ASSIGNMENT (REGRESSION)

SUBMITTED BY "SAMIYA JABBAR"

(Walpole) Eleven student teachers took part in an evaluation program designed to measure teacher effectiveness and determine what factors are important. The response measure was a quantitative evaluation of the teacher. The regressor variables were scores on four standardized tests given to each teacher. The data are as follows

y	*x* 1	*x* 2	*x* 3	*x* 4
410	69	125	59	55.66
569	57	131	31.75	63.97
425	77	141	80.5	45.32
344	81	122	75	46.67
324	0	141	49	41.21
505	53	152	49.35	43.83
235	77	141	60.75	41.61
501	76	132	41.25	64.57
400	65	157	50.75	42.41
584	97	166	32.25	57.95
434	76	141	54.5	57.9

Applying the multiple regression model using Excel ,so output will be as follows

SUMMARY OUTPUT 1

Regression Statistics	
Multiple R	0.85
R Square	0.73
Adjusted R Square	0.55
Standard Error	71.15
Observations	11.00

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4.00	81329.21	20332.30	4.02	0.06
Residual	6.00	30371.70	5061.95		
Total	10.00	111700.91			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-884.67	748.68	-1.18	0.28	-2716.62	947.28
x1	-0.84	1.41	-0.59	0.57	-4.30	2.62
x2	4.91	2.94	1.67	0.15	-2.28	12.09
x3	1.33	3.23	0.41	0.69	-6.57	9.23
x4	11.93	5.60	2.13	0.08	-1.77	25.64

Because P-value of x3 is greater than 0.05 and highest among others so we omit it first, and again run the model

y	*x* 1	*x* 2	*x* 4
410	69	125	55.66
569	57	131	63.97
425	77	141	45.32
344	81	122	46.67
324	0	141	41.21
505	53	152	43.83
235	77	141	41.61
501	76	132	64.57
400	65	157	42.41

584	97	166	57.95
434	76	141	57.9

SUMMARY OUTPUT 2

Regression Statistics	
Multiple R	0.85
R Square	0.72
Adjusted R Square	0.60
Standard Error	66.80
Observations	11.00

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3.00	80469.31	26823.10	6.01	0.02
Residual	7.00	31231.59	4461.66		
Total	10.00	111700.91			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-602.69	285.48	-2.11	0.07	-1277.74	72.36
x1	-0.42	0.93	-0.45	0.66	-2.61	1.77
x2	3.94	1.65	2.39	0.05	0.04	7.84
x4	9.92	2.60	3.82	0.01	3.78	16.07

The P-value of x1 is greater than 0.05, so we omit this ,and run again the regression model

y	***x*** 2	***x*** 4
410	125	55.66
569	131	63.97
425	141	45.32
344	122	46.67
324	141	41.21
505	152	43.83
235	141	41.61
501	132	64.57
400	157	42.41
584	166	57.95
434	141	57.9

SUMMARY OUTPUT FINAL

Regression Statistics	
Multiple R	0.84
R Square	0.71
Adjusted R Square	0.64
Standard Error	63.40
Observations	11.00

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2.00	79547.90	39773.95	9.90	0.01
Residual	8.00	32153.01	4019.13		
Total	10.00	111700.91			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-586.57	268.85	-2.18	0.06	-1206.54	33.40
x2	3.80	1.54	2.47	0.04	0.25	7.34
x4	9.45	2.26	4.18	0.00	4.23	14.67

Now all the p-values lie under our limit i.e <=0.05 (allow intercept values because of 10% margin) so, Eqn. will be

$y = -586.57 + 0x_1 + 3.80x_2 + 0x_3 + 9.45x_4$ OR $y = -586.57 + 3.80x_2 + 9.45x_4$

(shows Y is dependent on intercept, x2 & x4 only)