

**19ADC23 - BUSINESS DATA
ANALYTICS**

Assignment

on

**Stock Price Prediction and
Market Analysis**

Submitted By:

BHUVANA S

20BAD004

TABLE OF CONTENTS

S.No	Contents	Page No.
1	Introduction	1
2	History of the Companies	2
3	Data Extraction	7
4	Descriptive Analytics	14
5	Predictive Analytics	18
6	Prescriptive Analytics	26
7	Conclusion	28

1. Introduction:

Stock price prediction and market analysis are crucial aspects of the financial industry, offering investors valuable insights and strategies for informed decision-making. In this study, we focus on three prominent companies: Google, Tesla, and Twitter. These companies represent different sectors within the stock market, making it a diverse and interesting analysis.

The Companies I have chosen for Stock market analysis are: **Google, Tesla and Twitter.**

To perform our market analysis, we collected data from various sources, including financial news, stock exchange websites, and company reports. Web scraping techniques were employed to gather real-time and historical data related to these three companies, including stock prices, financial statements, and news sentiment.

Descriptive analytics involves uncovering insights from historical data. In our study, we utilize association rule mining to identify patterns and relationships among different stock market factors. This method helps us understand how various market conditions and events may influence the stock prices of Google, Tesla, and Twitter.

Predictive analysis is a crucial step in stock market analysis, as it involves using historical data to predict future stock price movements. Classification models are employed to make informed predictions about whether a stock's price will go up or down. These models are based on historical data, technical indicators, and other relevant factors.

Prescriptive analysis is the final stage of our study, where we aim to provide actionable recommendations for investors. We use predictive models to suggest which features of stocks (such as Google, Tesla, or Twitter) are likely to perform well in the near future, helping investors make informed decisions on buying, selling, or holding their stock positions.

This stock market analysis covers the history and data extraction of three prominent companies, employs descriptive analytics to identify patterns, utilizes predictive analysis to forecast stock price movements, and provides prescriptive analysis to guide investment decisions. By examining Google, Tesla, and Twitter, we aim to offer valuable insights for investors in today's dynamic financial market.

2. History of the Companies:

1) GOOGLE

Google, one of the world's leading technology companies, was founded by Larry Page and Sergey Brin, two Ph.D. students at Stanford University. The company's history is marked by its meteoric rise and its transformation from a simple search engine to a global conglomerate known as Alphabet Inc.

Under the name Google Inc., Alphabet was initially established in 1998 as a search engine company. Google entered the stock market on **August 19, 2004**. The company used an initial public offering (IPO) to enter the market. The stock was listed on the **NASDAQ** under the ticker symbol "**GOOG**." The initial stock price for

Google's IPO was set at **\$85 per share**. Now, there are 3 types of Shares offered by them. They are:

- Class A: Held by a regular investor with regular voting rights (GOOGL)
- Class B: Held by the founders, with 10 times the voting power of Class A shares.
- Class C: No voting rights, typically held by employees and some Class A stockholders (GOOG)

Today's Market price of Google is **129.56 USD**. (₹10,763.84). On **November 18, 2021**, Alphabet shares closed at an **all-time high price of 149.84**. The company's 52-week highest share price is 133.74 and its low stock price is 83.34 for the 52 weeks. It has an average price of 105.40 for an year. The current average share price is 109.2412. The year began at 89.1200 and ended at 132.7200. The yearly change is approximately 46.84% higher.

- **1 Year Ago (2022)**, The Average Stock price was **114.7604**. The Year Open and Year Close are 144.9915 and 88.2300 respectively. The Year High and Low was 148.000 and 83.4300 respectively. The annual Change was around **-39.09%**.
- **2 Years Ago (2021)**, The Average Stock price was **124.2249**. The Year Open and Year Close are 86.3065 and 144.8520 respectively. The Year High and Low was 149.8385 and 86.1440 respectively. The annual Change was around **65.30%**.
- **5 Years Ago (2018)**, The Average Stock price was **56.1022**. The Year Open and Year Close are 53.6605 and 52.2480 respectively. The Year High and Low was 64.2750 and 49.2335 respectively. The annual Change was around **-0.80%**.

Allied Industries:

Google Inc. is a holding corporation with subsidiary firms that operate in a variety of connected sectors. These products target a variety of use cases, including Gmail, Waze & Maps, Cloud, Chrome, YouTube, Drive, Translate, Photos, Meet, Nest, Pixel, wearables, and so on.

The **Major Competitors** of Google in Stock Market are Microsoft Corporation (MSFT), Apple Inc. (AAPL), Amazon.com Inc. (AMZN), Facebook, Inc. (Meta Platforms, Inc.) (FB), Netflix Inc. (NFLX), Tencent Holdings Limited (TCEHY) and Alibaba Group Holding Limited (BABA).

Google's stock has played a pivotal role in the company's remarkable history and financial success. Following its initial public offering (IPO) in 2004, Google's stock has consistently demonstrated exceptional performance, establishing itself as one of the most valuable and heavily traded stocks on the market. The company is traded on the NASDAQ stock exchange under two ticker symbols: "GOOGL" for Class A shares, which come with voting rights, and "GOOG" for Class C shares, which do not have voting rights.

The IPO, which took place on August 19, 2004, marked a significant milestone in Google's journey. The initial offering price for Google's shares was set at \$85 per share, resulting in the raising of over \$1.6 billion in capital. Subsequently, the company's market capitalization experienced rapid growth, firmly establishing Google (now a subsidiary of Alphabet Inc.) as one of the world's largest companies by market value.

Despite periodic fluctuations, Google's stock performance has been highly impressive since its IPO, offering substantial returns to investors. As of my last knowledge update in September 2021, Alphabet Inc.'s market capitalization had surpassed \$1.5 trillion, further solidifying its position as a heavyweight in the global market. Please note that stock market data is subject to change, and up-to-date financial information should be consulted if you are considering investing in Alphabet Inc. or monitoring its stock performance.

2) TESLA

Tesla, Inc., a pioneering American electric vehicle and clean energy company, has made a profound impact on the automotive and energy sectors since its founding. The company was established in 2003 by a group of engineers, including Martin Eberhard and Marc Tarpenning, but it was the entry of entrepreneur Elon Musk in the early stages that transformed Tesla into a global phenomenon. Tesla's journey began with the introduction of the groundbreaking Tesla Roadster in 2008, an all-electric sports car that showcased impressive range and performance. Subsequent milestones included the launch of the Model S in 2012, a luxury electric sedan that redefined the EV industry, and the Model 3 in 2017, which made electric vehicles more accessible to a wider audience. Beyond vehicles, Tesla expanded into the production of clean energy solutions and battery technology, solidifying its role in the renewable energy sector. Tesla's relentless innovation in autonomous driving technology and its stock market performance have made it a symbol of modern electric mobility.

Tesla's stock has captured the attention of investors and financial markets worldwide due to its extraordinary performance and disruptive influence. Traded under the ticker symbol "TSLA" on the NASDAQ stock exchange, Tesla went public in 2010,

initially raising approximately \$226 million with an IPO price of \$17 per share. Since then, Tesla's market capitalization has experienced remarkable growth, propelling it to become one of the world's most valuable car manufacturers. As of my last knowledge update in September 2021, Tesla's market capitalization exceeded \$700 billion. Tesla's stock performance has been marked by significant volatility, with periods of rapid growth and occasional corrections. It has garnered a dedicated and passionate investor base, making it a focal point of discussions and analysis within the financial markets. It's important to note that stock market data is subject to change, and investors should refer to the most current financial information when considering investments in Tesla, Inc.

Tesla, Inc. debuted in the stock market on June 29, **2010**, with its first public offering (IPO). Tesla was recognised for its **electric vehicles (EVs)** and pioneering work in the creation of electric automobiles at the time of the IPO, notably the Tesla Roadster, which was the company's flagship product in the early days.

The initial stock price for Tesla's IPO was **\$17.00** per share. The **Tesla Roadster** was Tesla's first production electric vehicle (EV) and marked the company's entry into the automotive industry and also into the Stock Market.

The current share price of Tesla is **\$242.65**. The **highest stock closing price** ever recorded (all-time highest) was **409.97** on November 4, **2021**. The highest stock price for 52 weeks is 314.67. The lowest stock price for Tesla stock is 101.81 for 52 weeks. The average Tesla stock price during the preceding 52 weeks was 210.65.

- **1 Year Ago (2022)**, The Average Stock price was **263.0931**. The Year Open and Year Close are 399.9267 and 123.1800 respectively. The Year High and Low was 399.9267 and 109.1000 respectively. The annual Change was around **-65.03%**.
- **2 Years Ago (2021)**, The Average Stock price was **259.9982**. The Year Open and Year Close are 243.2567 and 352.2600 respectively. The Year High and Low was 409.9700 and 187.6667 respectively. The annual Change was around **49.76%**.
- **5 Years Ago (2018)**, The Average Stock price was **21.1540**. The Year Open and Year Close are 21.3687 and 22.1867 respectively. The Year High and Low was 25.3047 and 16.7040 respectively. The annual Change was around **6.89%**.

Allied Industries:

Tesla, Inc. is a multi-industry company best recognised for its electric cars (EVs) and innovative energy solutions. Energy Storage, Autonomous Driving, Energy Generation and Distribution, Electric Vehicle Components, and Automotive Manufacturing and Innovation are all part of the programme.

The **Potential Competitors** of Tesla in Stock Market includes, Ford (NYSE: F), General Motors (NYSE: GM), Rivian (NASDAQ: RIVN), NIO Inc. (NYSE: NIO),

Lucid Group (NASDAQ: LCID) , Xpeng Motors (NYSE: XPEV), Li Auto Inc (NASDAQ: LI), Porsche (VWAGY), Nissan (NSANY), Hyundai (HYMTF) and Kia (KIMTF).

3) TWITTER

Twitter, a globally recognized social media platform, has played a pivotal role in shaping online communication and sharing since its inception. The company was founded in 2006 by Jack Dorsey, Noah Glass, Biz Stone, and Evan Williams, who sought to create a platform for real-time, short-form messaging. In March 2006, the very first tweet was sent by Jack Dorsey, marking the birth of this microblogging service. Over the years, Twitter grew in popularity, serving as a platform for individuals, organizations, and even governments to share news, updates, and opinions. The platform introduced the now-iconic 280-character tweet limit, hashtags, and features like retweets and likes. Twitter has become a vital tool for information dissemination, public discourse, and real-time updates, making it an integral part of the digital landscape.

Twitter's journey on the stock market has been closely watched, reflecting the company's role as a prominent tech and social media player. The company is traded on the New York Stock Exchange (NYSE) under the ticker symbol "TWTR." Twitter went public with its initial public offering (IPO) on November 7, 2013. The IPO was priced at \$26 per share, raising over \$1.8 billion and giving the company a market capitalization of around \$18 billion. Since then, Twitter's stock has seen fluctuations and challenges, influenced by user growth, engagement metrics, and competition in the social media space. Its market capitalization has fluctuated as well, reflecting the evolving dynamics of the platform and its user base. Twitter remains a significant player in the tech industry and social media landscape, impacting global conversations and trends. Please note that stock market data is subject to change, and for the most up-to-date financial information, investors should consult current sources when considering investments in Twitter, Inc.

Twitter, Inc. went public on November 7, **2013**, with an initial public offering (IPO). The company's main product is Twitter, a social media platform that allows users to post and engage with brief messages known as "tweets." Twitter has grown into an important medium for real-time news, information exchange, and social networking.

The initial stock price for Twitter's IPO was **\$26.00** per share.

The **primary product** that it used to attract investors was its social media platform was "**Twitter.**" It was a prominent venue for breaking news, conversations, and interactions in a variety of sectors when it went public. It was well recognised for its quick distribution of news, trends, and user-generated material at the time.

In October 2022, Twitter became a private company by which all its shares were removed from the Stock Market.

Because TWTR was delisted after October 27, 2022, the most recent closing stock price for Twitter on that day is **53.70**. On March 1, 2021, the closing stock price of Twitter was **77.63**, which was its **all-time highest**.

- **1 Year Ago (2022)**, The Average Stock price was **41.0348**. The Year Open and Year Close are 42.6600 and 53.7000 respectively. The Year High and Low was 53.7000 and 32.4200 respectively. The annual Change was around **24.25%**.
- **2 Years Ago (2021)**, The Average Stock price was **59.9917**. The Year Open and Year Close are 54.5300 and 43.2200 respectively. The Year High and Low was 77.6300 and 42.0700 respectively. The annual Change was around **-20.18%**.
- **5 Years Ago (2018)**, The Average Stock price was **32.8285**. The Year Open and Year Close are 24.5100 and 28.7400 respectively. The Year High and Low was 28.7400 and 22.1600 respectively. The annual Change was around **19.70%**.

Allied Industries:

Twitter, Inc. is largely involved in the social media and digital advertising industries. While Twitter's primary product is its social media platform, the company's activities and stock market impact extend to several allied industries and products, including Digital Advertising, Technology and Internet Services, Media and News, Public Relations and Brand Management, E-commerce and Influencer Marketing, and Entertainment.

The **Potential Competitors** of Twitter in Stock Market includes, Trade Desk (TTD), Baidu (BIDU), Zoom Video Communications (ZM), Pinterest (PINS), Snap (SNAP), Dell Technologies (DELL), TE Connectivity (TEL), MSCI (MSCI), Autodesk (ADSK), and ON Semiconductor (ON).

3. Data Extraction:

Web scraping is a powerful technique for extracting valuable stock market data from various online sources. Stock market data is abundantly available on financial news websites, stock exchanges, and financial data providers, offering a treasure trove of information, including stock prices, historical data, company financials, news, and more. To perform web scraping for stock market data, programming languages like Python, coupled with libraries such as BeautifulSoup and Scrapy, are commonly used. These tools enable you to specify the websites and specific elements (e.g., tables, charts) from which you want to extract data.

This process allows you to collect a wide array of stock market information, such as real-time stock prices, historical price charts, company financial statements, earnings reports, dividend histories, and news articles relevant to specific stocks or market trends. While web scraping provides valuable data, it comes with its share of challenges, including potential countermeasures implemented by websites to prevent or limit scraping, ethical considerations, and legal compliance. Once the data is scraped, it typically requires processing and cleaning to convert unstructured information into structured formats and ensure data accuracy.

Web scraped stock market data serves a range of purposes, from building financial models and market analysis to backtesting trading strategies and tracking stock or portfolio performance. However, for real-time data, accessing application programming interfaces (APIs) from stock exchanges or data providers is often necessary, as web scraping might not provide up-to-the-minute information. Nevertheless, web scraping remains a versatile tool for obtaining stock market data and conducting detailed financial analysis.

The scraped Data from **Google Stock Market** is saved as csv file. The dataset looks like

	Date	Open	High	Low	Close	Adj Close	Volume
0	05-10-2022	99.830002	101.930000	98.800003	101.430000	101.430000	22176900
1	06-10-2022	100.680000	102.930000	100.589996	101.419998	101.419998	22324000
2	07-10-2022	99.849998	100.529999	98.300003	98.680000	98.680000	27502800
3	10-10-2022	99.000000	99.099998	97.010002	97.860001	97.860001	18225100
4	11-10-2022	97.430000	99.250000	96.309998	97.180000	97.180000	26507100
5	12-10-2022	97.300003	98.750000	96.739998	97.559998	97.559998	21876400
6	13-10-2022	95.150002	99.779999	94.379997	99.059998	99.059998	34574500
7	14-10-2022	99.989998	100.690002	96.370003	96.559998	96.559998	25410700
8	17-10-2022	98.860001	100.930000	98.830002	99.970001	99.970001	28419200
9	18-10-2022	103.129997	103.470001	99.970001	100.769997	100.769997	24060700
10	19-10-2022	100.010002	101.019997	98.919998	99.629997	99.629997	23811100
11	20-10-2022	100.000000	102.320000	99.389999	99.970001	99.970001	25323200
12	21-10-2022	97.849998	101.309998	97.720001	101.129997	101.129997	32458700
13	24-10-2022	101.800003	102.750000	99.980003	102.519997	102.519997	27176400
14	25-10-2022	102.900002	104.820000	102.720001	104.480003	104.480003	40611400

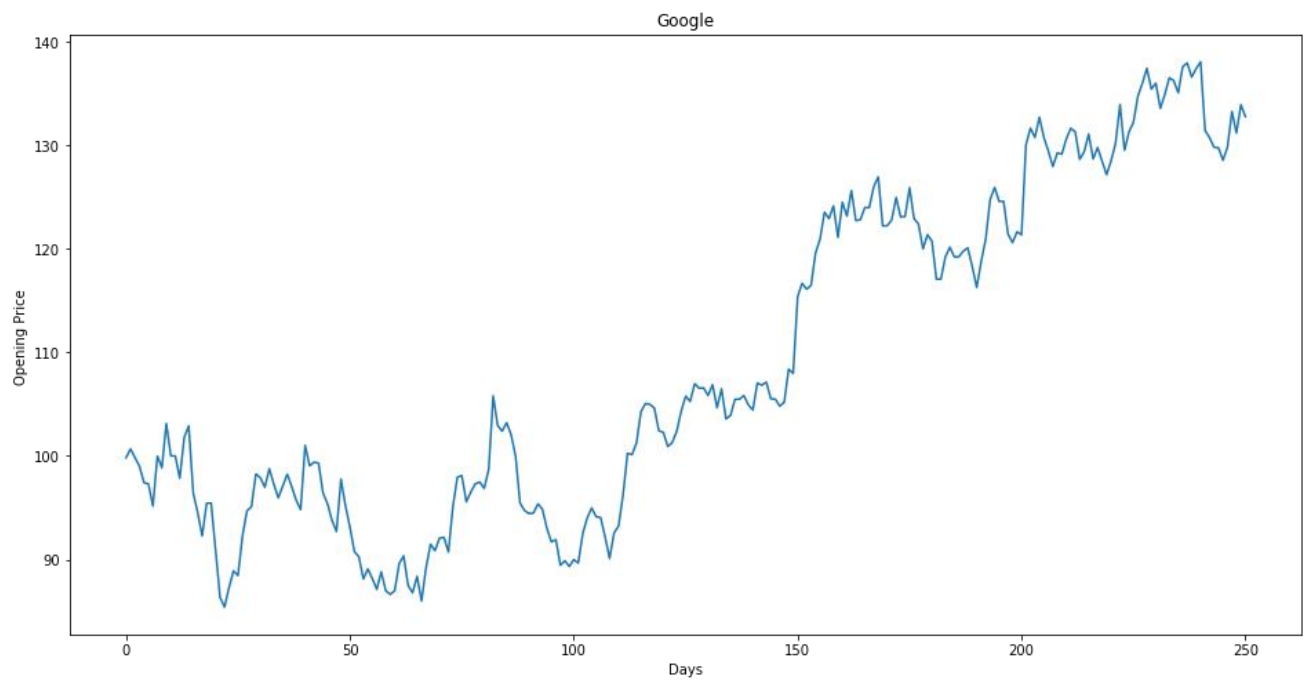
The Columns in the dataset and their data types are

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 251 entries, 0 to 250
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Date        251 non-null    object
1   Open        251 non-null    float64
2   High        251 non-null    float64
3   Low         251 non-null    float64
4   Close       251 non-null    float64
5   Adj Close   251 non-null    float64
6   Volume      251 non-null    int64
dtypes: float64(5), int64(1), object(1)
memory usage: 15.7+ KB
```

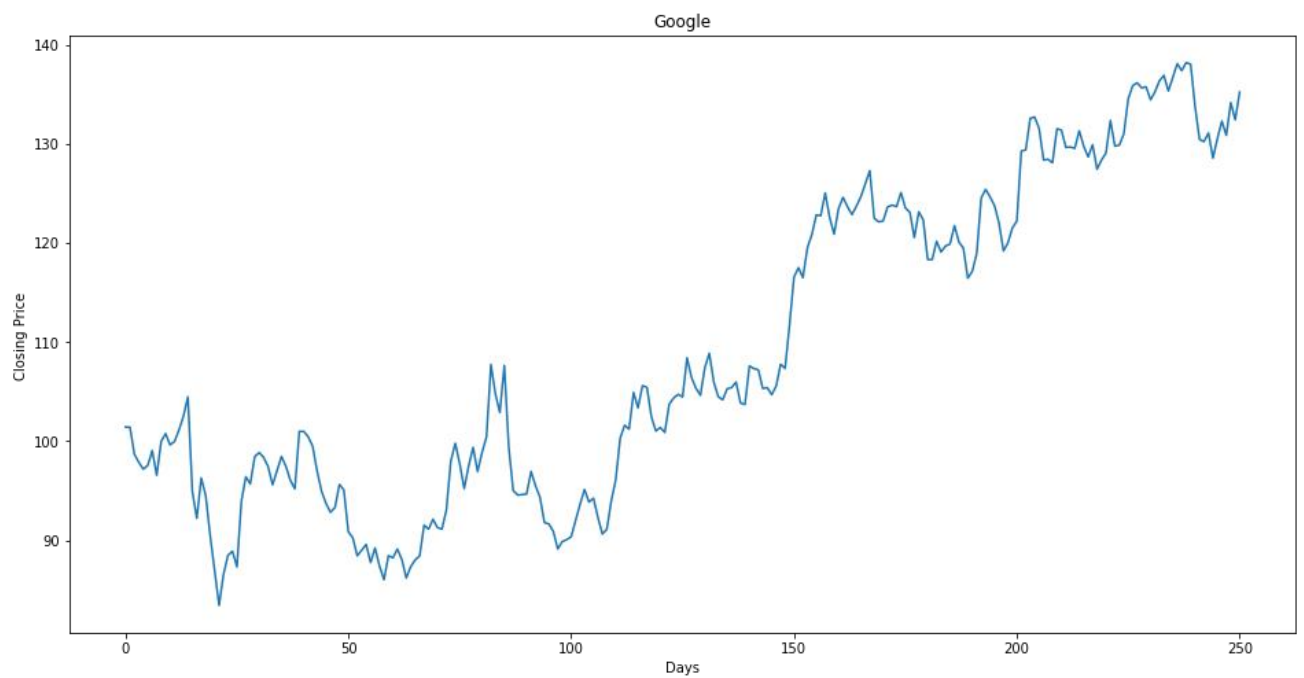
The detailed description of the dataset is

	Open	High	Low	Close	Adj Close	Volume
count	251.000000	251.000000	251.000000	251.000000	251.000000	2.510000e+02
mean	109.065578	110.571673	107.884900	109.254940	109.254940	3.331596e+07
std	15.827003	15.877364	15.820469	15.842716	15.842716	1.345568e+07
min	85.400002	86.519997	83.339996	83.430000	83.430000	9.701400e+06
25%	95.399998	96.930000	94.130001	95.555000	95.555000	2.512575e+07
50%	104.660004	106.160004	103.709999	104.720001	104.720001	2.992750e+07
75%	123.770001	125.220001	122.305001	123.669998	123.669998	3.658815e+07
max	138.080002	139.160004	136.619995	138.210007	138.210007	1.194550e+08

The Opening Stock Price of Google trends as:



The Closing Stock Price of Google trends as:



The scraped Data from **Tesla Stock Market** is saved as csv file. The dataset looks like

	Date	Open	High	Low	Close	Adj Close	Volume
0	05-10-2022	245.009995	246.669998	233.270004	240.809998	240.809998	86982700
1	06-10-2022	239.440002	244.580002	235.350006	238.130005	238.130005	69298400
2	07-10-2022	233.940002	234.570007	222.020004	223.070007	223.070007	83916800
3	10-10-2022	223.929993	226.990005	218.360001	222.960007	222.960007	67925000
4	11-10-2022	220.949997	225.750000	215.000000	216.500000	216.500000	77013200
5	12-10-2022	215.330002	219.300003	211.509995	217.240005	217.240005	66860700
6	13-10-2022	208.300003	222.990005	206.220001	221.720001	221.720001	91483000
7	14-10-2022	224.009995	226.259995	204.160004	204.990005	204.990005	94124500
8	17-10-2022	210.039993	221.860001	209.449997	219.350006	219.350006	79428800
9	18-10-2022	229.500000	229.820007	217.250000	220.190002	220.190002	75891900
10	19-10-2022	219.800003	222.929993	217.779999	222.039993	222.039993	66571500
11	20-10-2022	208.279999	215.550003	202.000000	207.279999	207.279999	117798100
12	21-10-2022	206.419998	214.660004	203.800003	214.440002	214.440002	75713800
13	24-10-2022	205.820007	213.500000	198.589996	211.250000	211.250000	100446800
14	25-10-2022	210.100006	224.350006	210.000000	222.419998	222.419998	96507900

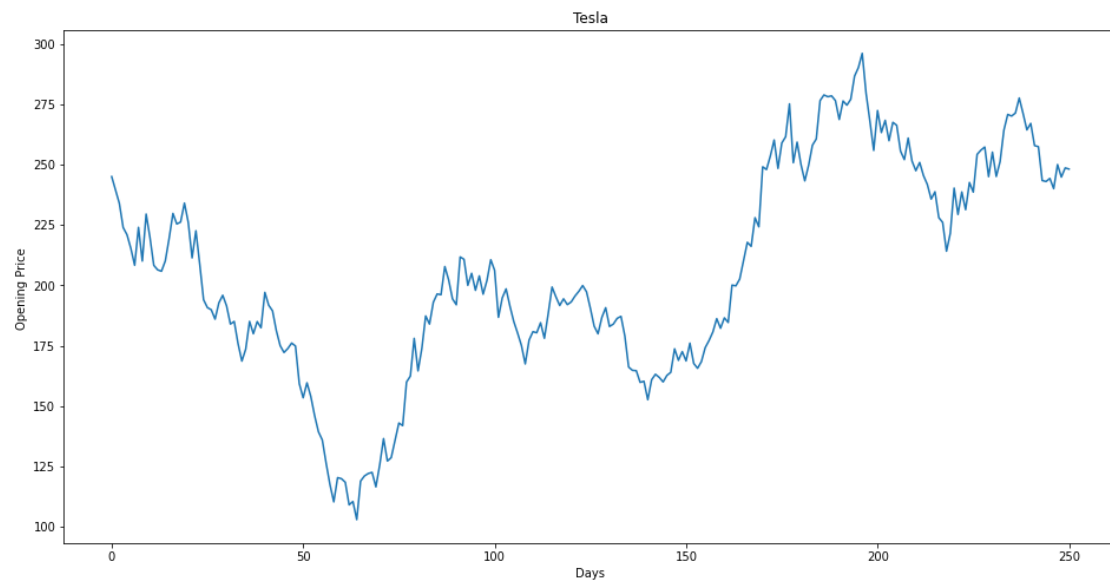
The Columns in the dataset and their data types are

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 251 entries, 0 to 250
Data columns (total 7 columns):
 #   Column      Non-Null Count  Dtype
---  -
 0   Date        251 non-null    object
 1   Open        251 non-null    float64
 2   High        251 non-null    float64
 3   Low         251 non-null    float64
 4   Close       251 non-null    float64
 5   Adj Close   251 non-null    float64
 6   Volume      251 non-null    int64
dtypes: float64(5), int64(1), object(1)
memory usage: 15.7+ KB
```

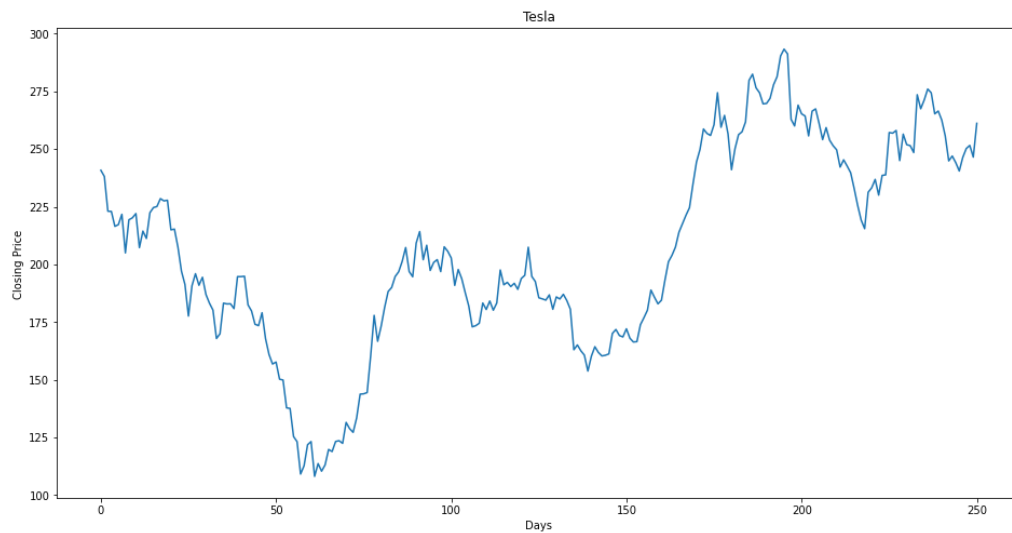
The detailed description of the dataset is

	Open	High	Low	Close	Adj Close	Volume
count	251.000000	251.000000	251.000000	251.000000	251.000000	2.510000e+02
mean	205.490558	210.022470	200.759084	205.547649	205.547649	1.334926e+08
std	43.930415	44.290499	43.563631	43.951559	43.951559	4.212314e+07
min	103.000000	111.750000	101.809998	108.099998	108.099998	5.067270e+07
25%	176.635002	182.419998	172.504998	178.474998	178.474998	1.019894e+08
50%	198.539993	203.000000	192.889999	197.580002	197.580002	1.270627e+08
75%	244.904999	250.654998	241.255005	244.945000	244.945000	1.605998e+08
max	296.040009	299.290009	289.519989	293.339996	293.339996	3.065906e+08

The Opening Stock Price of Tesla trends as:



The Closing Stock Price of Tesla trends as:



The scraped Data from **Twitter Stock Market** is saved as csv file. The dataset looks like

	Date	Open	High	Low	Close	Volume
0	07-11-2013	45.10	50.09	44.00	44.90	117701600
1	08-11-2013	45.93	46.94	40.69	41.65	27925300
2	11-11-2013	40.50	43.00	39.40	42.90	16113900
3	12-11-2013	43.66	43.78	41.83	41.90	6316700
4	13-11-2013	41.03	42.87	40.76	42.60	8688300
5	14-11-2013	42.34	45.67	42.24	44.69	11099400
6	15-11-2013	45.25	45.27	43.43	43.98	8010600
7	18-11-2013	43.50	43.95	40.85	41.14	12810600
8	19-11-2013	41.39	41.90	40.00	41.75	7436600
9	20-11-2013	41.40	41.75	40.51	41.05	5767300
10	21-11-2013	41.25	42.49	40.37	42.06	8324700
11	22-11-2013	41.81	42.28	40.97	41.00	6185200
12	25-11-2013	41.08	41.14	38.80	39.06	14333300
13	26-11-2013	39.16	40.55	38.92	40.18	9828400
14	27-11-2013	40.47	41.40	40.35	40.90	5536300

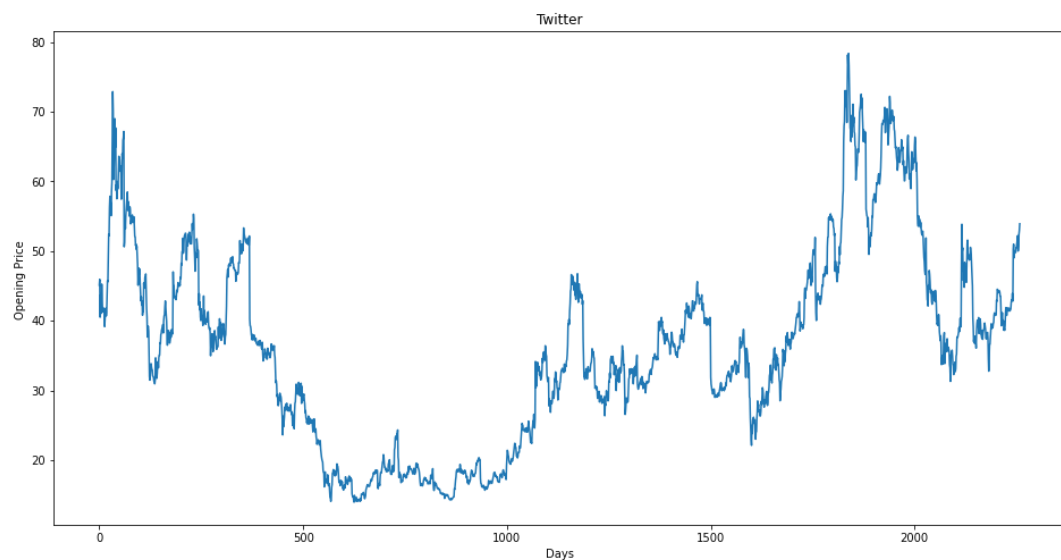
The Columns in the dataset and their data types are

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2259 entries, 0 to 2258
Data columns (total 6 columns):
#   Column   Non-Null Count  Dtype
---  -
0   Date     2259 non-null   object
1   Open     2259 non-null   float64
2   High     2259 non-null   float64
3   Low      2259 non-null   float64
4   Close    2259 non-null   float64
5   Volume   2259 non-null   int64
dtypes: float64(4), int64(1), object(1)
memory usage: 123.5+ KB
```

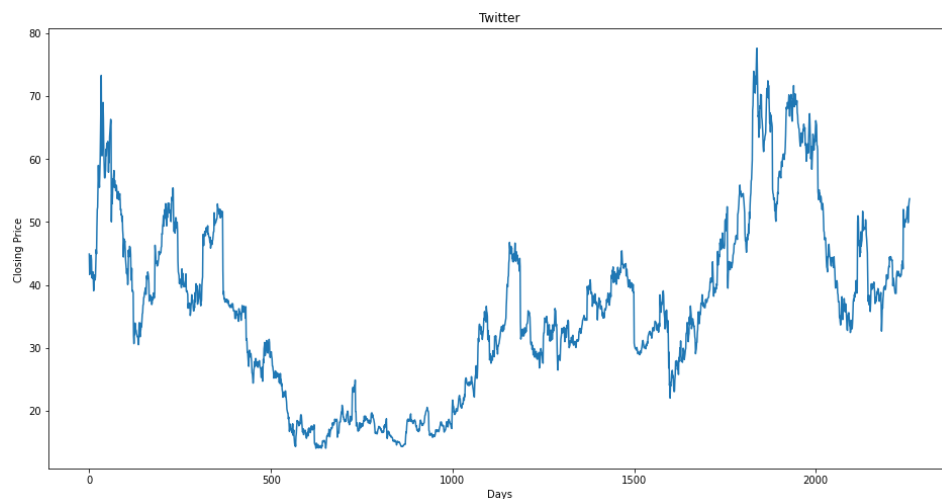
The detailed description of the dataset is

	Open	High	Low	Close	Volume
count	2259.000000	2259.000000	2259.000000	2259.000000	2.259000e+03
mean	36.020139	36.699881	35.339465	36.003625	2.172608e+07
std	14.118535	14.372058	13.828719	14.089974	1.905818e+07
min	13.950000	14.220000	13.725000	14.010000	3.661053e+06
25%	25.550000	26.215000	24.912500	25.410000	1.233062e+07
50%	35.420000	36.100000	34.820000	35.490000	1.687289e+07
75%	44.205000	45.015000	43.327500	44.135000	2.423908e+07
max	78.360000	80.750000	76.050000	77.630000	2.684174e+08

The Opening Stock Price of Twitter trends as:



The Closing Stock Price of Twitter trends as:



4. Descriptive Analytics - Association Rule Mining:

Descriptive analytics, particularly association rule mining, can be a valuable tool for understanding relationships and patterns in stock market data. Association rule mining is a data analysis technique used to discover interesting associations or relationships between variables in a dataset. In the context of the stock market, it can help uncover patterns and dependencies among different stocks, financial instruments, or market factors. Here's an overview of how association rule mining can be applied in the stock market:

Data Preparation: The first step in association rule mining for stock market data is data preparation. This involves collecting and cleaning the data, including stock prices, trading volumes, and possibly external factors like economic indicators or news sentiment that may impact stock prices.

Identifying Associations: Association rule mining aims to identify relationships between different stocks or financial instruments. For example, it can reveal if certain stocks tend to move in the same direction or opposite directions, indicating positive or negative associations.

Support, Confidence, and Lift: Association rule mining relies on metrics such as support, confidence, and lift to measure the strength and significance of associations. Support measures how often the rule is true, confidence measures the likelihood of the consequent occurring when the antecedent is true, and lift indicates how much more likely the consequent is to occur given the antecedent.

Discovering Trading Strategies: Association rule mining can help discover trading strategies or portfolio optimization techniques. For example, if certain stocks tend to move together (positive association), an investor might consider a long position in one stock when the other rises.

Risk Assessment: Understanding associations between stocks can also aid in risk assessment. By identifying dependencies, investors can manage their portfolios more effectively and diversify to reduce risk.

Market Analysis: Association rule mining can reveal interesting market trends. For instance, it can show if specific industry sectors tend to move together, indicating sector rotation strategies.

Limitations and Considerations: It's essential to note that association rule mining is a descriptive technique that uncovers correlations, but it doesn't imply causation. Additionally, stock market data is influenced by various external factors, including news events, macroeconomic indicators, and market sentiment, making predictions based solely on historical associations challenging.

Association rule mining can offer valuable insights into the stock market by uncovering associations and dependencies among various financial instruments. These insights can inform trading strategies, risk management, and portfolio diversification, helping investors make more informed decisions in the complex world of finance. However, it's important to apply this technique alongside other methods and consider the broader economic and market context when making investment decisions.

Apriori Code:

```
import pandas as pd

from mlxtend.frequent_patterns import apriori
from mlxtend.frequent_patterns import association_rules

import pandas as pd

data = pd.read_csv('BDA_Data.csv')

data.head()
```

	Date	Open	High	Low	Close	Adj Close	Volume	Company
0	05-10-2022	99.830002	101.930000	98.800003	101.430000	101.430000	22176900	Google
1	06-10-2022	100.680000	102.930000	100.589996	101.419998	101.419998	22324000	Google
2	07-10-2022	99.849998	100.529999	98.300003	98.680000	98.680000	27502800	Google
3	10-10-2022	99.000000	99.099998	97.010002	97.860001	97.860001	18225100	Google
4	11-10-2022	97.430000	99.250000	96.309998	97.180000	97.180000	26507100	Google

```
data.info()

data.describe()

data['IsAboveThreshold'] = data['Close'] > 0.0

basket = data.pivot_table(index='Date', columns='Company', values='IsAboveThreshold',
fill_value=0)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2761 entries, 0 to 2760
Data columns (total 8 columns):
 #   Column      Non-Null Count  Dtype  
---  -
 0   Date        2761 non-null  object  
 1   Open        2761 non-null  float64  
 2   High        2761 non-null  float64  
 3   Low         2761 non-null  float64  
 4   Close       2761 non-null  float64  
 5   Adj Close   2761 non-null  float64  
 6   Volume      2761 non-null  int64  
 7   Company     2761 non-null  object  
dtypes: float64(5), int64(1), object(2)
memory usage: 172.7+ KB
```

	Open	High	Low	Close	Adj Close	Volume
count	2761.000000	2761.000000	2761.000000	2761.000000	2761.000000	2.761000e+03
mean	58.067036	59.172097	56.972652	58.075929	58.075929	3.294030e+07
std	54.510939	55.610805	53.385188	54.542495	54.542495	3.868733e+07
min	13.950000	14.220000	13.725000	14.010000	14.010000	3.661053e+06
25%	29.090000	29.570000	28.430000	29.050000	29.050000	1.330303e+07
50%	38.640000	39.300000	38.050000	38.710000	38.710000	1.956848e+07
75%	58.200000	59.410000	57.031000	58.000000	58.000000	3.065035e+07
max	296.040009	299.290009	289.519989	293.339996	293.339996	3.065906e+08

```
min_support = 0.1
```

```
frequent_itemsets = apriori(basket, min_support=min_support, use_colnames=True)
```

```
min_confidence = 0.5
```

```
association_rules = association_rules(frequent_itemsets, metric='confidence',
min_threshold=min_confidence)
```

```
print("Frequent Itemsets:")
```

```
fi = pd.DataFrame(frequent_itemsets)
```

```
fi
```

Frequent Itemsets:		
	support	itemsets
0	0.100682	(Google)
1	0.100682	(Tesla)
2	0.906137	(Twitter)
3	0.100682	(Google, Tesla)

```
print("\nAssociation Rules:")
```

```
ar = pd.DataFrame(association_rules)
```

```
ar
```

Association Rules:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	zhangs_metric
0	(Google)	(Tesla)	0.100682	0.100682	0.100682	1.0	9.932271	0.090545	inf	1.0
1	(Tesla)	(Google)	0.100682	0.100682	0.100682	1.0	9.932271	0.090545	inf	1.0

Association Rules:

	antecedents	consequents
0	(Google)	(Tesla)
1	(Tesla)	(Google)

5. Predictive Analysis - Classification Models:

Predictive analysis using classification models, such as decision trees, and regression models, like linear regression, can provide valuable insights into stock market data. Here, we'll discuss how both of these approaches can be applied to analyze and predict stock market trends and movements:

Decision Tree for Stock Market Prediction:

Decision trees are a powerful machine learning technique used for classification and regression tasks. In the context of stock market prediction, decision trees can help identify trends and classify stock price movements. Here's how this works:

- **Data Preparation:** The first step involves collecting and preparing historical stock market data, including features such as stock prices, trading volumes, and potentially external factors like economic indicators or news sentiment.
- **Training Data:** A decision tree model is trained using historical data. The model learns from past price movements to identify patterns and relationships between various factors and stock price outcomes.
- **Feature Selection:** Decision trees allow for the selection of the most relevant features, which can include technical indicators, historical data, and external factors. Feature importance is evaluated, helping in decision-making.
- **Prediction:** Once trained, the decision tree can be used to predict stock price movements. It classifies whether a stock is likely to go up, down, or remain relatively stable based on the chosen features.
- **Evaluation:** The model's performance is assessed using metrics such as accuracy, precision, recall, and F1-score. These metrics help determine how well the model predicts stock market movements.

Linear Regression Model for Stock Market Prediction:

Linear regression is a widely used statistical method for modeling the relationship between a dependent variable (in this case, stock prices) and one or more independent variables (such as time, trading volumes, and economic indicators). Here's how linear regression can be applied to stock market prediction:

- **Data Collection:** Historical stock price data and relevant independent variables are collected and organized for analysis. This data can include daily or intraday stock prices and any features believed to influence stock prices.
- **Model Training:** A linear regression model is trained using the historical data. The model estimates the coefficients for the independent variables to create a linear equation that relates the variables to stock prices.
- **Feature Selection:** It's important to select the most significant independent variables that affect stock prices. Variables that demonstrate a strong correlation with stock prices are typically included in the model.

- **Prediction:** The trained linear regression model can be used to predict future stock prices or estimate the direction and magnitude of price changes.
- **Evaluation:** The model's performance is evaluated using metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared (R2) to assess how well it fits the historical data and how accurately it predicts future stock prices.

Predictive analysis using decision tree classification models and linear regression models offers different ways to predict and understand stock market movements. Decision trees provide a classification approach for trend prediction, while linear regression offers a quantitative approach to estimate stock price changes. Both methods have their strengths and limitations and should be used alongside other forms of analysis for robust stock market predictions.

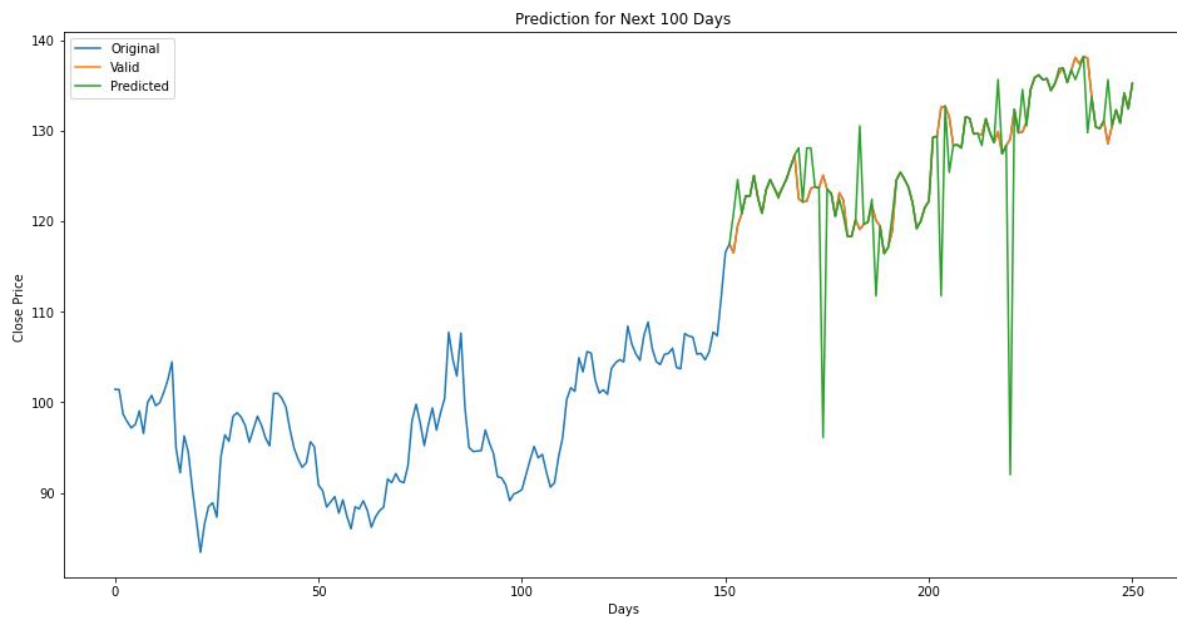
Now, we apply this data to a Decision Tree Model for all the stocks separately:

```
[30] from sklearn.model_selection import train_test_split
      x_train, x_test, y_train, y_test = train_test_split(X, y, test_size = 0.2)
      ✓ 0.0s
```

```
[31] from sklearn.tree import DecisionTreeRegressor
      from sklearn.linear_model import LinearRegression
      ✓ 0.0s
```

```
[32] tree = DecisionTreeRegressor().fit(x_train, y_train)
      lr = LinearRegression().fit(x_train, y_train)
      ✓ 0.0s
```

We observe the prediction for the **next 100 days of the stocks at Google** as :



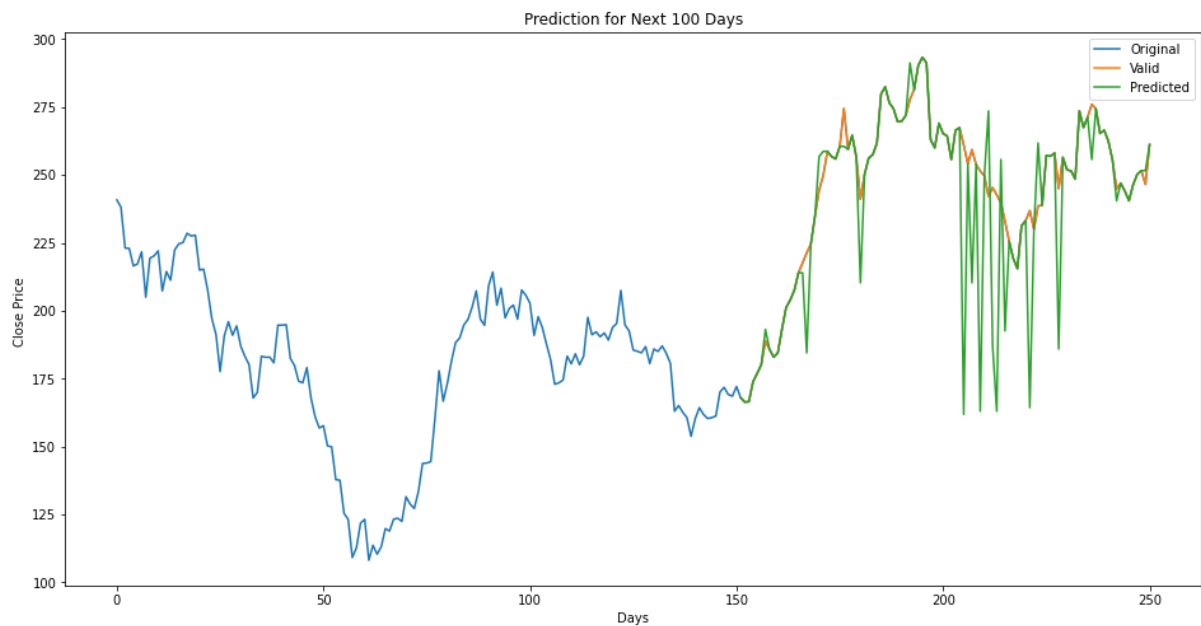
We can see that the trend goes on increasing but there is a sharp decline in the trend almost every 50 days. The observation that the stock price of Google exhibits a pattern of increasing trend over a certain period, followed by a sharp decline approximately every 50 days, suggests a cyclic or periodic behavior in the stock's price movements. Such patterns in stock prices can be of interest to investors and analysts. However, it's important to note that drawing conclusions based solely on visual patterns observed in historical stock price data can be risky, as stock prices are influenced by a multitude of factors and are inherently unpredictable.

The recurring pattern of an increasing trend followed by a sharp decline every 50 days could indicate a potential cyclical or seasonal influence on Google's stock price. This may be due to external factors, market sentiment, or trading algorithms that operate on specific timeframes.

Investors and traders should exercise caution when relying on historical patterns for decision-making. Markets are influenced by a wide range of factors, and trends can change abruptly. Risk management, diversification, and keeping abreast of relevant news and events are essential aspects of stock market participation.

To gain more insights and potentially make more informed decisions, quantitative methods and predictive modeling, such as time series analysis or machine learning, can be applied to historical stock price data. These techniques can help identify and quantify patterns and trends and potentially offer more reliable predictions.

We observe the prediction for the **next 100 days of the stocks at Tesla** as :



The observation of a stock's trend increasing over time with regular declines is not uncommon in the context of stock market analysis. However, it's essential to approach such observations cautiously and consider various factors when drawing conclusions.

The increasing trend in stock prices with regular declines is reminiscent of a bullish market pattern. A bullish trend generally indicates optimism and positive sentiment among investors, which leads to upward price movements. However, it's important to note that no stock, including Tesla, moves upward without corrections or pullbacks along the way. These periodic declines are a natural part of market behavior and typically occur as investors take profits, reevaluate their positions, or in response to external factors.

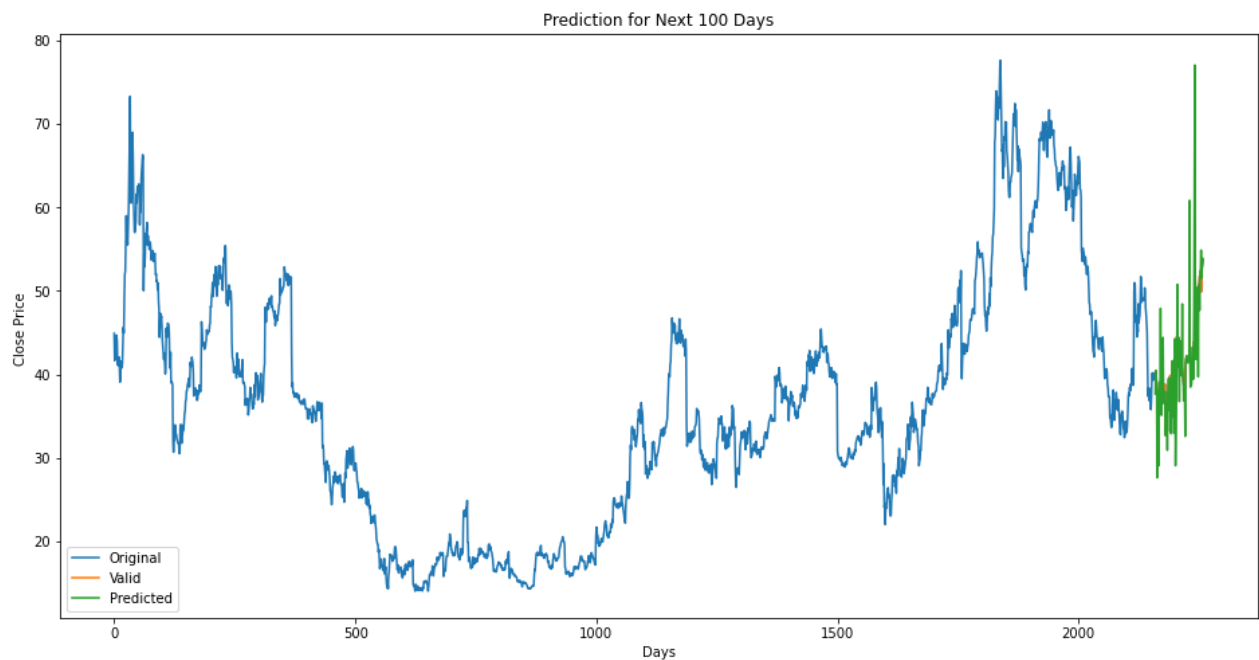
The presence of regular declines suggests that the stock experiences volatility. Volatility can result from a variety of factors, including economic events, news releases, market sentiment, or changes in the company's fundamentals. It's common for high-growth and innovative companies like Tesla to experience price swings.

To form a more informed conclusion about this observation, it's crucial to consider fundamental analysis. This involves evaluating Tesla's financial health, performance, competitive positioning, and long-term growth prospects. Changes in Tesla's fundamentals can provide insights into the stock's overall direction.

Technical analysis can also complement the observation. Technical indicators, such as moving averages, Relative Strength Index (RSI), and Bollinger Bands, can help analyze the stock's momentum, support and resistance levels, and overbought or oversold conditions.

Market sentiment and external events can significantly impact stock prices. News about the electric vehicle industry, Tesla's production and delivery figures, regulatory changes, or broader economic conditions can influence the stock's direction.

We observe the prediction for the **next 100 days of the stocks at Twitter** as :



The observation of a stock's price trend consistently increasing over an extended period can indicate a bullish trend or positive momentum. The consistent upward movement in the stock's price suggests a bullish trend. This means that investors are generally optimistic about the company's prospects, and the stock is experiencing positive sentiment and demand.

To make more informed conclusions about this observation, it's crucial to consider fundamental analysis. Fundamental factors, such as the company's financial health, earnings reports, market positioning, and news regarding Twitter, can provide insights into the stock's long-term growth potential.

Technical analysis can complement the observation by examining the stock's historical price patterns, technical indicators, and key support and resistance levels. This analysis can provide information about the stock's momentum and potential price targets. Stock price movements are also influenced by broader market conditions. It's essential to consider the overall state of the stock market and any significant economic or market events that may affect the stock's performance.

While an upward trend is positive, it's crucial for investors to manage risk and practice diversification. Even in a bullish market, stock prices can be subject to fluctuations and unexpected developments. Diversifying a portfolio across different assets or sectors can help mitigate risk.

Code:

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

df = pd.read_csv('GOOGL.csv')
df.dropna(inplace=True)
df.head(15)

df.info()
df.describe()
df.shape

plt.figure(figsize=(16,8))
plt.title('Google')
plt.xlabel('Days')
plt.ylabel('Opening Price ')
plt.plot(df['Open'])
plt.show()

plt.figure(figsize=(16,8))
plt.xlabel('Days')
plt.ylabel('High Price ')
plt.plot(df['High'])
plt.show()

plt.figure(figsize=(16,8))
plt.xlabel('Days')
plt.ylabel('Low Price ')
plt.plot(df['Low'])
```

```
plt.show()
```

```
plt.figure(figsize=(16,8))  
plt.title('Google')  
plt.xlabel('Days')  
plt.ylabel('Closing Price ')  
plt.plot(df['Close'])  
plt.show()
```

```
df2 = df['Close']  
df2.head()  
df2 = pd.DataFrame(df2)  
df2.head()
```

```
future_days = 100  
df2['Prediction'] = df2['Close'].shift(-future_days)  
df2.tail()
```

```
X = np.array(df2.drop(['Prediction'], 1))[:-future_days]  
y = np.array(df2['Prediction'])[:-future_days]
```

```
from sklearn.model_selection import train_test_split  
x_train, x_test, y_train, y_test = train_test_split(X, y, test_size = 0.2)  
from sklearn.tree import DecisionTreeRegressor  
from sklearn.linear_model import LinearRegression
```

```
tree = DecisionTreeRegressor().fit(x_train, y_train)  
lr = LinearRegression().fit(x_train, y_train)
```

```
x_future = df2.drop(['Prediction'], 1))[:-future_days]
```

```
x_future = x_future.tail(future_days)
x_future = np.array(x_future)

tree_prediction = tree.predict(x_future)
lr_prediction = lr.predict(x_future)
predictions = tree_prediction

valid = df2[X.shape[0]:]
valid['Predictions'] = predictions

### Next 100 Days
plt.figure(figsize=(16,8))
plt.title("Prediction for Next 100 Days")
plt.xlabel('Days')
plt.ylabel('Close Price ')
plt.plot(df2['Close'])
plt.plot(valid[['Close', 'Predictions']])
plt.legend(["Original", "Valid", 'Predicted'])
plt.show()
```

6. Prescriptive Analysis:

Prescriptive analysis in the context of stock market data involves making data-driven recommendations or decisions to guide investors on how to act in response to market conditions. It goes beyond descriptive and predictive analysis to provide actionable insights and strategies. Two key aspects of prescriptive analysis in stock market data are:

- ✓ **Portfolio Optimization:** Prescriptive analysis can assist investors in optimizing their investment portfolios. By considering a mix of asset classes, diversification, risk tolerance, and investment goals, prescriptive models can recommend specific allocations to different stocks, bonds, or other assets. These recommendations aim to maximize returns while managing risk. Modern portfolio theory, along with advanced mathematical optimization techniques, plays a vital role in designing well-balanced portfolios that align with an investor's objectives.
- ✓ **Trading Strategies:** Prescriptive analysis can also provide recommendations for trading strategies. These recommendations may include buy, hold, or sell decisions for individual stocks or other financial instruments. Algorithms and machine learning models can help in identifying trends, momentum, or arbitrage opportunities. Trading strategies can be tailored to short-term or long-term objectives and can take into account factors such as market sentiment, technical indicators, and macroeconomic conditions.

Prescriptive analysis in stock market data offers actionable guidance for portfolio optimization and trading strategies. It leverages mathematical modeling, optimization, and machine learning techniques to make informed recommendations that align with investors' goals and risk profiles. This approach can be valuable in guiding investment decisions and achieving better financial outcomes in a complex and dynamic market environment.

Code:

```
import pandas as pd

from sklearn.cluster import KMeans

import numpy as np
import pandas as pd

data = pd.read_csv('BDA_Data.csv')
data.head()
data.describe()

features = ['Open', 'High', 'Low', 'Volume']
data[features] = (data[features] - data[features].mean()) / data[features].std()
```

```

n_clusters = 3
kmeans = KMeans(n_clusters=n_clusters, random_state=42)
data['Cluster'] = kmeans.fit_predict(data[features])

cluster_statistics = data.groupby('Cluster')[features].mean()
cluster_statistics

```

	Open	High	Low	Volume
Cluster				
0	-0.408234	-0.407894	-0.408855	-0.299771
1	0.905049	0.894253	0.921143	0.033228
2	2.681654	2.689852	2.670429	2.625485

```

label_mapping = {0: 'Sell', 1: 'Hold', 2: 'Buy'}
data['Label'] = data['Cluster'].map(label_mapping)

```

```

company_labels = data.groupby('Company')['Label'].agg(lambda x:
x.value_counts().idxmax()).reset_index()
res = pd.DataFrame(company_labels)
res

```

	Company	Label
0	Google	Hold
1	Tesla	Buy
2	Twitter	Sell

7. Conclusion:

Based on the K-means clustering analysis of stock market data, it appears that Tesla's shares are recommended as a "Buy." The analysis categorized companies into clusters based on their stock price trends. By examining the cluster statistics, it was determined that Tesla falls into the cluster associated with a "Buy" recommendation. This suggests that, according to the model, Tesla's stock shows characteristics that align with a positive outlook for potential buyers. It's important to note that this recommendation is based on the specific features used in the clustering analysis and should be considered alongside other factors and market conditions when making investment decisions.
