Yash Scrum Reports

November 1st
I worked on trying to help scrape tweets. I applied for and set up my twitter developer account. We are currently trying to figure out how to use the API to scrape tweets. That is what we plan to work on for the next week

November 8th
I worked on analyzing a week's worth of tweets that we have managed to scrape. We are experimenting using the NLTK intensity analyzer. The next step is to try and see if we can find any meaningful insights with this.

November 15th
We found a kaggle dataset that has tweets from big tech companies from 2015-2020. We ran NLTK intensity analyzer on it to see if we could find any correlation. So far we found an 11% correlation between if the stock price went up/down and the mean sentiment score of tweets for that day. Current problem is trying to find a higher correlation score.

November 22th
This week I looked into working with other metrics. Currently we have been doing a binary increase/decrease as 1 or 0 if the stock increased or decreased. This week I looked into high-low but that didn't improve the correlation. We did however find a high correlation between the number of tweets and stock volume, showing that people tweet more when the stock is being traded a lot. Current issue is trying to find a higher correlation between the sentiment and stock price.

November 29th
Patrick and Connor scraped a dataset that is a month big. We are working on running analysis on that month. The data they collected is formatted a little differently that our kaggle dataset so we have to refactor our code that performs the analysis to fit their dataset. So far we did not find a significant correlation between the sentiment and stock price on their dataset yet.

December 6th
This week, I worked on the project report with the rest of the team. After switching to google sentiment analysis, we were able to find over 60% correlation. I wrote about our challenges and all the approaches Carlos and I had originally tried that had failed to find a significant correlation. Looking forward to presenting our results with the team soon!