Assignment #7 Predictive Modeling #3 by Joshua Troup

Q1: For the model you selected, what is the RMSE on the validation data and the test set?

**The validation data RMSE is 16862. The test data RMSE is 16709.**

Q2: What is the average error on the validation data and test set? What does this suggest?

**The average error for validation is 622. The average error for test is 172. Validation is used to tune the parameters of a classifier. Test is used to assess the performance of a trained classifier. The test is a test of the validation which is a good indicator it is a good model with the average error being relatively low. These small error values in both validation and test suggest the linear regression method has created an accurate predictor.**

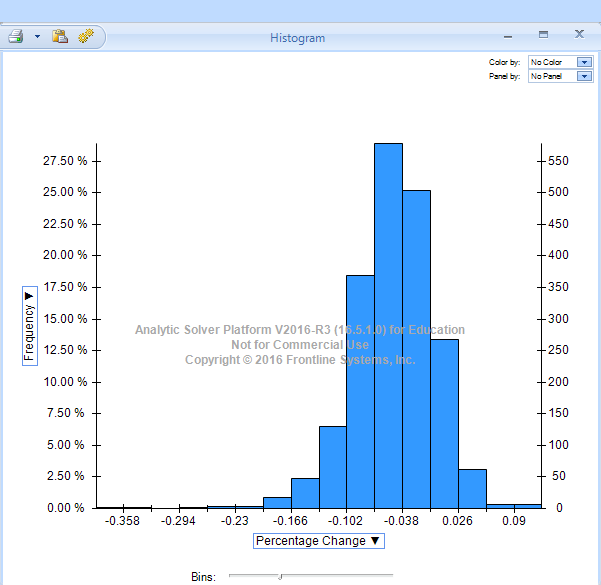
Q3: For the model you selected, what is the RMSE on the validation data and the test set?

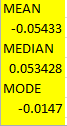
**The validation data RMSE is 18403. The test data RMSE is 17321.**

Q4: What is the average error on the validation data and test set? What does this suggest?

**The average error for validation is -692. The average error for test is -70. Validation is used to tune the parameters of a classifier. Test is used to assess the performance of a trained classifier. The test is a test of the validation which is a good indicator it is a good model with the average error being relatively low. These small negative error values in both validation and test suggest the linear regression method has created an accurate predictor.**

Q5: What is the average % change in predicted price between the pre and post-crisis model?

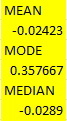
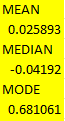
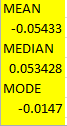




**The housing bubble affected the market by a 5.4% average difference. The histogram shows approximately 1/3 (30-33%) of the market was affected by the 5.4% difference.**

Q6: Finally, compare the average % change in predicted price using all the 3 supervised

algorithms (k-NN, Decision Tree and MLR (Multiple Linear Regression)

k-NN:  Decision Tree:  MLR: 

**k-NN and Decision tree average percentage change in predicted price are very similar. The MLR is slightly higher with the 5.4%. The medians are relatively close to each other in all three models. Mode is much more spread out between the models.**