

# Web Scraping & Time-Series Analysis

## Team Gamma

### SUMMARY

In this project, we explored how to collect and analyze pricing data over time using web scraping and time-series techniques. By extracting live data from a target website and analyzing its movement, we gained insights into price trends, volatility, and overall behavior.

### INTRODUCTION

The internet hosts vast and ever-changing price information—whether it's products, stocks, or services. Our goal was to build a tool that could automatically collect such data over a period of time and help us analyze the underlying patterns. This combination of **web scraping** and **time-series analysis** forms the core of many real-world applications, from e-commerce pricing strategies to financial forecasting.

### SCRAPING

#### METHOD 1 :

We used a combination of **Selenium** and **BeautifulSoup**, two powerful Python tools for web scraping.

- **Selenium** helped us **automate a real web browser**. This was important because Yahoo Finance loads its historical stock data dynamically using JavaScript — and tools like BeautifulSoup alone can't access such content.
- We configured Selenium to run in the background (headless mode) and visit the target page that lists Apple Inc. (AAPL) stock's historical prices.
- After giving the page a few seconds to fully load, we used **BeautifulSoup** to **read and extract the HTML content** of the page. We specifically looked for the table rows that contain the price data.
- From each row, we extracted:
  - The **Date**
  - The **Closing Price**
  - The **Trading Volume**
- We filtered and cleaned the data to only collect valid rows, then **stored everything in a Pandas DataFrame**, which is a structured table format in Python.
- Finally, we saved the data into a **CSV file (aapl\_prices.csv)**, which we later used for time-series analysis and visualization.

This approach allowed us to automate the collection of price data without manually visiting the website or copying values. It ensured accuracy and made it easy to repeat or scale the process for more stocks or different date ranges.

## **METHOD 2 :**

While Selenium and BeautifulSoup are effective for scraping, they can be **slower, more error-prone, and harder to maintain**—especially when dealing with dynamic websites like Yahoo Finance.

### **Better Approach: yfinance**

To simplify and improve our data collection, we used the **yfinance Python library**, which is a wrapper around the Yahoo Finance API. It allows us to **directly download stock price data** without needing to scrape the website.

### **How It Works (In Everyday Terms)**

- We used yfinance to fetch **Apple Inc. (AAPL)** stock data from the last 60 days.
- The function `yf.download("AAPL", period="60d")` pulled historical price data, including:
  - Date
  - Open, High, Low, Close prices
  - Volume
  - Adjusted Close
- The result was immediately available in a clean table format (a **Pandas DataFrame**).
- We then saved this data to a CSV file (`aapl_stock_data.csv`) for future analysis.

### **Why It's Better**

- **Fast** – Fetches data instantly from Yahoo Finance without needing to load a browser.
- **Reliable** – No issues with changing HTML structures or dynamic loading.
- **Simple** – Requires just a few lines of code.
- **Scalable** – Can be easily used for multiple stocks or larger datasets.

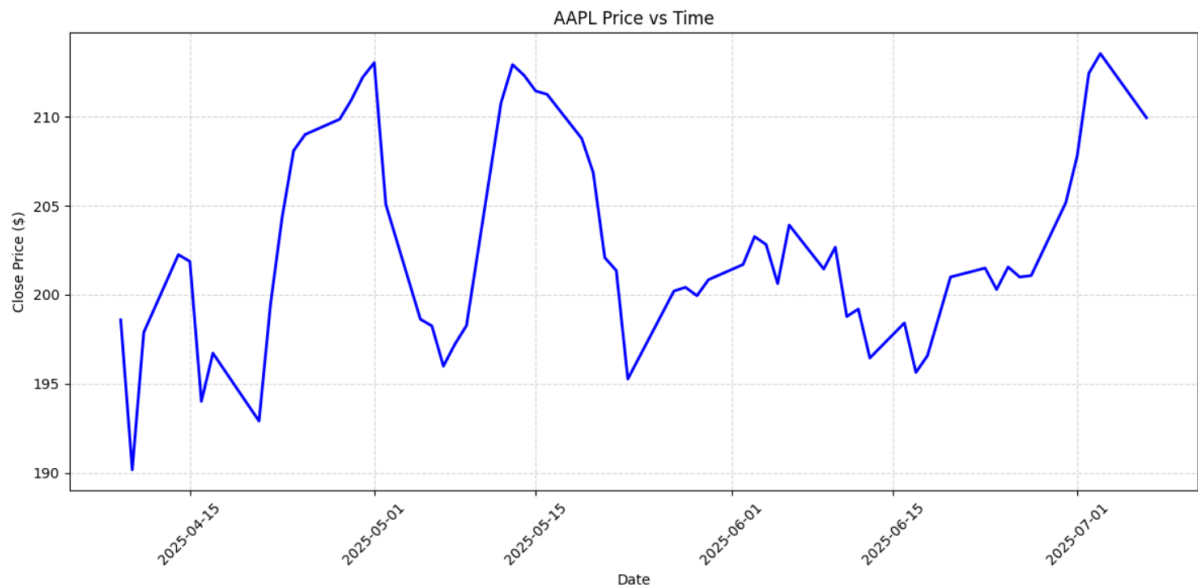
## **DATA VISUALIZATION**

### **AAPL Price vs Time – Summary**

The graph shows Apple Inc.'s (AAPL) closing stock prices over the last 60 days. The x-axis represents dates (April to July 2025), and the y-axis shows closing prices in USD.

The stock experienced several ups and downs, with peaks above \$210 around mid-May and early July, and lows near \$190 in mid-April and mid-June. This reflects short-term volatility with no strong upward or downward trend.

This graph effectively highlights how AAPL's stock price behaves over a two-month period. The observed fluctuations can help investors identify **potential trading opportunities**, analyze **market sentiment**, and make **data-driven decisions**. Additionally, this type of visualization is crucial for **time-series forecasting models** and understanding short-term price movements.



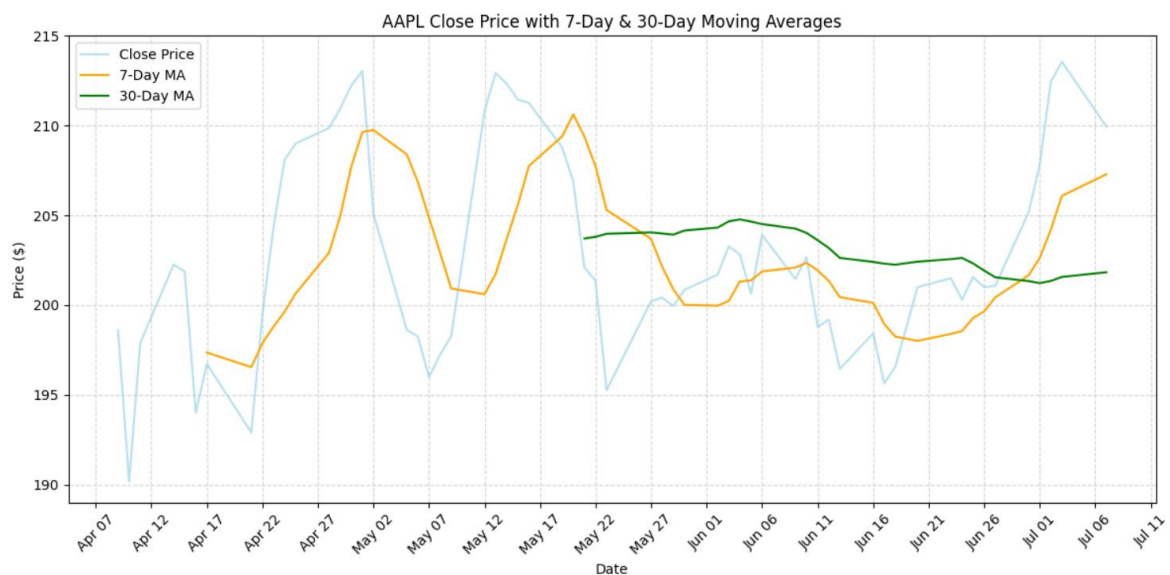
### **AAPL Close Price with 7-Day & 30-Day Moving Averages**

This graph shows Apple's closing stock prices over the last 60 days, along with two moving averages:

- **7-Day Moving Average (orange):** Reflects short-term trends and reacts quickly to price changes.
- **30-Day Moving Average (green):** Smooths out short-term fluctuations to show long-term trends.

The 7-day MA closely follows recent price movements, while the 30-day MA provides a more stable view. Crossovers between the two lines can signal potential trend changes.

The plot helps identify both short-term volatility and long-term trends in AAPL's stock performance, making it useful for timing investment decisions and understanding market sentiment.

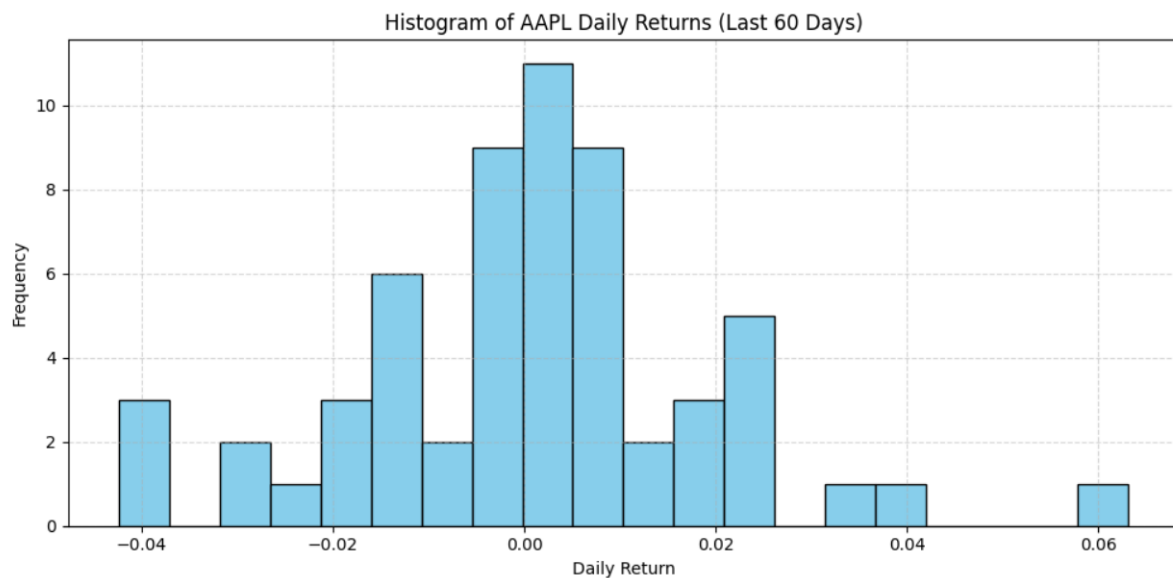


### **Histogram of AAPL Daily Returns (Last 60 Days)**

This histogram shows the distribution of Apple's daily stock returns over the past 60 days. The x-axis represents daily percentage changes in price, while the y-axis shows how frequently those changes occurred.

Most returns are centered around 0%, indicating that small daily changes are most common. However, there are a few days with larger gains or losses, highlighting occasional volatility.

The distribution is roughly bell-shaped with a slight skew, suggesting that while AAPL stock is generally stable, there are periods of sharp movement that investors should be aware of.



### **CONCLUSION**

Based on the analysis of AAPL's stock over the last 60 days, the overall trend shows moderate price fluctuations without a clear long-term direction. The line chart indicates short-term volatility, while the moving average plot highlights temporary shifts in momentum, with the 7-day MA reacting faster than the more stable 30-day MA. The histogram of daily returns shows most returns clustered around zero, suggesting general stability with occasional spikes in volatility. Together, these graphs indicate that while AAPL remains relatively stable, short-term opportunities and risks exist. Moving forward, this analysis can be extended with forecasting models, volatility metrics, or event-based studies to support better investment decisions.