MAANG Stock Prices Analysis

Cloud Final Report

Group 8

Pooja Ramesh

Sai Varun Kumar Namburi

Ramesh.po@northeastern.edu

Namburi.sai@northeastern.edu

Submission Date: Dec 8st, 2023

Problem Setting:

The volatile nature of stock prices and the various factors that affect the market behavior create a confounding environment for analysts, traders, scientists, and investors. The challenge lies in maneuvering the intricacies of the stock market, especially the technology sector like the SV companies, including MAANG (Meta, Amazon, Apple, Netflix, Google) hold great importance. The technology sector is known for its rapid growth and vulnerability to external factors like the advancements in technology, global economy, geopolitical events, etc. and understanding the impact on dynamic stock price is essential to analyze the investment decisions, set and assess investor expectation and risk involved.

Problem definition:

In rapidly evolving financial market, staying ahead of the trend analysis is highly essential for making informed investment decisions. Uncover and analyzing the long-term and short-term trends in stock prices of each MAANG company is highly essential for trend analysis and help greatly influence the investment decisions. We also aim to explore the external factors that impact the stock price and the level of risk associated with their securities.

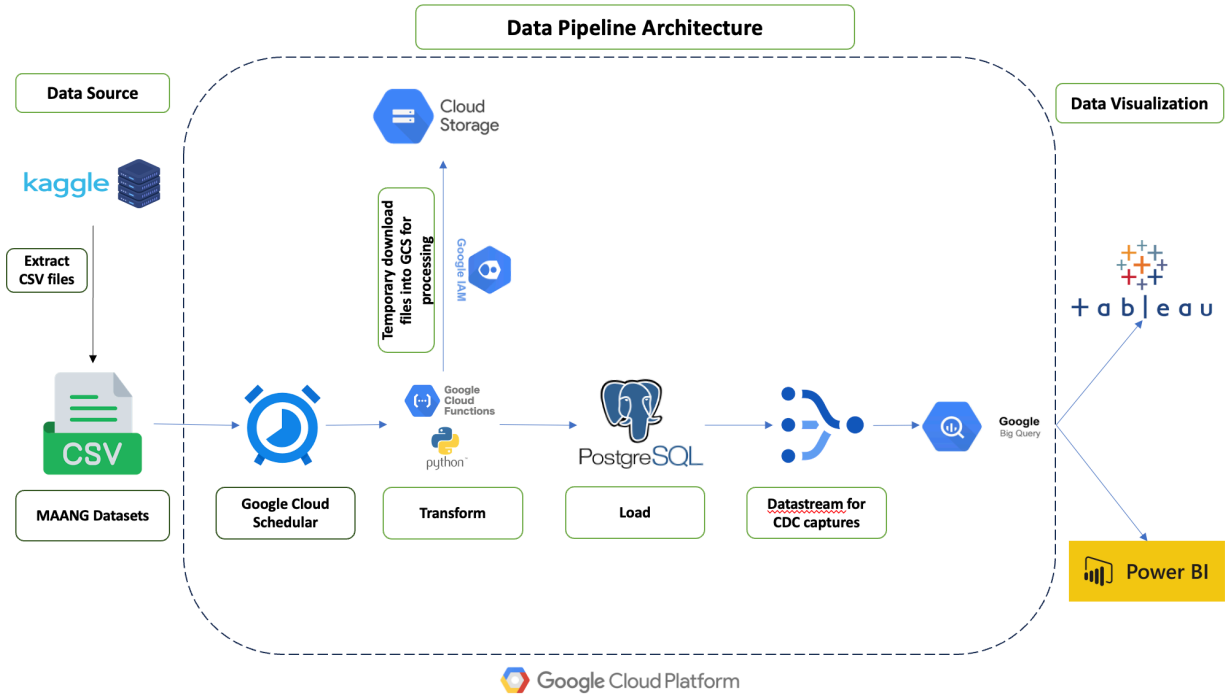
Objective:

This project seeks to develop a data driven solution to address the complexities of analyzing the historical stock prices of MAANG companies and identify the market conditions and external factors that could help make informed investment decisions. With this, the project aims to contribute to a broader understanding of MAANG stock prices, providing investors with niche perspectives of market dynamics and risks involved.

End-Goal:

The ultimate end goal of this project is to ensure we get valuable insights to empower and help investors, analysts and traders make informed decisions. The project aims to predict the effectiveness of trading strategies by back tracing with historical data insights and provide actionable insights to make data driven decisions about the trends, opportunities, and volatility. Furthermore, by developing an efficient cloud-based data engineering solution which can handle large volumes of data for real time analysis.

Data Pipeline Architecture:



The MAANG stocks dataset from Kaggle is updated on a daily, weekly, and monthly basis. Here our aim is to retrieve the data daily and transform it to the database (Postgres) using the Kaggle API. After this, we read the data in cloud functions using python and this is triggered by Goggle cloud scheduler based on the requirements. As the cloud function is serverless, we need a temporary storage which can be used to hold the extracted files in google cloud storage. Then we transform the data (data types, renaming columns, adding columns) and then files are ingested to three tables based on daily, weekly, and monthly into the Postgres DB. Then we use DataStream to capture for the changes (Auto incremental load) and move it into BigQuery.

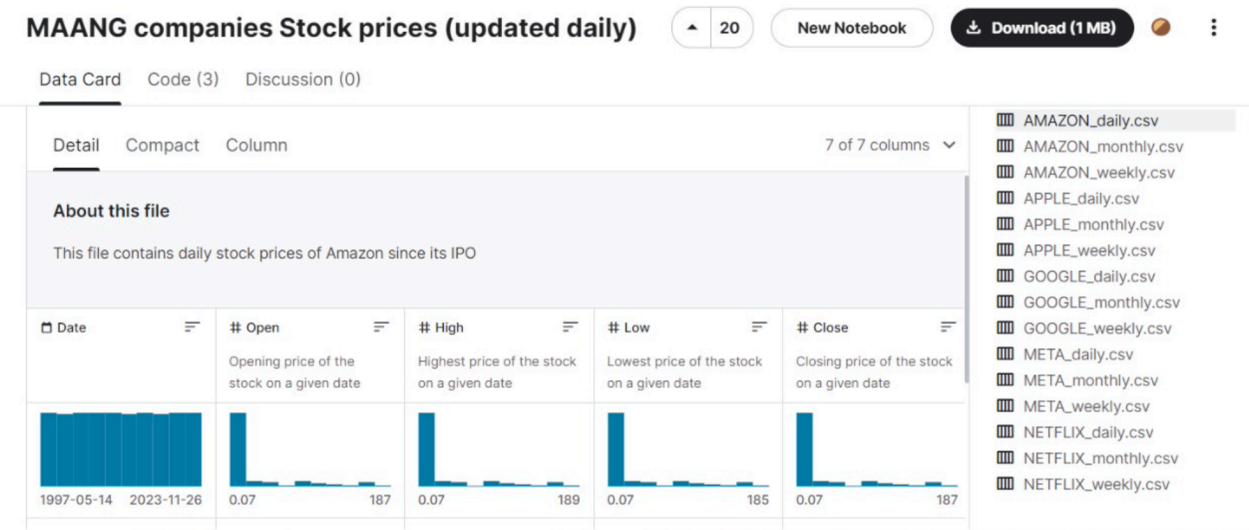
Data source:

This dataset includes the historical data of stock prices for MAANG companies. It contains the daily, weekly, and monthly stock prices for each company, and they automatically updated. Source is Yahoo finance. This dataset has 15 datasets, and each file has around 7 columns. Here we have daily, monthly, and weekly data for each of the companies from year 2000 -2023. We aim to create a combined analysis of stock prices on the daily, weekly, monthly, and quarterly basis which will help predict the trends. The features for each company include:

- Date of the stock price
- Opening prices of the company stock
- Highest trading price
- Lowest trading price

- Closing price
- Adjusted closing price
- Total number shares

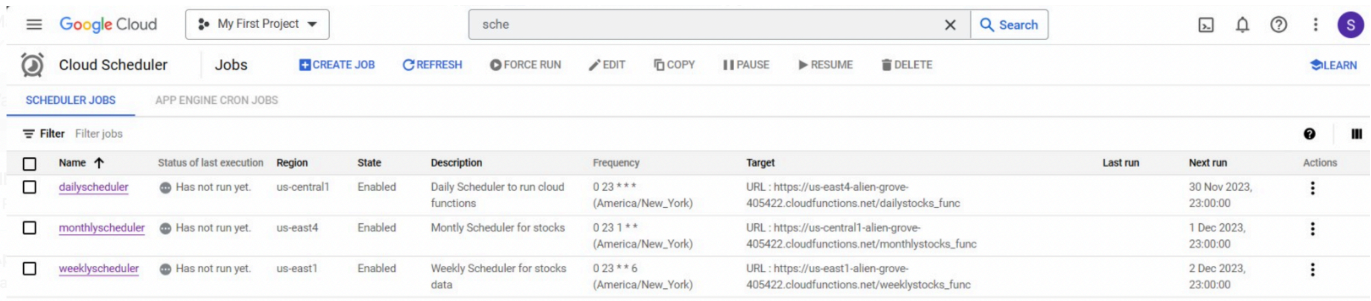
https://www.kaggle.com/datasets/nikhil1e9/netflix-stock-price



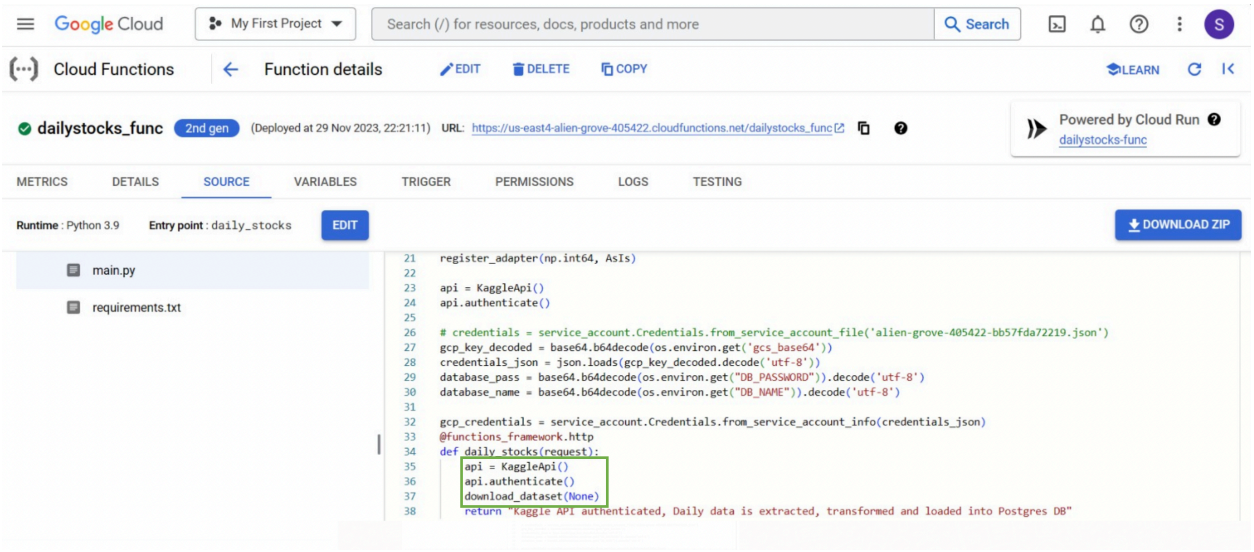
Data Ingestion:

We have adopted py scripts for daily, weekly, and monthly to transform and insert the 15 files into cloud postgres. Here, we used the Kaggle API instead of downloading the 15 files into the local or cloud bucket, as the data is very dynamic and changes daily.

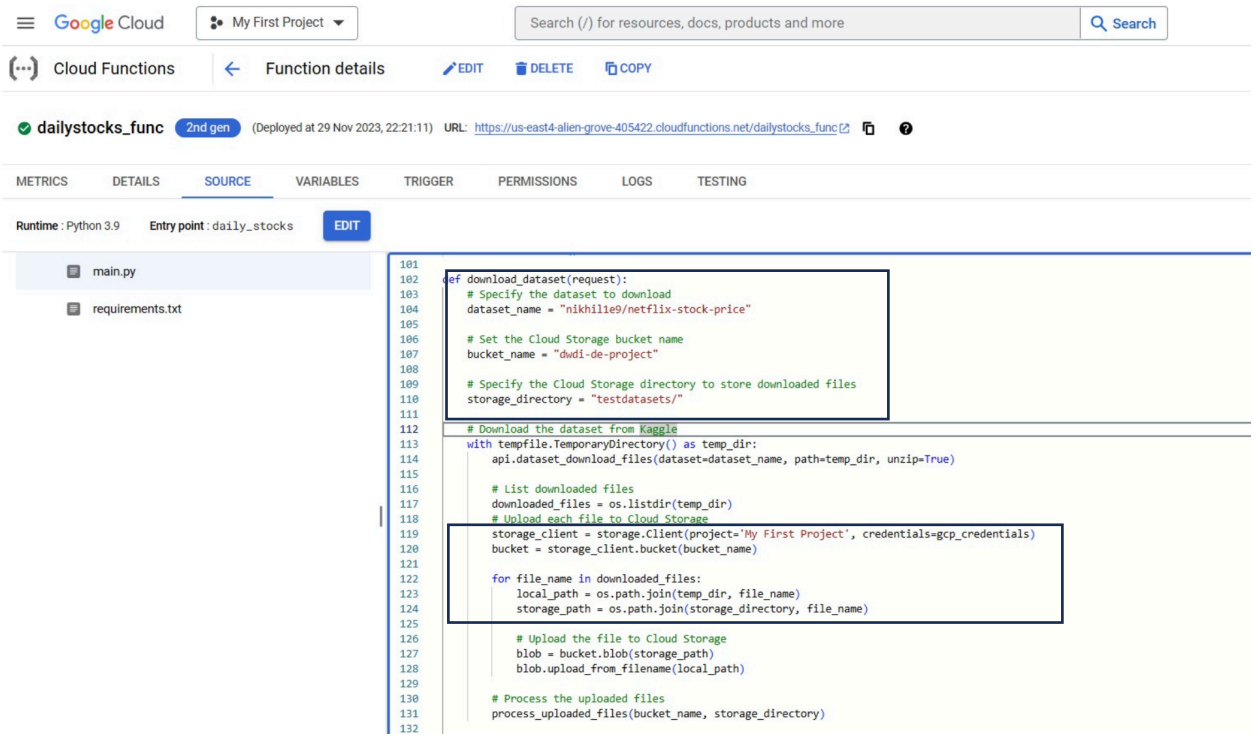
We are cloud scheduler to trigger the cloud functions.



Authentication of Kaggle API by providing credentials:

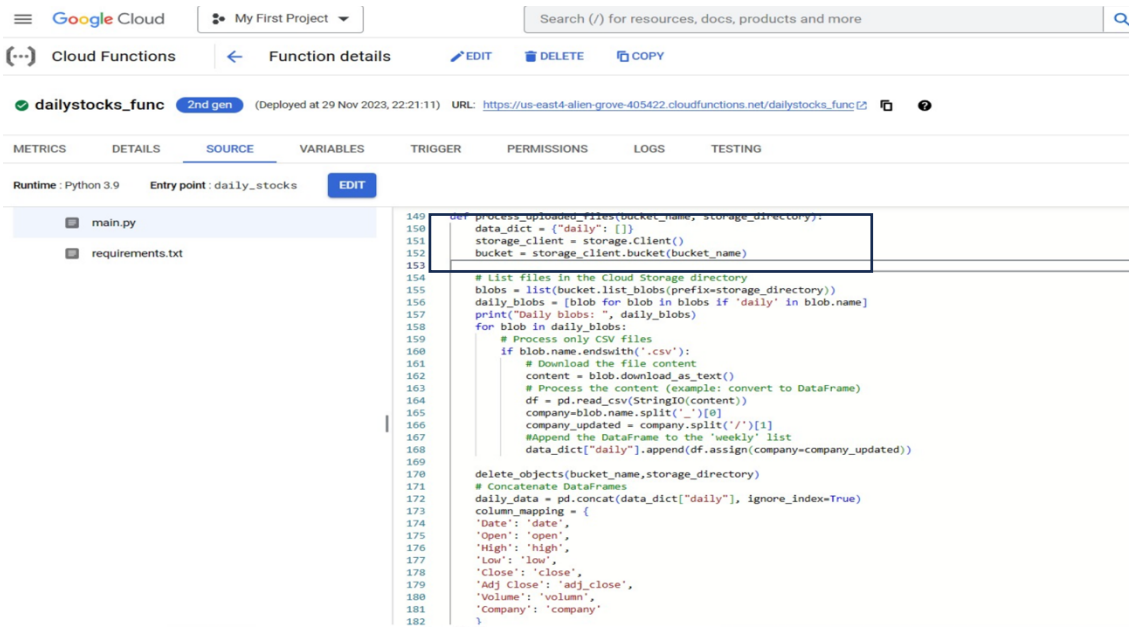


Providing the dataset link to download from Kaggle. Furthermore, the details of temporary storage bucket are also provided. It automatically authenticates using GCP default credentials to get the bucket access:

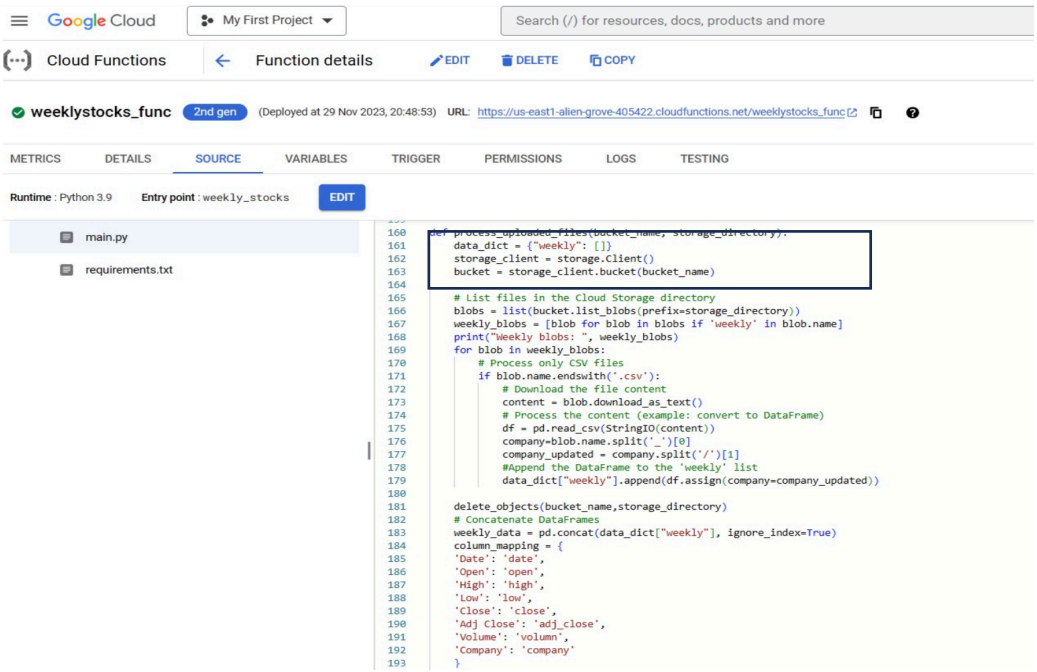


Now we call the process to upload files into cloud storage by providing the temporary path. Below is the screenshot for Daily, weekly, and monthly. Here the data is read and transformed. Now here we deleted the data stored in the bucket i.e., blobs.

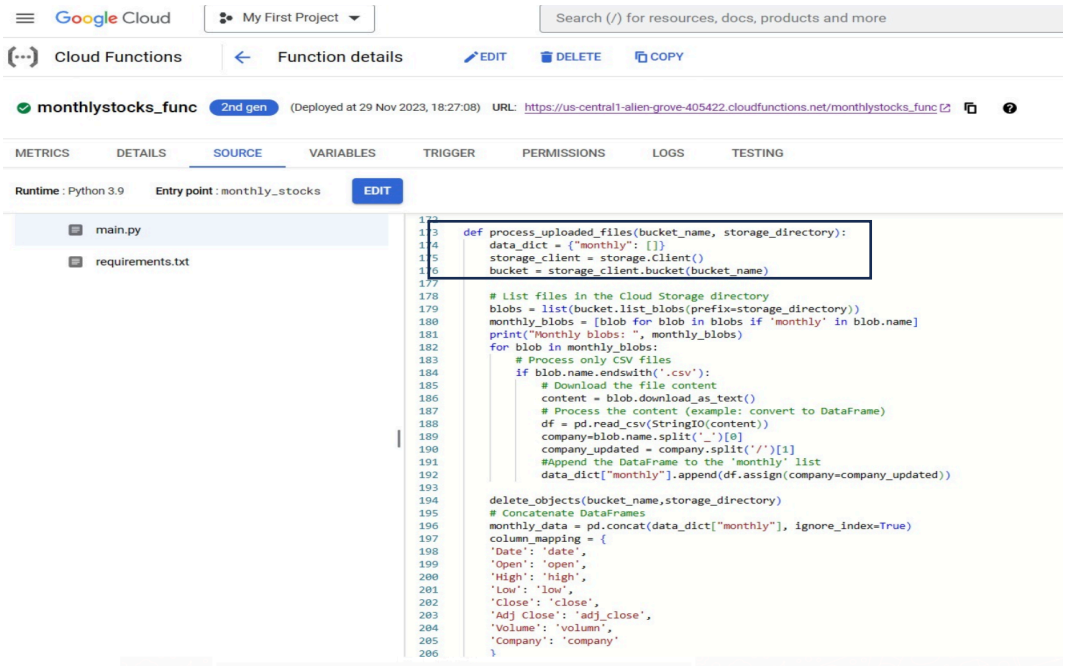
Daily -



Weekly –

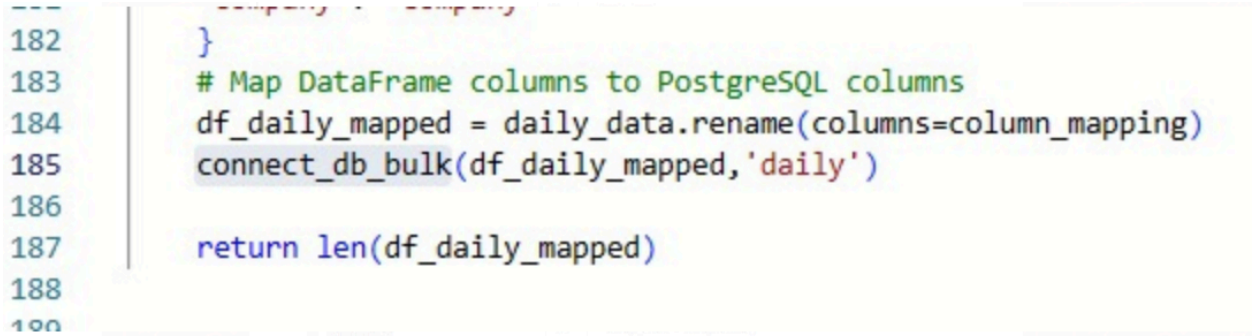


Monthly -

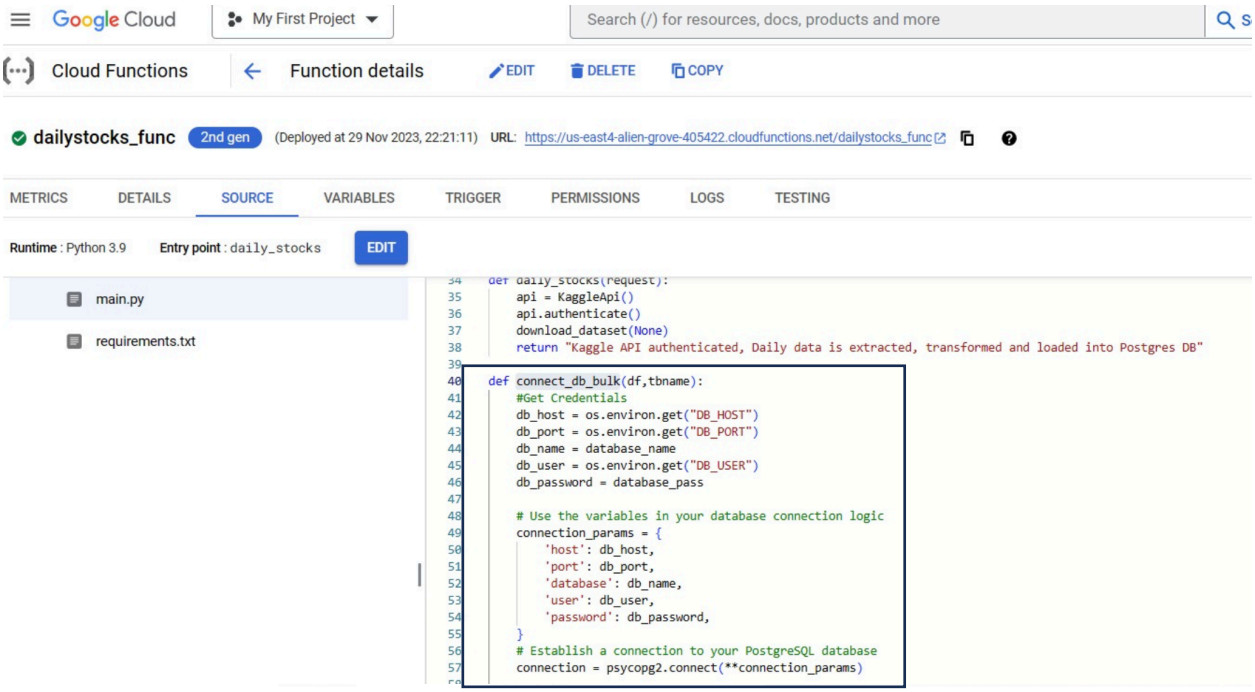


Now, after this the data gets downloaded into the cloud storage bucket for temporary purpose. We read the extracted files using the py script and read the files based on the files name and then read it as dataframes to transform such as renaming columns, adding column (Company name) and modifying the datatypes. Then we check now if the data is already present in PostGres Cloud using the py script. In case it is not present, we insert the data required to avoid data redundancy.

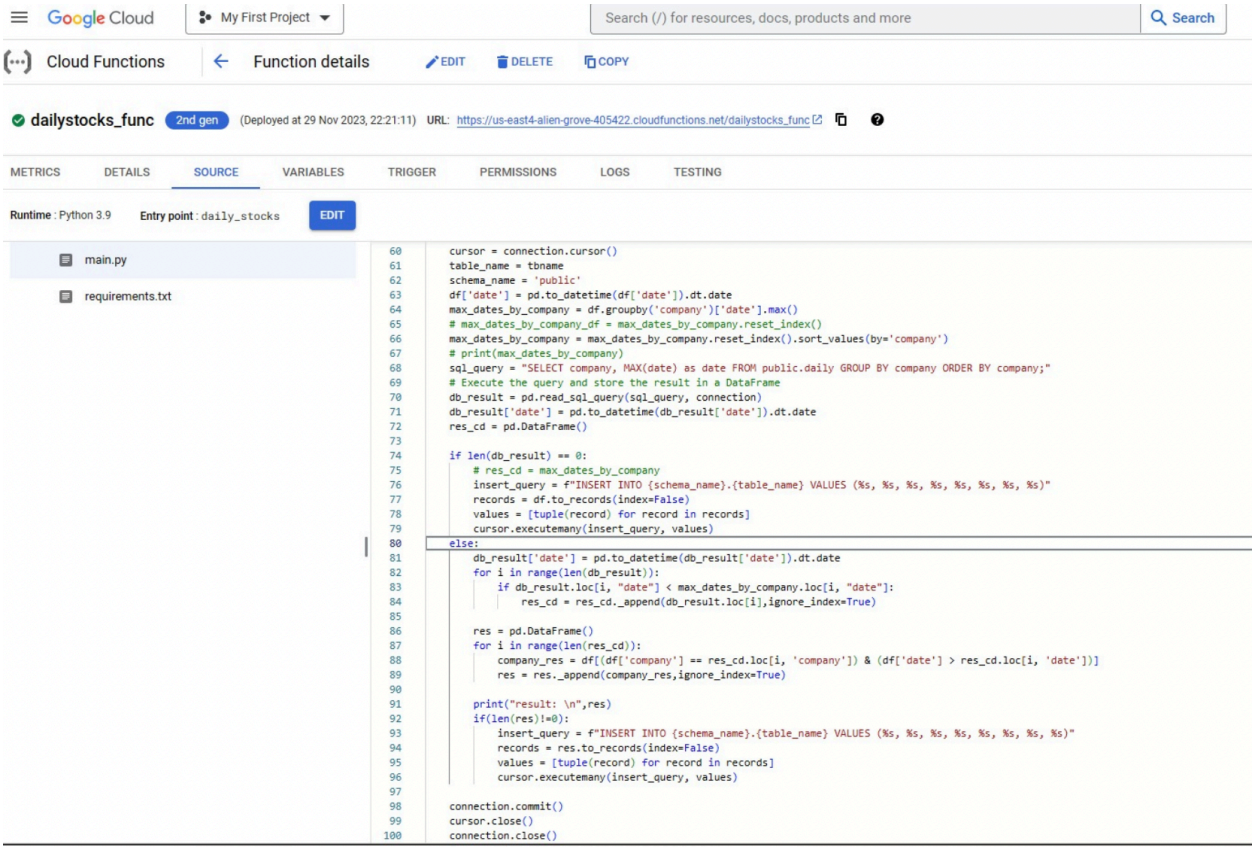
After reading the data, we call the connectDB function to insert it into Cloud Postgres. Now the transformed data is fed into cloud postgres.



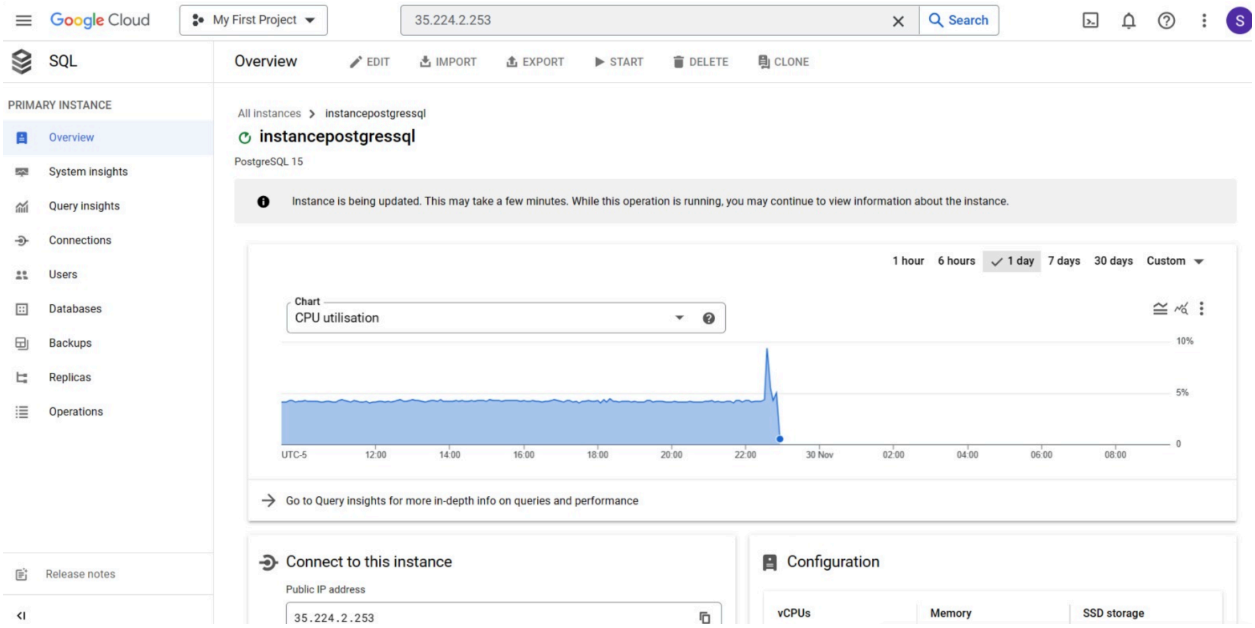
Now we provide the DB connections to connect to postgres DB and the password and DB name has been encoded in Base 64 format to ensure security.



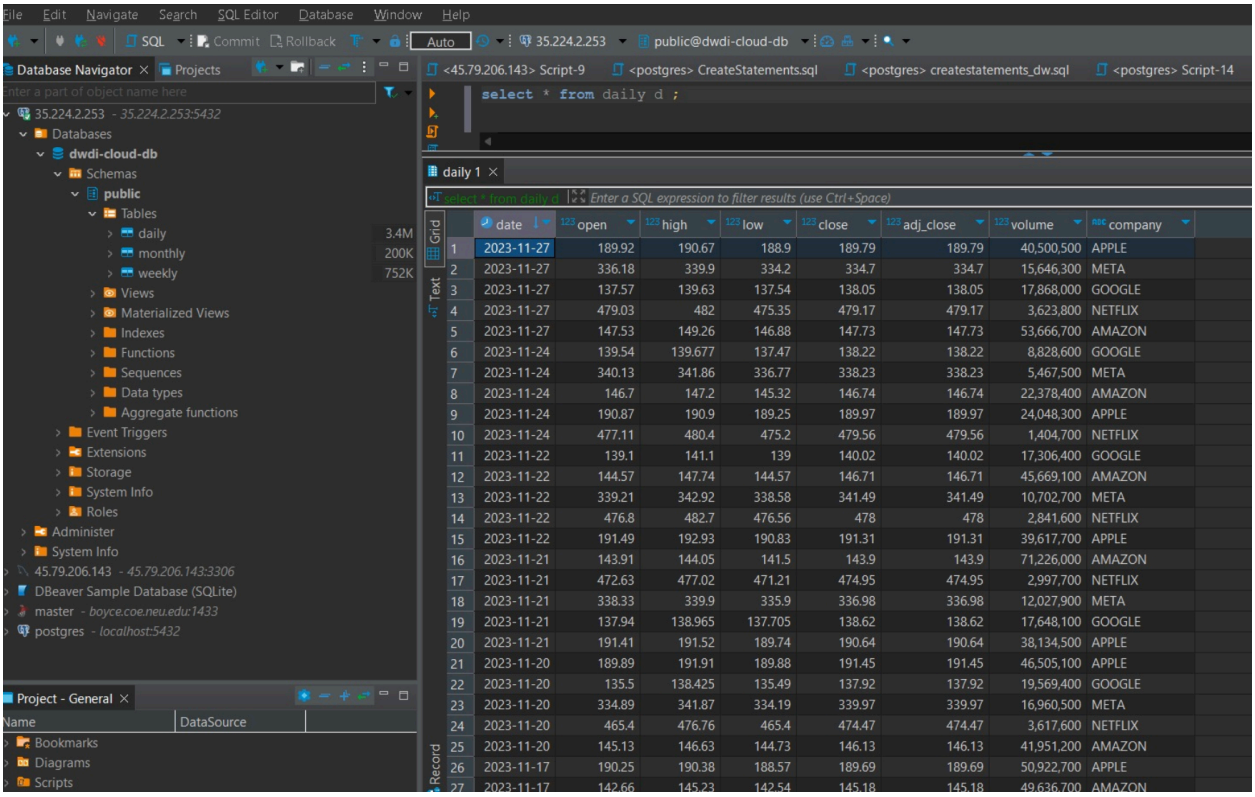
We are now checking of the data is present or not, if not we insert the data in order to avoid data redundancy.



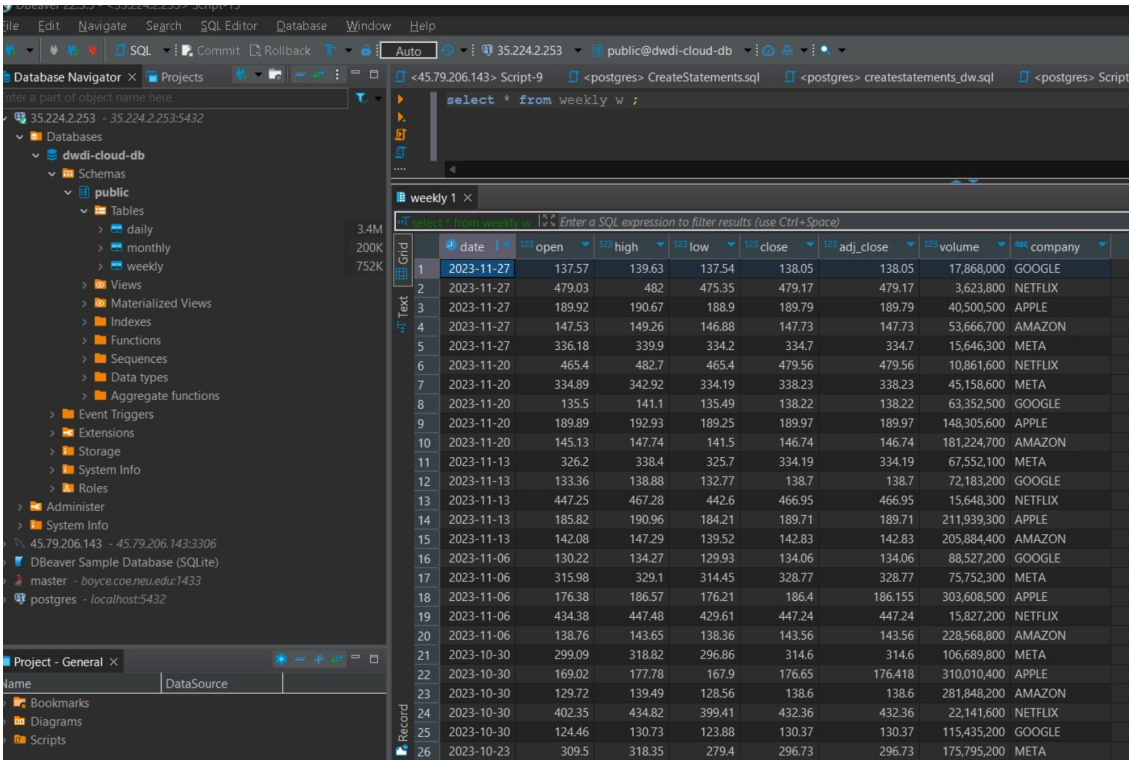
Loading into Postgres:



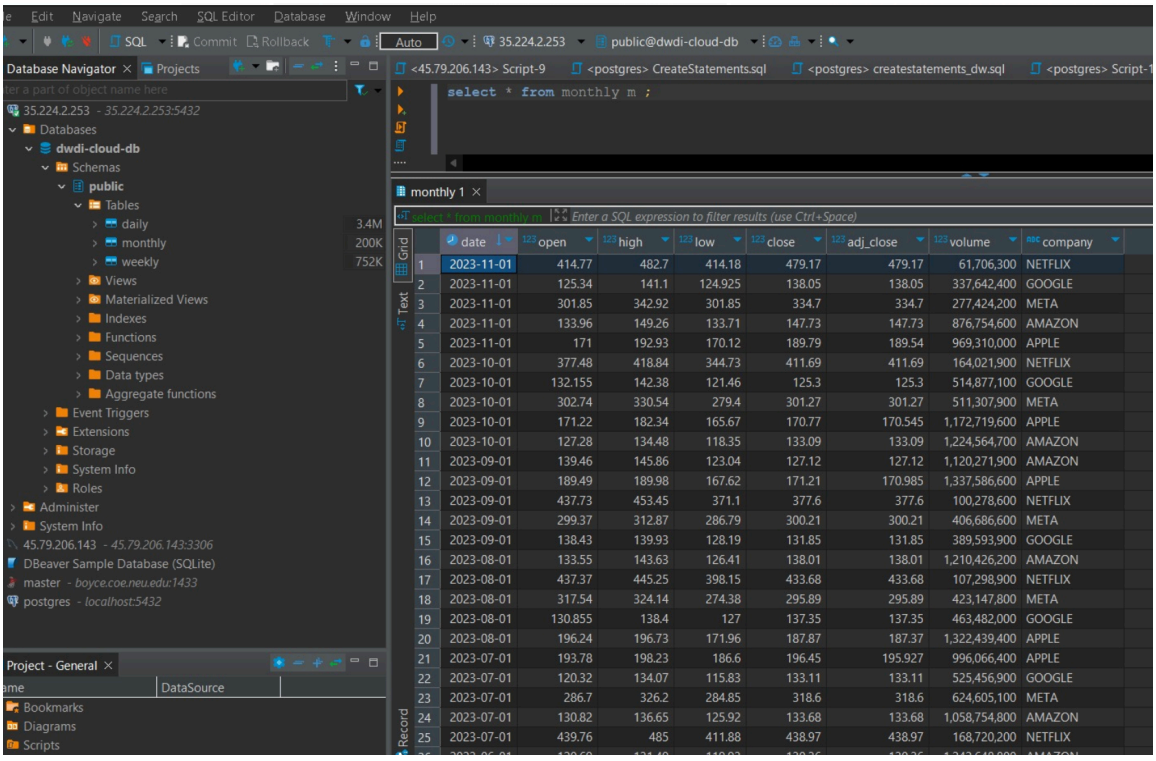
Daily data:



Weekly data:

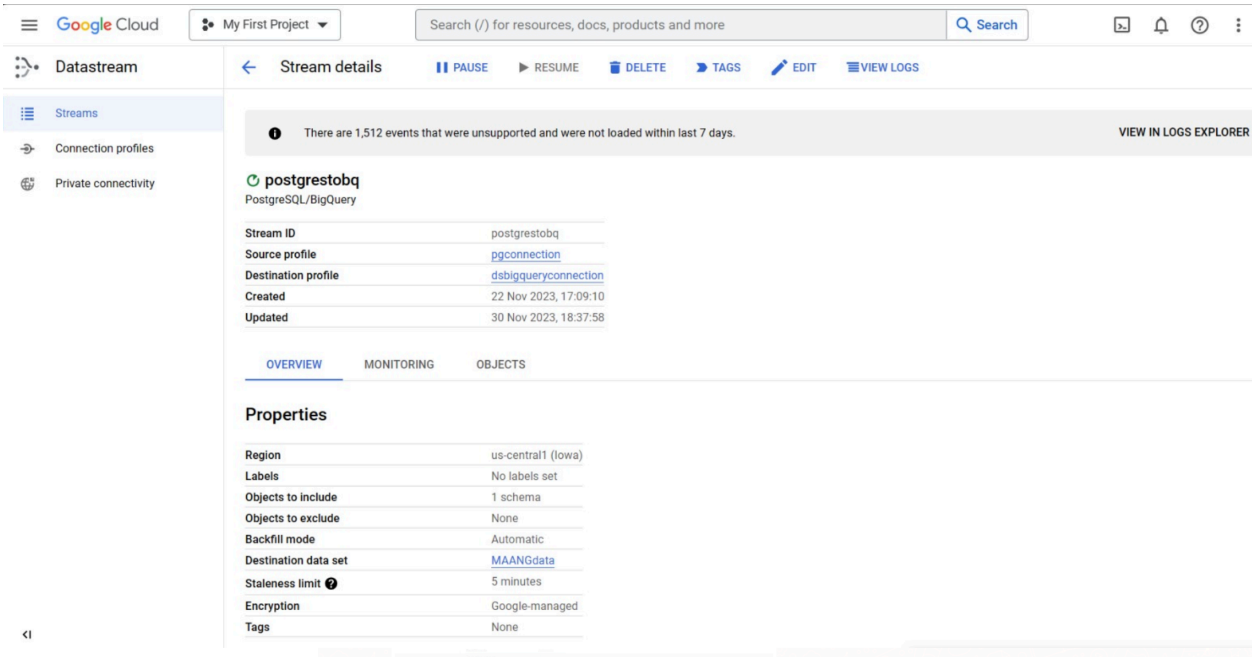


Monthly data:

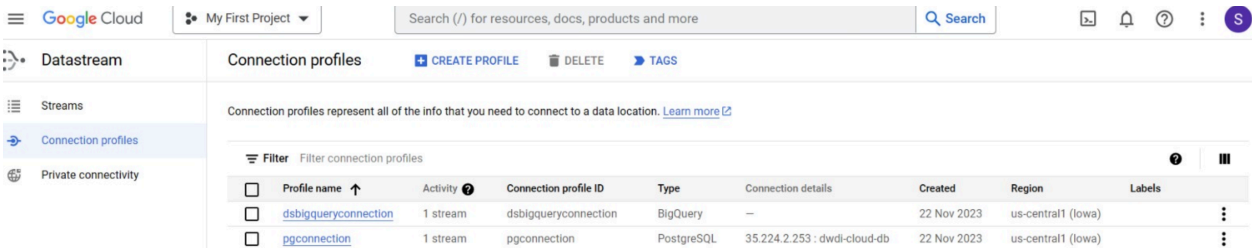


Data Ingestion into BigQuery:

Now the updated data in postgres will be inserted into BigQuery using DataStream.

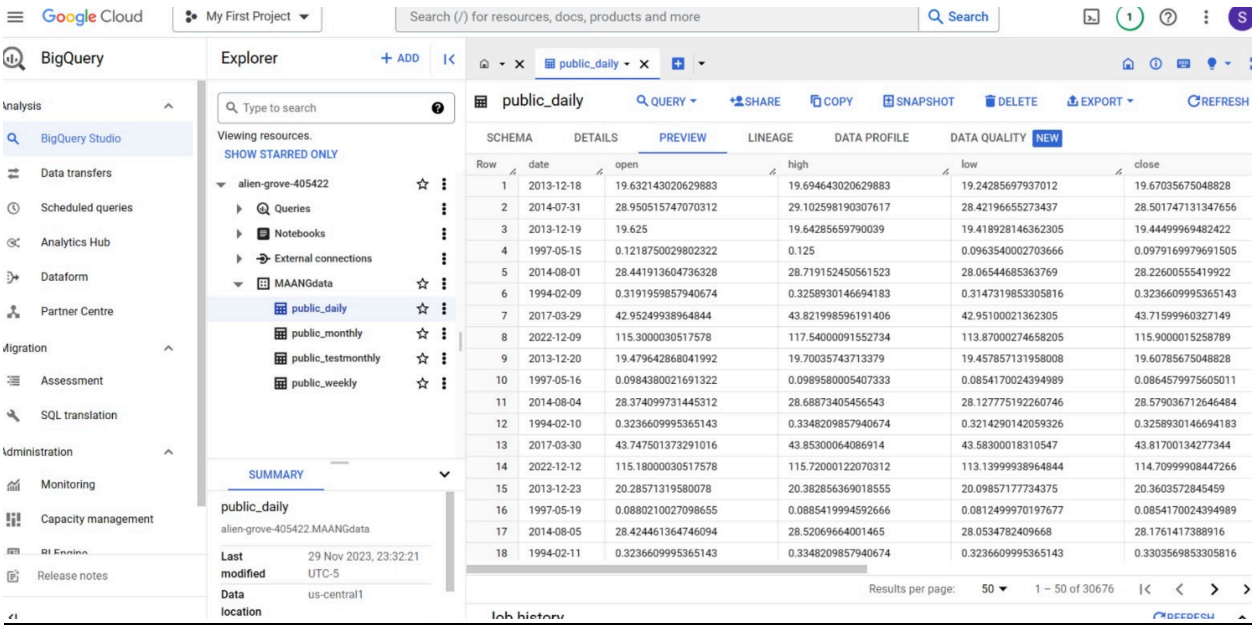


Below is the connection profile in DataStream:

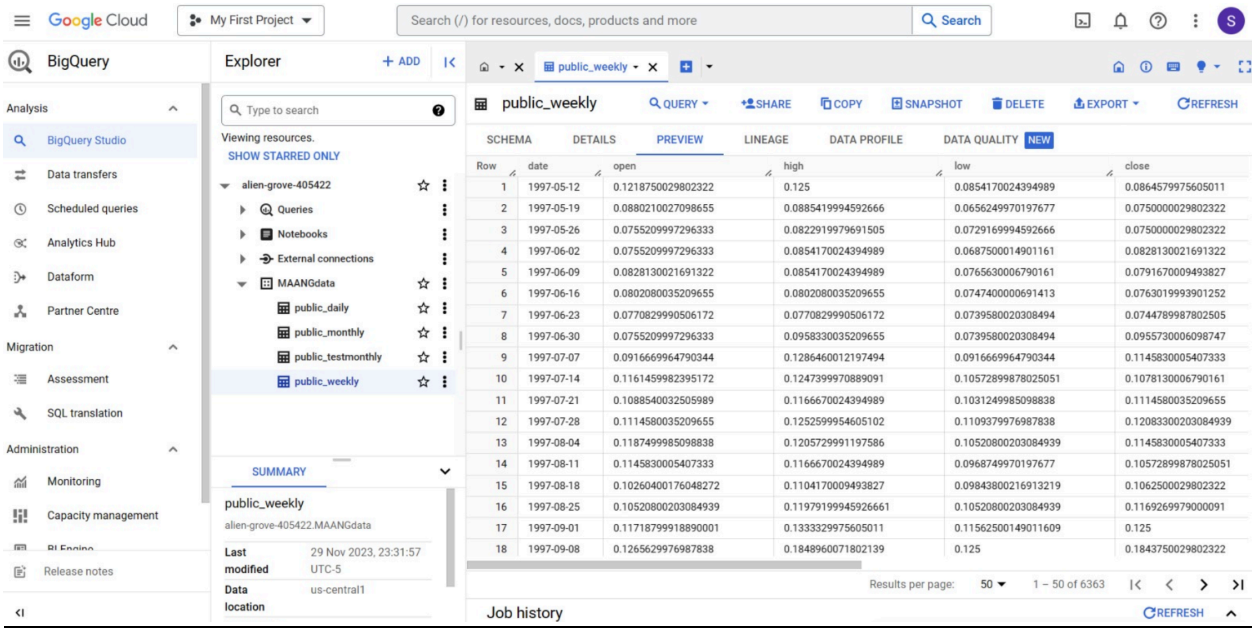


In order to capture the changes in the postgres database, we implemented DataStream to capture the incremental load in data in BigQuery.

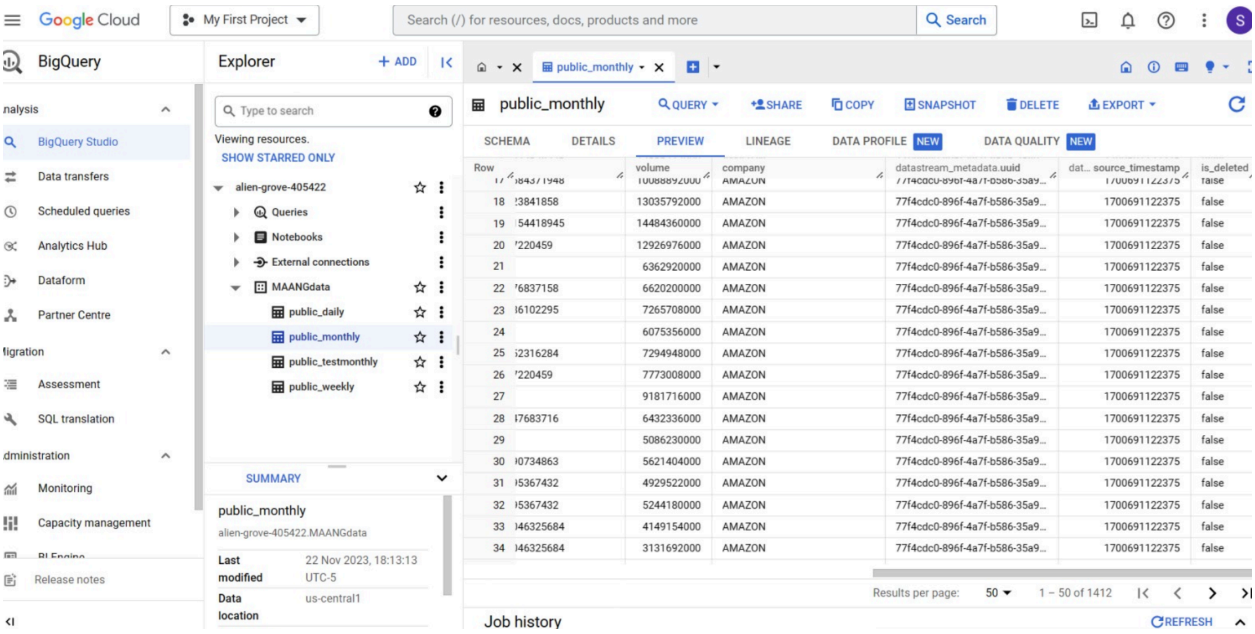
Daily:



Weekly:

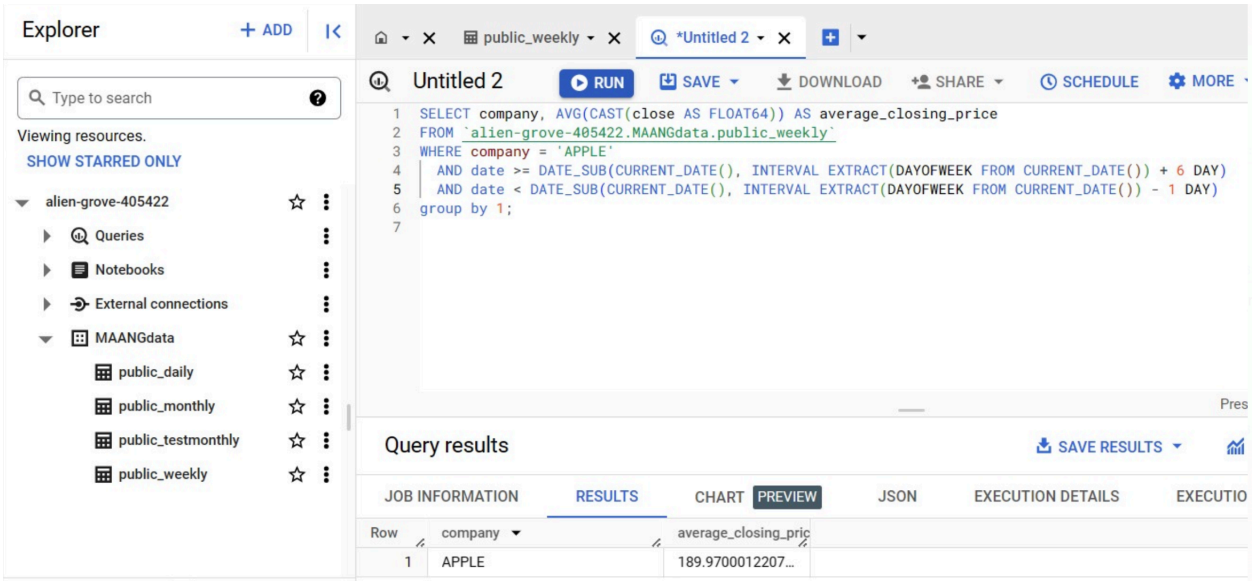


Monthly:



Analytical Queries:

1. Calculate the average closing price for last week for Apple:



2. What is the opening stock price for Meta in the month of October?

```
16
17  SELECT company, open
18  FROM `alien-grove-405422.MAANGdata.public_monthly`
19  WHERE company = 'META'
20    AND EXTRACT(MONTH FROM date) = 10
21    AND EXTRACT(YEAR FROM date) = EXTRACT(YEAR FROM CURRENT_DATE());
22
```

Query results

JOB INFORMATION | RESULTS | CHART PREVIEW | JSON | EXECUTION

Row	company	open
1	META	302.739990234375

3. Which is the lowest trading price and company in the month of November?

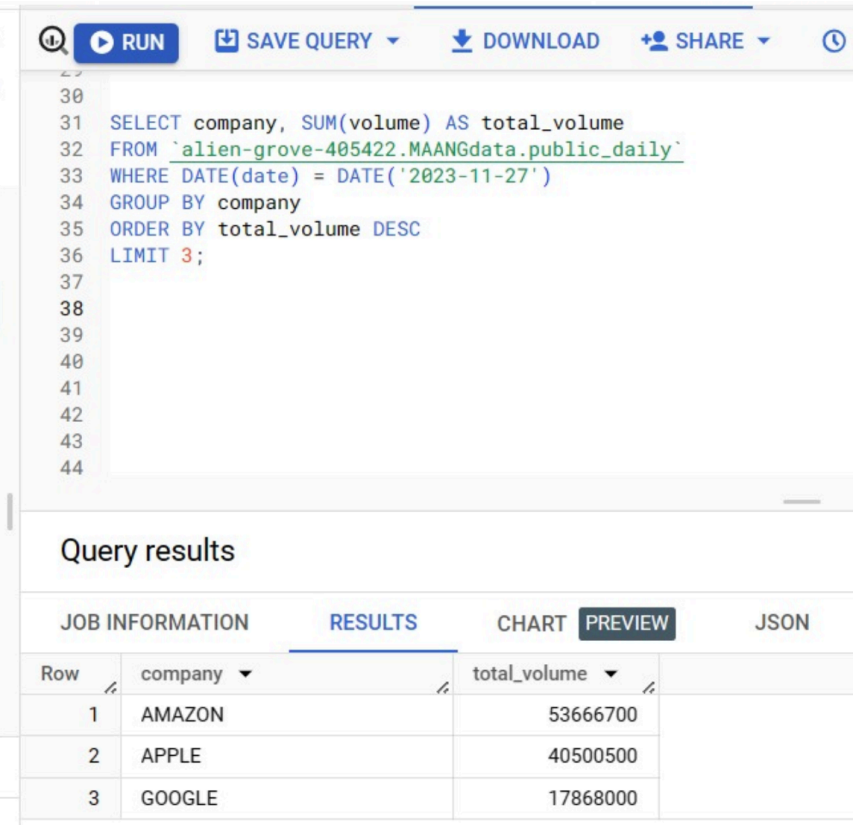
```
14
15  WITH MonthlyLowestPrices AS (
16    SELECT company, MIN(low) AS lowest_trading_price
17    FROM `alien-grove-405422.MAANGdata.public_monthly`
18    WHERE EXTRACT(MONTH FROM date) = 11
19      AND EXTRACT(YEAR FROM date) = EXTRACT(YEAR FROM CURRENT_DATE())
20    GROUP BY company
21  )
22
23  SELECT company, lowest_trading_price
24  FROM MonthlyLowestPrices
25  ORDER BY lowest_trading_price ASC
26  LIMIT 1;
27
28
```

Query results

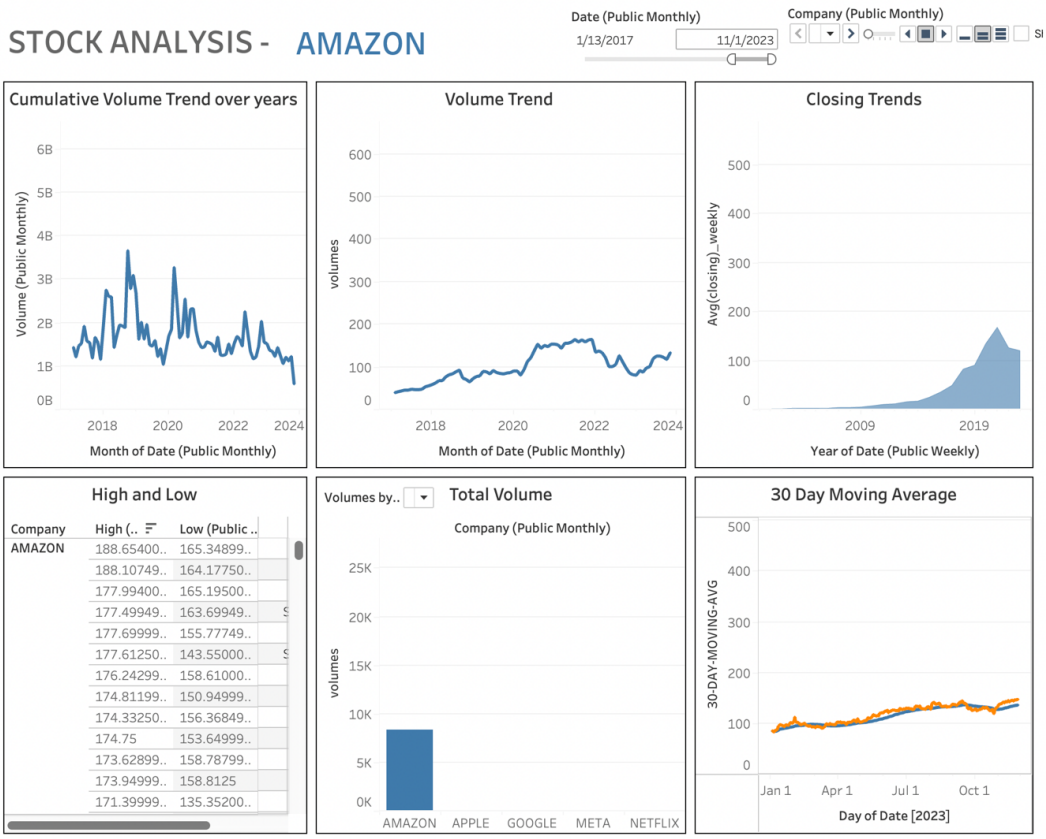
JOB INFORMATION | RESULTS | CHART PREVIEW | JSON | EXECUT

Row	company	lowest_trading_price
1	GOOGLE	124.9250030517578

4. On the 27th of November, what were the top 3 stocks sold and by which company?

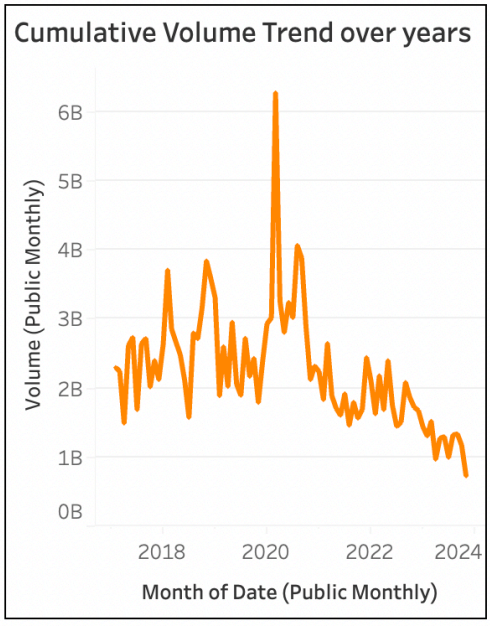


Cloud Dashboard:

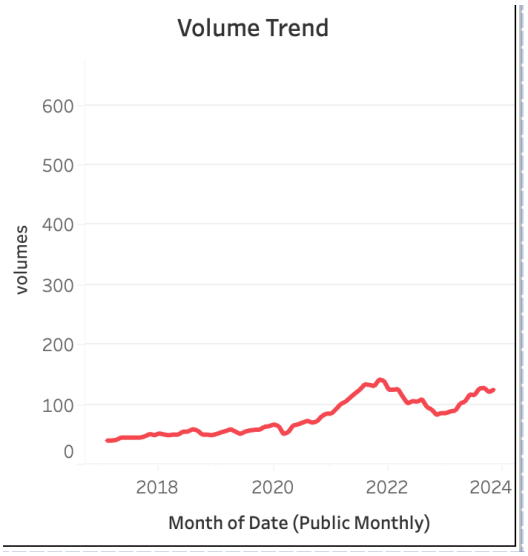


KPI Metrics:

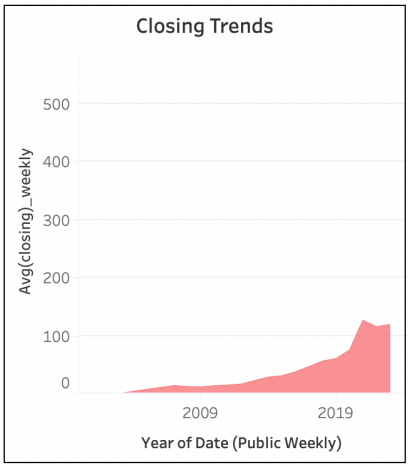
1. **Cumulative volume trend over years** – Determining the trends in monthly trading volume offers insightful information for financial choices. Swing trading techniques might benefit from strategically timing entry and exit opportunities by tracking patterns in volume fluctuations over time. Additionally, investors can make well-informed judgments based on market dynamics when volume spikes are correlated with notable price swings. Knowing the causes of volume spikes makes it easier to predict changes in market sentiment, which enables investors to modify their approaches for the best results in changing financial environments.



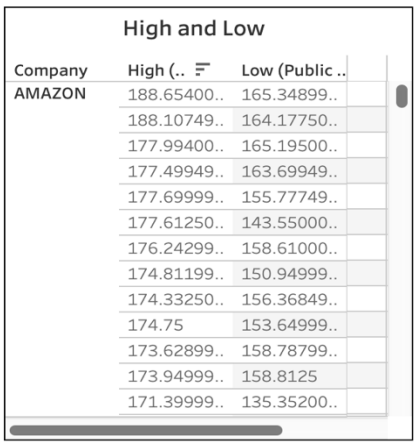
2. **Volume Trend** - The largest monthly trading volume analysis offers important information for making well-informed investing selections. Investors can take advantage of market dynamics by identifying months with unusually high trading activity and modifying portfolio weightings accordingly. Being able to use volume spikes as indicators of possible trend reversals or continuations requires an understanding of the relationship between volume peaks and company news or events. With the help of this all-encompassing strategy, investors may move through the market with accuracy and make well-informed choices based on reliable indicators.



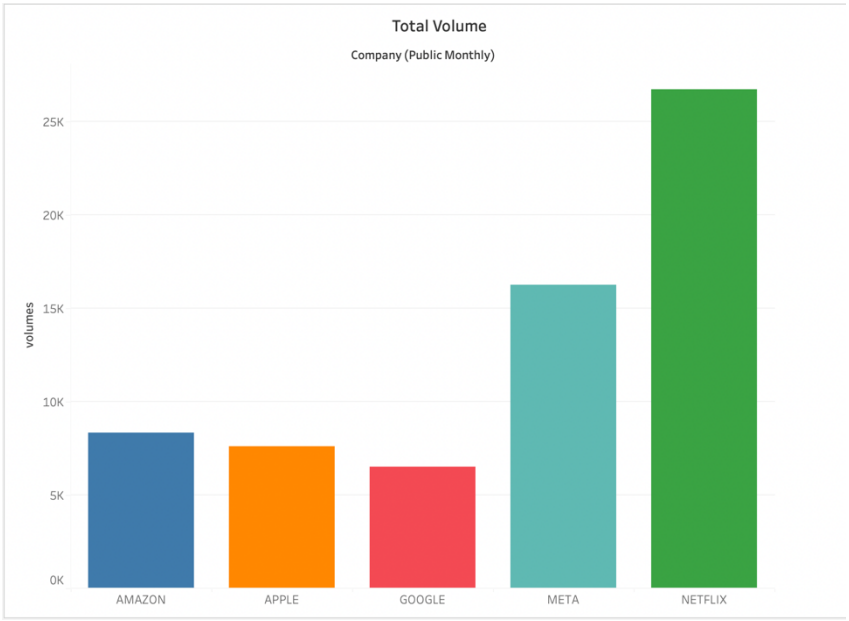
3. **Closing Trends:** Strategic investment decisions may be made with important information from an analysis of the annual closing price patterns. Investors can deploy more capital to stocks exhibiting persistent rising annual trends by identifying the overall annual performance patterns for each MAANG stock. Investors can alter their risk exposure by identifying trends, such as steady growth or volatile periods, based on the closing prices' historical volatility.



4. **High and Low:** Analyzing high and low prices offers important information about each stock's price extremes. By identifying highs and lows, investors may establish target prices for well-timed profit-taking or stop-loss orders. Investors are better equipped to make judgments and maximize their trading methods thanks to this study.



5. **Total Volume:** Trading activity of MAANG stocks may be thoroughly compared by examining the volume per firm. Increased portfolio liquidity can be achieved by investors allocating more capital to equities with greater average trading volumes. Investors can deliberately diversify portfolio weightings based on liquidity concerns by finding firms that exhibit consistently greater or lower volumes. Making wise investment decisions and improving portfolio management are made possible by the insightful information our research offers.



6. **30-day Moving Average:** The 30-day moving average offers important information about stock patterns and possible turning points. Investors using trend-following techniques might utilize crossovers as cues to enter and leave the market. Trading choices are guided by the distance between the stock price and its moving average, which acts as a dynamic support or resistance level. Furthermore, evaluating each MAANG stock's overall momentum enables trading tactics to be strategically adjusted based on the stock's position in relation to its moving average. By maximizing risk and reward, this strategy improves the efficacy of investment choices.