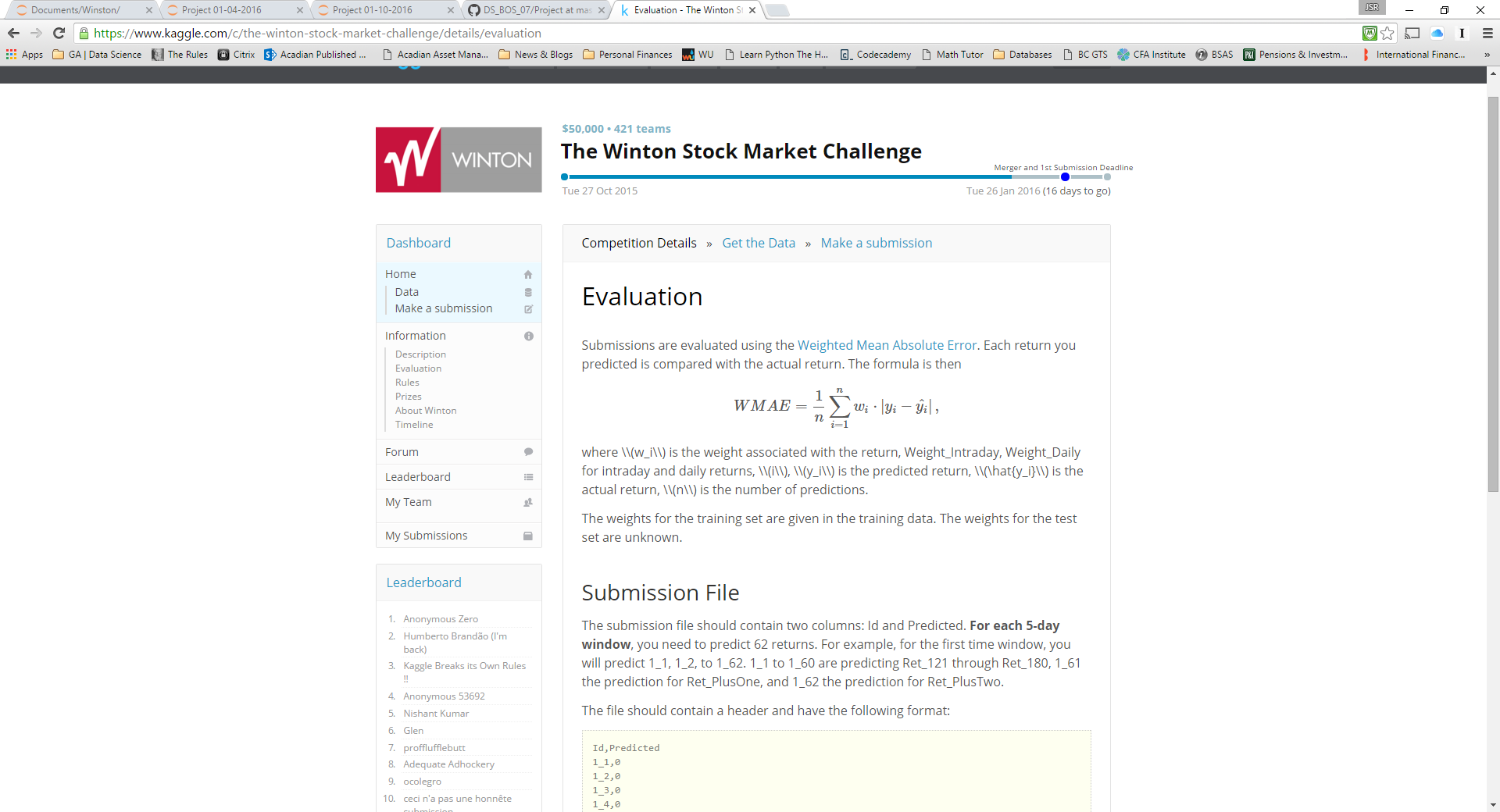
[Kaggle: Winston Stock Market Challenge](https://www.kaggle.com/c/the-winton-stock-market-challenge)

Goal: Minimize the weighted mean absolute error of predicted stock returns.



There weights for the training data set are known the weights for the test data set are unknown. There are essentially three times windows:

1. Intraday returns, t\_1 to t\_180. The training data set contains returns for time 1 to 180 which are one minute intervals, or three hours. The test data set provides returns for time 1 to 120 and the model should predict returns from time periods t = 121 to 180.
2. Day plus one, day\_1. This is the end of day return on the second day or day plus 1; t\_180 + 1.
3. Day plus two, day\_2. This is the end of day return on the third day or day plus 2; t\_180 + 2 or t\_180 + 1 + 1.

There are two sets of weights applied to the forecasted returns, an intraday weight applied to the forecasted returns for t\_121 to t\_180 and an end of day weight applied to t\_180 + 1 and t\_180 + 2.

The training data set also contains 25 features for each of the 40,000 hypothetical stocks. Some of the features are populated and some are not.

Hypothesis:

Intraday Return Null: Intraday returns are best be explained by the recent history/movement of the stock.

Test of Intraday Return Null: create various AR models.

End of Day Returns Null: Day\_1 and Day\_2 are best be explained by a sub set of the features in the data set.

Test of End of Day Returns Null: create a multivariable linear or non-linear regression model to predict the two returns