Introduction

Exploring the dynamic realm of commodities trading, particularly in gold and silver futures, unveils intricate patterns and trends within daily market data. Motivated by a personal drive to enhance my investment strategies, this project delves into the trends and patterns of daily data in gold and silver markets, aiming to extract valuable insights that I can directly apply to make informed and strategic investment decisions in precious metals. This project aims to uncover valuable insights relevant to traders seeking strategic advantages, academic researchers delving into historical behavior, or businesses managing risk in the precious metals industry. By focusing on gold and silver, the analysis promises a comprehensive understanding of the specific dynamics within these markets.

Data

Primary Data Source:

The primary source of data for this project is a comprehensive dataset obtained from Kaggle, a reputable platform for sharing and discovering datasets. This dataset comprises 28,215 observations and 8 variables. This set includes daily futures data for precious metals, specifically Gold, Silver, Platinum, Copper, and Palladium ("all_commodititesdata.csv"). The variables in the dataset are as follows:

Date: The date the data was recorded, formatted as YYYY-MM-DD.

Open: Market opening price.

High: Highest price during the trading day.

Low: Lowest price during the trading day.

Close: Market closing price.

Volume: Number of contracts traded during the day.

Ticker: Market quotation symbol for the future.

Commodity: Name of the precious metal the future refers to.

Gold, Silver, & Precious Metals Futures Daily Data

Additional Data Source (Yahoo Finance):

For the additional data source, I used Yahoo Finance because it can provide detailed historical data for stocks and commodities, including daily price movements, opening/closing prices, and trading volumes. Specifically, I downloaded the adjusted closes for both gold and silver in 2023.

Yahoo Finance (Gold)

Yahoo Finance (Silver)

Data Merging Process:

In this data merging process, the objective was to amalgamate three datasets pertaining to gold and silver markets: 'gold_silver_2023_data.csv', 'GC=F.csv' (Gold data), and 'SI=F.csv' (Silver data). The merging operation was executed using the "date" column as a key identifier. Initially, the script installed and loaded the essential libraries—'dplyr' for data manipulation and 'readr' for reading CSV files. The working directory was then set to the location of the data files. Subsequently, the data files were read into R using the 'read_csv' and 'read.csv' functions. To ensure uniformity in column names, the "Date" column in both the gold and silver datasets was renamed to "date."

The merging process employed the 'left_join' function from the 'dplyr' package, executed twice: first with the gold data and then with the silver data. The "date" column served as the key identifier, and suffixes "_gold" and "_SI" were appended to the columns from the gold and silver datasets, respectively. The resulting dataset, now containing comprehensive information from all three sources, was printed to the console using the 'print' function. Finally, the combined dataset was saved as a CSV file named 'combined_gold_silver_data.csv' using the 'write_csv' function. This unified dataset provides a foundation for further analysis, allowing for a holistic examination of trends and patterns within the gold and silver markets.

Data Dictionary:

Field Name	**Data Type**	**Description**
Date	Date	The date the data was recorded (Format: YYYY-MMDD)
Open	Numeric	Market opening price
High	Numeric	Highest price during the trading day
Low	Numeric	Market closing price
Close	Numeric	Market closing price
Volume	Numeric	Number of contracts traded during the day
Ticker	Text	Market quotation symbol for the future
Commodity	Text	Name of the precious metal the future refers to (gold, silver, platinum, copper, palladium)
Adj Close_gold	Numeric	The adjusted closing of gold, reflecting any corporate actions such as dividends, on the given date.
Adj Close_SI	Numeric	The adjusted closing of silver, reflecting any corporate actions such as dividends, on the given date.

Analysis:

Research Question #1: What is the correlation between gold and silver?

Reasoning for question:

Understanding the correlation between gold and silver prices is essential for investors and analysts as these precious metals often exhibit similar market behavior. It can provide insights into potential investment strategies, risk management, and broader economic trends.

Hypothesis:

When there is a positive correlation between gold and silver prices. This means that when one metal's price increases, the other metal's price tends to increase as well.

Data Used:

The dataset combined_gold_silver_data.csv is utilized.

The dataset gold_silver_correlation_data was created in the process.

The relevant variables are date, Adj.Close_gold, and Adj.Close_SI.

Analysis Process and Method(s):

Data Preprocessing:

Reading the dataset and converting column names to snake case for consistency.

Ensuring the date column is in the Date format.

Data Wrangling:

Extracting the required variables (date, Adj.Close_gold, Adj.Close_Sl).

Calculating the daily percentage change in closing prices for both gold and silver.

Visualization 1: Line Plot:

Plotting a line chart to visualize the daily percentage change over time.

Using two lines, one for gold and one for silver, to compare their trends.

Coloring the lines to distinguish between the two commodities.

Visualization 2: Density Plot:

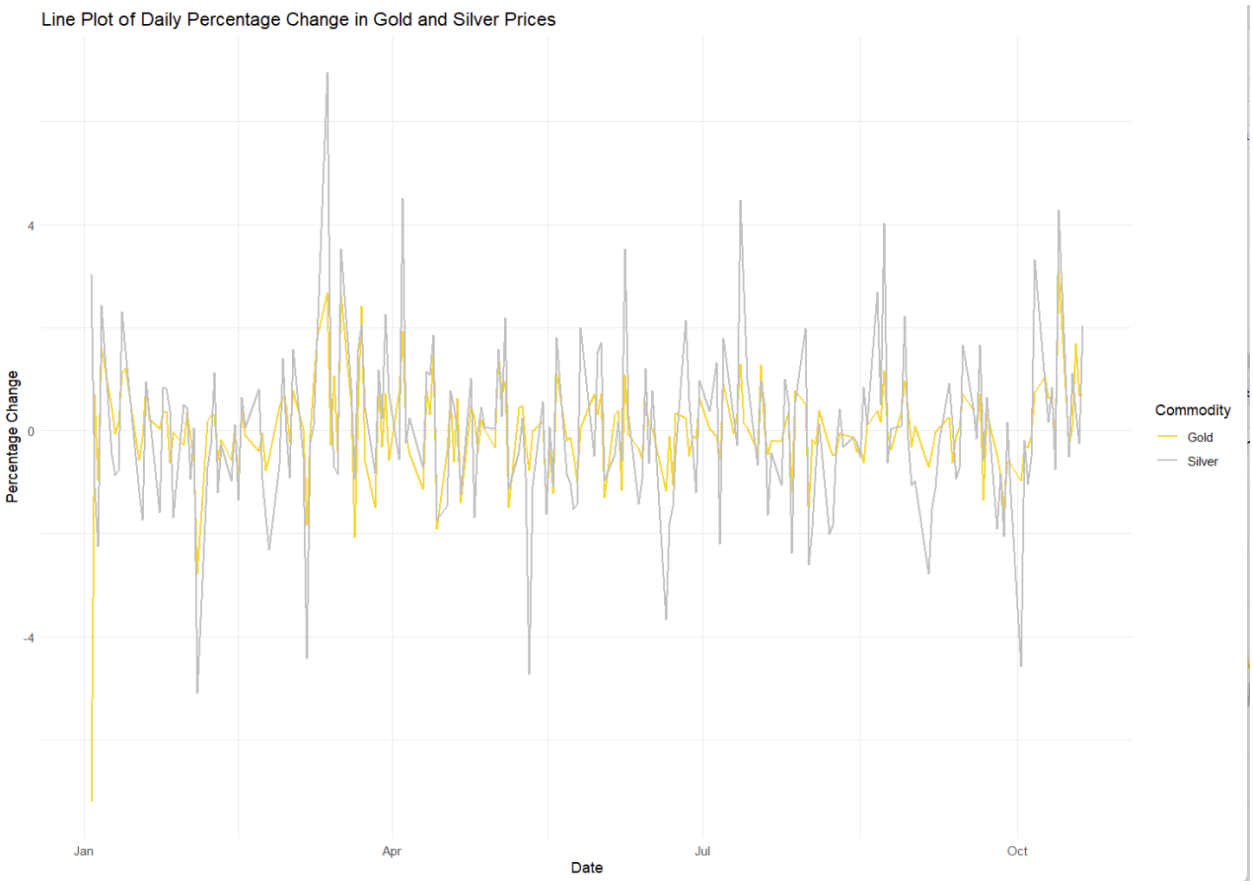
Creating a density plot to show the distribution of daily percentage changes for gold and silver.

Using different fill colors for gold and silver to highlight their distributions.

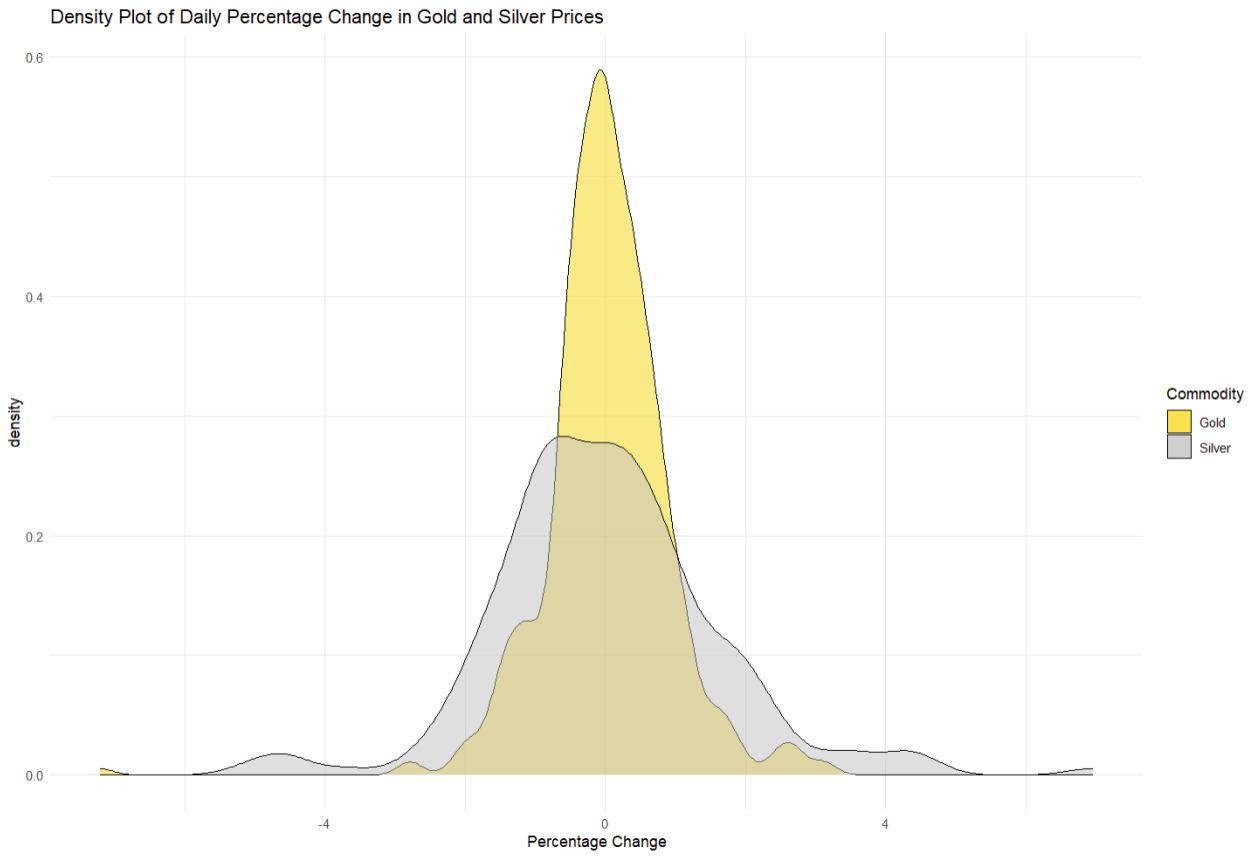
Visualization 3: Boxplot:

Constructing a boxplot to illustrate the central tendency and spread of daily percentage changes.

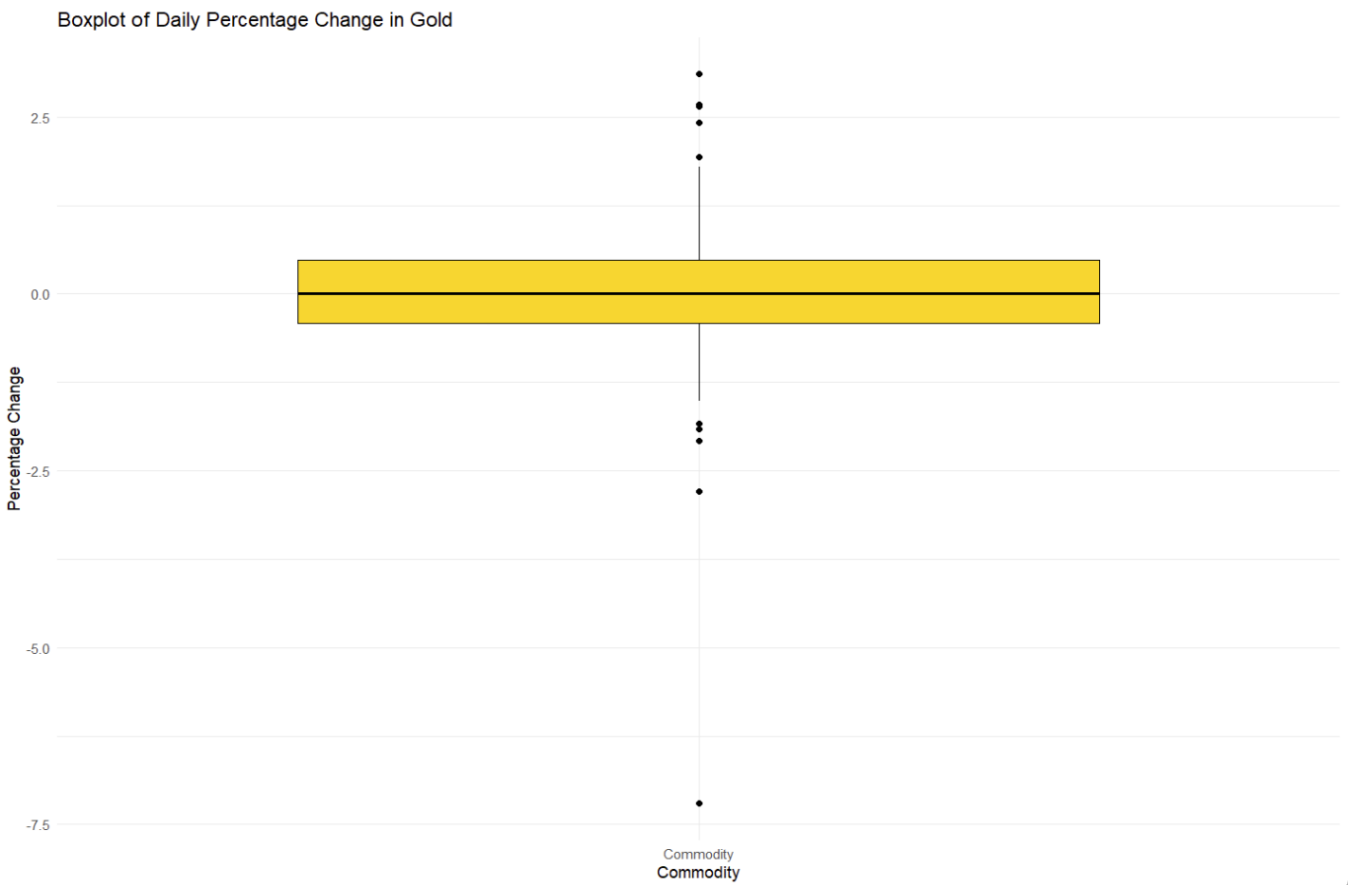
Grouping the boxplots by commodity (gold and silver).



While examining the monthly trends, there is a notably strong correlation between gold and silver, with an R value of **0.65**, indicating a moderately positive correlation. In January, gold experienced a high negative percentage change of **-4%**, while silver's change was less negative. Throughout 2023, silver consistently outperformed gold, particularly evident in July with silver having a significant positive change of over **4%** compared to gold's slight positive change. The trend persists until October, when both gold and silver show upward trajectories.



In the density plot, gold exhibited a concentrated pattern of percentage change, reaching a peak density of **0.6** and a low of (**-3.5**). This suggests relatively stable price fluctuations within that range. In contrast, silver displayed a less concentrated pattern with a peak density just under **3** and a low of **-5**, indicating a more diverse and variable range of percentage changes. The summary highlights the contrasting nature of price movements, with gold showing more stability and silver exhibiting greater variability.



The box plot suggests that while most of the data is centered around **0%**, there are some extreme values (outliers) indicating significant negative and positive percentage changes in gold prices. The minimum value, which is near **-7.5%**, might be worth investigating further due to its deviation from the overall trend.

Research Question #2: How do the monthly trading volumes of Gold and Silver in 2023 compare, and what are the factors influencing the highest trading volumes in specific months?

Reasoning for the Question:

I wanted to explore the monthly trading volumes for Gold and Silver in 2023 to understand the variations and identify any patterns or trends.

Hypothesis: I hypothesized that there might be certain months with higher trading volumes, possibly influenced by external factors like economic events or market sentiment.

Data Used:

I utilized the combined dataset of gold and silver prices (combined_gold_silver_data.csv). The dataset includes columns such as date, volume, and ticker, among others.

Analysis Process and Methods:

Data Preprocessing:

The date column was converted to the Date format.

Month and Year were extracted from the date column to create a new column month_year.

The data was filtered for Gold (GC=F) and Silver (SI=F) using the filter function.

Visualization 1: Scatter Plot:

Monthly Trading Volumes for Gold and Silver:

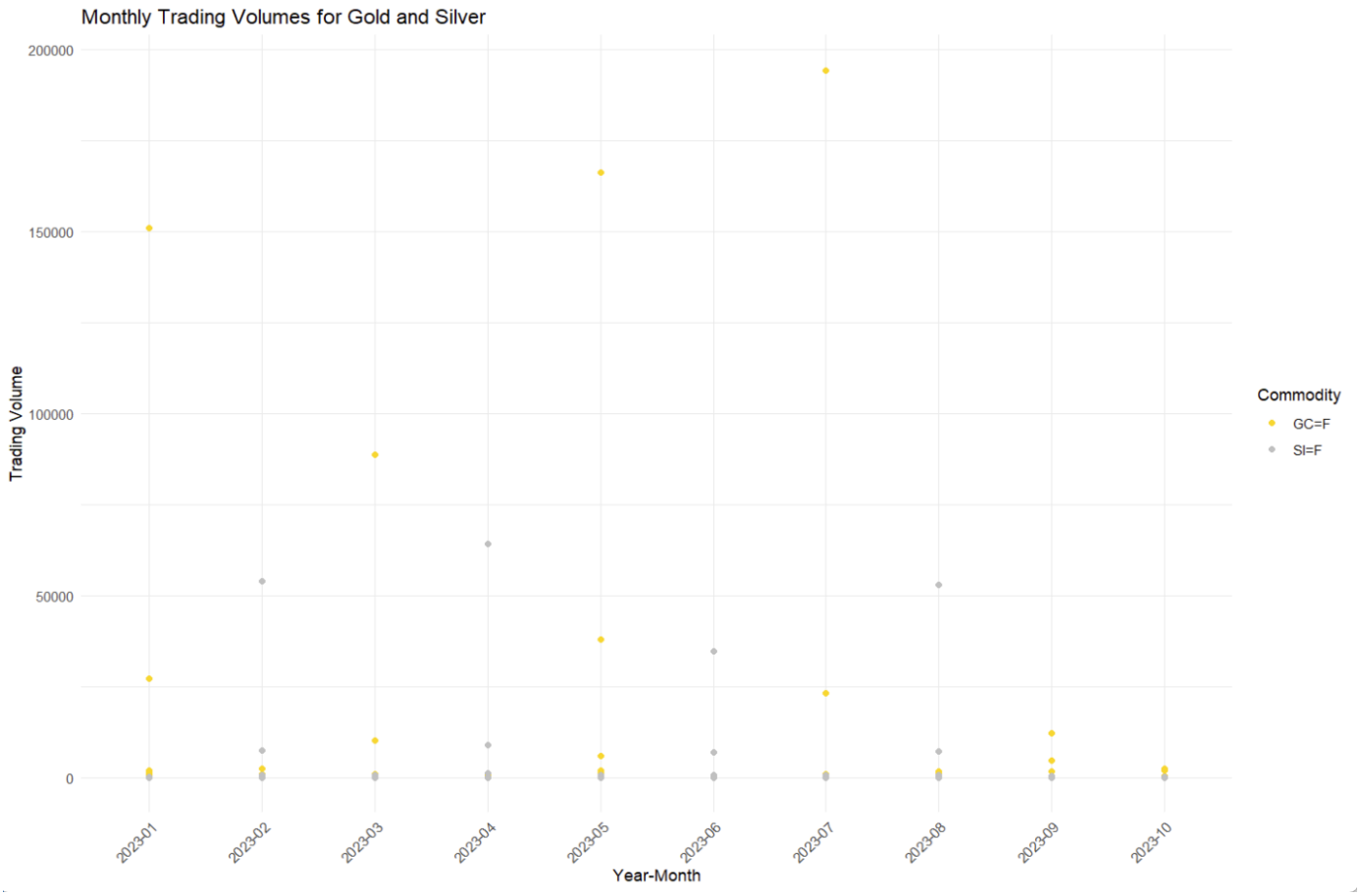
A scatter plot (geom_point()) was created to visualize the monthly trading volumes for Gold and Silver.

Different colors were assigned to Gold and Silver using the color aesthetic.

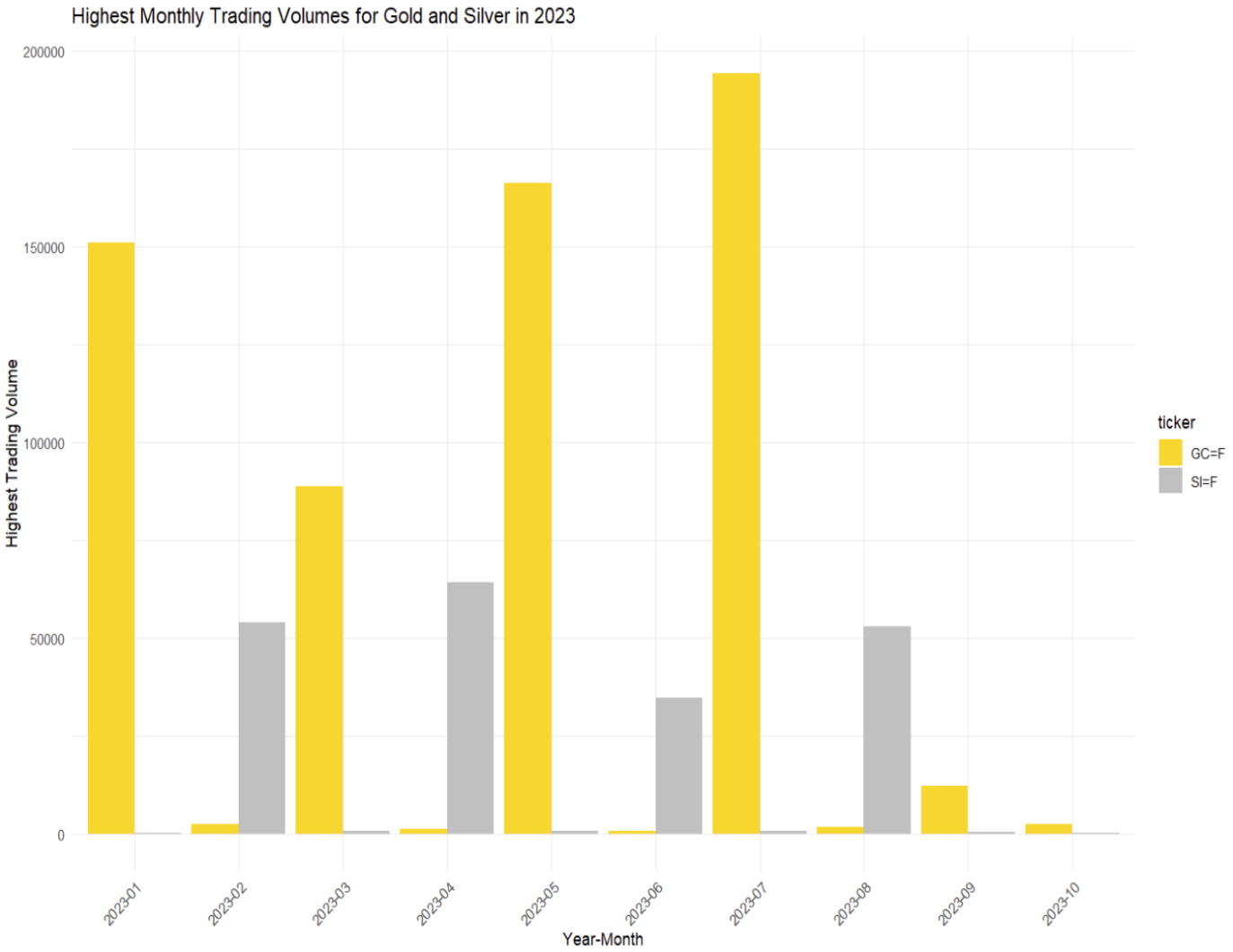
Highest Monthly Trading Volumes:

Calculated the highest monthly trading volumes for Gold and Silver using group_by and summarise.

A bar plot (geom_bar()) was generated to display the highest monthly trading volumes for Gold and Silver.



Observing outliers, gold displayed significant monthly trading volumes, reaching peaks of **194,253** in **July**, **166,066** in **May**, and **150,902** in **January**. In contrast, silver had a comparatively lower peak of **64,251**. The data suggests a higher trading interest in gold throughout 2023. Despite gold's outliers, both metals showed a consistent, highly correlated trading pattern in specific months.



The bar plot highlights the substantial disparity in monthly trading volumes between gold and silver. Notable instances include **February**, where silver surpasses gold with over **50,000** volume. March sees gold rebound with nearly **100,000** volume, while silver diminishes. This alternating trend continues in **April**. **May** and **July** stand out with gold peaking at over **150,000** and close to **200,000**. However, **September** and **October** show gold maintaining only a slight lead over silver.

Research Question #3: How do the candlestick patterns and area plots for gold and silver, represented by open, high, low, and close prices, evolve over time?

Reasoning for the Question:

This question aims to explore how the candlestick patterns, represented by open, high, low, and close prices, evolve over time for both gold and silver. Candlestick patterns are crucial in technical analysis, providing insights into market sentiment and potential price movements.

Hypothesis:

Given the historical use of candlestick patterns in identifying trend reversals and market sentiment, I hypothesized that there would be distinct patterns in the evolution of candlesticks for gold and silver. These patterns could reveal potential shifts in market dynamics.

Data Used:

The data used for this analysis includes daily candlestick data for gold (ticker: GC=F) and silver (ticker: SI=F) over a specified time (combined_gold_silver_data).

Analysis Process and Method(s):

Data Preprocessing:

The date column is converted to the Date format to facilitate accurate time-based analysis.

Candlestick Chart:

A candlestick chart is generated using plotly to visually represent the open, high, low, and close prices for both gold and silver.

Visualization 1:

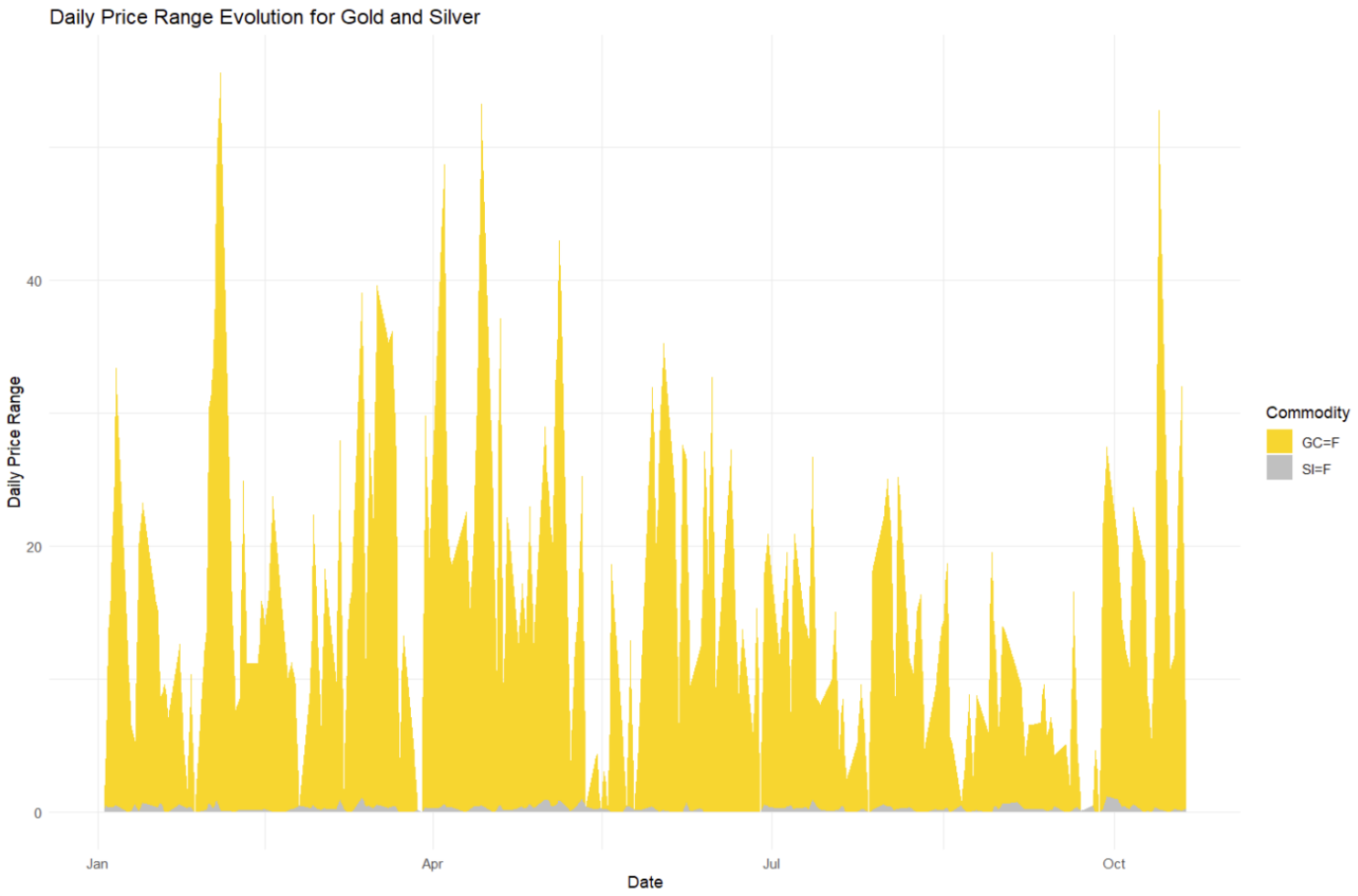
Candlestick Chart:

This visualization provides an overall view of how candlestick patterns evolve over time for both metals, highlighting potential trends and reversals.

Visualization 2:

Area Chart for Daily Price Range: An area chart illustrates the daily price range evolution, offering insights into volatility and potential market dynamics.

In the context of gold and silver analysis, the candlestick chart becomes a valuable tool for understanding the price dynamics of these precious metals. For gold and silver traders, recognizing specific candlestick patterns can offer insights into potential reversals, continuations, or indecision in the market. Bullish candlestick patterns, like the hammer or bullish engulfing, may indicate upward momentum, suggesting favorable conditions for buying gold or silver. Conversely, bearish patterns such as the shooting star or bearish engulfing might signal potential downtrends, prompting traders to consider selling or adjusting their positions. Additionally, the candlestick chart helps traders assess the volatility and intensity of price movements, allowing them to adapt strategies based on market conditions. The analysis of candlestick patterns in the context of gold and silver provides a nuanced understanding of market sentiment, aiding investors in making more informed decisions in the precious metals market.



The area map depicting the daily price evolution of gold and silver provides a clear visual representation of the inherent price range differences between the two precious metals. As expected, gold, being traditionally priced higher than silver, exhibits a broader high-low daily price range. This visualization is particularly useful for investors and analysts as it allows for a quick and intuitive comparison of the volatility levels between gold and silver. Traders can utilize this information to gauge the magnitude of potential price fluctuations in both metals, aiding in risk assessment and strategic decision-making. Understanding the distinct daily price patterns helps investors tailor their trading strategies, considering the varying volatility and risk profiles associated with gold and silver in the precious metals market.

Conclusion:

Through comprehensive analysis of gold and silver markets, this project has uncovered valuable insights that can inform strategic investment decisions. The exploration of the correlation between gold and silver prices revealed a moderately positive correlation, indicating that, on average, when one metal's price rises, the other tends to follow suit. The visualizations, including line plots, density plots, and box plots, provided a nuanced understanding of the daily percentage change distribution, identifying periods of divergence and convergence in their price movements.

The investigation into monthly trading volumes showcased the dominance of gold over silver, with consistently higher volumes and notable outliers, particularly in July, May, and January. The alternating trend in trading volumes between gold and silver each month highlighted dynamic shifts in trader preferences and market dynamics. This information is valuable for investors looking to capitalize on fluctuations in trading volumes for gold and silver.

The analysis of candlestick patterns offered a technical perspective, providing insights into potential trend reversals and market sentiment. The candlestick chart illuminated distinct patterns in the open, high, low, and close prices for gold and silver. Additionally, the area chart depicting daily price range evolution showcased the inherent volatility differences between the two metals, aiding traders in risk assessment and decision-making.

Limitations and Suggestions for Future Work:

Despite the valuable insights gained, there are limitations to this analysis. The dataset's timeframe is limited to 2023, and extending the analysis over a more extended period could reveal additional long-term trends. External factors influencing gold and silver prices, such as geopolitical events or economic indicators, were not explicitly incorporated, presenting an opportunity for more in-depth analysis. Additionally, considering a broader set of precious metals beyond gold and silver could provide a comprehensive understanding of the entire market.

Future work could involve incorporating sentiment analysis from financial news or social media to assess the impact of external factors on precious metal prices. Further, machine learning models could be employed to predict price movements based on historical data, enhancing the project's predictive capabilities. The analysis could also benefit from exploring intraday patterns and trading strategies for more granular insights into market behavior.

In conclusion, this project serves as a foundation for understanding the intricate dynamics within gold and silver markets, offering actionable insights for investors, analysts, and businesses operating in the precious metals industry.

Sources:

Listed in the data section.