>data
```

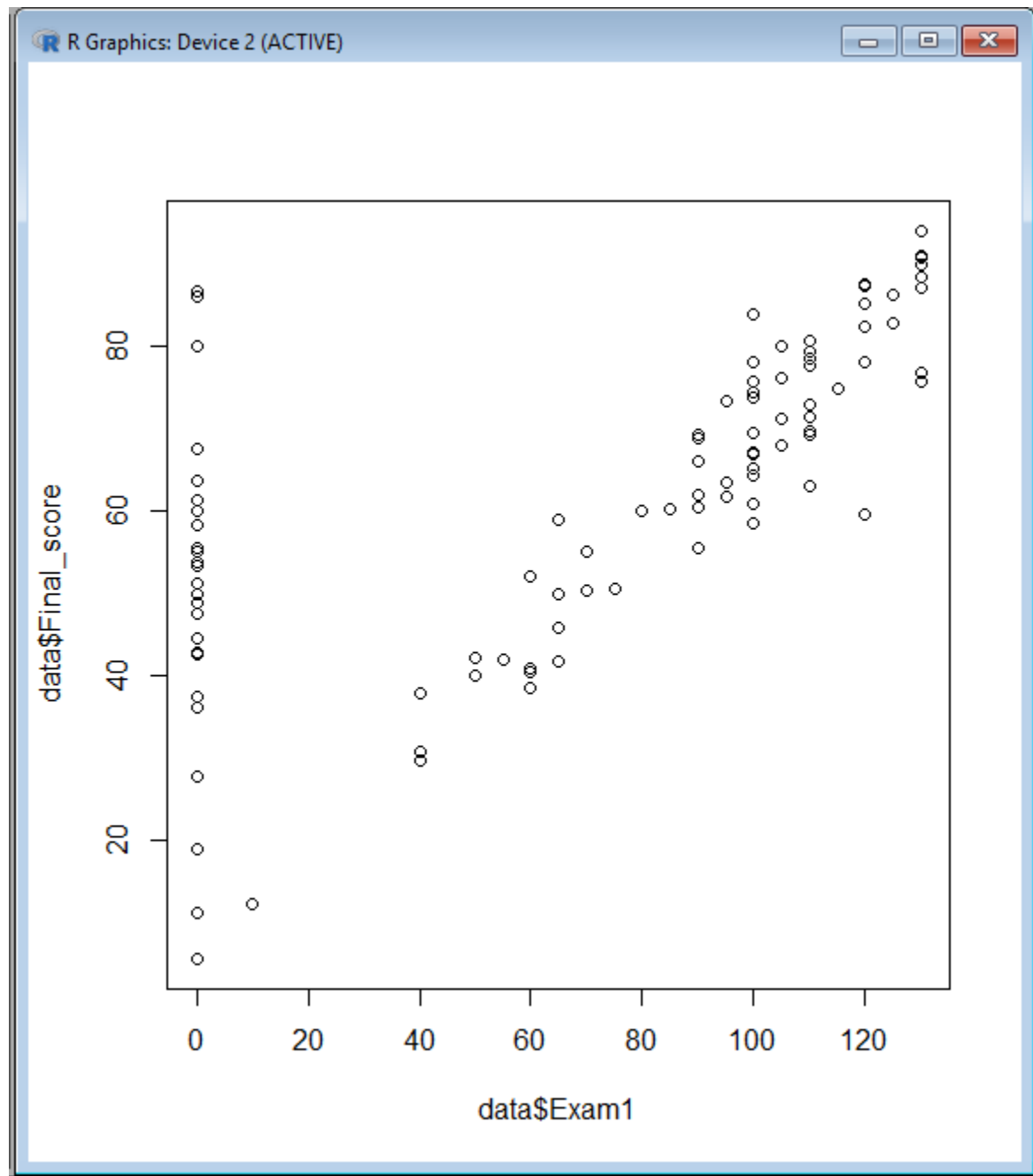


#PLOT THE DATASET:

COMMAND:

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

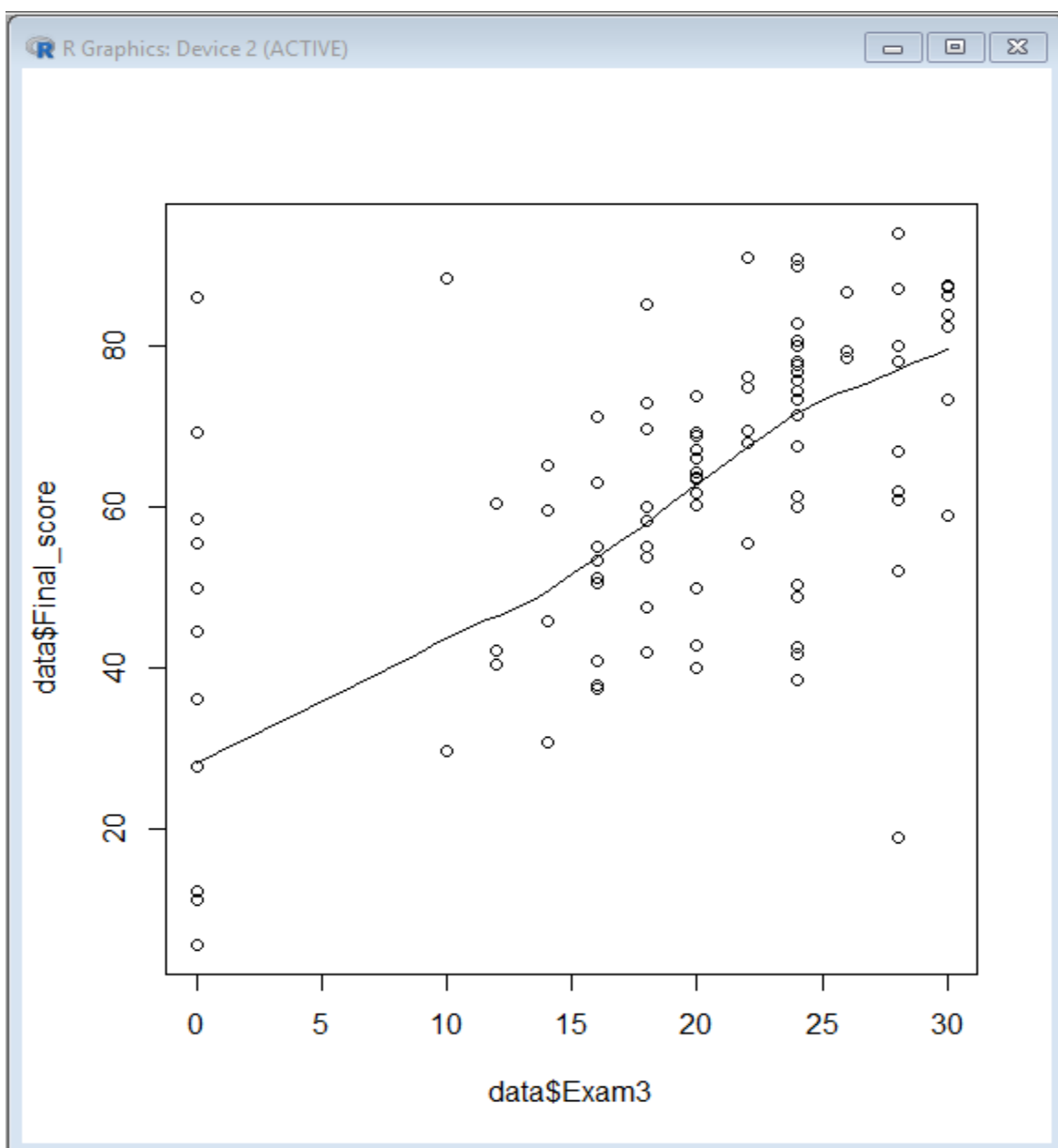
```
> plot(x=data$Exam1,y=data$Final_score)  
> |
```



#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
> s=sample(nrow(data),.7*nrow(data))
> score_tr=data[s,]
> score_test=data[-s,]
> score_tr
```

	Exam1	Exam2	Exam3	Exam4	Final_score	Grade
6	100	30	20	3.0	67.11	B
24	65	5	24	1.0	41.67	C
9	120	20	30	18.0	82.46	A
97	0	0	26	0.0	86.67	A
86	115	20	22	0.0	74.76	B
51	0	0	0	15.5	86.11	A
46	0	25	18	0.0	53.75	C
102	95	20	20	9.5	63.38	B
74	60	15	12	5.0	40.35	C
92	0	0	0	6.5	36.11	D
27	95	40	24	8.0	73.25	B
25	55	15	18	0.0	41.90	C
70	65	15	14	10.5	45.83	C
67	0	0	0	5.0	27.78	D
57	120	40	30	9.5	87.50	A
93	85	30	20	2.5	60.31	B
64	125	30	30	11.5	86.18	A
95	0	25	16	0.0	51.25	C
94	0	30	20	12.5	63.78	B
35	130	45	10	16.5	88.38	A
66	75	15	16	0.0	50.48	C
72	100	15	28	9.5	66.89	B
23	110	0	24	9.5	80.62	A
39	0	25	16	0.0	51.25	C
68	0	0	0	10.0	55.56	C
60	125	25	24	0.0	82.86	A
54	0	0	16	0.0	53.33	C
82	0	25	0	0.0	50.00	C
77	0	15	24	0.0	48.75	C

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

```
R Console
> actual_predict=data.frame(cbind(actuals=score_test$Final_score,predicteds=pdata))
> actual_predict
  actuals predicteds
2    69.23    39.53669
4    75.66    66.38965
11   68.86    61.91416
12   87.06    70.86515
16   60.11    66.38965
19   90.00    66.38965
21   30.70    55.20092
22   62.06    70.86515
26   83.99    73.10290
30   50.00    61.91416
32   90.79    66.38965
34   50.44    66.38965
36   58.33    59.67641
38   42.11    52.96317
40   61.84    61.91416
43   86.67    68.62740
50   94.08    70.86515
55   69.30    61.91416
58   79.39    68.62740
61   60.95    70.86515
62   11.11    39.53669
71   11.11    39.53669
75   37.94    57.43867
76   75.71    66.38965
78   64.29    61.91416
83   78.07    70.86515
85    5.56    39.53669
89   11.11    39.53669
90   67.50    66.38965
91   71.49    66.38965
98   58.43    39.53669
```

#PREDICTING THE OUTPUT ON TEST DATASET

```
> cor(actual_predict$actual,actual_predict$predict)
[1] 0.7674963
> |
```

```
> mape= mean(abs((actual_predict$predicted - actual_predict$actual))/ actual_predict$actual)*100
> mape
[1] 60.6191
> mape= mean(abs((actual_predict$predicted - actual_predict$actual))/ actual_predict$actual)
> mape
[1] 0.606191
> |
```

Conclusion : Successfully completed the simple and multiple linear regression

