Ahmed Al Noor
ID#109300511
AMS 315
Professor Finch

2nd AMS 315 Project

Introduction

In this assignment we were tasked to find the model of the data with the generated data that was given to us. From the given data ,we had to use some of the variables from the 5 environmental and 20 genetic independent variables to find the model. This is because some variables have a very miniscule effect on our model.

Methodology

First, I opened my file using read.table() so that I can read the .csv file that was assigned to me. I used the lm function and the summary function to construct a linear regression for the E(i) variables. This lets me see the P-values for my E(i), which was low. I used the command adj.r.squared function to find the adjusted r-squared, which in this case was 0.4295817.
Then I decided to check if my data was appropriate from 1st order interaction upto the 3rd order interaction. To do so I used
"lm(E1+E2+E3+E4+G1+G2+G3+G4+G5+G6+G7+G8+G9+G10+G11+G12+G13+G14+G15+G16+G17+G18+G19+G20)" with all my variables to construct the linear regressing and the summary() command to check if my values were significant , which can be determined by the stars.

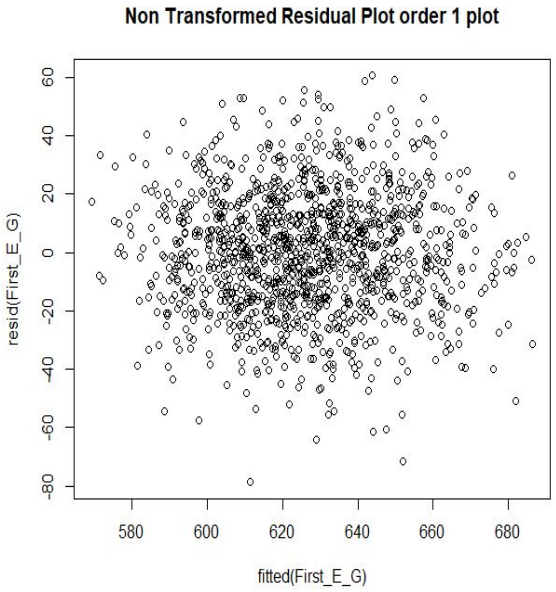
```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 64.57328   18.94521   3.408 0.000674 ***
E1           5.79895    0.41421  14.000  < 2e-16 ***
E2           8.40898    0.40522  20.752  < 2e-16 ***
E3           9.58670    0.41030  23.365  < 2e-16 ***
E4           0.45514    0.41492   1.097 0.272875
G1           2.49856    1.20616   2.072 0.038518 *
G2           9.72922    1.20781   8.055 1.84e-15 ***
G3          -1.62691    1.21600  -1.338 0.181169
G4           0.09506    1.21692   0.078 0.937748
G5           0.61032    1.20571   0.506 0.612815
G6          -0.27288    1.21956  -0.224 0.822986
G7          15.29282    1.21350  12.602  < 2e-16 ***
G8          -1.57521    1.20203  -1.310 0.190283
G9          -0.49196    1.21375  -0.405 0.685310
G10          2.04145    1.21212   1.684 0.092395 .
G11          0.34594    1.22694   0.282 0.778028
G12         -0.69175    1.22369  -0.565 0.571973
G13          2.34042    1.20136   1.948 0.051621 .
G14         -0.11708    1.21411  -0.096 0.923189
G15         -0.08856    1.20929  -0.073 0.941631
G16         -1.25676    1.22337  -1.027 0.304483
G17          0.54128    1.21421   0.446 0.655830
G18          0.75236    1.21561   0.619 0.536087
G19         -0.81286    1.22216  -0.665 0.506108
G20          0.92861    1.21365   0.765 0.444334
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 21 on 1245 degrees of freedom
Multiple R-squared:  0.5238,    Adjusted R-squared:  0.5146
F-statistic: 57.06 on 24 and 1245 DF,  p-value: < 2.2e-16
```

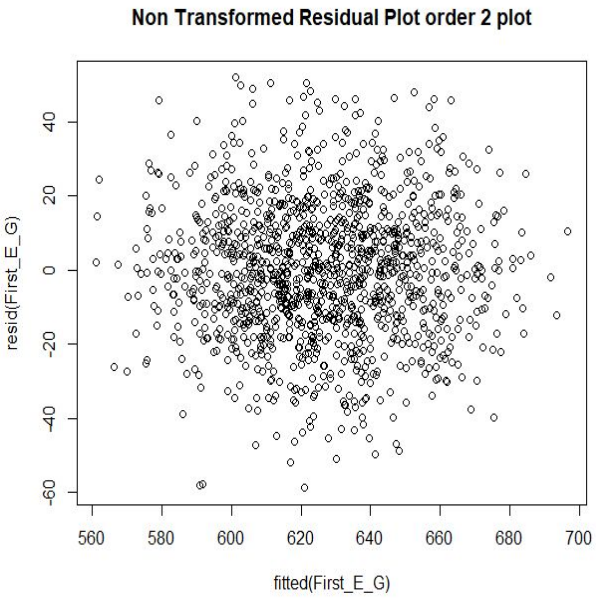
We can see that E1,E2,E3,G2 and G7 are significant for the first order.

For the second order I squared the variables in my “lm(E1+E2+E3+E4+G1+G2+G3+G4+G5+G6+G7+G8+G9+G10+G11+G12+G13+G14+G15+G16+G17+G18+G19+G20)^2” to find my 2nd order interactions between my G and E variables. Since the output is very long I attached the output at the end of the report labeled “2nd order output”. The output for the 2nd order doesn't have any of the “***”’s(3 stars), which means the T values aren't significant at the 0.001 level. This suggests that there might be little to no interactions in the 2nd order based on the low t-values .I used the code “lm(E1+E2+E3+E4+G1+G2+G3+G4+G5+G6+G7+G8+G9+G10+G11+G12+G13+G14+G15+G16+G17+G18+G19+G20)^3” and the summary() to find the third order.

```
Coefficients: (1055 not defined because of singularities)
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -3.114e+06         NA      NA       NA
E1           2.447e+03         NA      NA       NA
E2           2.195e+05         NA      NA       NA
E3           8.908e+04         NA      NA       NA
E4           5.565e+03         NA      NA       NA
G1          -1.074e+06         NA      NA       NA
G2           8.918e+05         NA      NA       NA
G3           2.446e+05         NA      NA       NA
G4           9.406e+05         NA      NA       NA
G5          -3.341e+05         NA      NA       NA
G6           2.458e+05         NA      NA       NA
G7           1.461e+06         NA      NA       NA
G8           2.079e+06         NA      NA       NA
G9           1.791e+05         NA      NA       NA
G10          5.777e+05         NA      NA       NA
```

The output was just a continuation of NAs for all our variables as we can see above, which suggest 3rd order interaction is not possible. I used plot() function to create my residual plot for all 3 orders.


Non Transformed Residual Plot order 1 plot
resid(First_E_G)
fitted(First_E_G)
-80
-60
-40
-20
0
20
40
60
580
600
620
640
660
680


Non Transformed Residual Plot order 2 plot
resid(First_E_G)
fitted(First_E_G)
-60
-40
-20
0
20
40
560
580
600
620
640
660
680
700


Non Transformed Residual Plot order 3 plot
resid(First_E_G)
fitted(First_E_G)
-1.0
-0.5
0.0
0.5
1.0
550
600
650
700

From the plots above we can see that there isn't much of a difference between the first and second order, while 3rd order looks completely different. From here on we can omit the 3rd order since it gave us no significant values and from its residual graph we can tell there is no interaction. To make my model better I installed the MASS package so that I can use the boxcox function to determine which values I should be using to transform my data. The boxcox for my first order linear regression suggests using no transformation for my variables, because the line is closing in more towards one. For the second order transformation the line is closing in more towards 0.5, which suggests I should be using 0.5 for my transformation. I applied the transformations as mentioned.

<u>From here on I will focus more on the 2nd order transformation:</u>
I created a new lm() again with my transformed dependent variables. Then I used the adj.r.squared again to find my new r squared so that I can determine if my model will be good enough. My adjusted r squared pre transformation was 0.5137398 and after the transformation it was 0.5133349, which means the transformation had little effect. I used the plot() function to plot my new transformed data, which gave me a flat eclipse. This means my transformation is fine for the model for which I will be making. I installed the leaps package so that i can use the regsubsets() command to create my stepwise regression. Then I used the kable() command to get the table with suggested models. I chose the model E1+E2:E3+E3:G2+E4:G7, because the suggested models that appeared before this model had huge differences in the BIC and adjusted R^2 values before this suggested model. Additionally, this model had high R^2 and low BIC values.But the difference after this model was very little, which is why I thought this model to be the appropriate model. To check the significance of the variables in the suggested model I used the command “kable(Check_Interaction$coefficients[abs(Check_Interaction$coefficients[,4]) <= 0.001,]”. All the values, but E4 was found to be significant. All the other selected variables from our model had high T-values, which can be seen in the result section of this report. After removing E4 from my model, my model was E1+E2:E3+E3:G2. To check if my model had second order interaction I plotted my model variables and this time there was no interaction in the 2nd order since no values were displayed when I ran the code :” M_2nd <- lm(I(Y^.5) ~ (.)^2, data=Source)
Check_Interaction <- summary(M_2nd)
kable(Check_Interaction$coefficients[abs(Check_Interaction$coefficients[,4]) <= 0.001,], caption='2nd order Interaction')” as it was said on the guide on blackboard.

```
> #probably right
> M_2nd <- lm( I(Y^.5) ~ (.)^2, data=Source)
> Check_Interaction  <- summary(M_2nd)
> kable(Check_Interaction$coefficients[ abs(Check_Interaction$coefficients[,4]) <= 0.001, ], caption='2nd order Interaction')


| Estimate| Std. Error| t value| Pr(>&#124;t&#124;)|
|--------:|----------:|-------:|------------------:|
```

I wanted to check if at least there was an interaction between my selected variable from my model so I ran the code “M_2stage <- lm(I(Y^.5) ~ (E1+E2+E3+G2+G7)^2, data=Source)
Check_Interaction <- summary(M_2stage)
Check_Interaction$coefficients[abs(Check_Interaction$coefficients[,3]) >= 4,]” as suggested from the guide. A t-values >=4 no variables appeared again. When I set it to look for t-values at

>=1 , >=2, >=3 the t-values that showed were very low as you can see in the result section of this report, which led me to conclude that there was little to no interactions in the second order.

From here on I will focus on the 1st order transformation since its the only one left.

Since the 1st order is non transformed so the R^2 stayed as 0.5146208. Since there was no transformation I had no residual plot to redo. Using the regsubsets() command I was given a table with suggested model. I decided to choose E1+E2+E3+G2+G7 over E1+E2+E3+G7, because the difference in BIC values seemed to be big as you can see below.

model	adjR2	BIC
(Intercept)+E3	0.20464279446149	-277.491854814956
(Intercept)+E2+E3	0.355417633713541	-538.286724310109
(Intercept)+E1+E2+E3	0.429202356908219	-686.533893444877
(Intercept)+E1+E2+E3+G7	0.490138496340871	-823.768126151598
(Intercept)+E1+E2+E3+G2+G7	0.514064729275593	-878.66660278039

To check the significance of my variables I used the command
kable(Check_Interaction$coefficients[abs(Check_Interaction$coefficients[,4]) <= 0.001,], caption='Check for Sig Coefficients'), which included all my variables in my model with high T-value. Then finally i used
“end <- lm(I(Y) ~ (E1+E2+E3+G2+G7), data=Source)
output <- summary(end)”
To get the values for my model.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	73.993232	16.4075362	4.509710	7.094305e-06
E1	5.812546	0.4120754	14.105540	4.551172e-42
E2	8.482580	0.4019458	21.103789	6.220615e-85
E3	9.620037	0.4079263	23.582784	3.345183e-102
G2	9.564096	1.2022425	7.955214	3.943731e-15
G7	15.353516	1.2019937	12.773374	3.139742e-35

(p-values for my result are here)

This leads me to conclude that the model selected can be obtained from the results above. After obtaining the model I used the command confint() to find my confidence intervals for my decision.

Results

I concluded my model to be Y = 73.993232 + 5.812546E1 + 8.482580E2 + 9.620037E3 + 9.564096G2+ 15.353516G7. The confidence levels are given below.

```
> confint(M_Check, '(Intercept)', level = 0.99)
                0.5 %   99.5 %
(Intercept) 31.66631 116.3202
> confint(M_Check, 'E1', level = 0.99)
       0.5 %   99.5 %
E1 4.749505 6.875587
> confint(M_Check, 'E2', level = 0.99)
       0.5 %  99.5 %
E2 7.445671 9.51949
> confint(M_Check, 'E3', level = 0.99)
     0.5 %   99.5 %
E3 8.5677 10.67237
> confint(M_Check, 'G2', level = 0.99)
       0.5 %   99.5 %
G2 6.462642 12.66555
> confint(M_Check, 'G7', level = 0.99)
      0.5 %   99.5 %
G7 12.2527 18.45433
```

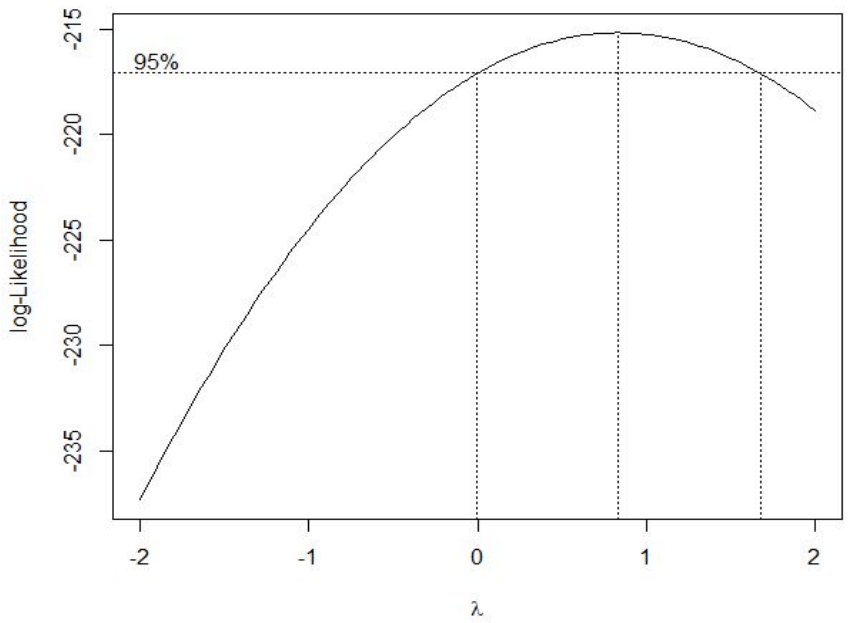
The confidence interval does not include 0 and from the picture above the confint picture we can see that the p-values are less than the alpha = 0.01. Which leads me to conclude to reject the null hypothesis that all the variables in my model are = 0 at alpha = 0.01.

Conclusion

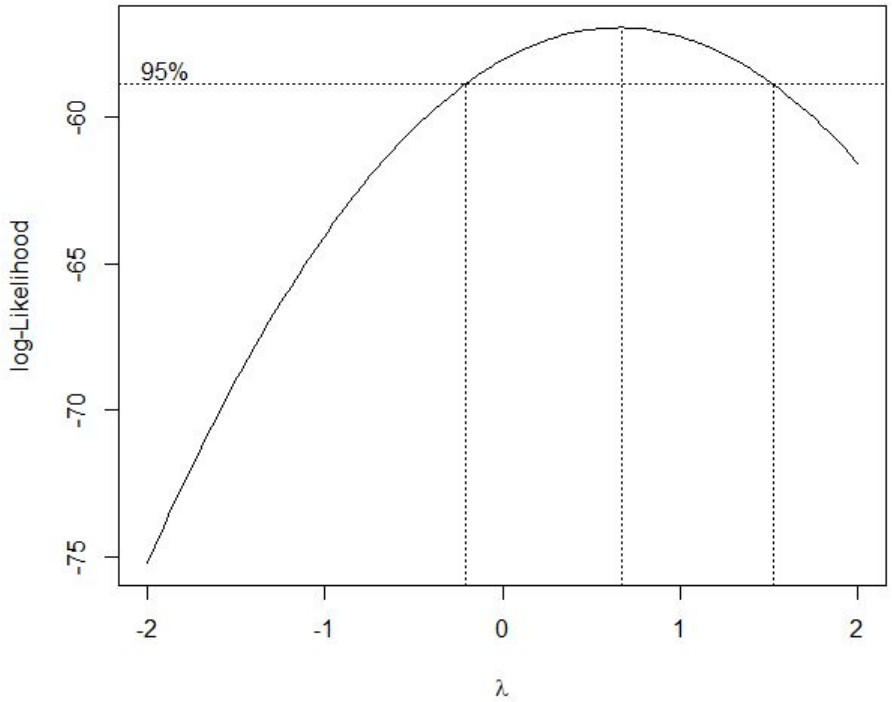
I found my model was Y = 73.993232 + 5.812546E1 + 8.482580E2 + 9.620037E3 + 9.564096G2+ 15.353516G7 based on my code and decisions. I decided to reject the null hypothesis because the p-value was less than alpha and 0 is not in our confidence interval. Our R^2 remained the same at 0.5146208 since there was no transformation due to our boxcox.

Screenshots of things mentioned in the report.

Boxcox first order



Boxcox second order



2nd order non transformed variable output (its very long)

```
> First_E_G <- lm( Y ~ (E1+E2+E3+E4+G1+G2+G3+G4+G5+G6+G7+G8+G9+G10+G11+G12+G13+G14+G15+G16+G17+G18+G19+G20)^2, data=Source )
> summary(First_E_G)

Call:
lm(formula = Y ~ (E1 + E2 + E3 + E4 + G1 + G2 + G3 + G4 + G5 +
    G6 + G7 + G8 + G9 + G10 + G11 + G12 + G13 + G14 + G15 + G16 +
    G17 + G18 + G19 + G20)^2, data = Source)

Residuals:
    Min      1Q  Median      3Q     Max
-58.747 -11.991  -0.138  11.626  51.825

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -3.607e+02  4.369e+02  -0.826  0.40923
E1           1.190e+01  1.284e+01   0.927  0.35428
E2           2.379e+01  1.255e+01   1.895  0.05844 .
E3           2.242e+01  1.279e+01   1.754  0.07981 .
E4          -2.694e-01  1.386e+01  -0.019  0.98450
G1           7.805e+01  4.412e+01   1.769  0.07721 .
G2          -5.162e+01  4.399e+01  -1.173  0.24089
G3           2.878e+01  4.439e+01   0.648  0.51689
G4           6.548e+01  4.455e+01   1.470  0.14197
G5          -3.213e+01  4.263e+01  -0.754  0.45115
G6          -3.069e+01  4.530e+01  -0.678  0.49819
G7          -4.302e+01  4.347e+01  -0.990  0.32262
G8           2.400e+01  4.436e+01   0.541  0.58867
G9           3.370e+01  4.529e+01   0.744  0.45703
G10          5.365e+00  4.421e+01   0.121  0.90344
G11         -1.619e+01  4.336e+01  -0.373  0.70889
G12         -2.901e+01  4.452e+01  -0.652  0.51477
G13          8.332e+00  4.402e+01   0.189  0.84992
G14         -2.471e+00  4.356e+01  -0.057  0.95477
G15          1.297e+01  4.453e+01   0.291  0.77087
G16          1.510e+00  4.476e+01   0.034  0.97309
G17          5.299e+01  4.528e+01   1.170  0.24215
G18         -6.190e+01  4.490e+01  -1.378  0.16841
G19          1.230e+02  4.521e+01   2.721  0.00662 **
G20          1.534e+01  4.383e+01   0.350  0.72647
E1:E2       -6.684e-01  3.179e-01  -2.103  0.03576 *
E1:E3       -4.937e-03  3.230e-01  -0.015  0.98781
E1:E4        4.433e-01  3.241e-01   1.368  0.17165
E1:G1       -1.266e+00  9.575e-01  -1.322  0.18640
E1:G2        4.244e-01  9.617e-01   0.441  0.65911
E1:G3       -3.579e-01  9.983e-01  -0.359  0.72004
E1:G4       -5.657e-01  9.693e-01  -0.584  0.55963
E1:G5       -8.581e-01  9.642e-01  -0.890  0.37372
E1:G6       -6.027e-02  9.748e-01  -0.062  0.95071
E1:G7        9.620e-01  9.618e-01   1.000  0.31746
E1:G8       -2.330e-01  9.553e-01  -0.244  0.80736
E1:G9        2.513e-01  9.677e-01   0.260  0.79512
E1:G10      -1.033e+00  9.507e-01  -1.086  0.27765
E1:G11       6.022e-01  9.714e-01   0.620  0.53544
E1:G12       1.242e+00  9.699e-01   1.280  0.20075
E1:G13       9.705e-01  9.721e-01   0.998  0.31839
E1:G14       2.549e-01  9.722e-01   0.262  0.79319
E1:G15      -8.311e-01  9.665e-01  -0.860  0.39002
E1:G16      -7.239e-02  9.805e-01  -0.074  0.94116
E1:G17      -7.444e-01  9.750e-01  -0.763  0.44541
E1:G18       6.188e-01  9.603e-01   0.644  0.51949
E1:G19      -6.386e-01  9.755e-01  -0.655  0.51286
E1:G20      -1.304e-01  9.495e-01  -0.137  0.89081
E2:E3       -3.262e-02  3.217e-01  -0.101  0.91925
E2:E4        1.145e-02  3.230e-01   0.035  0.97172
E2:G1       -6.851e-01  9.440e-01  -0.726  0.46818
E2:G2        8.641e-01  9.243e-01   0.935  0.35010
E2:G3       -1.164e-01  9.345e-01  -0.125  0.90091
E2:G4        1.338e-01  9.511e-01   0.141  0.88818
E2:G5        6.366e-01  9.258e-01   0.688  0.49188
E2:G6        1.076e+00  9.381e-01   1.147  0.25148
E2:G7        2.278e-01  9.299e-01   0.245  0.80649
E2:G8       -2.117e-02  9.283e-01  -0.023  0.98181
E2:G9        1.553e-01  9.578e-01   0.162  0.87120
E2:G10      -2.697e-01  9.418e-01  -0.286  0.77464
E2:G11      -5.057e-01  9.345e-01  -0.541  0.58855
E2:G12      -4.706e-01  9.653e-01  -0.488  0.62601
E2:G13      -1.078e+00  9.306e-01  -1.158  0.24702
E2:G14       5.728e-01  9.427e-01   0.608  0.54357
E2:G15       5.507e-01  9.488e-01   0.580  0.56176
E2:G16      -4.667e-01  9.493e-01  -0.492  0.62308
E2:G17      -2.129e-01  9.471e-01  -0.225  0.82220
E2:G18       1.536e+00  9.447e-01   1.625  0.10441
E2:G19      -2.333e+00  9.525e-01  -2.449  0.01448 *
E2:G20       1.235e+00  9.684e-01   1.276  0.20239
E3:E4       -4.172e-01  3.290e-01  -1.268  0.20512
E3:G1        3.485e-02  9.445e-01   0.037  0.97057
E3:G2        9.474e-01  9.451e-01   1.003  0.31635
E3:G3       -1.228e+00  9.618e-01  -1.277  0.20207
E3:G4       -8.240e-01  9.626e-01  -0.856  0.39217
E3:G5        1.104e+00  9.376e-01   1.177  0.23931
E3:G6        4.307e-02  9.759e-01   0.044  0.96480
E3:G7       -3.595e-02  9.538e-01  -0.038  0.96994
E3:G8       -1.094e+00  9.490e-01  -1.153  0.24926
E3:G9       -9.661e-01  9.569e-01  -1.010  0.31293
E3:G10       2.171e-02  9.719e-01   0.022  0.98218
E3:G11      -3.652e-01  9.691e-01  -0.377  0.70636
E3:G12       7.983e-02  9.498e-01   0.084  0.93304
E3:G13      -1.015e+00  9.412e-01  -1.079  0.28107
E3:G14      -2.525e-01  9.529e-01  -0.265  0.79110
E3:G15       6.575e-01  9.482e-01   0.693  0.48818
E3:G16       1.242e-02  9.648e-01   0.013  0.98974
E3:G17      -1.290e+00  9.527e-01  -1.354  0.17620
E3:G18       7.891e-01  9.721e-01   0.812  0.41710
E3:G19      -1.238e-01  9.765e-01  -0.127  0.89911
E3:G20      -8.267e-01  9.414e-01  -0.878  0.38006
E4:G1       -1.128e+00  9.618e-01  -1.173  0.24104
E4:G2        5.292e-01  9.606e-01   0.551  0.58184
E4:G3        2.823e-01  9.803e-01   0.288  0.77340
E4:G4       -1.158e+00  9.664e-01  -1.198  0.23114
E4:G5        5.081e-01  9.554e-01   0.532  0.59495
E4:G6        8.700e-02  9.971e-01   0.087  0.93049
E4:G7        2.079e+00  9.732e-01   2.136  0.03293 *
E4:G8        7.955e-02  9.510e-01   0.084  0.93335
E4:G9       -8.282e-01  9.856e-01  -0.840  0.40098
E4:G10       1.314e+00  9.569e-01   1.373  0.16995
```

```
E4:G17       5.664e-02  9.896e-01   0.057  0.95437
E4:G18       1.018e-01  9.821e-01   0.104  0.91749
E4:G19      -2.024e+00  9.760e-01  -2.074  0.03838 *
E4:G20      -1.230e+00  9.728e-01  -1.264  0.20651
G1:G2        3.539e+00  2.783e+00   1.272  0.20381
G1:G3       -2.333e+00  2.830e+00  -0.824  0.40990
G1:G4       -5.465e+00  2.812e+00  -1.943  0.05228 .
G1:G5        1.930e+00  2.745e+00   0.703  0.48227
G1:G6       -3.681e+00  2.821e+00  -1.305  0.19224
G1:G7       -3.437e+00  2.760e+00  -1.245  0.21335
G1:G8        5.129e+00  2.724e+00   1.883  0.05995 .
G1:G9       -3.056e+00  2.780e+00  -1.099  0.27193
G1:G10      -1.124e-01  2.761e+00  -0.041  0.96755
G1:G11       2.182e+00  2.857e+00   0.764  0.44507
G1:G12       6.066e-01  2.821e+00   0.215  0.82976
G1:G13      -5.182e-01  2.847e+00  -0.182  0.85560
G1:G14      -3.248e+00  2.798e+00  -1.161  0.24603
G1:G15       3.905e+00  2.835e+00   1.378  0.16863
G1:G16      -5.427e+00  2.866e+00  -1.894  0.05859 .
G1:G17      -1.855e+00  2.848e+00  -0.651  0.51494
G1:G18      -6.855e-01  2.773e+00  -0.247  0.80481
G1:G19      -5.239e-01  2.788e+00  -0.188  0.85096
G1:G20       1.062e+00  2.834e+00   0.375  0.70781
G2:G3        6.779e+00  2.781e+00   2.438  0.01496 *
G2:G4        1.752e-01  2.779e+00   0.063  0.94973
G2:G5       -1.605e+00  2.789e+00  -0.575  0.56528
G2:G6        3.831e+00  2.812e+00   1.362  0.17347
G2:G7       -4.996e+00  2.793e+00  -1.789  0.07395 .
G2:G8       -1.557e+00  2.754e+00  -0.565  0.57196
G2:G9       -3.229e+00  2.835e+00  -1.139  0.25492
G2:G10      -5.846e+00  2.792e+00  -2.094  0.03651 *
G2:G11       1.694e+00  2.885e+00   0.587  0.55730
G2:G12       5.978e+00  2.843e+00   2.102  0.03578 *
G2:G13      -2.753e+00  2.837e+00  -0.970  0.33214
G2:G14      -3.600e+00  2.798e+00  -1.287  0.19853
G2:G15       1.468e+00  2.814e+00   0.522  0.60194
G2:G16       1.588e+00  2.841e+00   0.559  0.57624
G2:G17      -3.527e+00  2.754e+00  -1.281  0.20064
G2:G18       1.823e+00  2.790e+00   0.654  0.51351
G2:G19      -2.314e+00  2.845e+00  -0.814  0.41609
G2:G20       1.527e+00  2.846e+00   0.536  0.59174
G3:G4       -2.367e+00  2.826e+00  -0.838  0.40246
G3:G5        4.614e+00  2.778e+00   1.661  0.09709 .
G3:G6        3.398e+00  2.808e+00   1.210  0.22651
G3:G7       -2.279e+00  2.779e+00  -0.820  0.41235
G3:G8        7.117e-01  2.774e+00   0.257  0.79754
G3:G9       -2.185e+00  2.806e+00  -0.779  0.43632
G3:G10       2.229e+00  2.845e+00   0.784  0.43338
G3:G11       3.153e+00  2.836e+00   1.112  0.26645
G3:G12      -8.589e-01  2.841e+00  -0.302  0.76247
G3:G13      -6.095e+00  2.782e+00  -2.191  0.02871 *
G3:G14       5.692e+00  2.860e+00   1.990  0.04687 *
G3:G15      -4.117e+00  2.767e+00  -1.488  0.13718
G3:G16       3.692e+00  2.888e+00   1.278  0.20146
G3:G17      -4.846e+00  2.842e+00  -1.705  0.08846 .
G3:G18      -1.985e-01  2.824e+00  -0.070  0.94399
G3:G19       1.449e+00  2.840e+00   0.510  0.61020
G3:G20      -5.265e+00  2.786e+00  -1.890  0.05905 .
G4:G5       -4.655e+00  2.740e+00  -1.699  0.08965 .
G4:G6       -2.516e+00  2.825e+00  -0.891  0.37340
G4:G7        2.555e-01  2.808e+00   0.091  0.92750
G4:G8        5.411e+00  2.827e+00   1.914  0.05586 .
G4:G9        4.124e+00  2.830e+00   1.457  0.14539
G4:G10      -1.828e+00  2.768e+00  -0.660  0.50918
G4:G11       4.082e+00  2.785e+00   1.466  0.14304
G4:G12      -4.570e+00  2.886e+00  -1.584  0.11358
G4:G13       1.533e+00  2.880e+00   0.532  0.59465
G4:G14       2.277e+00  2.834e+00   0.804  0.42188
G4:G15      -8.806e-01  2.843e+00  -0.310  0.75680
G4:G16       6.125e-01  2.822e+00   0.217  0.82820
G4:G17      -2.587e+00  2.788e+00  -0.928  0.35369
G4:G18      -3.925e+00  2.790e+00  -1.407  0.15975
G4:G19      -6.068e+00  2.871e+00  -2.113  0.03483 *
G4:G20      -8.501e-01  2.789e+00  -0.305  0.76059
G5:G6       -1.286e-01  2.808e+00  -0.046  0.96347
G5:G7       -3.057e+00  2.736e+00  -1.117  0.26412
G5:G8        2.366e+00  2.722e+00   0.869  0.38498
G5:G9        1.397e+00  2.808e+00   0.497  0.61903
G5:G10       1.272e+00  2.809e+00   0.453  0.65083
G5:G11      -1.155e+00  2.847e+00  -0.406  0.68502
G5:G12       3.186e+00  2.834e+00   1.124  0.26129
G5:G13       5.550e+00  2.763e+00   2.009  0.04483 *
G5:G14      -5.369e+00  2.772e+00  -1.937  0.05303 .
G5:G15       4.055e+00  2.762e+00   1.468  0.14242
G5:G16      -3.295e+00  2.881e+00  -1.144  0.25300
G5:G17      -6.046e+00  2.771e+00  -2.182  0.02935 *
G5:G18      -9.655e-01  2.741e+00  -0.352  0.72477
G5:G19       3.704e+00  2.783e+00   1.331  0.18343
G5:G20       2.372e+00  2.816e+00   0.842  0.39997
G6:G7       -2.617e+00  2.817e+00  -0.929  0.35318
G6:G8       -6.218e+00  2.742e+00  -2.268  0.02355 *
G6:G9        1.811e+00  2.850e+00   0.635  0.52536
G6:G10       2.137e+00  2.828e+00   0.756  0.45003
 [ reached getOption("max.print") -- omitted 101 rows ]
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 21.02 on 969 degrees of freedom
Multiple R-squared:  0.6287,    Adjusted R-squared:  0.5137
F-statistic: 5.469 on 300 and 969 DF,  p-value: < 2.2e-16
```

Stepwise regression

```
Data_Transformed <- lm( I(Y) ~ (.), data=Source )
summary(First_E_G)$adj.r.square

summary(Data_Transformed)$adj.r.square

plot(resid(Data_Transformed) ~ fitted(Data_Transformed), main='Transformed Residual Plot')


M <- regsubsets( model.matrix(Data_Transformed)[,-1], I(Source$Y),
                 nbest = 1, nvmax=5,
                 method = 'forward', intercept = TRUE )
Check_Interaction <- summary(M)

Var <- colnames(model.matrix(Data_Transformed))
M_Select <- apply(Check_Interaction$which, 1,
                  function(x) paste0(Var[x], collapse='+'))
kable(data.frame(cbind( model = M_Select, adjR2 = Check_Interaction$adjr2, BIC = Check_Interaction$bic)),
      caption='Model Suggestions')
```

P values for only E(i) values at the beginning of the code.

```
> library(MASS)
> library(knitr)
> library(leaps)
> Source <- read.csv('P2_00511.csv', header=TRUE)
> First_E <- lm(Y ~ E1+E2+E3+E4, data=Source)
> summary(First_E)

Call:
lm(formula = Y ~ E1 + E2 + E3 + E4, data = Source)

Residuals:
    Min      1Q  Median      3Q     Max
-81.720 -15.217   0.249  15.743  67.448

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  78.5796    20.1209   3.905 9.91e-05 ***
E1            5.7083     0.4467  12.778  < 2e-16 ***
E2            8.2152     0.4348  18.892  < 2e-16 ***
E3            9.8354     0.4415  22.279  < 2e-16 ***
E4            0.6071     0.4473   1.357    0.175
---
```

When I set the t to be <= 1,2,3,4

```
> Check_Interaction$coefficients[ abs(Check_Interaction$coefficients[,3]) >= 4, ] #### what is the 2.5?
    Estimate Std. Error t value Pr(>|t|)
```

```
> Check_Interaction$coefficients[ abs(Check_Interaction$coefficients[,3]) >= 3, ]
    Estimate Std. Error t value Pr(>|t|)
>
```

```
> Check_Interaction$coefficients[ abs(Check_Interaction$coefficients[,3]) >= 2, ] #### what is the 2.5?
   Estimate Std. Error  t value    Pr(>|t|)
E2 0.4823111  0.1778880 2.711319 0.006793129
E3 0.3752843  0.1875209 2.001293 0.045575965
>
```

```
> Check_Interaction$coefficients[ abs(Check_Interaction$coefficients[,3]) >= 1, ]
               Estimate  Std. Error    t value    Pr(>|t|)
(Intercept)  6.595015089 5.246824222  1.256954 0.209004470
E1           0.285785498 0.189698256  1.506527 0.132183858
E2           0.482311075 0.177887958  2.711319 0.006793129
E3           0.375284300 0.187520900  2.001293 0.045575965
E1:E2       -0.006682625 0.005627332 -1.187530 0.235243530
E1:G7        0.020700751 0.016849408  1.228574 0.219461957
E2:E3       -0.006949169 0.005643825 -1.231287 0.218446407
G2:G7       -0.071597510 0.049453335 -1.447779 0.147928767
> |
```