Project 2: Team Delta Proposal

Our team decided to pursue the financial track of data analysis for our ETL project. With that in mind, we used Kaggle to search for various financial datasets and decided on a dataset called “Top Tech Companies Stock Price”. The dataset can be found at :  
<https://www.kaggle.com/datasets/tomasmantero/top-tech-companies-stock-price> .

This dataset had CSV files which contained stock data from the top 100 tech companies. With that in mind, we plan on selecting a few of the tech companies CSV’s and extracting the data out of them through Jupyter Notebook. Once the data has been extracted and we see what columns / data we are working with, we want to transform the data to make it more readable. For example, one way that we would like to transform the data is by adding a “Ticker” column to be able to differentiate between the respective stocks once they’re loaded into a database.

Once we have the data transformed to make it as readable and user friendly as possible, we plan on loading these datasets into a database using PostGreSQL. We planned on using MongoDB but after discussing with the team, we have decided that it would make more sense to use a relational database since the stock’s are all related. Once we have loaded the datasets into PostGreSQL, we would be able to query against the databases and further analyze the stock data.