UNIVERSITY OF WESTMINSTER

INFORMATICS INSTITUTE OF TECHNOLOGY

5BUIS006C - Data Visualisation and Communication

PORTFOLIO

Lecturer Name:

Mr Fouzul Hassan

Name : ***Shahly Fayeek***

UOW ID : ***20521398***

IIT ID : ***20231076***

Table of Contents

Research Question & Data Sourcing

Research Question

What are the key factors influencing the closing prices of Apple's stocks in the historical stock market data, and how do these factors vary over time?

Relevance

This research question is relevant because understanding the factors that influence Apple's stock prices is crucial for investors, financial analysts, and policymakers. By identifying these factors, we can gain insights into market behaviour, the impact of economic events, and seasonal trends. This knowledge can help make informed investment decisions, manage risks, and understand the stock market dynamics.

Data Sourcing

The dataset used for this analysis is the "Stock Market Historical Data," which contains daily stock prices and trading volumes for various companies. The data includes columns such as Date, Month, Open, High, Low, Close, and Volume. This dataset was sourced from Kaggle, providing a comprehensive view of stock market trends over time.

For this research, we are focusing specifically on the stock prices of **Apple Inc**. This targeted approach allows us to gain detailed insights into the factors influencing Apple's stock prices and how these factors vary over time. **The data on Apple's stock prices was extracted and saved in a separate CSV file, "apple_stock_prices.csv" for detailed analysis.**

Data Preparation

The dataset used for this analysis, "apple_stock_prices.csv," contained daily stock price information for Apple. The initial inspection revealed inconsistencies in date formats and potential non-numeric values in key columns. To ensure the dataset was tidy and ready for analysis, several steps were undertaken:

Variables

1. **Company**: The name of the company, which is "AAPL" for Apple Inc.
2. **Date**: The date of the stock price data, formatted as YYYY-MM-DD.
3. **Close**: The closing price of the stock on that particular date.
4. **Volume**: The number of shares traded on that date.
5. **Open**: The opening price of the stock on that date.
6. **High**: The highest price of the stock reached on that date.
7. **Low**: The lowest price of the stock reached on that date.
8. **Month**: The month of the year when the data was recorded, represented as a number (1 for January, 2 for February, etc.).

Handling Date Format

```r
# Load necessary libraries
library(tidyverse)
library(lubridate)
library(zoo)

# Load the dataset
df <- read.csv("apple_stock_prices.csv")

# Inspect the first few rows of the Date column
head(df$Date)

# Convert the 'Date' column to Date format using multiple formats
df$Date <- parse_date_time(df$Date, orders = c("ymd", "mdy", "dmy"))
```

The Date column was initially in mixed formats, which could lead to errors in time-based analyses. Using the parse_date_time() function from the lubridate library, the dates were converted to a consistent Date format, accommodating multiple formats such as "ymd," "mdy," and "dmy." Rows with invalid or missing dates were identified and dropped, ensuring the dataset retained only valid entries for time series analysis.

Cleaning Numerical Data

```r
# Check the data type of the 'Close' column
str(df$Close)

# Convert the 'Close' column to numeric
df$Close <- as.numeric(df$Close)
```

```r
# Check for any NA values introduced during conversion
sum(is.na(df$Date))

# Inspect rows with NA dates
df %>% filter(is.na(Date))

# Drop rows with NA dates
df <- df %>% drop_na(Date)
```

The Close column, representing daily closing prices, contained potential non-numeric values. These values were identified using pattern matching, and the column was converted to numeric format. Any non-numeric or missing values were handled by correcting or dropping the affected rows.

Feature Engineering

```r
# Create the 'Month' column
df$Month <- month(df$Date, label = TRUE)

# Check the first few rows of the dataset after conversion
head(df)
```

A Month column was created from the cleaned Date column using the month() function to facilitate time-based analysis. This new feature enabled monthly aggregations and comparisons, which are essential for exploring seasonal trends.

Moving Averages

```r
# Calculate moving averages for the 'Close' prices
df_company <- df

# Calculate moving averages for the 'Close' prices
df_company <- df_company %>%
  arrange(Date) %>%
  mutate(
    MA5 = rollmean(Close, k = 5, fill = NA, align = "right"),
    MA10 = rollmean(Close, k = 10, fill = NA, align = "right"),
    MA20 = rollmean(Close, k = 20, fill = NA, align = "right"),
    Quarterly_MA = rollmean(Close, k = 63, fill = NA, align = "right"),  # Approx. 63 trading days in a quarter
    Four_Quarter_MA = rollmean(Close, k = 252, fill = NA, align = "right")  # Approx. 252 trading days in four quarter:
  )
```

Several moving averages were calculated for the Close prices to smooth short-term fluctuations and identify underlying trends. These included averages of 5-day, 10-day, 20-day, and quarterly (63 trading days) and four-quarter (252 trading days). These features provided insights into both short-term and long-term market behavior.

Data Integrity Checks

```
# Save the cleaned dataset to a CSV file
write.csv(df, "/Users/shahlyfayeek/Desktop/DVC/DVC_CW/cleaned_apple_stock_prices.csv", row.names = FALSE)
```

Throughout the cleaning and preparation process, checks were performed to ensure the integrity of the data. For example, the dataset was scanned for missing or anomalous values, and the consistency of transformations was validated.

After completing these steps, the dataset was saved as a cleaned file for further exploratory and visual analysis. This preparation ensured that the dataset was tidy, reliable, and appropriately structured for addressing the research question.

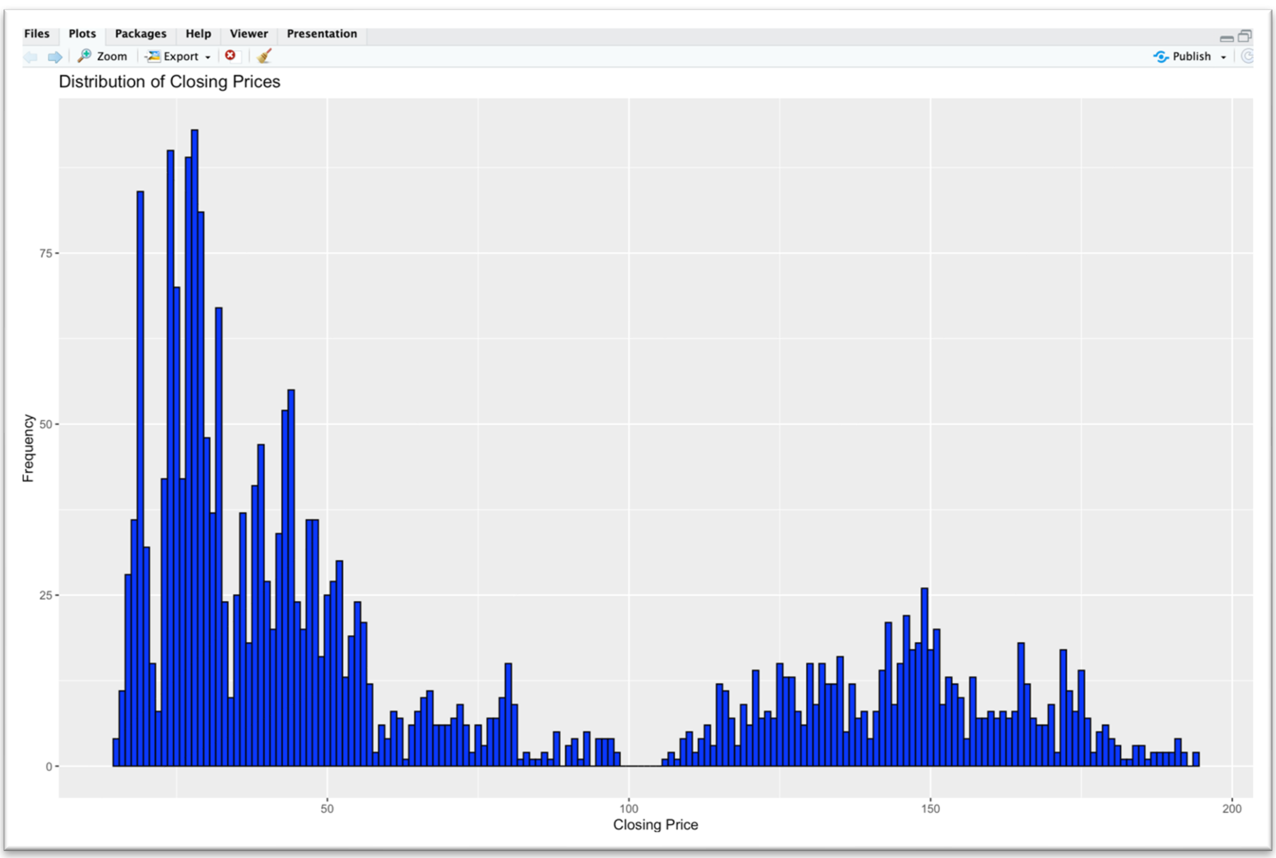
Exploratory Data Analysis

The exploratory data analysis (EDA) focused on understanding the distribution, patterns, and relationships within Apple's stock price data. A range of univariate, bivariate, and multivariate analyses was conducted to uncover insights and trends relevant to the research question.

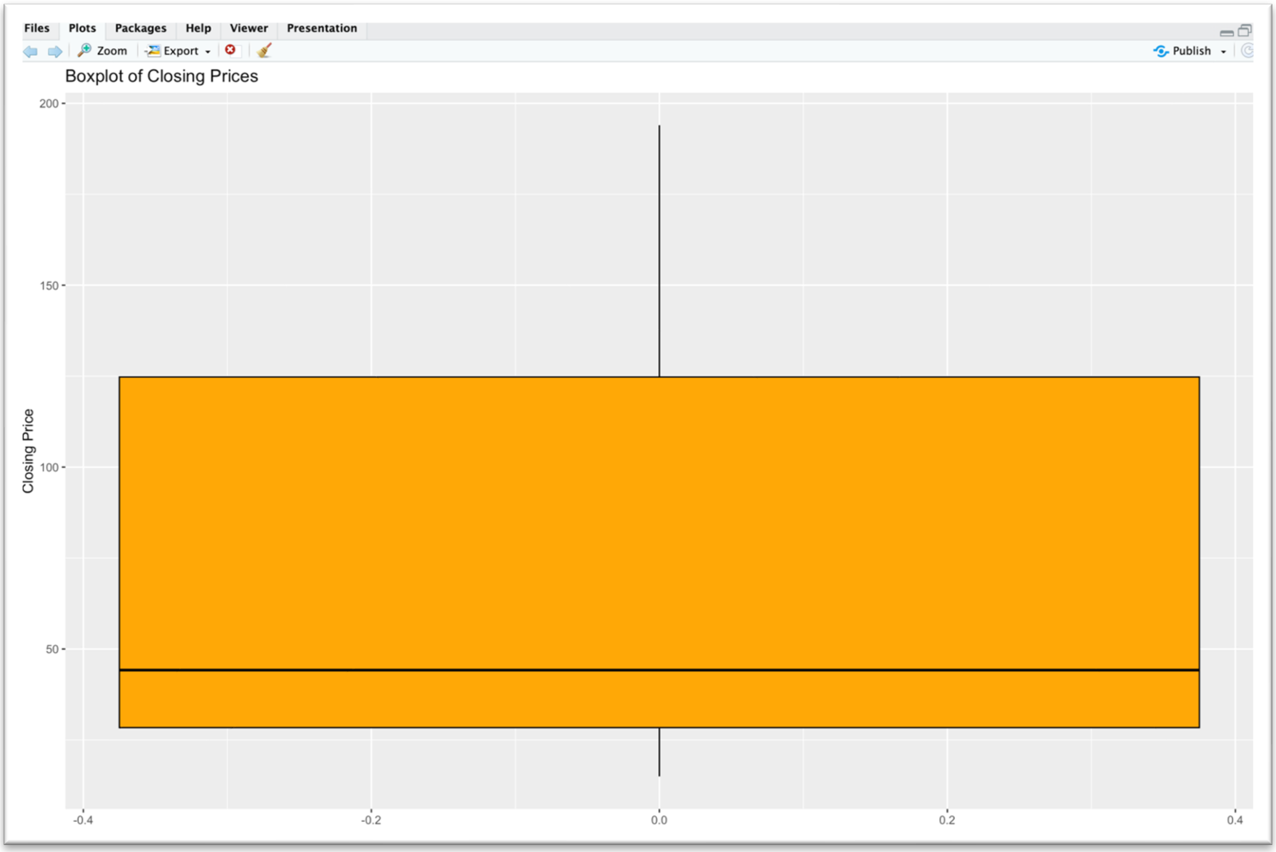
Univariate Analysis

The univariate analysis began with a detailed examination of the Close prices, representing Apple's daily closing stock prices.

- **Distribution Analysis**: A histogram of the Close prices revealed the overall distribution and frequency of stock price levels. This visual helped identify key ranges and potential outliers.



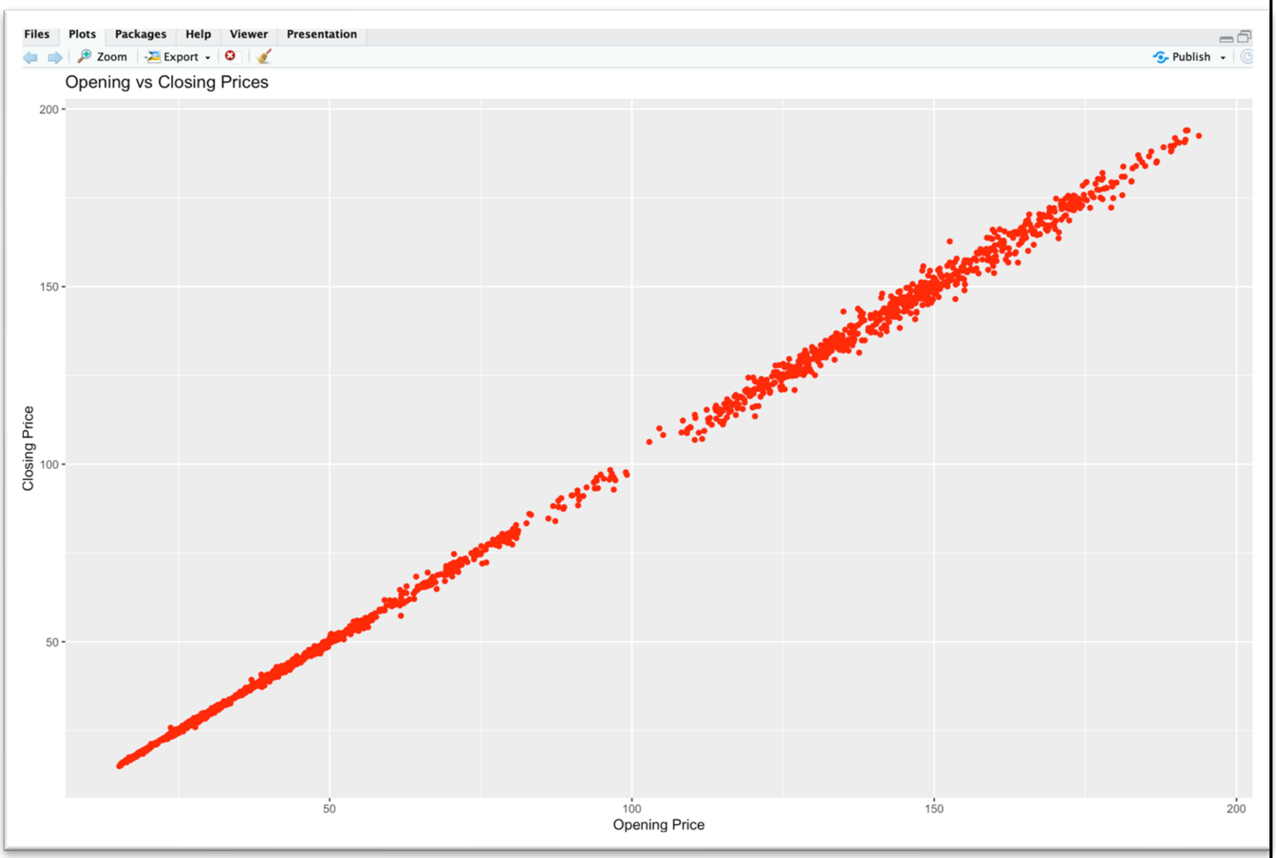
- **Outlier Detection**: A boxplot of Close prices highlighted the presence of outliers, providing insights into significant deviations from typical stock price behaviour.



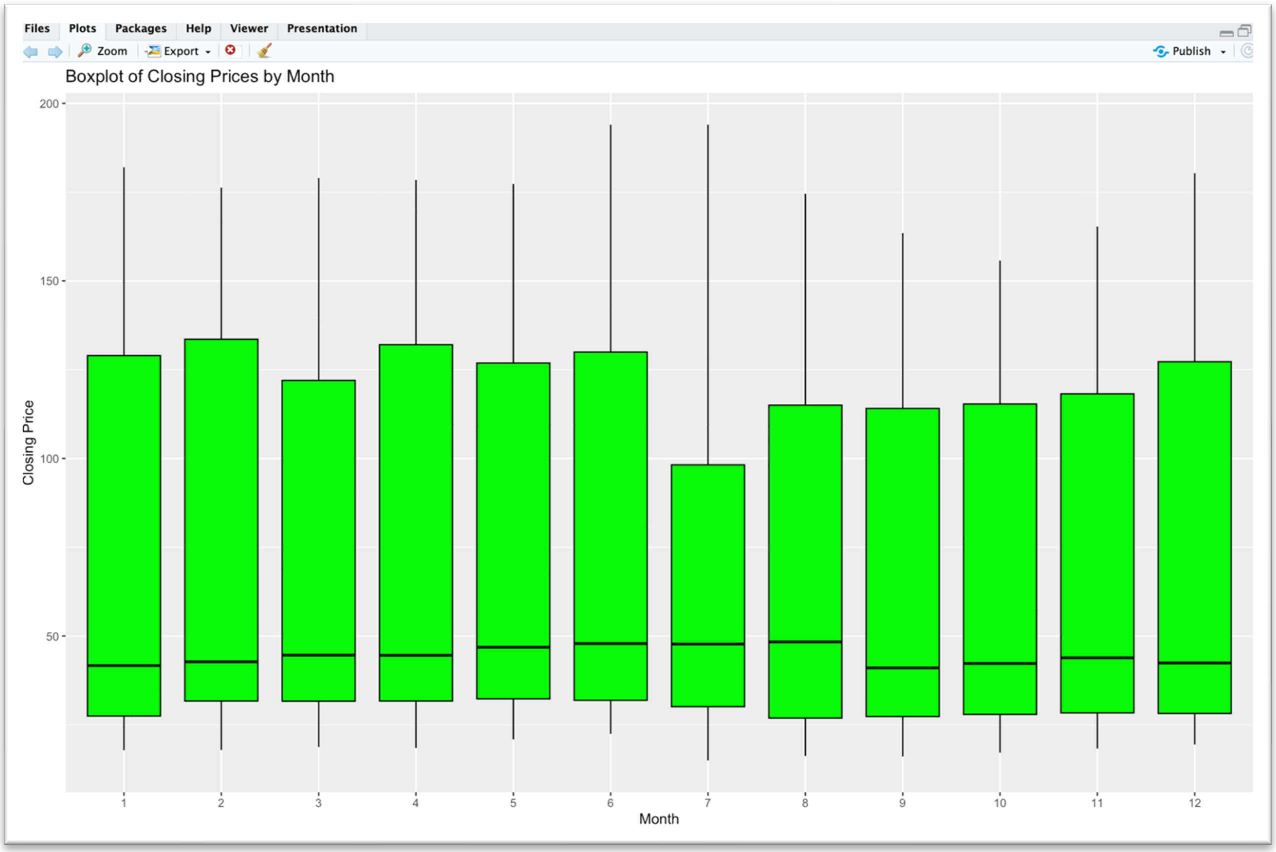
Bivariate Analysis

To explore relationships between key variables, several bivariate analyses were conducted:

- **Scatter Plot (Open vs. Close Prices)**: A scatter plot was created to examine the relationship between opening and closing prices. As expected, this revealed a strong positive correlation, indicating that the opening price often predicts the closing price within a trading day.

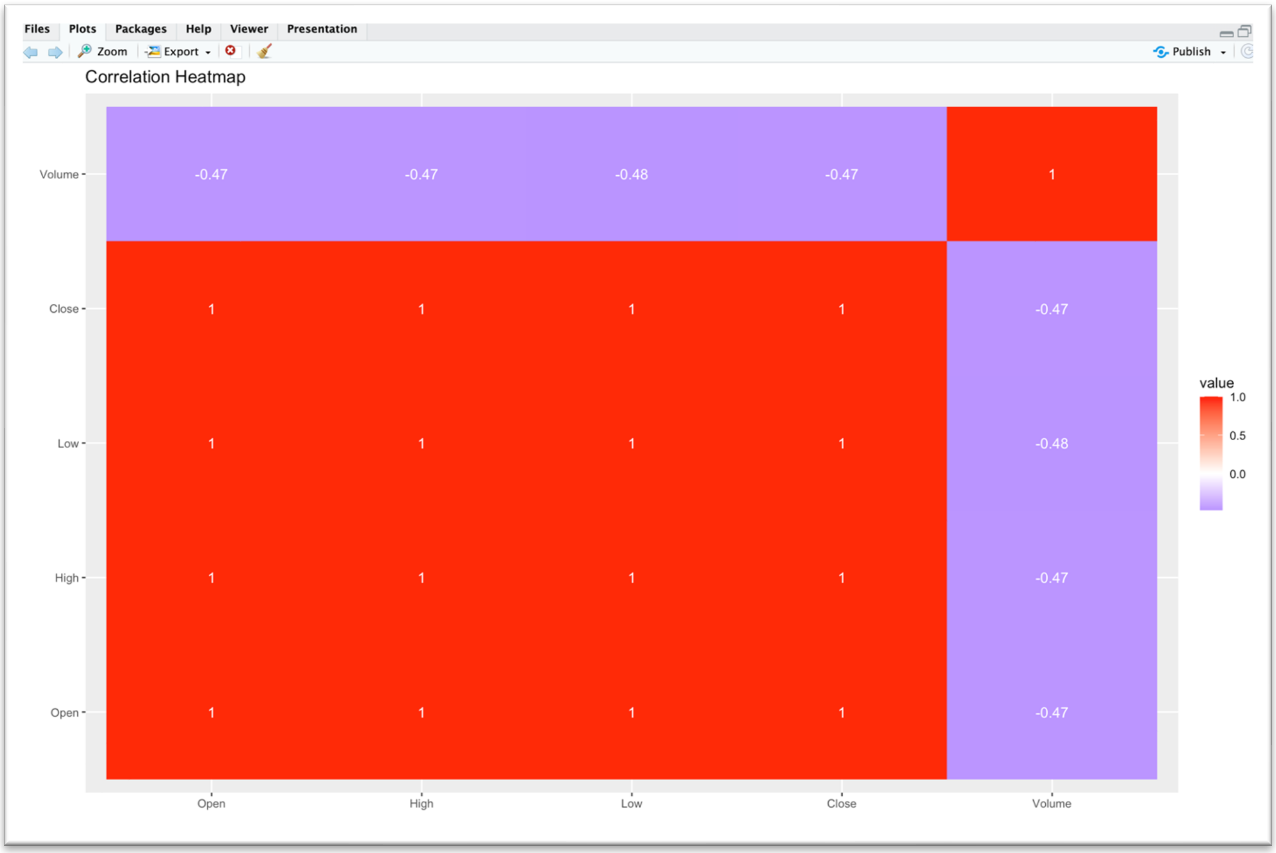


Boxplot by Month: Monthly trends were analysed by visualising closing prices grouped by months. This analysis highlighted seasonal variations and patterns, such as higher or lower average prices during certain months, which could indicate periodic market behaviour.



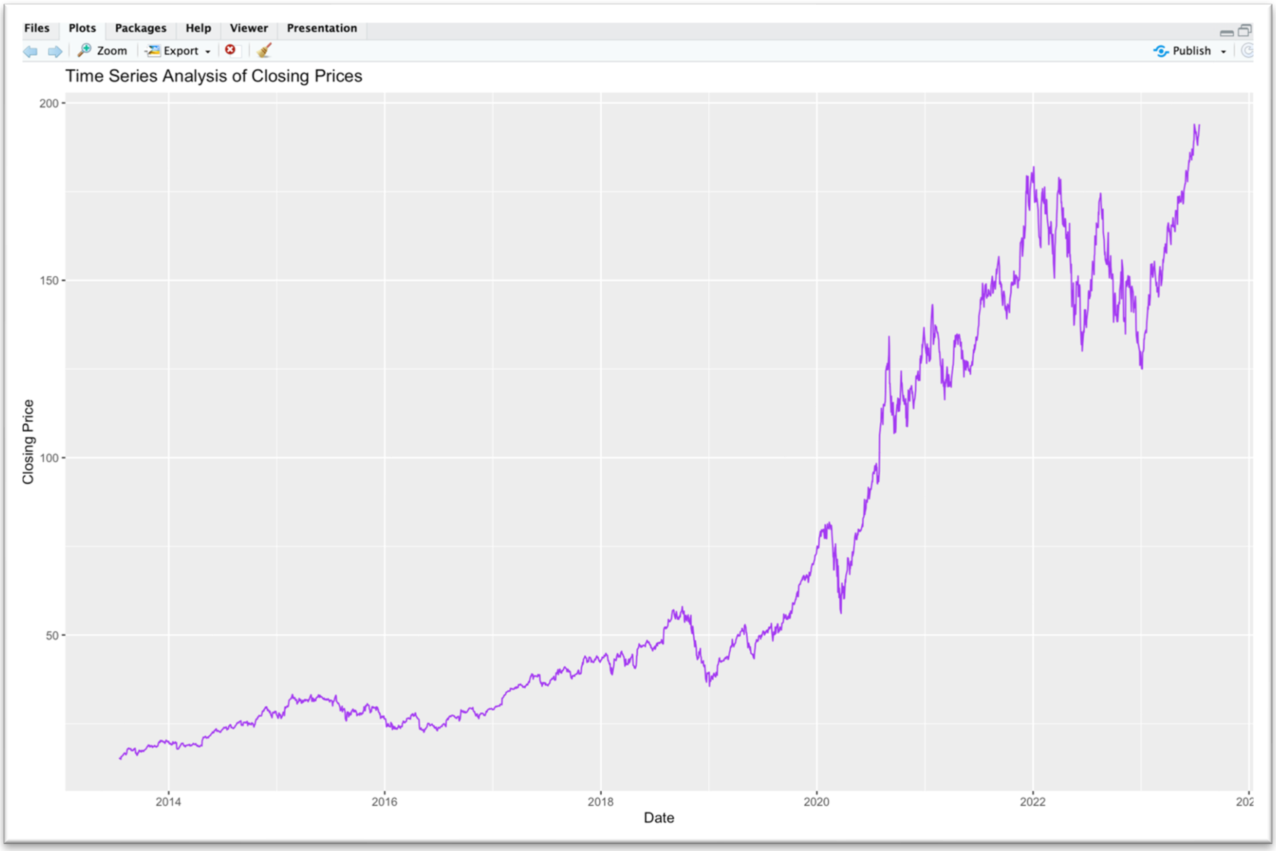
Correlation Analysis

A correlation matrix was computed for numerical variables, including Open, High, Low, Close, and Volume. This matrix was visualised as a heatmap to highlight the strength of relationships between variables. Strong correlations between Open, High, Low, and Close prices were observed, reflecting the interdependence of these metrics in daily stock trading.

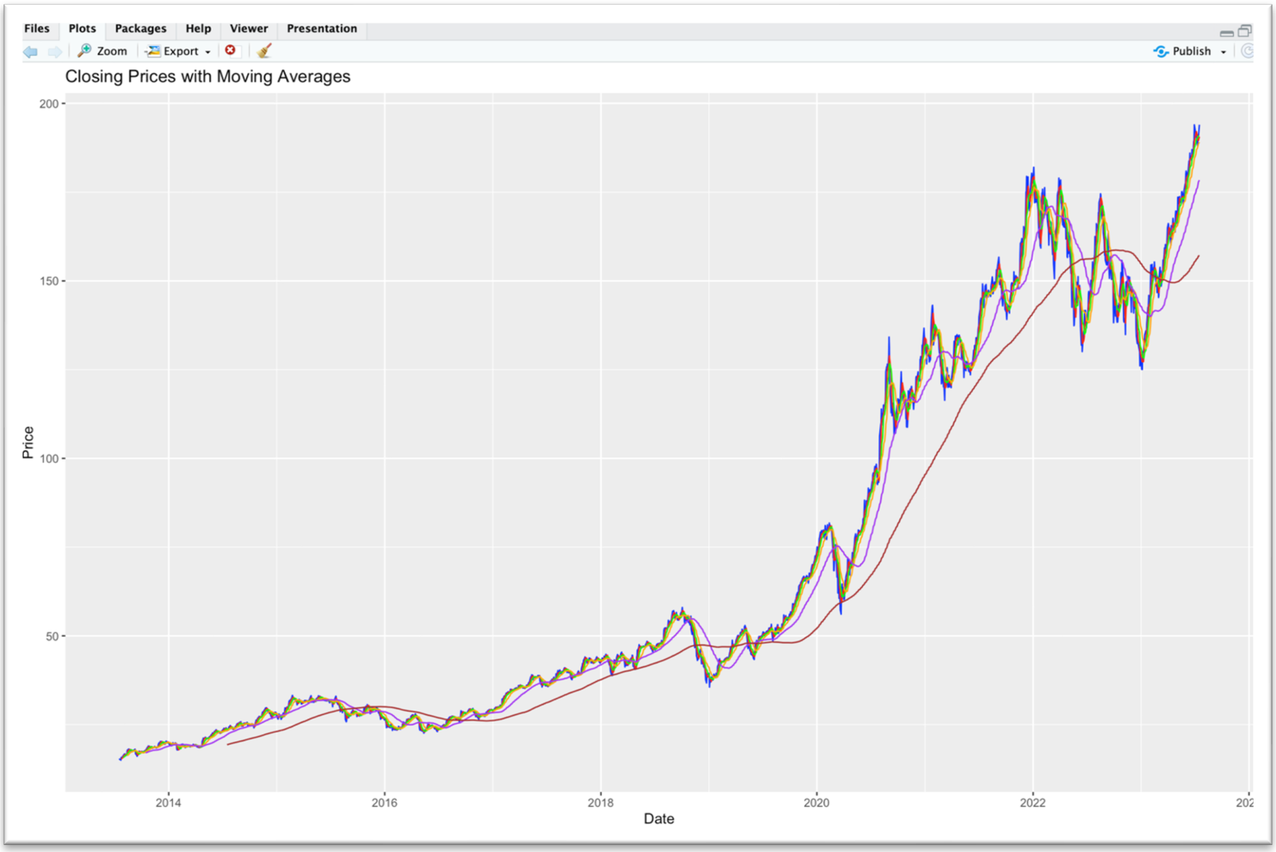


Time Series Analysis

A time series plot of the Close prices was created to visualise trends over the dataset's time span. This analysis revealed patterns of growth, dips, and potential volatility in Apple's stock performance.



- **Moving Averages**: To smooth short-term fluctuations, multiple moving averages (5-day, 10-day, 20-day, quarterly, and four-quarter) were calculated and overlaid on the time series plot. These moving averages provided insights into both short-term and long-term price trends, helping to identify periods of stability and significant market shifts.



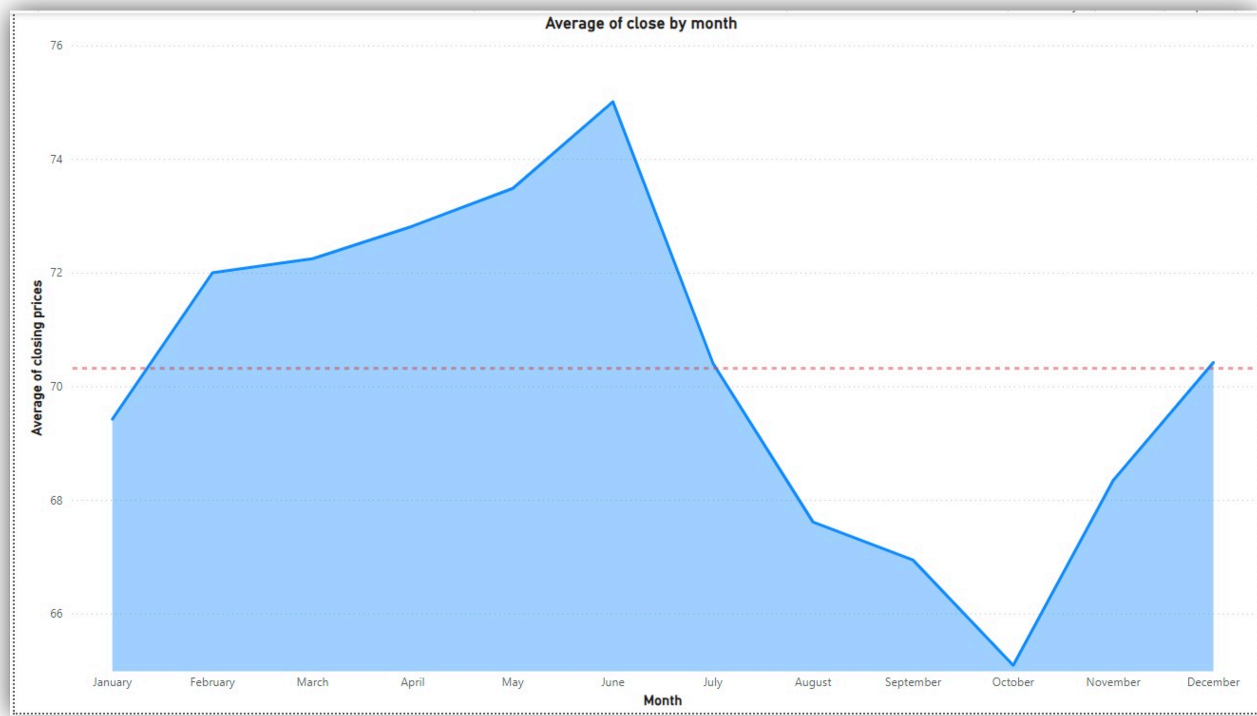
Key Insights

The dataset's structure, trends, and inter-variable relationships were thoroughly examined through EDA. This analysis forms the foundation for deeper insights and data storytelling, allowing for a comprehensive understanding of Apple's stock price movements over time.

Data Storytelling

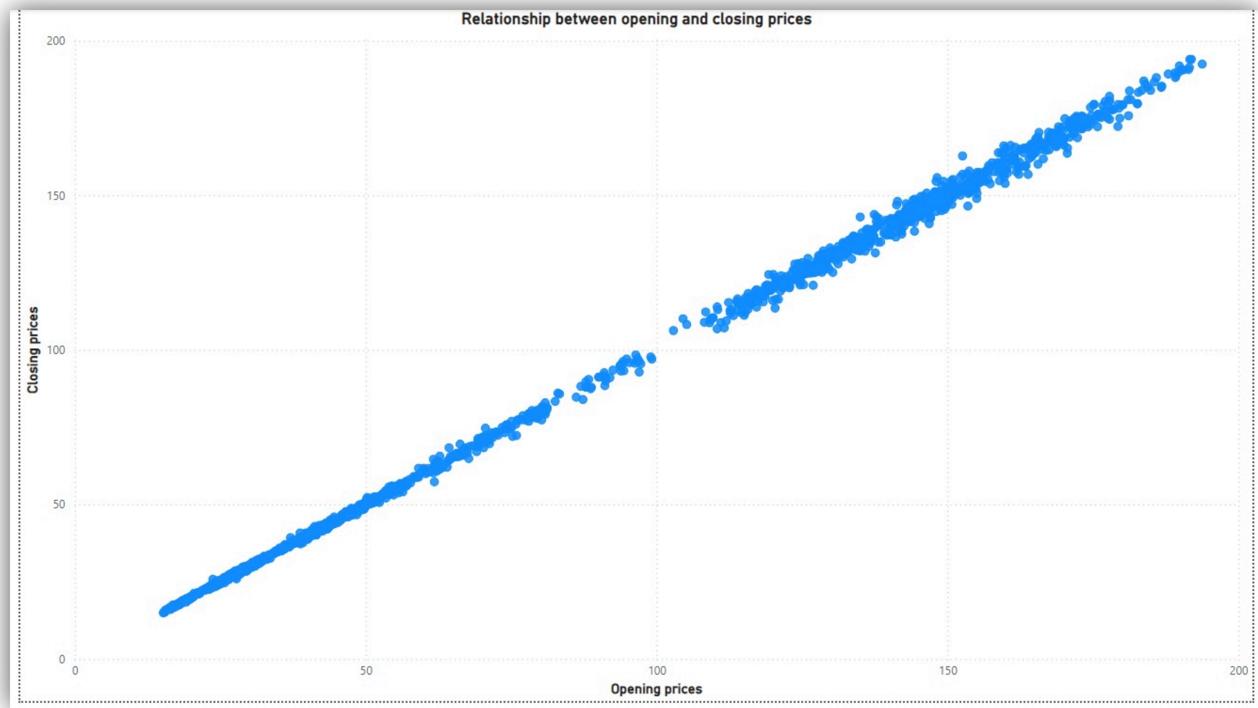
Apple Inc.’s journey in the stock market is a compelling story of innovation, resilience, and market dynamics. By delving into historical stock data, we uncover patterns, correlations, and trends that reveal the forces shaping Apple’s financial performance. This narrative, supported by detailed infographics, takes us through the highs and lows of Apple's market story and provides actionable insights for investors.

Seasonal Patterns and Investor Confidence



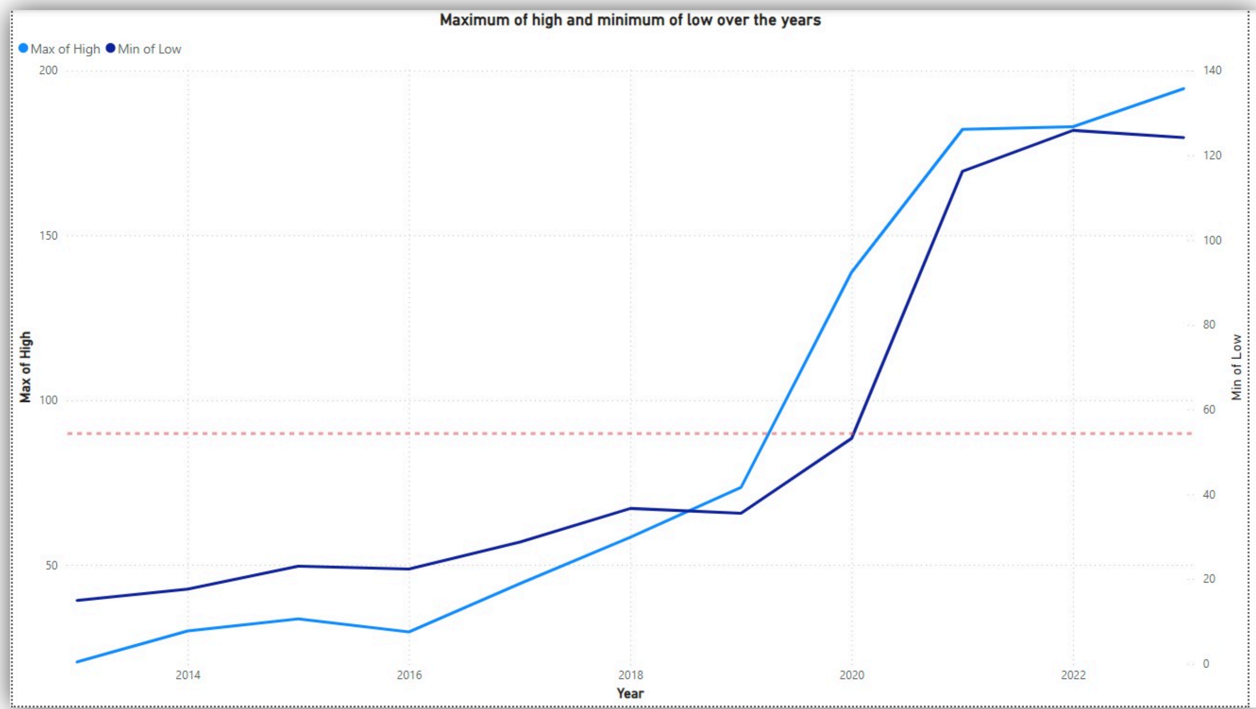
Apple’s stock prices show a fascinating seasonal rhythm. Monthly trends highlight that certain months, particularly those preceding the holiday season, witness consistently higher average closing prices. This surge is driven by heightened consumer confidence and investor anticipation of strong sales during major product launches and festive shopping periods. As illustrated in the monthly boxplot, peaks in closing prices around November and December underline the company’s strategic alignment with market opportunities.

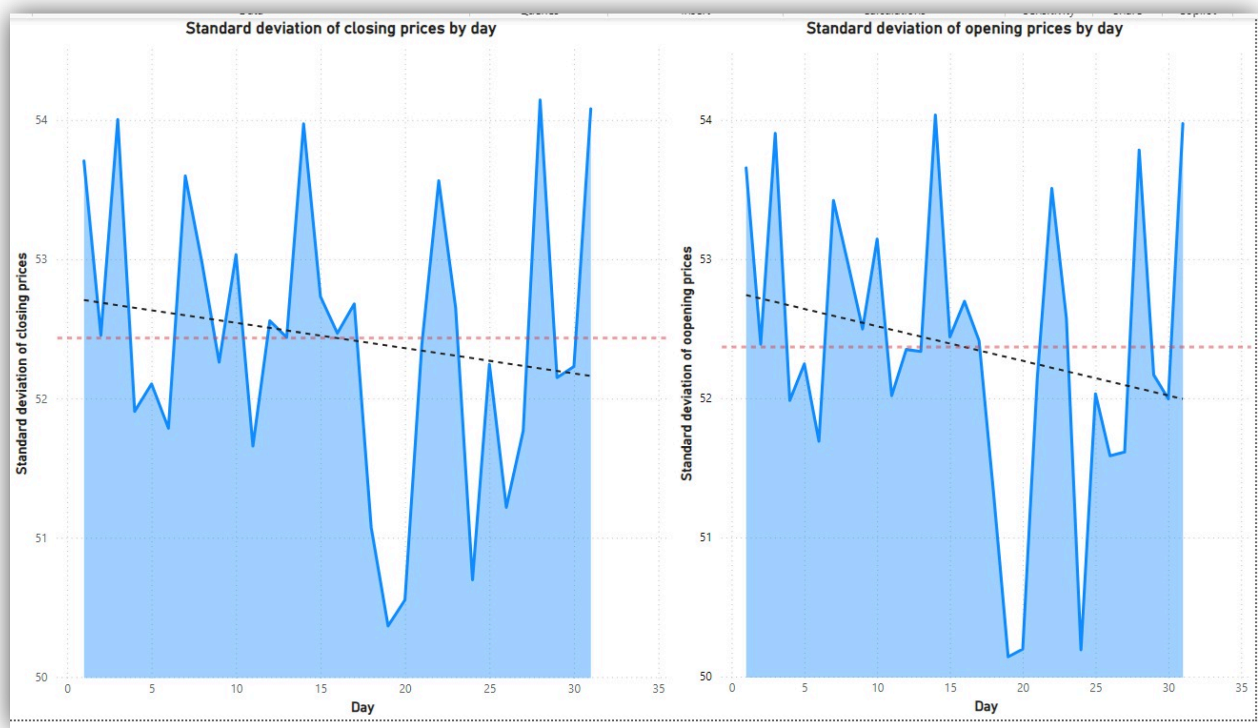
The Predictability of Opening and Closing Prices



A closer look at the relationship between opening and closing prices reveals a strong positive correlation. The scatter plot demonstrates that the opening price is often a reliable predictor of the day's closing price. This stability suggests that significant intraday fluctuations are relatively rare, providing a sense of predictability. For long-term investors, this consistency reinforces the appeal of holding Apple's stock as a dependable asset.

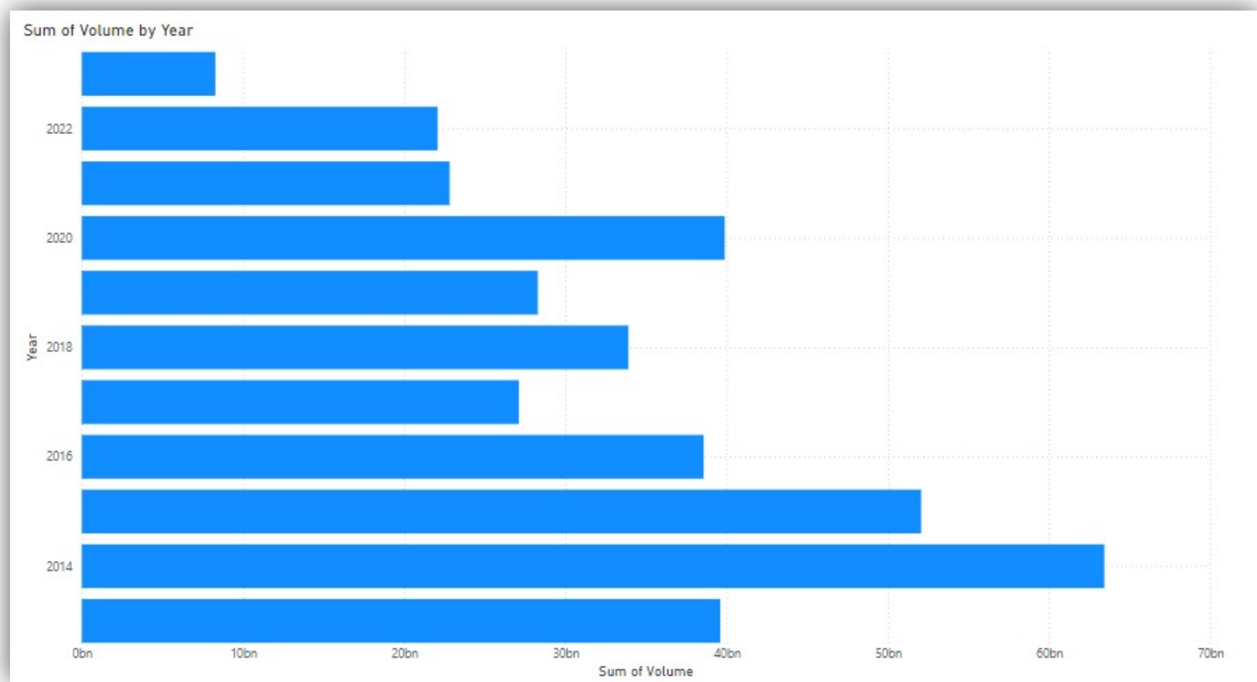
Volatility: Opportunities and Challenges





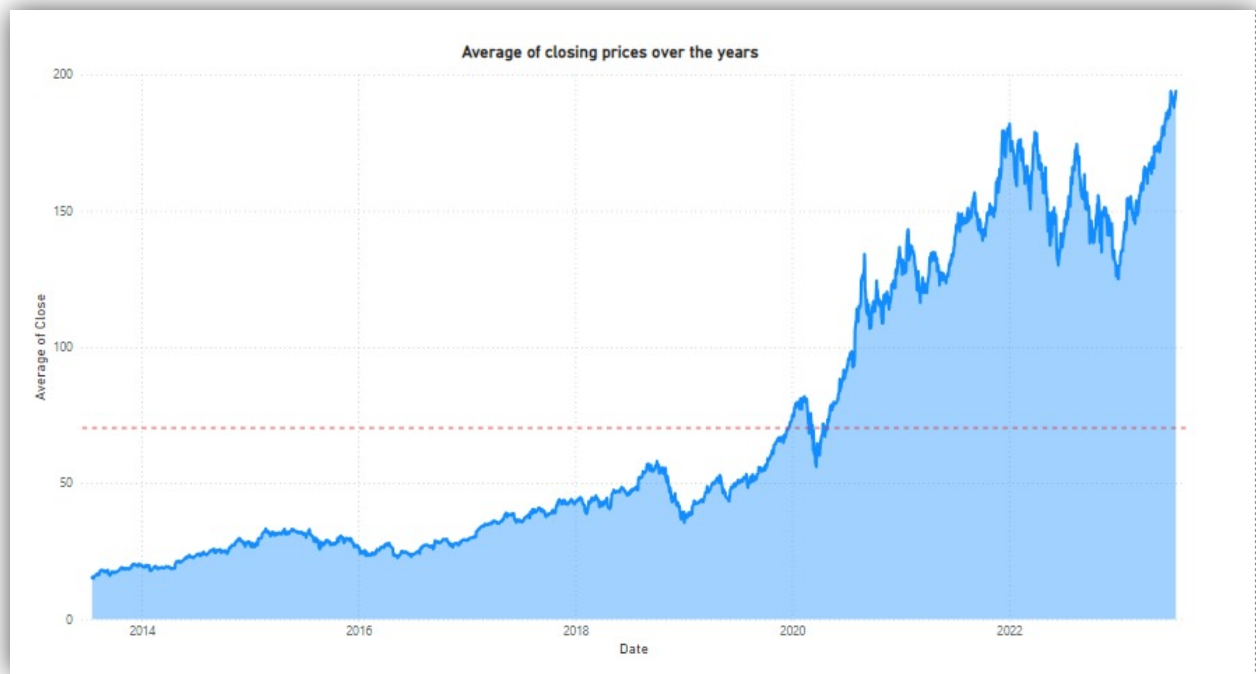
Over the years, Apple's stock has experienced varying levels of volatility. The range between daily high and low prices, depicted through detailed range plots, showcases periods of market turbulence influenced by economic downturns, major product announcements, or global events. For instance, sharp dips during certain quarters align with broader market corrections or competitive pressures, while sudden spikes correspond with successful product launches. Understanding these patterns helps investors balance risk and opportunity, leveraging periods of volatility for potential gains.

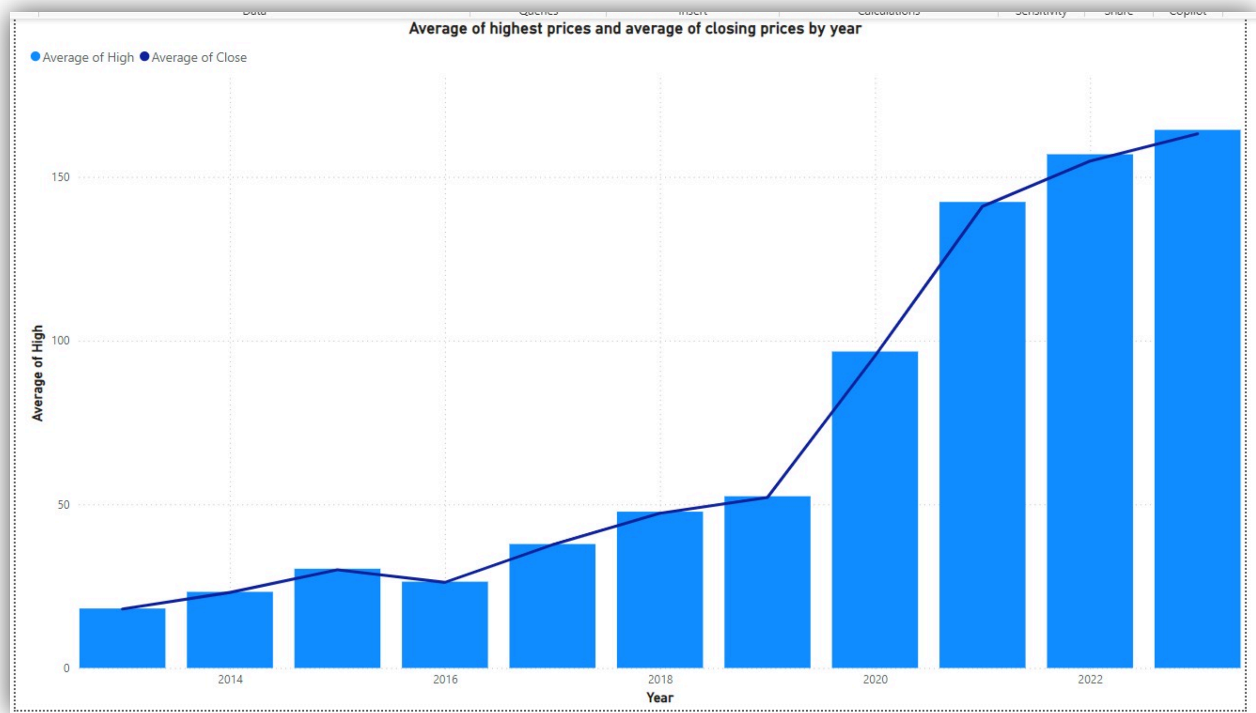
Insights from Trading Volumes



Trading volumes provide a window into market sentiment and investor activity. Infographics on trading volume trends highlight years with heightened trading, often linked to Apple's major announcements or external market factors. For example, the release of a new iPhone model or a game-changing software update often coincides with spikes in trading volume. High volumes signal increased investor interest, offering a cue for traders to closely monitor market reactions.

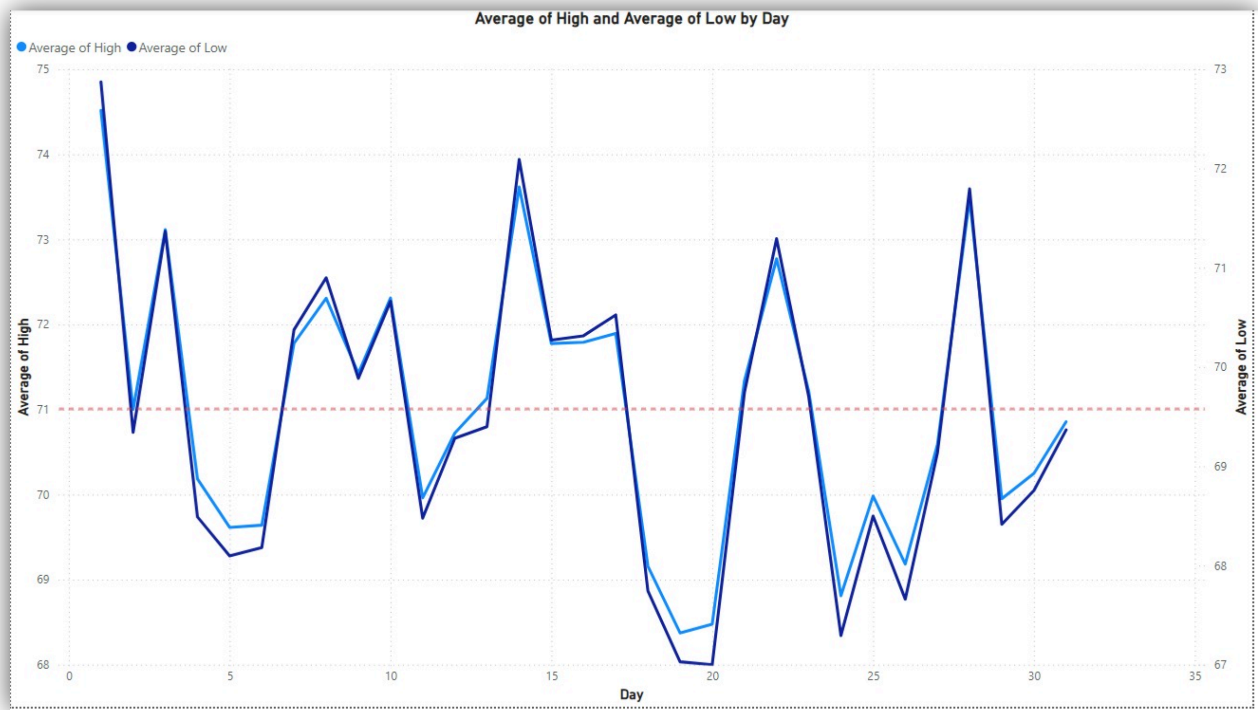
Long-Term Growth and Market Resilience





Apple's average yearly closing prices reveal a steady upward trajectory, reflecting the company's growth and its ability to adapt to market dynamics. This trend is a testament to Apple's resilience and innovative strategies in maintaining its competitive edge. Visual comparisons between yearly high prices and average closing prices further emphasize Apple's market strength, with consistent growth offering reassurance to long-term investors.

Strategic Insights for Investors



Apple's stock journey offers valuable lessons. Day traders can leverage daily high-low ranges and standard deviations to identify optimal entry and exit points, capitalizing on short-term fluctuations. Meanwhile, long-term investors can draw confidence from the steady appreciation in average yearly prices, reinforcing the benefits of holding Apple's stock through market cycles.

The Bigger Picture

The story of Apple’s stock is more than just numbers. It’s a narrative of a company’s relentless pursuit of excellence, navigating market challenges, and capitalizing on opportunities. From seasonal trends to long-term growth, each insight reflects a piece of a larger puzzle that defines Apple’s market presence. For investors and analysts alike, this journey underscores the importance of aligning market strategies with the underlying story of a stock.

In conclusion, Apple’s historical stock data provides a rich tapestry of insights. From the predictability of daily prices to the dynamics of trading volumes and long-term trends, the data tells a compelling story of growth, innovation, and resilience. For investors, these insights not only illuminate Apple’s past but also serve as a guiding light for navigating its future in the ever-evolving financial landscape.