First Status Report

Date: September 26, 2021

Accomplishments:

- The project GitHub repository was set up at https://github.com/mjlaw/FIT_capstone.
- The project proposal was uploaded to the GitHub repository.
- Air Quality (AQ) datasets were downloaded from the EPA AQS website to include daily summary data for PM2.5 for years 1999, 2012, and 2020; daily summary data for ozone for years 1999, 2012, and 2020; and annual summary for AQI by county for years 1999, 2012, and 2020.
- The datasets were uploaded to a dedicated folder for original datasets on the project's GitHub repository.
- Extracted information on content and format of the Annual Summary Files and the Daily Summary Files from https://aqs.epa.gov/aqsweb/airdata/FileFormats.html, saved the information to PDF documents, and uploaded the documents to the original datasets folder on the GitHub project repository.
- Read in 1999 and 2012 PM2.5 daily summary csv files, extracted arithmetic mean columns, and printed brief summaries in a Jupyter notebook file.
- Compared the summaries to Dr. Peng's summaries of 1999 and 2012 PM2.5 raw text files.
- Uploaded the Jupyter notebook file to the GitHub project repository.
- Uploaded the first status report to the GitHub project repository.

Current Activities: I am currently working on the PM2.5 AQ datasets as well as a synopsis for my GitHub project repository readme file. In addition, I am looking at a better way to incorporate textual content to my Jupyter Notebook files.

Challenges: Interpreting air quality datafiles is the most time-consuming area. Trying to reproduce research results using files with different parameters and numbers of observations is problematic.

Work to be Completed: For the next project milestone, I will complete the PM2.5 studies and visualizations, the ozone studies and visualizations, and the AQI studies and visualizations.