NAME : ANBUSELVAN MAHALINGAM

SUBJECT : DSC 630 PREDICTIVE ANALYSIS - WEEK 1 ASSIGNMENT

Choosen dataset: stock NVDA dataset.

NVDA stock has been one of hot stock in the market, and it is part of AI boom. Nvidia Corporation is an American multinational corporation and technology company headquartered in Santa Clara, California, and incorporated in Delaware. NVIDIA engineers the most advanced chips, systems, and software for the AI factories of the future. We build new AI services that help companies create their own AI factories. Referece: https://www.nvidia.com/en-us/about-nvidia/#the-iphone-moment-for-ai

Question to visually explore the dataset.

1.How has the stock price of NVIDIA changed over the five-year period?
2.What is the volatility of NVIDIA's stock prices as observed from the high and low prices, and how it relates to trading volume?

In [11]:
```
import pandas as pd

# To load the data from the uploaded csv file
nvda_data = pd.read_csv("NVDA.csv")

# To sisplay the first 5 rows of the dataframe
nvda_data.head(n=5)
```

Out[11]:

	Date	Open	High	Low	Close	Adj Close	Volume
0	2019-03-18	42.770000	43.312500	41.965000	42.237499	41.950645	48699200
1	2019-03-19	43.189999	44.375000	43.020000	43.927502	43.629166	85497200
2	2019-03-20	44.220001	44.757500	43.250000	43.599998	43.303886	71914800
3	2019-03-21	43.832500	46.250000	43.782501	45.985001	45.672695	82431200
4	2019-03-22	45.709999	46.200001	44.235001	44.375000	44.073624	74764400

In [12]:
```
# To display the basic inforamtion about the dataset
nvda_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1259 entries, 0 to 1258
Data columns (total 7 columns):
 #   Column     Non-Null Count  Dtype  
---  ------     --------------  -----  
 0   Date       1259 non-null   object 
 1   Open       1259 non-null   float64
 2   High       1259 non-null   float64
 3   Low        1259 non-null   float64
 4   Close      1259 non-null   float64
 5   Adj Close  1259 non-null   float64
 6   Volume     1259 non-null   int64  
dtypes: float64(5), int64(1), object(1)
memory usage: 69.0+ KB
```

In [13]:
```
# To get comprehensive statistial overview of the dataset
nvda_data.describe()
```

Out[13]:

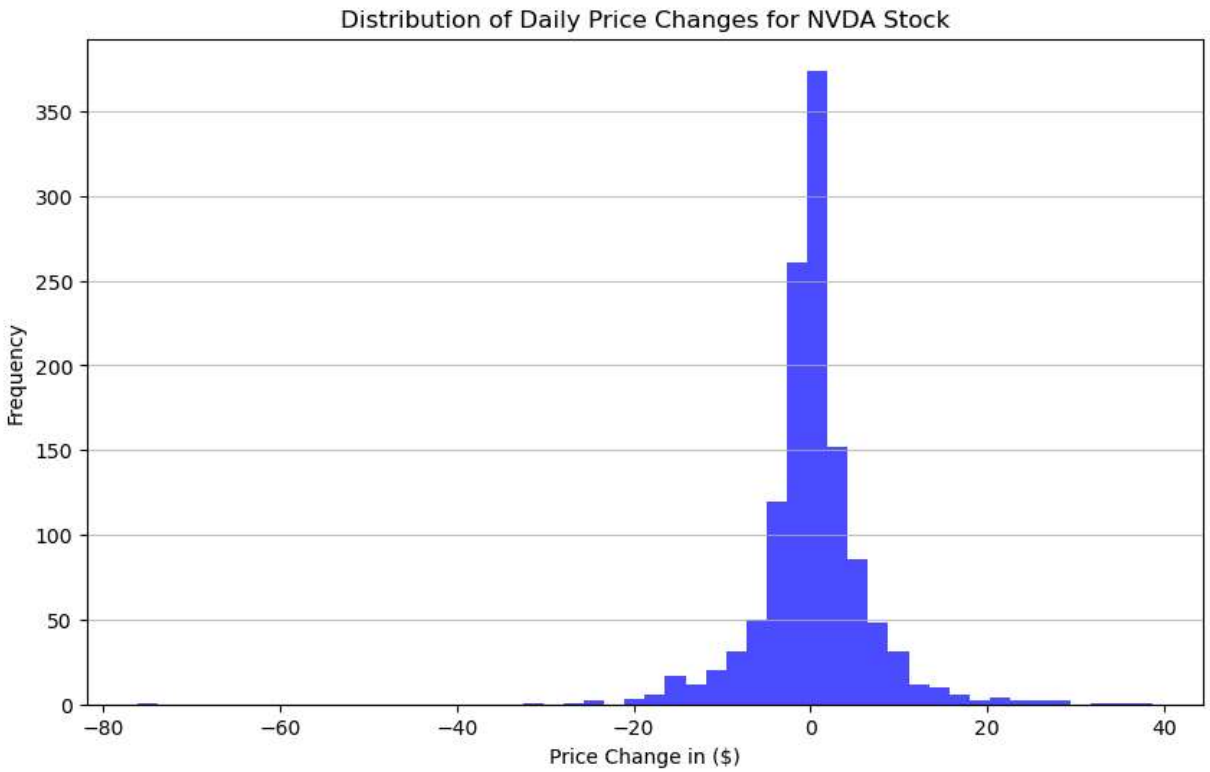
	Open	High	Low	Close	Adj Close	Volume
count	1259.000000	1259.000000	1259.000000	1259.000000	1259.000000	1.259000e+03
mean	204.178250	207.967573	200.315804	204.380348	204.158055	4.573561e+07
std	160.002276	162.798162	157.010112	160.160127	160.216328	1.862956e+07
min	33.977501	34.367500	33.150002	33.445000	33.255768	9.788400e+06
25%	91.651249	92.640003	88.943748	91.643753	91.379029	3.238370e+07
50%	156.997498	160.020004	154.119995	157.000000	156.693680	4.283480e+07
75%	260.370011	265.755004	255.635002	260.649994	260.445389	5.567165e+07
max	951.380005	974.000000	896.020020	926.690002	926.690002	1.543911e+08

In [14]:
```
# Create a histogram or bar graph from your data.
# To plot a histogram of the daily price changes to understand the distribution of

import matplotlib.pyplot as plt

# To extract  new eature from the dataset, calculate daily price changes
nvda_data['Price Change'] = nvda_data['Close'] - nvda_data['Open']

# To plot the histogram of daily price changes
plt.figure(figsize=(10, 6))
plt.hist(nvda_data['Price Change'], bins=50, color='blue', alpha=0.7)
plt.title('Distribution of Daily Price Changes for NVDA Stock')
plt.xlabel('Price Change in ($)')
plt.ylabel('Frequency')
plt.grid(axis='y', alpha=0.75)
plt.show()
```

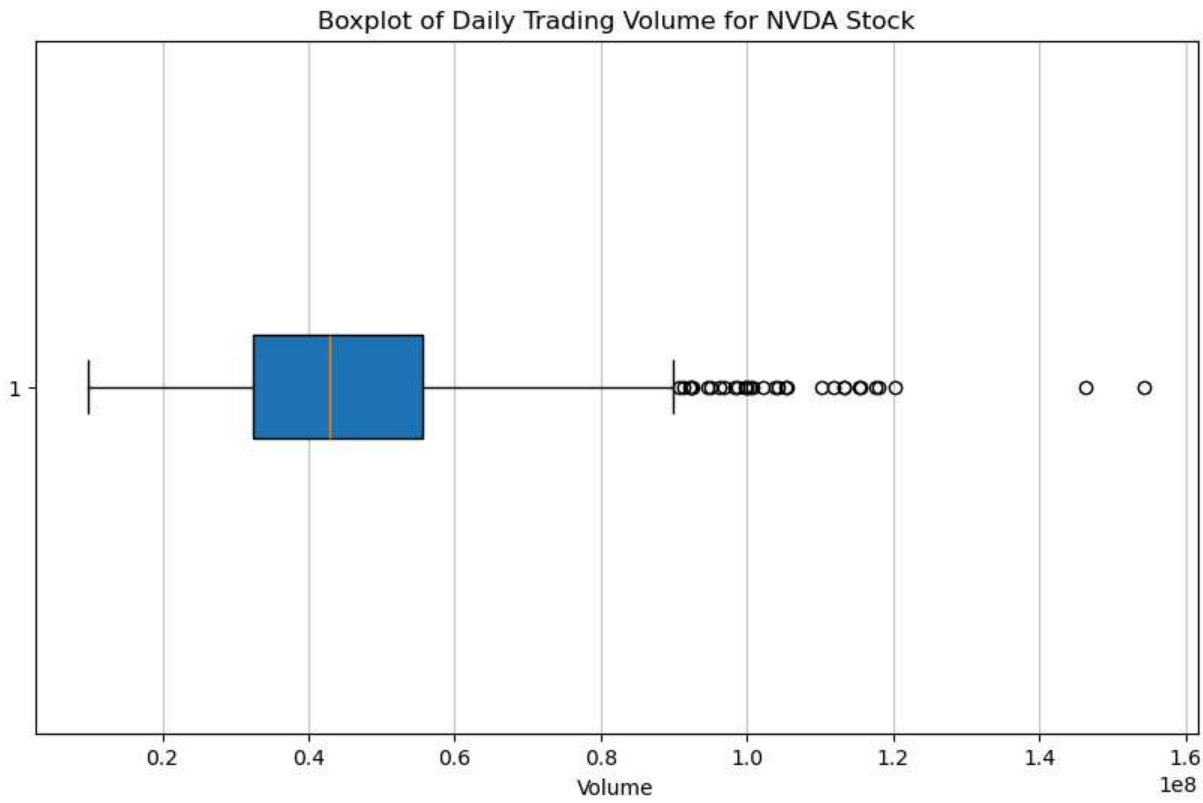


The histogram of daily price changes for NVDA stock shows the distribution indicating days where the opening and closing prices were very close. The distribution appears roughly symmetrical with tails extending to both gains and losses, suggesting days with significant price moves in both directions.

In [15]:
```
# Create a boxplot from your data.

# To plot a boxplot and identify any potential outliers - trading volume
plt.figure(figsize=(10, 6))
plt.boxplot(nvda_data['Volume'], vert=False, patch_artist=True)
plt.title('Boxplot of Daily Trading Volume for NVDA Stock')
plt.xlabel('Volume')
plt.grid(axis='x', alpha=0.75)

plt.show()
```

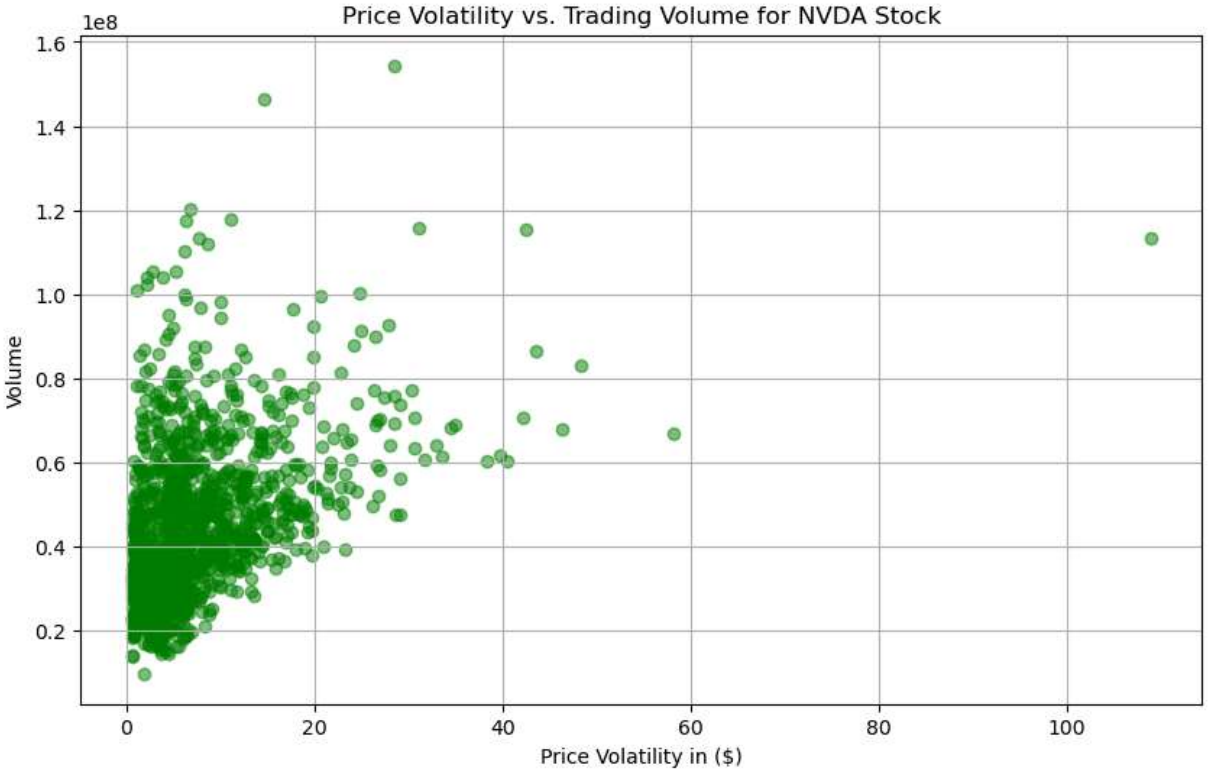


The boxplot of daily trading volume for NVDA stock shows that the bulk of trading volumes are clustered within a certain range, as shown by the box, with the line inside the box representing the median trading volume. The "whiskers" extend to cover the majority of the data points, indicating the typical range of trading volumes. Points outside the whiskers are considered outliers, representing days with unusually high trading volumes.

In [16]:
```
# Create a bivariate plot from your data.
# The relationship between price volatility and trading volume.

# To extract new feature, to calculate daily price volatility as the difference bet
nvda_data['Price Volatility'] = nvda_data['High'] - nvda_data['Low']

# To Plot the scatter plot for price volatility vs trading volume
plt.figure(figsize=(10, 6))
plt.scatter(nvda_data['Price Volatility'], nvda_data['Volume'], alpha=0.5, color='g
plt.title('Price Volatility vs. Trading Volume for NVDA Stock')
plt.xlabel('Price Volatility in ($)')
plt.ylabel('Volume')
plt.grid(True)
plt.show()
```



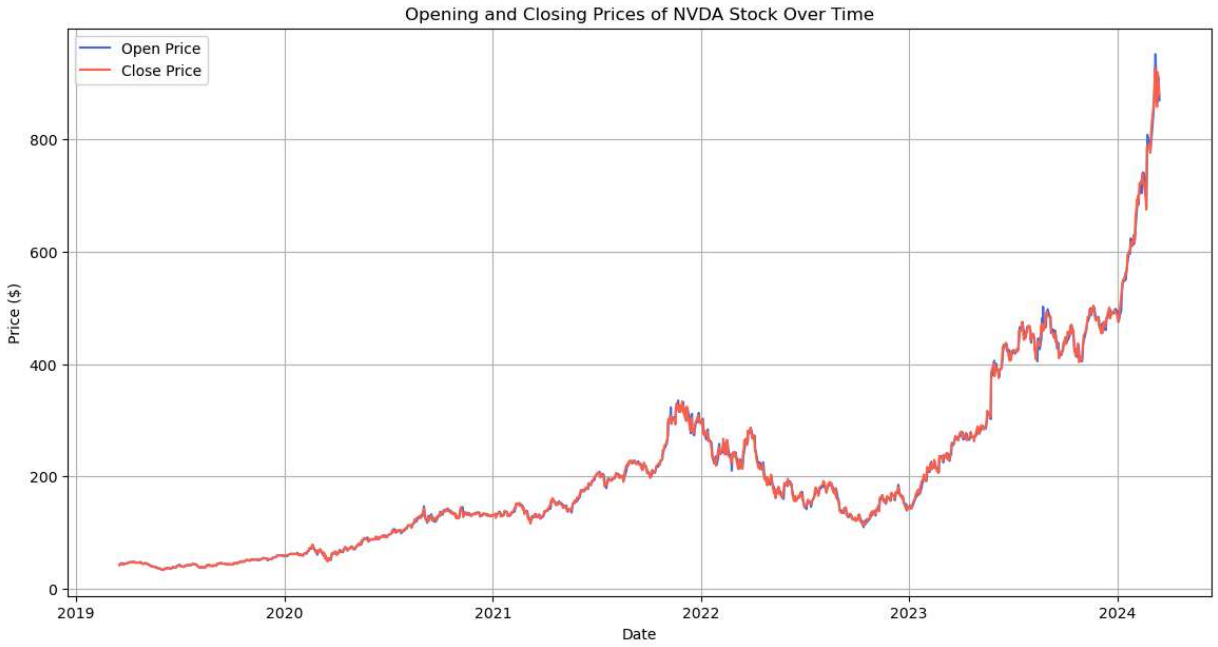
The scatter plot illustrating the relationship between price volatility and trading volume NVDA stock, it shows how these two factors interact. Each point on the plot represents a trading day, with the position determined by the day's price volatility and trading volume. the plot indicates that there is no clear linear relationship, the plot suggests that days with higher price volatility tend to have higher trading volumes.

In [17]:
```
# Create any additional visualizations that will help to answer the question(s) you

# To observe the overall trend and performance of NVIDIA's stock over time - line g

# to convert the date column to datetime format
nvda_data['Date'] = pd.to_datetime(nvda_data['Date'])

# to plot the line graph for opening and closing prices over time
plt.figure(figsize=(14, 7))
plt.plot(nvda_data['Date'], nvda_data['Open'], label='Open Price', color='royalblue
plt.plot(nvda_data['Date'], nvda_data['Close'], label='Close Price', color='tomato'
plt.title('Opening and Closing Prices of NVDA Stock Over Time')
plt.xlabel('Date')
plt.ylabel('Price ($)')
plt.legend()
plt.grid(True)
plt.show()
```



The line graph of opening and closing prices over time shows that NVDA stock has experienced periods of growth,volatility and some stability.In a 5 year period we can conlude that the stock has uptrend and good investor return. During the the period of 2022 to 2023 there was a slight down trend and then started moving up.

In [ ]:

Summarize the results and make a conclusion

Statistical Information - The average (mean) values for Open, High, Low, Close, and Adjusted Close prices suggest the stock traded around the 200 dollars mark on average, with typical daily high-low fluctuations of around 7 to 8 dollars.The minimum and maximum values highlight the stock's price range from as low as approximately 34 to as high as around 974, showing a wide variation in stock price and its massive growth over the period.The trading volume varied between roughly 9.79 million to 154.39 million shares per day, indicating days of both low and extremely high trading activity, this is the indication of high number of invstor's interest for the stock.

The histogram of daily price changes for NVDA stock shows the distribution indicating days where the opening and closing prices were very close. The distribution appears roughly symmetrical with tails extending to both gains and losses, suggesting days with significant price moves in both directions.

The boxplot of daily trading volume for NVDA stock shows that the bulk of trading volumes are clustered within a certain range, as shown by the box, with the line inside the box representing the median trading volume. The "whiskers" extend to cover the majority of the data points, indicating the typical range of trading volumes. Points outside the whiskers are considered outliers, representing days with unusually high trading volumes.

The scatter plot illustrating the relationship between price volatility and trading volume NVDA stock, it shows how these two factors interact. Each point on the plot represents a trading day, with the position determined by the day's price volatility and trading volume. the plot indicates that there is no clear linear relationship, the plot suggests that days with higher price volatility tend to have higher trading volumes.

The line graph of opening and closing prices over time shows that NVDA stock has experienced periods of growth,volatility and some stability.In a 5 year period we can conlude that the stock has uptrend and good investor return. During the the period of 2022 to 2023 there was a slight down trend and then started moving up.

In [ ]: