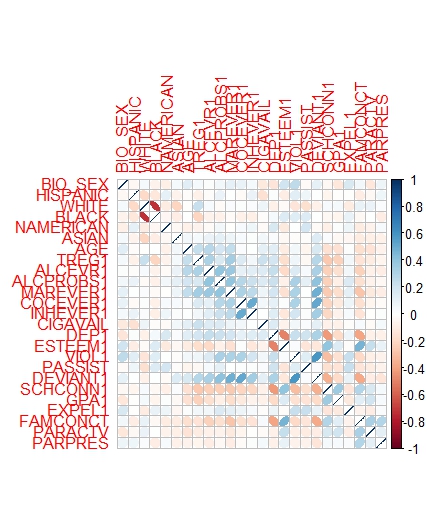
**Q1.**

Figure (correlation matrix)

Figure Tracer Plot

The variables are extracted based on the color of the matrix filler. In the case of SCHCONN1, the three explanatory parameters are colored from teal to light blue. ESTEEM1, GPA1 and FAMCONCT are the three variables with the max correlation to SCHCONN1.

**Q2.**

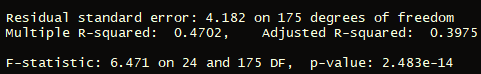
The adjusted R-Squared is quite low along with the multiple R-Squared. The model is not accurate. The value of R-squared and multiple R-squared falls below 50%.

Figure . Summary of the OLS test.

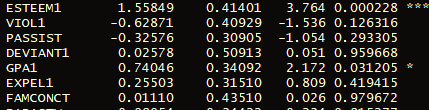
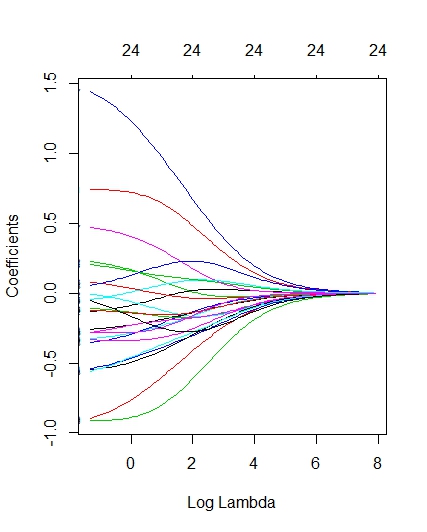


Figure Assessed explanatory variables.

The three correlated variables as selected in Q1 have a positive correlation. ESTEEM 1 has the highest, strongest correlation out of the three.

**Q3.**

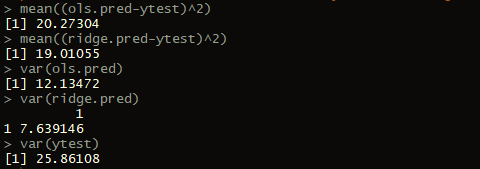
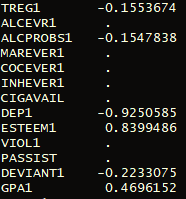
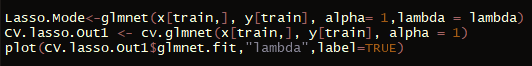


Figure Mean squared error and variance of Ridge regression/OLS

The value, ‘mean’ in the case of OLS is lower than the mean in the case of the ridge test. The variance is smaller in the ridge test compared to the variance of the OLS. As shown in figure 5. A smaller MSE value indicates that the ridge model is more accurate.

Similarly, a smaller variance value of ridge regression suggests that the correlation between variables on the x and y coordinates is stronger.

**Q4.**

Figure. 6 Code to generate the lasso plot.

The lasso trace plot in figure 6 and the Elnet trace plot in figure 9. demonstrates the correlation between the explanatory variables. The right side demonstrating the highest level of correlation.

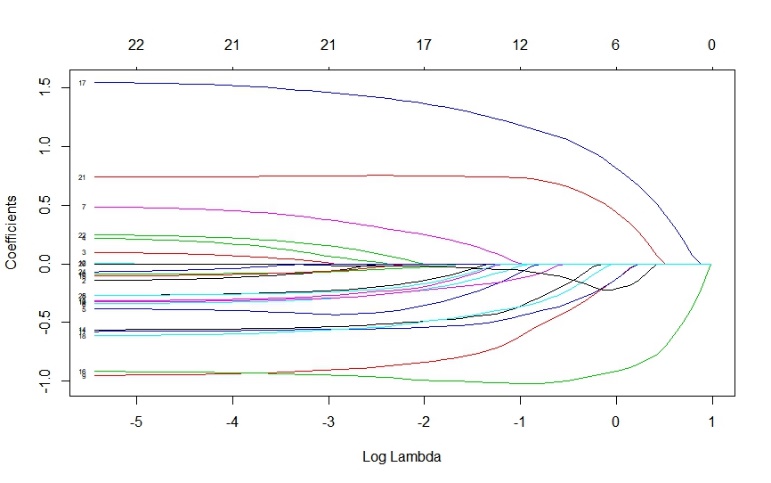
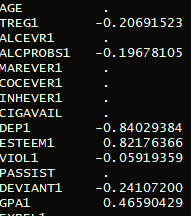


Figure. 10 Correlation using lasso

Figure. 7 Lasso tracer plot

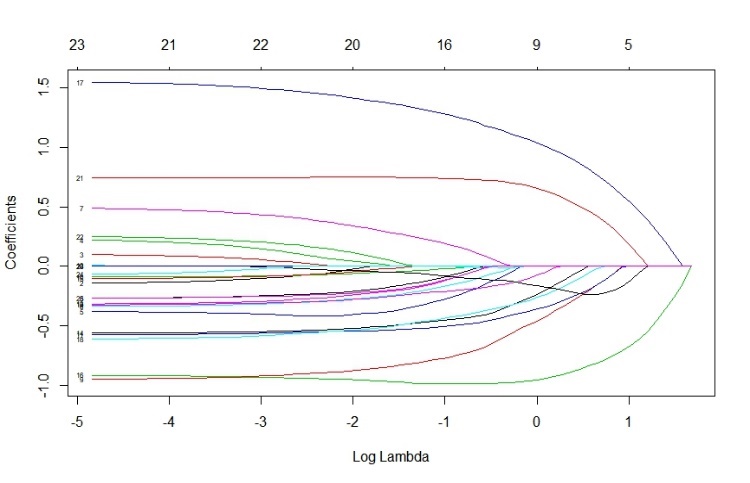


Figure. 11 Correlation using Elastic net

Figure. 9 Elastic Net tracer plot

Figure. 8 Code to generate elastic net tracer plot

The correlation between the elastic net and lasso regression is shown in figure 10 and figure 11, respectively. The variable ESTEEM1, is more strongly related to SCHCONN1 in the lasso. The GPA1 in lasso is more strongly correlated to SCHCONN1 than the GPA1 in elastic net. However, the third term FAMCONC has no correlation in either lasso or elastic net regression. Figure 7 and 9 shows the trace plot. Lasso trace plot has more variables with a positive correlation than the elastic net regression.

**Q5.**

* **Advantages of OLS regression**
* Easily executable.
* Easily implementable.
* Easy to interpret results.
* Fast processing.

**Disadvantages of OLS regression**

* Assumptions about residuals and predictors.
* Overfits the sample
* Unstable during collinearity.
* Unable to solve large data sets with a unique solution
* Works with only with linear models.

**Advantages of Ridge regression.**

* Ridge regression avoids overfitting.
* Ridge regression reduces the standard errors.
* Ridge regression solves the multi-collinearity problems.
* Although biased but has a smaller variance and mean squared error(MSE)
* Good in grouped selection

**Disadvantages of Ridge regression.**

* Reduces coefficient to zero. However, is unable to demonstrate an optimum level of prediction or explanation with as few predictor variables as possible.
* Not good for removing trivial genes.

**Advantages of Lasso.**

* LASSO is a better method of automatic variable selection which involves forward, backward and stepwise. These variable selection techniques do show wrong results; however, lasso results are a lot better.
* Sparsity in parameters.

**Disadvantages of Lasso**

* Lasso is an automatic variable selection method; hence, it has problems.
* It is known to put the analyst at ease compromising the thinking power.
* It is known for creating models which do not make sense.
* Lack of hierarchy principle and it ignored variables which could be useful.

**Advantages of elastic net**

* Elastic Net can deal with scenarios when the number of features is greater than the number of samples, and with correlated features, where LASSO behaves erratically.
* Can select various variables without a limitation.
* Uses grouping even though there are highly correlated predictors.
* Uses a combined method which utilizes penalties of ridge and lasso.

**Disadvantages of Elastic net**

* Results in double shrinkage.
* Might select redundant variables.
* Two regularization parameters to validate increasing running time.

**Q6.**

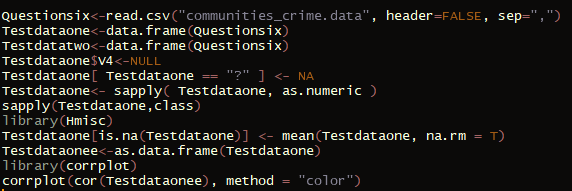
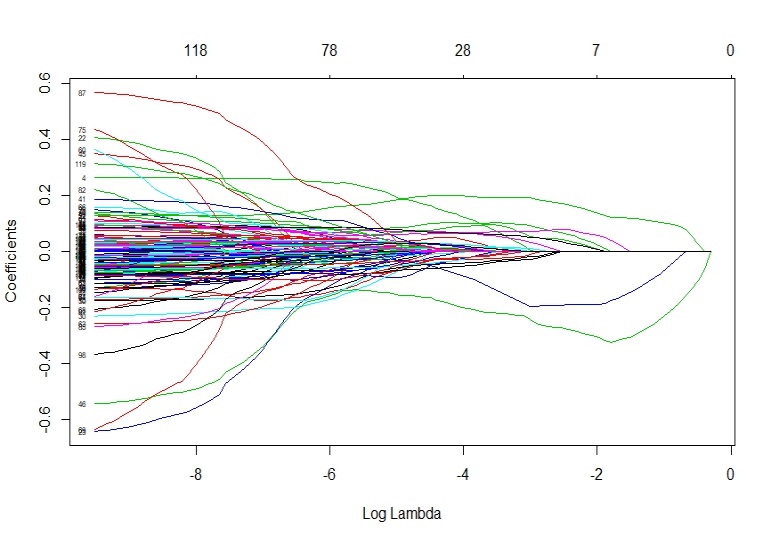


Figure 12. Code to load up the file and plot the correlation matrix.



Figure 13. Correlation matrix

From figure. 16, it can be inferred that the number of variables that are to be plotted against, are quite large. This makes the graph cluttered blocking a clear view of the variables and the correlation.

**Q7.**

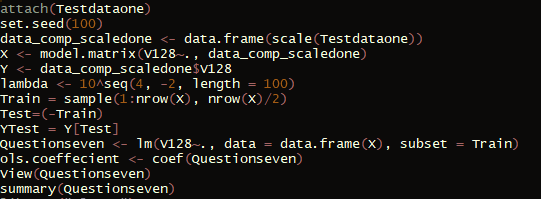


Figure. 14 Splitting data in half

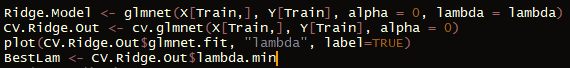
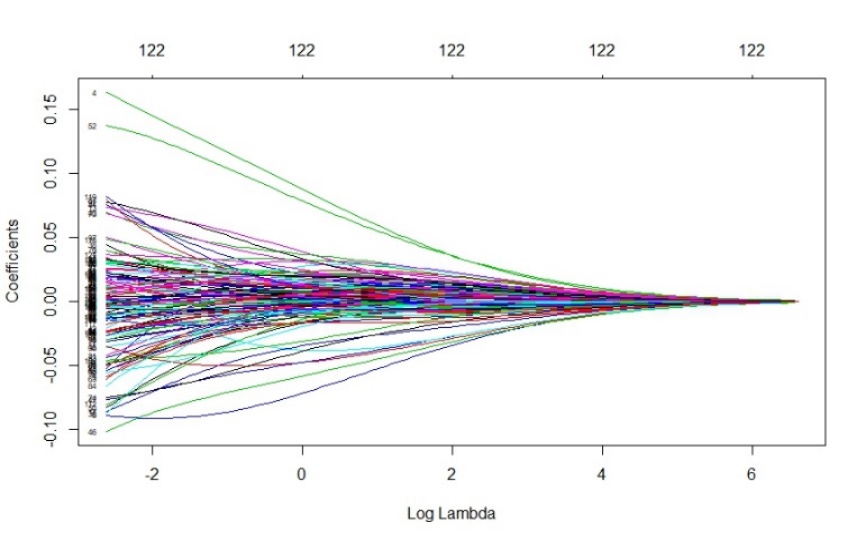


Figure. 21 Elastic net plot code

Figure 18. Ridge tracer plot

Figure. 17 Ridge tracer plot code

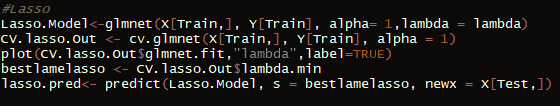
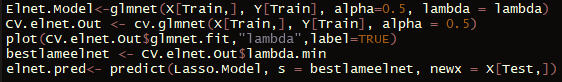
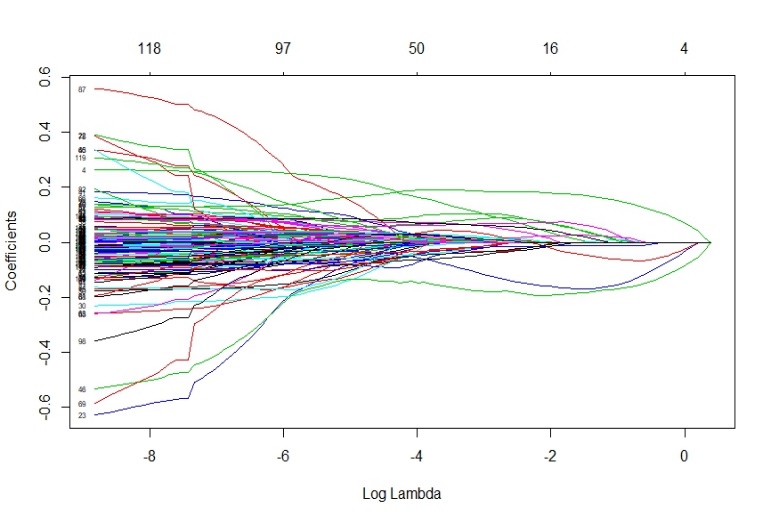


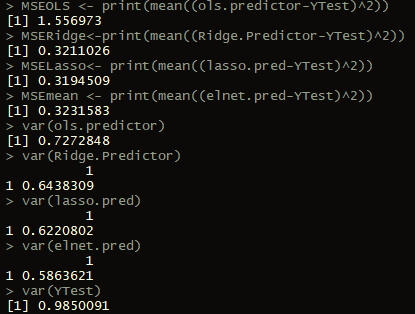
Figure. 22 Elastic net tracer plot

Figure. 19 Lasso regression plot code

Figure. 23 Mean squared error and variance for OLS,Ridge,Lasso and Elastic Net.

 Figure. 20 Lasso regression tracer plot.



From the output of the MSE and variance in figure. 23, it can be concluded that best result is from lasso. This is followed by the ridge, elastic net and then OLS. Figure 18. Shows the shrinkage due to penalization of the coefficients. The least efficient ones shrink the fastest.

In LASSO as shown in figure 20, all the coefficients can be seen in the model. The larger the value of lambda the fewer the number of variables. Due to the large difference between the lasso and the OLS value, it can be stated that the lambda value is not small. Also due to lasso having the smallest MSE value it can be determined that lasso is good at picking up small signals through a lot of noise.

It can also be extracted from the ridge trace plot that without an iterative process it is difficult to solve the problem of multicollinearity.

The values of the variances are the highest for ridge and lowest for elastic net. Having a smaller value, such in the case of LASSO basically tells us how sparse the data is, in terms of making interpretation and variable selection a lot easier.

The elastic net uses the attributes of lasso in terms of variable selection and the ridge attributes in terms of shrinkage of correlated predictors.

The ridge regression includes all the predictors in the final model. Most the penalty terms have gone close to zero but not zero. This makes it difficult to accurately predict the model and interpret the results. Lasso overcomes this by pushing the coefficients to zero shrinking and allowing variable selection. never exactly to zero. This isn’t generally a problem for prediction accuracy, but it can make the model more difficult to interpret the results. Lasso overcomes this.

Figure. 23 MSE and Variances of OLS,Ridge,Lasso and Elastic-net