

Stock Market ETL Project

Group Five

Chris Haub, Ryan Permenter, and Jim Worlein

21 December 2020

Overview

Project to extract, transform, and load stock data into a PostgreSQL database and perform queries on this database.

Data & Modeling Approach

Data was scraped from three tables found on Yahoo Finance and combined with a CSV file from NASDAQ. The three Yahoo tables contain daily stock updates on the most active, highest gainers, and biggest losers. The CSV file contains basic information on each listed company by stock symbol.

These four tables were uploaded to a PostgreSQL database with serial primary IDs. The stock symbol acted as a primary key for all the tables.

NASDAQ CSV File:

(<https://www.nasdaq.com/market-activity/stocks/screener?exchange=NYSE&render=download>)

Yahoo Finance:

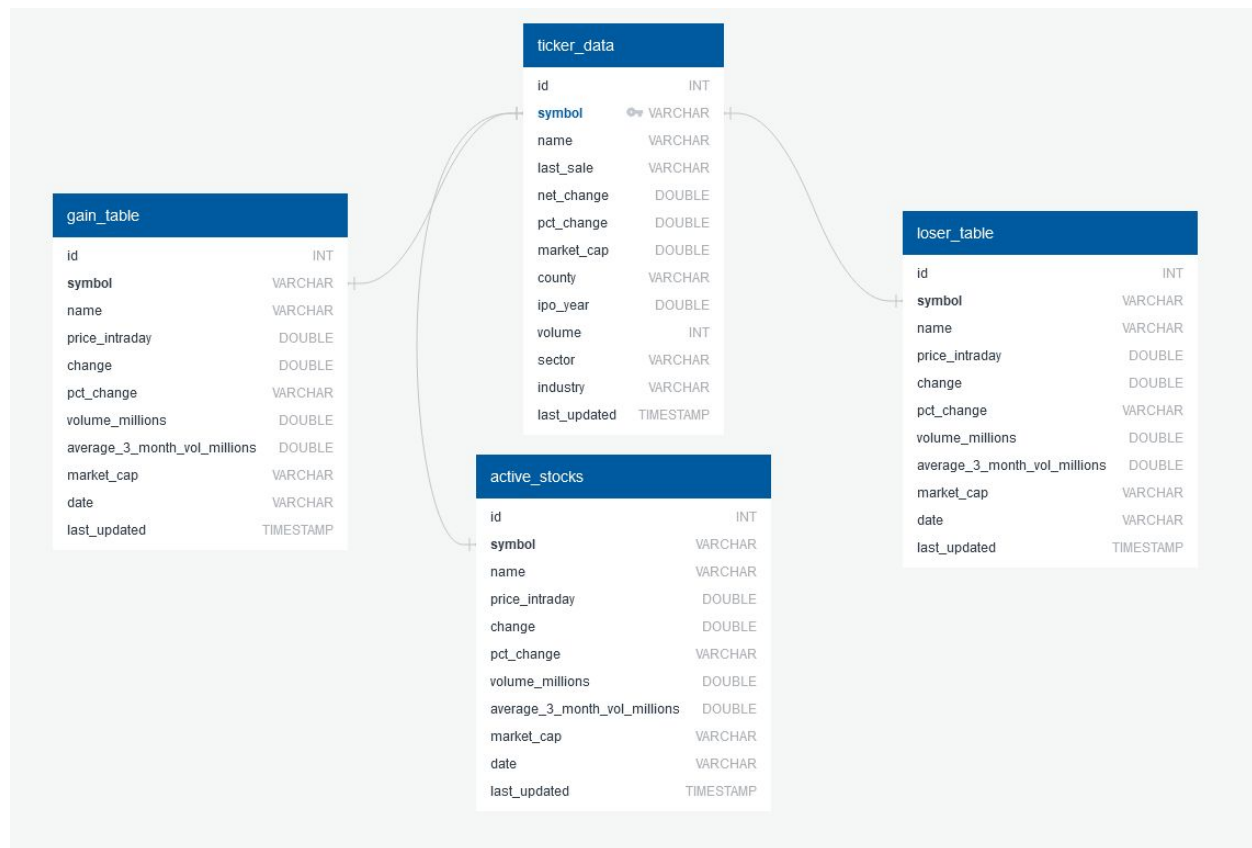
url = "https://finance.yahoo.com/most-active"

url2 = "https://finance.yahoo.com/gainers"

url3 = "https://finance.yahoo.com/losers"

Group 5ETL Project







ERD diagram



SQL Queries

```

SELECT *
FROM active_stocks a
INNER JOIN gain_table g ON g.symbol = a.symbol
;
    
```

Data Output		Explain	Messages	Notifications				
 index bigint 	Symbol text 	Name text 	Price_Intraday double precision 	Change double precision 	PCT_Change text	Volume_Millions double precision	Avg_3_Month_VoL_Millions double precision	
1	16 LAZR	Lumina...	35.15	7.53	+27.26%	36.438	7.903	
2	19 FUBO	fuboTV...	49.48	10.23	+26.06%	32.168	4.175	
3	24 QS	Quantu...	94.9	21.17	+28.71%	26.559	5.771	

Group 5ETL Project

```
SELECT *
FROM active_stocks a
INNER JOIN loser_table l ON l.symbol = a.symbol
;
```

Data Output	Explain	Messages	Notifications
index bigint	Symbol text	Name text	Price_Intraday double precision
1	7	TSLA	Tesla, I...
2	23	OXY	Occide...

```
SELECT *
FROM ticker_data t
INNER JOIN active_stocks a ON a.symbol = t.symbol;
```

Data Output	Explain	Messages	Notifications
id integer	symbol text	name text	last_sale text
1	4588	NIO	NIO Inc...
2	17	AAPL	Apple I...
3	716	BAC	Bank of...
4	10	AAL	Americ...
5	2746	GE	General...
6	5224	PLTR	Palanti...
7	2369	F	Ford M...
8	6583	TSLA	Tesla I...
9	3456	INTC	Intel Co...
10	6999	WFC	Wells F...
11	1229	CCL	Carniva...
12	2448	FEYE	FireEye...

```
SELECT *
FROM active_stocks a
LEFT JOIN ticker_data t on t.symbol = a.symbol
WHERE t.ipo_year = 1999;
```

Data Output	Explain	Messages	Notifications
id integer	symbol text	name text	price_intraday double precision
1	19	PLUG	Plug P...

