**Introduction**: I am a PhD student in finance and so am naturally interested in quantitative finance. One problem I was curious to better understand is how the market values volatility throughout time. Volatility is referring to the changes in prices for stocks on the market. Typically, these price movements are relatively small-think the stock tickers you see on news channels. But sometimes these movements can be very large and seemingly discontinuous. Options traders value this volatility inherent in the market to make educated trades. A popular proxy for volatility is a measure called VIX or the Volatility Index. It serves as a metric to measure the expectations of the market regarding future volatility. It is often called the *Fear Index* because, when it has high values, there is often panic within the market. Take for instance our recent circumstances. The Index peaked in mid-March at the peak of uncertainty relating to the pandemic. Additionally, it has stayed at an elevated value since then. The index is generated by looking at the forward volatility within the options traded on the stocks that make up the S&P500 index. This project seeks to better understand how options traders (who are the ones who dictate the value of the future volatility when they trade) come up with their volatility measures (which they use to trade). Mainly, I want to see if a few stocks having an uncommonly bad/good day have a disproportionate effect on market expectations of volatility.

**Why this project is useful**: Plainly put, if I were to create a model with significant predictive power, I could develop a trading strategy around it and make money. Not that I will be doing this as I doubt my model would be able to compete with the hyper competitive quantitative finance industry’s models (this would also be a violation of my agreement with the data provider). Indeed, there are many firms within Wallstreet that develop these types of models for various industries, stocks, derivatives, and pretty much any financial instrument possible. One of the most prosperous of which would be Renaissance Technologies, the founder of which was originally a mathematician who worked with the US government developing cryptography methods. As for how lucrative this is when done well, that mathematician is named Jim Simons. He has a net worth of $23.5 billion.

**What datasets I will be using**: I will be using WRDS or Wharton Research Database System which is the academic finance data standard from which most academic finance papers published derive their data. Not to be overly specific, I will be using their data on stock returns and index composition to generate my novel data.

**Methods of classification used**: As it is still early into the year and we have not been fully exposed to all the tools this class has to offer, I would currently say I will be using simple OLS regressions, possibly a ridge regression, and hopefully a neural network when we learn about that. I will also be using holdout data and cross validation to hopefully develop a more robust model.

**Reason these methods will be implemented**: Regressions are a fundamental tool in developing a straightforward model, and ridge regressions will allow me to keep in check a propensity to throw too much data at the problem. As for why I hope to try my hand at developing a neural network, that’s mainly a selfish aim because I have wanted to create one for several years now. Beyond that, neural networks are commonplace in the quantitative finance industry and becoming more prevalent in academic finance literature, so it would be good for me to develop my ability.

**Timeline**: So far as I can tell there are only a few broad things that I have to accomplish, that said there are several tasks required within the broad item for those items to be completed.

1. Gather data (November 1st)
2. Clean and develop novel data (November 12th)
3. Specify model (November 19th)
4. Implement model (November 25th)
5. Refine model (November 30th)
6. Test with new data (December 5th)
7. Summarize work (December 12th)

**GitHub link**: <https://github.com/trentmckinnon/Matrix-Methods-of-Machine-Learning/>