Project 2 Summary

Allison Brown

After completing this project, I feel more confident in my ability to clean and transform datasets in Jupyter Notebook and SQL. In retrospect, I would like to go back and find different datasets that had more compatibility and similar datas in them. Of the four datasets that I started with, only three of them ended up making it into the final database in SQL. This is due to all the datasets only having one to two columns of similar data. While initially it seemed as though I had found rich sources of lots of data, after diving into it, there was not as much compatibility across them as I thought, drastically limiting the amount of data available in the database. Future opportunities for improvement could include identifying and removing rows of incomplete data as long as it does not dramatically reduce the size of the final dataset.