MD Mahmudur Rahman

Data Analysis Project 2: Multiple Regression Computing Project

Introduction

In this project, we are first given a dataset containing one dependent variable and twenty four independent variables. We regarded variables *E*1 to *E*4 as environmental (or control) variables. We also regarded the variables *G*1 to *G*20 as genetic risk variables. The *G* variables are indicator variables and the data that we are analyzing is synthetic. Our task is to find the model that the TA used for my data. The background for this project is that there was a paper written by Caspi et al. that reports a finding of a gene-environment interaction. The paper used multiple regression techniques as the methodology for its findings. The author was determined to figure out why stressful experiences lead to depression in some people but not in others? It was hypothesized that the (5-HTT) gene is found to moderate the influence of stressful life events on depression. There were Three lines of experimental research that suggest this hypothesis of a gene-by-environment (G X E) interaction. It was tested in mice, in rhesus macaques and in human neuroim- aging research that suggests that the stress response is mediated by variations in the 5-HTTLPR (Caspi et al). Also, after the hypothesis was made, it was tested "among members of the Dunedin Multidisciplinary Health and Development Study (*16*). This representative birth cohort of 1037 children (52% male) has been assessed at ages 3, 5, 7, 9, 11, 13, 15, 18, and 21 and was virtually intact (96%) at the age of 26 years" (Caspi et al). The two big questions were to figure out whether there are any associations of the synthetic *Y* variable with any of the *G* variables and whether there are any interactions of *G* and *E* variables or *G* and *G* variables.

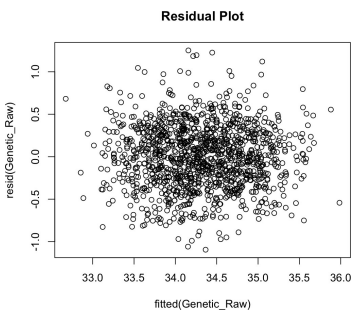
Method

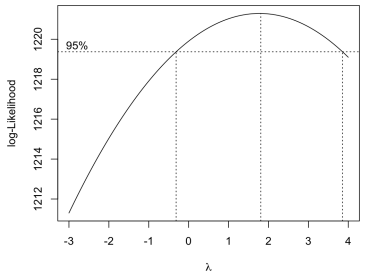
The first step that was done was to import my dataset into R Studios and realized that there was no missing data. My dataset is complete. I had 1,253 observations and 25 variables. There was 1 dependent variable which is Y and 24 Independent variables (*E*1 to *E*4 and *G*1 to *G*20). Next what I did was model the Environmental variables and got the Summary portion. Then, it was time to work with the Genetic Variables. I used the lm function to model it and I assumed this model includes all interaction terms up to the 2nd order. A Residual plot was created and saw that it was adequate because all I saw was a flat ellipse. The summary was then created. I felt like the residual plot could be a bit better. Therefore, a transformation was made. I did the boxcox and saw that it's not going to be a log transformation. I used this line of code best.lam = bc$x[which(bc$y==max(bc$y))] and got the lambda value to be 1.81 which I rounded it to 2. Therefore, I used the transformation y^2. The new residual plot was created and the summary. Then, we performed Stepwise Regression and the R Studios produced a proposed model. The model summary table is shown below with adjusted R^2 and the Bayesian Information Criterion (BIC). The code for that is written below. I chose variables selected in the 3rd model as candidates; namely E3, E1, E3, G9, and G14. That is, I split the E1E3 and G9G14 interactions. Besides this, I want to also make sure main effects that are significant are in the

model. The code for that is also shown below in the appendix. The code allowed me to get sig coefficients and the variables with all of the p-values. The E1:E3 interaction was eliminated.

Results:

After modeling the Environmental variables, the adjusted R squared value was 0.5899. As for the additional contribution of the Genetic Variables, the residual plot was created and the adjusted R squared value after assuming that this model includes all interaction terms up to the 2nd order in the model is 0.6002644. This was already a good model, however the transformation was found and it was a little bit better. I used the y^2 transformation using boxcox. The lambda value that I got was 1.81 which I rounded it to 2. I got the new adjusted R squared value to be 0.6006907. Then, after doing the stepwise regression, according to the model summary table below we see that, there is an obvious increase in the adjR2 from the 2nd model to the 3rd model. There is only a very small increase from the 3rd to the 4th model, which may not be significant. Another measure that researchers use to assess models is the Bayesian Information Criterion (BIC). The BIC values consistently decrease. The decrease from the 3rd to the 4th and the 5th model is much smaller than the other decreases. I ended up choosing the 3rd model as candidates namely E3, E1, E3, G9, and G14. That is, I split the E1E3 and G9G14 interactions. However, the table titled Sig Coefficient shows variables E1, E2 and G14 that are significant in the main effect model. I then changed my model to include all interaction terms up to the 1st order. The sig coefficient table showed E1, E3, and G14 to be significant in both the 2nd order and the first order interaction. Based on the last table where all the p-values are shown, the E1:E3 interaction was eliminated. We only had G9:G14 interaction. In the end, adding the G variables only made a little improvement since the adjusted R squared value went from 0.5899 to 0.6002644.





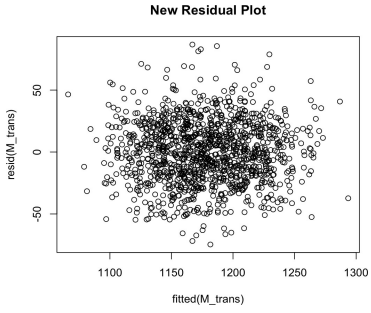


Table: Model Summary

model	adjR2	BIC
(Intercept)+E1:E3	0.577745343507559	-1067.00441622411
(Intercept)+E1:E3+G9:G14	0.590059481964278	-1097.95754934016
(Intercept)+E3+E1:E3+G9:G14	0.601626601448881	-1127.69081955919
(Intercept)+E3+E1:E3+G2:G8+G9:G14	0.603590146881052	-1127.75231891966
(Intercept)+E3+E1:E3+G2:G8+G9:G14+G17:G18	0.605381471055271	-1127.29840653493

Table: Sig Coefficients

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	410.085406	25.9487656	15.803658	0.000000
E1	14.546641	0.5620162	25.882956	0.000000
E3	19.417578	0.5687037	34.143574	0.000000
G14	6.722998	1.8180401	3.697937	0.000227

Conclusion

Our analysis found associations with genetic variables after the environmental variables had been controlled. Specifically the G9-G14 interaction with t-value 4.4014975 significantly associated with the y^2. Including these variables significantly increased the r-squared value from 0.5912 to 0.6044. The model that the TA used for my data is $y^2 = \beta_0 + \beta_{13} E_3 + \beta_{21} E_1 + \beta_{39} G_9 G_{14} + \epsilon$. The final model with estimated parameters is:

$y^2 = 513.311353 + 15.319715 E_3 + 9.908984 E_1 + 17.808464 G_9 G_{14}$. This is shown in the table below.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	513.311353	200.292591	2.5628075	1.049997e-02
E3	15.319715	8.869358	1.7272630	8.436905e-02
E1	9.908984	8.883250	1.1154683	2.648656e-01
G9	12.824424	40.920885	0.3133956	7.540328e-01
G14	-26.413922	39.205027	-0.6737381	5.006032e-01
G9:G14	17.808464	4.046001	4.4014975	1.167364e-05

Citation:

Caspi, Avshalom, et al. “Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene.” *Science*, American Association for the Advancement of Science, 18 July 2003, science.sciencemag.org/content/301/5631/386/tab-figures-data.

Code for Modeling the Environmental Variables

```
> library(readr)
> SP21_P2_949881 <- read_csv("~/Downloads/Project 2 Data Set SP2021/SP21_P2_949881.csv")

── Column specification ────────────────────────────────────────────────────────
cols(
  .default = col_double()
)
ℹ Use `spec()` for the full column specifications.

> View(SP21_P2_949881)
> Environmental_fit <- lm(Y ~ E1+E2+E3+E4+E5, data=SP21_P2_949881)
Error in eval(predvars, data, env) : object 'E5' not found
> Environmental_fit <- lm(Y ~ E1+E2+E3+E4, data=SP21_P2_949881)
> summary(Environmental_fit)

Call:
lm(formula = Y ~ E1 + E2 + E3 + E4, data = SP21_P2_949881)

Residuals:
     Min       1Q   Median       3Q      Max
-1.36359 -0.30496 -0.00475  0.29390  1.51875

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 23.144364   0.375954  61.562   <2e-16 ***
E1           0.212730   0.008210  25.911   <2e-16 ***
E2           0.002782   0.008264   0.337    0.736
E3           0.281853   0.008274  34.066   <2e-16 ***
E4          -0.001355   0.008354  -0.162    0.871
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.425 on 1248 degrees of freedom
Multiple R-squared:  0.5912,    Adjusted R-squared:  0.5899
F-statistic: 451.3 on 4 and 1248 DF,  p-value: < 2.2e-16
```

Code for Modeling the Additional Contribution of the Genetic Variables, Getting the Residual Plot and Finding the Transformation

```
> Genetic_Raw <- 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=
SP21_P2_949881 )
> plot(resid(Genetic_Raw) ~ fitted(Genetic_Raw), main='Residual Plot')
> boxcox(Genetic_Raw)
> Genetic_trans <- lm( I(log(Y)) ~ (.)^2, data=SP21_P2_949881 )
> summary(Genetic_Raw)$adj.r.square
[1] 0.6002644
> summary(Genetic_trans)$adj.r.square
[1] 0.5996779
> Genetic_trans <- lm( I(Y)^(1/2) ~ (.)^2, data=SP21_P2_949881 )
> summary(Genetic_trans)$adj.r.square
[1] 0.5999911
> Genetic_trans <- lm( I(Y)^(1/2) ~ (.)^2, data=SP21_P2_949881 )
> Genetic_trans <- lm( I(Y)^(2) ~ (.)^2, data=SP21_P2_949881 )
> summary(Genetic_trans)$adj.r.square
[1] 0.6006907
```

Code for Getting the Summary of the Additional Contribution of the Genetic Variables Without Transformation

MD Mahmudur Rahman

```
> summary(Genetic_Raw)

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 = SP21_P2_949881)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.09452 -0.25063  0.00228  0.25494  1.25136 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)   
(Intercept) 24.2032327  8.1042194   2.986  0.00289 **
E1           0.1632109  0.2492668   0.655  0.51278   
E2          -0.0893265  0.2613321  -0.342  0.73257   
E3           0.2630954  0.2563671   1.026  0.30504   
E4           0.1035604  0.2583491   0.401  0.68862   
G1           0.1899178  0.9572110   0.198  0.84277   
G2           0.0654399  0.9429390   0.069  0.94469   
G3          -0.2766230  0.9837082  -0.281  0.77861   
G4          -1.4893894  0.9591759  -1.553  0.12081   
G5          -0.2022686  0.9660415  -0.209  0.83420   
G6          -0.7066605  0.9995759  -0.707  0.47976   
G7          -2.4758661  0.9933794  -2.492  0.01286 * 
G8          -0.9661723  0.9898797  -0.976  0.32929   
G9           1.4265073  1.0083603   1.415  0.15749   
G10          0.1276429  0.9379250   0.136  0.89178   
G11          1.7731926  0.9442782   1.878  0.06071 . 
G12          0.8261814  1.0050599   0.822  0.41127   
G13         -0.2717844  0.9791398  -0.278  0.78140   
G14         -0.0492792  0.9608174  -0.051  0.95911   
G15          0.0109222  0.9534520   0.011  0.99086   
G16          0.7283826  1.0085192   0.722  0.47033   
G17         -0.3275600  0.9915678  -0.330  0.74121   
G18         -0.5959492  0.9535185  -0.625  0.53212   
G19         -1.7462114  0.9731420  -1.794  0.07307 . 
G20          0.0095157  0.9178548   0.010  0.99173   
E1:E2        0.0036022  0.0064525   0.558  0.57680   
E1:E3        0.0053444  0.0066567   0.803  0.42226   
E1:E4       -0.0071957  0.0065432  -1.100  0.27173   
E1:G1       -0.0210337  0.0207295  -1.015  0.31052   
E1:G2        0.0338809  0.0200545   1.689  0.09146 . 
E1:G3        0.0111566  0.0210796   0.529  0.59675   
E1:G4        0.0433236  0.0211074   2.053  0.04039 * 
E1:G5       -0.0224192  0.0207398  -1.081  0.27998   
E1:G6        0.0405533  0.0211553   1.917  0.05555 . 
E1:G7        0.0077659  0.0213720   0.363  0.71641   
E1:G8        0.0176027  0.0214573   0.820  0.41222   
E1:G9        0.0091288  0.0213256   0.428  0.66870   
E1:G10      -0.0372398  0.0212667  -1.751  0.08025 . 
E1:G11      -0.0338431  0.0210227  -1.610  0.10776   
E1:G12      -0.0492745  0.0215619  -2.285  0.02252 * 
E1:G13       0.0004711  0.0213883   0.022  0.98243   
E1:G14       0.0101498  0.0207821   0.488  0.62539   
E1:G15       0.0093713  0.0200876   0.467  0.64095   
E1:G16       0.0207276  0.0218584   0.948  0.34324   
E1:G17       0.0090420  0.0209000   0.433  0.66538   
E1:G18       0.0322940  0.0205915   1.568  0.11714   
E1:G19       0.0166097  0.0207429   0.801  0.42348   
E1:G20      -0.0534555  0.0199034  -2.686  0.00736 **
E2:E3       -0.0026037  0.0066249  -0.393  0.69440   
E2:E4        0.0029273  0.0066502   0.440  0.65991   
E2:G1       -0.0092246  0.0212859  -0.433  0.66485   
E2:G2        0.0198411  0.0214527   0.925  0.35526   
E2:G3       -0.0138310  0.0207179  -0.668  0.50456   
E2:G4        0.0033951  0.0214203   0.158  0.87410   
E2:G5        0.0132364  0.0210987   0.627  0.53058   
E2:G6        0.0119058  0.0211131   0.564  0.57295   
E2:G7        0.0119021  0.0213121   0.558  0.57666   
E2:G8       -0.0155846  0.0211354  -0.737  0.46108   
E2:G9       -0.0101087  0.0214386  -0.472  0.63738   
E2:G10       0.0258389  0.0209404   1.234  0.21754   
E2:G11      -0.0157578  0.0212306  -0.742  0.45814   
E2:G12      -0.0371624  0.0214645  -1.731  0.08371 . 
E2:G13       0.0116578  0.0216920   0.537  0.59110   
E2:G14       0.0171878  0.0207523   0.828  0.40774   
E2:G15      -0.0272725  0.0204131  -1.336  0.18186   
E2:G16      -0.0271922  0.0220863  -1.231  0.21856   
E2:G17       0.0026479  0.0209441   0.126  0.89942   
E2:G18       0.0065012  0.0209522   0.310  0.75641   
E2:G19       0.0177641  0.0216241   0.821  0.41157   
E2:G20       0.0345078  0.0210364   1.640  0.10126   
E3:E4       -0.0015339  0.0064680  -0.237  0.81259   
E3:G1        0.0014516  0.0206278   0.070  0.94391   
E3:G2       -0.0465267  0.0209022  -2.226  0.02625 * 
E3:G3       -0.0149187  0.0205172  -0.727  0.46733   
E3:G4       -0.0131214  0.0214032  -0.613  0.53999   
E3:G5        0.0106557  0.0207532   0.513  0.60776   
E3:G6       -0.0167859  0.0209650  -0.801  0.42353   
```

```
G3:G7        0.0489207  0.0693335   0.706  0.48062
G3:G8       -0.0762264  0.0703958  -1.083  0.27916
G3:G9       -0.0820956  0.0672419  -1.221  0.22243
G3:G10       0.0771642  0.0692639   1.114  0.26553
G3:G11       0.0053330  0.0675484   0.079  0.93709
G3:G12      -0.0795181  0.0714112  -1.114  0.26576
G3:G13      -0.0009792  0.0669437  -0.015  0.98833
G3:G14       0.0887053  0.0689812   1.286  0.19878
G3:G15      -0.0893085  0.0671225  -1.331  0.18366
G3:G16      -0.0182270  0.0695817  -0.262  0.79342
G3:G17       0.0317662  0.0705801   0.450  0.65276
G3:G18       0.0124986  0.0671750   0.186  0.85244
G3:G19       0.0580081  0.0676586   0.857  0.39146
G3:G20       0.0631692  0.0703168   0.898  0.36923
G4:G5        0.0019520  0.0667221   0.029  0.97667
G4:G6        0.1101162  0.0680457   1.618  0.10594
G4:G7        0.0766008  0.0686973   1.115  0.26511
G4:G8        0.1250608  0.0721622   1.733  0.08341 .
G4:G9       -0.0270300  0.0687989  -0.393  0.69449
G4:G10       0.0213548  0.0663559   0.322  0.74766
G4:G11      -0.0639321  0.0673599  -0.949  0.34280
G4:G12       0.0648314  0.0723671   0.896  0.37055
G4:G13       0.0490973  0.0677568   0.725  0.46887
G4:G14       0.0857092  0.0658415   1.302  0.19332
G4:G15       0.0179232  0.0665023   0.270  0.78759
G4:G16       0.0134290  0.0697239   0.193  0.84731
G4:G17       0.0254842  0.0722091   0.353  0.72422
G4:G18       0.0297874  0.0705311   0.422  0.67288
G4:G19      -0.0333188  0.0684265  -0.487  0.62642
G4:G20       0.1358131  0.0664229   2.045  0.04116 *
G5:G6        0.0642481  0.0703443   0.913  0.36130
G5:G7       -0.0298923  0.0707585  -0.422  0.67279
G5:G8        0.0737206  0.0734092   1.004  0.31552
G5:G9        0.0448837  0.0696234   0.645  0.51930
G5:G10       0.0591099  0.0669099   0.883  0.37723
G5:G11      -0.0314655  0.0697970  -0.451  0.65223
G5:G12      -0.0198348  0.0707792  -0.280  0.77936
G5:G13       0.0666785  0.0691935   0.964  0.33546
G5:G14      -0.0139349  0.0679491  -0.205  0.83756
G5:G15      -0.0380960  0.0669689  -0.569  0.56959
G5:G16       0.0187042  0.0713402   0.262  0.79324
G5:G17       0.0912430  0.0704469   1.295  0.19556
G5:G18       0.1115888  0.0702722   1.588  0.11263
G5:G19      -0.0450347  0.0689877  -0.653  0.51405
G5:G20      -0.0356792  0.0669387  -0.533  0.59415


G6:G7        0.0700004  0.0676090   1.035  0.30076
G6:G8        0.0715260  0.0716130   0.999  0.31815
G6:G9       -0.1101537  0.0702140  -1.569  0.11702
G6:G10       0.1076522  0.0696508   1.546  0.12253
 [ reached getOption("max.print") -- omitted 101 rows ]
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4196 on 952 degrees of freedom
Multiple R-squared:  0.696,     Adjusted R-squared:  0.6003
F-statistic: 7.267 on 300 and 952 DF,  p-value: < 2.2e-16
```

MD Mahmudur Rahman

Code for Getting the Summary of the Additional Contribution of the Genetic Variables With Transformation

```
> summary(Genetic_trans)

Call:
lm(formula = I(Y)^(4) ~ (.)^2, data = SP21_P2_949881)

Residuals:
    Min      1Q  Median      3Q     Max
-173736  -40998    -285   40959  208627

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  142594.56 1310970.64   0.109  0.91341
E1             6724.35   40322.38   0.167  0.86759
E2           -14526.01   42274.12  -0.344  0.73121
E3            23982.64   41470.95   0.578  0.56320
E4            20555.12   41791.58   0.492  0.62294
G1            34096.54  154842.24   0.220  0.82576
G2            16626.00  152533.54   0.109  0.91323
G3           -39422.70  159128.54  -0.248  0.80439
G4          -255969.90  155160.09  -1.650  0.09933 .
G5           -33651.39  156270.70  -0.215  0.82955
G6          -122302.16  161695.35  -0.756  0.44961
G7          -401252.49  160692.99  -2.497  0.01269 *
G8          -146146.67  160126.86  -0.913  0.36164
G9           227360.44  163116.36   1.394  0.16369
G10           10860.97  151722.46   0.072  0.94295
G11          309905.52  152750.18   2.029  0.04275 *
G12          134648.20  162582.47   0.828  0.40777
G13          -38796.09  158389.53  -0.245  0.80655
G14          -19953.96  155425.64  -0.128  0.89787
G15            7068.01  154234.17   0.046  0.96346
G16          130965.19  163142.07   0.803  0.42231
G17          -55156.46  160399.94  -0.344  0.73102
G18          -85075.41  154244.93  -0.552  0.58138
G19         -305062.63  157419.30  -1.938  0.05293 .
G20             809.35  148475.82   0.005  0.99565
E1:E2           630.01    1043.78   0.604  0.54627
E1:E3          1714.15    1076.82   1.592  0.11175
E1:E4         -1192.83    1058.45  -1.127  0.26004
E1:G1         -3516.97    3353.28  -1.049  0.29453
E1:G2          5248.93    3244.09   1.618  0.10599
E1:G3          1881.15    3409.92   0.552  0.58131
E1:G4          7265.49    3414.42   2.128  0.03360 *
E1:G5         -3483.29    3354.95  -1.038  0.29942
E1:G6          6763.10    3422.17   1.976  0.04841 *
E1:G7          1372.37    3457.21   0.397  0.69149
E1:G8          2602.76    3471.02   0.750  0.45353
E1:G9          1700.15    3449.71   0.493  0.62224
E1:G10        -6020.69    3440.19  -1.750  0.08042 .
E1:G11        -5763.67    3400.71  -1.695  0.09043 .
E1:G12        -8167.66    3487.95  -2.342  0.01940 *
E1:G13           39.39    3459.86   0.011  0.99092
E1:G14         2023.52    3361.80   0.602  0.54737
E1:G15         1485.76    3249.44   0.457  0.64761
E1:G16         3229.18    3535.91   0.913  0.36134
E1:G17         1612.04    3380.87   0.477  0.63360
E1:G18         5130.39    3330.96   1.540  0.12384
E1:G19         3098.13    3355.45   0.923  0.35608
E1:G20        -8793.70    3219.66  -2.731  0.00643 **
E2:E3          -382.71    1071.68  -0.357  0.72108
E2:E4           404.49    1075.77   0.376  0.70700
E2:G1         -1501.63    3443.29  -0.436  0.66286
E2:G2          3314.73    3470.27   0.955  0.33973
E2:G3         -2441.43    3351.41  -0.728  0.46650
E2:G4           932.95    3465.03   0.269  0.78780
E2:G5          2037.67    3413.01   0.597  0.55063
E2:G6          2196.95    3415.33   0.643  0.52021
E2:G7          1578.07    3447.54   0.458  0.64725
E2:G8         -2798.73    3418.94  -0.819  0.41322
E2:G9         -1943.10    3467.99  -0.560  0.57541
E2:G10         4208.94    3387.41   1.243  0.21435
E2:G11        -2924.42    3434.35  -0.852  0.39469
E2:G12        -5904.27    3472.18  -1.700  0.08937 .
E2:G13         1843.38    3508.98   0.525  0.59948
E2:G14         2671.33    3356.97   0.796  0.42637
E2:G15        -4619.12    3302.10  -1.399  0.16219
E2:G16        -4478.25    3572.77  -1.253  0.21035
E2:G17          580.08    3388.00   0.171  0.86409
E2:G18         1234.94    3389.32   0.364  0.71567
E2:G19         2951.46    3498.00   0.844  0.39902
E2:G20         5619.52    3402.93   1.651  0.09899 .
E3:E4          -318.62    1046.29  -0.305  0.76080
E3:G1           124.21    3336.84   0.037  0.97031
E3:G2         -7566.36    3381.22  -2.238  0.02547 *
E3:G3         -2369.55    3318.95  -0.714  0.47543
E3:G4         -2288.16    3462.27  -0.661  0.50885
E3:G5          1679.74    3357.12   0.500  0.61694
E3:G6         -2897.35    3391.38  -0.854  0.39314
E3:G7          7771.74    3551.91   2.188  0.02891 *
E3:G8            68.53    3554.46   0.019  0.98462
```

```
E3:G9         -5027.09    3432.47  -1.465  0.14337
E3:G10         5075.22    3336.55   1.521  0.12857
E3:G11        -3142.23    3416.93  -0.920  0.35801
E3:G12         1199.70    3529.65   0.340  0.73401
E3:G13        -3420.65    3485.06  -0.982  0.32659
E3:G14         2128.05    3369.46   0.632  0.52782
E3:G15         2042.76    3359.72   0.608  0.54332
E3:G16        -2821.97    3446.93  -0.819  0.41317
E3:G17         1433.75    3427.25   0.418  0.67579
E3:G18        -3027.03    3491.17  -0.867  0.38613
E3:G19         4958.28    3490.57   1.420  0.15580
E3:G20         1789.05    3381.76   0.529  0.59691
E4:G1          3609.92    3405.63   1.060  0.28942
E4:G2         -1695.02    3458.81  -0.490  0.62421
E4:G3          4632.51    3505.37   1.322  0.18664
E4:G4          3930.65    3429.79   1.146  0.25207
E4:G5          1039.27    3458.50   0.300  0.76386
E4:G6         -1534.09    3486.22  -0.440  0.66001
E4:G7          6688.78    3495.41   1.914  0.05597 .
E4:G8          6155.51    3598.31   1.711  0.08747 .
E4:G9         -4564.74    3570.56  -1.278  0.20141
E4:G10        -3148.39    3402.01  -0.925  0.35497
E4:G11        -1661.33    3524.90  -0.471  0.63753
E4:G12         6877.21    3587.53   1.917  0.05554 .
E4:G13         2884.61    3447.48   0.837  0.40295
E4:G14        -6214.29    3513.11  -1.769  0.07723 .
E4:G15         1069.31    3469.99   0.308  0.75803
E4:G16        -1879.90    3549.06  -0.530  0.59645
E4:G17        -2179.49    3586.59  -0.608  0.54355
E4:G18         -436.64    3510.19  -0.124  0.90103
E4:G19         2755.39    3460.45   0.796  0.42608
E4:G20          860.45    3304.31   0.260  0.79461
G1:G2         11508.73   11151.55   1.032  0.30232
G1:G3          5169.80   10636.98   0.486  0.62706
G1:G4         10932.16   11096.60   0.985  0.32479
G1:G5          3558.43   11105.91   0.320  0.74873
G1:G6          5973.17   10965.20   0.545  0.58606
G1:G7          4851.33   11044.44   0.439  0.66058
G1:G8           826.06   11701.48   0.071  0.94374
G1:G9         -4919.17   11074.42  -0.444  0.65700
G1:G10        -8221.62   10825.63  -0.759  0.44777
G1:G11        -9516.01   11176.61  -0.851  0.39475
G1:G12       -14933.82   11447.19  -1.305  0.19235
G1:G13       -11645.43   10769.31  -1.081  0.27981
G1:G14         8236.35   10663.06   0.772  0.44006
G1:G15       -11810.05   10885.93  -1.085  0.27824
G1:G16        13792.60   11290.11   1.222  0.22214
G1:G17         2356.03   11080.08   0.213  0.83166
G1:G18        -6278.34   10809.60  -0.581  0.56150
G1:G19        -8992.23   11029.16  -0.815  0.41510
G1:G20        -3691.39   10852.93  -0.340  0.73384
G2:G3          8978.99   11014.50   0.815  0.41516
G2:G4           839.27   11205.89   0.075  0.94031
G2:G5          -465.26   10941.96  -0.043  0.96609
G2:G6        -13998.40   10662.62  -1.313  0.18955
G2:G7        -12975.59   10948.13  -1.185  0.23624
G2:G8         -6868.30   11569.76  -0.594  0.55289
G2:G9        -14731.71   11126.62  -1.324  0.18582
G2:G10       -13123.47   10586.88  -1.240  0.21543
G2:G11        -9292.69   10674.35  -0.871  0.38421
G2:G12         3975.46   12013.47   0.331  0.74078
G2:G13        11410.38   10860.11   1.051  0.29368
G2:G14         -997.63   10860.26  -0.092  0.92683
G2:G15         -529.07   10797.24  -0.049  0.96093
G2:G16         -854.60   11159.09  -0.077  0.93897
G2:G17        -8550.26   11455.38  -0.746  0.45561
G2:G18        24335.38   11246.68   2.164  0.03073 *
G2:G19        -5663.53   10982.07  -0.516  0.60618
G2:G20        -5659.43   10773.06  -0.525  0.59947
G3:G4         17962.26   11046.03   1.626  0.10425
G3:G5        -27568.25   11222.07  -2.457  0.01420 *
G3:G6         -4681.77   10925.54  -0.429  0.66837
G3:G7          8805.62   11215.67   0.785  0.43258
G3:G8        -12098.98   11387.50  -1.062  0.28829
G3:G9        -13697.49   10877.32  -1.259  0.20824
G3:G10        11776.38   11204.41   1.051  0.29350
G3:G11          287.36   10926.90   0.026  0.97902
G3:G12       -13272.62   11551.75  -1.149  0.25086
G3:G13         -756.46   10829.08  -0.070  0.94432
G3:G14        14058.88   11158.68   1.260  0.20801
G3:G15       -15168.44   10858.00  -1.397  0.16274
G3:G16        -3062.61   11255.82  -0.272  0.78561
G3:G17         4671.24   11417.32   0.409  0.68253
G3:G18         2958.83   10866.49   0.272  0.78546
G3:G19        10260.59   10944.72   0.937  0.34874
G3:G20         9257.61   11374.72   0.814  0.41592
G4:G5           311.47   10793.23   0.029  0.97698
G4:G6         17720.53   11007.34   1.610  0.10775
G4:G7         13373.33   11112.75   1.203  0.22911
G4:G8         20139.68   11673.24   1.725  0.08480 .
G4:G9         -4744.36   11129.19  -0.426  0.66999
G4:G10         3768.21   10733.99   0.351  0.72563
G4:G11       -10465.08   10896.40  -0.960  0.33709
G4:G12        10783.86   11706.38   0.921  0.35718
G4:G13         8195.59   10960.61   0.748  0.45481
G4:G14        14077.04   10650.79   1.322  0.18659
G4:G15         2166.22   10757.67   0.201  0.84046
G4:G16         1274.36   11278.82   0.113  0.91006
G4:G17         4156.69   11680.84   0.356  0.72203
G4:G18         3756.80   11409.38   0.329  0.74202
G4:G19        -5039.92   11068.95  -0.455  0.64898
G4:G20        22358.76   10744.83   2.081  0.03771 *
G5:G6         11237.62   11379.18   0.988  0.32362
G5:G7         -4887.45   11446.17  -0.427  0.66948
G5:G8         12509.58   11874.96   1.053  0.29241
G5:G9          7012.20   11262.56   0.623  0.53369
G5:G10        10349.59   10823.62   0.956  0.33921
G5:G11        -5290.90   11290.64  -0.469  0.63946
G5:G12        -3870.03   11449.52  -0.338  0.73543
G5:G13        11288.06   11193.01   1.008  0.31347
G5:G14        -2445.10   10991.72  -0.222  0.82401
G5:G15        -6926.59   10833.16  -0.639  0.52272
G5:G16         4181.92   11540.27   0.362  0.71715
G5:G17        14889.87   11395.78   1.307  0.19166
G5:G18        17612.40   11367.50   1.549  0.12163
G5:G19        -6931.07   11159.72  -0.621  0.53470
G5:G20        -5994.74   10828.27  -0.554  0.57997
G6:G7         11116.54   10936.70   1.016  0.30968
G6:G8         10996.79   11584.40   0.949  0.34272
G6:G9        -18145.06   11358.10  -1.598  0.11048
G6:G10        17299.87   11266.99   1.535  0.12501
 [ reached getOption("max.print") -- omitted 101 rows ]
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 67880 on 952 degrees of freedom
Multiple R-squared:  0.6967,    Adjusted R-squared:  0.6011
F-statistic: 7.288 on 300 and 952 DF,  p-value: < 2.2e-16
```

MD Mahmudur Rahman

Code for the Stepwise Regression, Modeling the Summary and the Table for Sig Coefficient

```
> library(leaps)
> M <- regsubsets( model.matrix(M_trans)[,-1], I(SP21_P2_949881$Y^(2))),
Error: unexpected ',' in "M <- regsubsets( model.matrix(M_trans)[,-1], I(SP21_P2_949881$Y^(2))),"
>                  nbest = 1 , nvmax=5,
Error: unexpected ',' in "                 nbest = 1 ,"
>                  method = 'forward', intercept = TRUE )
Error: unexpected ',' in "                 method = 'forward',"
> M <- regsubsets( model.matrix(Genetic_trans)[,-1], I(SP21_P2_949881$Y^2)), nbest = 1 , nvmax=5, method = 'forward',
 intercept = TRUE )
Error: unexpected ',' in "M <- regsubsets( model.matrix(Genetic_trans)[,-1], I(SP21_P2_949881$Y^2)),"
> M <- regsubsets( model.matrix(Genetic_trans)[,-1], I(SP21_P2_949881$Y^2),nbest = 1 , nvmax=5,
+                  method = 'forward', intercept = TRUE )
> temp <- summary(M)
> Var2 <- colnames(model.matrix(M_trans))
> M_select2 <- apply(temp$which, 1,
+                    function(x) paste0(Var2[x], collapse='+'))
> kable(data.frame(cbind( model = M_select2, adjR2 = temp$adjr2, BIC = temp$bic)),
+       caption='Model Summary')
```

Table: Model Summary

model	adjR2	BIC
(Intercept)+E1:E3	0.578146731094707	-1068.19606167482
(Intercept)+E1:E3+G9:G14	0.590578080300737	-1099.54366989543
(Intercept)+E3+E1:E3+G9:G14	0.602374779809626	-1130.04627028489
(Intercept)+E3+E1:E3+G2:G8+G9:G14	0.604331544130127	-1130.09797341833
(Intercept)+E3+E1:E3+G2:G8+G9:G14+G17:G18	0.606129517267023	-1129.67587069904

```
> M_main <- lm( I(Y^2) ~ ., data=SP21_P2_949881)
> temp <- summary(M_main)
> kable(temp$coefficients[ abs(temp$coefficients[,4]) <= 0.001, ], caption='Sig Coefficients')
```

Table: Sig Coefficients

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	410.085406	25.9487656	15.803658	0.000000
E1	14.546641	0.5620162	25.882956	0.000000
E3	19.417578	0.5687037	34.143574	0.000000
G14	6.722998	1.8180401	3.697937	0.000227

```
> M_1st <- lm( I(Y^2) ~ (.), data=SP21_P2_949881)
> temp  <- summary(M_1st)
> kable(temp$coefficients[ abs(temp$coefficients[,4]) <= 0.001, ], caption='1st Interaction')


Table: 1st Interaction

|            |   Estimate| Std. Error|   t value| Pr(>&#124;t&#124;)|
|:-----------|----------:|----------:|---------:|------------------:|
|(Intercept) | 410.085406| 25.9487656| 15.803658|           0.000000|
|E1          |  14.546641|  0.5620162| 25.882956|           0.000000|
|E3          |  19.417578|  0.5687037| 34.143574|           0.000000|
|G14         |   6.722998|  1.8180401|  3.697937|           0.000227|
> M_2stage <- lm( I(Y^2) ~ (E3+E1+E3+G9+G14)^2, data=SP21_P2_949881)
> temp <- summary(M_2stage)
> temp$coefficients[ abs(temp$coefficients[,3]) >= 4, ]
    Estimate   Std. Error      t value     Pr(>|t|)
1.780846e+01 4.046001e+00 4.401497e+00 1.167364e-05
> temp$coefficients[ abs(temp$coefficients[,2]) >= 4, ]
              Estimate Std. Error    t value     Pr(>|t|)
(Intercept) 513.311353 200.292591  2.5628075 1.049997e-02
E3           15.319715   8.869358  1.7272630 8.436905e-02
E1            9.908984   8.883250  1.1154683 2.648656e-01
G9           12.824424  40.920885  0.3133956 7.540328e-01
G14         -26.413922  39.205027 -0.6737381 5.006032e-01
G9:G14       17.808464   4.046001  4.4014975 1.167364e-05
>
```