Project 2 – ETL (Group 4: Sam Clements, Matt Lett, Nathan Weinberg)

Extract:

Stock Exchange data came from a Kaggle website. There were 3 files, and all were .csv

<https://www.kaggle.com/mattiuzc/stock-exchange-data/version/2?select=indexInfo.csv>

Foreign Currency Exchange info was retrieved from:

<https://excelrates.com/>

Transform:

* Loaded all the csv files from the listed sources
* Shortened all df to only contain data from 2000 to present
* Dropped all null values
* Converted all dates that were in string format to date objects
* Renamed columns, dropped columns as needed
* Remove string characters from numerics
* Concatenated all Exchange rate dataframes into one

Load:

Used Postgres for the database.

* Created stock\_markets\_db in pgAdmin
* Used SQL statements to create the 3 tables
* Used Jupyter Notebook to connect to db
* Loaded 3 dataframes into Postgres from jupyter notebook
* Confirmed loaded data into tables by retrieving the data in jupyter notebook