Firstly, I can say right off the bat, that the process of completing each Milestone every two weeks was very enriching and fulfilling. I am really happy that I had the opportunity to do this term project. As I completed each milestone, that is from Milestone 1-5, each week I got to learn something new and unique in Python. I came to know how Python is so versatile and how much one can use Python for conducting various kind of data analyzing tasks.

Regarding on what I learned; I would like to take this opportunity to break it down by milestones.

**Milestone 1** – This was the very beginning of the term project and we had to choose a topic that had to have three resources – a flat file, website, and an API. This time researching for a topic was very different compared to a typical data science project where one has to choose a topic/dataset, have csv file or xlsx connected to it and start the cleaning process. This time I had to think outside the box for a topic. After a lot of research, I decided to choose a topic that I thought had relevance in today’s time and had access to all 3 resources, that was dealing with COVID 19 data.

**Milestone 2**  - In this course, the emphasise on cleaning a dataset thoroughly was put on heavily and rightly so. I learned that cleaning a dataset is one of the important pillars in data analysis. As the author of our textbook mentioned that cleaning a dataset takes 80% of the time when analyzing data (McKinney, 2018). So I put the words into action and I made sure I did a through job in cleaning the CSV file associated with the topic. In this process, I also learned and applied new forms of cleaning methods.

**Milestone 3** – This milestone I had learned a lot. First, I learned how to connect to a website using Beautiful Soup in Python. Next, I learned how load and parse the website data. This is also needed a bit of HTML knowledge in order to apply it to the code. Another step I learned is to convert the website data I extracted into data frames in order to make the process of cleaning the data a little bit simpler.

**Milestone 4 –** In this milestone, I learned how to extract API data. First, I learned how to connect to an API using urlib, reuqests and json packages in Python. Next, I had to load and parse through the API information and made sure the information is in an organized and readable format in JSON. Furthermore, I had to learn how convert JSON information into data frames and then insert the data into a csv file in order for me to clean the data.

**Milestone 5 -** This would be the last milestone of the term project. In this milestone, I had the opportunity to merge all three datasets that I created from Milestone 2 -4. This is was a refresher, as I had learned how to merge datasets in the Week 5-6 assignments. One of the new elements I got to learn this milestone was connecting to SQL and then creating and loading the merged data all in Python. The second element I learned is how visualize data using another visualization package in Python called Seaborn.

All in all, I learned a lot by doing this term project and unlocked a lot of Python’s capabilities. I learned that Python can be very useful and can be an all-one programming language in the world of data science.

**Reference**

McKinney, W. (2018). Data Cleaning and Preparation. In *Python for data analysis data wrangling with Pandas, NumPy, and IPython* (p. 191). Beijing: O'Reilly.