**Ryan Malani**

**IT-586-A**

**Correlation**

1. Prior to performing analysis on the file, my intuition based on the features in the dataset would be that the price action is dictated in part by volume.
2. Of the options in the data provided, there were no strong indicators for the ‘Tomorrow Percent Change’ feature. Of the data collected the highest correlation was the Low price of the day, with +0.005395, which is statistically insignificant, particularly when making financial decisions.
3. From my knowledge of stocks, there are a variety of features missing from the dataset that would likely make for stronger indicators. Some suggestions I would make for additions to the dataset would be ‘Yesterdays Percent Change’, for trading in more real time scenarios I believe it would be helpful to have Level II order book data (highest bid/ask), as well as qualitative sources turned quantitative (sentiment analysis of news, articles, public forums).
4. I added a feature for change\_on\_volume which serves as a proportion of the change in price (Close – Open) against the day’s Volume. This feature showed stronger correlation than the existing features, while still insignificant, more than double the previous best correlation. This feature performed stronger because it analyzed two major features and market drivers at once – price action and volume. Increasing prices on strong volume typically indicates a strong market for that security.