

Jackson Haiz  
3/24/2025  
Prof. Williamson  
DS 2002

### Data Project Reflection

The most challenging part of my data project would have to be writing the functions to switch file types. The logic was simple. If the user inputs a “CSV” file using the JSON file transformation function, he will get a CSV file. However, getting data to transition seamlessly between all 3 different types of files is harder than I expected. I would frequently get “Table not found” or “Trailing Line” errors with the SQL database and the JSON files that would just stump me.

The aspect of my project that was easier than expected was the data scraping. In a separate data science class, we used a gadget selector in order to data scrape off of websites. This turned out to be the most simple part of my project.

Aspects that turned out more difficult than I expected was the syntax and many options for the files and figuring out which options worked and which didn't. I was using the “new line” feature for the JSON file and trying to turn it into a SQL database. Unbeknownst to me and according to the internet, it turns into a JSONL, which is the new-line delimited format. This format could not be read by the JSON functions which ended up giving me the “Trailing Line Error”. I spent a while trying to figure out this issue but after I did, it was smooth sailing

This utility can be used for future data projects because of how helpful it is. I felt like an NFL data scientist doing the project. The ability to data scrape and write code to transition files is incredibly useful. This can be used to create data sets and run complex tests like linear regression in order to retrieve data and find out what data means. In other classes, I am doing projects where we try to retrieve data in order to run complex tests, but there aren't always options. This data project showed me the utility and usefulness of creating your own data set using CSV files and data scraping. As I continue my data science journey, I will always come back to this project as the one that taught me the practicalness of ETL data methods.