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.
   This is a solo team. My NetID is huiyinz2.

2. What is your free topic?
   My free topic is to build an intelligent market risk analyzer.

   Please give a detailed description. What is the task?
   The task is to perform topic modeling and sentiment analysis on the "Risk Factors" portion of the filed financial report of companies within the same industry and provide assessments and conclusions regarding the potential risk of investing in certain industries.

   Why is it important or interesting?
   Nowadays, aside from basic salaries and other fixed income, people also seek opportunities to make investments in areas that would generate desirable returns as a secondary source of income. However, it is rather difficult for laymen to fully understand the potential risks of investing in certain industries, the current market trend, or more specifically, what are the risks tied to certain companies. It is required that each publicly traded company should file an annual report, and there is always a section where the systematic and idiosyncratic risks are discussed. However, there are few attempts to computationally analyze and extract the main points from that section of the report, which can be very valuable in evaluating the attractiveness of this industry and help people make better and more informed investment decisions. It is also rare to see sentiment analysis being performed in such a setting, which can also tell the investors more about how the companies themselves feel about the level of risk. Therefore, I find it both important and interesting to devote time to this project.

   What is your planned approach?
   My approach breaks down into three main portions. Firstly, I am going to use a web-scraping approach to extract the risk factors section from financial reports of companies. Secondly, I will apply topic models and perform sentiment analysis on the extracted information. Thirdly, I will present the conclusion and recommendations in a user-friendly way (i.e. an interactive UI, if time allows).

   What tools, systems or datasets are involved?
   I would use web-scraping tools (including the basic libraries), the edgar downloader api for downloading financial reports, topic modeling and sentiment analysis packages, and possibly some other packages for visualization.

   What is the expected outcome?
   It is expected that my system will be able to 1. Extract the expected data, 2. Analyze the expected data as expected, 3. Display the results

   How are you going to evaluate your work?
   I aim to deliver a minimally viable product that meets my basic requirements and establishes the basic foundation that can be incrementally improved in the future.

3. Which programming language do you plan to use?
   Python and maybe R

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.
   Web-scraping and data cleaning task: 8 hours
   Topic modeling and sentiment analysis: 8 hours
   UI construction: 6-8 hours
   Sample use case demo and presentation: 4 hours