

The `SP500History.csv` file contains 5 years (from 2013-02-08 to 2018-02-07) of historical stock price information for the companies tracked by S&P 500.

Data Format

Each of the record in the CSV file contains the following information:

- **Timestamp:** date for which the stock price was recorded
- **Open:** the strike price of the first trade of the day (USD, same below)
- **High:** the highest strike price during the day
- **Low:** the lowest strike price during the day
- **Close:** the strike price of the last trade of the day
- **Volume:** the total number of shares traded during the day
- **Symbol:** the alphabetic string representing the company, e.g., AAPL represents Apple Inc.

In total, there are 619040 records in the file.

Loading into the Database

A `COPY` statement can be used to load the csv file into a database in PostgreSQL. The following steps provide an example of doing so:

- Upload the csv file to the VM (say, under `/data`), e.g., using `scp`.
- Create a table in the database:

```
CREATE TABLE Stocks(timestamp DATE, open REAL, high REAL, low REAL, close REAL, volume INT, symbol VARCHAR(5), PRIMARY KEY(symbol, timestamp));
```
- Load all the records in the csv file into the table:

```
COPY Stocks(timestamp, open, high, low, close, volume, symbol) FROM '/data/SP500History.csv' DELIMITER ',' CSV HEADER;
```

Your database design may have a different schema for storing these records, so you should load data accordingly.