Answer the following three SCRUM prompts:

What have you done so far on your project? Please be as detailed as possible. List any code you wrote or tested, papers/documentation you read, data you attempted something with, etc. Note any changes to project direction that resulted from your work.

What are you doing next? What is your plan to complete the required project components on time?

Do you have any road blocks? Please describe things that are preventing you from making progress such as computing resources or difficulties using software/algorithms. Note that time itself is not a road block.

So far, I have gotten down to trying to understand the web API as well as scrape some data on the top 500 large-cap US Companies. The short-term goal I am going for now is to do a quick run through with 1 stock with high interest and 1 stock with low interest to familiarize myself with the algorithm a am planning to do as well as fine tune it. Once that is done I intend to try to scale it upwards to more companies if possible.

One issue I am facing now is identifying the levels of interest in the us companies and ranking them.

As I need to summarize and average 500 gdelt queries, I think it is too labour and time intensive to be carried out.

For now I am simply arbitrarily selecting the company number 1- Apple.inc, number 250 – Mettler-Toledo-International Inc, and number 500 – Fox Corporation Class B as a sample dataset for now.

I selected number 499 – Perrigo Co.Plc as I I think Class B shares would affect the interest in the share and skew the data

The time period I am selecting are 2 years so I would standardize it to May1 2019 – May1 2021

The source for the stock price data is going to be yahoo finance as I can download

Some road blocks I am facing is trying to use the gdelt api properly. I am still pretty unclear of exactly how to make it do what I want.

Some changes made to project direction so far has been to guess interest in stock I was not sure on calculating interest in each of the 500 stocks fairly due to the existance of class A,B,C stocks which I think skewed the data set. Another assumption I had to make was that stock splits and dividends would not affect the the stock prediction. I think this is pretty inaccurate unless I am using a time period of over 20 years but due to the unfeasibility of correcting each stock split/dividend/withheld dividend for each stock over the time period as I would have to manually calculate and change the data set, I decided to just leave it as it is. Thus the assumption of stock price not being affected by dividends etc.