Sources of Data:

JSON : stock.json <https://datahub.io/core/s-and-p-500-companies>

CSV: all\_stocks\_5yr.csv <https://www.kaggle.com/camnugent/sandp500>

CSV: wiki.csv <https://en.wikipedia.org/wiki/S%26P_500_Index>

Type of Transformation Needed for Data:

JSON: From the stock.json we obtained the name, sector, and symbol. We took the open and close from the second csv and added it to the first data. We then merged this data frame with the all\_stocks\_5yr.csv data frame.

CSV: all\_stocks\_5yr.csv We transformed the CSV into a data frame. We filtered out the open and close for the stock. We used the open and close to create a new data frame called df2.

We then loaded wiki.csv to the headquarter location and merged it using the symbol using an inner join.

We cleaned the data by replacing the empty cells with “No Data Available” instead of deleting rows of the empty cells.

We then connected to postgressql and cleated and empty database, stocks3, in postgres. we right clicked on the server we were using and clicked 'create database' to do this.

We then brought back out data from postgres to csv and into the jupyter notebook.