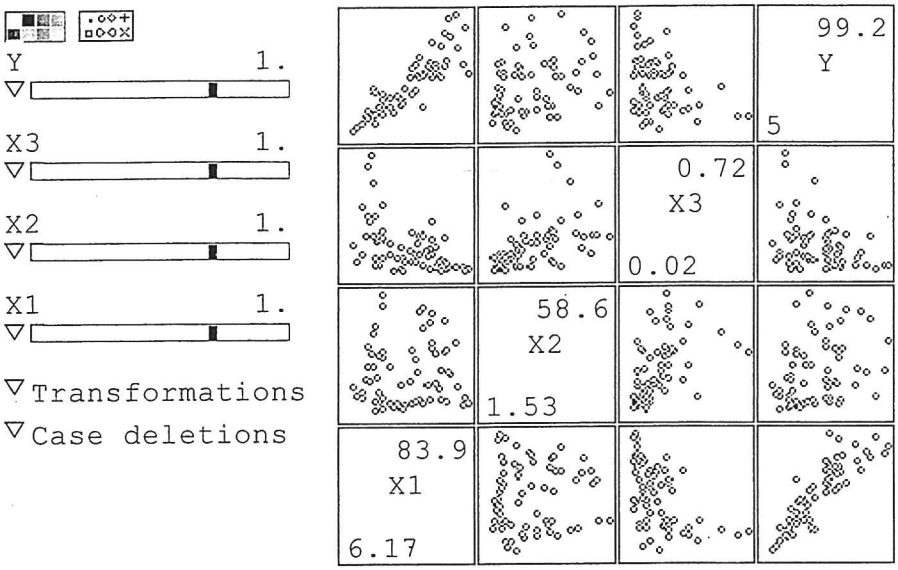
Math 484 TAKEHOME Quiz 7 Fall 2016 Name_____



1) Use the above plot to give the transformations, if any, that you would use to remove strong nonlinearities from variables.

$\log(X_1)$, $\log(X_2)$, $\log(X_3)$, $\log(Y)$

ask which ones need log trans or log rule

$\frac{\max}{\min} > 10$

25

	L1	L2	L3	L4
# of predictors	6	4	3	2
# with $0.01 \leq$ p-value ≤ 0.05	0	0	0	0
# with p-value > 0.05	3	2	0	0
R_I^2	0.913	0.913	0.913	0.900
$\text{corr}(\hat{Y}, \hat{Y}_I)$	1.0	0.9999	0.9996	0.9926
$C_p(I)$	6.0	2.190	0.793	14.389
$\sqrt{MSE}$	20.17	20.00	19.96	21.28
p-value for partial F test	1.0	0.910	0.851	0.004

2) The above table gives summary statistics for 4 MLR models considered as final submodels after performing variable selection. The data set had 112 cases and the response and residual plots for the full model L1 was good. Which model should be used as the final submodel? Explain briefly why each of the other 3 submodels should not be used.

L3 should be used $L_3 = I_{min} = I_I$

L1 L2 both have too many predictors

L4 has Cp too high ($C_p > 2k$) and partial F test pval too low

1

25

Backward

```
Current terms: (x3 x4 x6)
               df    RSS          |   k    C_I
Delete: x4    109   43441.6       |   3    0.793
Delete: x6    109   48569.8       |   3   13.399
Delete: x3    109   194726.       |   3  372.694

Current terms: (x3 x6)
               df    RSS          |   k    C_I
Delete: x6    110   49786.2       |   2   14.389
Delete: x3    110   196210.       |   2  374.343
```

$-I_{min}$

constant x_3, x_6

x_4 is deleted

Forward

```
Base terms: Intercept
               df    RSS          |   k    C_I
Add: x3       110   49786.2       |   2   14.389
Add: x6       110   196210.       |   2  374.343
Add: x2       110   205931.       |   2  398.241
Add: x4       110   431391.       |   2  952.489
Add: x5       110   461290.       |   2 1025.988

Base terms: (x3)
               df    RSS          |   k    C_I
Add: x6       109   43441.6       |   3    0.793
Add: x4       109   48569.8       |   3   13.399
Add: x5       109   49618.9       |   3   15.978
Add: x2       109   49772.8       |   3   16.356
```

$-I_{min}$

constant, x_3 x_6 is added

3) The above output is for backward elimination and forward selection. The full model has a constant x_1 and predictors x_2 = *height while sitting*, x_3 = *height while kneeling*, x_4 = *head length*, x_5 = *nasal breadth*, and x_6 =*arm span*. Hence the full model has 6 terms.

a) What is the model I_{min} that has the smallest $C_p(I)$?

constant, x_3, x_6 Note $k=3$

b) Suppose the backward elimination or forward selection model with $C_p(I) = 14.398$ had $C_p(I) = 1.440$. Would this model then be I_I? Explain briefly.

yes since $k=2<3$ and

$C_p(I) = 1.44 < C_p(I_{min}) + 1 = 1.793$

50