UNIVERSITY OF MIAMI
MIAMI HERBERT
BUSINESS SCHOOL

MACHINE LEARNING METHODS TO FORECAST STOCK MARKETS

Authors:

- Axel Adriel Peluso Nieto
- Brian Zurita
- Carlos Enrique Egues Dubuc
- Liping Song

FIN 686-57
Advanced Financial Modeling
Daniel McGibney
May 5, 2022

Abstract

In this report, we detail the analysis of five ETFs, four of which stand for different countries and the last which represents the volatility index. We created correlation matrices for each feature variable to determine which variable was the most highly correlated to y_Today. Following this, using eight different methods, we compute this information using the feature variables in relation to the response variables. Each model is represented by a corresponding graph that depicts its results. The outcome is to determine which is the optimal model and compare the Root Mean Square Error and Mean Absolute Error for each response variable. It is found that for response variables of y_Today and y_High, the optimal method is Random Forest, while the optimal method for y_Low is Elastic Net.

Introduction

In order to help us mimic the roles of a Hedge Fund analyst, we analyzed the stock market data of varying ETFs, which will be further considered as the "response variables." The purpose of this report is not only to highlight the relationships between the varying response variables, but the "feature variables" such as the previous day's gain, the percent increase in the volume of the previous day, the previous day's price range, etc. The feature variables are compiled from various statistics of relative importance to the current day's percentage increase over the previous day's close of the S&P 500. Each of the feature variable and response variables was compiled from the information of the S&P 500.

The S&P 500, or the Standard and Poor's 500, is equity indices that follow the performance of the 500 largest companies on the US stock exchange. This is the most commonly used benchmark for valuations of stock portfolios and the market as a whole. Since it is a combination of varying companies, it is considered to be a safe investment due to its diversification. While it may be considered safe, it will not provide the highest returns, most likely only returning its growth rate. The index's versatility can offer the possibility of providing optimal results to predicting the future, which would allow investors the ability to have even higher returns.

Our report as a whole provides value to the data being analyzed by offering a detailed look at how various models interact. The models are reviewed by focusing on the Root Mean Squared Error (RMSE) and the Mean absolute error (MAE) which determine how the distance of actual and predicted values and the accuracy of the values, respectively. This led us to determining which models guaranteed the best results. As will be seen below, some models do not have the optimization that is needed to analyze large data sets, which leads to errors. Certain models are designed to process data in the most efficient way possible.

The data that was provided and the combination of methods that were tested provide an accurate picture of how the world of business analytics works. Of course, the data was engineered for ease of use, but there could have been more interactions included if needed. The models listed are only a small portion of what could be used to analyze the leading market benchmark.

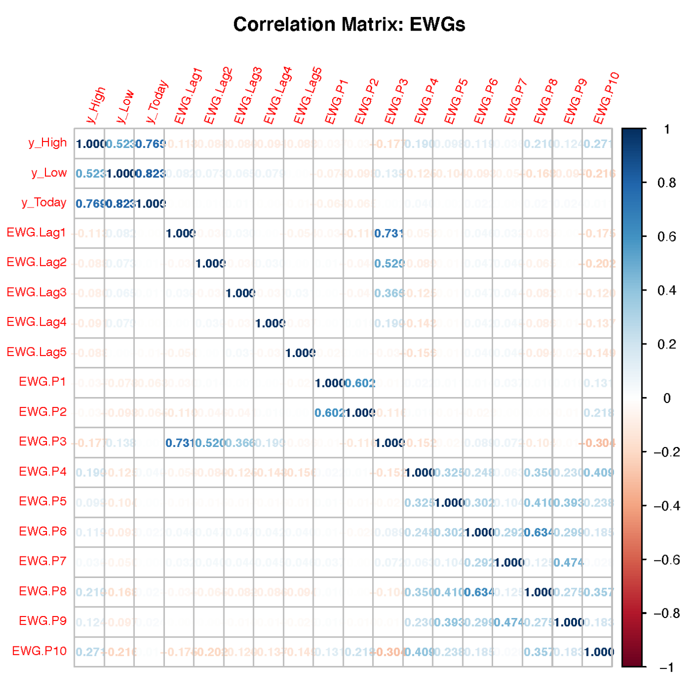
Analysis (describe the analysis of the data or after the regression modeling, i.e. correlation matrix, scatter plots, bar charts, outlier testing, supplemental modeling)

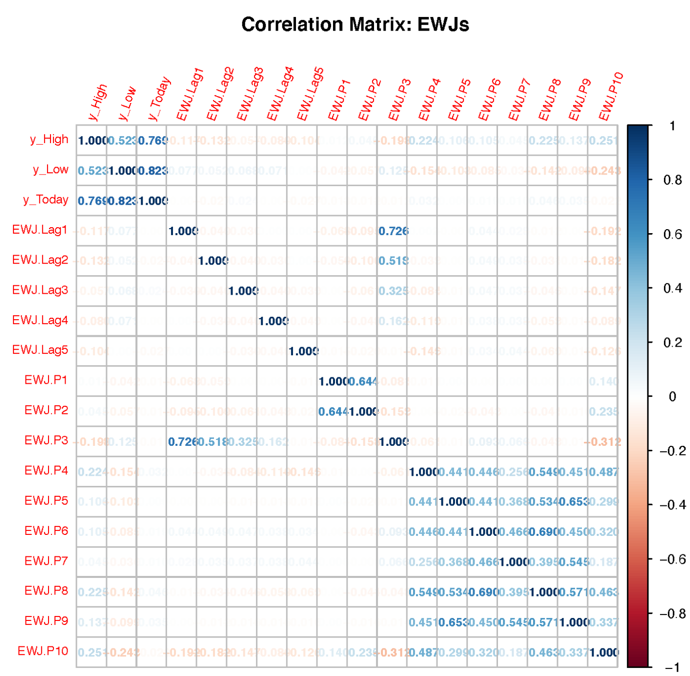
Bullet 1: Which feature variables are most correlated with the response variable ('y_Today')?

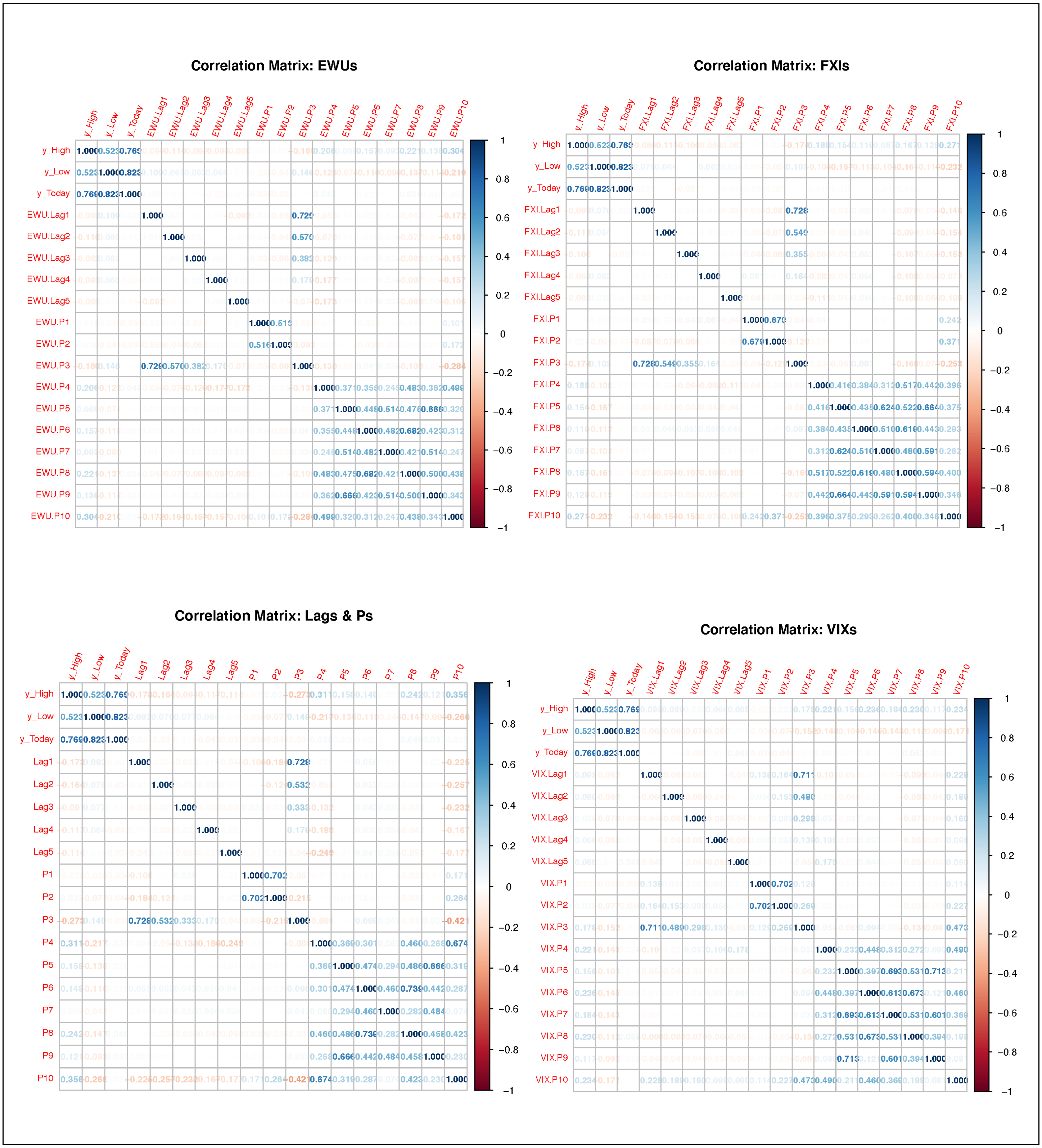
FEATURE	MOST CORRELATED
Lags & Ps	+P8 (0.04814)
FXIs	-Lag3 (-0.03659)
EWUs	+P4 (0.04281)
EWJs	+P8 (0.04576)
EWGs	-P2 (-0.06498)
VIXs	+P8 (0.05295)

As detailed, we reviewed each feature variable and its correlation to y_Today. The list includes both positive and negative correlations; our key to choosing the correct variables is which had the highest value. Listed below are each matrix for the various features, which numerical representation of the correlations. Y_Low and Y_High were included into the matrix to offer a better picture and to determine if the features had similar correlations to the features.

Correlations graphs by features:



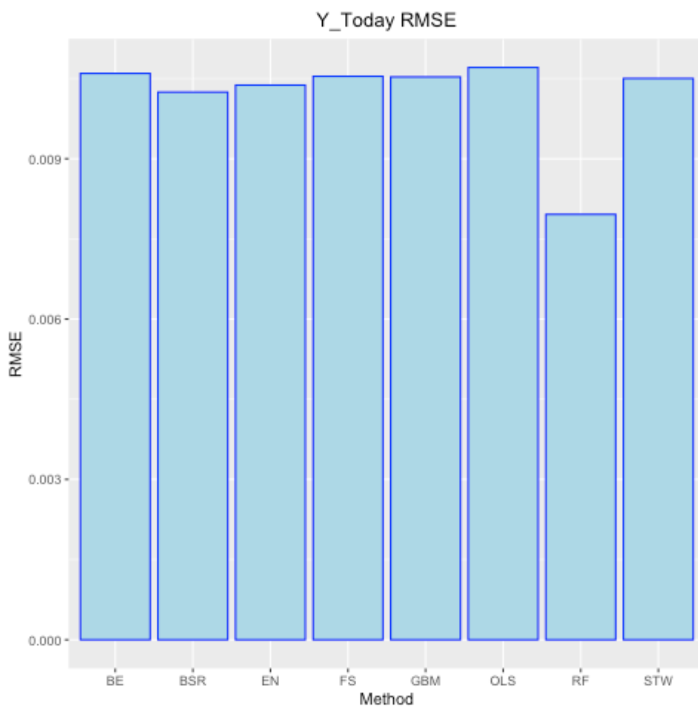



Correlation Matrix: EWUs
Correlation Matrix: FXIs
Correlation Matrix: Lags & Ps
Correlation Matrix: VIXs

Bullet 2: *Fit 4 different regression models on the 'y_Today' response variable and compare the resulting RMSE and MAE values on the test data set.*

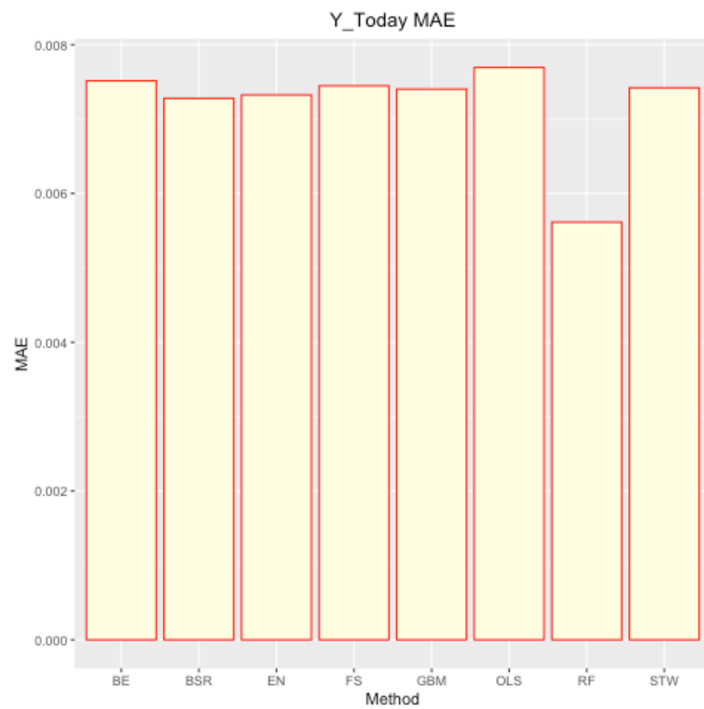
Y_Today RMSE

It can be concluded that the optimal method to produce the lowest Root Mean Square error is Random Forest. Using this method will provide us with the best quality of predictions.



y_Today MAE

It can be concluded that the optimal method to produce the lowest Mean Square Error is Random Forest. This information confirms that this method will provide us with the most accurate results.



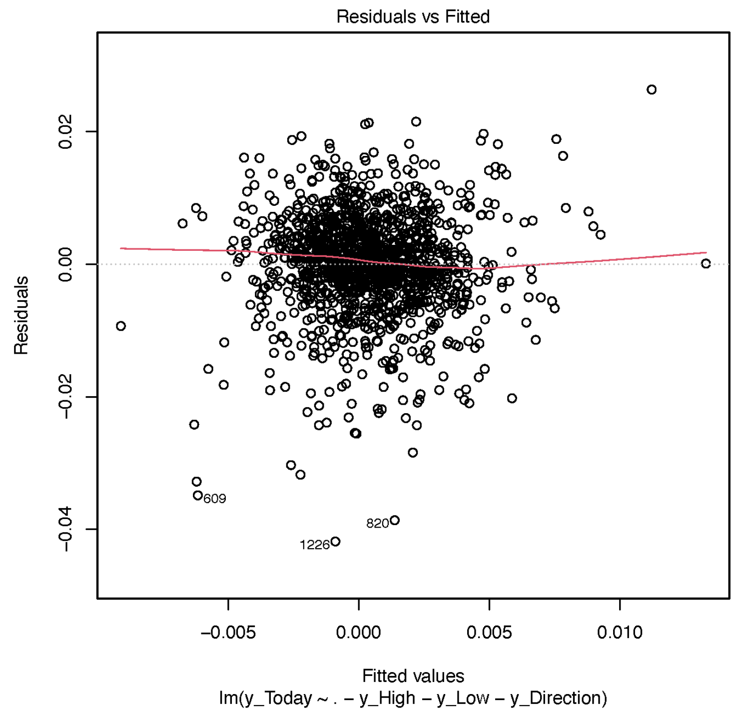
<u>Methods Used</u>

1. **Min Square Root**
2. **Backward Elimination**
3. **Forward Selection**
4. **Stepwise**
5. **Random Forest**
6. **Gradient Boosting Method**
7. **Elastic Net**
8. **Best Subsets**

Min Square Root

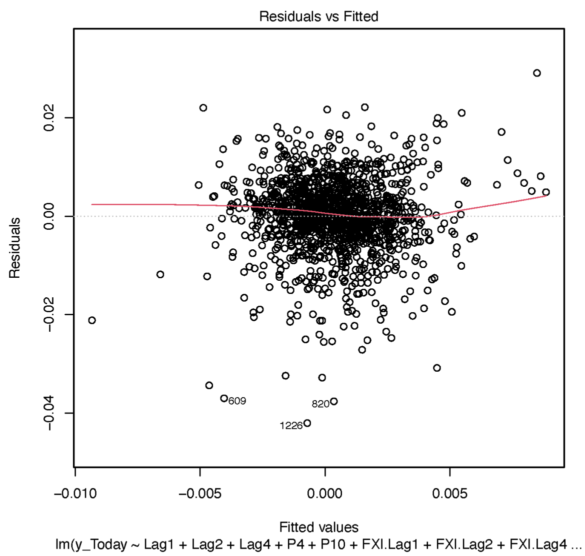
With this model, we can test the strength of the linear correlation. After analyzing the data, the R-Squared states that it has a low positive linear relationship due to its value being close to 0. As can be seen, the only predictor chosen in this instance is P10. Expanded table of predictors can be found within Appendix 1A.)

	y_Today		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	-0.00 – 0.00	0.244
P10	-0.09	-0.26 – 0.07	0.263
Observations	259		
R^2 / R^2 adjusted	0.005 / 0.001		



Linear Regression Model: Backward Elimination

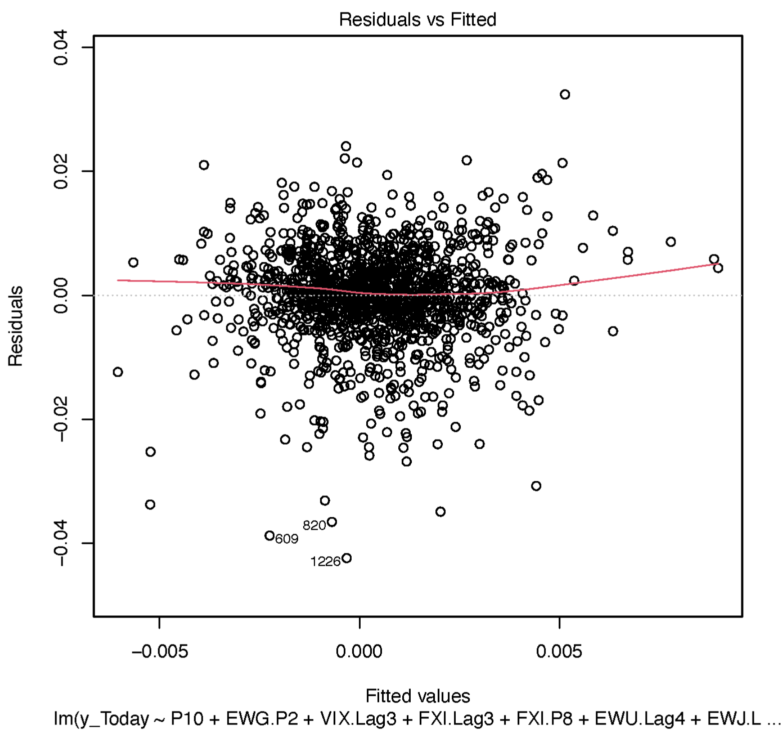
With certain variable selection procedures, only one independent variable is reviewed at a time as variables are eliminated in the process. As such, the regression equation provides us with an output that sits close to (0.00, 0.00) which tells us that there are most variables are correlated to the positive side of regression line, yet since the value is low, it is not strongly correlated. Based upon the adjusted R-Square, this model does not fit the data very well, but of the regression models, this is considered to be the best.



		y_Today	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	-0.00 – 0.00	0.813
Lag1	-0.17	-0.27 – -0.06	**0.002**
Lag2	-0.17	-0.26 – -0.07	**0.001**
Lag4	-0.07	-0.15 – 0.02	0.139
P4	-0.14	-0.32 – 0.03	0.11
P10	0.17	0.05 – 0.29	**0.005**
FXI Lag1	0.16	0.07 – 0.24	**<0.001**
FXI Lag2	0.11	0.04 – 0.17	**0.003**
FXI Lag4	0.04	0.00 – 0.09	**0.047**
FXI P3	-0.16	-0.26 – -0.07	**0.001**
FXI P8	-0.03	-0.05 – -0.01	**0.003**
EWU Lag1	0.18	0.05 – 0.30	**0.006**
EWU Lag3	0.07	-0.02 – 0.17	0.128
EWU Lag5	0.06	0.00 – 0.11	**0.038**
EWU P3	-0.12	-0.25 – 0.01	0.064
EWU P4	0.15	0.04 – 0.27	**0.01**
EWJ Lag3	0.04	-0.01 – 0.09	0.127
EWG Lag1	1.73	0.08 – 3.37	**0.04**
EWG Lag2	1.45	0.22 – 2.69	**0.021**
EWG Lag3	0.9	0.07 – 1.73	**0.033**
EWG Lag4	0.47	0.05 – 0.89	**0.03**
EWG P2	0	-0.00 – -0.00	**0.035**
EWG P3	-2.23	-4.28 – -0.17	**0.034**
VIX Lag5	0.01	0.00 – 0.02	**0.02**
VIX P3	-0.01	-0.02 – -0.01	**0.002**
VIX P7	0	-0.00 – -0.00	**0.021**
VIX P8	0.01	0.00 – 0.02	**0.014**
Observations			1400
R^2 / R^2 adjusted	0.060 / 0.043		

Linear Regression Model: Forward Selection

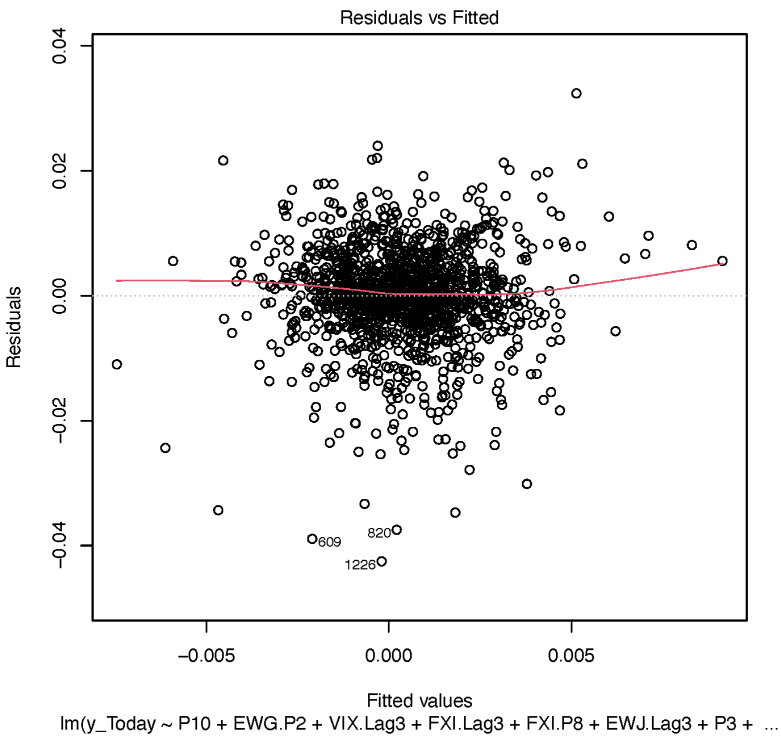
This model starts with no independent variables and slowly adds variables if a reduction in SEE can be found. As seen below, the regression is like that of Backward elimination as the regression line sits close to 0.00 with variables spread across the centre points. Based upon the adjusted R-Square, there is a weak positive linear relationship.



	y_Today		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.00	-0.00 – 0.00	0.256
P10	0.13	0.04 – 0.22	**0.006**
EWG P2	-0.00	-0.00 – -0.00	**0.024**
VIX Lag3	-0.01	-0.02 – -0.00	**0.006**
FXI Lag3	-0.06	-0.09 – -0.02	**0.003**
FXI P8	-0.03	-0.05 – -0.01	**0.002**
EWU Lag4	-0.05	-0.10 – 0.01	0.089
EWJ Lag3	0.06	0.00 – 0.11	**0.036**
FXI Lag1	0.03	-0.01 – 0.06	0.174
P3	-0.22	-0.33 – -0.11	**<0.001**
VIX P3	-0.01	-0.02 – -0.00	**0.004**
EWG Lag2	0.10	0.03 – 0.16	**0.003**
EWU Lag1	0.10	0.02 – 0.17	**0.013**
VIX P7	-0.00	-0.00 – -0.00	**0.007**
VIX P8	0.01	0.00 – 0.02	**0.017**
EWJ Lag4	0.04	-0.01 – 0.09	0.113
EWU Lag2	-0.06	-0.13 – 0.01	0.087
EWG Lag1	-0.06	-0.12 – 0.01	0.099
Observations	1400		
R^2 / R^2 adjusted	0.052 / 0.041		

Linear Regression Model: Stepwise

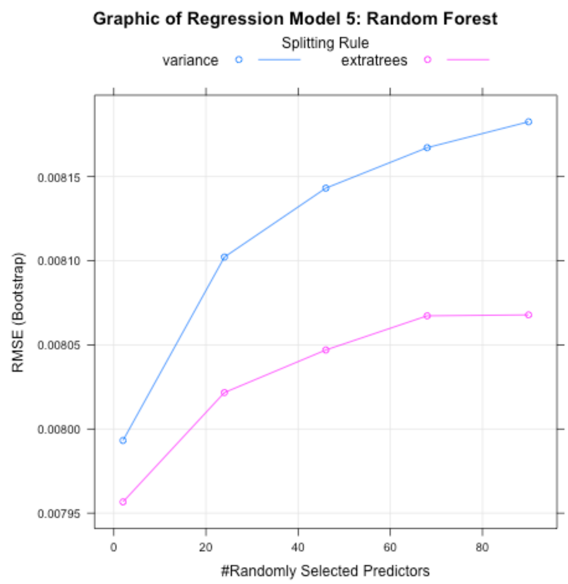
In this model, starting with no variables, variables are added and certain variables may be deleted. The majority of the spread sits in between both the positive and negative sides of the regression line. According to the R-Squared, this is a weak positive correlation.



		y_Today	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.00	-0.00 – 0.00	0.222
P10	0.13	0.04 – 0.22	**0.005**
EWG P2	-0.00	-0.00 – -0.00	**0.024**
VIX Lag3	-0.01	-0.02 – -0.00	**0.004**
FXI Lag3	-0.06	-0.09 – -0.02	**0.002**
FXI P8	-0.03	-0.05 – -0.01	**0.001**
EWJ Lag3	0.05	-0.00 – 0.10	0.057
P3	-0.22	-0.32 – -0.11	**<0.001**
VIX P3	-0.01	-0.02 – -0.00	**0.004**
EWG Lag2	0.09	0.03 – 0.16	**0.005**
EWU Lag1	0.11	0.04 – 0.19	**0.002**
VIX P7	-0.00	-0.00 – -0.00	**0.007**
VIX P8	0.01	0.00 – 0.02	**0.017**
EWU Lag2	-0.06	-0.13 – 0.01	0.096
EWG Lag1	-0.05	-0.12 – 0.02	0.140
Observations	1400		
R^2 / R^2 adjusted	0.049 / 0.039		

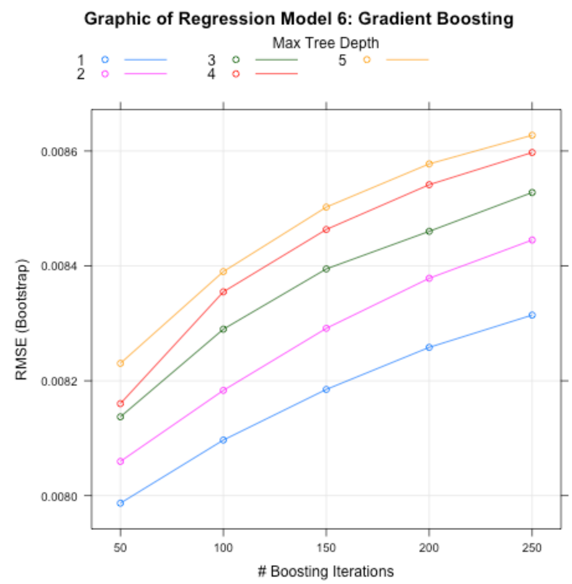
Random Forest

This model is based on decision trees, where variables are split by minimizing RSS. As seen below, the relationship between predictors and response is relatively linear with the RMSE being rather high as the amount of predictors increases.



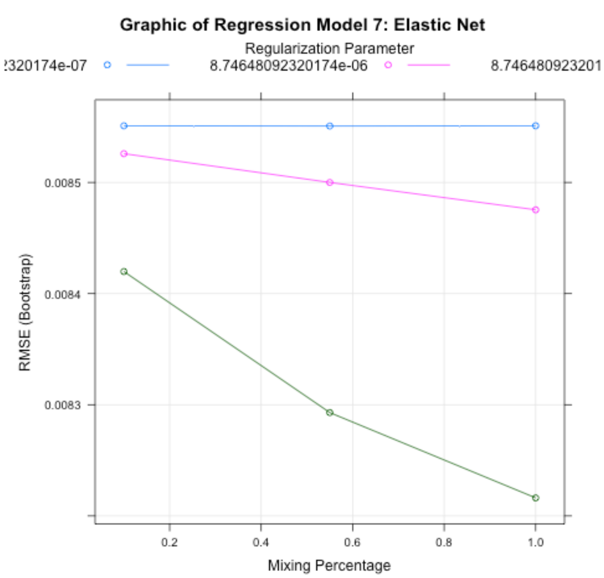
Gradient Boosting Method

In this model, weak variables are strengthened by combining to strong variables. The best possible outcome sits above 0.0086, which indicates this has the highest probability.



Elastic Net

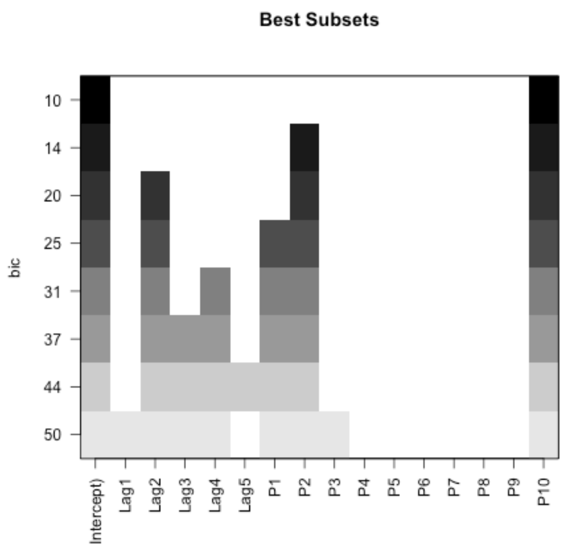
This model focuses on minimizing beta, where variable selection is automized by combining the ridge and lasso methods. We have the mixing percentage with RMSE scores with high values of the regularization parameters.

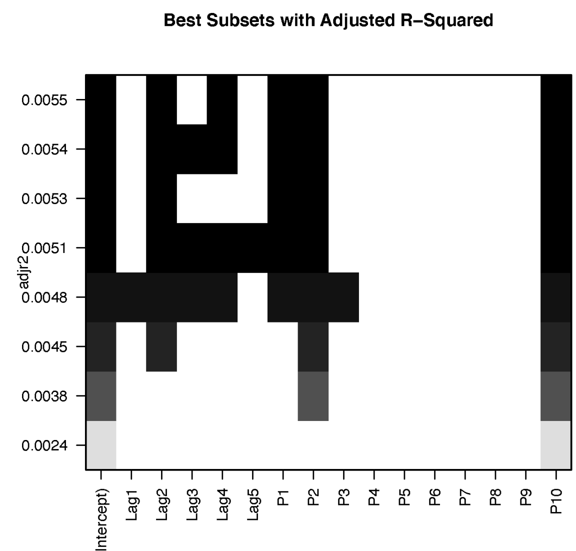


Best Subset

This model analyses different subsets of independent variables, which helps guarantee the best model will be found. In the data provided below, the chosen variables in the subset is P10, but when we use an adjusted R-Square, the variables increase to Lag2, Lag4, P1, P2, and finally P10.

	y_Today		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	-0.00 – 0.00	0.244
P10	-0.09	-0.26 – 0.07	0.263
Observations	259		
R^2 / R^2 adjusted	0.005 / 0.001		

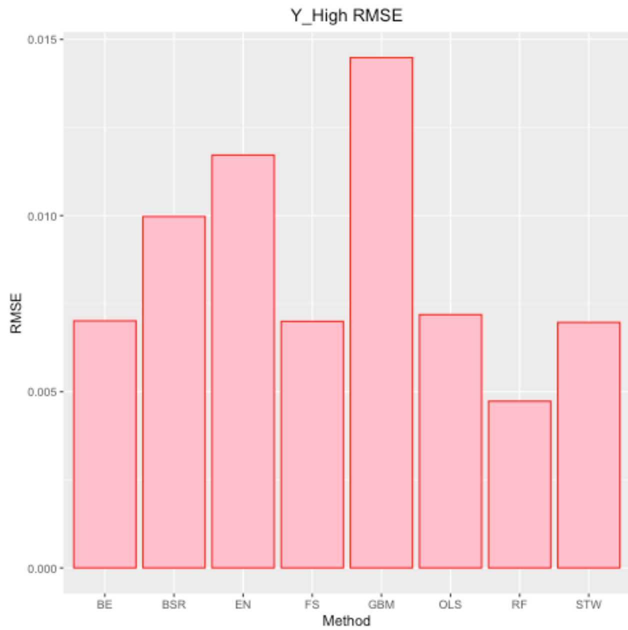




Bullet 3: Fit 4 regression models on both the 'y_High' and 'y_Low' response variables and compare the resulting RMSE and MAE values on the test data set.

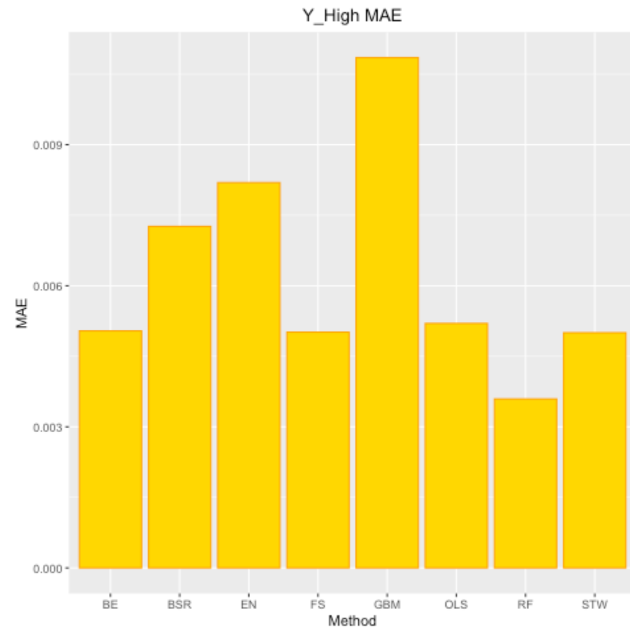
y_High RMSE

Within the graph below, it is determined thar Random Forest will provide the optimal method to obtain a low RMSE.



y_High MAE

As can be seen below, the method that would provide the most accurate results would be Random Forest.

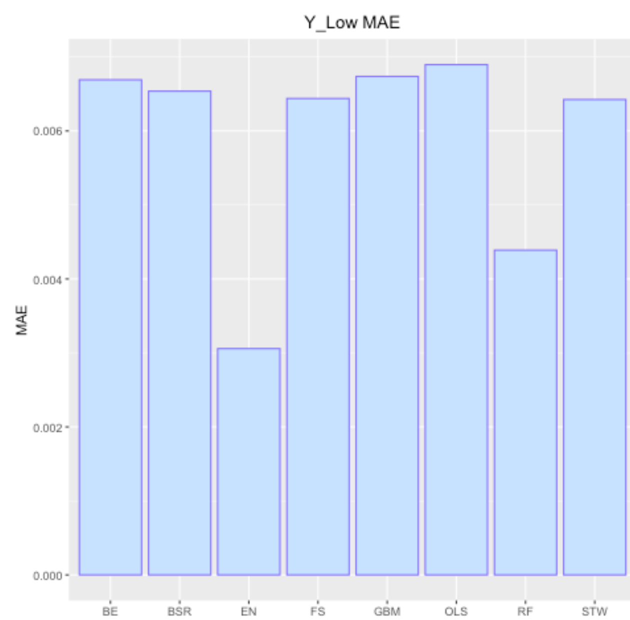


y_Low RMSE

As per the graph below, for y_Low, Elastic Net would provide with the lowest RMSE.

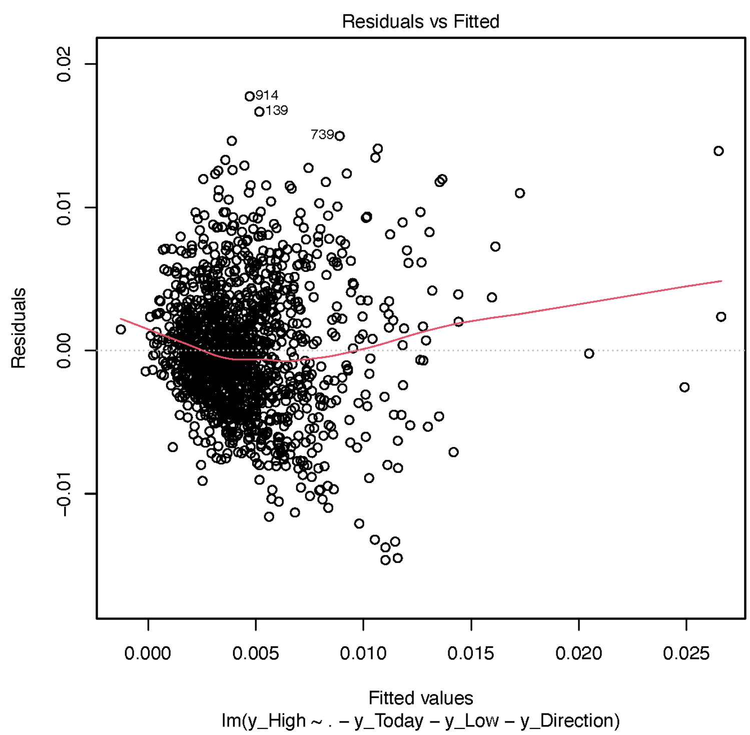
y_Low MAE

As the graph depicts, the Elastic Net method would provide the most accurate results.



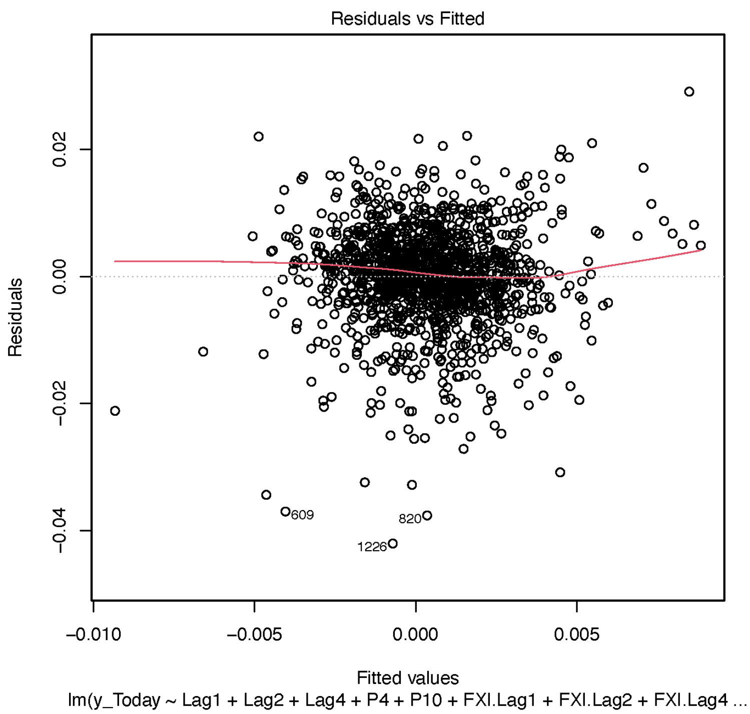
Min Square Root

After analyzing the data, the R-Squared states that it has a low linear correlation due to its value being close to 0, with an r-squared of .0271. It could be considered to have close to no linear correlation, but in comparison to y_Today, this has the higher correlation. (Expanded predictors listed in Appendix 2A)



Linear Regression Model: Backward Elimination

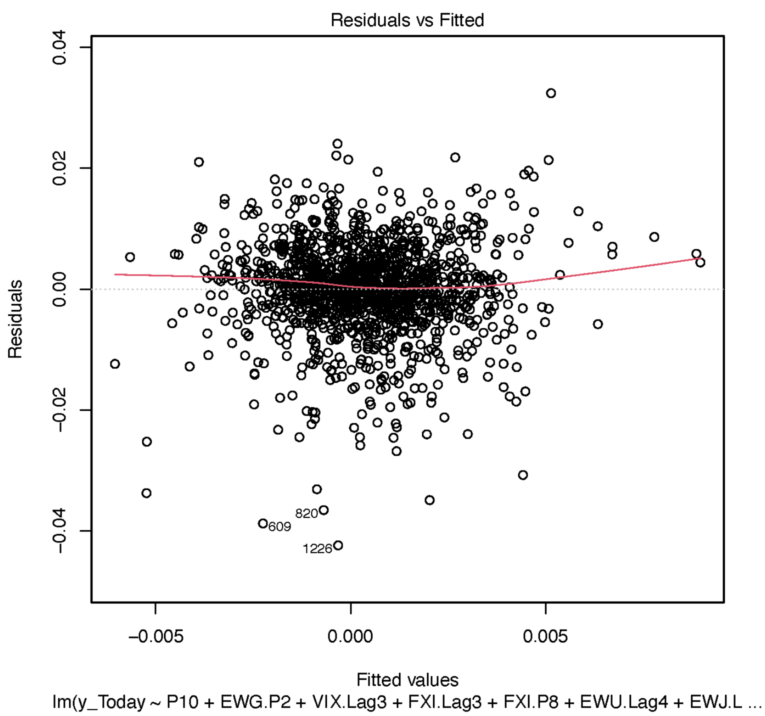
As such, the regression equation provides us with an output that sits close to (0.00, 0.00) which tells us that there are most variables are strongly correlated to the positive side of regression line. Its R-Squared is .255, while positive, would only be considered weak, since it does not pass the 0.5 threshold. This has a lower correlation than Backward Elimination.



		y_High	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	0.00 – 0.00	**0.002**
Lag1	-2.48	-4.86 – -0.10	**0.041**
Lag2	-1.9	-3.69 – -0.11	**0.037**
Lag3	-1.16	-2.35 – 0.03	0.056
Lag4	-0.63	-1.22 – -0.04	**0.038**
Lag5	-0.09	-0.13 – -0.04	**0.001**
P3	2.87	-0.10 – 5.85	0.059
P8	-0.03	-0.06 – 0.00	0.068
P10	0.13	0.04 – 0.21	**0.005**
FXI Lag3	-0.03	-0.06 – -0.01	**0.003**
FXI P4	0.07	0.02 – 0.12	**0.003**
FXI P8	-0.02	-0.03 – -0.01	**0.005**
EWU Lag2	-0.08	-0.13 – -0.02	**0.006**
EWU Lag4	-0.03	-0.08 – 0.01	0.138
EWU Lag5	0.05	0.01 – 0.09	**0.006**
EWU P3	0.1	0.05 – 0.15	**<0.001**
EWU P8	0.03	0.01 – 0.05	**0.006**
EWU P10	0.08	0.00 – 0.16	**0.048**
EWJ P8	0.02	-0.01 – 0.04	0.145
EWJ P9	0.02	0.00 – 0.03	**0.012**
EWG Lag1	1.12	0.22 – 2.03	**0.015**
EWG Lag2	0.94	0.26 – 1.62	**0.007**
EWG Lag3	0.55	0.10 – 1.01	**0.017**
EWG Lag4	0.29	0.06 – 0.52	**0.014**
EWG P2	0	-0.00 – -0.00	**0.009**
EWG P3	-1.45	-2.58 – -0.32	**0.012**
EWG P5	-0.24	-0.38 – -0.11	**<0.001**
EWG P7	-0.02	-0.03 – -0.00	**0.009**
VIX P3	-0.01	-0.01 – -0.00	**0.022**
VIX P8	0.01	0.00 – 0.01	**<0.001**
VIX P9	0	0.00 – 0.01	**0.036**
Observations			1400
R^2 / R^2 adjusted	0.255 / 0.239		

Linear Regression Model: Forward Selection

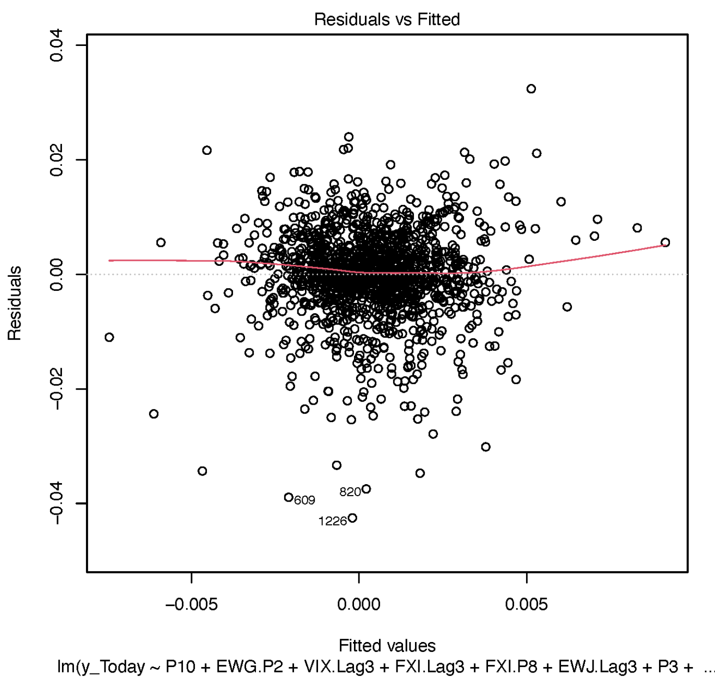
As seen below, the regression is like that of Backward elimination as the regression line sits close to 0.00 with variables spread across the line leaning towards the negative side of the plot. Per the R Squared, this model provides us with a weak positive correlation.



		y_High	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	0.00 – 0.00	**0.014**
P10	0.14	0.06 – 0.23	**0.001**
P3	-0.26	-0.32 – -0.19	**<0.001**
VIX P8	0.01	0.01 – 0.02	**<0.001**
P4	0.08	-0.00 – 0.17	0.063
EWU P3	-0.02	-0.07 – 0.04	0.531
EWG P2	0	-0.00 – -0.00	**0.005**
EWU Lag4	-0.01	-0.03 – 0.02	0.625
VIX Lag5	0	-0.00 – 0.01	0.329
EWJ P9	0.02	0.00 – 0.03	**0.007**
VIX Lag3	0	-0.01 – 0.00	0.437
FXI Lag3	-0.03	-0.05 – -0.00	**0.017**
EWG P5	-0.15	-0.27 – -0.04	**0.009**
EWU P10	0.07	-0.01 – 0.15	0.088
EWG P6	-0.03	-0.05 – -0.01	**0.014**
EWU P8	0.02	-0.00 – 0.04	0.074
EWG Lag5	0.04	0.01 – 0.07	**0.014**
Lag5	-0.05	-0.11 – 0.01	0.123
Lag3	0.09	0.03 – 0.16	**0.005**
VIX P3	0	-0.01 – 0.00	0.138
EWG Lag2	0.06	0.03 – 0.10	**<0.001**
EWU Lag1	0.09	0.04 – 0.15	**<0.001**
Observations			1400
R^2 / R^2 adjusted	0.245 / 0.234		

Linear Regression Mode: Stepwise

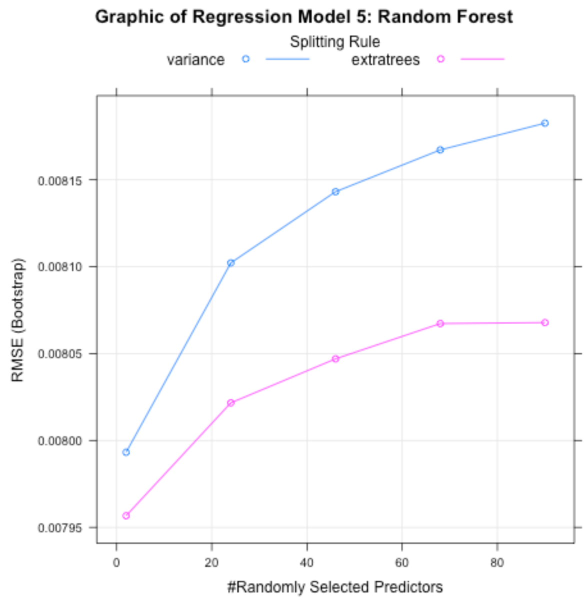
The majority of the spread sits in between both the positive and negative sides of the regression line. This is the weakest of the three regression models presented and would be considered to have a weak positive correlation.



		y_High	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	0.00 – 0.00	**0.01**
P10	0.14	0.05 – 0.22	**0.001**
P3	-0.27	-0.33 – -0.21	**<0.001**
VIX P8	0.01	0.01 – 0.01	**<0.001**
P4	0.09	0.00 – 0.18	**0.044**
EWG P2	0	-0.00 – -0.00	**0.005**
EWJ P9	0.02	0.00 – 0.03	**0.008**
FXI Lag3	-0.03	-0.05 – -0.01	**0.012**
EWG P5	-0.16	-0.27 – -0.04	**0.006**
EWU P10	0.08	-0.00 – 0.16	0.063
EWG P6	-0.03	-0.05 – -0.01	**0.014**
EWU P8	0.02	-0.00 – 0.04	0.061
EWG Lag5	0.04	0.01 – 0.07	**0.017**
Lag5	-0.07	-0.12 – -0.02	**0.003**
VIX P3	-0.01	-0.01 – 0.00	0.055
Lag3	0.11	0.06 – 0.15	**<0.001**
EWG Lag2	0.06	0.03 – 0.08	**<0.001**
EWU Lag1	0.08	0.05 – 0.11	**<0.001**
Observations			1400
R^2 / R^2 adjusted	0.244 / 0.234		

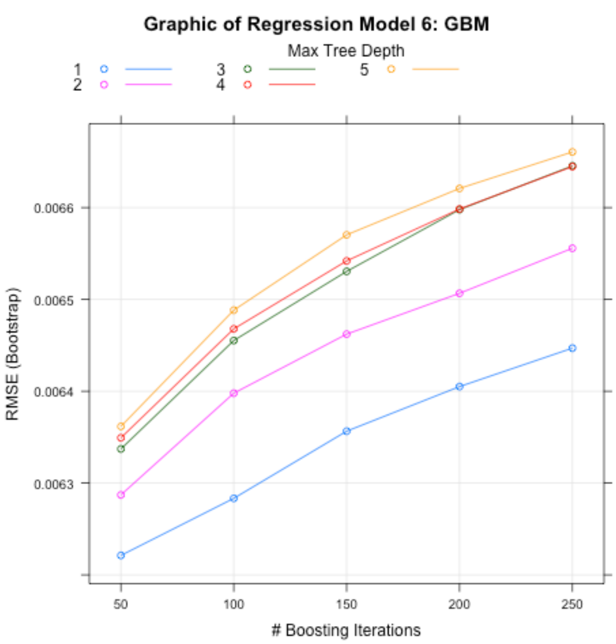
Random Forest

As seen below, the relationship between predictors and response is relatively linear with variances tree having a higher RMSE. This leads us to consider that we have a high accuracy.



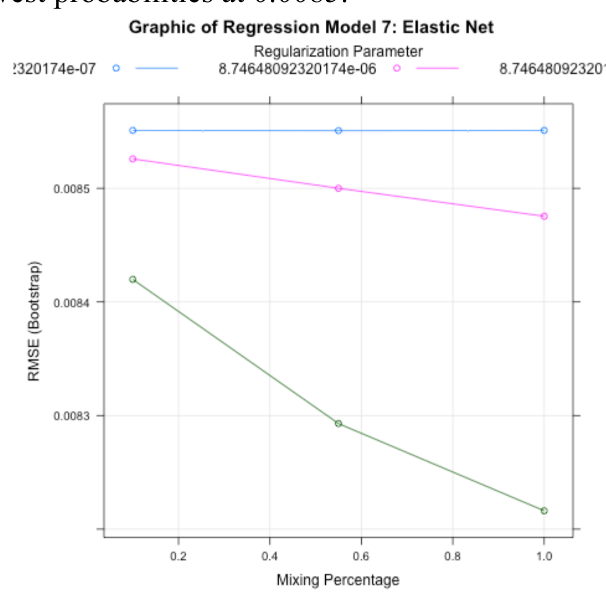
Gradient Boosting Method

The best possible outcome sits above 0.0086, which is amongst our highest predictabilities.



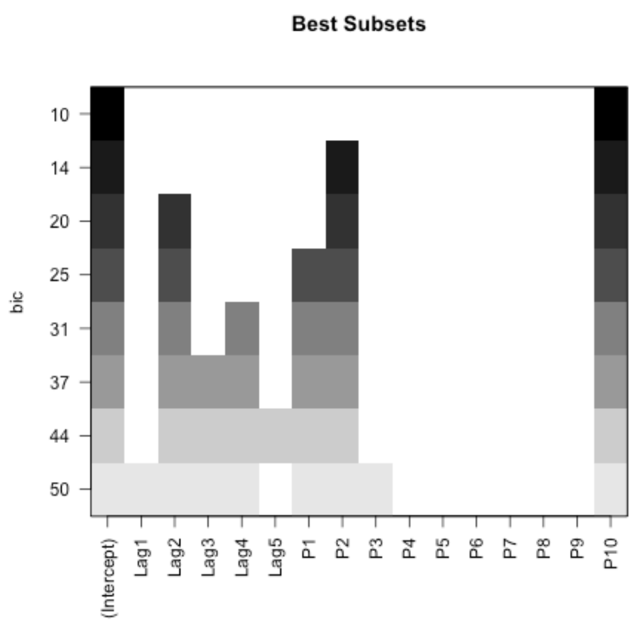
Elastic Net

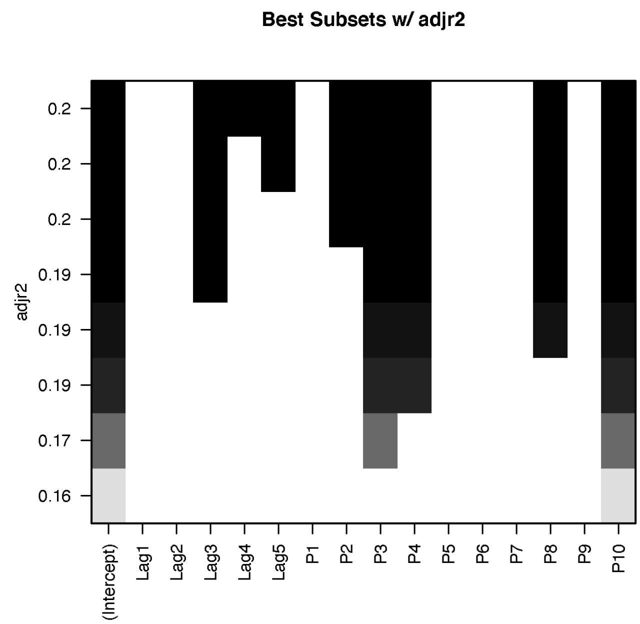
We have the mixing percentage with RMSE scores with high values of the regularization parameters. In comparison with other models, this produces the lowest probabilities at 0.0085.



Best Subset

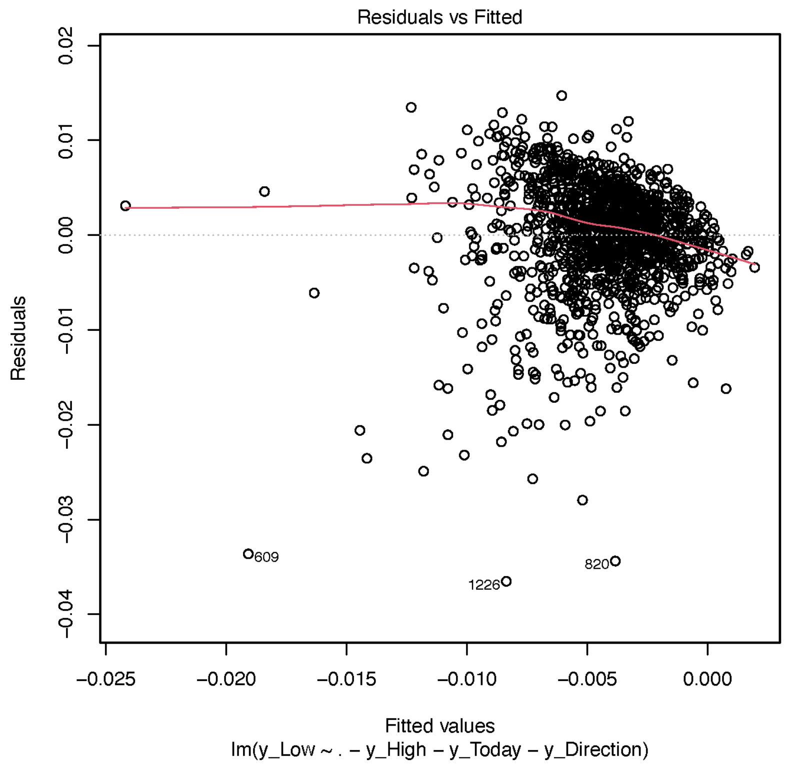
In the data provided below, the best variables to include is P10, but using adjusted R Squared the variables grow to include Lag3, Lag4, Lag5, P2, P3, P4, P8, and P10.





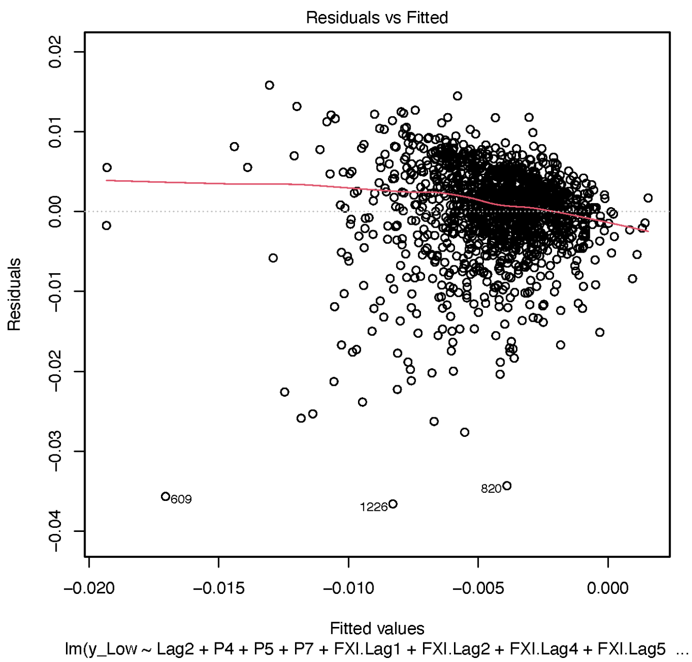
Min Square Root

After analyzing the data, the R-Squared states that it has a low linear correlation due to its value being close to 0. It could be considered to have close to no linear correlation. In comparison to y_Today and y_High, this has the lowest correlation. (Expanded predictor calculations in Appendix 3A).

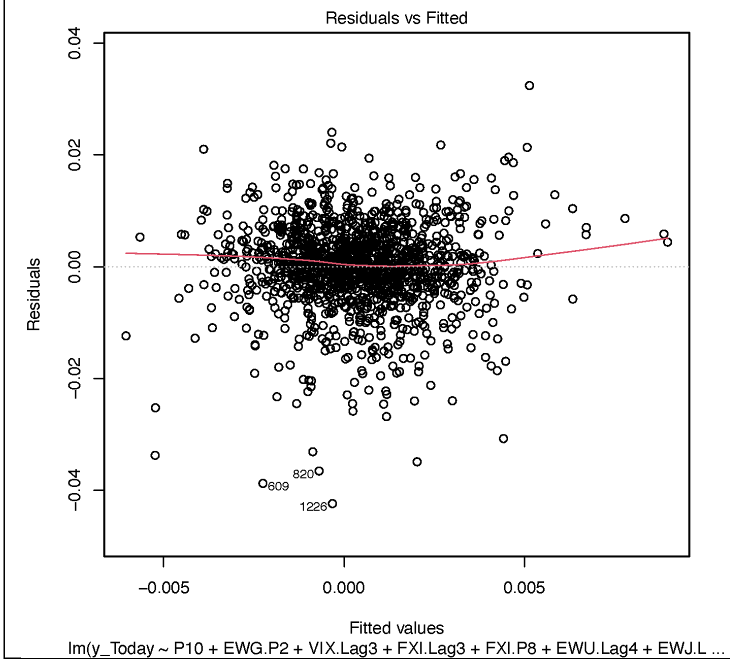


Linear Regression Model: Backward Elimination

The majority of the fitted values trend towards the negative end, but the R-Square states that there is a weak positive correlation.



		y_Low	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	-0.00 – 0.00	0.777
Lag2	-0.1	-0.18 – -0.02	**0.019**
P4	-0.31	-0.43 – -0.18	**<0.001**
P5	-0.59	-1.00 – -0.18	**0.004**
P7	-0.02	-0.05 – 0.00	0.078
FXI Lag1	0.09	0.03 – 0.15	**0.004**
FXI Lag2	0.08	0.03 – 0.13	**0.002**
FXI Lag4	0.02	-0.01 – 0.04	0.157
FXI Lag5	0.02	-0.00 – 0.05	0.069
FXI P3	-0.12	-0.18 – -0.05	**<0.001**
FXI P4	0.08	0.02 – 0.15	**0.013**
FXI P8	-0.03	-0.04 – -0.01	**0.005**
EWU Lag1	0.05	0.00 – 0.10	**0.037**
EWU P4	0.07	-0.01 – 0.15	0.102
EWU P8	0.02	-0.01 – 0.05	0.147
EWJ Lag1	-0.08	-0.16 – -0.00	**0.041**
EWJ Lag2	-0.06	-0.12 – 0.01	0.079
EWJ P3	0.07	-0.01 – 0.16	0.099
EWJ P9	0.02	-0.00 – 0.04	0.125
EWG Lag2	0.04	-0.00 – 0.09	0.067
EWG P1	0	-0.00 – -0.00	**0.038**
EWG P8	-0.02	-0.05 – 0.01	0.111
EWG P9	-0.01	-0.03 – 0.00	0.136
EWG P10	-0.07	-0.15 – 0.02	0.114
VIX Lag1	0.01	0.00 – 0.02	**0.026**
VIX Lag5	0	-0.00 – 0.01	0.067
VIX P3	-0.02	-0.03 – -0.01	**<0.001**
VIX P7	0	-0.00 – 0.00	0.058
VIX P9	0.02	0.00 – 0.03	**0.009**
VIX P10	0.01	-0.00 – 0.01	0.06
Observations			1400
R^2 / R^2 adjusted	0.129 / 0.111		

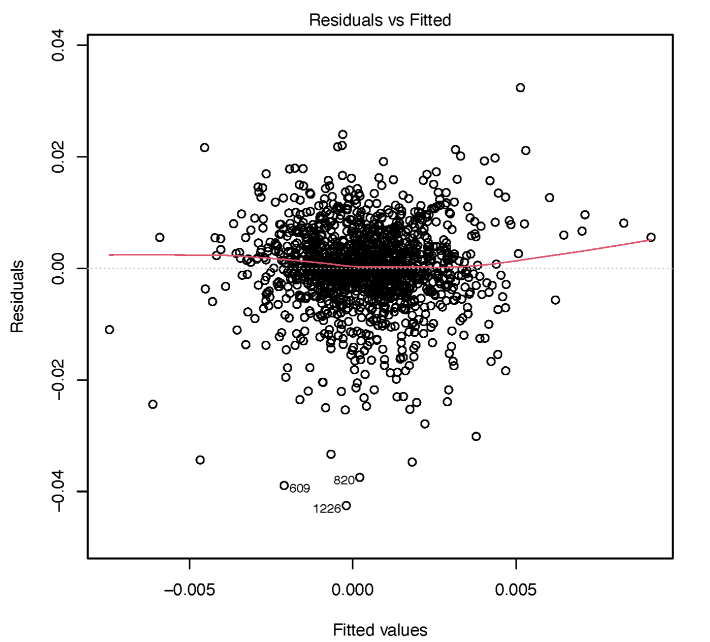


Linear Regression Model: Forward Selection

As seen below, the regression is like that of Backward elimination as the regression line sits close to 0.00 with variables spread across the line leaning towards the negative side of the plot. Of the three regression model, this holds the highest adjusted R-Squared. (Expanded predictors listed in Appendix 3C).

Linear Regression Mode: Stepwise

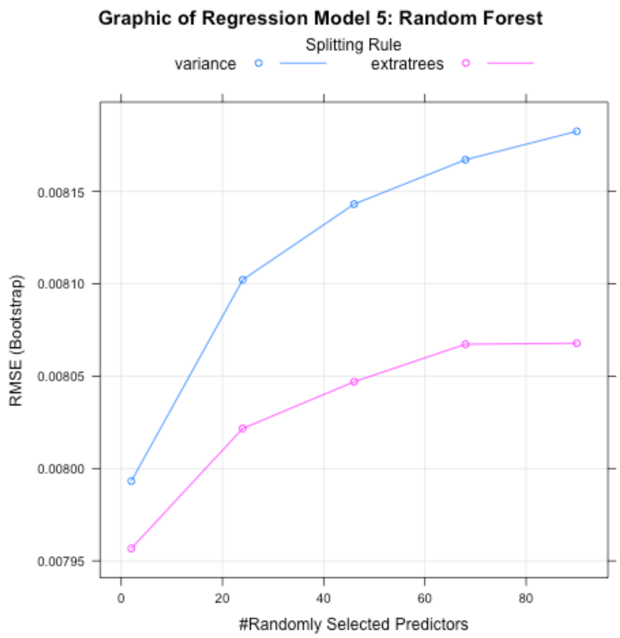
The majority of the spread sits in between both the positive and negative sides of the regression line. While this shows a positive correlation, it is weak.



		y_Low	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	-0.00 – 0.00	0.321
FXI P5	-0.17	-0.27 – -0.06	**0.002**
VIX P3	-0.01	-0.01 – -0.01	**<0.001**
P4	-0.31	-0.40 – -0.21	**<0.001**
FXI Lag3	-0.05	-0.08 – -0.03	**<0.001**
VIX Lag3	0	-0.01 – 0.00	0.094
EWG P2	0	-0.00 – -0.00	**0.015**
FXI P8	-0.03	-0.04 – -0.01	**0.002**
FXI P4	0.09	0.03 – 0.16	**0.004**
EWJ Lag3	0.03	-0.01 – 0.07	0.121
P7	-0.02	-0.04 – 0.00	0.123
Observations		1400	
R^2 / R^2 adjusted		0.111 / 0.104	

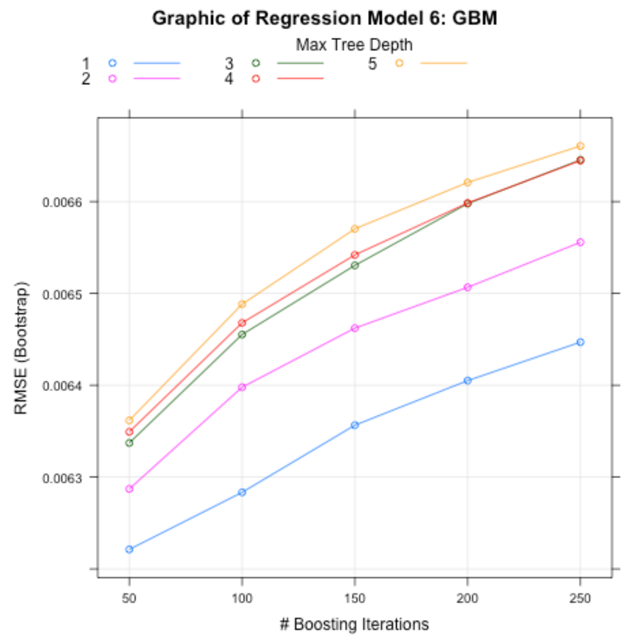
Random Forest

In comparison to y_Today and y_High, the variance for y_Low only rises to just above 0.00815. This is significantly lower than 0.0085.



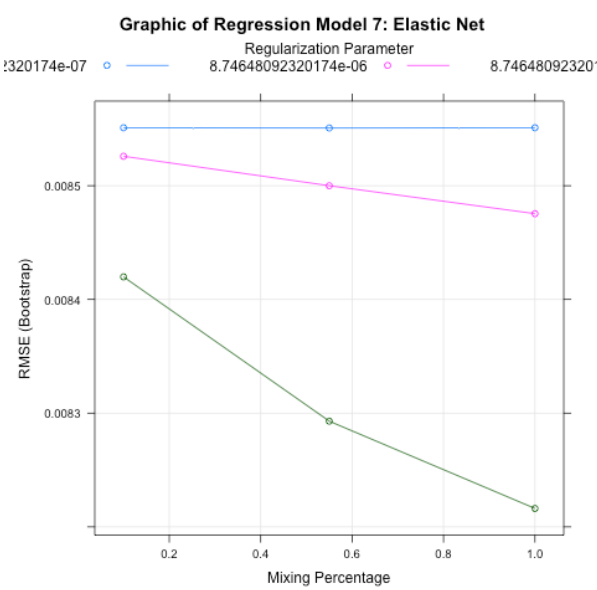
Gradient Boosting Method

In relation to y_Today and y_Low, this method provides the best outcomes with the max probability sitting higher than 0.0066.



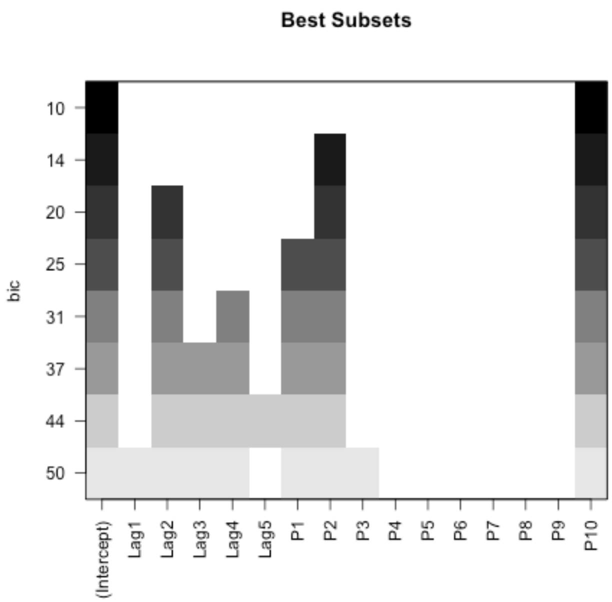
Elastic Net

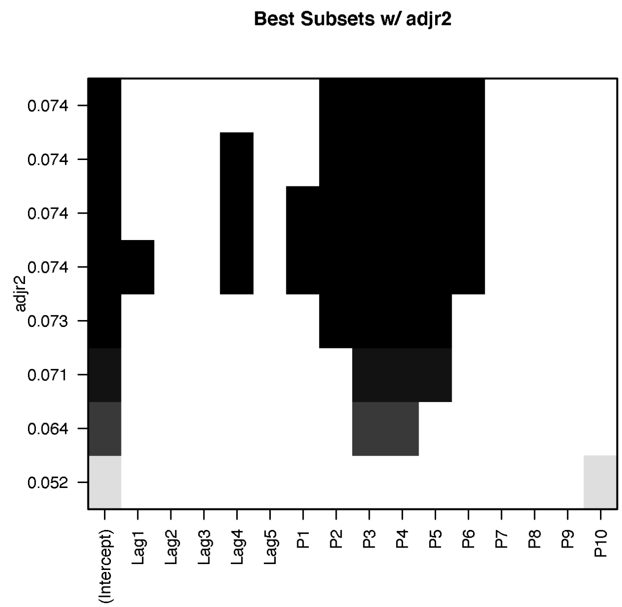
Of the three methods outside of the regression models, this model produces the highest results, which sit above 0.0087. Due to the nature of the model, this model is considered to be the best for y_Low, which is starkly different from y_Today and y_High, which used Random Forest to produce the highest results.



Best Subset

Within the base Best Subsets model, the best variable is P10, but when R Squares are adjusted, the new variables become P2, P3, P3, P4, P5, & P6.





Conclusion:

In conclusion, the dataset provided offered a good opportunity to predict the next day's performance using variables built from the S&P 500. Since the S&P 500 is the most common benchmark, this was the optimal choice for prediction. The dataset provided three response variables, (y_Today, y_High, y_Low), which represents the current day's percentage movement in terms of regular, high, and low increase, respectively. Fifteen feature variables were provided to compare against. The data was first analyzed with a correlation matrix, to give a picture of which feature variables were most correlated to the response variable of y_Today.

Within this report, a total of eight methods that were used to predict the next day's performance instead of the required four, in tandem with a combination of features, provided within the case, interactions, and transformations. Four additional methods were chosen due to the variable selection procedures (Stepwise, Forward Selection, and Backward Elimination) did not provide strong nor accurate results. This can be attributed due to the model being iterative, meaning one independent variable at a time is added or deleted. The addition of these method offered the best results; as for y_Today and y_High the best method is random forest and for y_Low the best method is elastic net. The results are determined by the Root Mean Square Error and the Mean Absolute Error as these two outputs determine the accuracy of the methods. The methods with the lowest error are the optimal choice. While the methods used produced fruitful results and had its potential maximized, there are certain aspects that could be expanded on further.

A potential way to build upon this is to adjust the feature variables and explore more models. For example, it would interesting to include more time periods, such as volume and price gain over the 30 day period and the price ranges for more than just the previous day. The prominent change would be to use more ETFs, including more countries would offer a better picture, a rather worldly view. This would certainly increase the number of observations, but the results would be more accurate. As for the model exploration, there are of course more methods that could be used, such as the KNN method could be used since it can handle many levels of qualitative responses, but this is primarily used for classification and not regression as this report is based upon.

Appendix

Appendix 1: Methods for (y_Today) w/o Interactions

1A: Min Square Root

	y_Today		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.00	-0.00 – 0.01	0.393
Lag1	-2.93	-9.26 – 3.40	0.365
Lag2	-2.27	-7.03 – 2.50	0.351
Lag3	-1.41	-4.57 – 1.76	0.383
Lag4	-0.75	-2.34 – 0.83	0.351
Lag5	-0.06	-0.19 – 0.07	0.344
P1	0.00	-0.00 – 0.01	0.207
P2	-0.01	-0.01 – 0.00	0.066
P3	3.46	-4.48 – 11.40	0.393
P4	-0.30	-0.56 – -0.03	**0.029**
P5	-0.67	-1.74 – 0.39	0.215
P6	0.04	-0.05 – 0.13	0.384
P7	-0.04	-0.09 – 0.02	0.227
P8	-0.02	-0.09 – 0.05	0.513
P9	0.02	-0.02 – 0.07	0.332
P10	0.20	-0.00 – 0.41	0.051
FXI Lag1	0.78	-0.61 – 2.17	0.273
FXI Lag2	0.57	-0.47 – 1.61	0.285
FXI Lag3	0.31	-0.39 – 1.00	0.384
FXI Lag4	0.20	-0.15 – 0.54	0.268
FXI Lag5	0.01	-0.03 – 0.05	0.729
FXI P1	-0.00	-0.00 – 0.00	0.405
FXI P2	0.00	-0.00 – 0.00	0.313
FXI P3	-0.94	-2.67 – 0.80	0.291
FXI P4	0.08	-0.03 – 0.18	0.177
FXI P5	-0.05	-0.49 – 0.38	0.807
FXI P6	0.00	-0.03 – 0.03	0.884
FXI P7	-0.00	-0.02 – 0.02	0.944
FXI P8	-0.03	-0.06 – 0.01	0.114
FXI P9	-0.00	-0.03 – 0.02	0.906
FXI P10	-0.04	-0.17 – 0.08	0.503
EWU Lag1	0.69	-1.64 – 3.02	0.562
EWU Lag2	0.39	-1.36 – 2.14	0.659
EWU Lag3	0.33	-0.84 – 1.51	0.580
EWU Lag4	0.10	-0.50 – 0.69	0.748
EWU Lag5	0.06	-0.02 – 0.14	0.147
EWU P1	-0.00	-0.00 – 0.00	0.603
EWU P2	0.00	-0.00 – 0.00	0.734
EWU P3	-0.76	-3.69 – 2.18	0.613
EWU P4	0.10	-0.09 – 0.28	0.314
EWU P5	0.27	-0.21 – 0.75	0.267
EWU P6	-0.03	-0.09 – 0.04	0.381
EWU P7	-0.02	-0.07 – 0.03	0.431
EWU P8	0.05	-0.01 – 0.10	0.115
EWU P9	0.00	-0.06 – 0.06	0.959
EWU P10	-0.01	-0.20 – 0.18	0.918
EWJ Lag1	1.22	-1.75 – 4.20	0.419
EWJ Lag2	0.92	-1.31 – 3.15	0.417
EWJ Lag3	0.66	-0.82 – 2.15	0.382
EWJ Lag4	0.35	-0.40 – 1.09	0.361
EWJ Lag5	0.01	-0.05 – 0.07	0.707
EWJ P1	0.00	-0.00 – 0.00	0.879
EWJ P2	0.00	-0.00 – 0.00	0.409
EWJ P3	-1.56	-5.28 – 2.16	0.410
EWJ P4	0.03	-0.13 – 0.18	0.718
EWJ P5	0.10	-0.25 – 0.45	0.582

EWJ P6	-0.01	-0.06 – 0.04	0.727
EWJ P7	-0.01	-0.04 – 0.03	0.783
EWJ P8	0.02	-0.03 – 0.07	0.441
EWJ P9	0.04	-0.01 – 0.08	0.109
EWJ P10	0.05	-0.14 – 0.24	0.620
EWG Lag1	1.21	-1.98 – 4.39	0.457
EWG Lag2	1.07	-1.32 – 3.46	0.382
EWG Lag3	0.63	-0.96 – 2.23	0.438
EWG Lag4	0.35	-0.44 – 1.15	0.384
EWG Lag5	0.02	-0.06 – 0.09	0.672
EWG P1	-0.00	-0.00 – 0.00	0.352
EWG P2	-0.00	-0.00 – 0.00	0.216
EWG P3	-1.57	-5.55 – 2.41	0.438
EWG P4	0.09	-0.09 – 0.26	0.323
EWG P5	-0.49	-1.27 – 0.30	0.222
EWG P6	-0.01	-0.07 – 0.05	0.769
EWG P7	-0.01	-0.05 – 0.03	0.617
EWG P8	-0.02	-0.07 – 0.03	0.493
EWG P9	-0.02	-0.07 – 0.02	0.353
EWG P10	-0.03	-0.19 – 0.13	0.714
VIX Lag1	0.00	-0.05 – 0.05	0.981
VIX Lag2	-0.00	-0.04 – 0.03	0.898
VIX Lag3	-0.00	-0.03 – 0.02	0.719
VIX Lag4	0.00	-0.01 – 0.02	0.760
VIX Lag5	0.01	-0.00 – 0.02	0.264
VIX P3	-0.01	-0.07 – 0.04	0.628
VIX P4	0.00	-0.02 – 0.02	0.736
VIX P5	0.02	-0.07 – 0.10	0.677
VIX P6	-0.00	-0.01 – 0.00	0.733
VIX P7	-0.00	-0.00 – 0.00	0.397
VIX P8	0.01	-0.00 – 0.02	0.175
VIX P9	0.02	-0.00 – 0.04	0.079
VIX P10	0.00	-0.01 – 0.01	0.872
Observations	1400		
R^2 / R^2 adjusted	0.082 / 0.020		

1B: Backward Elimination

	y_Today		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	-0.00 – 0.00	0.813
Lag1	-0.17	-0.27 – -0.06	**0.002**
Lag2	-0.17	-0.26 – -0.07	**0.001**
Lag4	-0.07	-0.15 – 0.02	0.139
P4	-0.14	-0.32 – 0.03	0.11
P10	0.17	0.05 – 0.29	**0.005**
FXI Lag1	0.16	0.07 – 0.24	**<0.001**
FXI Lag2	0.11	0.04 – 0.17	**0.003**
FXI Lag4	0.04	0.00 – 0.09	**0.047**
FXI P3	-0.16	-0.26 – -0.07	**0.001**
FXI P8	-0.03	-0.05 – -0.01	**0.003**
EWU Lag1	0.18	0.05 – 0.30	**0.006**
EWU Lag3	0.07	-0.02 – 0.17	0.128
EWU Lag5	0.06	0.00 – 0.11	**0.038**
EWU P3	-0.12	-0.25 – 0.01	0.064
EWU P4	0.15	0.04 – 0.27	**0.01**
EWJ Lag3	0.04	-0.01 – 0.09	0.127
EWG Lag1	1.73	0.08 – 3.37	**0.04**
EWG Lag2	1.45	0.22 – 2.69	**0.021**
EWG Lag3	0.9	0.07 – 1.73	**0.033**
EWG Lag4	0.47	0.05 – 0.89	**0.03**
EWG P2	0	-0.00 – -0.00	**0.035**
EWG P3	-2.23	-4.28 – -0.17	**0.034**
VIX Lag5	0.01	0.00 – 0.02	**0.02**
VIX P3	-0.01	-0.02 – -0.01	**0.002**
VIX P7	0	-0.00 – -0.00	**0.021**
VIX P8	0.01	0.00 – 0.02	**0.014**
Observations			1400
R^2 / R^2 adjusted	0.060 / 0.043		

1C: Forward Selection

	y_Today		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.00	-0.00 – 0.00	0.256
P10	0.13	0.04 – 0.22	**0.006**
EWG P2	-0.00	-0.00 – -0.00	**0.024**
VIX Lag3	-0.01	-0.02 – -0.00	**0.006**
FXI Lag3	-0.06	-0.09 – -0.02	**0.003**
FXI P8	-0.03	-0.05 – -0.01	**0.002**
EWU Lag4	-0.05	-0.10 – 0.01	0.089
EWJ Lag3	0.06	0.00 – 0.11	**0.036**
FXI Lag1	0.03	-0.01 – 0.06	0.174
P3	-0.22	-0.33 – -0.11	**<0.001**
VIX P3	-0.01	-0.02 – -0.00	**0.004**
EWG Lag2	0.10	0.03 – 0.16	**0.003**
EWU Lag1	0.10	0.02 – 0.17	**0.013**
VIX P7	-0.00	-0.00 – -0.00	**0.007**
VIX P8	0.01	0.00 – 0.02	**0.017**
EWJ Lag4	0.04	-0.01 – 0.09	0.113
EWU Lag2	-0.06	-0.13 – 0.01	0.087
EWG Lag1	-0.06	-0.12 – 0.01	0.099
Observations	1400		
R^2 / R^2 adjusted	0.052 / 0.041		

1D: Stepwise

	y_Today		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.00	-0.00 – 0.00	0.222
P10	0.13	0.04 – 0.22	**0.005**
EWG P2	-0.00	-0.00 – -0.00	**0.024**
VIX Lag3	-0.01	-0.02 – -0.00	**0.004**
FXI Lag3	-0.06	-0.09 – -0.02	**0.002**
FXI P8	-0.03	-0.05 – -0.01	**0.001**
EWJ Lag3	0.05	-0.00 – 0.10	0.057
P3	-0.22	-0.32 – -0.11	**<0.001**
VIX P3	-0.01	-0.02 – -0.00	**0.004**
EWG Lag2	0.09	0.03 – 0.16	**0.005**
EWU Lag1	0.11	0.04 – 0.19	**0.002**
VIX P7	-0.00	-0.00 – -0.00	**0.007**
VIX P8	0.01	0.00 – 0.02	**0.017**
EWU Lag2	-0.06	-0.13 – 0.01	0.096
EWG Lag1	-0.05	-0.12 – 0.02	0.140
Observations	1400		
R^2 / R^2 adjusted	0.049 / 0.039		

Appendix 2: Methods for y_High w/ Interactions

2A: Min Square Root

	y_High		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.00	-0.00 – 0.00	0.479
Lag1	-2.41	-6.13 – 1.31	0.205
Lag2	-1.84	-4.64 – 0.96	0.198
Lag3	-1.14	-3.00 – 0.72	0.229
Lag4	-0.60	-1.53 – 0.33	0.205
Lag5	-0.06	-0.14 – 0.01	0.085
P1	0.00	-0.00 – 0.00	0.348
P2	-0.00	-0.01 – 0.00	0.264
P3	2.79	-1.87 – 7.46	0.240
P4	-0.00	-0.16 – 0.15	0.951
P5	-0.27	-0.90 – 0.35	0.395
P6	-0.02	-0.08 – 0.03	0.376
P7	-0.01	-0.04 – 0.02	0.586
P8	-0.01	-0.05 – 0.03	0.518
P9	0.01	-0.02 – 0.03	0.619
P10	0.09	-0.03 – 0.21	0.156
FXI Lag1	-0.56	-1.38 – 0.25	0.178
FXI Lag2	-0.43	-1.04 – 0.18	0.168
FXI Lag3	-0.31	-0.72 – 0.09	0.132
FXI Lag4	-0.15	-0.35 – 0.06	0.164
FXI Lag5	0.00	-0.02 – 0.03	0.799
FXI P1	0.00	-0.00 – 0.00	0.858
FXI P2	-0.00	-0.00 – 0.00	0.494
FXI P3	0.69	-0.33 – 1.71	0.182
FXI P4	0.03	-0.04 – 0.09	0.398
FXI P5	0.09	-0.17 – 0.34	0.489
FXI P6	0.01	-0.01 – 0.03	0.425
FXI P7	-0.01	-0.02 – 0.01	0.296
FXI P8	-0.02	-0.04 – -0.00	**0.045**
FXI P9	0.00	-0.01 – 0.02	0.906
FXI P10	0.00	-0.07 – 0.08	0.955
EWU Lag1	-0.28	-1.65 – 1.09	0.687
EWU Lag2	-0.29	-1.32 – 0.74	0.583
EWU Lag3	-0.17	-0.86 – 0.52	0.634
EWU Lag4	-0.11	-0.45 – 0.24	0.550
EWU Lag5	0.04	-0.01 – 0.09	0.110
EWU P1	0.00	-0.00 – 0.00	0.375
EWU P2	-0.00	-0.00 – 0.00	0.786
EWU P3	0.46	-1.26 – 2.19	0.597
EWU P4	0.01	-0.10 – 0.12	0.802
EWU P5	0.01	-0.27 – 0.29	0.937
EWU P6	-0.01	-0.05 – 0.03	0.643
EWU P7	0.01	-0.02 – 0.04	0.488
EWU P8	0.03	0.00 – 0.07	**0.049**
EWU P9	0.02	-0.01 – 0.05	0.275
EWU P10	0.04	-0.07 – 0.16	0.445
EWJ Lag1	1.24	-0.50 – 2.99	0.162
EWJ Lag2	0.93	-0.38 – 2.24	0.164
EWJ Lag3	0.64	-0.23 – 1.51	0.149
EWJ Lag4	0.32	-0.11 – 0.76	0.147
EWJ Lag5	-0.02	-0.05 – 0.02	0.371
EWJ P1	-0.00	-0.00 – 0.00	0.626
EWJ P2	0.00	-0.00 – 0.00	0.189
EWJ P3	-1.57	-3.76 – 0.61	0.159
EWJ P4	0.02	-0.07 – 0.11	0.690
EWJ P5	0.04	-0.17 – 0.24	0.739
EWJ P6	-0.00	-0.03 – 0.03	0.832
EWJ P7	0.01	-0.01 – 0.03	0.314
EWJ P8	0.02	-0.01 – 0.05	0.170

EWJ P9	0.02	-0.01 – 0.04	0.204
EWJ P10	0.04	-0.07 – 0.15	0.508
EWG Lag1	1.30	-0.57 – 3.18	0.172
EWG Lag2	1.08	-0.32 – 2.49	0.130
EWG Lag3	0.66	-0.28 – 1.59	0.169
EWG Lag4	0.34	-0.13 – 0.81	0.161
EWG Lag5	0.03	-0.02 – 0.07	0.205
EWG P1	-0.00	-0.00 – 0.00	0.704
EWG P2	-0.00	-0.00 – 0.00	0.057
EWG P3	-1.67	-4.01 – 0.67	0.162
EWG P4	0.04	-0.06 – 0.14	0.472
EWG P5	-0.42	-0.88 – 0.04	0.072
EWG P6	-0.01	-0.04 – 0.03	0.629
EWG P7	-0.02	-0.05 – -0.00	**0.040**
EWG P8	-0.00	-0.03 – 0.03	0.815
EWG P9	-0.00	-0.03 – 0.02	0.778
EWG P10	0.02	-0.07 – 0.12	0.667
VIX Lag1	0.00	-0.02 – 0.03	0.835
VIX Lag2	0.00	-0.02 – 0.02	0.793
VIX Lag3	0.00	-0.01 – 0.01	0.998
VIX Lag4	0.00	-0.01 – 0.01	0.663
VIX Lag5	0.00	-0.00 – 0.01	0.358
VIX P3	-0.01	-0.04 – 0.02	0.614
VIX P4	0.00	-0.01 – 0.01	0.824
VIX P5	0.02	-0.03 – 0.07	0.451
VIX P6	-0.00	-0.00 – 0.00	0.413
VIX P7	0.00	-0.00 – 0.00	0.545
VIX P8	0.01	0.00 – 0.02	**0.006**
VIX P9	0.01	-0.01 – 0.02	0.326
VIX P10	0.00	-0.00 – 0.01	0.618
Observations	1400		
R^2 / R^2 adjusted	0.271 / 0.222		

2B: Backward Elimination

		y_Low	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	-0.00 – 0.00	0.777
Lag2	-0.1	-0.18 – -0.02	**0.019**
P4	-0.31	-0.43 – -0.18	**<0.001**
P5	-0.59	-1.00 – -0.18	**0.004**
P7	-0.02	-0.05 – 0.00	0.078
FXI Lag1	0.09	0.03 – 0.15	**0.004**
FXI Lag2	0.08	0.03 – 0.13	**0.002**
FXI Lag4	0.02	-0.01 – 0.04	0.157
FXI Lag5	0.02	-0.00 – 0.05	0.069
FXI P3	-0.12	-0.18 – -0.05	**<0.001**
FXI P4	0.08	0.02 – 0.15	**0.013**
FXI P8	-0.03	-0.04 – -0.01	**0.005**
EWU Lag1	0.05	0.00 – 0.10	**0.037**
EWU P4	0.07	-0.01 – 0.15	0.102
EWU P8	0.02	-0.01 – 0.05	0.147
EWJ Lag1	-0.08	-0.16 – -0.00	**0.041**
EWJ Lag2	-0.06	-0.12 – 0.01	0.079
EWJ P3	0.07	-0.01 – 0.16	0.099
EWJ P9	0.02	-0.00 – 0.04	0.125
EWG Lag2	0.04	-0.00 – 0.09	0.067
EWG P1	0	-0.00 – -0.00	**0.038**
EWG P8	-0.02	-0.05 – 0.01	0.111
EWG P9	-0.01	-0.03 – 0.00	0.136
EWG P10	-0.07	-0.15 – 0.02	0.114
VIX Lag1	0.01	0.00 – 0.02	**0.026**
VIX Lag5	0	-0.00 – 0.01	0.067
VIX P3	-0.02	-0.03 – -0.01	**<0.001**
VIX P7	0	-0.00 – 0.00	0.058
VIX P9	0.02	0.00 – 0.03	**0.009**
VIX P10	0.01	-0.00 – 0.01	0.06
Observations			1400
R^2 / R^2 adjusted	0.129 / 0.111		

2C: Forward Selection

		y_Low	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	-0.00 – 0.01	0.686

Lag1	-3.47	-8.35 – 1.41	0.163
Lag2	-2.65	-6.32 – 1.02	0.157
Lag3	-1.76	-4.20 – 0.67	0.156
Lag4	-0.83	-2.05 – 0.39	0.18
Lag5	0.03	-0.07 – 0.13	0.533
P1	0	-0.00 – 0.00	0.192
P2	0	-0.01 – -0.00	**0.027**
P3	4.31	-1.80 – 10.43	0.167
P4	-0.41	-0.62 – -0.20	**<0.001**
P5	-0.74	-1.56 – 0.08	0.076
P6	0.02	-0.05 – 0.09	0.623
P7	-0.03	-0.07 – 0.02	0.208
P8	0	-0.05 – 0.06	0.947
P9	0.01	-0.02 – 0.04	0.546
P10	0.07	-0.09 – 0.23	0.382
FXI Lag1	0.41	-0.66 – 1.48	0.454
FXI Lag2	0.32	-0.48 – 1.12	0.43
FXI Lag3	0.16	-0.38 – 0.69	0.561
FXI Lag4	0.11	-0.16 – 0.37	0.441
FXI Lag5	0.03	-0.01 – 0.06	0.119
FXI P1	0	-0.00 – 0.00	0.719
FXI P2	0	-0.00 – 0.00	0.387
FXI P3	-0.51	-1.85 – 0.83	0.453
FXI P4	0.11	0.02 – 0.19	**0.014**
FXI P5	-0.06	-0.39 – 0.28	0.747
FXI P6	-0.01	-0.03 – 0.02	0.508
FXI P7	0	-0.01 – 0.02	0.863
FXI P8	-0.02	-0.04 – 0.01	0.183
FXI P9	0	-0.03 – 0.02	0.634
FXI P10	-0.04	-0.14 – 0.06	0.448
EWU Lag1	0.04	-1.75 – 1.84	0.961
EWU Lag2	-0.05	-1.40 – 1.29	0.937
EWU Lag3	0.02	-0.89 – 0.92	0.966
EWU Lag4	-0.02	-0.48 – 0.44	0.924
EWU Lag5	0.01	-0.05 – 0.08	0.657
EWU P1	0	-0.00 – 0.00	0.531
EWU P2	0	-0.00 – 0.00	0.209
EWU P3	0.02	-2.23 – 2.28	0.983
EWU P4	0.08	-0.07 – 0.22	0.307
EWU P5	0.24	-0.14 – 0.61	0.213
EWU P6	-0.04	-0.09 – 0.01	0.116
EWU P7	-0.01	-0.05 – 0.03	0.512
EWU P8	0.04	-0.01 – 0.08	0.103
EWU P9	-0.02	-0.07 – 0.02	0.329
EWU P10	0.01	-0.14 – 0.16	0.93
EWJ Lag1	-0.92	-3.21 – 1.37	0.431
EWJ Lag2	-0.69	-2.40 – 1.03	0.433
EWJ Lag3	-0.42	-1.57 – 0.72	0.468
EWJ Lag4	-0.21	-0.78 – 0.36	0.468
EWJ Lag5	-0.01	-0.05 – 0.04	0.809
EWJ P1	0	-0.00 – 0.00	0.901
EWJ P2	0	-0.00 – 0.00	0.601
EWJ P3	1.12	-1.74 – 3.99	0.442
EWJ P4	-0.05	-0.17 – 0.07	0.379
EWJ P5	0.23	-0.05 – 0.50	0.102
EWJ P6	0.01	-0.03 – 0.05	0.663
EWJ P7	-0.01	-0.04 – 0.02	0.511
EWJ P8	0	-0.04 – 0.04	0.934
EWJ P9	0.03	-0.01 – 0.06	0.147
EWJ P10	-0.05	-0.19 – 0.10	0.518
EWG Lag1	1	-1.45 – 3.46	0.423
EWG Lag2	0.83	-1.02 – 2.67	0.379
EWG Lag3	0.52	-0.71 – 1.75	0.409
EWG Lag4	0.28	-0.34 – 0.89	0.38
EWG Lag5	-0.01	-0.07 – 0.04	0.695
EWG P1	0	-0.00 – 0.00	0.286
EWG P2	0	-0.00 – 0.00	0.261
EWG P3	-1.28	-4.35 – 1.78	0.412
EWG P4	0.05	-0.09 – 0.18	0.486
EWG P5	-0.27	-0.88 – 0.33	0.378
EWG P6	0.01	-0.03 – 0.06	0.598
EWG P7	-0.01	-0.04 – 0.02	0.58
EWG P8	-0.04	-0.08 – 0.00	0.066
EWG P9	-0.01	-0.05 – 0.03	0.605
EWG P10	-0.06	-0.18 – 0.07	0.38
VIX Lag1	0.04	-0.00 – 0.07	0.051
VIX Lag2	0.02	-0.00 – 0.05	0.076
VIX Lag3	0.01	-0.01 – 0.03	0.223
VIX Lag4	0.01	-0.00 – 0.02	0.061
VIX Lag5	0.01	-0.00 – 0.01	0.102
VIX P3	-0.05	-0.10 – -0.01	**0.013**
VIX P4	0.01	-0.01 – 0.03	0.226
VIX P5	0.02	-0.04 – 0.08	0.54
VIX P6	0	-0.01 – 0.00	0.573
VIX P7	0	-0.00 – 0.00	0.225

VIX P8	0.01	-0.01 – 0.02	0.353
VIX P9	0.02	-0.00 – 0.03	0.076
VIX P10	0.01	-0.00 – 0.02	0.057
Observations			1400
R^2 / R^2 adjusted	0.151 / 0.094		

2D: Stepwise

	y_Low		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	-0.00 – 0.00	0.321
FXI P5	-0.17	-0.27 – -0.06	**0.002**
VIX P3	-0.01	-0.01 – -0.01	**<0.001**
P4	-0.31	-0.40 – -0.21	**<0.001**
FXI Lag3	-0.05	-0.08 – -0.03	**<0.001**
VIX Lag3	0	-0.01 – 0.00	0.094
EWG P2	0	-0.00 – -0.00	**0.015**
FXI P8	-0.03	-0.04 – -0.01	**0.002**
FXI P4	0.09	0.03 – 0.16	**0.004**
EWJ Lag3	0.03	-0.01 – 0.07	0.121
P7	-0.02	-0.04 – 0.00	0.123
Observations			1400
R^2 / R^2 adjusted	0.111 / 0.104		

Appendix 3: Methods for y_Low w/ Interactions

3A: Min Square Root

		y_Low	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	-0.00 – 0.00	0.777
Lag2	-0.1	-0.18 – -0.02	**0.019**
P4	-0.31	-0.43 – -0.18	**<0.001**
P5	-0.59	-1.00 – -0.18	**0.004**
P7	-0.02	-0.05 – 0.00	0.078
FXI Lag1	0.09	0.03 – 0.15	**0.004**
FXI Lag2	0.08	0.03 – 0.13	**0.002**
FXI Lag4	0.02	-0.01 – 0.04	0.157
FXI Lag5	0.02	-0.00 – 0.05	0.069
FXI P3	-0.12	-0.18 – -0.05	**<0.001**
FXI P4	0.08	0.02 – 0.15	**0.013**
FXI P8	-0.03	-0.04 – -0.01	**0.005**
EWU Lag1	0.05	0.00 – 0.10	**0.037**
EWU P4	0.07	-0.01 – 0.15	0.102
EWU P8	0.02	-0.01 – 0.05	0.147
EWJ Lag1	-0.08	-0.16 – -0.00	**0.041**
EWJ Lag2	-0.06	-0.12 – 0.01	0.079
EWJ P3	0.07	-0.01 – 0.16	0.099
EWJ P9	0.02	-0.00 – 0.04	0.125
EWG Lag2	0.04	-0.00 – 0.09	0.067
EWG P1	0	-0.00 – -0.00	**0.038**
EWG P8	-0.02	-0.05 – 0.01	0.111
EWG P9	-0.01	-0.03 – 0.00	0.136
EWG P10	-0.07	-0.15 – 0.02	0.114
VIX Lag1	0.01	0.00 – 0.02	**0.026**
VIX Lag5	0	-0.00 – 0.01	0.067
VIX P3	-0.02	-0.03 – -0.01	**<0.001**
VIX P7	0	-0.00 – 0.00	0.058
VIX P9	0.02	0.00 – 0.03	**0.009**
VIX P10	0.01	-0.00 – 0.01	0.06
Observations			1400
R^2 / R^2 adjusted	0.129 / 0.111		

3B: Backward Elimination

		y_Low	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	-0.00 – 0.00	0.777
Lag2	-0.1	-0.18 – -0.02	**0.019**
P4	-0.31	-0.43 – -0.18	**<0.001**
P5	-0.59	-1.00 – -0.18	**0.004**
P7	-0.02	-0.05 – 0.00	0.078
FXI Lag1	0.09	0.03 – 0.15	**0.004**
FXI Lag2	0.08	0.03 – 0.13	**0.002**
FXI Lag4	0.02	-0.01 – 0.04	0.157
FXI Lag5	0.02	-0.00 – 0.05	0.069
FXI P3	-0.12	-0.18 – -0.05	**<0.001**
FXI P4	0.08	0.02 – 0.15	**0.013**
FXI P8	-0.03	-0.04 – -0.01	**0.005**
EWU Lag1	0.05	0.00 – 0.10	**0.037**
EWU P4	0.07	-0.01 – 0.15	0.102
EWU P8	0.02	-0.01 – 0.05	0.147
EWJ Lag1	-0.08	-0.16 – -0.00	**0.041**
EWJ Lag2	-0.06	-0.12 – 0.01	0.079
EWJ P3	0.07	-0.01 – 0.16	0.099
EWJ P9	0.02	-0.00 – 0.04	0.125
EWG Lag2	0.04	-0.00 – 0.09	0.067
EWG P1	0	-0.00 – -0.00	**0.038**
EWG P8	-0.02	-0.05 – 0.01	0.111
EWG P9	-0.01	-0.03 – 0.00	0.136
EWG P10	-0.07	-0.15 – 0.02	0.114
VIX Lag1	0.01	0.00 – 0.02	**0.026**
VIX Lag5	0	-0.00 – 0.01	0.067
VIX P3	-0.02	-0.03 – -0.01	**<0.001**
VIX P7	0	-0.00 – 0.00	0.058
VIX P9	0.02	0.00 – 0.03	**0.009**
VIX P10	0.01	-0.00 – 0.01	0.06
Observations			1400
R^2 / R^2 adjusted	0.129 / 0.111		

3C: Forward Selection

		y_Low	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0	-0.00 – 0.01	0.686
Lag1	-3.47	-8.35 – 1.41	0.163
Lag2	-2.65	-6.32 – 1.02	0.157
Lag3	-1.76	-4.20 – 0.67	0.156

Lag4	-0.83	-2.05 – 0.39	0.18
Lag5	0.03	-0.07 – 0.13	0.533
P1	0	-0.00 – 0.00	0.192
P2	0	-0.01 – -0.00	**0.027**
P3	4.31	-1.80 – 10.43	0.167
P4	-0.41	-0.62 – -0.20	**<0.001**
P5	-0.74	-1.56 – 0.08	0.076
P6	0.02	-0.05 – 0.09	0.623
P7	-0.03	-0.07 – 0.02	0.208
P8	0	-0.05 – 0.06	0.947
P9	0.01	-0.02 – 0.04	0.546
P10	0.07	-0.09 – 0.23	0.382
FXI Lag1	0.41	-0.66 – 1.48	0.454
FXI Lag2	0.32	-0.48 – 1.12	0.43
FXI Lag3	0.16	-0.38 – 0.69	0.561
FXI Lag4	0.11	-0.16 – 0.37	0.441
FXI Lag5	0.03	-0.01 – 0.06	0.119
FXI P1	0	-0.00 – 0.00	0.719
FXI P2	0	-0.00 – 0.00	0.387
FXI P3	-0.51	-1.85 – 0.83	0.453
FXI P4	0.11	0.02 – 0.19	**0.014**
FXI P5	-0.06	-0.39 – 0.28	0.747
FXI P6	-0.01	-0.03 – 0.02	0.508
FXI P7	0	-0.01 – 0.02	0.863
FXI P8	-0.02	-0.04 – 0.01	0.183
FXI P9	0	-0.03 – 0.02	0.634
FXI P10	-0.04	-0.14 – 0.06	0.448
EWU Lag1	0.04	-1.75 – 1.84	0.961
EWU Lag2	-0.05	-1.40 – 1.29	0.937
EWU Lag3	0.02	-0.89 – 0.92	0.966
EWU Lag4	-0.02	-0.48 – 0.44	0.924
EWU Lag5	0.01	-0.05 – 0.08	0.657
EWU P1	0	-0.00 – 0.00	0.531
EWU P2	0	-0.00 – 0.00	0.209
EWU P3	0.02	-2.23 – 2.28	0.983
EWU P4	0.08	-0.07 – 0.22	0.307
EWU P5	0.24	-0.14 – 0.61	0.213
EWU P6	-0.04	-0.09 – 0.01	0.116
EWU P7	-0.01	-0.05 – 0.03	0.512
EWU P8	0.04	-0.01 – 0.08	0.103
EWU P9	-0.02	-0.07 – 0.02	0.329
EWU P10	0.01	-0.14 – 0.16	0.93
EWJ Lag1	-0.92	-3.21 – 1.37	0.431
EWJ Lag2	-0.69	-2.40 – 1.03	0.433
EWJ Lag3	-0.42	-1.57 – 0.72	0.468
EWJ Lag4	-0.21	-0.78 – 0.36	0.468
EWJ Lag5	-0.01	-0.05 – 0.04	0.809
EWJ P1	0	-0.00 – 0.00	0.901
EWJ P2	0	-0.00 – 0.00	0.601
EWJ P3	1.12	-1.74 – 3.99	0.442
EWJ P4	-0.05	-0.17 – 0.07	0.379
EWJ P5	0.23	-0.05 – 0.50	0.102
EWJ P6	0.01	-0.03 – 0.05	0.663
EWJ P7	-0.01	-0.04 – 0.02	0.511
EWJ P8	0	-0.04 – 0.04	0.934
EWJ P9	0.03	-0.01 – 0.06	0.147
EWJ P10	-0.05	-0.19 – 0.10	0.518
EWG Lag1	1	-1.45 – 3.46	0.423
EWG Lag2	0.83	-1.02 – 2.67	0.379
EWG Lag3	0.52	-0.71 – 1.75	0.409
EWG Lag4	0.28	-0.34 – 0.89	0.38
EWG Lag5	-0.01	-0.07 – 0.04	0.695
EWG P1	0	-0.00 – 0.00	0.286
EWG P2	0	-0.00 – 0.00	0.261
EWG P3	-1.28	-4.35 – 1.78	0.412
EWG P4	0.05	-0.09 – 0.18	0.486
EWG P5	-0.27	-0.88 – 0.33	0.378
EWG P6	0.01	-0.03 – 0.06	0.598
EWG P7	-0.01	-0.04 – 0.02	0.58
EWG P8	-0.04	-0.08 – 0.00	0.066
EWG P9	-0.01	-0.05 – 0.03	0.605
EWG P10	-0.06	-0.18 – 0.07	0.38
VIX Lag1	0.04	-0.00 – 0.07	0.051
VIX Lag2	0.02	-0.00 – 0.05	0.076
VIX Lag3	0.01	-0.01 – 0.03	0.223
VIX Lag4	0.01	-0.00 – 0.02	0.061
VIX Lag5	0.01	-0.00 – 0.01	0.102
VIX P3	-0.05	-0.10 – -0.01	**0.013**
VIX P4	0.01	-0.01 – 0.03	0.226
VIX P5	0.02	-0.04 – 0.08	0.54
VIX P6	0	-0.01 – 0.00	0.573
VIX P7	0	-0.00 – 0.00	0.225
VIX P8	0.01	-0.01 – 0.02	0.353
VIX P9	0.02	-0.00 – 0.03	0.076
VIX P10	0.01	-0.00 – 0.02	0.057
Observations			1400

R^2 / R^2 adjusted	0.151 / 0.094

3D: Stepwise

		y_Low	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.00	-0.00 – 0.00	0.321
FXI P5	-0.17	-0.27 – -0.06	**0.002**
VIX P3	-0.01	-0.01 – -0.01	**<0.001**
P4	-0.31	-0.40 – -0.21	**<0.001**
FXI Lag3	-0.05	-0.08 – -0.03	**<0.001**
VIX Lag3	-0.00	-0.01 – 0.00	0.094
EWG P2	-0.00	-0.00 – -0.00	**0.015**
FXI P8	-0.03	-0.04 – -0.01	**0.002**
FXI P4	0.09	0.03 – 0.16	**0.004**
EWJ Lag3	0.03	-0.01 – 0.07	0.121
P7	-0.02	-0.04 – 0.00	0.123
Observations	1400		
R^2 / R^2 adjusted	0.111 / 0.104		