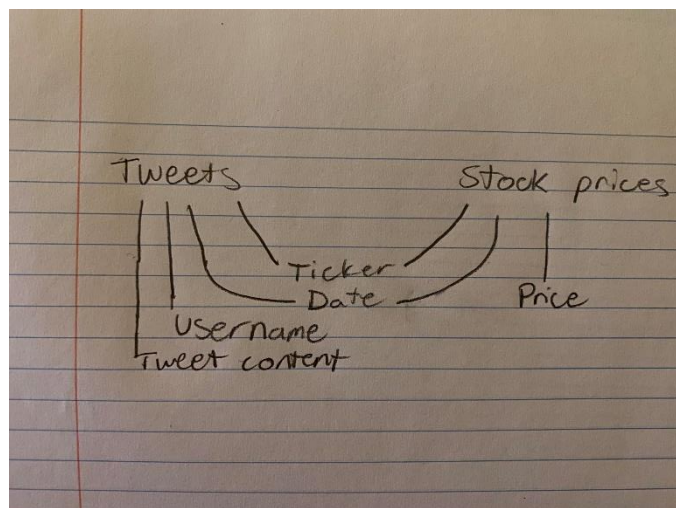Thomas McShane

- Dataset selection

  o I would like to work on a dataset that has to do with the stock market and its relation to twitter. I want to see how direct of a relationship there is between a company's stock price and tweets about that company to see if there is some sort of correlation. I have selected two separate datasets that I may combine into one dataset. The first describes several tweets about companies and the second describes stock prices over the same time period.

  o https://www.kaggle.com/omermetinn/tweets-about-the-top-companies-from-2015-to-2020?select=Company.csv

  o https://www.kaggle.com/yashj302/stock-market-data-by-symbol

- Problem formulation

  o In my project I would like to see if there is some correlation between a tweet that has been put out about a company and the company's stock price.

- Dataset description

  o I have two separate datasets that I would like to compare for this project and I found both of them on Kaggle. The first dataset is full of tweets about the top companies in the stock market. It is super cool because it

has hundreds to thousands of separate tweets about big companies in the stock market. It mainly focuses on Apple, Facebook, Amazon, Tesla, and Microsoft. But this dataset ranges from 2015 to 2020 and has every piece of metadata about the tweet. This ranges from the time to the user who posted them. There is a lot of possibility with this dataset because there is a lot of correlation with my other dataset. My other dataset has the stock prices of hundreds of different stock tickers and their stock prices for an extended period of time. This is really nice because I can track the common dates from the tweets and the stock prices of the day. It might be beneficial to combine the two datasets into one dataset. I could also potentially look at if certain users tweet about a company if that has more effect on the stock price than the tweet itself. Like if two similar tweets were made if one person had more impact than another.

- Project organization

- Project repository

  - I have added everyone on the instructional team for this class as collaborators to this repository for my project:

  - https://github.com/trm4/ENG-498-Final-Project