<u>PubH 7405</u>: BIOSTATISTICS REGRESSION, 2010
PRACTICE PROBLEMS FOR MULTIPLE LINEAR REGRESSION
(Some are new & Some from Old exams; #10 was 2010 exam)

<u>Problem 1:</u>

There are two measures of the maximum amount of air that can be exhaled after a maximum inhalation. One is the "forced expiratory volume" – or FEV, the forced expiratory volume in the first second is the FEV1 (also called forced vital capacity); the other "slow expiratory volume, or briefly the "vital capacity" (VC). They are measured using spirometry; spirometry (meaning *the measuring of breath*) is the most common of the Pulmonary Function Tests measuring lung function, specifically the measurement of the amount (volume) and/or speed (flow) of air that can be inhaled and exhaled. The difference between VC and the total Lung Capacity depends on the general condition of lung tissue. The FEV is lower than the VC either in normal individuals or in obstructive patients; the level of difference increases with the degree of obstruction. The following Table shows data observed from 10 hospitalized patients.

X1, VC(liters)	X2, Total Lung Capacxity (liters)	Y, FEV (liters/second)
2.2	2.5	1.6
1.5	3.2	1.0
1.6	5.0	1.4
3.4	4.4	2.6
2.0	4.4	1.2
1.9	3.3	1.3
2.2	3.2	1.6
3.3	3.3	2.3
2.4	3.7	2.1
0.9	3.6	0.7

(1) Fit the regression model of Y on X1 and X2 and interpret the results (values of the estimated regression coefficients).
(2) Calculate the coefficient multiple determination and interpret the result.
(3) Obtain the ANOVA table that decomposes the Sum of Squares Regression (SSR) into (a) Sum of Squares associated with X1 and extra Sum of Squares associated with X2 given X1, and (b) Sum of Squares associated with X1 and extra Sum of Squares associated with X2 given X1. Is it true that:
SSR(X1) + SSR(X2|X1) = SSR(X2)+SSR(X1|X2)?
(4) Taken collectively, do total lung capacity and vital capacity contribute significantly to the prediction of forced expiratory volume?
(5) Does total lung capacity have any value added to the explanation of the variation in forced expiratory volume over and above that achieved by Vital capacity? (i.e. Testing whether X2 can be dropped from the regression model given that X1 is retained)
(6) Calculate the coefficient of partial determination between Y and X2 measuring the marginal reduction in the variation of Y associated with the addition of X2, when X1 is in the model, and give your interpretation.
(7) Plot the residuals against values of each of the two predictors; what do the plots suggest, any clear departures from the model?
(8) Form the added-variable plot to see if "the part of X2 not contained in X1" can further explained "the part of Y not explained by X1" (Should X2 be added to the regression model? If so, any transformation is necessary?)

(9) Examine the correlation between X1 and X2; any sign of multi-collinearity supported by this correlation?
(10) Perform a stepwise regression and explain the results.

Problem 2:

Much research has been devoted to the etiology of hypertension. One general problem is to determine to what extent hypertension is a genetic phenomenon, whether we could explain variation in blood pressure among children from the blood pressures of their parents, especially the mothers. The following Table shows data observed from 20 families with systolic blood pressures from the mother, father, and the first-born child.

X1, Mother SBP	X2, Father SBP	Y, Child SBP	X1, Mother SBP	X2, Father SBP	Y, Child SBP
130.0	140.0	90.0	125	150.0	100
125.0	120.0	85.0	110	125.0	80
140.0	180.0	120.0	90	140.0	70
110.0	150.0	100.0	120	170.0	115
145.0	175.0	105.0	150	150.0	90
160.0	120.0	100.0	145	155.0	90
120.0	145.0	110.0	130	160.0	115
180.0	160.0	140.0	155	115.0	110
120.0	190.0	115.0	110	140.0	90
130.0	135.0	105.0	125	150.0	100

(1) We would expect from genetic principles that the correlation between the mothers's SBP and the child's SBP exist; can we test this expectation at the level of 0.05? Is there any other biological explanation besides the genetic principles?
(2) Fit the regression model of Y on X1 and X2 and interpret the results (values of the estimated regression coefficients).
(3) Calculate the coefficient multiple determination and interpret the result.
(4) Obtain the ANOVA table that decomposes the Sum of Squares Regression (SSR) into (a) Sum of Squares associated with X1 and extra Sum of Squares associated with X2 given X1, and (b) Sum of Squares associated with X1 and extra Sum of Squares associated with X2 given X1. Is it true that:
SSR(X1) + SSR(X2|X1) = SSR(X2)+SSR(X1|X2)?
(5) Taken collectively, do blood pressures of the parents contribute significantly to the prediction of their first-born child blood pressure?
(6) Does the father's SBP have any value added to the explanation of the variation in first-born children blood pressures over and above that achieved by mother's SBP? (i.e. Testing whether X2 can be dropped from the regression model given that X1 is retained)
(7) Calculate the coefficient of partial determination between Y and X2 measuring the marginal reduction in the variation of Y associated with the addition of X2, when X1 is in the model, and give your interpretation.
(8) Plot the residuals against values of each of the two predictors; what do the plots suggest, any clear departures from the model?
(9) Form the added-variable plot to see if "the part of X2 not contained in X1" can further explained "the part of Y not explained by X1" (Should X2 be added to the regression model? If so, any transformation is necessary?)
(10) Examine the correlation between X1 and X2, is the lack of correlation expected? any sign of multi-collinearity supported by this correlation? Perform a stepwise regression and explain the results.

Problem 3:

Let FEV (a measure of Lung Health) be the dependent variable and Age, Height, and Weight are three potential predictors; we have the following results:

Model 1A: FEV = Age/

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	2.2573846	2.257385	3.111695	0.091012196
Residual	23	16.685391	0.725452		
Total	24	18.942776			

Model 1B: FEV = Height/

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	10.28899	10.28899	27.34604	2.64568E-05
Residual	23	8.653786	0.376252		
Total	24	18.94278			

Model 1C: FEV = Weight/

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	6.031565	6.031565	10.74462	0.003301129
Residual	23	12.91121	0.561357		
Total	24	18.94278			

Model 2A: FEV = Age Height/

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	12.62367	6.311837	21.97471	5.69284E-06
Residual	22	6.319101	0.287232		
Total	24	18.94278			

Model 2B: FEV = Age Weight/

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	8.485928	4.242964	8.926707	0.00145054
Residual	22	10.45685	0.475311		
Total	24	18.94278			

Model 2C: FEV = Height Weight/

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	10.47598	5.237991	13.61032	0.000142242
Residual	22	8.466794	0.384854		
Total	24	18.94278			

Model 3: FEV = Age Height Weight/

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	12.767	4.255666	14.47088	2.4569E-05
Residual	21	6.175779	0.294085		
Total	24	18.94278			

Model 4: FEV = Age Height Weight Age*Age Height*Height Weight*Weight/

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	6	15.07969	2.513281	11.71059	2.23806E-05
Residual	18	3.863089	0.214616		
Total	24	18.94278			

(1) Is Height predictive of FEV? Do you have enough information to calculate Coefficient of Correlation measuring the strength of that relationship?

(2) Does the addition of Height, as an independent variable, add significantly to the prediction of FEV over and above that achieved by Age?"

(3) Does the addition of Height, as an independent variable, add significantly to the prediction of FEV over and above that achieved by Age and Weight?"

(4) Suppose we have Model 1A and are considering adding either Height or Weight to that model. Calculate the Coefficients of Partial Determination associated with each of the two potential predictors; then use the results to decide which factor would add *more* significantly to the prediction of FEV over and above that achieved by Age.

(5) Does the addition of the three quadratic terms, as independent variables, add significantly to the prediction of FEV over and above that achieved by Age, Height, and Weight?"

Problem 4:

Consider a multiple regression model with two predictors:

$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

Form the Sum of Squared Errors, derive the Normal Equations, and then use the three normal equations to show that the Fitted Value and the Residual are uncorrelated.

Problem 5:

(1) In a multiple regression model of Y on (X_1, X_2, X_3, X_4}, what does the Intercept β_0 represent?

(2) Let the response Y be "FEV" (a lung health measure) and potential predictors consist of Age, Gender (Male/Female), Race (White/Black/Others), and Weight. What should we do so as to have the Intercept representing the "Mean FEV" of a white woman who is 27 years old and weights 129 pounds?

(3) Explain each of the following terms (what it is and/or what it does):
   a) Design Matrix for regression model without intercept
   b) Coefficient of Partial Determination due to X_2 in regression model of Y on (X_1, X_2, X_3}
   c) DFBETAS
   d) Correlation Transformation
   e) Type I Sum of Squares
   f) First-order autoregressive error model

Problem 6:

In an assay of Cod-liver oil (test) against Vitamin D3 (standard), let:

Y = Response

P = Preparation (=0 for Standard, =1 for Test)
X = log(Dose)
PX = X*P

(1) From the following results:

Term	Regression Coefficient	Standard Error	"t" Statistic	"p" Value
Intercept	24.1098	3.6601	6.59	<.0001
X	5.8398	1.5663	3.73	0.0011
P	-16.5821	5.5459	-2.99	0.0065
PX	1.2001	1.8302	0.66	0.5185

Is it reasonable to assume that the response is linearly related to log of the dose, why or why not? (Hint: why does log scale have anything to do with the product term?)

(2) From the following results (of a different model):

Term	Regression Coefficient	Standard Error	"t" Statistic	"p" Value
Intercept	22.0844	1.9399	11.38	<.0001
X	6.7188	0.8005	8.39	<.0001
P	-13.1556	1.8347	-7.17	<.0001

Is the strength of the Test Preparation stronger or weaker than the strength of the Standard Preparation? Find a point estimate of the log relative potency (call it “m”).

(3) Given the following results from using option “COVB”:

Term	Intercept	X	P
Intercept	3.7631	-1.4767	2.6808
X	-1.4767	0.6409	-1.3197
P	2.6808	-1.3197	3.3659

(a) Why is this matrix symmetric?
(b) Find an approximate 95% Confidence Interval for the Intercept
(c) Show “how” you would do to obtain the value of the Standard Error of the estimated log relative potency “m” (in question b above); numerical answer is not very important!

Problem 7:

In order to predict Survival in patients (Y: Survival time) undergoing some liver operation, we consider the following four potential predictors:

X_1: Blood clotting score
X_2: Some prognostic index
X_3: Enzyme function test result
X_4: Liver function test result

We fitted and obtained the following four optimal models of different size.

Best 1-variable Model:

	df	*SS*	*MS*	*F*
Regression	1	583808.873	583808.873	56.731
Residual	52	535122.627	10290.820	
Total	53	1118931.500		

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-71.921	38.300	-1.878	0.066
X4	98.055	13.018	7.532	<0.001

Best 2-variable Model:

ANOVA

	df	*SS*	*MS*	*F*
Regression	2	737504.604	368752.302	49.305
Residual	51	381426.896	7478.959	
Total	53	1118931.500		

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-424.564	63.750	-6.660	<0.001
X2	4.883	0.703	6.945	<0.001
X3	4.058	0.559	7.259	<0.001

Best 3-varable Model:

ANOVA

	df	*SS*	*MS*	*F*
Regression	3	931546.037	310515.346	82.855
Residual	50	187385.463	3747.709	
Total	53	1118931.500		

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-659.179	55.674	-11.840	<0.001
X1	38.323	5.326	7.196	<0.001
X2	4.568	0.500	9.144	<0.001
X3	4.485	0.400	11.208	<0.001

Best 4-variable Model:

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	936264.538	234066.134	62.788	0.000
Residual	49	182666.962	3727.897		
Total	53	1118931.500			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*
Intercept	-621.598	64.800	-9.592	0.000	-751.8189
X1	33.164	7.017	4.726	0.000	19.06209
X2	4.272	0.563	7.582	0.000	3.139696
X3	4.126	0.511	8.071	0.000	3.098522
X4	14.092	12.525	1.125	0.266	-11.07903

(1) How much does Liver function test result help in explaining the variation among survival times?
(2) Does the addition of Blood clotting score, as an independent variable, add significantly to the prediction of patients' Survival time over and above that achieved by Prognostic index and Enzyme function test result? Does the t-test agree with the F-test?
(3) Does the addition of Blood clotting score and Liver function test result, as independent variables, add significantly to the prediction of patients' Survival time over and above that achieved by Prognostic index and Enzyme function test result?
(4) If we used, in different computer runs, Type II then Type III Sum of Squares, would we get the same two sets of results, why or why not?

Problem 8:

The purpose of this study was to examine the data for 44 physicians working for an emergency room at a major hospital so as to determine which of the following four independent variables, if any, are related to the number of complaints received during the previous year. In addition to the number of complaints, served as the "response variable", data available consist of two continuous independent variables, the revenue (dollars per hour) and work load at the emergency service (hours) and two binary variables, gender (1=Female/0=Male) and residency training in emergency services (0=No/1=Yes). Because "the number of complaints in a year" would follow the Poisson distribution, a number of assumptions of the "error normal regression model" would be violated, I took the square root of that number - denoted by Y - and used as the dependent variable; this transformation would stabilize the variance and improves normality. I then fitted 4 regression models with the following results:

Model 1: All 4 independent variables: Revenue, Gender, Residency, and Hours

Regression Statistics	
Multiple R	0.5499804
R Square	0.3024784
Adjusted R Square	0.2309378
Observations	44

ANOVA

	df	*SS*	*MS*	*F*	*p-value*
Regression	4	6.91983886	1.72996	4.22806	0.00614
Residual	39	15.9572928	0.409161		
Total	43	22.87713166			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-0.6965082	0.928866193	-0.74985	0.45785
Revenue	0.0033458	0.003340798	1.00149	0.32276
Gender	-0.210065	0.228204671	-0.92051	0.36296
Residency	-0.0468979	0.238624724	-0.19653	0.84521
Hours	0.0011178	0.000334052	3.346055	0.00182

Model 2: Two independent variables: Residency and Hours

Regression Statistics	
Multiple R	0.50746603
R Square	0.25752177
Adjusted R Square	0.22130332
Observations	44

	df	*SS*	*MS*	*F*	*p-value*
Regression	2	5.891359401	2.9456797	7.110237	0.00223
Residual	41	16.98577226	0.4142871		
Total	43	22.87713166			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	0.06845447	0.526779053	0.1299491	0.897242
Residency	0.00484189	0.214823797	0.0225389	0.982127
Hours	0.00113512	0.000330914	3.4302508	0.001388

Model 3: Only Hours (Simple Regression)

Regression Statistics	
Multiple R	0.507456963
R Square	0.257512569
Adjusted R Square	0.239834297
Observations	44

	df	*SS*	*MS*	*F*	*p-value*
Regression	1	5.891148942	5.8911489	14.566614	0.00044
Residual	42	16.98598272	0.4044282		
Total	43	22.87713166			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	0.075104188	0.431183894	0.1741813	0.8625598
Hours	0.001131979	0.000296592	3.8166234	0.0004384

Model 4: Two independent variables: Revenue and Hours

Regression Statistics	
Multiple R	0.532502222
R Square	0.283558616
Adjusted R Square	0.248610256
Observations	44

	df	*SS*	*MS*	*F*	*p-value*
Regression	2	6.487007791	3.2435039	8.1136458	0.0011
Residual	41	16.39012387	0.3997591		
Total	43	22.87713166			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-0.884573898	0.895352443	-0.9879617	0.3289663
Revenue	0.003609432	0.002956422	1.2208787	0.2291067
Hours	0.00114659	0.000295118	3.885197	0.0003655

(1) Refer to the computer output/results for Model 1:
   a) Verify the values of "R Square" (.5499804) and "Adjusted R Square" (.2309378) using numerical results in the ANOVA table.
   b) Give your interpretation for the value of "Multiple R" (.3024784)
   c) Use the p-value for the F-test and the p-values for the four t-tests, give your interpretation concerning the research question "which of the four independent variables, if any, are related to the number of complaints received during the previous year?".

(2) Start with Model 3 using "Hours" as the only predictor. Suppose you have to choose between "Residency" and "Revenue" to add in and form a multiple model with two predictors, which would you select and why?

(3) Does "Residency" (whether the physician had residency training in emergency services or not) add any value to explain the number of complaints beyond and above what already contributed by the number work hour? Why or why not?

(4) Refer to the computer output/results for Model 3, how are the F-test and the t-test related? How are the F-test and the four t-tests of Model 1 related?

(5) Does the addition of three factors (Residency, Gender, and Revenue) add significantly to the prediction or explanation of the number of complaints beyond and above what already contributed by the number work hour? Why or why not?

(6) What data transformation (on the independent variables) should I do so that, after fitting Model 1, the "intercept" would give me the average number of complaints (on the square root scale) of a woman physician who has no residency training (in Emergency Medicine), worked 1,500 hours last year, and made $300 an hour.

(7) Define the terms "Correlation Transformation" and "VIF" (for each, what it is and what it does).

(8) In the effort to explain the number of complaints, besides the four factors mentioned above (Hours, Residency, Gender, and Revenue) do you think of any other factor that we should consider and why?

(9) Do you think we could draw some inference on whether the revenue (dollars per hour) and work load at the emergency service (hours) are related given only the computer output we have above? Which would be the basis for that inference?

<u>Problem 9:</u>

Let X (Birth weight) and Y (Growth of the infant during the third month of life expressed as percentage of the birth weight), we have the following summarized data:

$n = 12$, $\Sigma x = 1{,}207$, $\Sigma y = 975$, $\Sigma x^2 = 123, 571$, $\Sigma y^2 = 86{,}487$, and $\Sigma xy = 94, 322$.

Calculate the entries of the matrices: **X'X**, **(X'X)$^{-1}$**, **X'Y**, and **B**; and from these matrices, give the point estimates of the intercept and the slope of the regression line.

<u>Note:</u> In the last problem, Problem 10, you are provided with the problem overall goal and a data set but no questions. . Play around with your scientific curiosity; form your own specific aims and work to answer your own questions.

Problem 10:

An investigator is interested to see to what extent "height" is genetically determined, whether we could explain variation in heights of different people from the heights of their parents. The following Table shows the height Y at age 18 for a random sample of 20 males. Factors considered are:
X1 = Length at birth,
X2 = Mother's height when she was 18,
X3 = Father's height when he was 18;
All heights and lengths are in inches.

Y	X1	X2	X3
67.2	19.7	60.5	70.3
69.1	19.6	64.9	70.4
67.1	19.4	65.4	65.8
72.4	19.4	63.4	71.9
63.6	19.7	65.1	65.1
72.7	19.6	65.2	71.1
68.5	19.8	64.3	69.7
69.8	19.7	65.3	68.8
68.4	19.7	64.5	68.7
72.4	19.9	63.4	70.3
67.5	21.9	60.3	70.4
70.2	20.3	64.9	68.8
69.8	19.7	63.5	70.3
63.6	19.9	67.1	65.5
64.3	19.6	63.5	65.2
68.5	21.3	66.1	65.4
70.5	20.1	64.8	70.2
68.1	20.2	62.6	68.6
73.3	20.8	66.2	70.3
66.2	19.3	62.4	67.5

We also calculated and have available:
X4 = X2 + X3
No new computer work are needed nor allowed; use computer output in subsequent pages to answer these questions; in each case, show your work/explanations to justify your answers:

(1) Form the 95% Confidence Interval for the Coefficient of Correlation measuring the strength of the relationship between Y (Height at age 18) and X1 (Length at birth).
(2) From the result in (1), can we decide if a man's Height and his Length at birth related? Calculate the Intercept and the Slope of the simple linear regression of Y on X1.
(3) Taken collectively, does the entire set of 3 independent variables X1, X2, and X3 contribute significantly to the prediction of the Dependent Variable Y? Calculate the values of "R Square" and "Adjusted R Square"
(4) Which of the three factors (X1 or X2 or X3) stand out as the most significant factor in the prediction of a man's Height at age 18 over and above that achieved by the other two factors?
(5) Does the addition of a man's Length at birth add significantly to the prediction of his own Height at age 18 over and above that achieved by the two genetic factors (Mother's height when she

was 18 and Father's height when he was 18)? Calculate the corresponding Coefficient of Partial Determination.

(6) Does the addition of the two genetic factors, Mother's height when she was 18 and Father's height when he was 18, add significantly to the prediction of a man's Height at age 18 over and above that achieved by his Length at birth? Calculate the corresponding Coefficient of Partial Determination.

(7) Is that scientifically reasonable to conclude that ,Mother's height when she was 18 and Father's height when he was 18 contribute equally to improving the prediction of a man's Height at age 18 over and above that achieved by his Length at birth?

(8) Using only the original 3 predictors (X1, X2, X3), perform a stepwise regression; provide and explain the final model.

(9) Suppose we perform multiple regression analysis for 2 models: (1) Model A: Response is Y and two predictors are X2 and X3; and (2) Model B: Response is X1 and the two predictors are also X2 and X3; residuals from both models were obtained and saved but computer results are not given below. Find the Pearson's Coefficient of Correlation measuring the strength of the relationship between these two sets of residuals.

(10) Explain each of the following terms (what it is and what it does): (a) Added value plot, (b) DFBETAS, (c) Correlation Transformation, (d) Type I Sum of Squares, and (e) Durbin-Watson test statistic

Note: All questions are equally weighted, 10 points each

THESE ARE PROVIDED FOR YOUR USE:

(1) Descriptive Statistics

Variable	Mean	St Deviation
X1	19.98	0.64
X2	64.17	1.72
X3	68.72	2.14
X4	132.89	2.19
Y	68.66	2.80

(2) Correlation Matrix

	X1	X2	X3	X4	Y
X1	1	-0.108	0.031	-0.054	0.103
X2	-0.108	1	-0.368	0.423	0.074
X3	0.031	-0.368	1	0.687	0.784
X4	0.103	0.423	0.687	1	0.822
Y	0.103	0.074	0.784	0.822	1

(3) Regression Results

Model: Y = X1;

ANOVA

	df	*SS*	*MS*	*F*	*p-value*
Regression	1	1.668	1.668	0.193	0.66546
Residual	18	155.320	8.629		
Total	19	156.988			

Model: Y = X2

ANOVA

	df	*SS*	*MS*	*F*	*p-value*
Regression	1	0.853	0.853	0.098	0.75743
Residual	18	156.135	8.674		
Total	19	156.988			

Model: Y = X3

ANOVA

	df	*SS*	*MS*	*F*	*p-value*
Regression	1	96.444	96.444	28.673	0.00004
Residual	18	60.544	3.364		
Total	19	156.988			

Model: Y = X4

ANOVA

	df	*SS*	*MS*	*F*	*p-value*
Regression	1	105.983	105.983	37.402	0.00001
Residual	18	51.005	2.834		
Total	19	156.988			

Model: Y = X1 X2;

ANOVA

	df	*SS*	*MS*	*F*	*p-value*
Regression	2	2.811	1.405	0.155	0.85764
Residual	17	154.177	9.069		
Total	19	156.988			

Model: Y = X1 X3;

ANOVA

	df	*SS*	*MS*	*F*	*p-value*
Regression	2	97.417	48.708	13.900	0.00026
Residual	17	59.571	3.504		
Total	19	156.988			

Model: Y = X2 X3

ANOVA

	df	*SS*	*MS*	*F*	*p-value*
Regression	2	120.303	60.151	27.874	<.00001
Residual	17	36.685	2.158		
Total	19	156.988			

Model: Y = X1 X4;

ANOVA

	df	*SS*	*MS*	*F*	*p-value*
Regression	2	109.408	54.704	19.545	0.00004
Residual	17	47.580	2.799		
Total	19	156.988			

Model: Y = X1 X2 X3;

ANOVA

	df	*SS*	*MS*	*F*	*p-value*
Regression	3	122.557	40.852	18.984	0.00002
Residual	16	34.431	2.152		
Total	19	156.988			

	Coefficients	*Std Error*	*t Stat*	*p-value*
Intercept	-71.752	23.467	-3.058	0.00752
X1	0.525	0.513	1.024	0.32126
X2	0.707	0.207	3.418	0.00352
X3	1.231	0.165	7.460	<.00001