

# Netflix Data Analysis-

## Objective of this project-

1. Volume of Stock Traded
2. Netflix Stock Price - High, Open, Close
3. Netflix Stock Price - Day, Month, Year wise
4. Top 5 Dates With Highest Stock Price
5. Top 5 Dates with Lowest Stock Price

## Importing libraries-

```
In [80]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px

from datetime import datetime
```

## Loading the Dataset-

```
In [3]: data = pd.read_csv("Netflix.csv")
data.head()
```

```
Out[3]:
```

	Date	Open	High	Low	Close	Adj Close	Volume
0	2002-05-23	1.156429	1.242857	1.145714	1.196429	1.196429	104790000
1	2002-05-24	1.214286	1.225000	1.197143	1.210000	1.210000	11104800
2	2002-05-28	1.213571	1.232143	1.157143	1.157143	1.157143	6609400
3	2002-05-29	1.164286	1.164286	1.085714	1.103571	1.103571	6757800
4	2002-05-30	1.107857	1.107857	1.071429	1.071429	1.071429	10154200

```
In [4]: data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4874 entries, 0 to 4873
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Date        4874 non-null   object
1   Open        4874 non-null   float64
2   High        4874 non-null   float64
3   Low         4874 non-null   float64
4   Close       4874 non-null   float64
5   Adj Close   4874 non-null   float64
6   Volume      4874 non-null   int64
dtypes: float64(5), int64(1), object(1)
memory usage: 266.7+ KB
```

## Convert Date datatype into Datetime-

```
In [5]: data["Date"] = pd.to_datetime(data["Date"])
data.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4874 entries, 0 to 4873
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype  
---  -
0   Date        4874 non-null   datetime64[ns]
1   Open        4874 non-null   float64
2   High        4874 non-null   float64
3   Low         4874 non-null   float64
4   Close       4874 non-null   float64
5   Adj Close   4874 non-null   float64
6   Volume      4874 non-null   int64   
dtypes: datetime64[ns](1), float64(5), int64(1)
memory usage: 266.7 KB

```

## Setting Date Column in Index-

```

In [7]: data = data.set_index("Date")
        data.head()

```

```

Out[7]:

```

	Open	High	Low	Close	Adj Close	Volume
<b>Date</b>						
<b>2002-05-23</b>	1.156429	1.242857	1.145714	1.196429	1.196429	104790000
<b>2002-05-24</b>	1.214286	1.225000	1.197143	1.210000	1.210000	11104800
<b>2002-05-28</b>	1.213571	1.232143	1.157143	1.157143	1.157143	6609400
<b>2002-05-29</b>	1.164286	1.164286	1.085714	1.103571	1.103571	6757800
<b>2002-05-30</b>	1.107857	1.107857	1.071429	1.071429	1.071429	10154200

## Discriptive Analysis-

```

In [8]: data.describe()

```

Out[8]:

	Open	High	Low	Close	Adj Close	Volume
<b>count</b>	4874.000000	4874.000000	4874.000000	4874.000000	4874.000000	4.874000e+03
<b>mean</b>	104.644811	106.243736	103.006660	104.686524	104.686524	1.681687e+07
<b>std</b>	155.886656	158.173217	153.546958	155.940368	155.940368	1.937625e+07
<b>min</b>	0.377857	0.410714	0.346429	0.372857	0.372857	2.856000e+05
<b>25%</b>	3.901429	3.977143	3.835357	3.905714	3.905714	6.301225e+06
<b>50%</b>	21.386429	21.816429	20.935714	21.233572	21.233572	1.084630e+07
<b>75%</b>	125.445000	126.592499	122.987499	125.097502	125.097502	2.016192e+07
<b>max</b>	608.049988	619.000000	608.049988	610.340027	610.340027	3.234140e+08

## Checking Null Values-

In [9]: `data.isnull().sum()`

Out[9]:

Open	0
High	0
Low	0
Close	0
Adj Close	0
Volume	0

dtype: int64

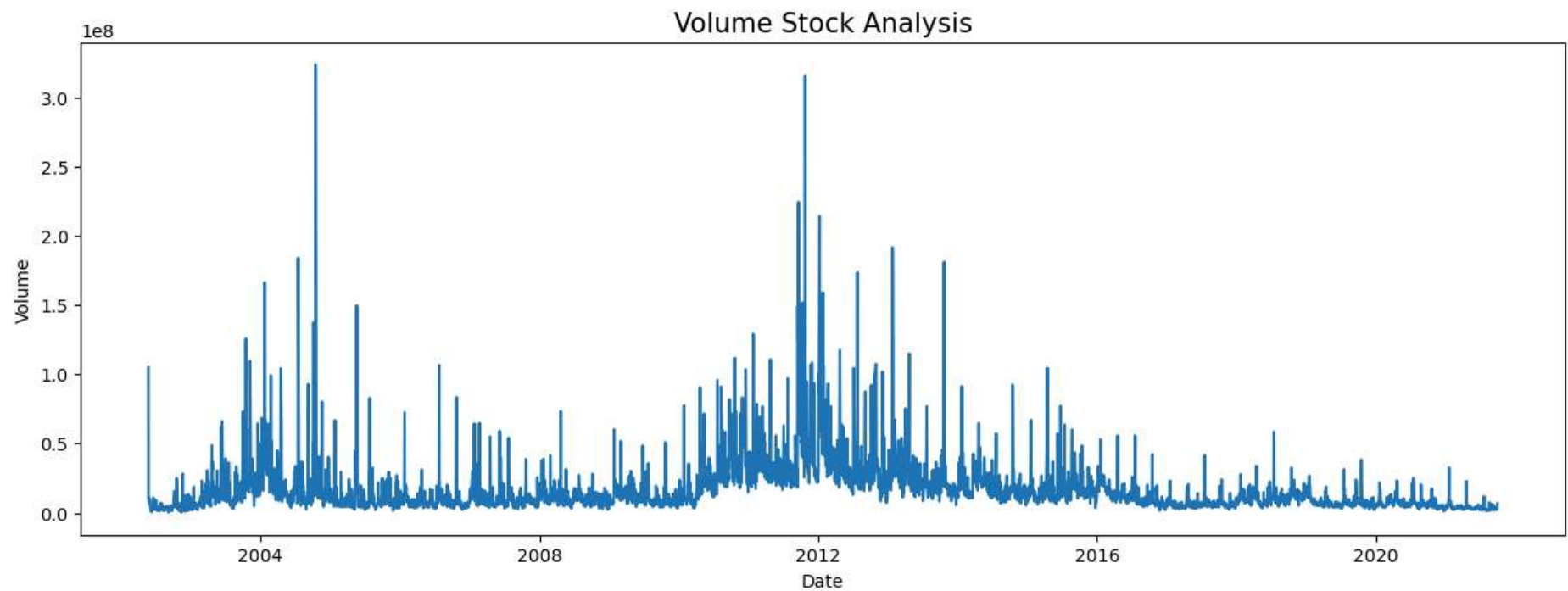
## Volume of Stock Traded-

In [10]: `data.head()`

Out[10]:

	Open	High	Low	Close	Adj Close	Volume
Date						
2002-05-23	1.156429	1.242857	1.145714	1.196429	1.196429	104790000
2002-05-24	1.214286	1.225000	1.197143	1.210000	1.210000	11104800
2002-05-28	1.213571	1.232143	1.157143	1.157143	1.157143	6609400
2002-05-29	1.164286	1.164286	1.085714	1.103571	1.103571	6757800
2002-05-30	1.107857	1.107857	1.071429	1.071429	1.071429	10154200

```
In [44]: plt.figure(figsize=[15,5])
sns.lineplot(data=data, x="Date", y="Volume")
plt.title("Volume Stock Analysis", fontsize=15)
plt.show()
```



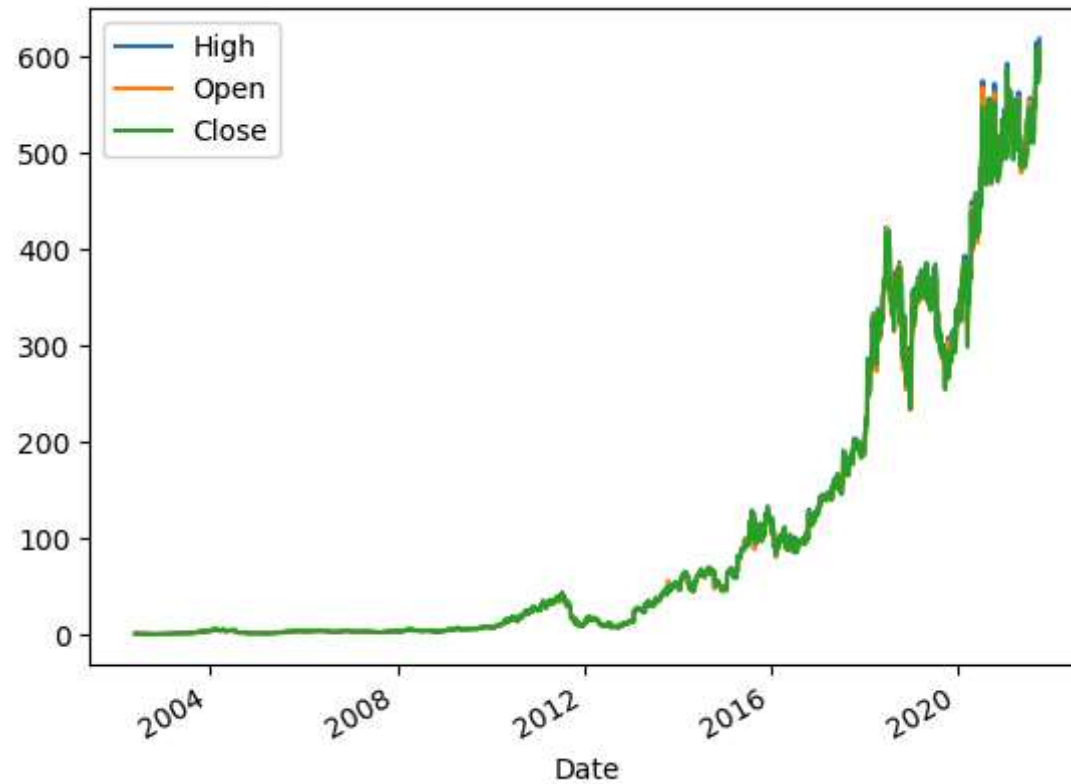
Netflix Stock Price - High, Open, Close-

```
In [48]: data.head()
```

Out[48]:

	Open	High	Low	Close	Adj Close	Volume
Date						
2002-05-23	1.156429	1.242857	1.145714	1.196429	1.196429	104790000
2002-05-24	1.214286	1.225000	1.197143	1.210000	1.210000	11104800
2002-05-28	1.213571	1.232143	1.157143	1.157143	1.157143	6609400
2002-05-29	1.164286	1.164286	1.085714	1.103571	1.103571	6757800
2002-05-30	1.107857	1.107857	1.071429	1.071429	1.071429	10154200

```
In [59]: data.plot(y=["High", "Open", "Close"])
plt.show()
```



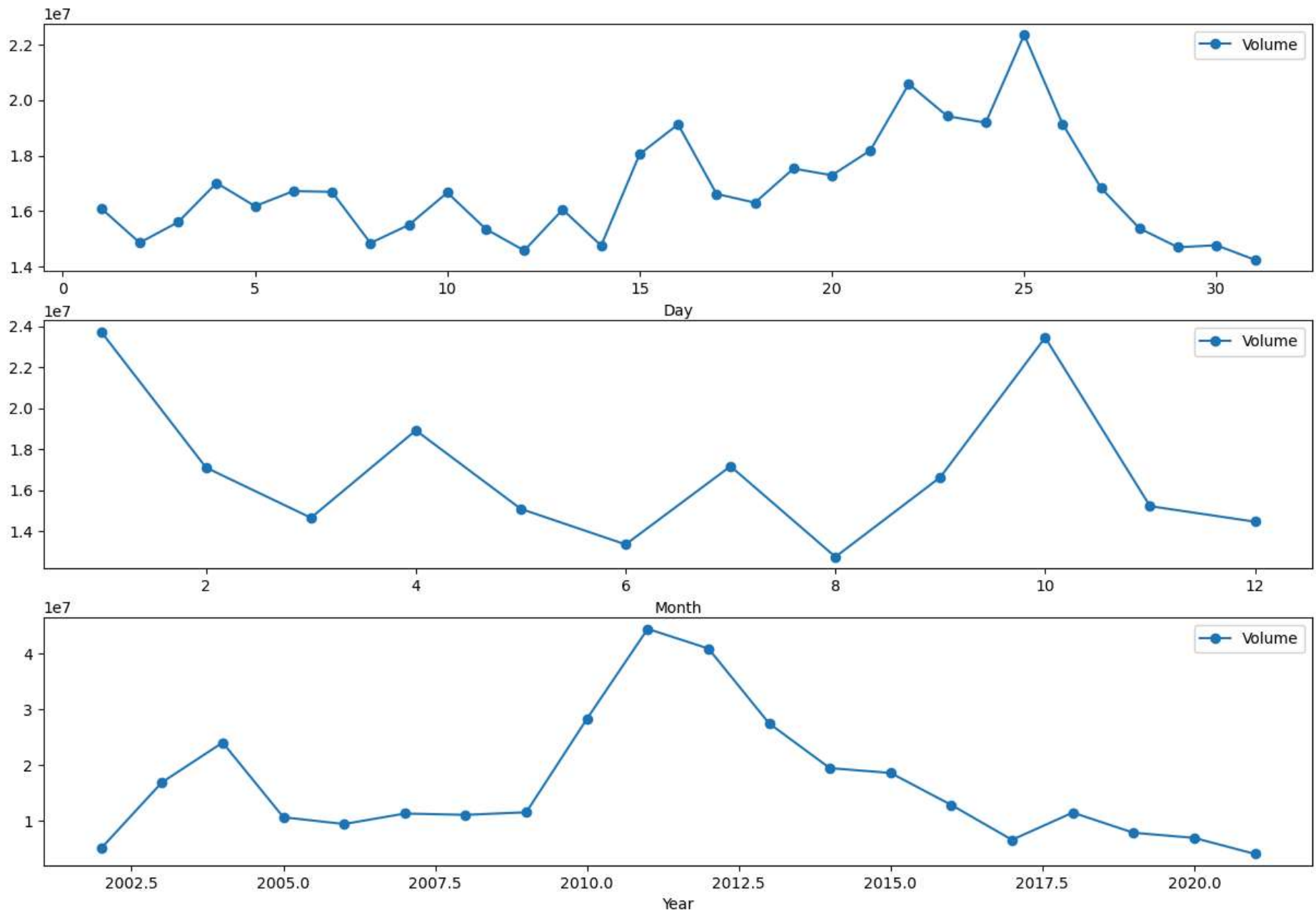
## Netflix Stock Price - Day, Month, Year wise-

```
In [77]: fig, (ax1, ax2, ax3) = plt.subplots(3, figsize=[15,10])

data.groupby(data.index.day).mean().plot(y="Volume", ax=ax1, xlabel="Day", marker="o")
data.groupby(data.index.month).mean().plot(y="Volume", ax=ax2, xlabel="Month", marker="o")
data.groupby(data.index.year).mean().plot(y="Volume", ax=ax3, xlabel="Year", marker="o")

plt.suptitle("Netflix Stock Price - Day, Month, Year wise", fontsize=15)
plt.show()
```

Netflix Stock Price - Day, Month, Year wise





## Top 5 Dates With Highest/ Lowest Stock Price-

```
In [60]: data.head()
```

```
Out[60]:
```

	Open	High	Low	Close	Adj Close	Volume
Date						
2002-05-23	1.156429	1.242857	1.145714	1.196429	1.196429	104790000
2002-05-24	1.214286	1.225000	1.197143	1.210000	1.210000	11104800
2002-05-28	1.213571	1.232143	1.157143	1.157143	1.157143	6609400
2002-05-29	1.164286	1.164286	1.085714	1.103571	1.103571	6757800
2002-05-30	1.107857	1.107857	1.071429	1.071429	1.071429	10154200

```
In [72]: a=data.sort_values(by="High", ascending=False).head()  
a["High"]
```

```
Out[72]:
```

Date	High
2021-09-30	619.000000
2021-09-08	615.599976
2021-09-07	613.849976
2021-09-29	609.880005
2021-09-10	609.450012

Name: High, dtype: float64

```
In [74]: b = data.sort_values(by="Low", ascending=True).head()  
b["Low"]
```

```
Out[74]:
```

Date	Low
2002-10-10	0.346429
2002-10-09	0.347143
2002-10-07	0.382143
2002-10-08	0.390714
2002-10-16	0.442857

Name: Low, dtype: float64

```
In [90]: fig,axes=plt.subplots(nrows=1, ncols=2, sharex = True, figsize=(15,5))  
fig.suptitle("High & Low Values Stock per period of time", fontsize=15)
```

```
sns.lineplot(ax=axes[0], y=data["High"], x=data.index, color='green')  
sns.lineplot(ax=axes[1], y=data["Low"], x=data.index, color='blue')
```

Out[90]: <Axes: xlabel='Date', ylabel='Low'>

High & Low Values Stock per period of time

