README

TEAM: CARBON AGGREGATOR

Members:
Ian Hash (ihash):
Hang Ruan: (hruan)
Yuxin Zheng (yuxinz3)
Ida Mattsson (imattsso)

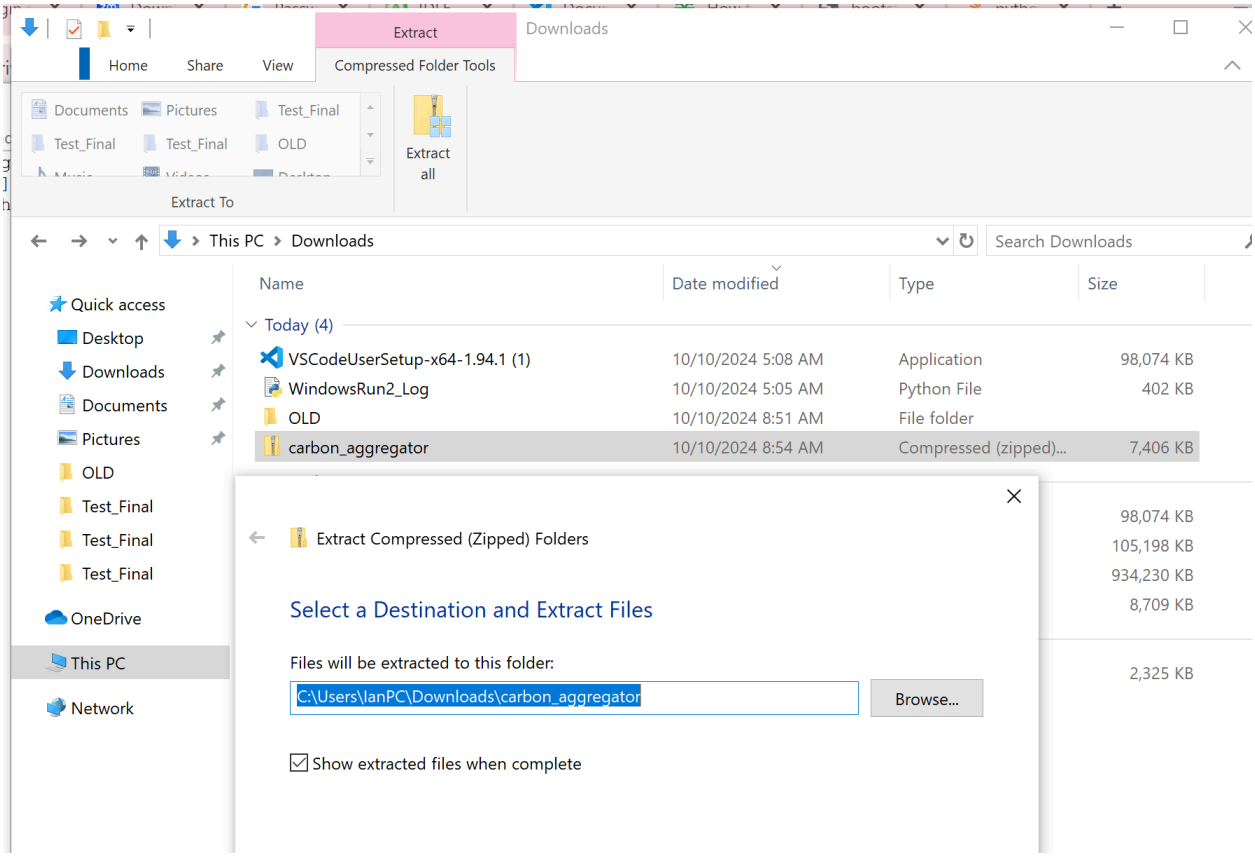
Welcome to the Team Carbon Aggregator's Final Project!

To run the Carbon Aggregator App Please follow these steps:

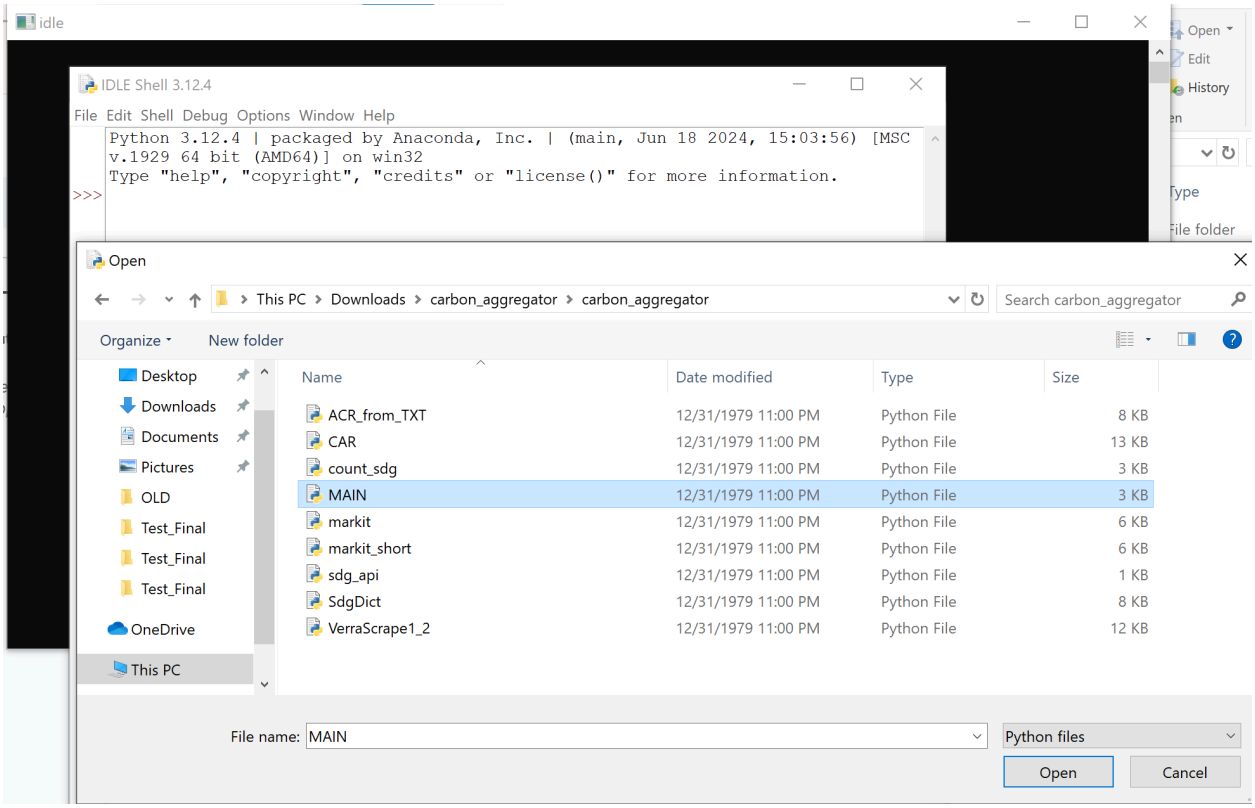
1. Make sure you have a stable internet connection.
2. If you don’t already have anaconda and python installed on your machine, download and install Anaconda, which will also do a fresh install of python. Allow Anaconda to install all the default packages.
3. We assume that you have a Chrome web browser installed. If not, please download the latest version to your local machine.
4. In addition to the default packages, install the following packages to your ‘base’ environment, using in the Anaconda>Environments>Packages, or a pip install in the command line.

REQUIRED IMPORTS:
Selenium: run 'pip install selenium' from your Command Line
Chrome web driver manager run: ‘pip install webdriver-manager’ from your command line
Pandas: run ‘pip install pandas’
Requests: run ‘pip install requests’

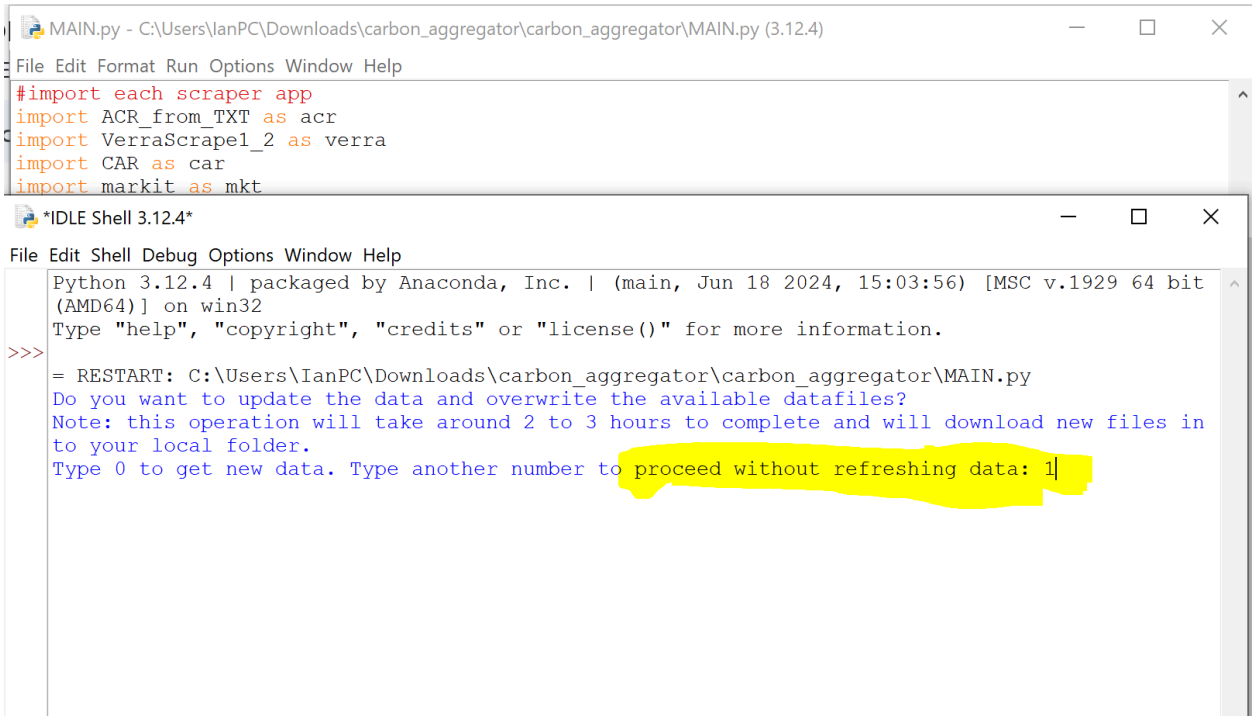
5. Download and unzip carbon_aggregator.zip to your downloads folder.
6. Unzip the folder



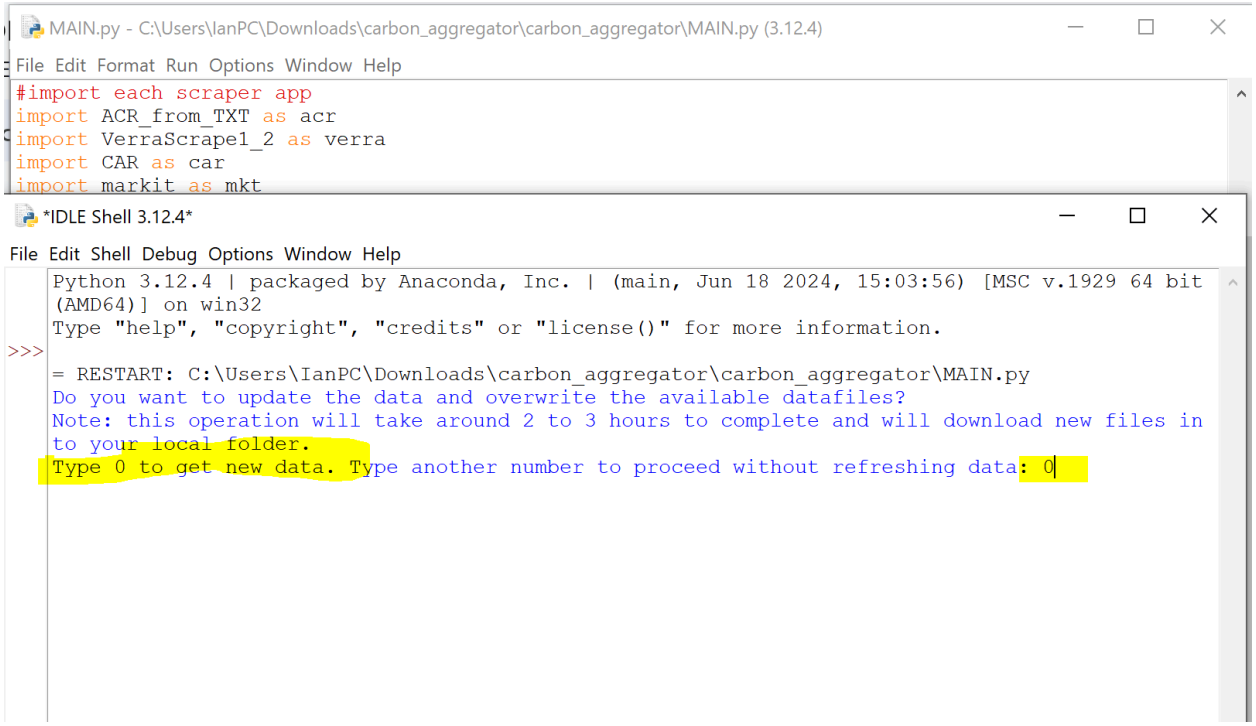
7. In idle, terminal, or your choice of an IDE, open or navigate to the folder and execute the MAIN.py python file (We have run successfully from VSCode, pyCharm, Spyder on a Mac, and in IDLE from a Windows machine).



8. You will be asked whether you want to clean local data or scrape fresh data before cleaning.
    a. OPTION 1: We recommend choosing this option on the first run. Enter any number key that is not 0 into the prompt. This will aggregate and format the data that has already been downloaded and is in the same folder as MAIN.py.



b. OPTION 2: If you want to download fresh data, before running MAIN.py you will need to move the sample data out of the carbon_aggregator folder so that the freshly scraped data can be placed there by the script. (This is not required on some machines, so you could try choosing '0' without removing the files).
   Move the following files out before choosing option '0' in the prompt:
   i. ACR_Projects.txt
   ii. ACR_RetiredCredits.txt
   iii. allprojects.csv
   iv. vcus.csv
   v. CAR_issued_projects_data.csv
   vi. CAR_retired_projects_data.csv
   vii. Mrk_projects.csv
   viii. Country_codes.csv
c. Once moved, you can select option '0' from the prompt to scrape fresh data. The rest of the application will run as normal, and output all_Markets.csv as described in step 7.



9. When either type of run is finished, you should see the output all_Markets.csv in the carbon_aggregator folder. This contains the aggregated data file containing data for all carbon credit markets, that has been scraped and formatted, and is the input to the data visualization tool (see the visualization section below for more details).

10. For ease of viewing, check out our data visualization tool at https://carbonoffsetregistries.streamlit.app/
11. Please reach out to us if you have any issues running the application, we'll be happy to assist you. Thank you!

VISUALIZATION:

You can interact with the visualization for our pre-compiled data on streamlit.io at this url:

https://carbonoffsetregistries.streamlit.app/

INCLUDES:

1. Table of data from all four registries
2. Bar graph showing distribution of SDG for each registry
3. Table of SDGs and their descriptions
4. Filters and search bar interactions

HOW TO USE:

1. TO START, SELECT THE DESIRED LOCATION(S) ON THE SIDE TO SEE THE ASSOCIATED PROJECTS
2. Filter the project table and SDG distribution chart by registries
3. Filter the project table by location
4. Search numbers 1 to 17 in search bar to view projects for each SDG

If you would like to update the visualization with new data locally, follow the steps below:

1. Download the repository from GitHub

   https://github.com/yxinzh/carbonoffset-dataset

2. Run count_sdg.py from carbon_aggregator.zip (Note: make sure all_Markets.csv from MAIN.py is in the local folder). This will create a new csv named sdg_counts.csv in your local folder.
3. Go into the data folder in the repository folder you just downloaded and replace the old all_Markets.csv file and the old sdg_counts.csv file with their respective new ones.

4. Go into your command interface and navigate to the repository folder and run:

```
streamlit run streamlit_app.py
```

Example:

```
C:\Users\yzhen\Desktop\FL2024\pycharm\carbonoffset-dataset>streamlit run streamlit_app.py
```

5. The app should run in your default browser