Financial Data Analysis App

(MySQL & Excel)

Yinan Zhou

March 25, 2024

Introduction

The Financial Data Analysis App is a comprehensive tool designed for both individual investors, portfolio manager, financial analysts, and market observers to gain valuable insights into stock and index data. The app's primary use cases include real-time data monitoring, trend analysis, personalized dashboards, and event-driven alerts. Users can monitor prices and volumes, analyze historical trends, create customized dashboards, and set alerts based on specific criteria.

One of the key business challenges of this app is ensuring data accuracy and security, integrating diverse data sources, providing real-time analysis, designing a user-friendly interface, ensuring performance and scalability, complying with regulations, and competing in a crowded market. These challenges require robust data management systems, advanced analytics capabilities, intuitive user interfaces, and strong security measures.

The cost model for this app could be a freemium model, offering basic features for free while charging for advanced functionalities or premium content. Alternatively, it could adopt a subscription-based model with different tiers based on the level of access and features provided. As for our personal connection to this app, as someone interested in finance and investments, we see great value in having a tool that can provide real-time data analysis and insights, helping us make informed investment decisions and stay updated with market trends. This app aligns with our interests and goals in financial management and investment strategies.

Business Analysis

User Persona 1: Individual Investor

This persona represents individual investors who use the app for personal investment activities. Their goal is to monitor real-time prices and trends of specific stocks, indices, and cryptocurrencies. They rely on the app to provide accurate and timely data to make informed investment decisions. They use personalized dashboards to track their favorite financial instruments and set event-driven alerts for price notifications. For them, the app is a valuable tool for staying updated on market fluctuations and identifying potential investment opportunities.

User Persona 2: Portfolio Manager

Portfolio managers use the application to manage diversified investment portfolios across different asset classes. They rely on features like portfolio tracking, risk assessment tools,

performance reporting, asset allocation optimization, and compliance monitoring. This persona values data accuracy, portfolio customization options, risk management functionalities, and integration with external systems for seamless portfolio management.

User Persona 3: Financial Analyst

Financial analysts are another key user group who use the app extensively. Their goal is to conduct in-depth research on market trends, generate reports, and provide data-driven recommendations to clients or organizations. They use the app to access historical price data, perform trend analysis, and extract insights for making strategic investment decisions. Their usage involves running complex queries, creating customized reports, and utilizing advanced analytical tools within the app.

User Persona 4: Market Observer

Market observers are individuals or professionals who need to stay informed about global financial markets. They use the app to access real-time data on various indices, commodities, and currencies. Their goal is to understand market movements, analyze macroeconomic indicators, and gain insights into global economic trends. They rely on the app's comprehensive data coverage and analytical capabilities to stay ahead of market developments and make informed observations.

Business Rules

1. **Financial Instrument Unique Identification:**

   Each record in financial_instrument_table must have a unique Instrument_Name (PK) and categorized by Instrument_Type.

2. **Stock Data Association:**

   Entries in stock_table link to financial instruments in financial_instrument_table via Instrument_Name (FK) and include Date, Price, and Volume.

3. **Cryptocurrency Data Integrity:**

   Cryptocurrency_table records are tied to financial instruments through Instrument_Name (FK) for data integrity with Date, Price, and Volume.

4. **Index Tracking:**

   Index_table entries reference financial instruments via Instrument_Name (FK) for tracking index data with Date, Price, and Volume.

5. **Commodity Price and Volume Tracking:**

   Commodity_table records are linked to financial instruments through Instrument_Name (FK) for monitoring commodity prices with Date, Price, and Volume.

Table Design and Analysis

Data Entities

1. **financial_instrument_table:**

   Attributes: Instrument_Name (PK), Instrument_Type

2. **stock_table:**

   Attributes: Stock_ID (PK), Date (PK), Instrument_Name (FK), Price, Volume

3. **cryptocurrency_table:**

   Attributes: Crypto_ID (PK), Date (PK), Instrument_Name (FK), Price, Volume

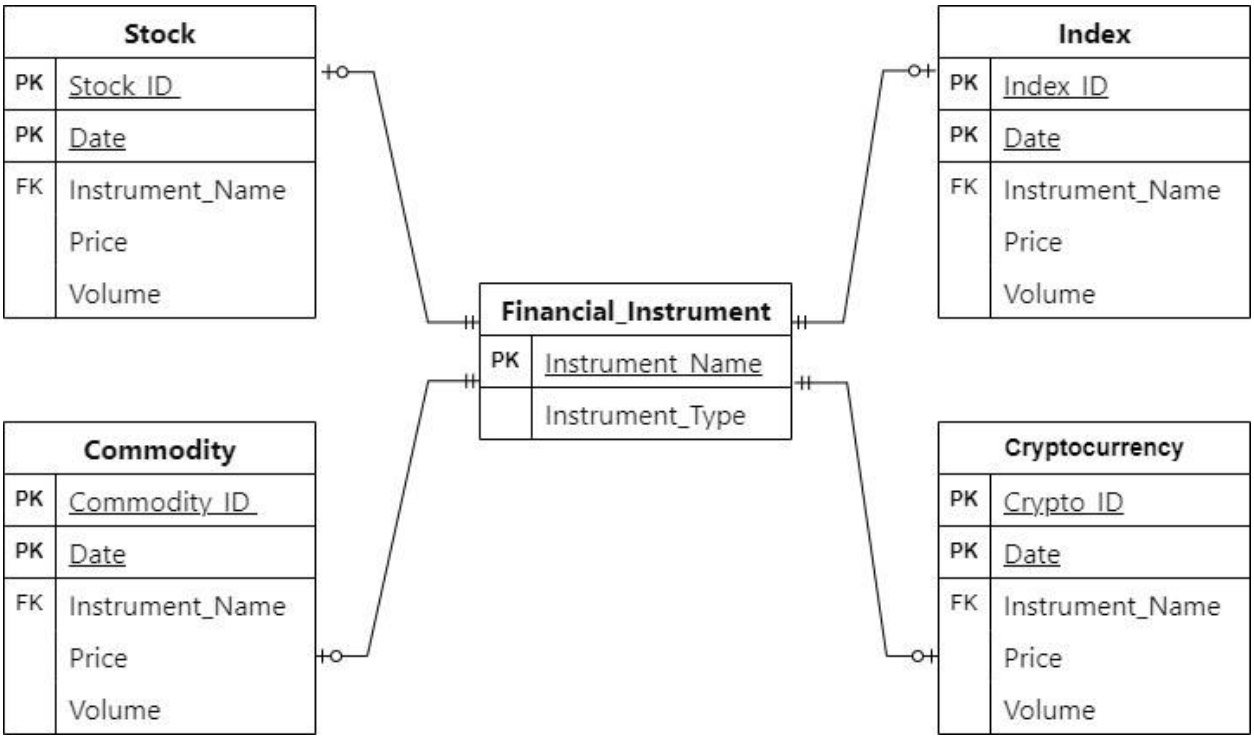
4. **index_table:**

   Attributes: Index_ID (PK), Date (PK), Instrument_Name (FK), Price, Volume

5. **commodity_table:**

   Attributes: Commodity_ID (PK), Date (PK), Instrument_Name (FK), Price, Volume

ER Diagram



Database Implementation and Report

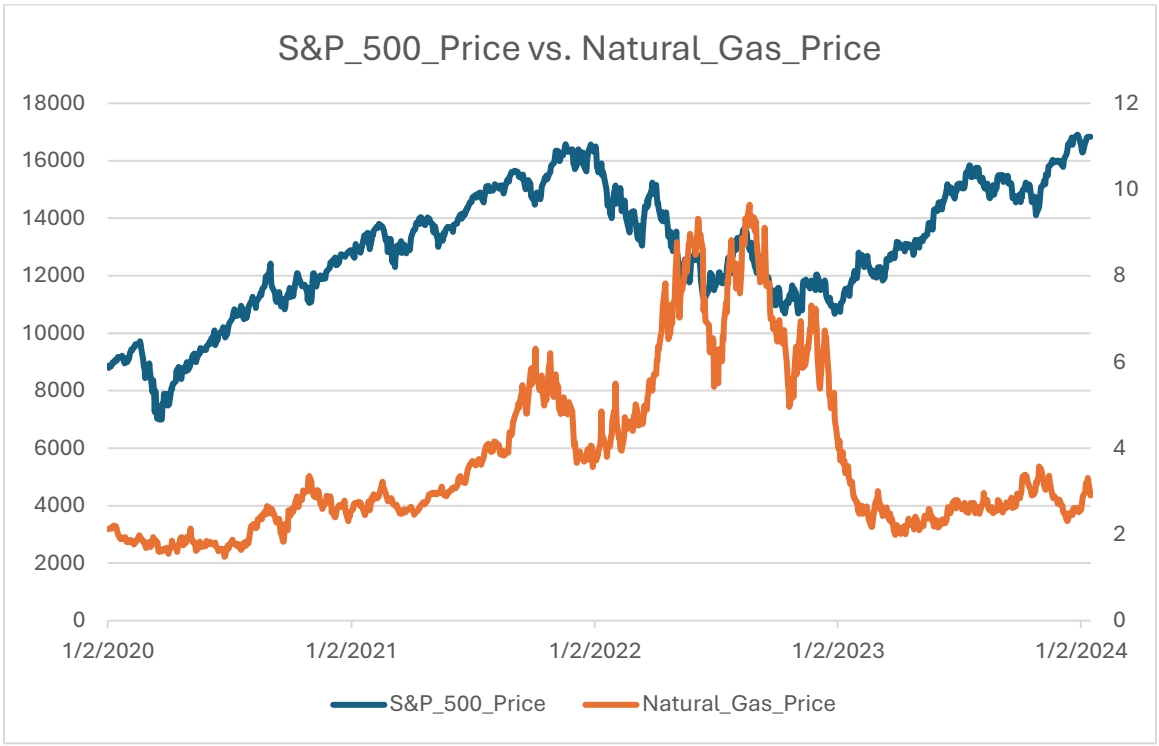
User registration is often the starting point for many applications. In this project, we assume that users provide interests in indices, commodities, cryptocurrencies, stocks, etc., during registration.

Query 1: Union Query for Index and Commodity Data:

This query demonstrates how to retrieve index and commodity data using a union query and return only relevant information for specific indices and commodities.

```
SELECT
i.Date, i.Index_ID, i.Instrument_Name AS Index_Name, i.Price AS Index_Price, i.Volume AS Index_Volume,
c.Commodity_ID, c.Instrument_Name AS Commodity_Name, c.Price AS Commodity_Price, c.Volume AS Commodity_Volume
FROM index_table i
JOIN commodity_table c
ON i.Date = c.Date
WHERE i.Index_ID = 'Index_1' AND c.Commodity_ID = 'Commodity_1';
```

Date	Index_ID	Index_Name	Index_Price	Index_Volume	Commodity_ID	Commodity_Name	Commodity_Price	Commodity_Volume
2020-01-02	Index_1	S&P_500	8872.22	152650000	Commodity_1	Natural_Gas	2.122	164570
2020-01-03	Index_1	S&P_500	8793.90	144750000	Commodity_1	Natural_Gas	2.130	144670
2020-01-06	Index_1	S&P_500	8848.51	142380000	Commodity_1	Natural_Gas	2.135	154680
2020-01-07	Index_1	S&P_500	8846.45	151390000	Commodity_1	Natural_Gas	2.162	163010
2020-01-08	Index_1	S&P_500	8912.37	164570000	Commodity_1	Natural_Gas	2.141	242560
2020-01-09	Index_1	S&P_500	8989.63	150730000	Commodity_1	Natural_Gas	2.166	230030
2020-01-10	Index_1	S&P_500	8966.64	134200000	Commodity_1	Natural_Gas	2.202	282650
2020-01-13	Index_1	S&P_500	9070.65	140340000	Commodity_1	Natural_Gas	2.182	232950
2020-01-14	Index_1	S&P_500	9033.42	154920000	Commodity_1	Natural_Gas	2.187	281850
2020-01-15	Index_1	S&P_500	9035.67	136690000	Commodity_1	Natural_Gas	2.120	218920
2020-01-16	Index_1	S&P_500	9125.00	143140000	Commodity_1	Natural_Gas	2.077	206420



The above line graph shows the prices of S&P 500 and natural gas from January 2, 2020, to January 26, 2024. By observing the data, the following conclusions can be drawn:

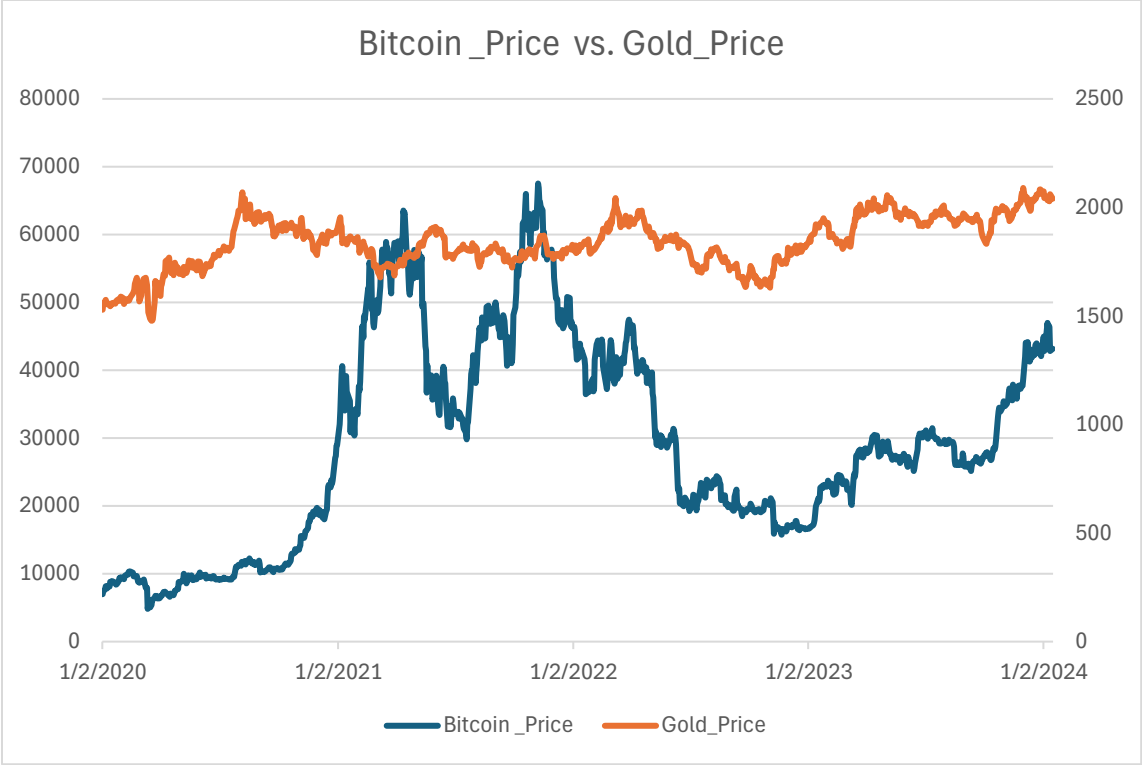
1. **S&P 500 Price Trends:** The S&P 500 Price shows fluctuating trends over the given period, with periods of growth, decline, and stability, indicating market volatility. In early 2020, a substantial decline was observed in late February and March, likely attributed to the impact of the COVID-19 pandemic on the market. After the initial drop, there was a recovery period, with the S&P 500 price gradually increasing again.
2. **Natural Gas Price Trends:** The Natural Gas Price also exhibits variability, showing periods of both increase and decrease. There are periods of significant price spikes and drops, indicating volatility in the natural gas market.
3. **Correlation:** There doesn't seem to be a clear direct correlation between the S&P 500 Price and Natural Gas Price based solely on the provided data. Further analysis and statistical techniques would be needed to determine if there's any significant correlation between these two variables.

Query 2: Union Query for Cryptocurrency and Commodity Data:

This query demonstrates how to perform a union query for cryptocurrency and commodity data, returning only relevant information for specific cryptocurrencies and commodities.

```
SELECT
r.Date, r.Crypto_ID, r.Instrument_Name AS Crypto_Name, r.Price AS Crypto_Price, r.Volume AS Crypto_Volume,
c.Commodity_ID, c.Instrument_Name AS Commodity_Name, c.Price AS Commodity_Price, c.Volume AS Commodity_Volume
FROM cryptocurrency_table r
JOIN commodity_table c
ON r.Date = c.Date
WHERE r.Crypto_ID = 'Crypto_1' AND c.Commodity_ID = 'Commodity_3';
```

Date	Crypto_ID	Crypto_Name	Crypto_Price	Crypto_Volume	Commodity_ID	Commodity_Name	Commodity_Price	Commodity_Volume
2020-01-02	Crypto_1	Bitcoin	6967.00	632780	Commodity_3	Gold	1528.100	270550
2020-01-03	Crypto_1	Bitcoin	7343.10	936290	Commodity_3	Gold	1552.400	436740
2020-01-06	Crypto_1	Bitcoin	7759.10	786750	Commodity_3	Gold	1568.800	558970
2020-01-07	Crypto_1	Bitcoin	8155.70	1010000	Commodity_3	Gold	1574.300	435870
2020-01-08	Crypto_1	Bitcoin	8059.60	1190000	Commodity_3	Gold	1560.200	813410
2020-01-09	Crypto_1	Bitcoin	7842.40	786500	Commodity_3	Gold	1554.300	372880
2020-01-10	Crypto_1	Bitcoin	8187.10	1050000	Commodity_3	Gold	1560.100	344340
2020-01-13	Crypto_1	Bitcoin	8111.40	610660	Commodity_3	Gold	1550.600	311730
2020-01-14	Crypto_1	Bitcoin	8829.20	1480000	Commodity_3	Gold	1544.600	340910
2020-01-15	Crypto_1	Bitcoin	8818.30	1170000	Commodity_3	Gold	1554.000	310070
2020-01-16	Crypto_1	Bitcoin	8726.90	748320	Commodity_3	Gold	1550.500	260050



The above line graph shows the prices of Bitcoin and gold from January 2, 2020, to January 26, 2024. By observing the data, the following conclusions can be drawn:

1. **Bitcoin Price Fluctuations**: The Bitcoin price shows significant fluctuations over time, ranging from lows around 4800 to highs exceeding 59000. This indicates high volatility in the cryptocurrency market during the period covered by the dataset.
2. **Gold Price Stability**: In contrast, the gold price remains relatively stable compared to Bitcoin. It fluctuates within a narrower range, with lows around 1475 and highs around 2100. Gold is traditionally considered a more stable investment compared to cryptocurrencies.
3. **Correlation Trends**: There are periods where Bitcoin and Gold prices seem to move in opposite directions, indicating a negative correlation. For example, when Bitcoin prices rise sharply, gold prices may dip, and vice versa. However, this correlation is not consistent throughout the entire dataset.

Query 3: Calculate the Average Price of Natural Gas for the Year 2023:

This query calculates the average price of natural gas for the year 2023, providing insights into market trends.

```sql
SELECT AVG(Price) AS Average_Price
FROM commodity_table
WHERE Instrument_Name = 'Natural_Gas'
AND Date BETWEEN '2023-01-01' AND '2023-12-31';
```

	Average_Price
▶	2.6710880

From the outcome, we can determine that the average price of natural gas for the year 2023 is approximately 2.67.

Query 4: Find the Day with the Highest Volume of Crude Oil Traded:

This query identifies the day with the highest trading volume of crude oil, which can be crucial information for investors and traders.

```sql
SELECT Date, Volume AS Highest_Volume
FROM commodity_table
WHERE Volume = (
    SELECT MAX(Volume)
    FROM commodity_table
    WHERE Instrument_Name = 'Crude_Oil'
);
```

	Date	Highest_Volume
▶	2020-03-09	1770000

From the outcome, we can identify March 9, 2020, as the day with the highest trading volume of crude oil, reaching 177000.

Query 5: Find the Stock with the Largest Percentage Change in Price from the Beginning to the End of the Dataset:

This query identifies the stock with the largest percentage change in price from the beginning to the end of the dataset, providing an important reference for investors.

```sql
SELECT Stock_ID, Instrument_Name AS Stock_Name, ((MAX(Price) - MIN(Price)) / MIN(Price)) * 100 AS Percentage_Change
FROM stock_table
GROUP BY Stock_ID, Instrument_Name
ORDER BY Percentage_Change DESC
LIMIT 1;
```

	Stock_ID	Stock_Name	Percentage_Change
▶	Stock_2	Tesla	1602.533223

Based on the results, it can be concluded that Tesla had the highest percentage change in price from January 1, 2020, to January 26, 2024, with an approximate increase of 1602.63%.

Project Wrap-up and Future Considerations

1. **Data Management:** Managing data across multiple tables and establishing relationships is crucial for database integrity and efficient data retrieval.
2. **Business Rules:** Defining clear business rules helps maintain consistency and accuracy in data handling and processing.
3. **User Personas:** Understanding different user personas and their specific needs is essential for designing user-friendly and effective applications.
4. **Application Design:** Designing applications that cater to diverse user requirements, from individual investors to financial analysts, requires careful planning and consideration of functionality.
5. **Continuous Learning:** Staying updated with industry trends, user feedback, and technological advancements is crucial for ongoing improvement and innovation in projects.

Reference

https://www.kaggle.com/datasets/dhavalpatel555/us-stock-market-2020-to-2024?resource=download