Real Time Forex Arbitrage Detection & Price Prediction

CS GY 6513: BIG DATA

Presented by
Pranshu Goyal (pg2592)
Vibhor Mechu (vm2491)
Suprateek Chatterjee (sc10344)

OVERVIEW

- INTRODUCTION
- TECHNOLOGIES
- SYSTEM ARCHITECTURE
- DAG SCHEDULING
- DATA PROCESSING
- ARBITRAGE
- PRICE PREDICTION
- DASHBOARD
- FUTURE WORK

INTRODUCTION

Objectives:

- Detect arbitrage opportunities in the forex market.
- Predict future forex prices accurately.
- Implement real-time data processing for timely insights.

Why This Is a Big Data Problem:

Handling voluminous and high-velocity forex market data requires robust Big Data solutions for real-time analytics and decision-making.

TECHNOLOGIES USED

Big Query

A fully-managed, serverless data warehouse that enables scalable analysis over petabytes of data with a SQL-like interface for running super-fast queries.

Apache Spark

An open-source unified analytics engine for large-scale data processing, known for its speed, ease of use, and sophisticated analytics capabilities.

GCP Cloud Storage

A highly durable and secure object storage service that offers global edge-caching and accessibility for hosting large-scale data in the cloud

Prefect

A data workflow management system designed to orchestrate and optimize data pipelines, ensuring fault-tolerant execution and streamlined automation of complex workflows.

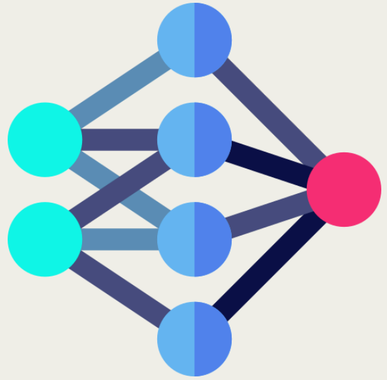
TECHNOLOGIES USED

Compute Engine

A highly customizable and scalable infrastructure service from Google Cloud that provides virtual machines (VMs) on demand, allowing you to run large-scale computing workloads on Google's infrastructure.

DataProc

A fast, easy-to-use, fully managed cloud service for running Apache Spark and Apache Hadoop clusters in a simpler, more cost-efficient way, supporting rapid processing of big data sets



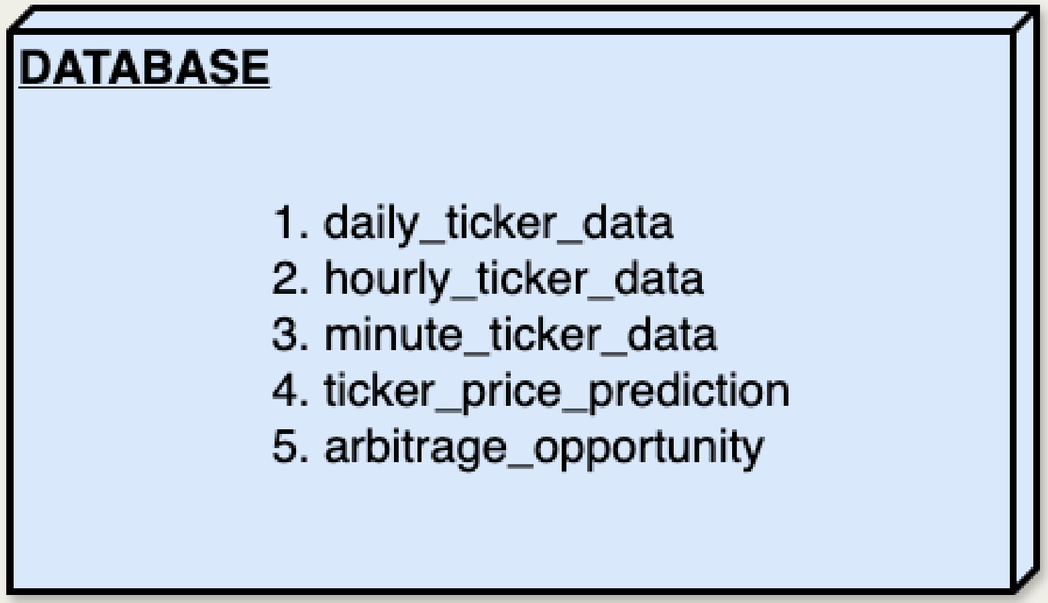
LSTM Model

Long Short-Term Memory is a type of recurrent neural network (RNN) architecture used in the field of deep learning that excels at learning from sequences of data, making it ideal for time series prediction, natural language processing, and other sequential tasks.

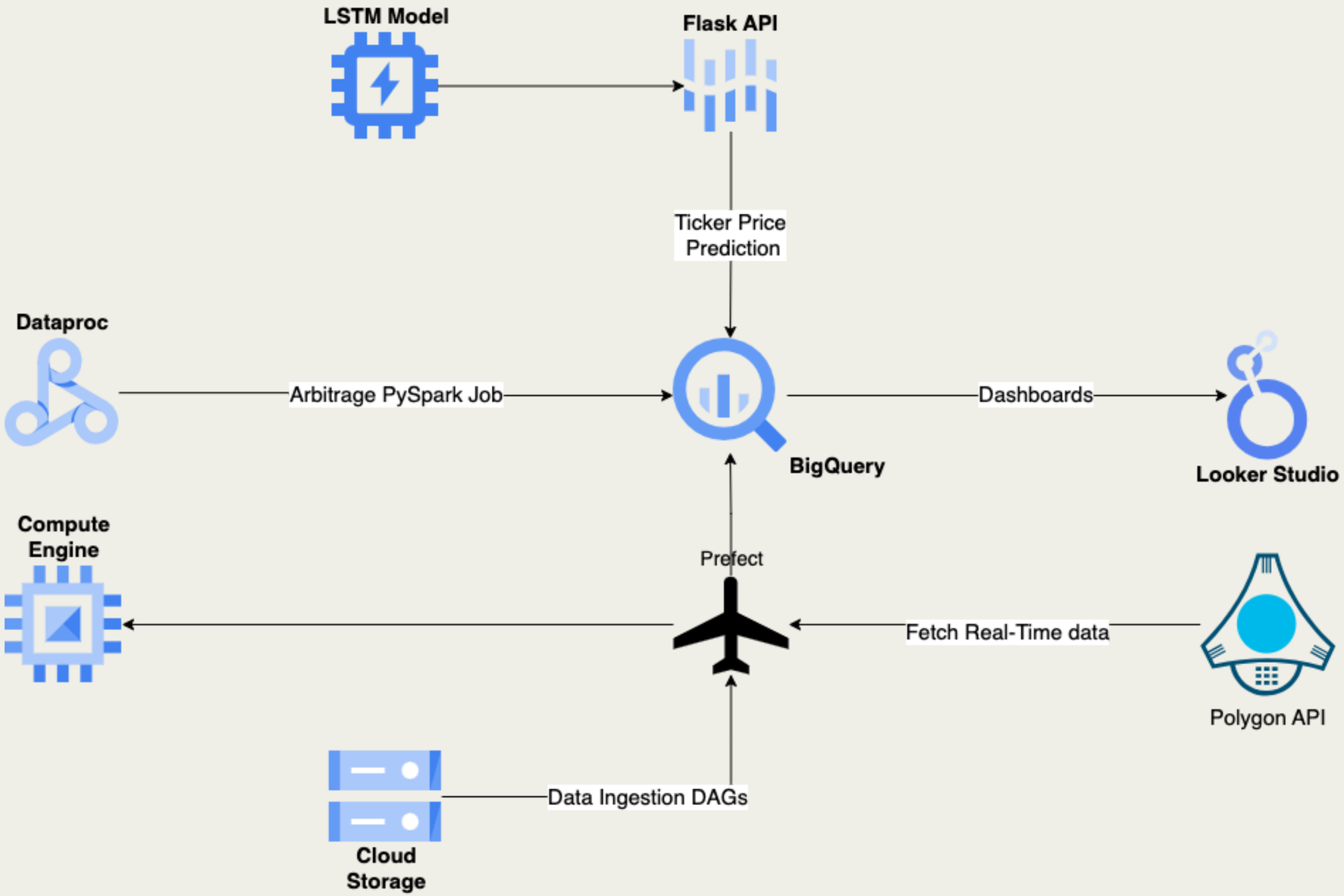
Looker Studio

A data visualization tool from Google that enables users to create interactive dashboards and engaging reports from various data sources, offering customizable views to derive insights and make data-driven decisions effectively.

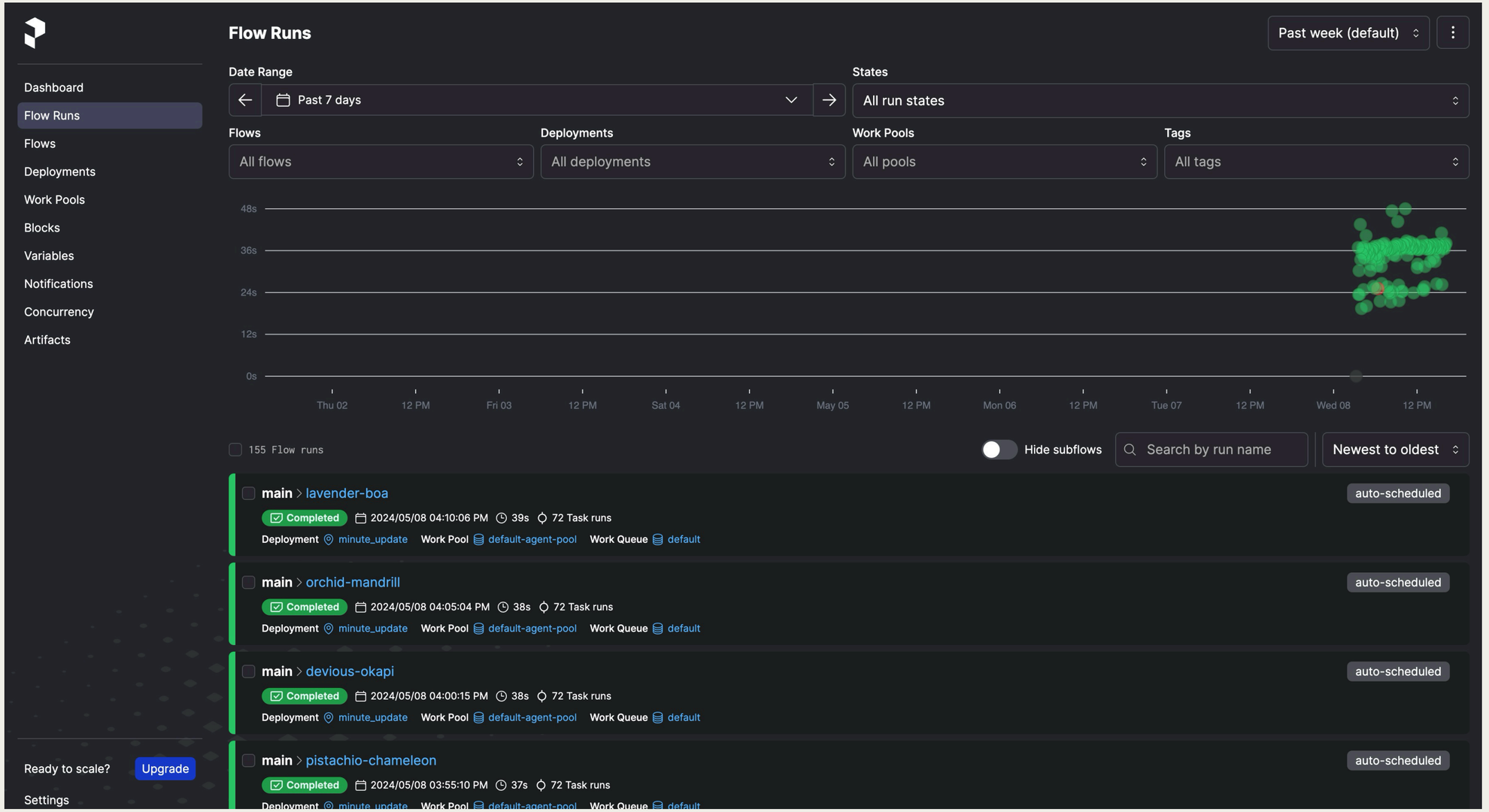
DATA MODEL


DATABASE
1. daily_ticker_data
2. hourly_ticker_data
3. minute_ticker_data
4. ticker_price_prediction
5. arbitrage_opportunity

SYSTEM ARCHITECTURE



DAG SCHEDULING



ARBITRAGE

What is Arbitrage?

Arbitrage involves exploiting the price differences of identical or similar financial instruments across different markets or in different forms to achieve a risk-free profit.

Forex Arbitrage

In forex markets, arbitrage is the simultaneous purchase and sale of a currency to profit from an imbalance in the price. It is a trade that profits by exploiting the price differences of identical or similar financial instruments on different markets or in different forms.

Triangular Arbitrage

Triangular arbitrage involves three trades, closing with a currency that one started with, and exploiting the price discrepancies in three different currencies in the forex market.

- Example: Consider currencies USD, EUR, and GBP. Triangular arbitrage might involve:
  - Buying EUR with USD.
  - Using EUR to buy GBP.
  - Finally, converting GBP back to USD at a favorable rate, resulting in more USD than initially started with.

Formula:

$$\text{Arbitrage Value} = (E_{AB} \times E_{BC} \times E_{CA})$$

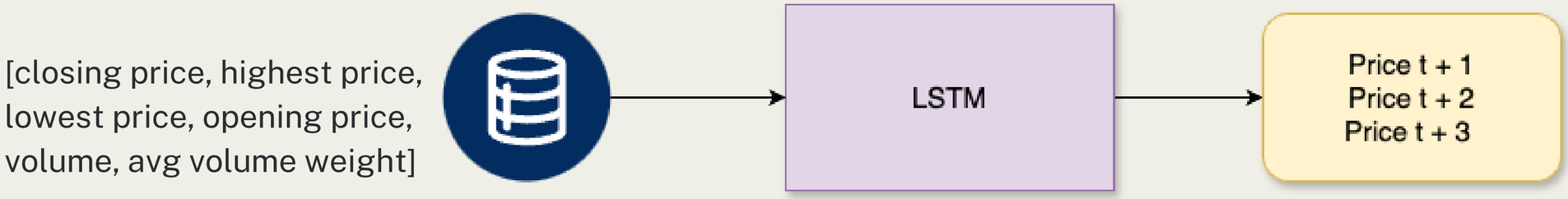
Where:

- E_{AB}, E_{BC}, and E_{CA} are the exchange rates for trading between the currencies A-B, B-C, and C-A, respectively.

Opportunity Condition:

- **Profitable Arbitrage**: When $\text{Arbitrage Value} > 1$, indicating the cycle through A, B, and C yields more of Currency A than started with.
- **No Arbitrage**: When $\text{Arbitrage Value} = 1$, the trades break even.
- **Loss**: When $\text{Arbitrage Value} < 1$, indicating a loss if the trades are executed.

PRICE PREDICTION

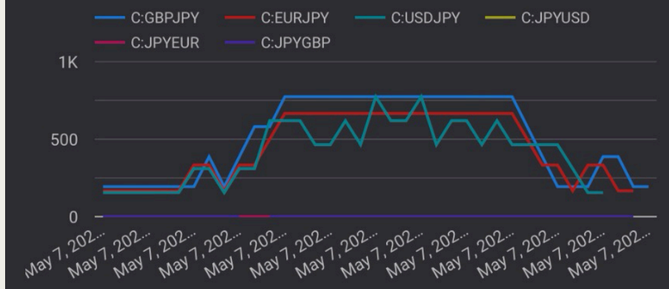


- Multivariate-Multistep LSTM model for Forex price prediction.
- Data Processing with Spark.
- Model Design and Training.
- Prediction.
- Hosting the Models.

DASHBOARD

REAL TIME FOREX ARBITRAGE & PRICE PREDICTION

Current Closing Price



Predicted Price Next 3 Hours

	ticker	prediction_time ▲	price_1	price_2	price_3
1.	C:USDEUR	May 8, 2024, 7:55:13 A...	0.914437651...	0.9143341779...	0.91438043...
2.	C:USDJPY	May 8, 2024, 8:02:48 A...	143.8854217...	143.90371704...	143.910018...
3.	C:USDJPY	May 8, 2024, 6:58:55 PM	143.8854217...	143.90371704...	143.910018...
4.	C:USDEUR	May 8, 2024, 6:59:18 PM	0.914437651...	0.9143341779...	0.91438043...
5.	C:USDGBP	May 8, 2024, 7:01:20 PM	0.788037121...	0.7879160642...	0.78806102...

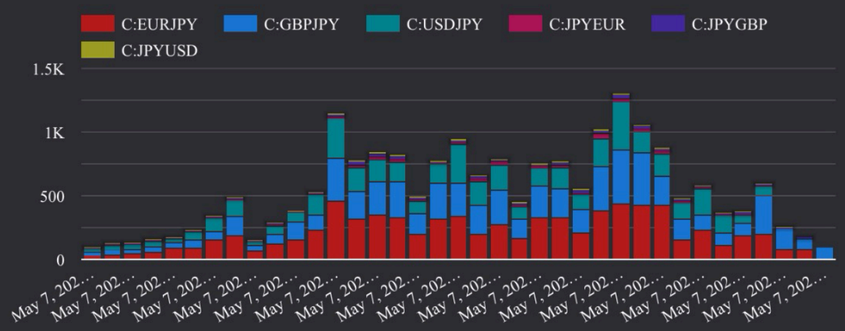
Arbitrage Opportunities

	arbitrage_value ▼	ticker_1	ticker_2	ticker_3
1.	0.9998041994007698	GBP	EUR	USD
2.	0.9997045859851658	GBP	JPY	USD
3.	0.9994804044739886	JPY	EUR	USD

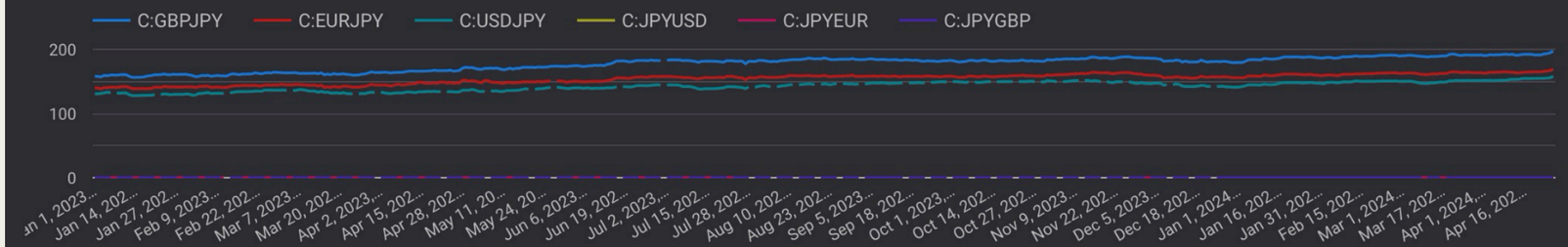
1 - 3 / 3

Current Volume Of Trades

1 - 5 / 5



Daily Closing Price



FUTURE WORK

- Expand Project Scope with Additional Ticker Values
- Integrate Kafka for Data Stream Processing
- Integrate a similar Dashboard for Stocks

Thank you!

NEW BUSINESS OPPORTUNITY

PRANSHU GOYAL
VIBHOR MECHU
SUPRATEEK CHATTERJEE