

ETL Project  
August 2019  
Heather, Mike, and Tommy

For our ETL Project, we used data found at <https://www.kaggle.com/dgawlik/nyse#securities.csv>, a page that listed NYSE stock exchange historical data.

We were able to merge our data using three of the CSV files available at that website, specifically 'securities.csv', 'fundamentals.csv' and 'prices.csv'. We merged these CSV files using their common primary key "Ticker Symbol".

Upon merging the files, we cleaned the dataframe created. The columns were also renamed to show a cleaner format, IE from 'ticker symbol' to 'ticker\_symbol'.

Using psycopg2, we were able to easily transfer the dataframe information into SQL and then interact with that dataframe in Postgres.

We then connected the data to a local database using Postgres. We created the dataframes on our computers within our local SQL environment first, then pushed the cleaned dataframe into SQL using pandas, using the 'to\_sql' command.

Now we were able to confirm the data was added by querying the securities and fund stock databases.

There is also an app we created via Flask that would use our session engine to query the data from the dataframes, which we hope would return all data for a ticker symbol. This also had the exception to print 'no such ticker symbol' if no information was available or an incorrect ticker symbol was given.

We were all able to contribute to this project, we feel we worked well together and are happy with the results.