PROJECT 1
STOCK-MARKET ANALYSIS

Sonu Sharma Mete Ozkazanc
Brandon Allen Ryan Goda

Introduction Summary

What type of data are we looking at and what is the timeframe of our data?

- In this project, we will be looking at a portion of NASDAQ stock data ranging from 2019 – 2024.

Where did we acquire our data?

- We acquired our CSV data set from "Yahoo-Finance" and acquired our outlier events from "Morningstar", "Google Finance", "New York Times", and "MarketWatch".

What Analysis / Plots / Time-Series information will we be running on our data?

- In this project, we utilized merging, forecasting, model-prediction, data cleaning, maximums, minimum, averages, correlations, to-datetime, and re-indexing.

Points regarding news of speculation vs. authentic news:

- When dealing with markets, it should be noted that not all news sources are accurate in their research, and many speculate news. One shouldn't make investments based on fabricated information.

Were there any expected outliers in the data?

- From our dataset, we expect outlier events to include COVID-19, Global Conflicts, and Interest Rate fluctuations.

NQ History

NQ = NASDAQ 100: Stock Market index made up of 100 of the largest non-financial companies listed in the NASDAQ stock exchange.

- Excludes the financial companies (banks) but includes the major technology companies like Google, Amazon, Meta, etc.
- Other areas of focus besides technology are retail, biotechnology & telecommunications; however the major tech companies (magnificent 7) have the majority impact on how the index performs.
- Created on January 31, 1985.
- Trading products included in NQ include ETFs (QQQ & TQQQ), options, futures, etc.
- The index is a important measure to not only the American economic status, but the global landscape and primary concerns trends in tech & innovation industries.
- In scope of our analysis, we are interested in creating a dashboard allowing users to input their required rates of return based on their "risk aversion" and evaluate these inputs against more API Keys and economic considerations. Along with this, we would like to program a "holiday-exemption" allowing the forecast to become aware and skip over forecasting data of known U.S. holidays, allowing for a more-accurate forecast.

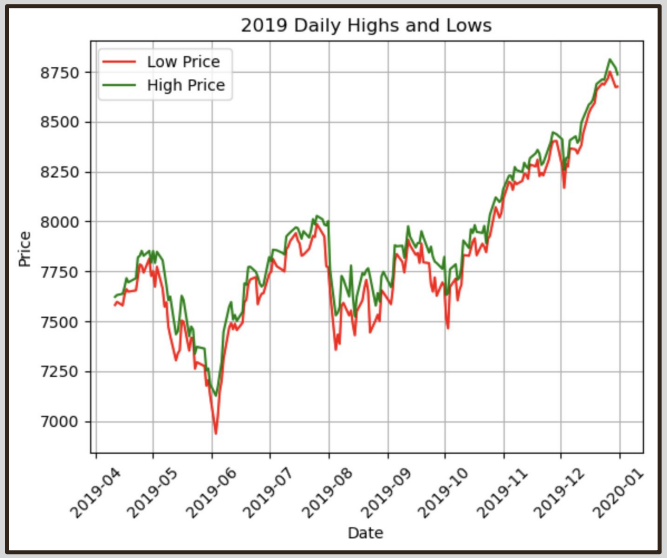
Correlating NQ with the New York Times

1. Connect to the New York Times API
   - Pull the last five years of front page headlines that relate to “United States Economy”
   - The headlines also came with between one and five additional associated keywords.
2. Compute the days with the largest Nasdaq drops and surges (largest daily deltas)
3. Combining the data, what were the headlines on a day of great change?

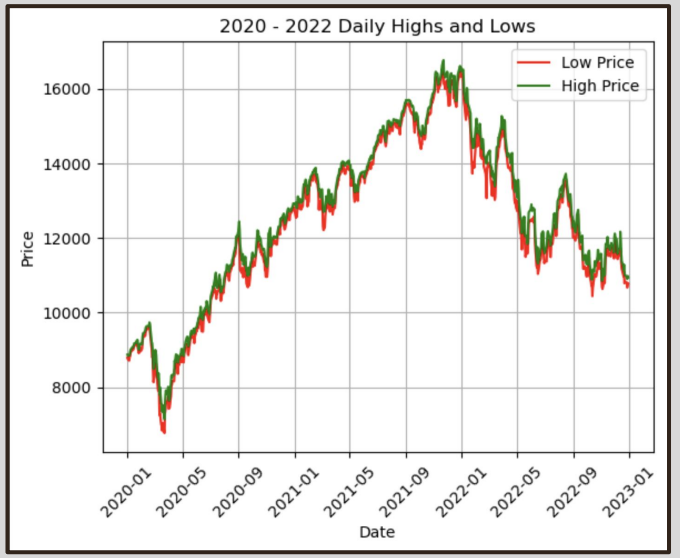
Date	Daily Change	headline
2020-02-24	-399.089844	Fears About Coronavirus Rattle World’s Markets
2020-03-02	388.919921	Economists Worldwide Cut Outlook On Growth
2020-03-12	-416.420410	Beyond Bears, Markets Show Strange Signs
2020-03-12	-416.420410	Worst Rout for Wall Street Since 1987 Crash
2020-03-16	-435.059571	Trump Urges Limits as Virus Batters Economy
2020-03-16	-435.059571	Markets Plunge as Investors Fear Cataclysmic ...
2020-03-16	-435.059571	Crisis Shutting Down Commerce, Likely for Months

keywords	
Inflation (Economics)	36
Interest Rates	29
Coronavirus (2019-nCoV)	23
Recession and Depression	18
Prices (Fares, Fees and Rates)	18
United States Politics and Government	16
Banking and Financial Institutions	16
Stocks and Bonds	15
Labor and Jobs	14
Standard & Poor's 500-Stock Index	11
Wages and Salaries	11
International Trade and World Market	7
Government Bonds	6
Shutdowns (Institutional)	6
Consumer Behavior	6
Credit and Debt	6
Mortgages	5
Consumer Price Index	5
Real Estate and Housing (Residential)	5
Quarantines	5

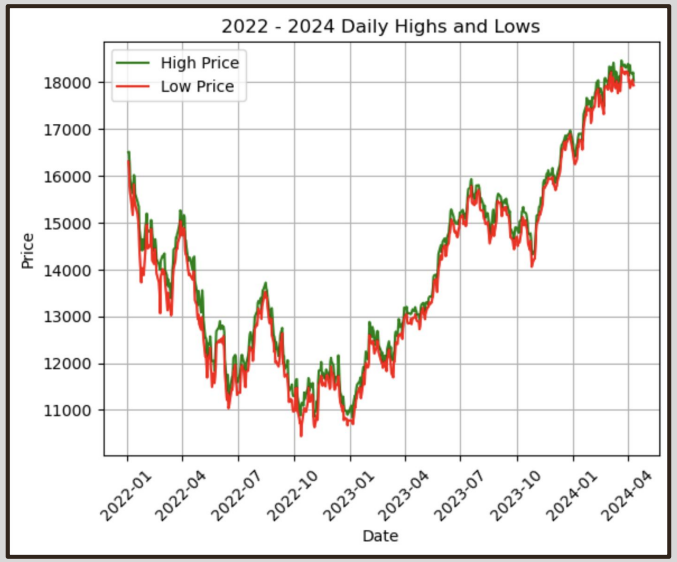
Plots & Analysis #1:



Our "2019 Daily Highs and Lows" data indicates a significant upward trend beginning in the month of October. Notably, the lowest price of all the data occurs in June.

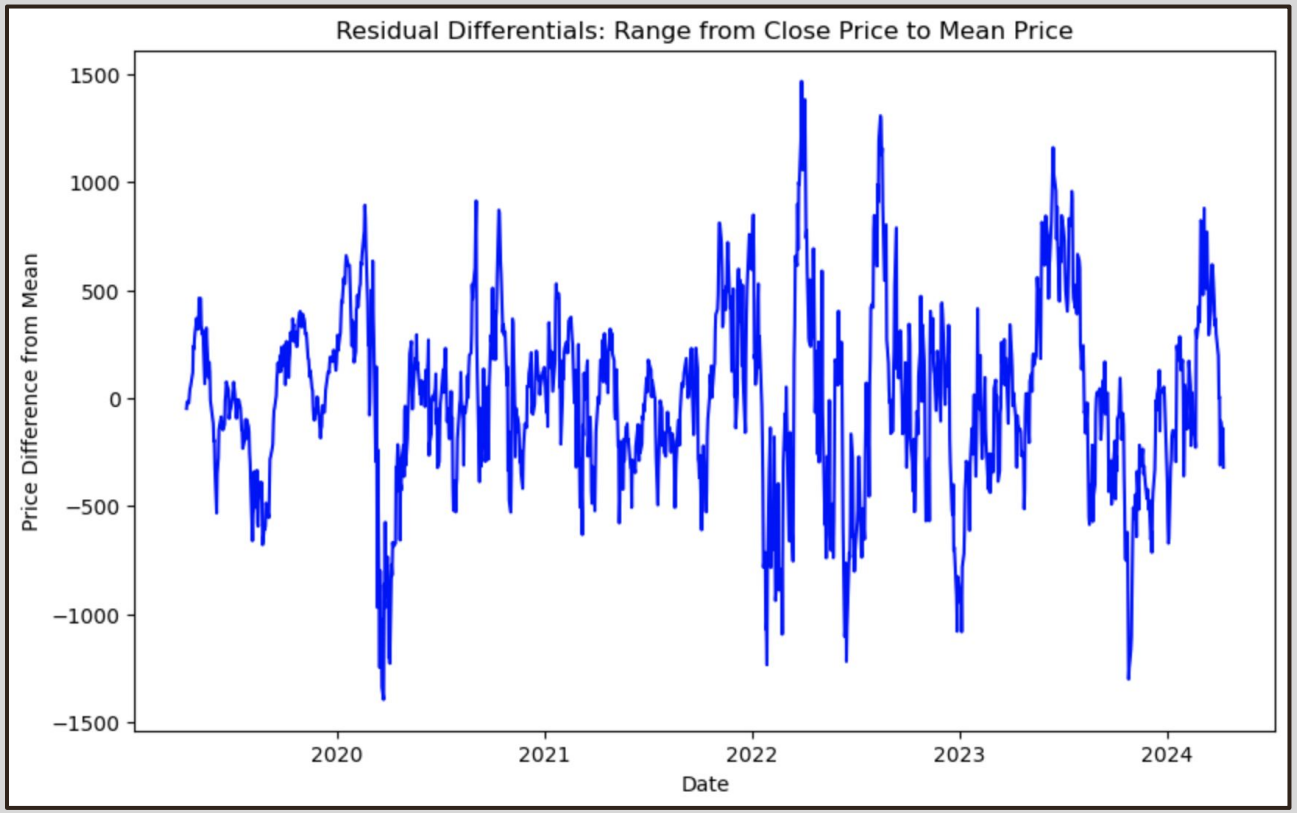


Our "2020 - 2022 Daily Highs and Lows" data continues from the 2019 data; the upward trend remaining through most of the two years. This unusual as historically, the trend upward begins in October instead of April.



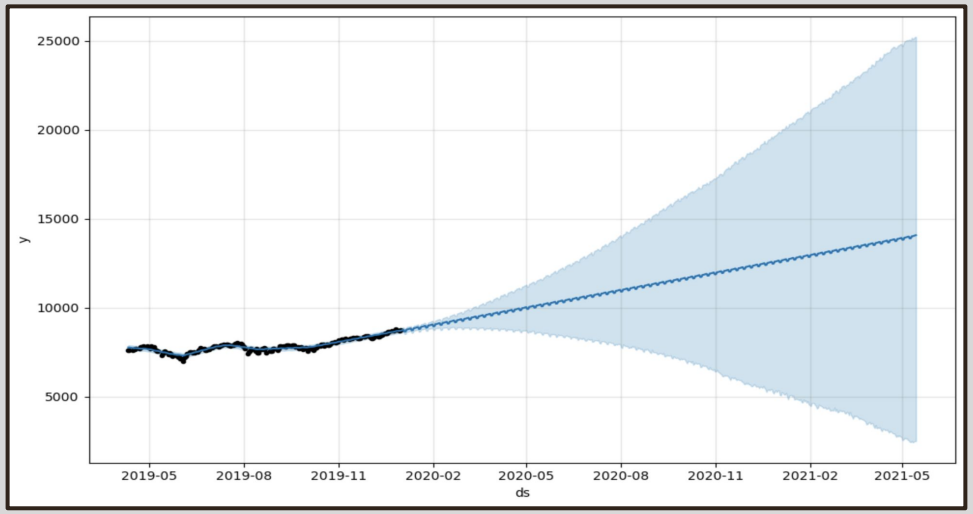
Our "2022 - 2024 Daily Highs and Lows" data continues from the 2020 - 2022 data; there is now a steep downwards trend for essentially the entirety of 2022, to which the market begins recovering at the beginning of 2023 to 2024. The NASDAQ also broke an all-time high in 04/2024.

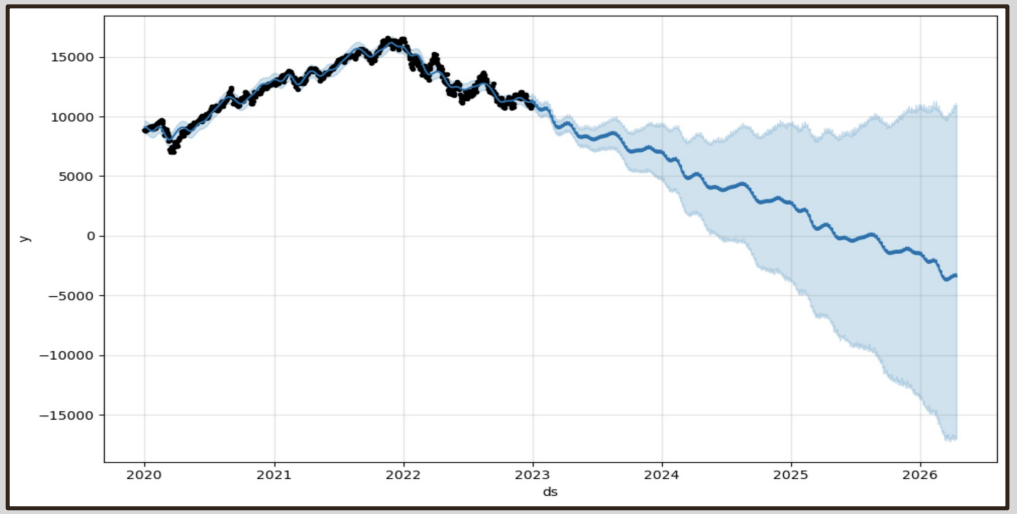
Plots and Analysis #2:

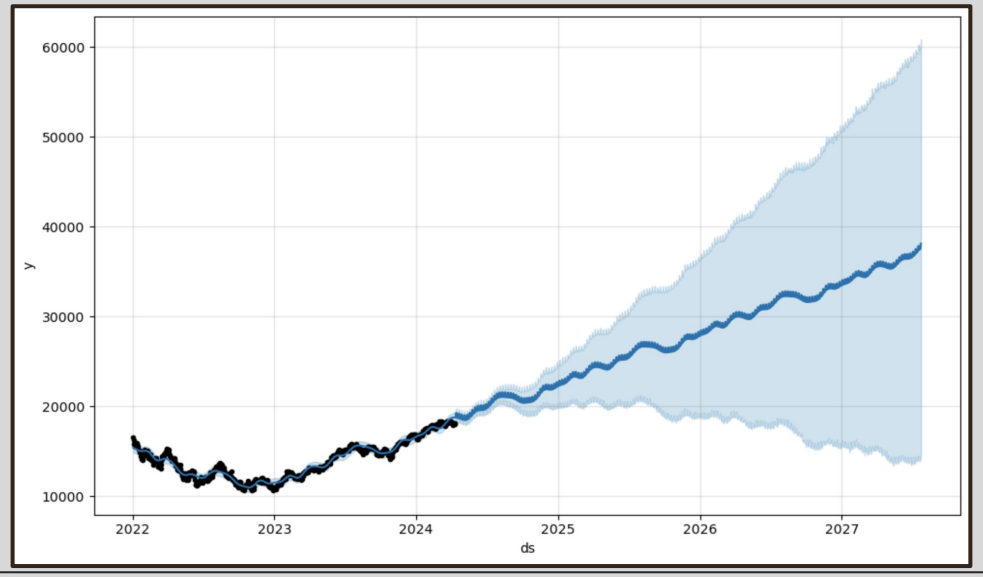


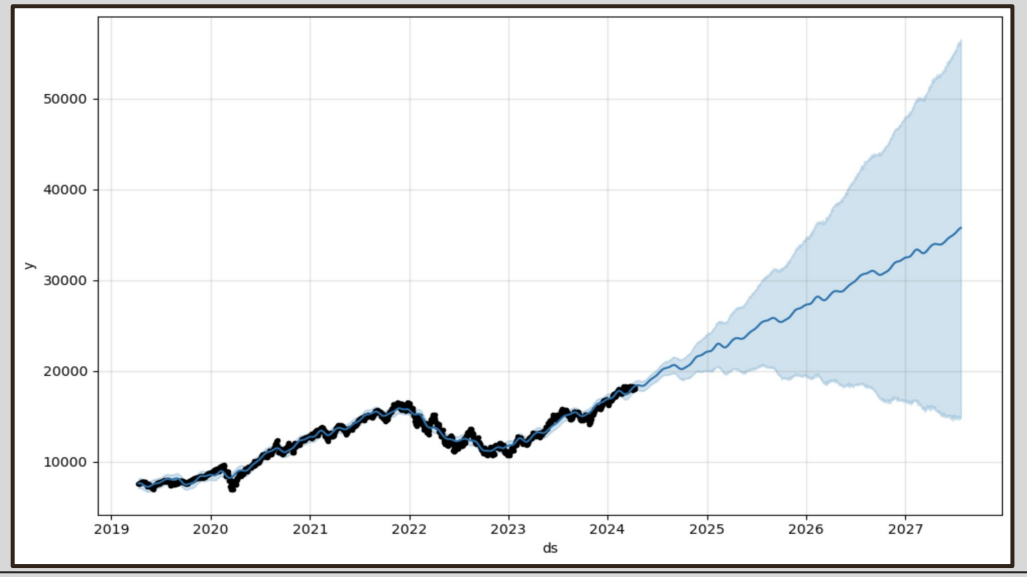
In this Residual Analysis, we can observe the difference in closing price from the mean price over 2019 - 2024. Along with this, we can see that the highest negative price difference from the mean occurs around 04/2020 while the highest positive difference from the mean price occurs around 03/2022.

Plots & Analysis #2





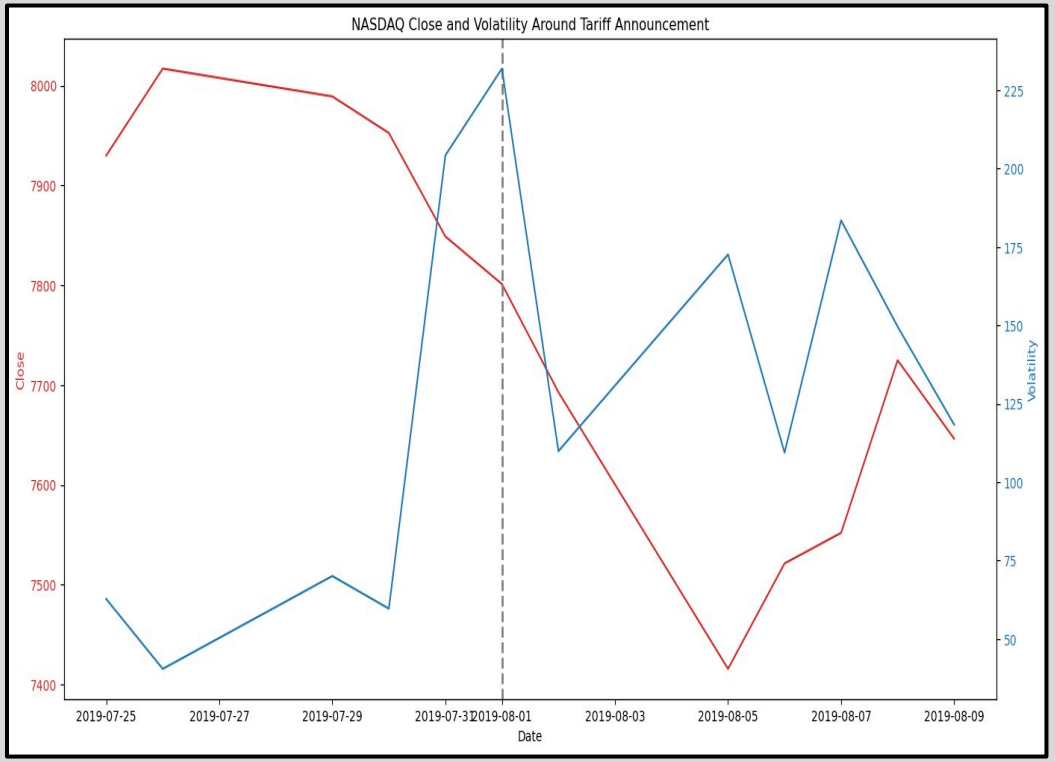




Assessing our forecast, we notice that the 2019 data (top left) has a stable forecasted price, and a very high and low limits. For 2020 - 2024, we can observe the limits becoming smaller with the addition of more data and notice more fluctuation in the average forecast. This remains the same for 2022 - 2024, it is just trending downward. For 2019 - 2024, we have the largest data set, and therefore our forecast accuracy increases due to breadth of data.

Global Markets - Major Geopolitical Events - 2019-2024:

- In August 2019, the U.S. announced new tariffs on Chinese goods, and China responded with tariffs on U.S. goods.



The NASDAQ graph shows a red line for closing prices sharply dropping at the tariff announcement (grey dashed line), then rebounding, while the blue volatility line spikes, indicating immediate market reaction and continued uncertainty.

```
import pandas as pd
import matplotlib.pyplot as plt
```

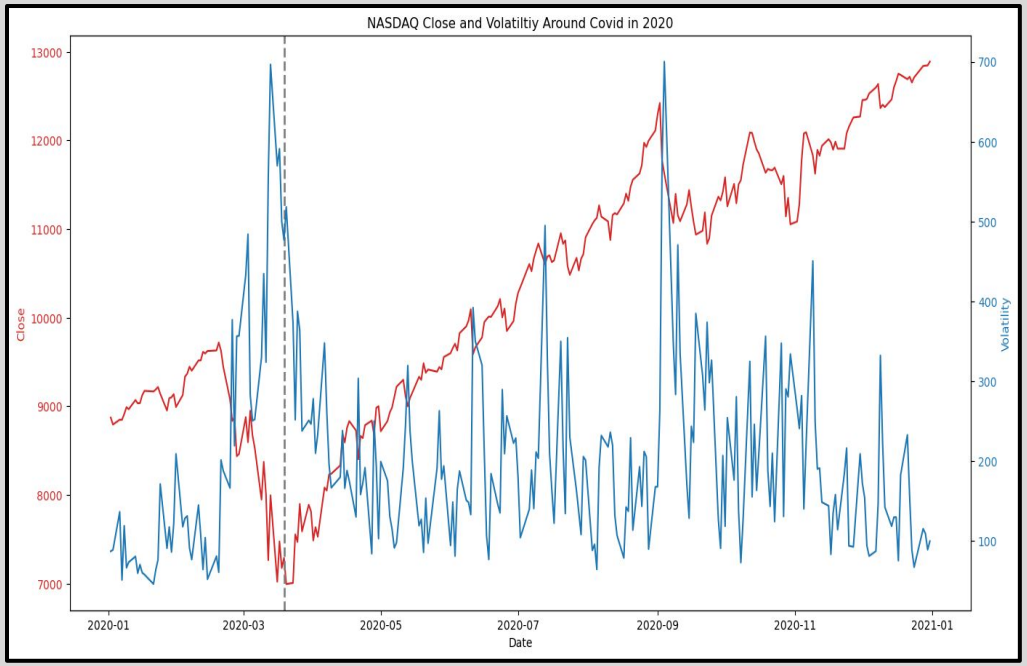
```
df = pd.read_csv('2019.csv')
df['Volatility'] = df['High'] - df['Low']
max_volatility_day = df.loc[df['Volatility'].idxmax()]
print(max_volatility_day)
```

```
Date            8/23/2019
Open          7658.47998
High          7722.430176
Low           7442.930176
Close         7464.990234
Adj Close     7464.990234
Volume         2214370000
Volatility          279.5
Name: 93, dtype: object
```

```
df['Date'] = pd.to_datetime(df['Date'])
start_date = '2019-07-25'
end_date = '2019-08-10'
announcement_date = '2019-08-01'
event_window = df[(df['Date'] >= start_date) & (df['Date'] <= end_date)]
fig, ax1 = plt.subplots(figsize=(14, 7))
color = 'tab:red'
ax1.set_xlabel('Date')
ax1.set_ylabel('Close', color=color)
ax1.plot(event_window['Date'], event_window['Close'], color=color)
ax1.tick_params(axis='y', labelcolor=color)
ax1.axvline(x=pd.to_datetime(announcement_date), color='grey', linestyle='--', lw=2)
ax2 = ax1.twinx()
color = 'tab:blue'
ax2.set_ylabel('Volatility', color=color)  # we already handled the x-label with ax1
ax2.plot(event_window['Date'], event_window['Volatility'], color=color)
ax2.tick_params(axis='y', labelcolor=color)
fig.tight_layout()
plt.title('NASDAQ Close and Volatility Around Tariff Announcement')
plt.show()
```

Global Markets - Major Geopolitical Events - 2019-2024:

- In 2020, a period marked by the global onset and impact of COVID-19.



The NASDAQ's performance in 2020, as depicted in the graph, illustrates the immediate and severe reaction to the onset of COVID-19, followed by a robust recovery and a return to lower levels of volatility. The market's response highlights the impact of the pandemic on investor sentiment and the overall resilience of the tech-heavy index in the face of unprecedented global health and economic crises.

```
# 2020 Covid Impact
import pandas as pd
import matplotlib.pyplot as plt
```

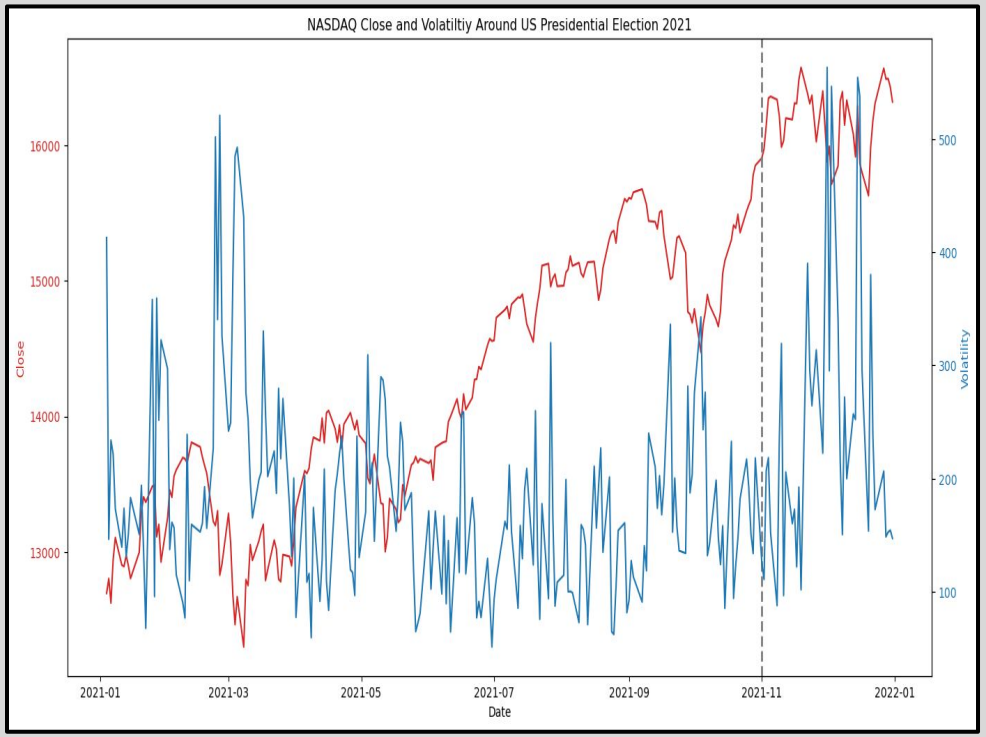
```
df = pd.read_csv('2020.csv')
df['Volatility'] = df['High'] - df['Low']
max_volatility_day = df.loc[df['Volatility'].idxmax()]
print(max_volatility_day)
```

```
Date            9/4/2020
Open        11686.24023
High        11846.17969
Low         11145.99023
Close       11622.12988
Adj Close   11622.12988
Volume       4284440000
Volatility     700.18946
Name: 171, dtype: object
```

```
df['Date'] = pd.to_datetime(df['Date'])
start_date = '2020-01-02'
end_date = '2020-12-31'
announcement_date = '2020-03-19'
event_window = df[(df['Date'] >= start_date) & (df['Date'] <= end_date)]
fig, ax1 = plt.subplots(figsize=(14, 7))
color = 'tab:red'
ax1.set_xlabel('Date')
ax1.set_ylabel('Close', color=color)
ax1.plot(event_window['Date'], event_window['Close'], color=color)
ax1.tick_params(axis='y', labelcolor=color)
ax1.axvline(x=pd.to_datetime(announcement_date), color='grey', linestyle='--', lw=2)
ax2 = ax1.twinx()
color = 'tab:blue'
ax2.set_ylabel('Volatility', color=color)  # we already handled the x-label with ax1
ax2.plot(event_window['Date'], event_window['Volatility'], color=color)
ax2.tick_params(axis='y', labelcolor=color)
fig.tight_layout()
plt.title('NASDAQ Close and Volatiltiy Around Covid in 2020')
plt.show()
```

Global Markets - Major Geopolitical Events - 2019-2024:

- In November 2021, the U.S. Presidential Election took place.



The NASDAQ's performance post-U.S. Presidential Election in 2021 highlights a typical pattern of initial market volatility due to political uncertainty, followed by a trend of adjustment and stabilization as investors gained clarity on the direction of future policies.

```
import pandas as pd
import matplotlib.pyplot as plt
```

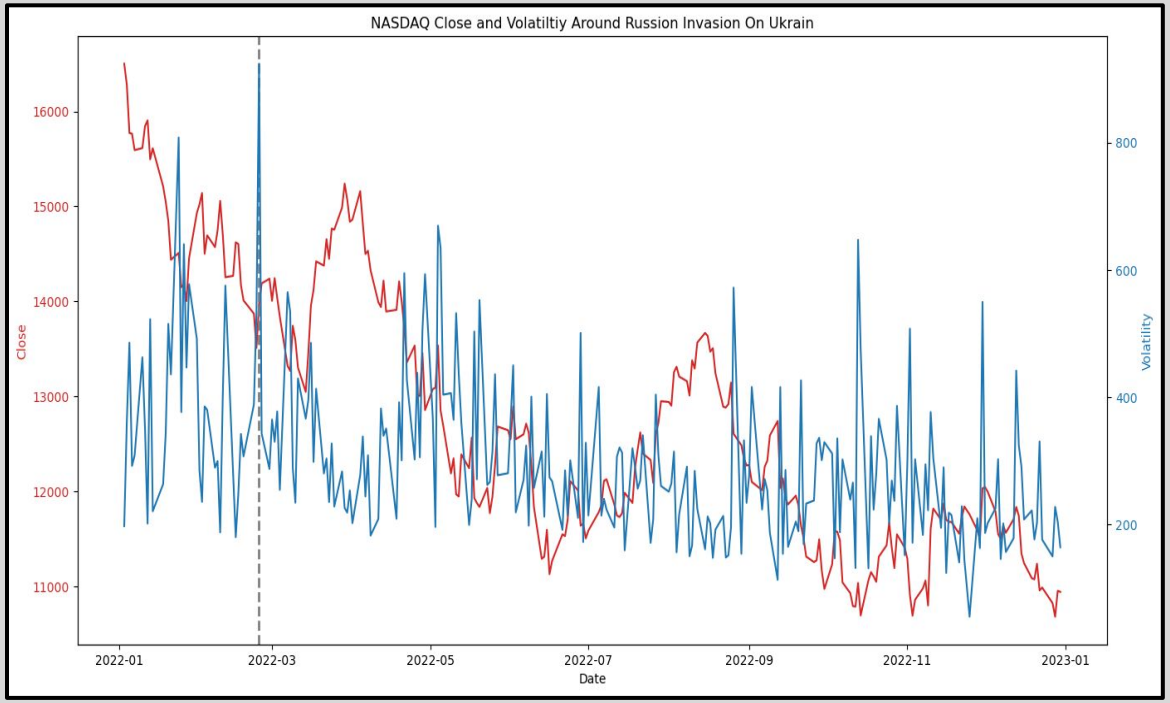
```
df = pd.read_csv('2021.csv')
df['Volatility'] = df['High'] - df['Low']
max_volatility_day = df.loc[df['Volatility'].idxmax()]
print(max_volatility_day)
```

```
Date            12/1/2021
Open          16347.95996
High          16427.18945
Low               15864.0
Close         15877.71973
Adj Close     15877.71973
Volume         6320260000
Volatility      563.18945
Name: 230, dtype: object
```

```
df['Date'] = pd.to_datetime(df['Date'])
start_date = '2021-01-02'
end_date = '2021-12-31'
announcement_date = '2021-11-01'
event_window = df[(df['Date'] >= start_date) & (df['Date'] <= end_date)]
fig, ax1 = plt.subplots(figsize=(14, 7))
color = 'tab:red'
ax1.set_xlabel('Date')
ax1.set_ylabel('Close', color=color)
ax1.plot(event_window['Date'], event_window['Close'], color=color)
ax1.tick_params(axis='y', labelcolor=color)
ax1.axvline(x=pd.to_datetime(announcement_date), color='grey', linestyle='--', lw=2)
ax2 = ax1.twinx()
color = 'tab:blue'
ax2.set_ylabel('Volatility', color=color)  # we already handled the x-label with ax1
ax2.plot(event_window['Date'], event_window['Volatility'], color=color)
ax2.tick_params(axis='y', labelcolor=color)
fig.tight_layout()
plt.title('NASDAQ Close and Volatiltiy Around US Presidential Election 2021')
plt.show()
```

Global Markets - Major Geopolitical Events - 2019-2024:

- In 2022, Russia began its Invasion of Ukraine.



The NASDAQ's performance around the Russia conflict, as shown by the dashed line, reflects a spike in market volatility without a significant corresponding drop in closing prices, suggesting a contained immediate impact on the market. The subsequent upward trend in closing prices alongside the gradual normalization of volatility indicates resilience and a market rebound from short-term geopolitical tensions.

```
import pandas as pd
import matplotlib.pyplot as plt
```

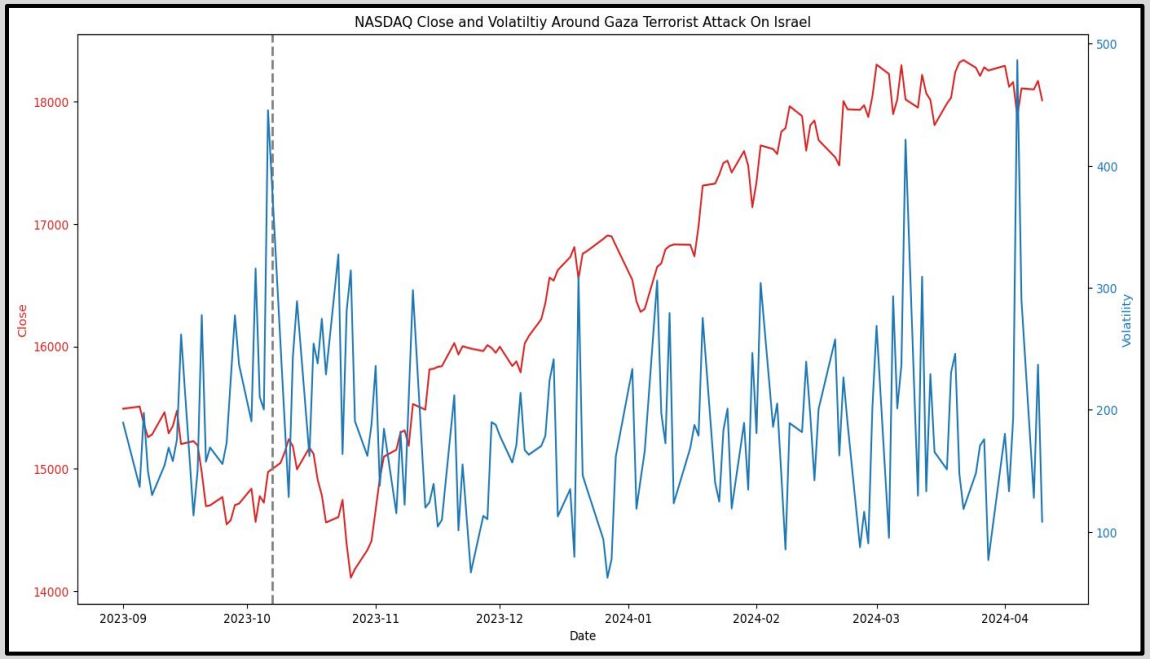
```
df = pd.read_csv('2022.csv')
df['Volatility'] = df['High'] - df['Low']
max_volatility_day = df.loc[df['Volatility'].idxmax()]
print(max_volatility_day)
```

```
Date            2/24/2022
Open          13065.44043
High          13989.09961
Low           13065.44043
Close         13974.66992
Adj Close     13974.66992
Volume         6180200000
Volatility      923.65918
Name: 36, dtype: object
```

```
df['Date'] = pd.to_datetime(df['Date'])
start_date = '2022-01-02'
end_date = '2022-12-31'
announcement_date = '2022-02-24'
event_window = df[(df['Date'] >= start_date) & (df['Date'] <= end_date)]
fig, ax1 = plt.subplots(figsize=(14, 7))
color = 'tab:red'
ax1.set_xlabel('Date')
ax1.set_ylabel('Close', color=color)
ax1.plot(event_window['Date'], event_window['Close'], color=color)
ax1.tick_params(axis='y', labelcolor=color)
ax1.axvline(x=pd.to_datetime(announcement_date), color='grey', linestyle='--', lw=2)
ax2 = ax1.twinx()
color = 'tab:blue'
ax2.set_ylabel('Volatility', color=color)  # we already handled the x-label with ax1
ax2.plot(event_window['Date'], event_window['Volatility'], color=color)
ax2.tick_params(axis='y', labelcolor=color)
fig.tight_layout()
plt.title('NASDAQ Close and Volatiltiy Around Russion Invasion On Ukrain')
plt.show()
```

Global Markets - Major Geopolitical Events - 2019-2024:

- Gaza and Israel conflict begins.



The NASDAQ's performance around the Gaza conflict, as shown by the dashed line, reflects a spike in market volatility without a significant corresponding drop in closing prices, suggesting a contained immediate impact on the market. The subsequent upward trend in closing prices alongside the gradual normalization of volatility indicates resilience and a market rebound from short-term geopolitical tensions.

```
import pandas as pd
import matplotlib.pyplot as plt
```

```
df = pd.read_csv('Sep 2023_April 10 2024.csv')
df['Volatility'] = df['High'] - df['Low']
max_volatility_day = df.loc[df['Volatility'].idxmax()]
print(max_volatility_day)
```

```
Date             4/4/2024
Open          18331.01953
High          18362.28906
Low              17875.75
Close          17878.7793
Adj Close      17878.7793
Volume         5368700000
Volatility      486.53906
Name: 147, dtype: object
```

```
df['Date'] = pd.to_datetime(df['Date'])
start_date = '2023-09-01'
end_date = '2024-04-10'
announcement_date = '2023-10-07'
event_window = df[(df['Date'] >= start_date) & (df['Date'] <= end_date)]
fig, ax1 = plt.subplots(figsize=(14, 7))
color = 'tab:red'
ax1.set_xlabel('Date')
ax1.set_ylabel('Close', color=color)
ax1.plot(event_window['Date'], event_window['Close'], color=color)
ax1.tick_params(axis='y', labelcolor=color)
ax1.axvline(x=pd.to_datetime(announcement_date), color='grey', linestyle='--', lw=2)
ax2 = ax1.twinx()
color = 'tab:blue'
ax2.set_ylabel('Volatility', color=color)  # we already handled the x-label with ax1
ax2.plot(event_window['Date'], event_window['Volatility'], color=color)
ax2.tick_params(axis='y', labelcolor=color)
fig.tight_layout()
plt.title('NASDAQ Close and Volatility Around Gaza Terrorist Attack On Israel')
plt.show()
```

Implementation:

- While you can never "predict" the market, you can analyze it's prior performance to help make better educated decisions
- From our data we can see that outlier events directly correlated with unusual fluctuations in the stock market. Through our 2019 stock market information, we were able to see what a regular market looks like, in comparison to our other datasets from different years.
- For new or seasoned traders, analyzing trends (peaks/high) is vital if your a position trader; or a more hands on (day trader)
- In preparation for scenarios such as outlier events, we believe it would be best to begin by taking precautionary steps in the event something happens and placing "stop losses" on your investments. It is also important that your information gathering is from an authentic source.
- Month by month, week by week, or even day by day, comparing the highs and lows will give any trader the range they can expect and help make better decision in terms where to place their trading

Appendix Information:

In the data, <u>what could be incorrect?</u>

- From our dataset, we believe no information to be incorrect. We do note that of all the available days; Monday – Friday, 9:30AM – 4:00PM the stock market could be open, it is not due to certain holidays. Therefore, any forecasting methods used may show a slightly misconstrued utilization rate.
- We must also ask if our dataset includes data outside of regular trading hours; which it does not.