KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DEPARTMENT OF MATHEMATICS AND STATISTICS
Term 152

STAT 212 BUSINESS STATISTICS II
Third Major Exam
Allowed time 75 minutes
Wednesday April 20, 2016

Name:_____ ID #:_____ Section #:_____ Srl #:___

Important Note:

1) You must **show all work** to obtain full credit for questions on this exam.
2) **DO NOT round** your answers at each step. Round answers only if necessary at your final step to **4 decimal places**.

Question No	Full Marks	Marks Obtained
Q1	17	
Q2	11	
Q3	22	
Total	**50**	

Question One (17 *points*):

The marketing department of a tool manufacturing company forecasts the quarterly demand for the company's products (Sales in $Millions). The predictors used are Car Sales (X_1 in Millions), Money Supply Index of the previous quarter (X_2), and Oil Price of previous quarter (X_3 in $/barrel). The results of the regression analysis are shown in the Minitab output below:

Regression Analysis: Y versus X1, X2, X3

```
The regression equation is
Y = 85.5 + 0.523 X1 - 33.3 X2 + 0.249 X3


Predictor     Coef  SE Coef      T      P    VIF
Constant     85.51    13.69   6.25  0.000
X1          0.5226   0.1131   4.62  0.001  1.891
X2         -33.274    6.360  -5.23  0.000  2.134
X3          0.2489   0.1277   1.95  0.080  1.393


S = 0.697823   R-Sq = 86.1%   R-Sq(adj) = 81.9%


Analysis of Variance

Source          DF      SS      MS      F      P
Regression       3  30.094  10.031  20.60  0.000
Residual Error  10   4.870   0.487
Total           13  34.964


Source  DF  Seq SS
X1       1   2.043
X2       1  26.200
X3       1   1.851
```

Regression Analysis: Y versus X1, X2, ...

```
The regression equation is
Y = - 20517 + 565 X1 + 8665 X2 + 1162 X3 - 241 X1*X2 - 32.0
X1*X3 - 488 X2*X3
    + 13.4 X1*X2*X3 + 0.128 X1^2 + 7 X2^2 + 0.0283 X3^2


Predictor     Coef  SE Coef      T      P
Constant    -20517    13077  -1.57  0.215
X1           565.2    318.4   1.78  0.174
X2            8665     5949   1.46  0.241
```

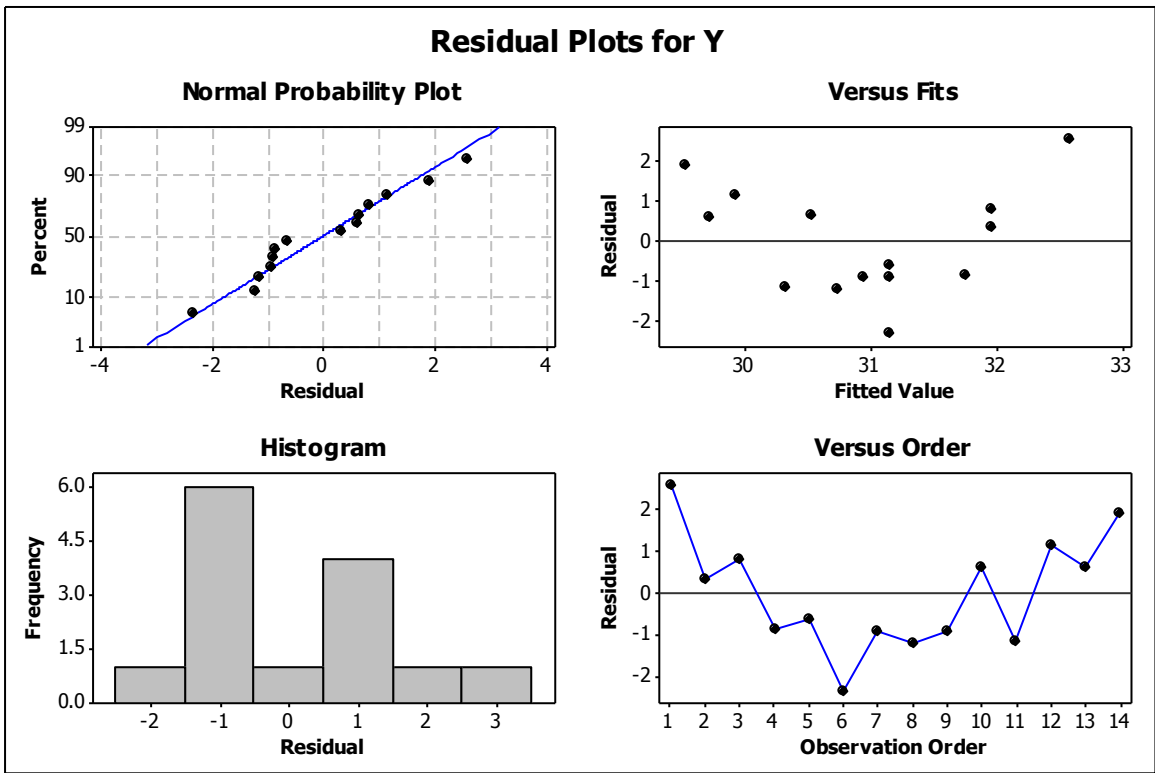
```
X3           1161.9     694.3    1.67   0.193
X1*X2        -240.7     138.1   -1.74   0.180
X1*X3        -31.97     18.06   -1.77   0.175
X2*X3        -487.8     294.2   -1.66   0.196
X1*X2*X3     13.412     7.661    1.75   0.178
X1^2         0.1275    0.1418    0.90   0.435
X2^2            6.9     247.0    0.03   0.979
X3^2        0.02827   0.09627    0.29   0.788

S = 0.571188   R-Sq = 97.2%   R-Sq(adj) = 87.9%

Analysis of Variance

Source           DF       SS       MS       F       P
Regression       10  33.9852   3.3985   10.42   0.040
Residual Error    3   0.9788   0.3263
Total            13  34.9639
```

a. Referring to the graphs below, what can you conclude about the regression model? Explain. (4 *points*)



b. Given that fitting the model `Y = 164 + 0.520 X1 - 102 squareroot(X2) + 0.247 X3` gave the previous graphs in part a , then what do you suggest based on your conclusion in part a above? Explain in detail. (2 *points*)

c. From the linear model, which predictor is significant to estimate the quarterly demand of the company's product? Why? (3 *points*)

d. From the nonlinear model, is the overall model significant in predicting the quarterly demand of the company's product? Why? (3 *points*)

e. Do you see any conflict between the two models? If yes, what (is/are) (that/these) conflict(s)? (5 *points*)

Question Two (11 *points*):

The Minitab output below shows the result of regressing the asking Price, for condos in some city, on six variables namely; number of Rooms, number of Bedrooms, number of Baths, Age of the condo, assessed Value ($), and Area ($ft^2$).

Correlations: Price, Rooms, Bedrooms, Baths, Age, Value, Area

	Price	Rooms	Bedrooms	Baths	Age	Value
Rooms	0.694					
	0.000					
Bedrooms	0.639	0.835				
	0.001	0.000				
Baths	0.325	0.301	0.137			
	0.113	0.144	0.513			
Age	0.429	0.440	0.475	0.114		
	0.032	0.028	0.016	0.588		
Value	0.555	0.354	0.228	0.331	0.018	
	0.004	0.082	0.274	0.106	0.931	
Area	0.672	0.727	0.621	0.365	0.327	0.546
	0.000	0.000	0.001	0.073	0.111	0.005

Cell Contents: Pearson correlation
P-Value

Best Subsets Regression: Price versus Rooms, Bedrooms, ...

Response is Price

Vars	R-Sq	R-Sq(adj)	Mallows Cp	S	Rooms	Bedrooms	Baths	Value	Area	Age * Value
1	48.2	46.0	10.8	8927.0	X					
1	45.2	42.8	12.7	9182.4					X	
2	62.0	58.5	4.4	7823.6			X			X
2	59.1	55.4	6.1	8112.1	X			X		
3	69.3	64.9	1.9	7197.8	X		X			X
3	67.3	62.6	3.1	7423.5		X	X			X
4	70.8	65.0	2.9	7184.9			X	X	X	X
4	70.8	64.9	2.9	7191.4	X		X	X		X
5	72.1	64.8	4.1	7204.1	X		X	X	X	X
5	71.1	63.5	4.8	7341.4		X	X	X	X	X
6	72.2	62.9	6.1	7392.5	X	X	X	X	X	X
6	72.2	62.9	6.1	7394.1	X		X	X	X	X
7	72.3	60.9	8.0	7591.1	X	X	X	X	X	X

Based on the output before, answer the following:

a. If you want to use forward selection regression, which variable you must start with? Why? (2 *points*)

b. If you want to use backward elimination regression, which variable you must start with? Why? (2 *points*)

c. If you want to use best subset regression, write down the model you are going to select? Why? (4 *points*)

d. Do you think there is severe collinearity in this study? Why? (3 *points*)

Question Three (22 *points*):

The following Minitab output resulted from a study conducted on a sample of 96 companies to predict the Return on capital (%) using four other variables which are Sales ($ millions), operating Margin (%), Debt to Capital (%), and company's Sector (Banking, Computers, Construction, Energy).

Regression Analysis: RETURN versus SALES, MARGIN, …

```
The regression equation is
RETURN = 14.6 + 0.000023 SALES + 0.0824 MARGIN - 0.0918 DEBTTOCAPITAL
+ 10.1 Computer + 2.81 Construction - 1.64 Energy

Predictor            Coef     SE Coef      T      P    VIF
Constant           14.619       2.515   5.81  0.000
SALES          0.00002284  0.00002572   0.89  0.377  1.247
MARGIN            0.08240     0.05528   1.49  0.140  1.221
DEBTTOCAPITAL    -0.09184     0.04439  -2.07  0.041  1.622
Computer           10.052       2.025   4.96  0.000  1.856
Construction        2.807       2.276   1.23  0.221  1.822
Energy             -1.641       1.873  -0.88  0.383  1.910

S = 6.30579   R-Sq = 46.8%

Analysis of Variance

Source           DF       SS      MS      F      P
Regression        6  3109.65          13.03  0.000
Residual Error   89  3538.90
Total            95  6648.55
```

Regression Analysis: RETURN versus SALES, Computer, Construction, Energy

```
Predictor            Coef     SE Coef      T      P    VIF
Constant           14.448       1.385  10.43  0.000
SALES          0.00002960  0.00002498   1.19  0.239  1.116
Computer           10.975       1.914   5.73  0.000  1.574
Construction        0.327       2.091   0.16  0.876  1.460
Energy             -3.000       1.847  -1.62  0.108  1.763

              R-Sq = 42.6%   R-Sq(adj) = 40.1%

Analysis of Variance

Source           DF       SS      MS      F      P
Regression        4  2835.53  708.88  16.92  0.000
Residual Error   91  3813.01   41.90
Total            95  6648.55

Source         DF   Seq SS
SALES           1     0.59
Computer        1  2675.19
Construction    1    49.16
Energy          1   110.58
```

Referring to the two models above, answer the following:

a. Explain in detail which of the two models is better in predicting the Return on capital of the company. (6 *points*)

b. Using the first model, interpret the coefficient of the Computer predictor.(4 *points*)

c. Compute the coefficient of partial determination for BOTH the operating Margin and the Debt to Capital. (4 *points*)

d. Using the second model, estimate the Return on capital for a company from the banking sector with sales of $ 6 million, an operating margin of 5%, and debt to capital of 7% . (2 *points*)

e. Do you think that BOTH the operating Margin and the Debt to Capital are significant to the Return on Capital in the presence of the other predictors? Explain in detail. (6 *points*)

With My Best Wishes