**Comparing Reddit Chatter on r/wallstreetbets to Trading Movement**

Gio Abou Jaoude | Alex ONeill | Noah Ponticiello

*CS666 - Enterprise Intelligence Development (Sp 2021) | Prof Barabasi*

Project Proposal

QUESTION: How much does the stock price of GameStop (GME) fluctuate in comparison to the number of posts on the subreddit Wall Street Bets(WSB)?

PLAN:

The pieces of data we need in order to create our project are current stock price, current trading volume, as well as the amount of posts that contain content regarding GME. In order to capture the stock market data we will use an API provided by IEX Cloud, which will allow us to extract stock information. In order to access this information we must turn it into a .JSON object, which will make it attribute value pairs. We are now able to access the data by data[‘desired attribute’]. We will then run this function every 60 seconds to get updated information.

base\_url = f'[https://cloud.iexapis.com/stable/stock/{ticker}/quote?token={token}](https://cloud.iexapis.com/stable/stock/%7Bticker%7D/quote?token=%7Btoken%7D)'

for i in range(15):

data = requests.get(base\_url).json()

print('{}: {}\tlatestVolume: {}\tvolume: {}\tiexVolume: {}'.format(data['symbol'], data['iexRealtimePrice'],

data['latestVolume'], data['volume'], data['iexVolume']))

time.sleep(60)

To achieve the amount of posted content regarding GME on WSB, we will use a web scraper called BeautifulSoup or a python wrapper called PRAW. BeautifulSoup will allow us to break down the subreddit page into HTML tags and we will search each HTML tag to see if it contains either GME, GameStop, or any other variations. PRAW, on the other hand, is a python library which aims to make interacting with Reddit much easier. PRAW allows you to create Reddit “objects'' which contain attributes such as rating, up/down votes, comment number, post-id, etc. Whether PRAW or BeautifulSoup is used, the corresponding function will be executed every 60 seconds as well in order to align the timing of both the number of posts within a minute and the stock data.

reddit = praw.Reddit(client\_id='', client\_secret='', user\_agent='', username='',

password='')

subreddit = reddit.subreddit('wallstreetbets')

hot\_posts = subreddit.hot(limit=10)

for post in hot\_posts:

if not post.stickied:

print('Title: {}\nUps: {}\n'.format(post.title, post.ups))

We connected a PostgreSQL database to our program. This will allow us to house our data as well timestamp the data every time it goes into the database. We will create one table for the information that will be composed of the timestamp as the primary key, stock price, trading volume, and amount of posts during that duration of time.

To visually show this relationship we will query all of the information in the database for the desired time then, connect it to Tableau to create the visualization. We will specifically use a line graph to display the data as Time will be on the x-axis and we will have numerical values on the y-axis to represent the trading volume, amount of posts, and stock price.