

The background is a dark blue gradient. It features several vertical white lines of varying lengths. Scattered throughout are small squares in teal, orange, and pink. Some squares are solid, while others are outlined in white.

STOCK MARKET DATA PIPELINE

Manveer Sadhal
October 15, 2021

NEED

Establish a data pipeline which retrieves, cleans, and stores stock market data to be deployed in a web application.



DATA AND SCOPE

DATA

- Daily price at market close
- General company information (e.g. products, sector)
- Financial metrics (e.g. market cap, P/E)

SCOPE

- Companies traded on US stock market exchanges - NASDAQ, NYSE, AMEX
- Largest 2,000 companies by market cap
- 15 years of market data

5,512,745

total records

(...and counting...)

TOOLS

DATA ACQUISITION

- Yahoo Finance API (yfinance library)

DATA CLEANING

- Pandas
- NumPy

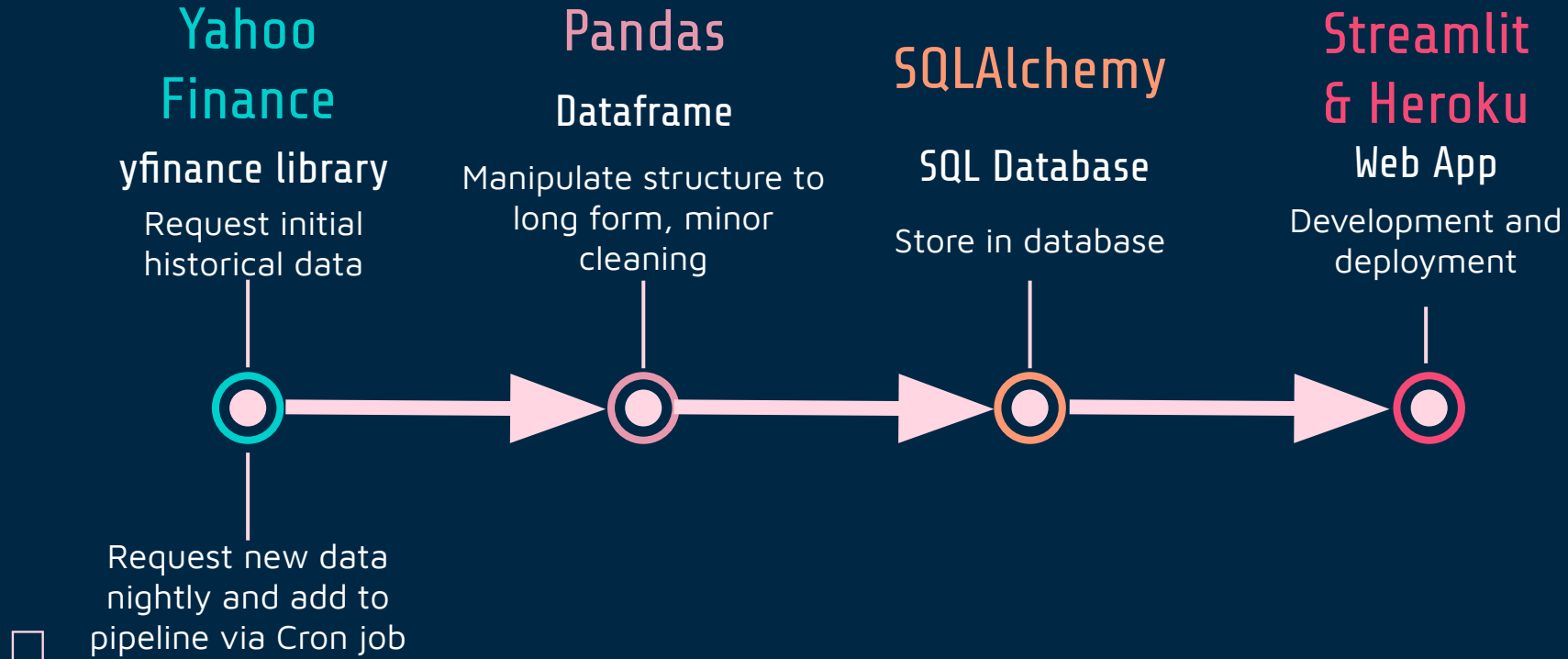
DATA STORAGE

- SQLAlchemy for SQL database connection
- Pandas to push to DB

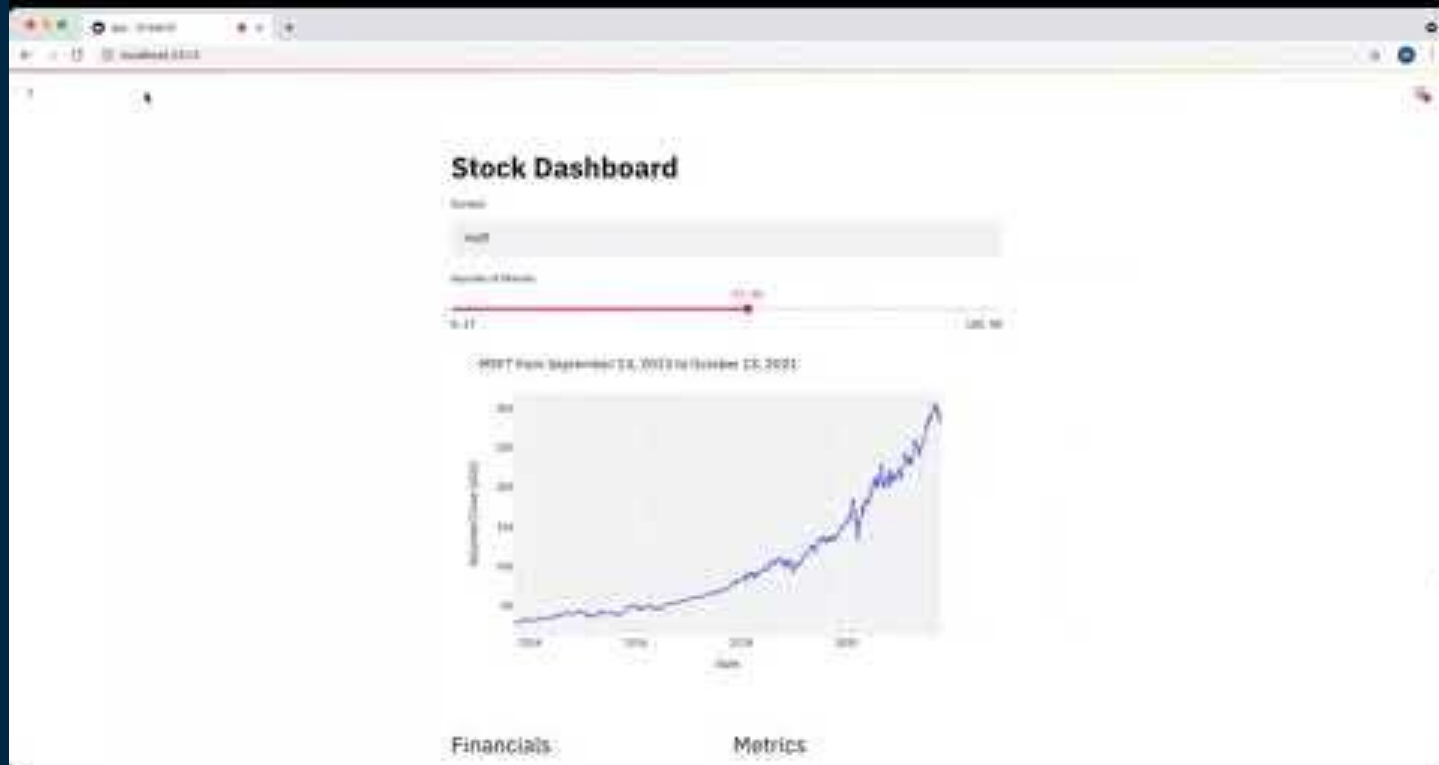
WEB APP DEPLOYMENT

- App developed with Streamlit
- Plotly for interactive chart
- Deployed via Heroku

PIPELINE OVERVIEW

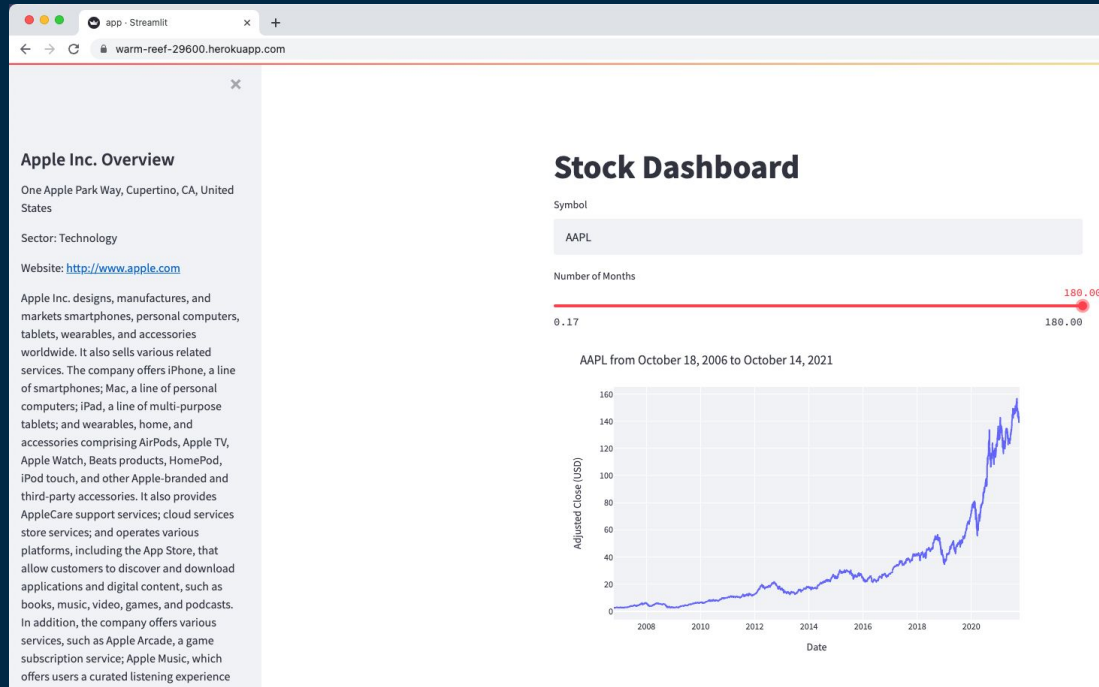


DEMO



HEROKU DEPLOYMENT

<https://warm-reef-29600.herokuapp.com/>



FURTHER WORK

- Additional data - market indices, cryptocurrency
- Functionality to allow user to build and track portfolio performance
- Data storage on a cloud platform
- Algorithmic analysis/modeling

APPENDIX

The background is a dark navy blue. It is decorated with various geometric elements: small squares in teal, orange, and pink, some of which are solid and others are hollow outlines. Thin, light-colored vertical lines of varying lengths are scattered across the page, some extending from the top or bottom edges.

Crontab Example

A terminal window with a title bar that says "final — crontab -e". The window contains a single line of text: "0 23 * * 1-5 /path/to/python3 /path/to/update_db.py". The text is in a monospaced font, and there is a cursor at the end of the line.

```
0 23 * * 1-5 /path/to/python3 /path/to/update_db.py
```

Run update_db.py M-F at 11:00PM

Load new historical data, update metrics

Do you have any questions?

youremail@freepik.com

+91 620 421 83

yourcompany.com

THANKS



CREDITS: This presentation template was created by [Slidesgo](#),
including icons by [Flaticon](#), and infographics & images by [Freepik](#)
Please keep this slide for attribution