1. What are the names and NetIDs of all your team members? Who is the captain? The captain will have more administrative duties than team members.

   The project team will be a team of just myself, and I will be the captain.
   NetId: feiyang3, Name: Takumi Li

2. What is your free topic? Please give a detailed description. What is the task? Why is it important or interesting? What is your planned approach? What tools, systems, or datasets are involved? What is the expected outcome? How are you going to evaluate your work?

   The topic of the project is Congress Financial Disclosures Sentiment Analysis. I was inspired by the news that "Four senators sold stocks before coronavirus threat crashed market" (https://thehill.com/homenews/senate/488593-four-senators-sold-stocks-before-coronavirus-threat-crashed-market), I would like to develop a tool to extract useful information from trades made by senators and representatives, to improve the data transparency into Congress and make insider trading less likely. It's crucially important because it helps the public to understand if and how politicians benefit from the information they know before the public, which is the purpose of financial disclosures. I plan to take advantage of the senatestockwatcher(https://senatestockwatcher.com/api) and housestockwatcher(https://housestockwatcher.com/api) APIs designed by Tim Carambat (https://ko-fi.com/rambat), get the disclosed trading data from the APIs and pre-process the data into tabular forms, make data visualization and Sentiment Analysis about their buy or sale actions in different industries or the whole market, get the S&P 500 and Nasdaq index data from Yahoo Finance API (https://www.yahoofinanceapi.com/), and correlates them by various similarity matrices. Due to the limits of the disclosed information (no specific trading price, volumes, and trading time, but only trading volume ranges and dates), I do not expect to develop trading algorithms based on disclosed information. However, I would expect to see the general bullish or bearish trends of the disclosed trading. To evaluate the work, I would like to consider the process of sentiment analysis about the trading actions and concluding the bullish or bearish trends of some specific periods (day/week/month) as the most important milestones. I do not consider finding out the correlation between the disclosed information and the S&P 500 and Nasdaq as an evaluation measure because there may or may not be an actual correlation. Finding the specific relationship would probably require extensive analysis and open research work.

3. Which programming language do you plan to use?

Python will be the primary programming language to be used, and **JavaScript** may also be used.

4. Please justify that the workload of your topic is at least 20*N hours, N being the total number of students in your team. You may list the main tasks to be completed and the estimated time cost for each task.

Takumi Li, feiyang3, will complete all the following tasks.

a. senatestockwatcher(https://senatestockwatcher.com/api) and housestockwatcher(https://housestockwatcher.com/api) APIs designed by Tim Carambat (https://ko-fi.com/rambat), get the disclosed trading data from the APIs, clean and pre-process the data into tabular forms

   3 hours

b. make data visualization and Sentiment Analysis about their buy or sale actions in different industries or the whole market

   4 hours

c. get the S&P 500 and Nasdaq index data from Yahoo Finance API (https://www.yahoofinanceapi.com/)

   3 hours

d. correlates them by various similarity matrices, analyze the disclosed information and the market index data

   4 hours

e. conclude the bullish or bearish trends about the disclosed trades in different periods

   4 hours