IST652 Final Project Proposal

Lakshya Gupta, Qingqing Hu

Group member:

Lakshya Gupta: lkgupta@syr.edu
Qingqing Hu: qhu100@syr.edu

Topic of Investigation:

"*In the long run, valuations may drive stock prices, but in the short term it is market sentiment that moves prices.***"**
In this project, we will try to study the effect of tweets sentiments on the stock prices of companies. We are analysing five well known companies (Amazon, Microsoft, Apple, Advanced Micro Devices(AMD), Tesla) for this study.

Dataset we will focus on:

Tweets were collected between April 9 and July 16, 2020 using not only the SPX500 tag but also the top 25 companies in the index and "#stocks". Due to lack of computational resources we will work on 5 companies by filtering them from data.
Besides, we want to search for the historical stock price of these five companies on Yahoo Finance at the same period. Therefore, we will get the new combined data sets by processing the different resources.

Primary dataset: https://www.kaggle.com/utkarshxy/stock-markettweets-lexicon-data/

Secondary dataset: Yahoo Finance (https://finance.yahoo.com/)

`$AMZN $TSLA $MSFT $AAPL $AMD SNP500`

Methods of data analysis:

- First we will find out and filter tweets of above top 5 companies using regex patterns. Then we will perform some descriptive analysis on these filtered tweets. For example, how many tweets correspond to each company? How do the counts vary with date?
- Then we will apply sentiment analysis on these filtered tweets using some pretrained models available in python. We will also analyse and try to find out how sentiment is changing with the date.
- Parallely we will perform correlation analysis among daily prices of these 5 companies.

- And finally we planned to showcase whether tweets sentiments and stock price of the company are following the same trend or not? This task can be done using correlation analysis.

Potential development tasks:

- We want to explore trends and cycles present in historical stock prices for few of the top companies in the market.
- For different units of descriptive analysis, we can make some comparisons for sentiments and companies.
- Finally, using visualizations and statistics, we will figure out the patterns and relationship between the stock price and the tweets sentiments.