The DS project was a very helpful project to my understanding of data ingestion, transformation, and the loading process. Throughout the project, I made a python script capable of fetching data from different origins, transforming it, and storing it in different formats.

One of the biggest challenges in this project was the process of handling different data formats and making sure that all formats could be converted to another format depending on the user's preference. It took me several days to develop a course of action and understand the steps of this part of the assignment. API integration also posed challenges--this topic was something that I felt unconfident in going into this project, so it took me several days to review the methods of fetching data from APIs.

However, there were certain aspects of the project that were easier than anticipated. Developing a method to modify the data that was inputted was a quick task. Similarly, creating summaries of the data pre-process and post-process were very easy to implement, which was surprising to me. Additionally, designing the code that implemented user input was very easy--I thought this would be the most challenging task of the project, but it was the easiest.

I did unexpectedly have last-minute troubles with organizing Github. I ended up creating a new repository entirely to neatly compile my three files for this project, but that was a curveball I did not expect on encountering.

A utility like this will be helpful for many future data projects, as it provides a foundation for automating data extraction and transformation, saving time and reducing human error. Having a reusable and customizable ETL pipeline will streamline integrating various data sources. As the demand for data-driven decision-making grows, being able to efficiently process and manipulate data from multiple formats will be a critical skill in any data-centric project.