Carlos Scrum Report

November 1

We mostly just looked for a way to get our data. Yash was able to get access to the Twitter API, so we are probably going to take that route. Today, I mainly just familiarized myself with the API and looked for alternative sources of Tweets. We also thought about expanding to other companies besides Apple. Throughout the week, we want to see if we can find ways to better extract data from Twitter.

November 8

Yash and I mainly tried to analyze the week of data we were able to gather with Twitter's API. I looked through a variety of sentiment analysis tools, but eventually settled with python's NLTK. We were mostly working on Jupyter notebook trying to put something together, and by the end, we did get a few numbers.

November 15

This week we had a pretty good find since we came across a Kaggle dataset containing a list of Tweets containing the names of several companies like Apple, Tesla, and Microsoft. Hence, Yash and I spent this week wrangling the data to only consider Apple tweets. Eventually we did get some numbers but they weren't the best. The best result we could find was about a 11% correlation between stock price and the mean sentiment of the Tweets for that company on that day.

November 22

I read up on a few different research papers that had goals similar to ours where they were trying to predict stock movements from Twitter data. One particular paper that stuck out to me was one from some students from UT Austin who had a project pretty similar to ours and appeared to get better results through using Spacey's sentiment analysis

November 29

Recently, we have mostly been working on deliverables and Powerpoint slides. We are still making progress, but I'll admit it does feel like they get in the way a bit since I feel like that's mainly what I've been doing. This week, Patrick and Connor proposed using Google Cloud to analyze Tweets, so we are currently working on reanalyzing data. It seems like Connor and Patrick are already getting promising results though.

December 6

This week, we found an astounding 60% correlation between stock volume and mean sentiment analysis. It appears that switching to Google Cloud really did fix a lot of issues. We are currently working on the final presentation by making slides and scripts for it, but it seems like it will go well.