1. This task involves two parts: Obtaining a list of Github users from a webpage and downloading the information of those Github users using the Github API.

   (a) The webpage http://45.79.253.243/index.html contains a list of Github login names. Scrape this webpage and save the list of login names in a file (You may use a csv file, even though you have only one column of data)

   (b) Make sure that you check the validity of the login names downloaded. The information you saved should match what you see when you browse the webpage using a browser.

   (c) To obtain the list, you only need to download this webpage once. However, you can imagine that in a real world task you may have to download from the webpage 10000 times. Therefore, you should write a program to download the list from this webpage instead of just manually copy and paste the content.

   (d) The list may contain identical login names. You need to remove the duplicates.

   (e) Using the login names you obtained, download the information of these Github users using the Github API. For each of the Github users, try to obtain their id,avatar_url,url,number of followers,number of following,numnber of starred,number of repos,name,company,blog,location,email,hirea time, last update time, and any other information you found interesting.

   (f) Bonus: You can look into the repository of these Github users and obtain more information. You can get an A even if you do not attempt this part (given that you complete all other parts excellently). You will get bonus points if you can get even a small amount of new information from the users' repositories.

   (g) Save the dataset you downloaded into a csv file.

   (h) You need to use a Github personal authorization token. Do NOT put the token into your Github repo. You should make use of the .gitignore file.

   (i) In a real world task, most likely, you are not supposed to commit the files you downloaded into your git repository. However, for the purpose of the coursework, you should include that in your Git repository so that I can see your work.

   (j) You must hand in your homework via Github. Create a repository named "ECON860_midterm". In your repository, you should have the code, the data downloaded, and a .gitignore file. You should also include a file named README, which includes step-by-step instructions on how to run your Python code to collect the data you collected. This is especially important if you have multiple Python files (And you should!).

   (k) You can commit and push to Github as many times as you like. Only your last commit before the deadline is graded. I can read your previous commits, but they will not be graded.