STAT 2509 A
Assignment #3

DUE: November 3rd, 2014 (to be handed in during the class)

1. Marketing Department for a large manufacturer of electronic games would like to measure the effectiveness of different types of advertising media in the promotion of its products. Specifically, two types of media are to be considered: radio and television advertising and newspaper advertising. A sample of 22 cities with approximately equal populations is selected for study during a test period of one month. The sales (in millions of dollars) for electronic games during the test month are given in the following table:

City	Sales ($ Millions), y	Radio & TV Advertising ($000), x_1	Newspaper Advertising($000), x_2
1	9.73	0	20
2	11.19	0	20
3	8.75	5	5
4	6.25	5	5
5	9.10	10	10
6	9.71	10	10
7	9.31	15	15
8	11.77	15	15
9	8.82	20	5
10	9.82	20	5
11	16.28	25	25
12	15.77	25	25
13	10.44	30	0
14	9.14	30	0
15	13.29	35	5
16	13.30	35	5
17	14.05	40	10
18	14.36	40	10
19	15.21	45	15
20	17.41	45	15
21	18.66	50	20
22	17.17	50	20

Consider the model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$.

(a) State all assumptions which are necessary for the statistical inference.

(b) Use matrices to compute the estimates of the population parameters β_0, β_1, β_2 and hence obtain the fitted least squares prediction line.

Hint: $\mathbf{X^T X} = \begin{bmatrix} 22 & 550 & 260 \\ 550 & 19250 & 6650 \\ 260 & 6650 & 4300 \end{bmatrix}$, $\mathbf{X^T Y} = \begin{bmatrix} 269.53 \\ 7667.20 \\ 3515.10 \end{bmatrix}$, $\mathbf{Y^T Y} = \sum y_i^2 = 3552.8893$

$$(\mathbf{X}^T\mathbf{X})^{-1} = \begin{bmatrix} 0.26048106 & -0.00429715 & -0.00910442 \\ -0.00429715 & 0.00018243 & -0.00002230 \\ -0.00910442 & -0.00002230 & 0.00081754 \end{bmatrix}, \quad \sum y_i = 269.53$$

(c) Set up the ANOVA table and hence test for the significance of the model. Use $\alpha = 0.05$.

(d) Test whether x_2 term (i.e. whether the newspaper advertising) contributes to the given model. Use t-test with $\alpha = 0.05$.

(e) Find the values of the coefficient of determination, r^2, and the adjusted r^2 and interpret their meanings in this problem.

(f) Run SAS to verify your above results and also use the SAS output to answer part (d) using partial F-test with $\alpha = 0.05$.

2. Consider the following model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3 + \varepsilon$$

where $x_2 = \begin{cases} 1, & if\ drug\ \ B \\ 0, & otherwise \end{cases}$ $x_3 = \begin{cases} 1, & if\ drug\ \ C \\ 0, & otherwise \end{cases}$

$x_1 = \ln(dose)$ y = potency of drug

Run **SAS** to test whether the 3 lines are parallel, i.e. test whether the slopes of these 3 lines are the same. Use $\alpha = 0.05$.

Use the following data:

	Drug Product		
Dose	**A**	**B**	**C**
0.2	2.0	1.8	1.3
0.4	4.3	4.1	2.0
0.8	6.5	4.9	2.8
1.6	8.9	5.7	3.4

Note: Remember to put a **FOOTNOTE** statement with your name and student number in the first line of your program.

Hand in both, your **program** and your **output**, with the appropriate measures **highlighted.**

Hand in only the **required** output when possible.