Lab no. 16

It was reported somewhere that children whenever playing the game on computer, they use the computer very roughly which may reduce the lifetime of a computer. The random access memory (RAM) of a computer also plays a crucial role in the lifetime of a computer. A researcher wanted to examine how the lifetime of a personal computer that is used by children is affected by the time (in hours) spent by the children per day playing games and the available random access memory (RAM) measured in megabytes (MB) of a used computer. The data is provided in the following table.

Lifetime(years)	Play time(hours)/day	RAM in Mb
5	2	8
1	8	2
7	1	6
2	5	3
3	6	2
4	3	4
6	2	7

A. Write the estimated regression equation for the lifetime.
B. Interpret the parameters of the regression model.
C. What percentage of variation in lifetime is explained by two independent variables?
D. Compute the standard error of the estimate.
E. Also compute adjusted R square.
F. Test the significance of each of the regression coefficients.
G. Test the overall goodness of fit of the model.

Solution:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.939861914
R Square	0.883340416
Adjusted R Square	0.825010625
Standard Error	0.903668681
Observations	7

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	24.73353166	12.36676583	15.143898	0.013609458
Residual	4	3.266468338	0.816617085		
Total	6	28			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*
Intercept	6.961325967	2.481625648	2.805147493	0.048556218	0.071228582
X Variable 1	-0.785380365	0.29455627	-2.666316917	0.056020402	-1.603199679
X Variable 2	0.014874628	0.307243511	0.048413156	0.963707852	-0.838170114

Let the regression equation be: Y=a+b1x1+b2x2
From the coefficient table,

A. Y= 6.91 – 0.78x1+0.01x2.

Regression

Lab no 15:

A computer manager needs to know how the efficiency of her new computer program depends on the size of incoming data and how many tables are used to arrange each data set. Efficiency will be measured by the number of processed requests per hour. Applying the program to data sets of different sizes and number of tables, she gets the following results.

Processed requests Y	Data size, (GB), X_1	Number of tables, X_2
16	15	1
26	10	10
41	8	10
50	7	20
55	7	20
40	6	4

A. **Write the regression equation for the processed request.**
B. **Interpret the parameters of the regression model.**
C. **What percentage of variation on processed requests is explained by two independent variables?**
D. **Compute the standard error of the estimate.**
E. **Also compute adjusted R square.**
F. **Test the significance of each of the regression coefficients.**
G. **Test the overall goodness of fit of the model.**

Solution

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.954350535
R Square	0.910784943
Adjusted R Square	0.851308238
Standard Error	5.651459145
Observations	6

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	978.1830286	489.0915143	15.31330537	0.026647547
Residual	3	95.8169714	31.93899047		
Total	5	1074			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*
Intercept	51.56781896	11.44130651	4.507161739	0.020403002	15.15647532
X Variable 1	-2.627727391	0.931963619	-2.819560054	0.066762648	-5.593651566
X Variable 2	0.890194431	0.390179791	2.281498046	0.106787936	-0.351531803

Let the regression equation be : Y=a+b1x1+b2x2
From the coefficient table,

a. **Y=51.56-2.62x1+0.89x2 which is a required equation.**

b. **Here,**
a= 51.56 i.e. if x1 and x2 become zero then efficiency becomes 51.56.
b1=-2.62 i.e. if we increase the data size by one unit then efficiency decreases by 2.62 units keeping the effect of several tables as constant.
B2=0.89 i.e. if we increase the number of tables by one unit then efficiency increases by 0.89 keeping the effect of data size constant.

Lab no. 17

1. **The Following table represents the layout of CRD of four treatments.**

A (9)	B (14)	D (11)	C (10)
D (8)	A (14)	B (13)	C (16)
B (7)	C (12)	D (5)	A (11)
A (14)	B (12)	C (6)	D (5)

At 5% level of significance test whether there is significant difference between mean of 4 treatments.

Solution:

Hypothesis:

H_0: There is no significant difference between treatment.

H_1: There is significant difference between treatment.

Alpha = 5%

Test statistics:

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
A	4	48	12	6
B	4	46	11.5	9.666667
C	4	44	11	17.33333
D	4	39	9.75	18.25

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups(treatment)	11.1875	3	3.729167	0.291057	0.831069	3.490295
ERROR	153.75	12	12.8125			
Total	164.9375	15				

Decision,
Since $F_{cal} < F_{tab}$ so we accept H_0.

Hence we conclude that there is no significance difference between mean of 4 treatments.

Lab no 18:

2. **There are three brands of computers namely dell, Lenovo, and HP. The following are the lifetime of 15 computers in years.**

Serial Number	Computer Brand	Lifetime in Years
1	Dell	15
2	Lenovo	10
3	HP	9
4	Dell	12
5	Lenovo	6
6	HP	7
7	Dell	4
8	Lenovo	8
9	HP	13
10	Dell	11
11	HP	5
12	Lenovo	7
13	Dell	3
14	HP	5
15	Lenovo	4

Apply appropriate statistical tests to identify whether the average lifetime (in years) is significantly different across three brands of computers at a 5% level of significance. You can again tabulate the data initially in the required format for statistical analysis.

Hypothesis:

H_0: There is no significance difference among average lifetime of three brands of computers.

H_1: There is significance difference among average lifetime of three brands of computers.

Alpha = 5%

Test statistics:

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Row 1	5	45	9	27.5
Row 2	5	35	7	5
Row 3	5	39	7.8	11.2

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	10.13333333	2	5.066666667	0.347826087	0.713111407	3.885294
ERROR	174.8	12	14.56666667			
Total	184.9333333	14				

Decision,

$F_{cal} < f_{tab}$, so we accept H_0,

Hence we conclude that There is no significance difference among average lifetime of three brands of computers.