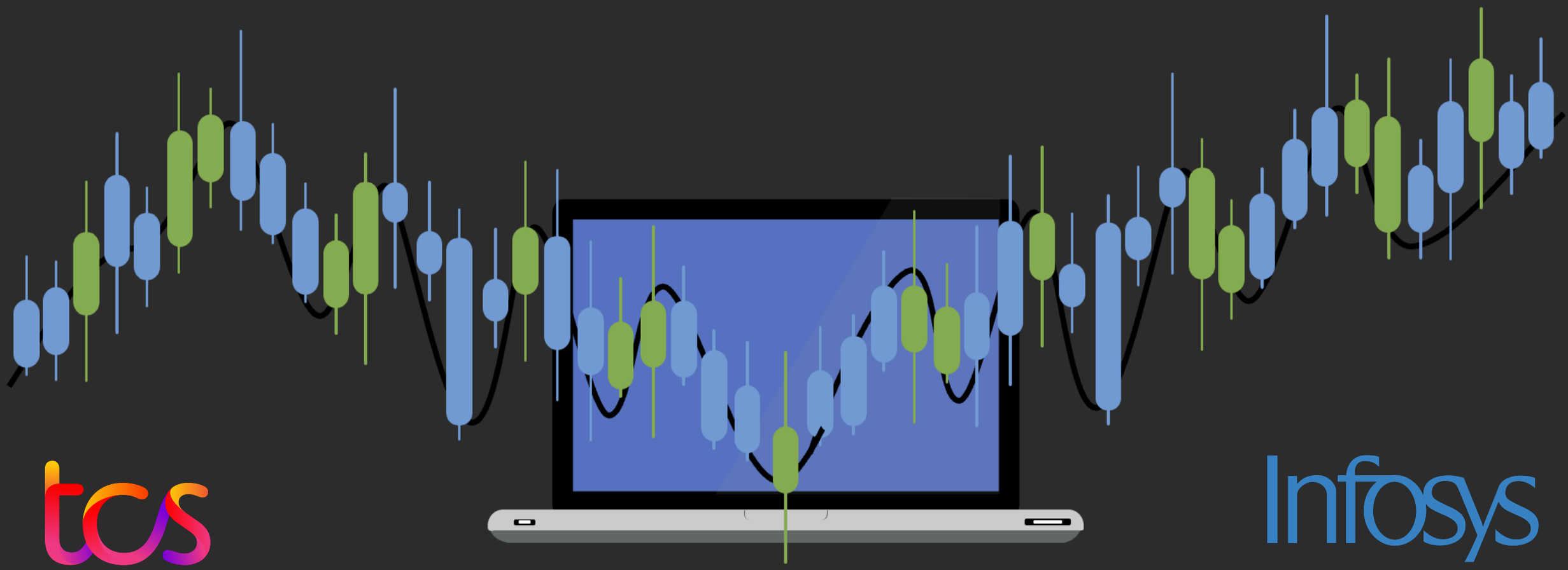


STOCK PREDICTIONS NSE AND S&P 500

DATA NINJA 7:

Akshay Deshmukh
Jai Trivedi
Nilabh Chaturvedy
Niyathi Aiella
Tanveer Singh Sapra

Problem Statement

The financial-investment market is a highly volatile and sensitive market which attracts many investors. These investors are looking for a reliable way to know the risks and returns associated with a stock listing.

We aim to predict the stock price and volume of 4 International Tech firms, namely: **Accenture, Infosys, IBM, and TCS**. We have implemented predictive and forecasting analysis to study the price variation of the stocks and predict their trends.

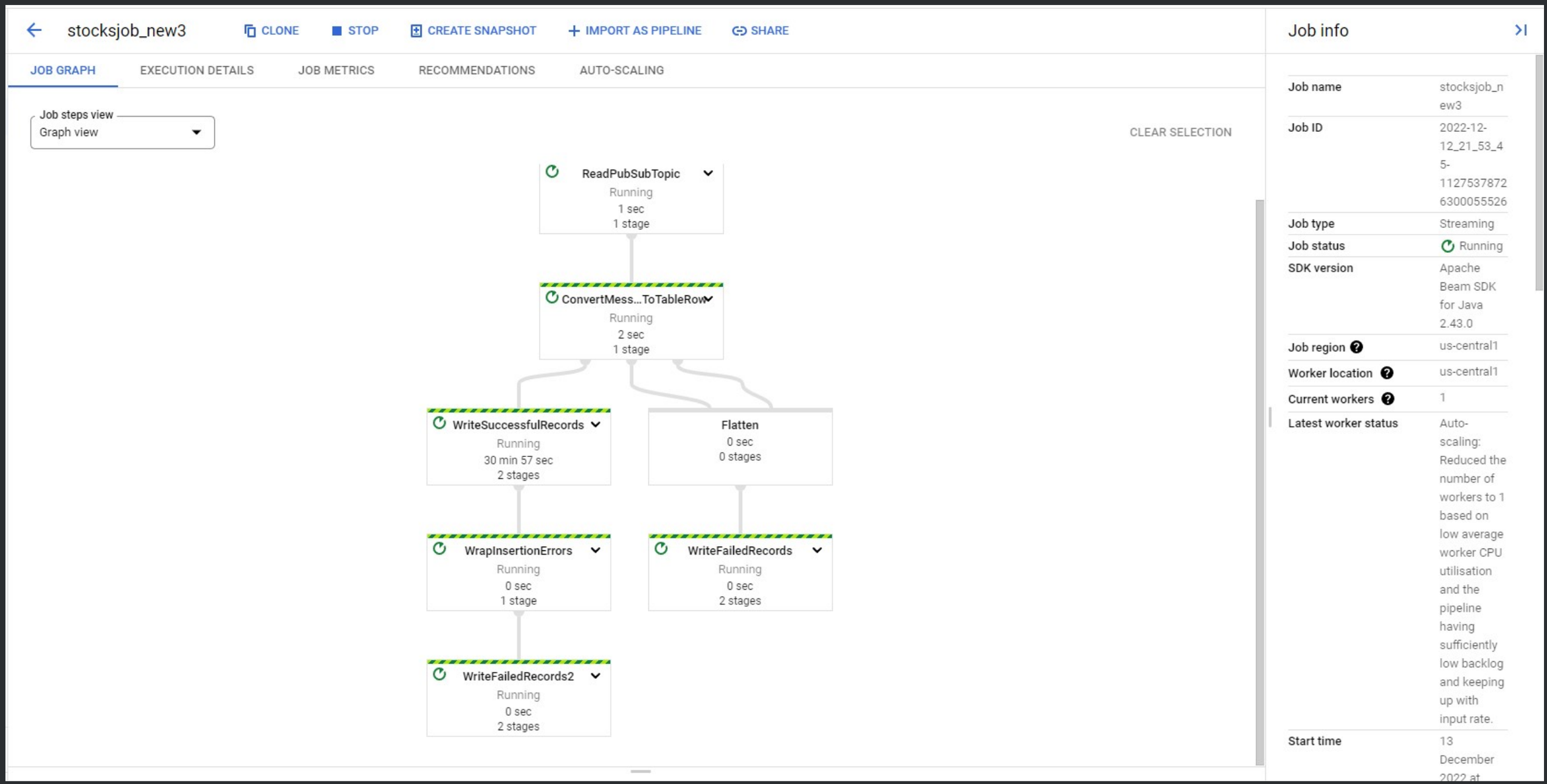
By predicting the price of the stocks, our clients are more informed of their investment decisions and have a clearer picture of the prospects of their investments.

Our data analysis revealed the trends in each of the stocks. We were also able to compare our predictive model with the actual results and retrain our model on regularly streamed data.
By studying the variability in the stock prices, we are also able to recommend stocks that are low and high risk.

Pipeline Formulation

Data flow named stocksjob_new3

-- Pumps data into bigquery and storage

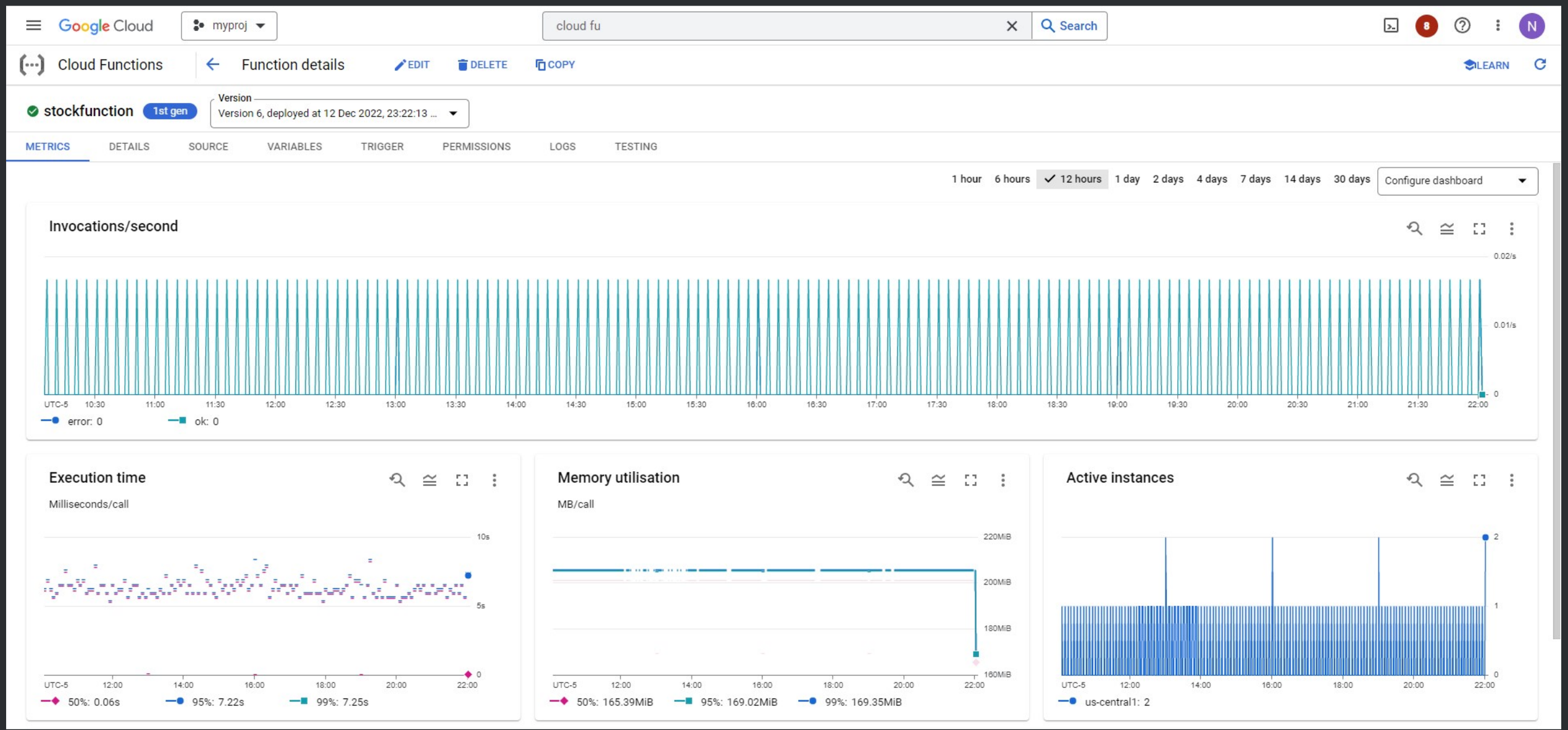


Pipeline Formulation

Cloud Function named stockfunction:

-- This inturn gets triggered by cloud Scheduler which runs every 5 minutes

-- Pulls data from YahooFinance API from the code



Pipeline Formulation

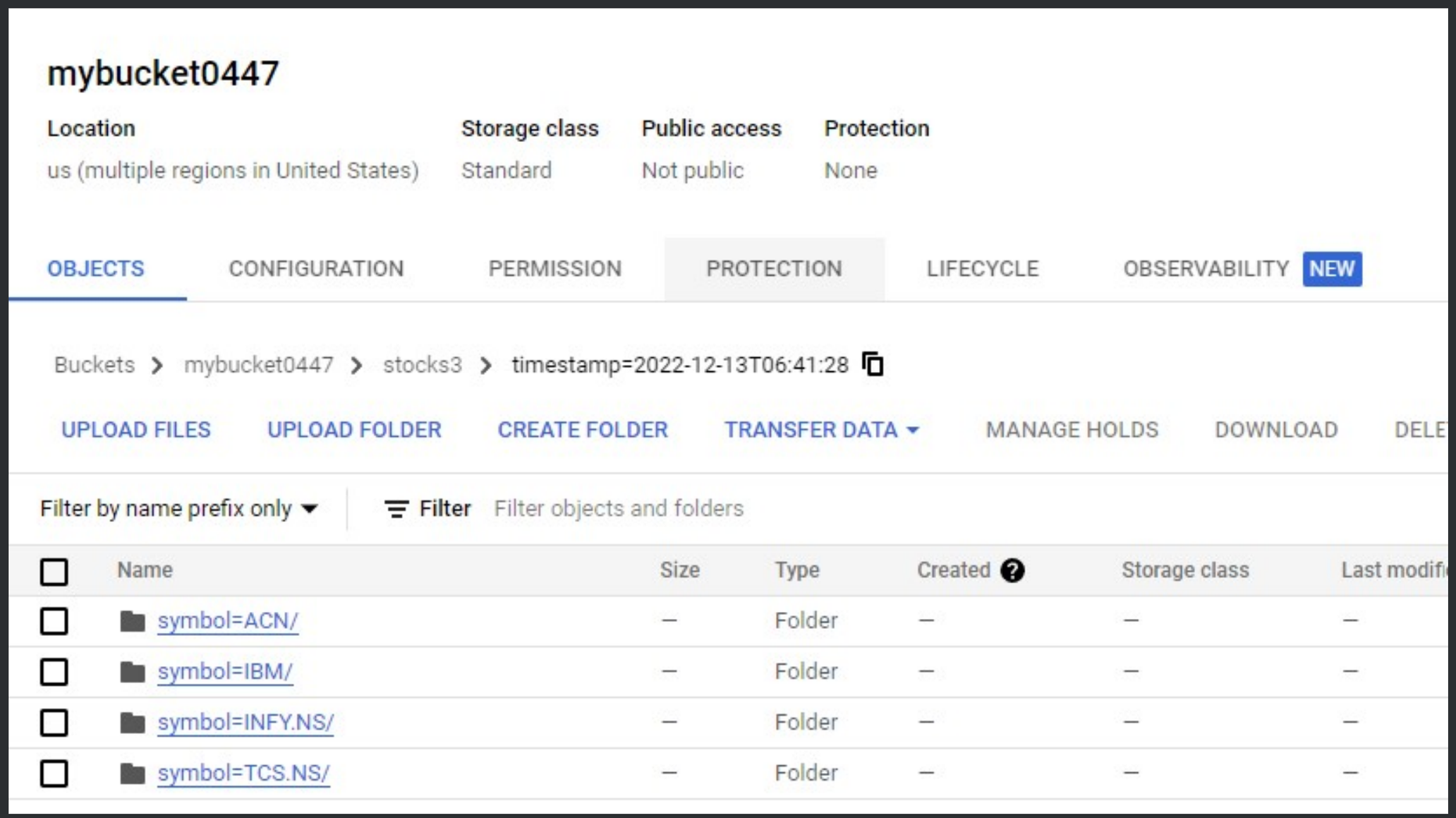
Cloud scheduer was enabled:

-- This scheduler runs every 5 minutes enabling the cloud function to run and pull the data

☐	stock_streaming_scheduler	✓ Success	us-central1	Enabled	*/5 * * * * (America/Indianapolis)

Cloud storage is populated with the data:

-- All the data that is coming from the archive and the yahooFinance is going into cloud storage bucket



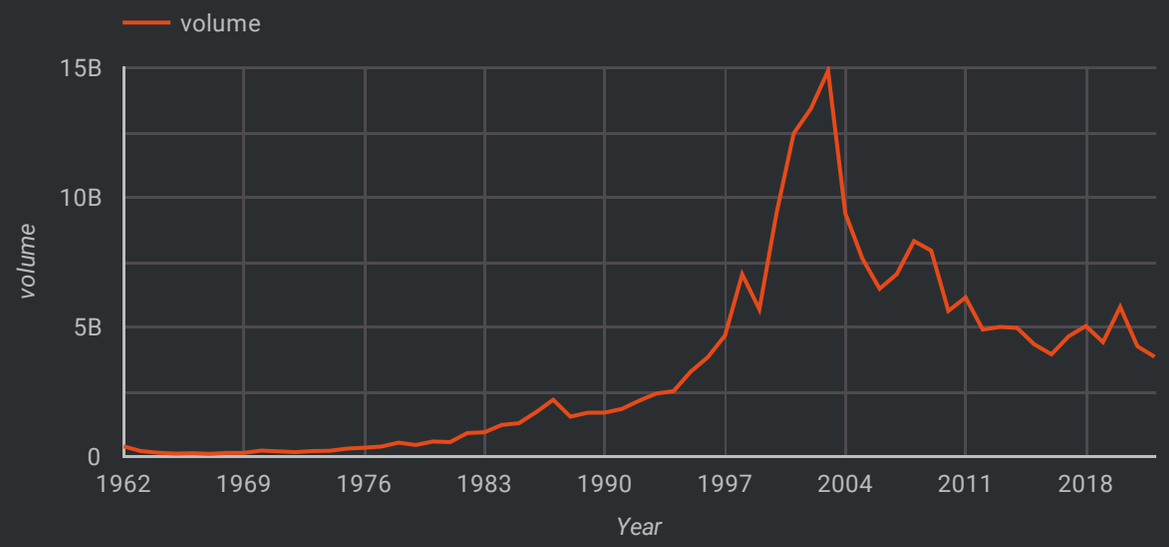
Historical Analysis

Average High

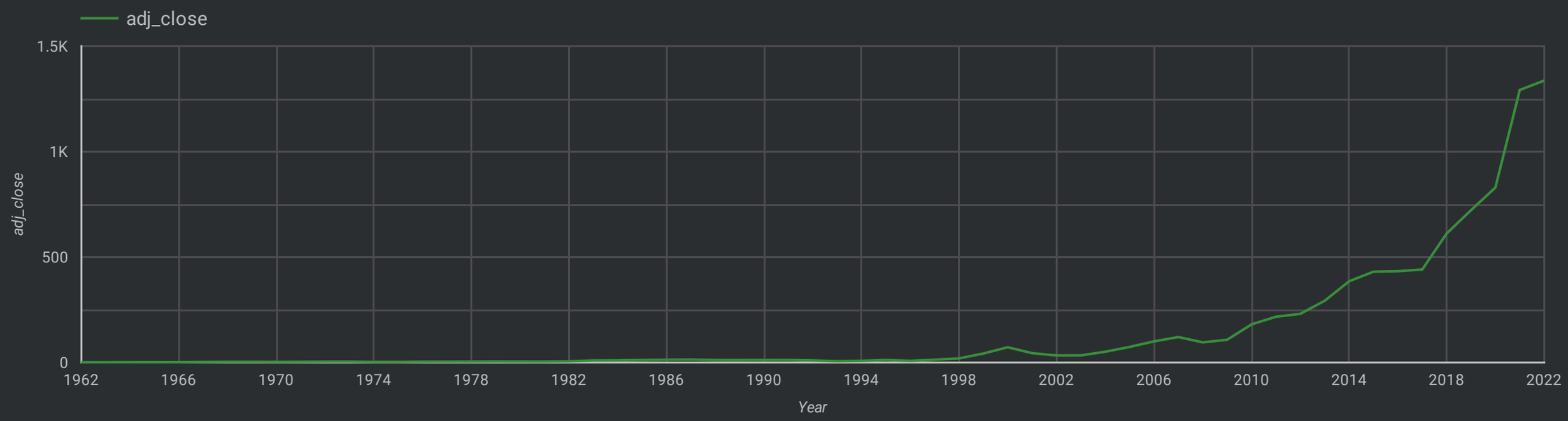
	symbol	high ▾
1.	TCS	1,055.7
2.	INFY	375.52
3.	ACN	96.34
4.	IBM	61.12

1 - 4 / 4

Volume Traded



Adjusted Close



Evaluation Of Models

ARIMA(Autoregressive Integrated Moving Average)

DETAILS | TRAINING | EVALUATION | SCHEMA

Time series ID	Non-seasonal P	Non-seasonal D	Non-seasonal Q	Has drift	Log likelihood	AIC	Variance	Seasonal period
ACN	1	2	4	False	-9,341.17	18,694.341	0.643	Weekly, Yearly
IBM	0	1	5	False	-12,837.109	25,686.219	0.186	Weekly, Yearly
INFY	0	2	5	False	-25,746.824	51,505.649	11.076	Weekly, Yearly
TCS	0	2	5	False	-26,681.341	53,374.681	78.789	Weekly, Yearly

REGRESSION - Volume Prediction

volume_prediction_boosting

DETAILS | TRAINING | EVALUATION | SCHEMA

Mean absolute error	2,489,842.0023
Mean squared error	43,586,161,573,481.52
Mean squared log error	2.6431
Median absolute error	1,053,025.5
R squared	0.5543

Boosting Regression

volume_prediction_dnn

DETAILS | TRAINING | EVALUATION | SCHEMA

Mean absolute error	6,630,523.1794
Mean squared error	141,764,249,078,145.53
Mean squared log error	122.659
Median absolute error	4,019,247.9788
R squared	-0.4495

DNN Regressor

volume_prediction_rf

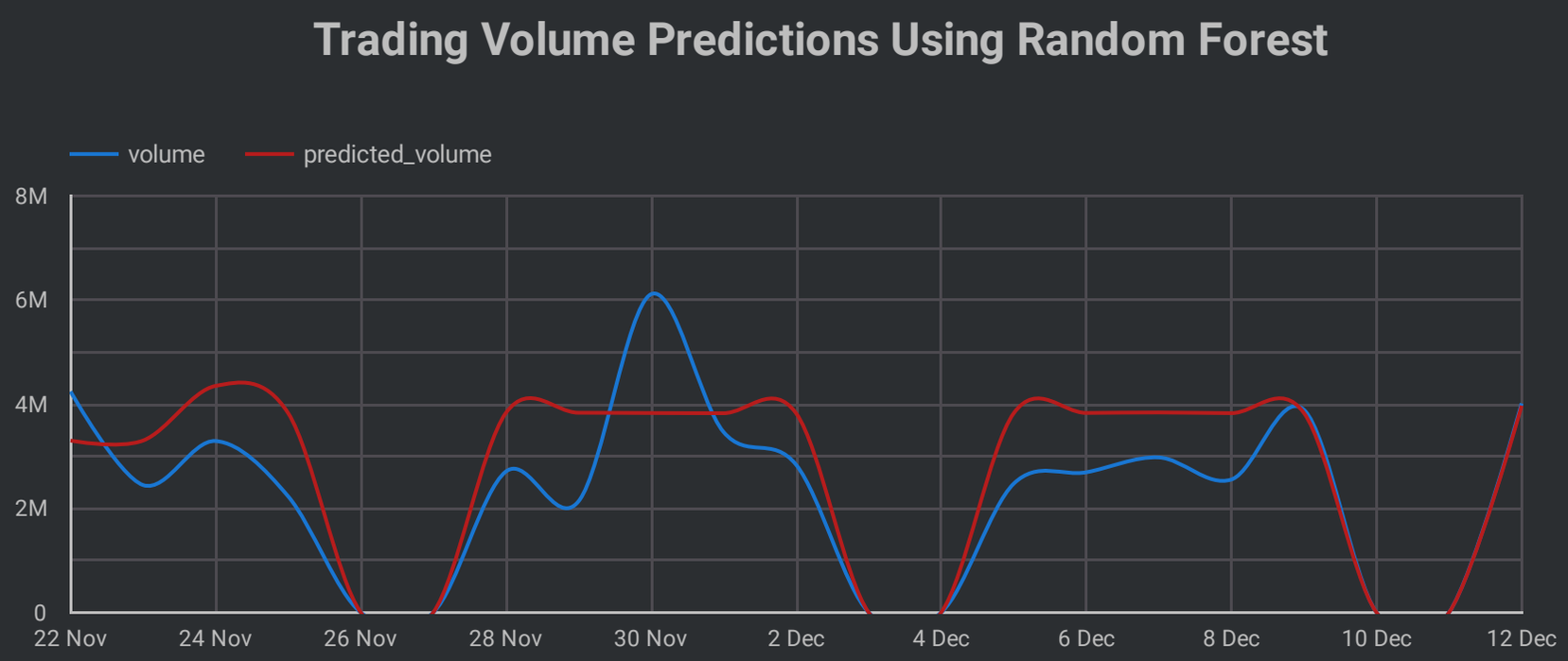
DETAILS | TRAINING | EVALUATION | SCHEMA

⚠ This model has no evaluation metrics.

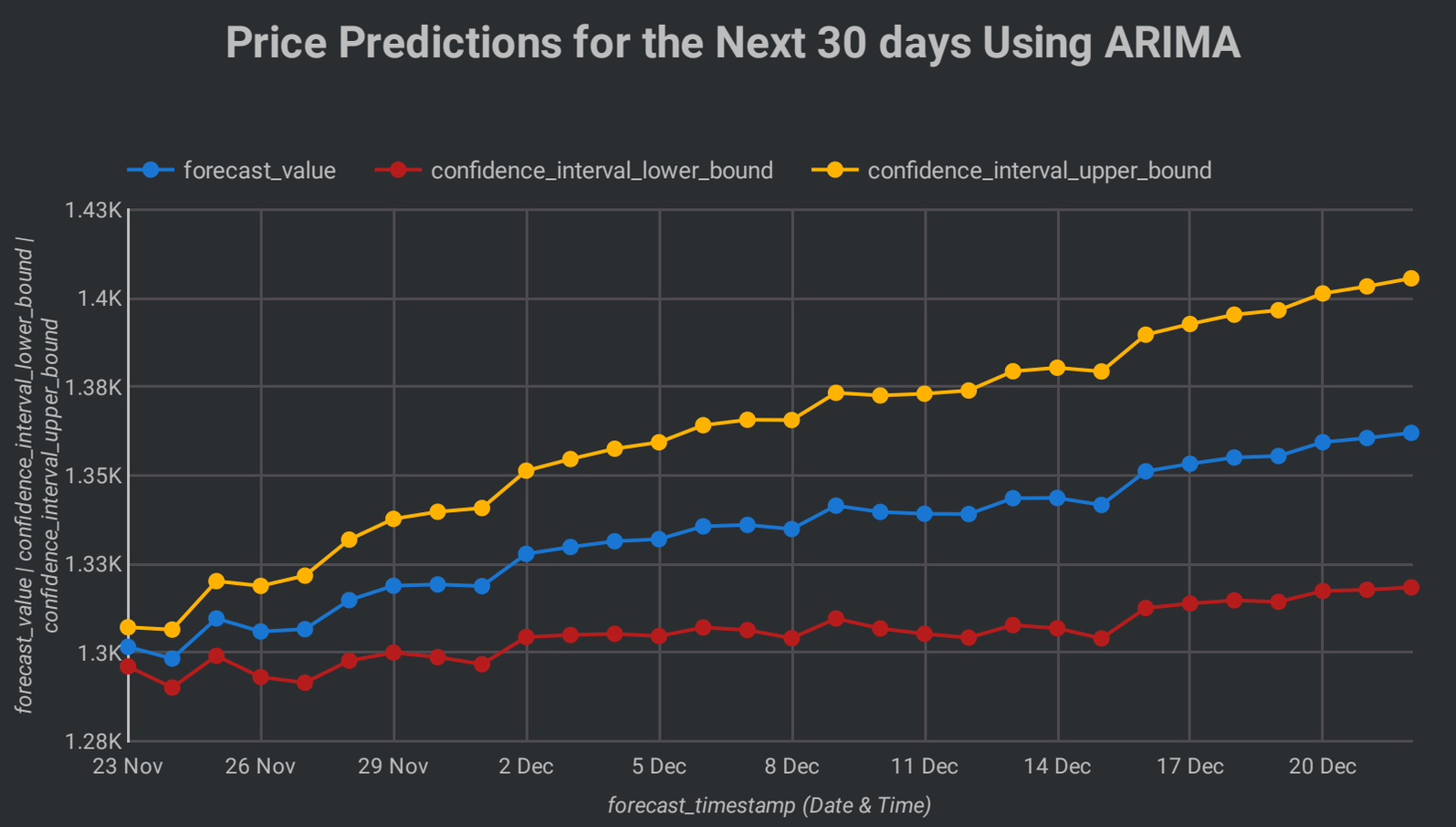
Mean absolute error	2,361,132.5923
Mean squared error	32,116,559,093,516.113
Mean squared log error	1.7534
Median absolute error	1,027,287.5
R squared	0.6063

Random Forest Regressor

Model Predictions

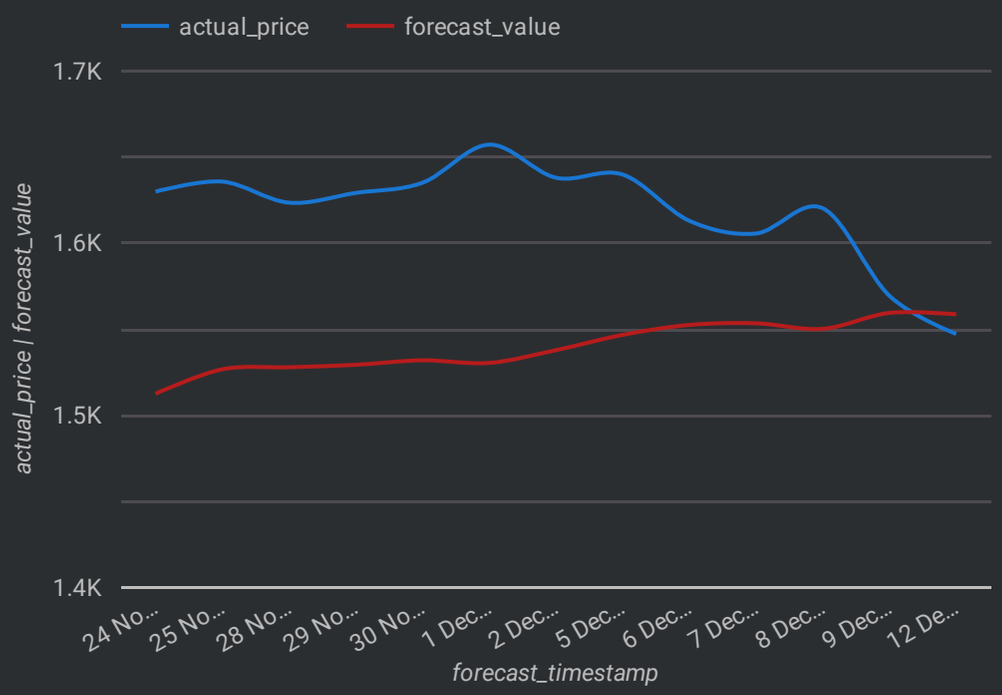


symbol
TCS
INFY
IBM
ACN

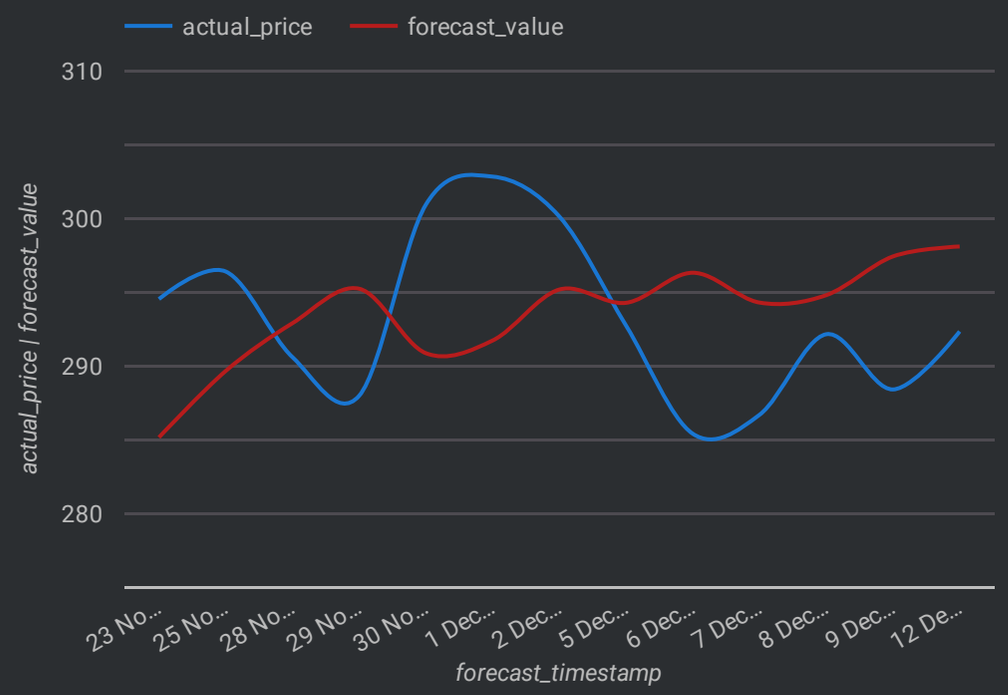


Forecast vs Actuals

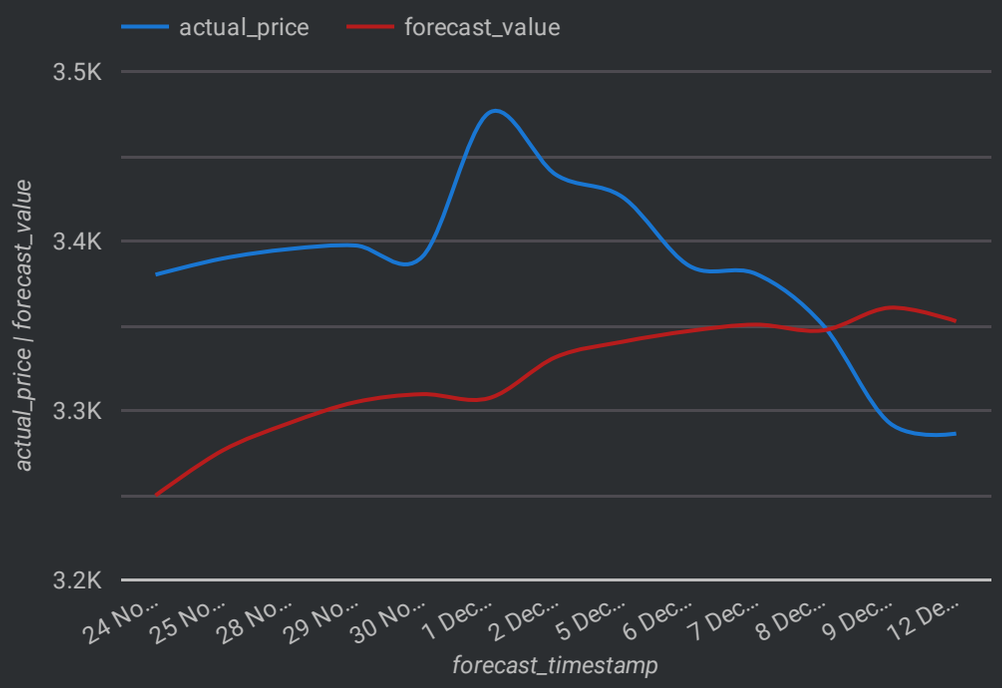
Infosys (INFY)



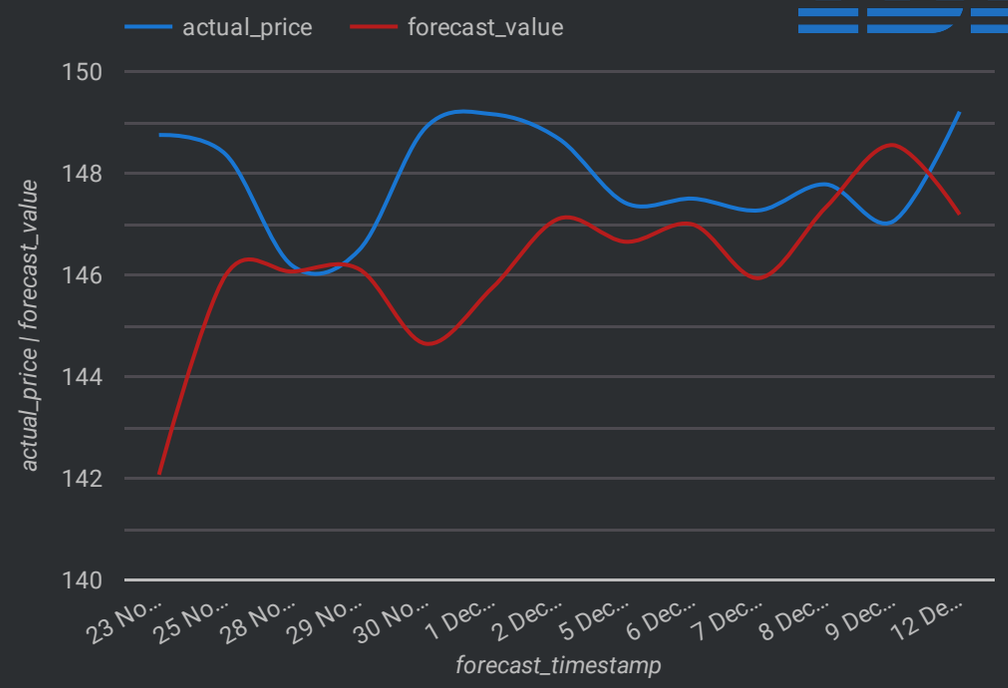
Accenture (ACN)



Tata Consultancy Services (TCS)



IBM (IBM)



Price Variation - Risks & Opportunities

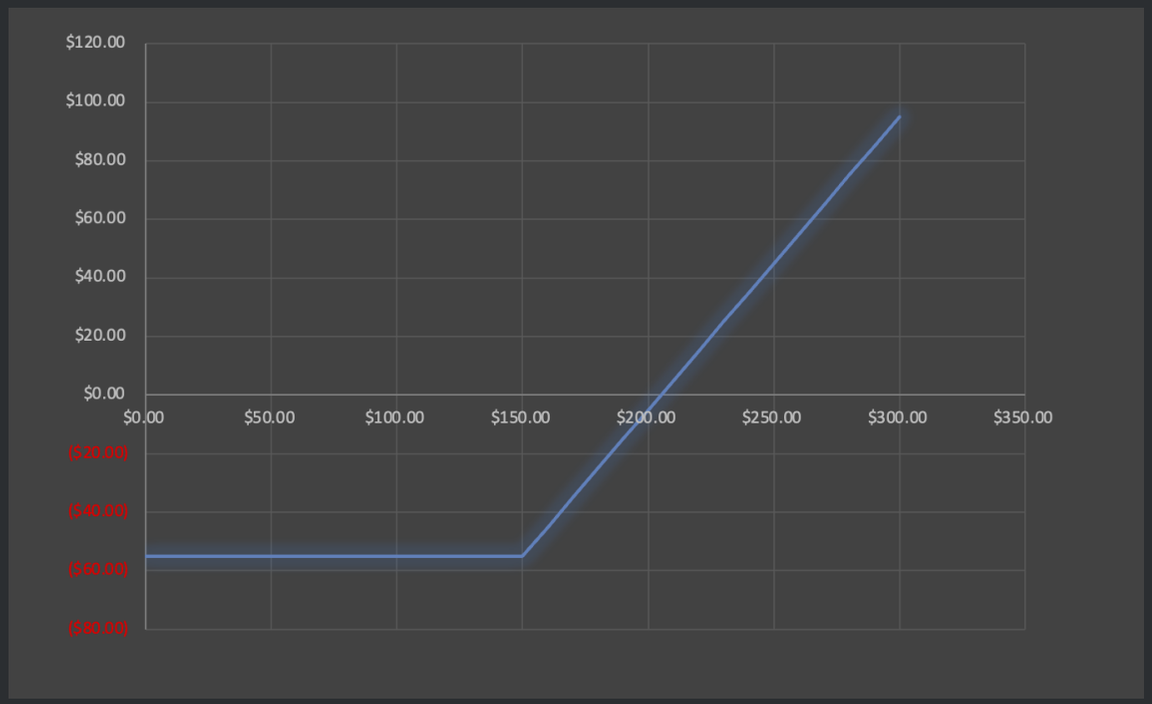
- Greater variation in stock prices means greater unpredictability and risk
- Prediction trends allow us to generalize the stock price variations
- Similarly, stocks with lesser variations would be more predictable and safer options
- **Infosys** and **TCS** are highly variable stocks. They are high-risk and less predictable
- **Accenture** and **IBM** are less variable stocks. They are low risk and more predictable

Recommendations

For institutional and retail investors, it becomes essential that they know the trends for the stocks and make their trades accordingly.

- Our predictions help such entities formulate strategies corresponding to the stocks we have chosen for the short run.

- As per the models that we implemented, it would be prudent for an investor to use the predictions for hedging by making use of a protective put in the short run



Assignment Feedback

The assignment was challenging and allowed us to work on all the aspects we covered in class. We were also excited to work on a new dataset. Yahoo Finance dataset was a good chance for us to learn more about the stock market.

The assignment did not take as long as we were well versed with the data pipeline creation. We only needed to spend time on creating our analysis. We spent nearly a week working on this project.

The skills we applied was GCS tools and material on financial investments.