Emilio Kiryakos, Adithya Jose, Bart Nowobilski, David Mendoza and Mustafa Ausaf

CS 418

4/30/24

Final Project Report

Data:

- Original Data Links:
  - Stock Data:
    - Apple Stock Data: https://github.com/BartNowo/LunarPhaseStockExploration/blob/main/data/apple.csv
    - NASDAQ Stock Data: https://github.com/BartNowo/LunarPhaseStockExploration/blob/main/data/nasdaq.csv
    - NVIDIA Stock Data: https://github.com/BartNowo/LunarPhaseStockExploration/blob/main/data/nvidia.csv
    - SPY Stock Data: https://github.com/BartNowo/LunarPhaseStockExploration/blob/main/data/spy.csv
    - Tesla Stock Data: https://github.com/BartNowo/LunarPhaseStockExploration/blob/main/data/tesla.csv
  - Moon Data: https://github.com/BartNowo/LunarPhaseStockExploration/blob/main/data/moon_phases.csv
- Scripts used to Clean the Data:

```
# Extract only the date component from 'timestamp' column
tesla_df['Date'] = tesla_df['timestamp'].str.split(' ').str[0]

# Merge the DataFrames on the 'Date' column
merged_df = pd.merge(moon_phases_df, tesla_df, on='Date')

# Drop the 'timestamp' column from the merged DataFrame
merged_df = merged_df.drop('timestamp', axis=1)

# Save the merged DataFrame to a new CSV file
merged_df.to_csv('merged_data.csv', index=False)
```

```
# converts the date columns to datetime type
moon_phases_df['Date'] = pd.to_datetime(moon_phases_df['Date']).dt.date
stock_data_df['Date'] = pd.to_datetime(stock_data_df['timestamp']).dt.date

# merges the dataframes on the data column the inner means it will include only dates
# present in both dataframes
merged_stock_df = pd.merge(
    moon_phases_df, stock_data_df, on='Date', how='inner')

merged_stock_df['Date'] = pd.to_datetime(merged_stock_df['Date'])

# Function to read the contents of a CSV file for the volume data
def read_volume_csv_file(file_path):
    with open(file_path, 'r') as csvfile:
        reader = csv.reader(csvfile)
        headers = next(reader)  # Skip the header row
        dates = []
        volume_values = []
        for row in reader:
            date = datetime.strptime(row[6], "%Y-%m-%d %H:%M:%S").date()
            volume = float(row[4])
            dates.append(date)
            volume_values.append(volume)
        return dates, volume_values

# Function to read the contents of a CSV file for the moon phases data
def read_moonphase_csv_file(file_path):
    with open(file_path, 'r') as csvfile:
        reader = csv.reader(csvfile)
        next(reader)  # Skip the header row
        dates = []
        moon_phases = []
        for row in reader:
            date = datetime.strptime(row[0], "%Y-%m-%d").date()
            moon_phase = row[1]
            dates.append(date)
            moon_phases.append(moon_phase)
        return dates, moon_phases
```

ML/Stats:

Correlation: Tesla Stock Prices vs Lunar Phases **(Adithya Jose):**

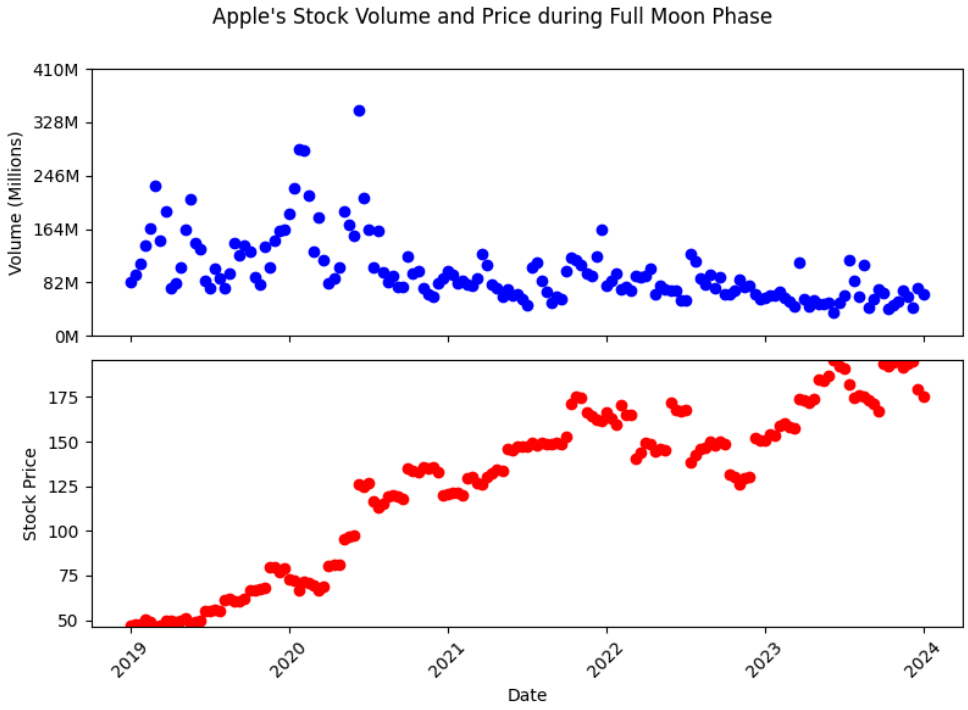
Used Linear Regression with the Mean Square Error, Mean Absolute Error and the R-squared values to find a correlation between the moon phases and the stock price of Tesla. The inferences uncovered were that the stock prices were consistent and didn't show any inconsistencies.

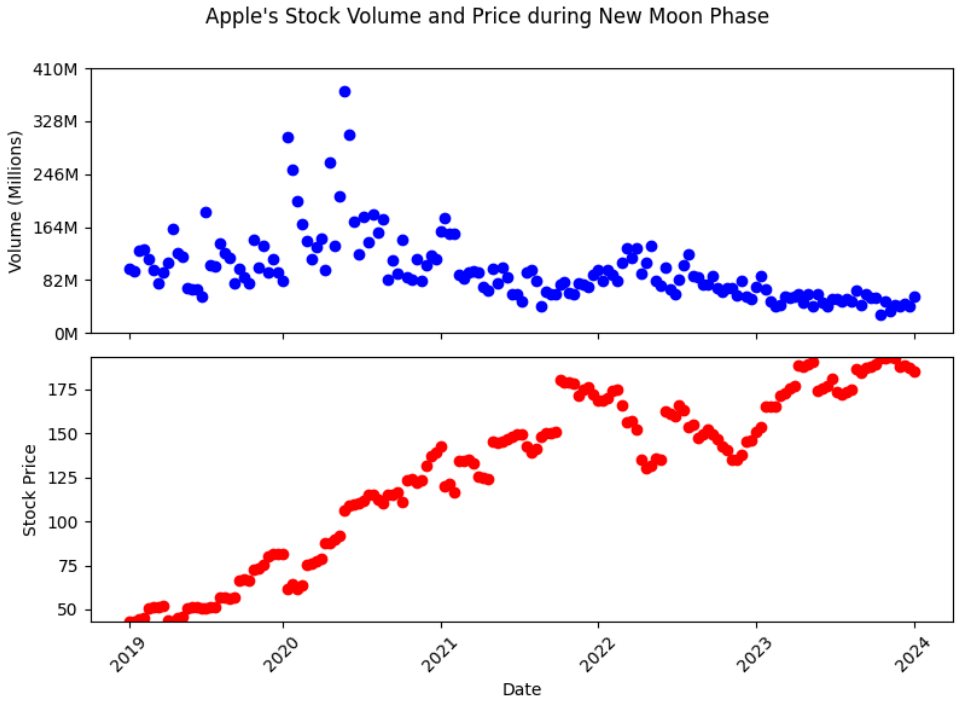
Top 5 Features Importance for ML Model **(Mustafa Ausaf):**

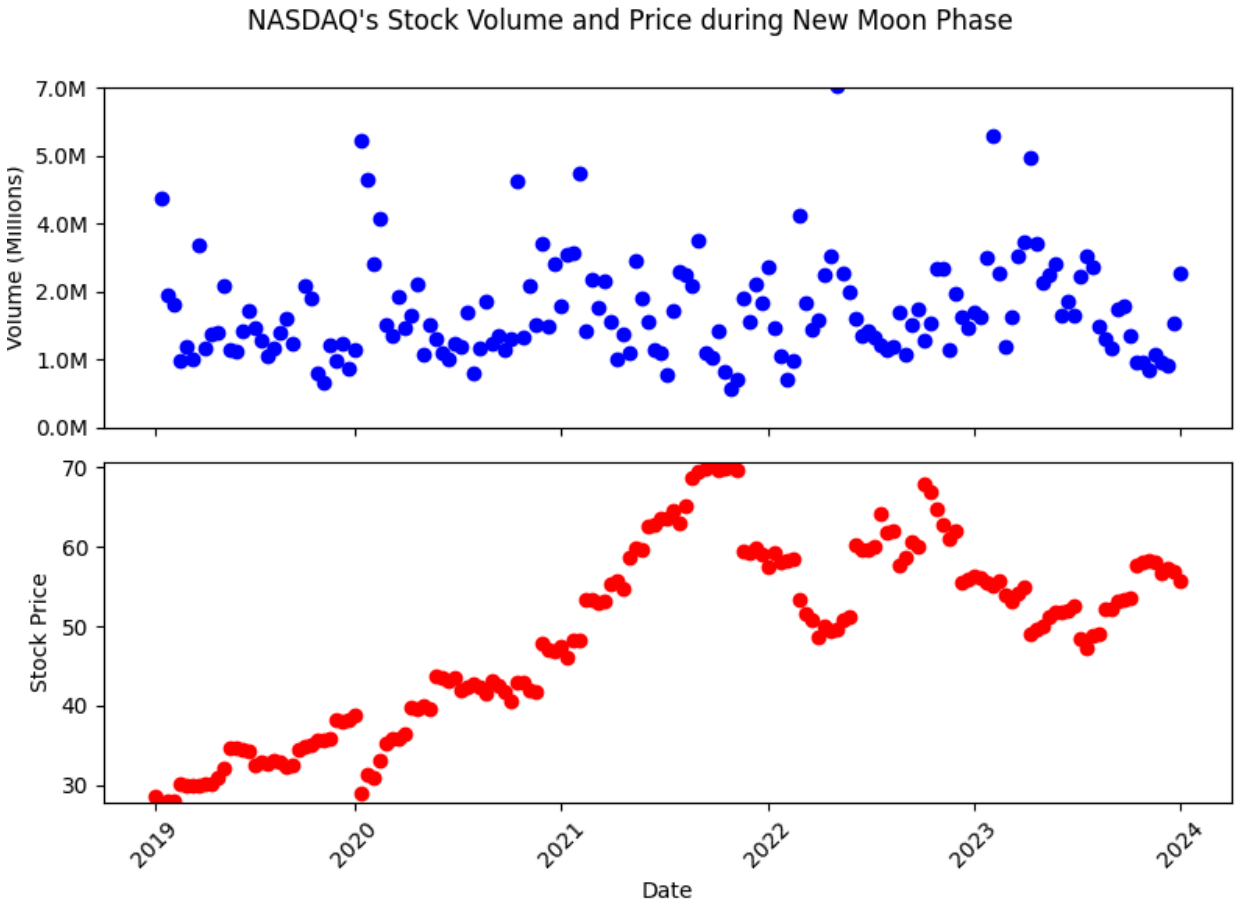
Used Random Forest Classifier with time series ranges to calculate the importance of stock features so that we know what stats are important when finding a correlation between any of them to the stock price and or moon phases and if they were accurate as well

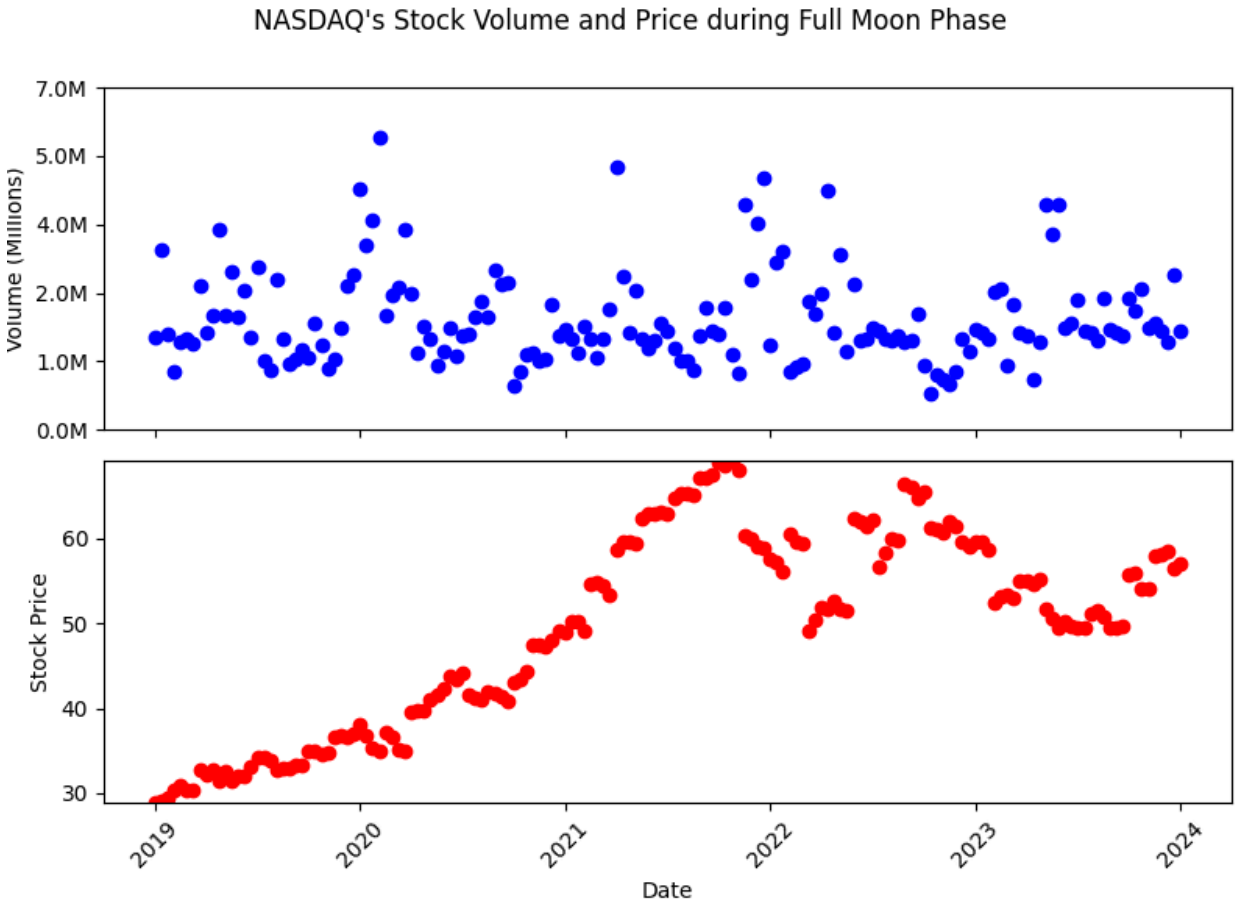
Visualizations:

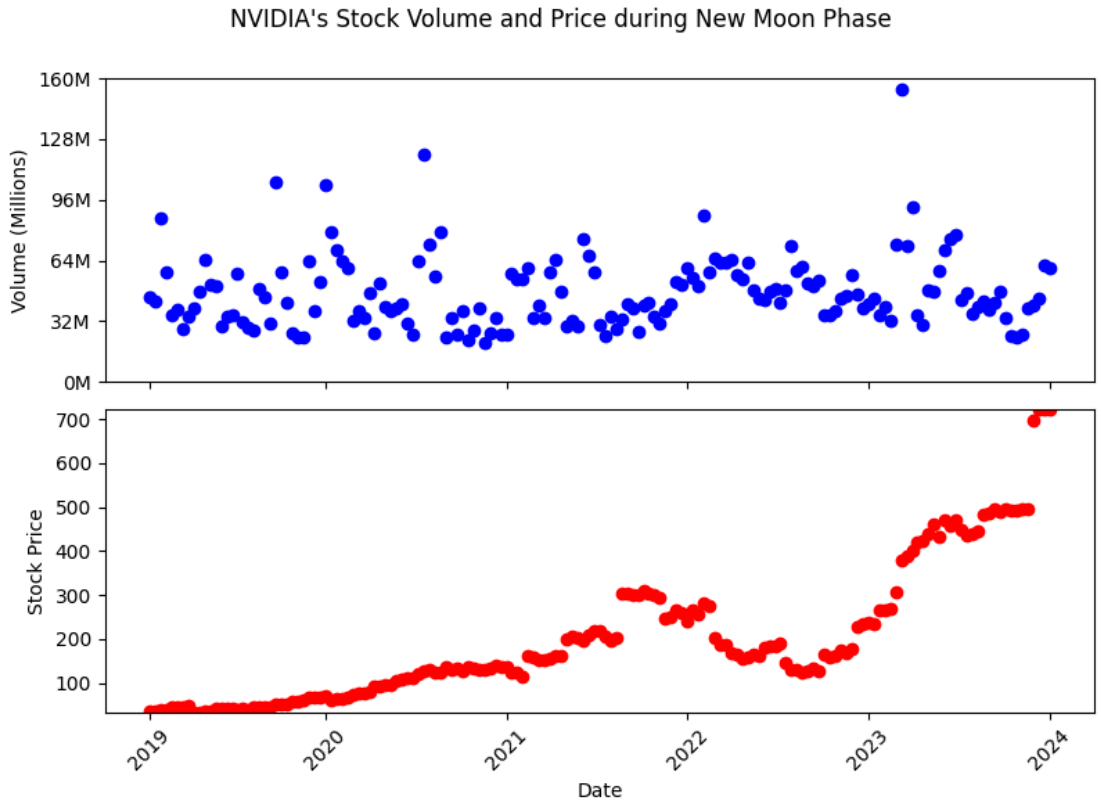
- Stock Volume vs Stock Price Correlation Visualizations **(Emilio Kiryakos)**:

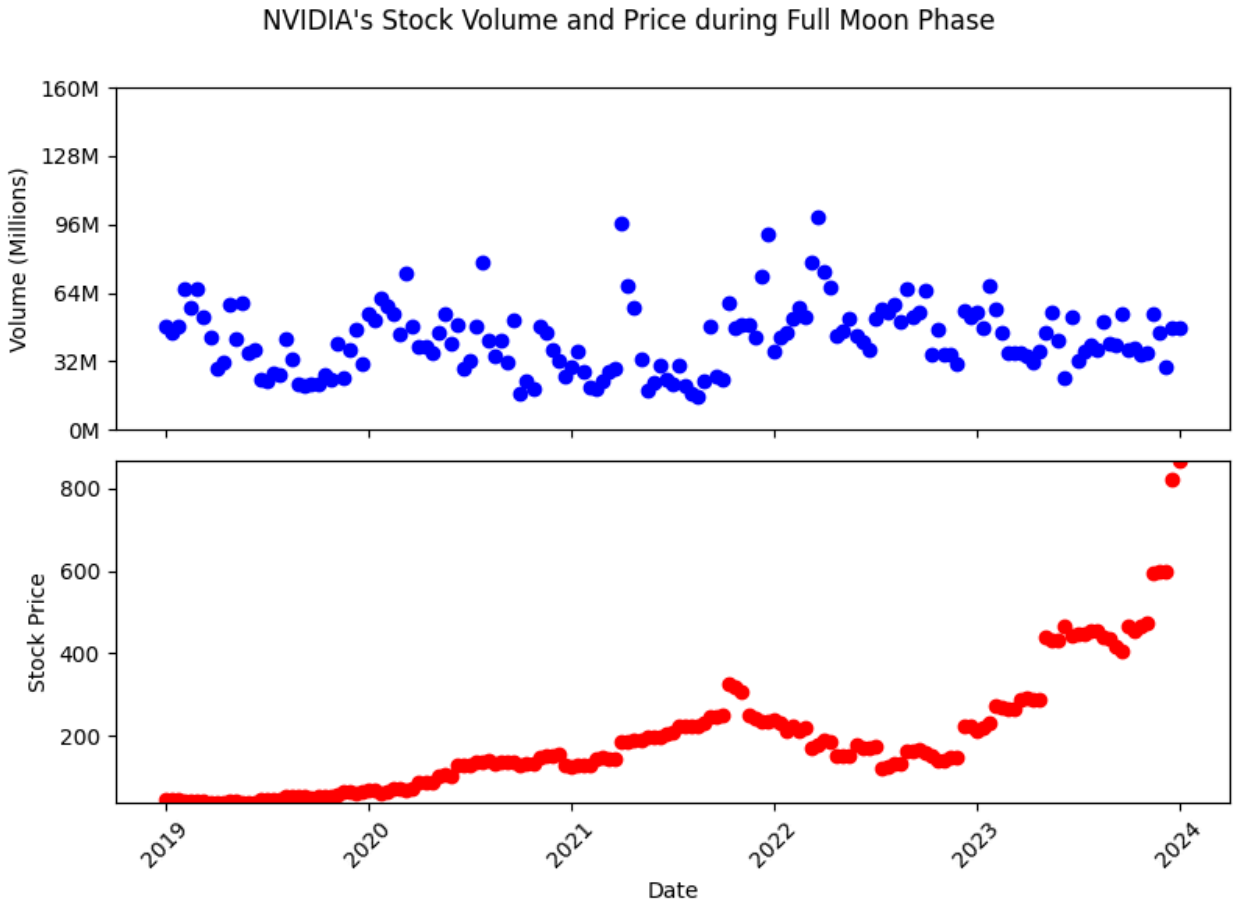


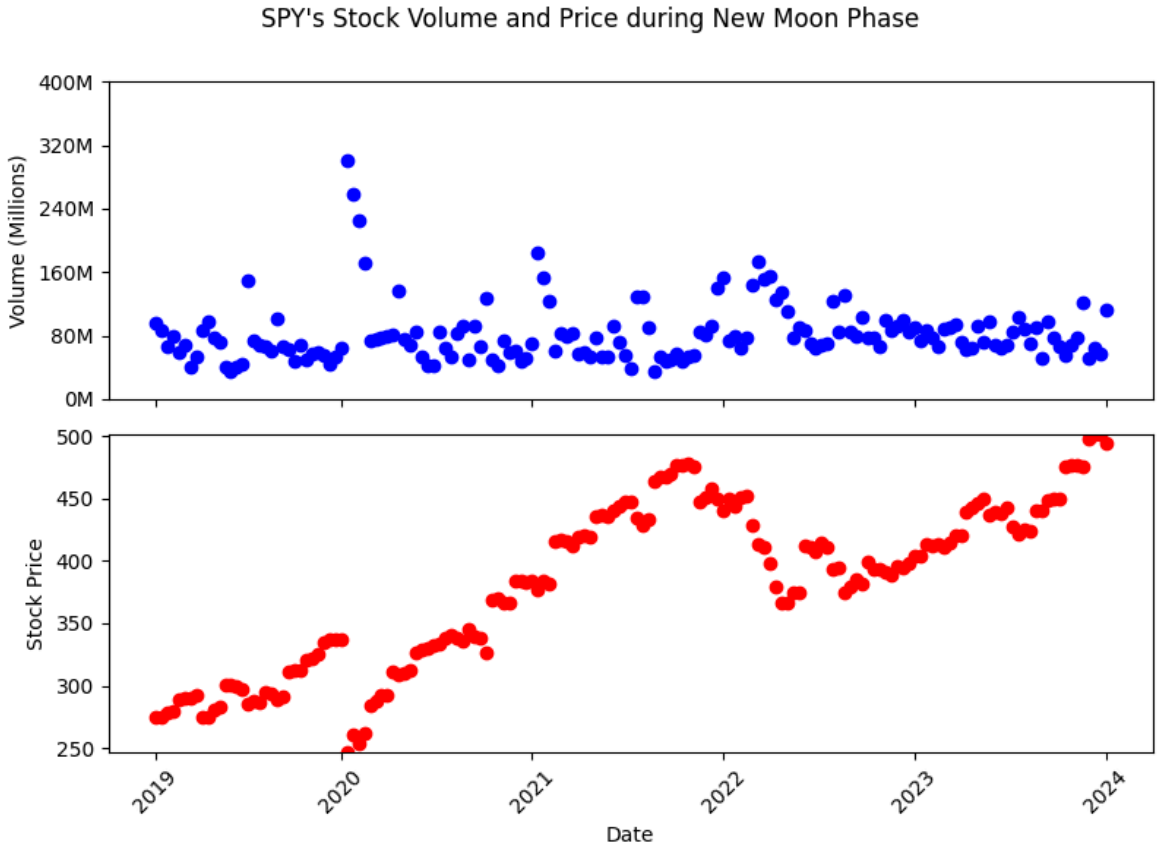


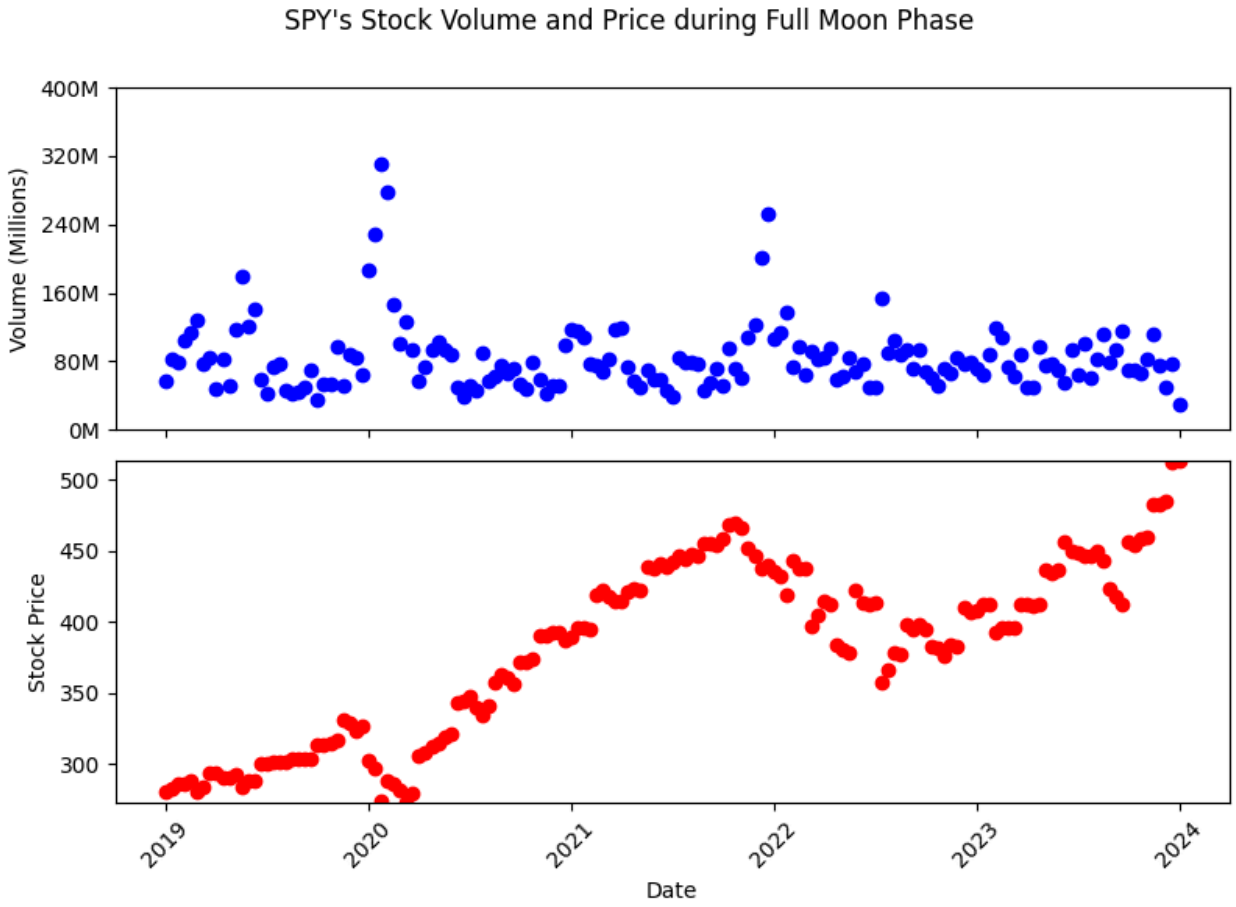

NASDAQ's Stock Volume and Price during New Moon Phase
Volume (Millions)
7.0M
5.0M
4.0M
2.0M
1.0M
0.0M
Stock Price
70
60
50
40
30
2019
2020
2021
2022
2023
2024
Date

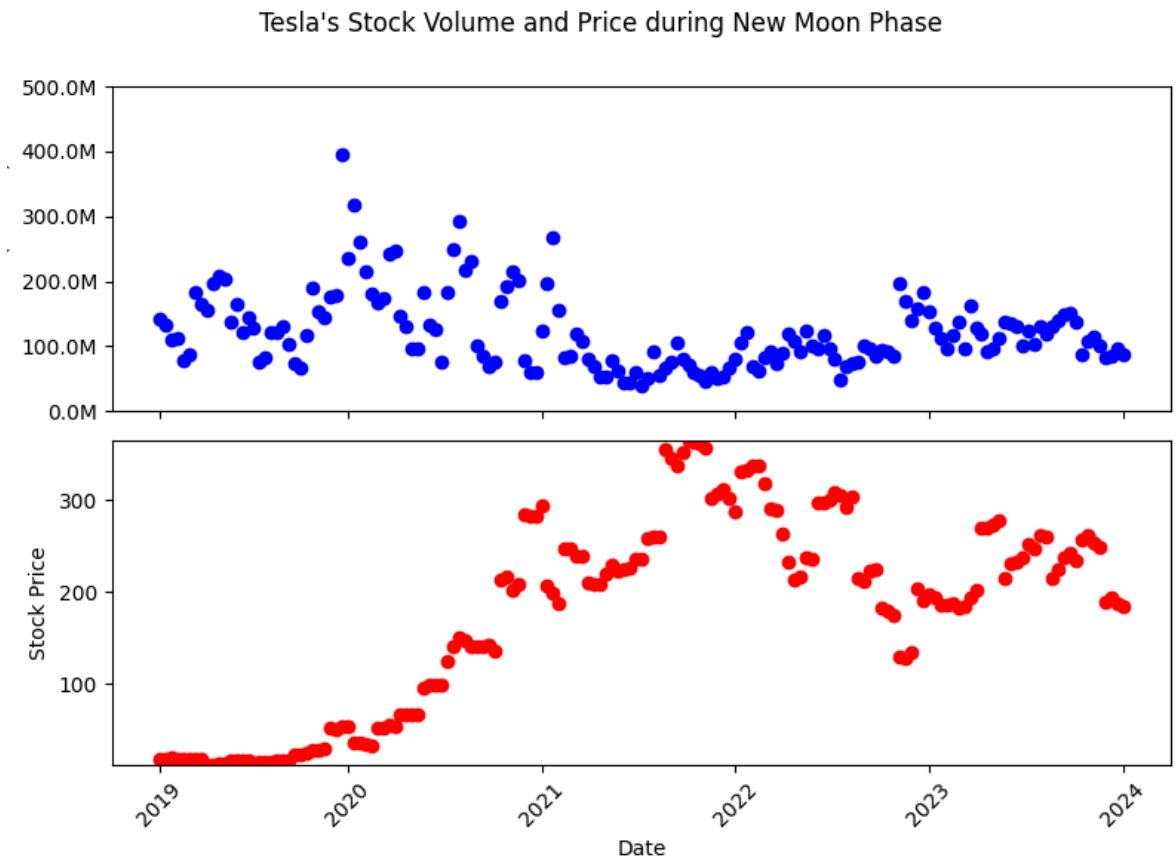

NASDAQ's Stock Volume and Price during Full Moon Phase
Volume (Millions)
7.0M
5.0M
4.0M
2.0M
1.0M
0.0M
Stock Price
60
50
40
30
2019
2020
2021
2022
2023
2024
Date

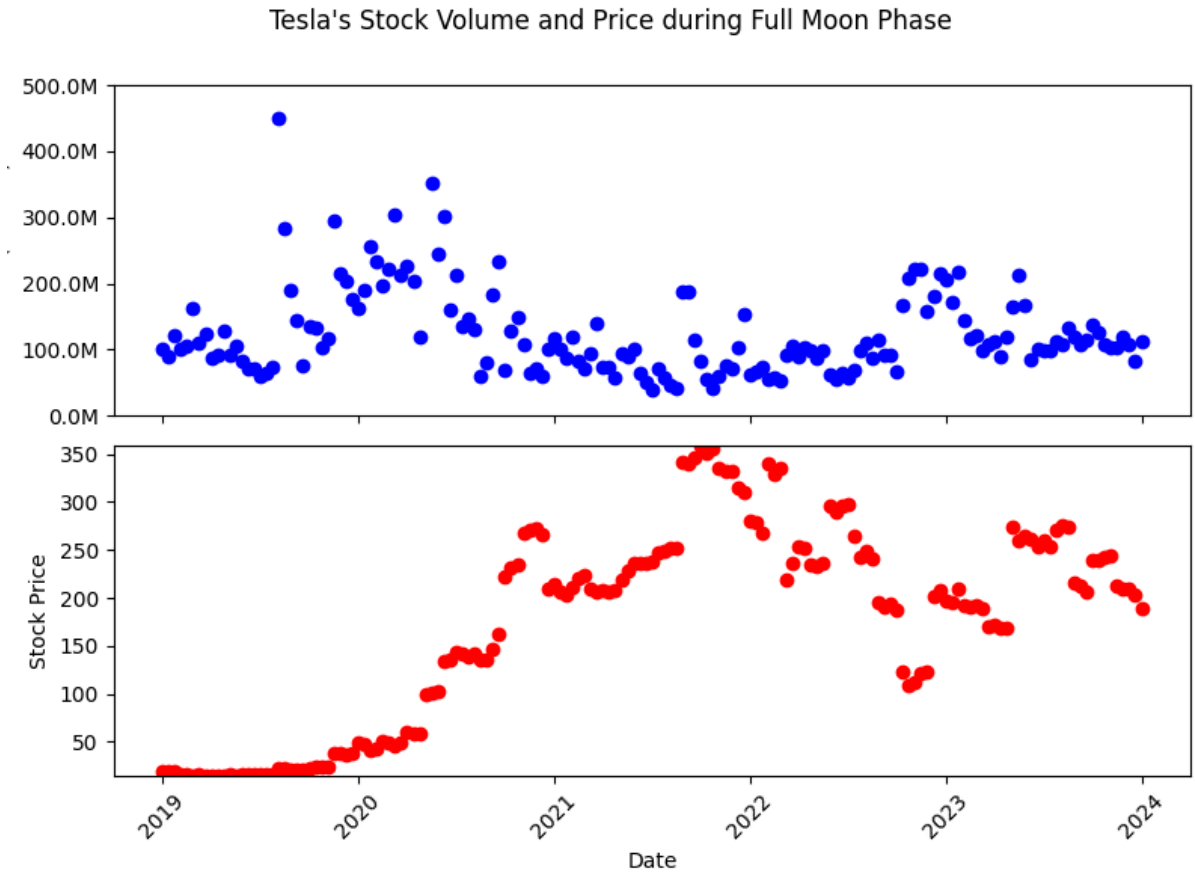

NVIDIA's Stock Volume and Price during New Moon Phase
Volume (Millions)
160M
128M
96M
64M
32M
0M
Stock Price
700
600
500
400
300
200
100
2019
2020
2021
2022
2023
2024
Date


NVIDIA's Stock Volume and Price during Full Moon Phase
Volume (Millions)
160M
128M
96M
64M
32M
0M
Stock Price
800
600
400
200
2019
2020
2021
2022
2023
2024
Date

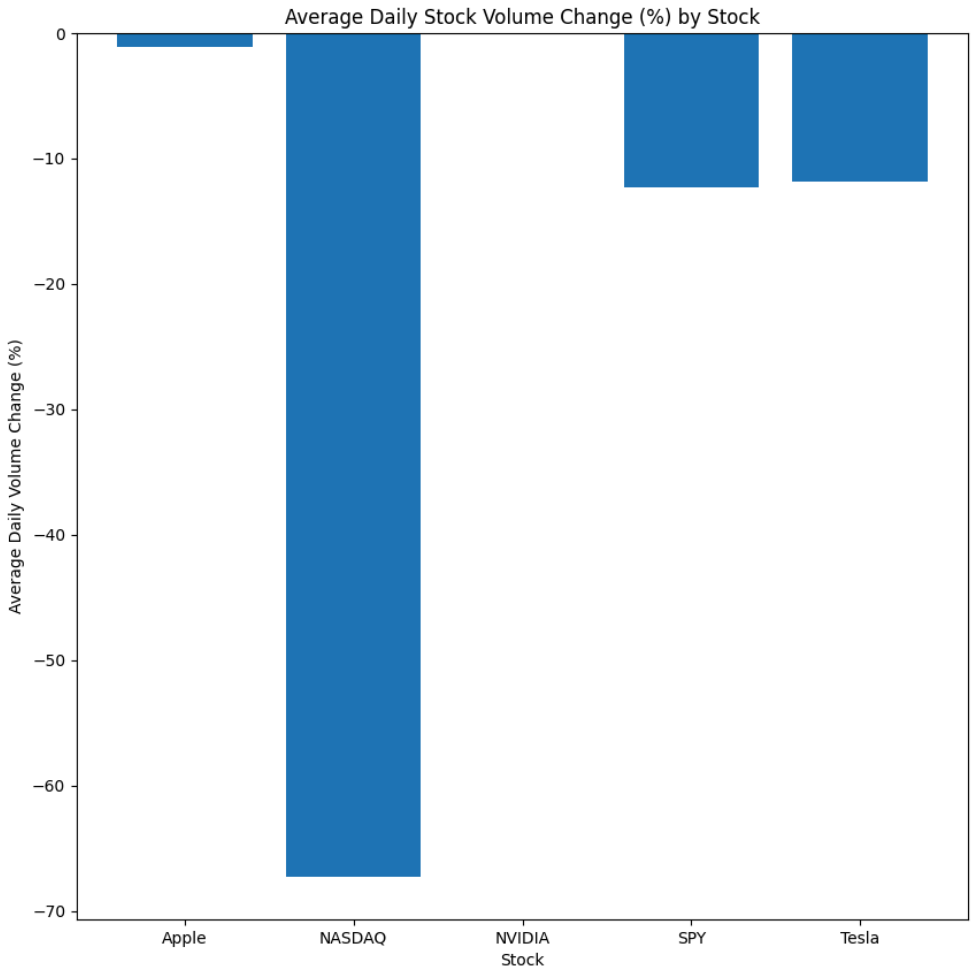

SPY's Stock Volume and Price during New Moon Phase
Volume (Millions)
400M
320M
240M
160M
80M
0M
Stock Price
500
450
400
350
300
250
2019
2020
2021
2022
2023
2024
Date


SPY's Stock Volume and Price during Full Moon Phase
Volume (Millions)
400M
320M
240M
160M
80M
0M
Stock Price
500
450
400
350
300
2019
2020
2021
2022
2023
2024
Date

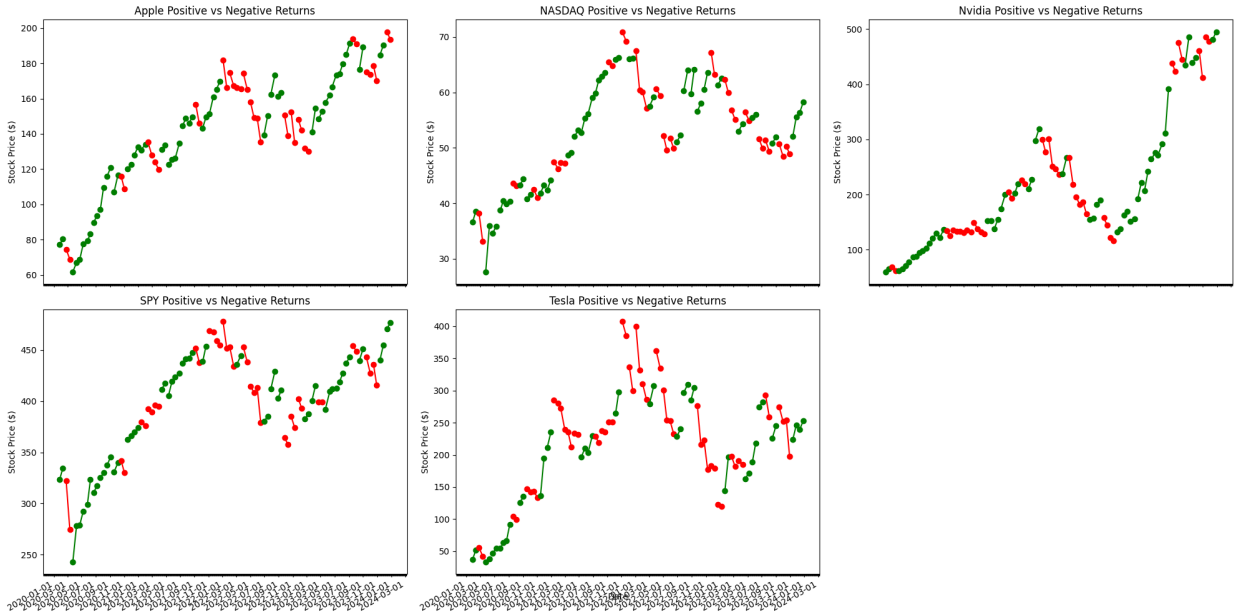

Tesla's Stock Volume and Price during New Moon Phase
500.0M
400.0M
300.0M
200.0M
100.0M
0.0M
Stock Price
300
200
100
2019
2020
2021
2022
2023
2024
Date


Tesla's Stock Volume and Price during Full Moon Phase
500.0M
400.0M
300.0M
200.0M
100.0M
0.0M
Stock Price
350
300
250
200
150
100
50
2019
2020
2021
2022
2023
2024
Date

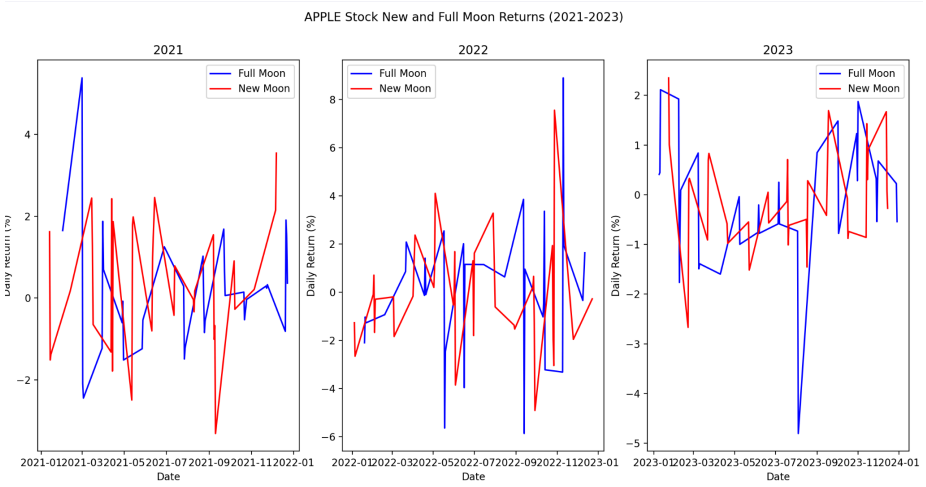
- Average Daily Stock Volume Change (%) by Stock **(Emilio Kiryakos)**:

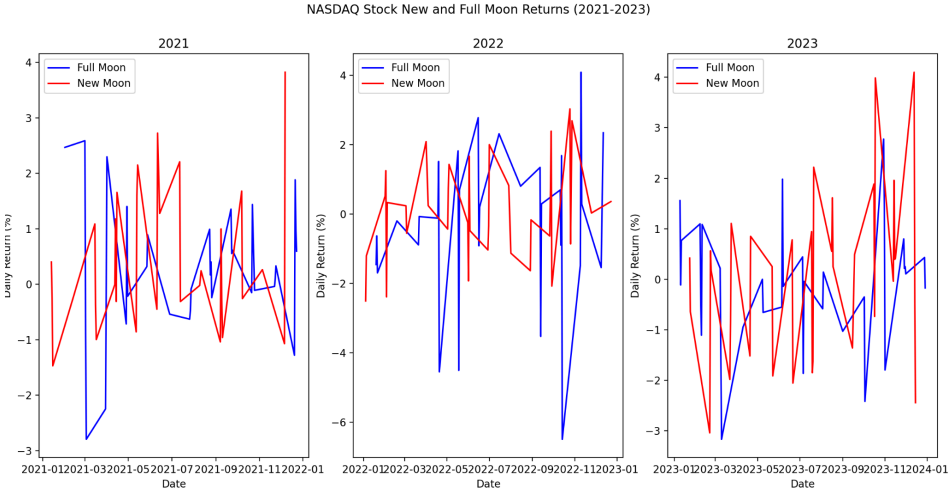


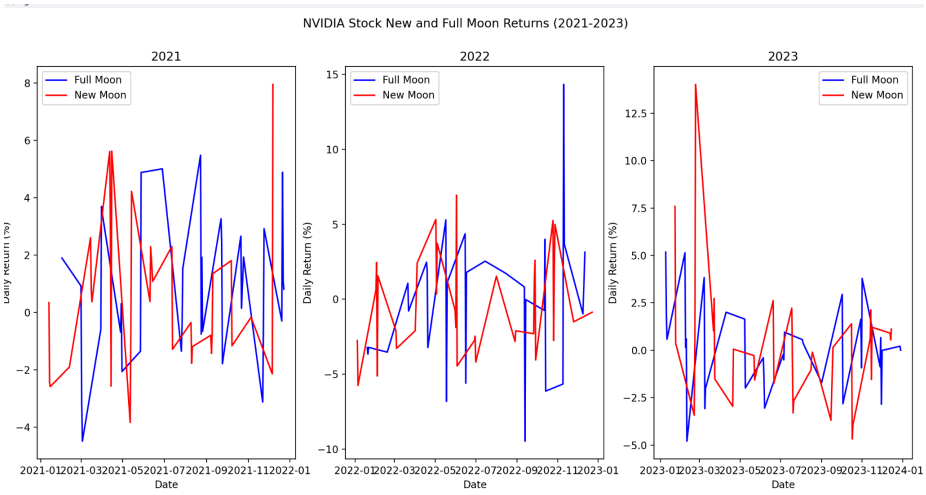
- Returns over lunar cycles (New Moon to Full Moon) **(Bart Nowobilski)**:

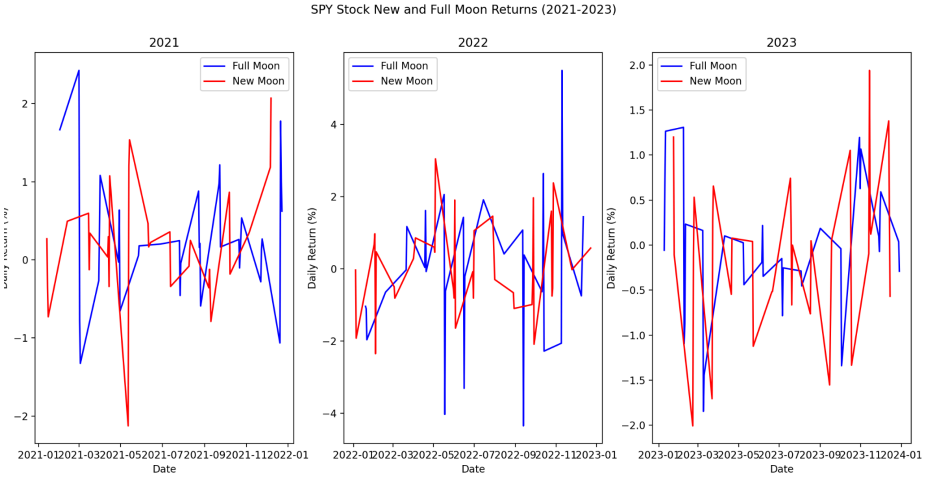


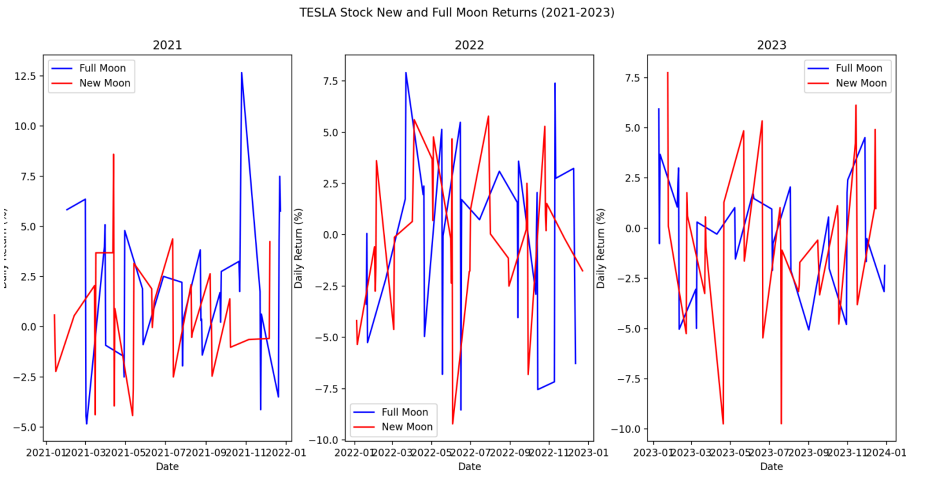
- Volatility and Profitability during a New Moon and Full Moon **(David Mendoza)**:



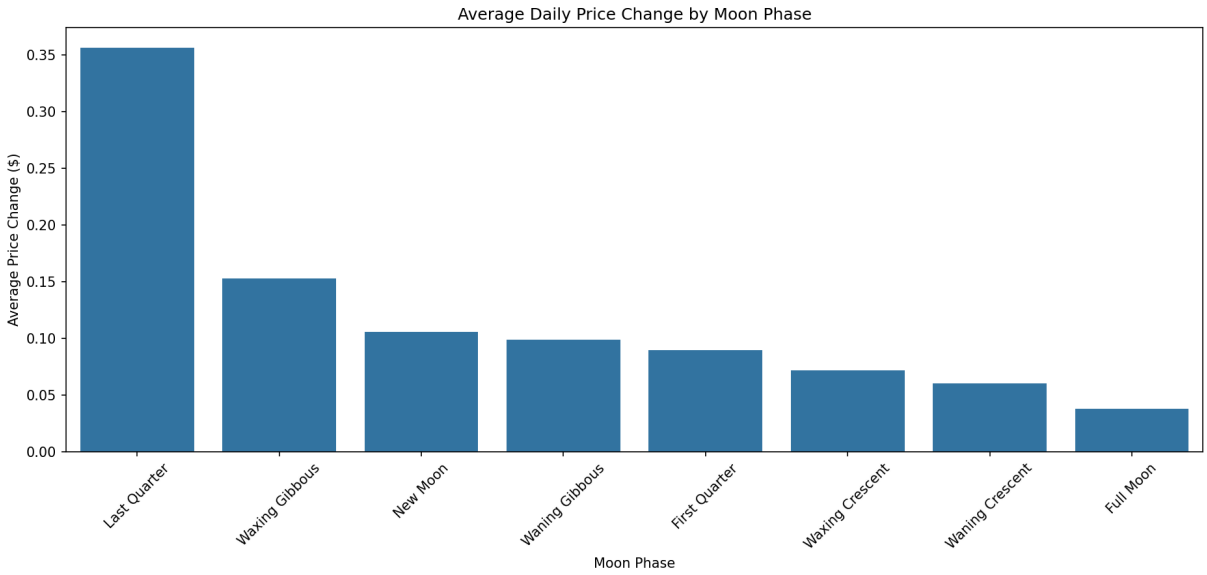




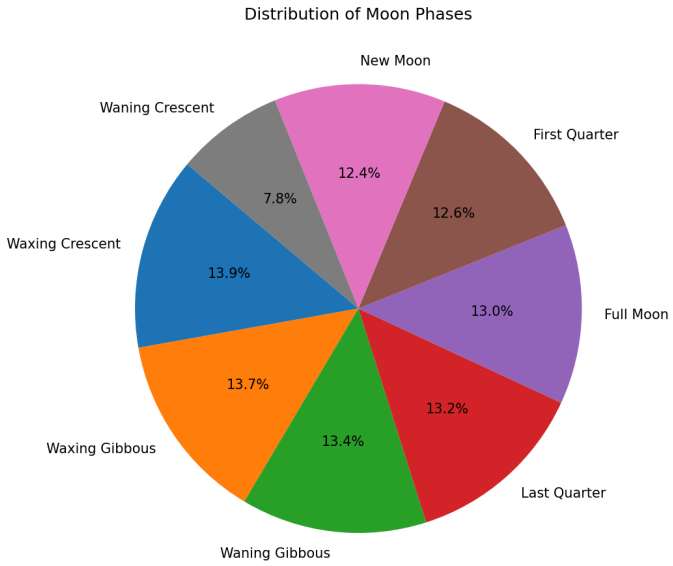




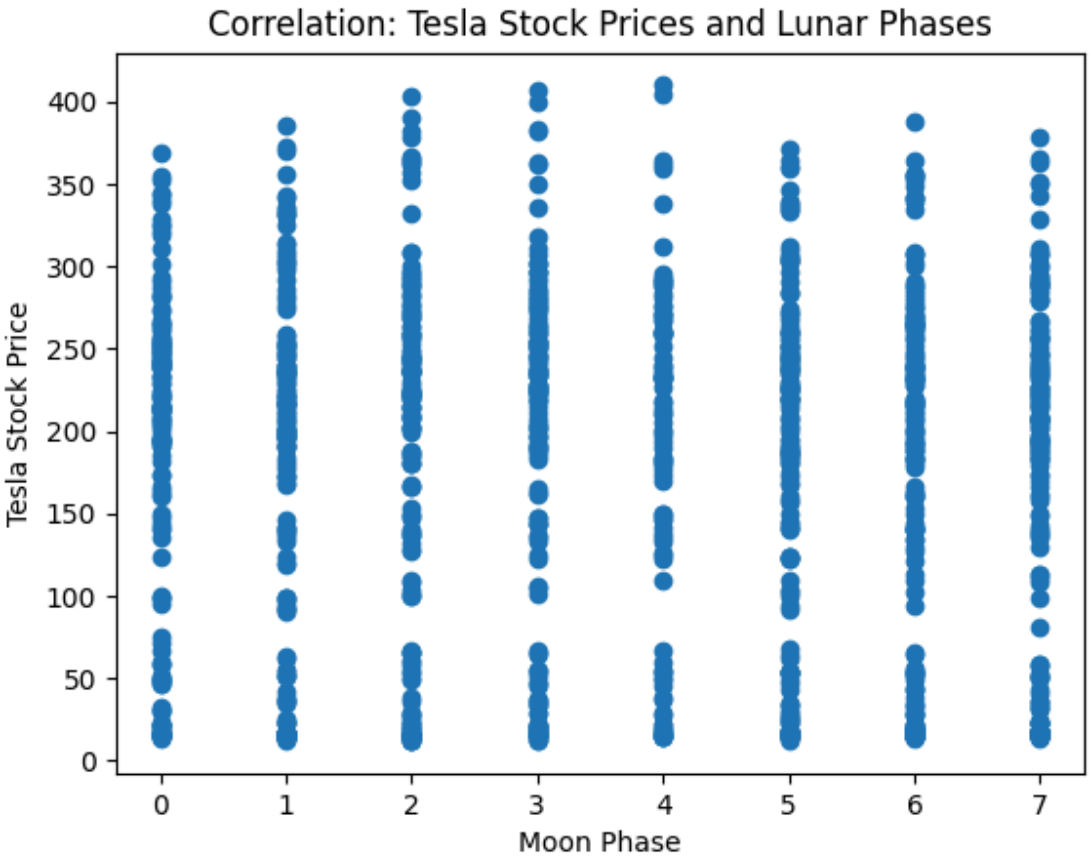
- Average Daily Price Change by Moon Phases **(Mustafa Ausaf)**:



- Distribution of Moon Phases **(Mustafa Ausaf)**:



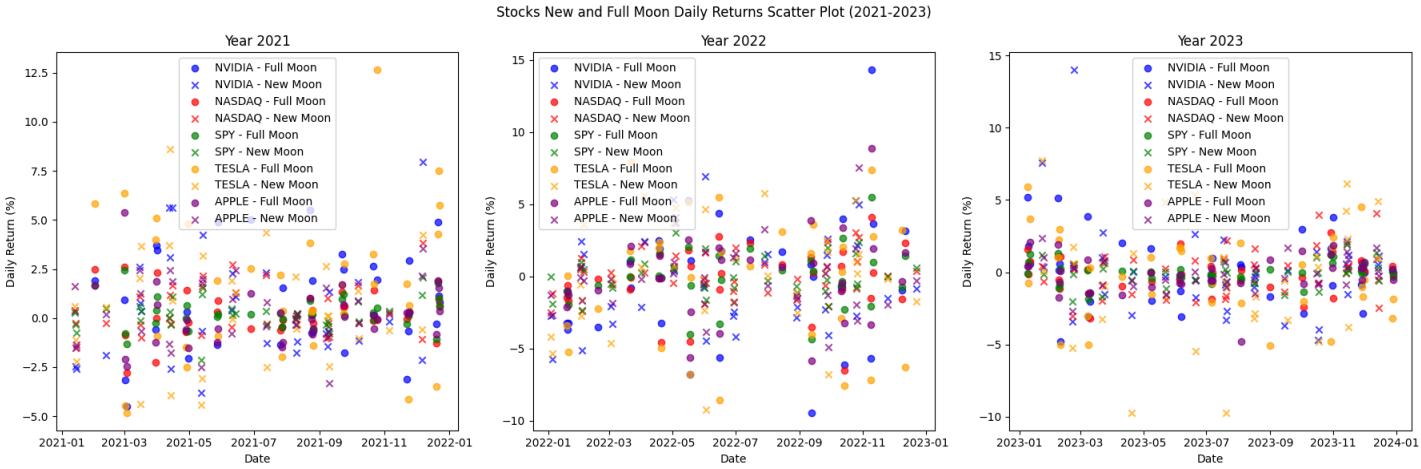
- Correlation: Tesla Stock Prices and Lunar Phases (**Adithya Jose**):

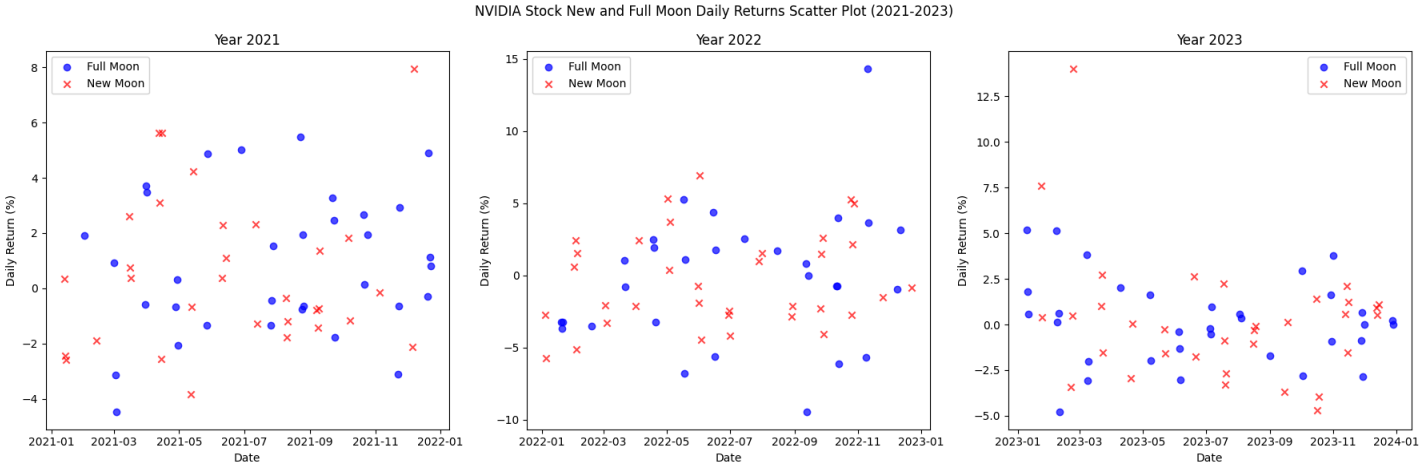


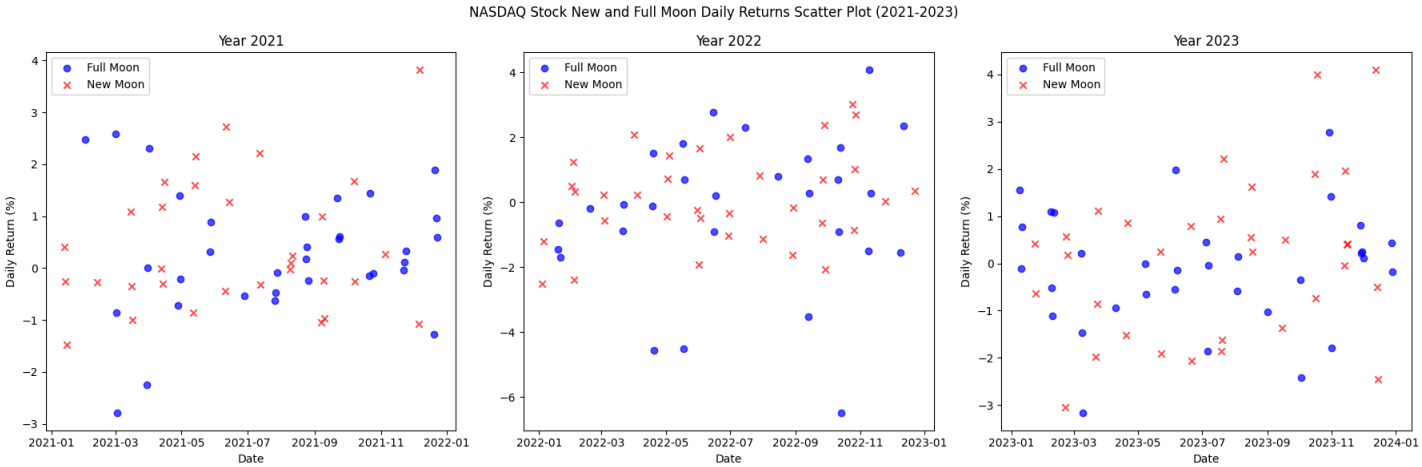
Additional Work:

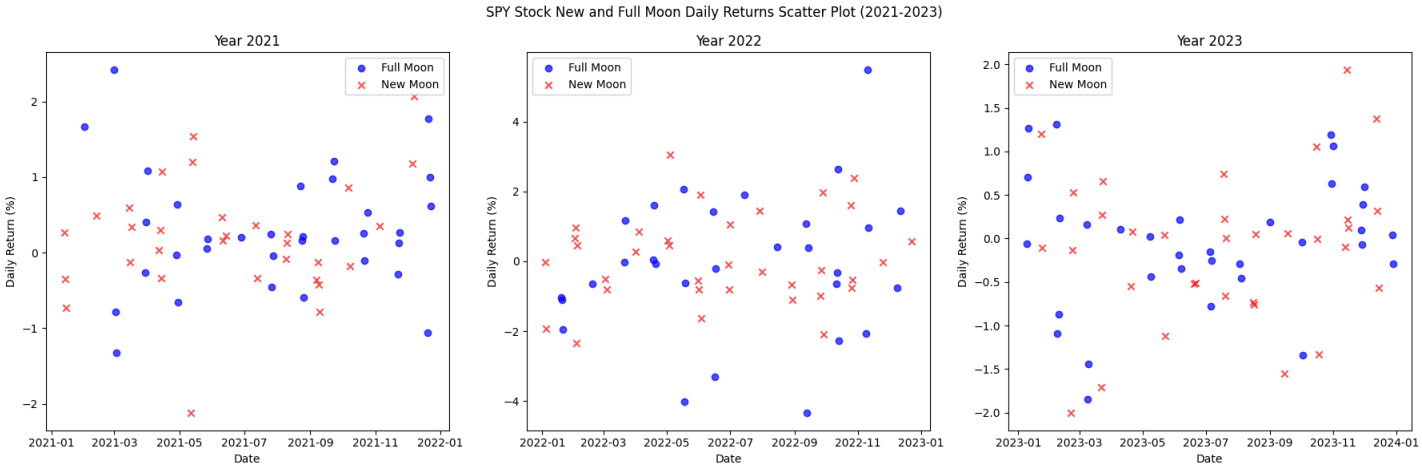
Comparison of Daily Returns During New and Full Moon Phases Across Multiple Stocks (2021-2023)

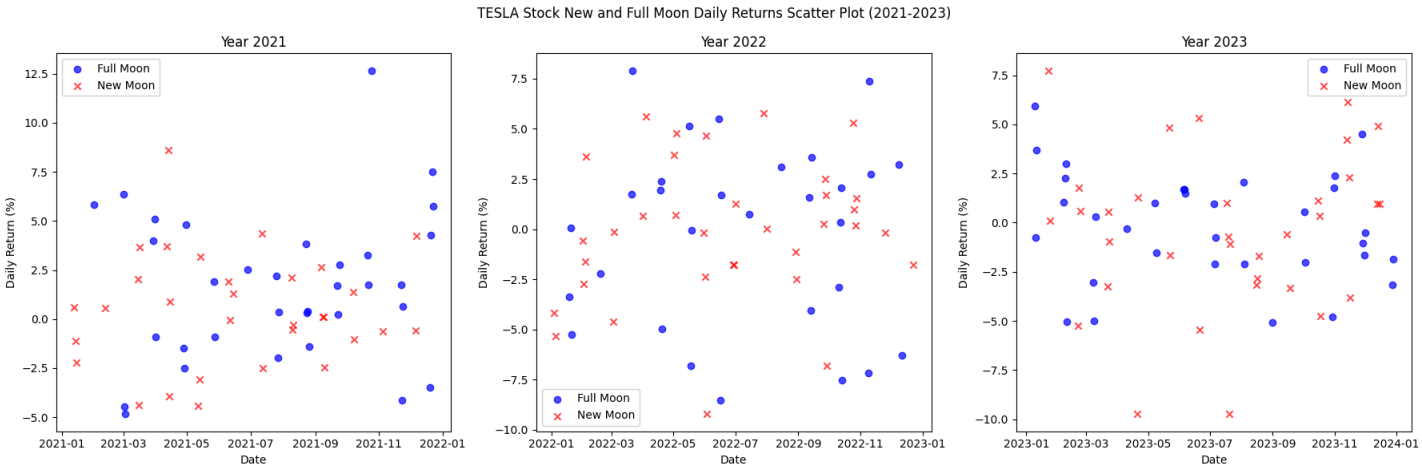
(Adithya Jose):

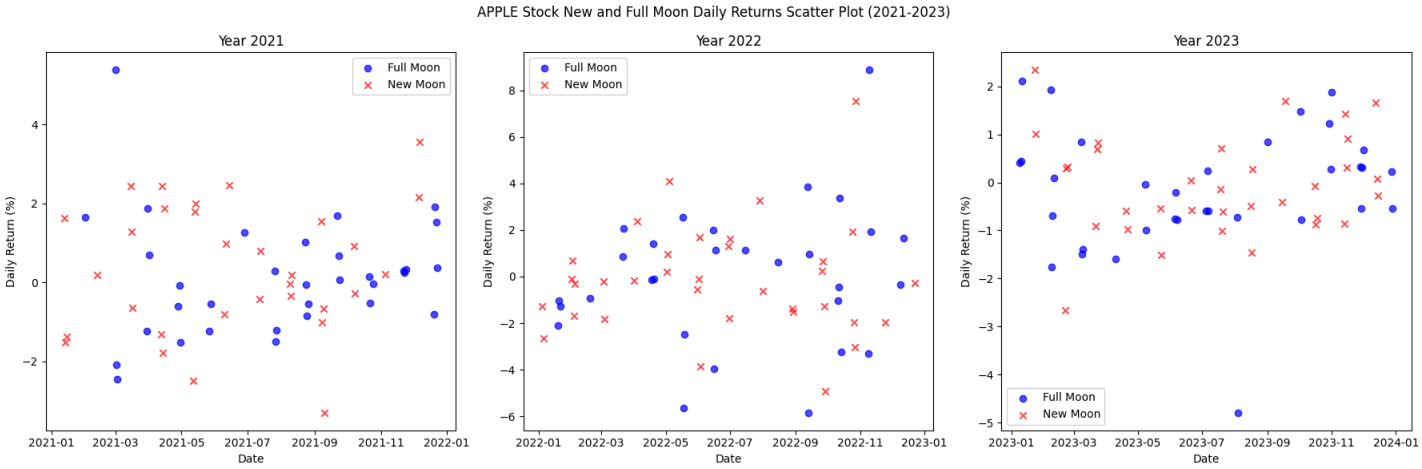



NVIDIA Stock New and Full Moon Daily Returns Scatter Plot (2021-2023)
Year 2021
Year 2022
Year 2023
Full Moon
New Moon
Daily Return (%)
Date

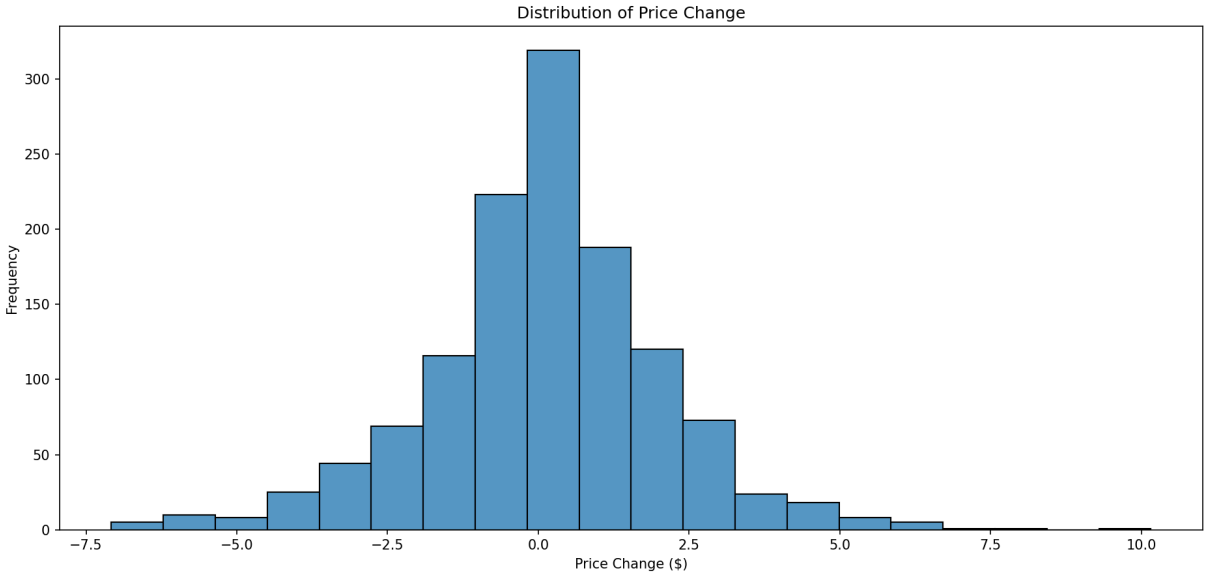

NASDAQ Stock New and Full Moon Daily Returns Scatter Plot (2021-2023)
Year 2021
Year 2022
Year 2023
Full Moon
New Moon
Daily Return (%)
Date


SPY Stock New and Full Moon Daily Returns Scatter Plot (2021-2023)
Year 2021
Year 2022
Year 2023
Full Moon
New Moon
Daily Return (%)
Date

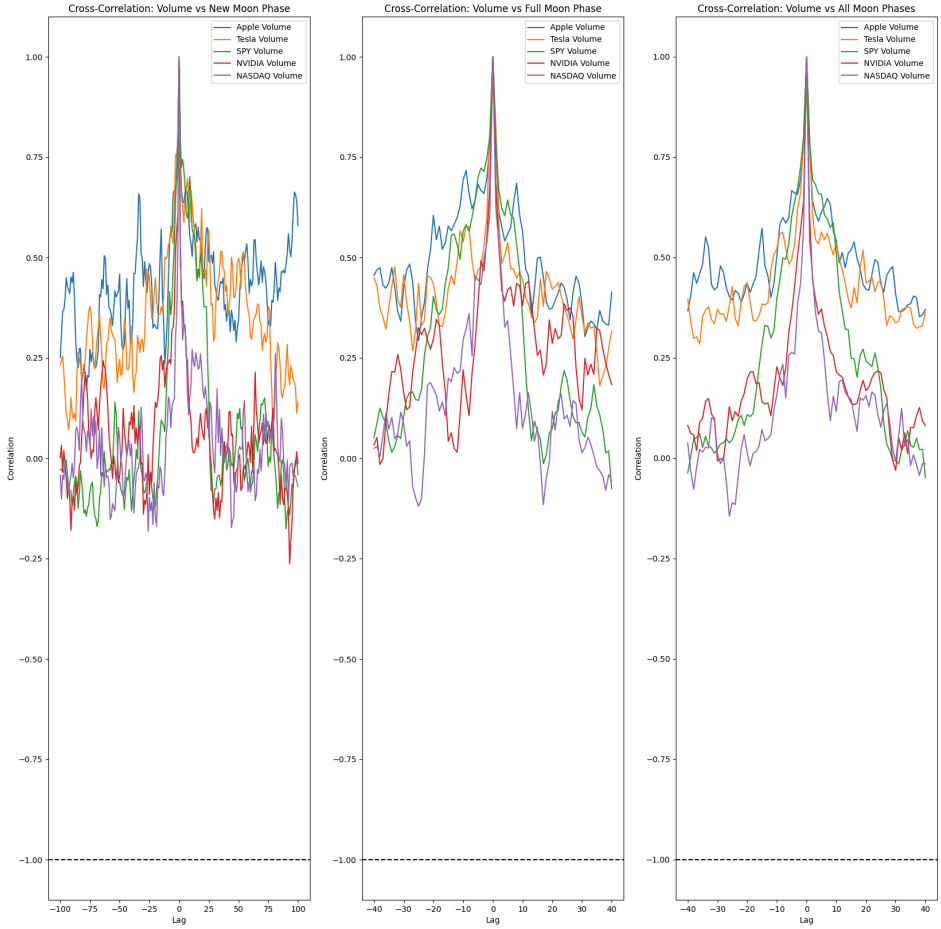




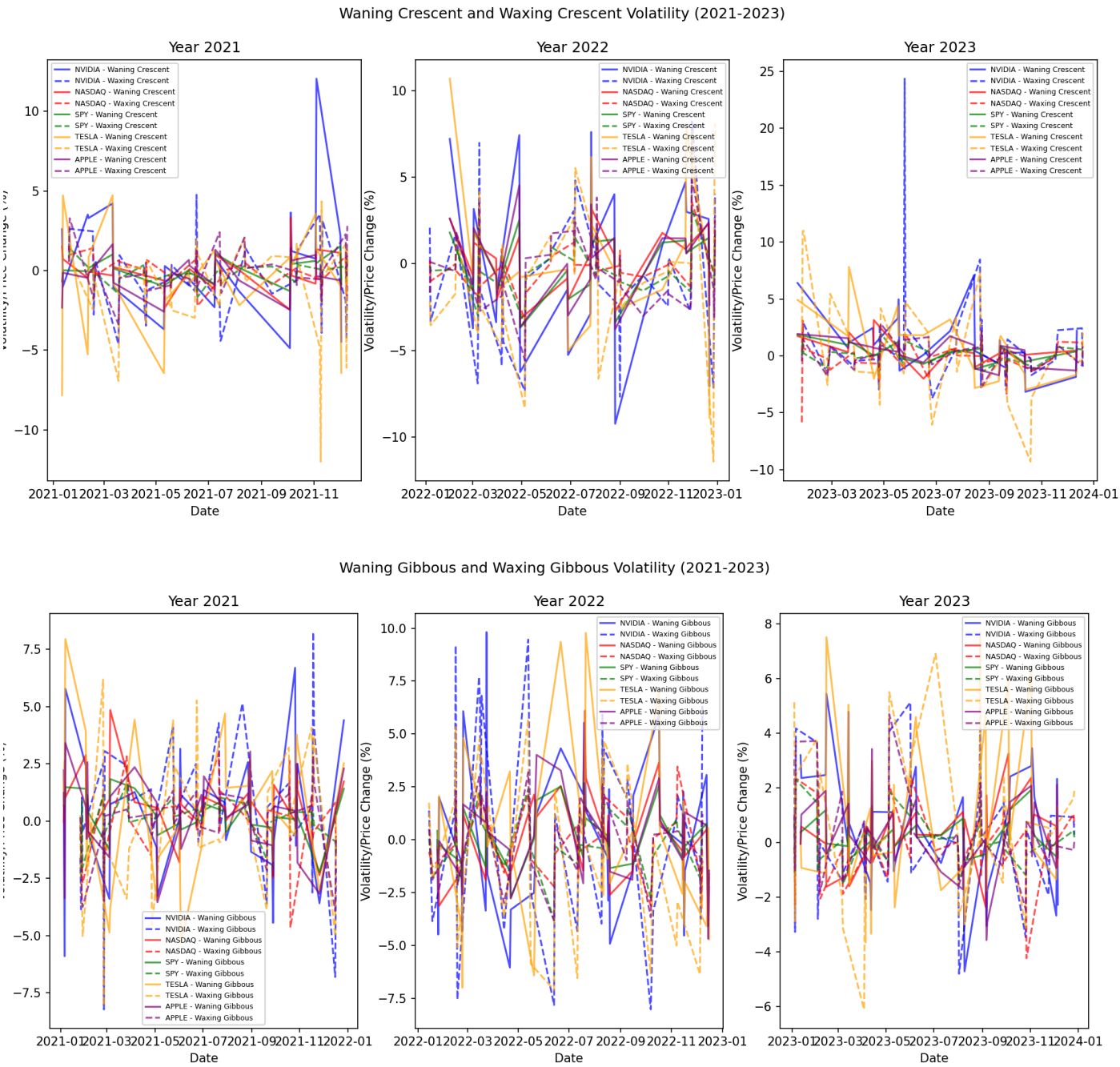
Distribution of Price Change **(Mustafa Ausaf):**



Cross-Correlation: Stocks vs New, Full and All Moon Phases over time **(Emilio Kiryakos):**



Crescent vs. Gibbous Phase Return and Volatility, All Stocks Plotted **(David Mendoza):**



Results:

- Key Findings:
  - Analyzed performance of five stocks (Apple, NASDAQ, NVIDIA, SPY, TESLA) over four years:
    - $1000 investment at each new moon and selling at each full moon yielded positive returns for all stocks we investigated.
    - Stock prices generally increased over time while stock volumes remained relatively constant, indicating sustained interest.

        - Stocks during the full moon were not only more profitable, there was also more volatility present during the full moon period than in new moon periods.
            - Stocks during Waning (after the full moon) moon phases yielded higher average daily returns, and only higher average daily returns than in Waxing (before the full moon) moon phases.
            - The price change/volatility was on average only positive during Waning moon phases.
        - Some stocks, like NVIDIA in 2022, show very high returns on specific moon phases, while in other years or different stocks, such as Tesla in 2021, the variation is less pronounced. This further supports the notion that stock-specific factors or broader market conditions are more influential than lunar cycles.
        -
    - Utilized machine learning techniques:
        - Found a small but noticeable relationship between Tesla stock prices and lunar phases, which suggests potential predictive value.
- Summary:
    - Investigation into lunar phases and stock behavior:
        - Indicates potential relationships worth further exploration.
        - Additional research and data refinement needed for validation and understanding.
- What worked and what we could have improved:
    - What worked:
        - Utilization of the Timeanddate and Polygon.io APIs for data collection of the moon and stocks within the past 5 years.
        - Found factors of stock such as volume that had positive correlation with stock price
    - What we could have improved:
        - Collect data for more stocks than the 5 we chose to see if the correlation that we concluded still holds to be true for other stocks.
        - Look into more factors and stock features and see if they also had any positive correlation or not to better determine the result.
        - Could have looked into more past data from over 5 years ago if we were able to get the subscriptions associated with the frameworks and libraries.

Link to Repository: https://github.com/BartNowo/LunarPhaseStockExploration