

TASK 1- STOCK PREDICTION

```
[10] pip install seaborn
Defaulting to user installation because normal site-packages is not writable
Requirement already satisfied: seaborn in c:\users\anu\shindevapdata\roaming\python\python311\site-packages (0.13.0)
Requirement already satisfied: numpy>=1.24.0, <=1.28 in c:\users\anu\shindevapdata\roaming\python\python311\site-packages (from seaborn) (1.24.2)
Requirement already satisfied: pandas in c:\users\anu\shindevapdata\roaming\python\python311\site-packages (from seaborn) (1.5.3)
Requirement already satisfied: matplotlib>=3.6.1, <=3.7.5 in c:\users\anu\shindevapdata\roaming\python\python311\site-packages (from seaborn) (3.8.0)
Requirement already satisfied: contourpy>=0.9.1, <=1.0 in c:\users\anu\shindevapdata\roaming\python\python311\site-packages (from matplotlib>=3.6.1, <=3.7.5->seaborn) (1.1.1)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\anu\shindevapdata\roaming\python\python311\site-packages (from matplotlib>=3.6.1, <=3.7.5->seaborn) (4.4.1)
Requirement already satisfied: kiwisolver>=0.1 in c:\users\anu\shindevapdata\roaming\python\python311\site-packages (from matplotlib>=3.6.1, <=3.7.5->seaborn) (0.1.4)
Requirement already satisfied: packaging>20.0 in c:\program files\python\python311\site-packages (from matplotlib>=3.6.1, <=3.7.5->seaborn) (23.0)
Requirement already satisfied: pillow>=6.2.0 in c:\users\anu\shindevapdata\roaming\python\python311\site-packages (from matplotlib>=3.6.1, <=3.7.5->seaborn) (10.0.1)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\anu\shindevapdata\roaming\python\python311\site-packages (from matplotlib>=3.6.1, <=3.7.5->seaborn) (3.1.1)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\anu\shindevapdata\roaming\python\python311\site-packages (from matplotlib>=3.6.1, <=3.7.5->seaborn) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in c:\users\anu\shindevapdata\roaming\python\python311\site-packages (from pandas>1.2->seaborn) (2023.2)
Requirement already satisfied: tzdata in c:\users\anu\shindevapdata\roaming\python\python311\site-packages (from pandas-dateutil>=2.7, matplotlib>=3.6.1, <=3.7.5->seaborn) (2016.4)
note: you may need to restart the terminal to use updated packages.
[notice] A new release of pip is available: 23.2 -> 23.2.1
[notice] To update, run: python.exe -m pip install --upgrade pip

[11] pip install statsmodels
Defaulting to user installation because normal site-packages is not writable
Collecting statsmodels
  Downloading statsmodels-0.14.0-cp311-cp311-win_amd64.whl (9.2 MB)
```

```
[37] pip install pandas
```

```
Note: you may need to restart the kernel to use updated packages; defaulting to user installation because normal site-packages is not writable  
Requirement already satisfied: pandas in c:\users\amu\shindevadptar\roaming\python\python311\site-packages (1.5.2)  
Requirement already satisfied: python-dateutil<2.8.1,>=2.8.1 in c:\users\amu\shindevadptar\roaming\python\python311\site-packages (from pandas) (2.8.2)  
Requirement already satisfied: numpy>=2020.1.9 in c:\users\amu\shindevadptar\roaming\python\python311\site-packages (from pandas) (2023.2)  
Requirement already satisfied: numba>=23.0 in c:\users\amu\shindevadptar\roaming\python\python311\site-packages (from pandas) (0.60.0)  
Requirement already satisfied: six>=1.16.0 in c:\users\amu\shindevadptar\roaming\python\python311\site-packages (from pandas) (1.16.0)
```

```
[notice] A new release of pip is available: 23.2 -> 23.2.1  
[notice] To update, run: python.exe -m pip install --upgrade pip
```

```
In [39]: pip install --upgrade pip
```

```
defaulting to user installation because normal site-packages is not writable  
Requirement already satisfied: pip in c:\users\amu\shindevadptar\roaming\python\python311\site-packages (23.2)  
Collecting pip  
Obtaining dependency information for https://files.pythonhosted.org/packages/5f/c2/e9d85e8cc2bdac7c15f473db8a33bc06dc704ca10931b3bd85de2968a/pip-23.2.1-py3-none-any.whl  
1.metadata  
Downloading pip-23.2.1-py3-none-any.whl.metadata (4.2 kB)  
Downloading pip-23.2.1-py3-none-any.whl (2.1 MB)  
  
-----  
0.0/2.1 MB 7 eta <-:-:  
-----  
2.2/2.1 MB 5.3 MB/s eta 0:00:01  
-----  
2.2/2.1 MB 3.1 MB/s eta 0:00:01  
-----  
3.2/2.1 MB 2.3 MB/s eta 0:00:01  
-----  
3.3/2.1 MB 2.2 MB/s eta 0:00:01  
-----  
4.2/2.1 MB 2.0 MB/s eta 0:00:01
```

```

..... 0.5/2.1 MB 1.7 MB/s eta 0:00:01
..... 0.6/2.1 MB 1.6 MB/s eta 0:00:01
..... 0.7/2.1 MB 1.6 MB/s eta 0:00:01
..... 0.7/2.1 MB 1.6 MB/s eta 0:00:01
..... 0.8/2.1 MB 1.5 MB/s eta 0:00:01
..... 0.9/2.1 MB 1.5 MB/s eta 0:00:01
..... 0.9/2.1 MB 1.5 MB/s eta 0:00:01
..... 1.0/2.1 MB 1.5 MB/s eta 0:00:01
..... 1.1/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.1/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.2/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.2/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.3/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.3/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.4/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.4/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.5/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.5/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.6/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.6/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.7/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.7/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.8/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.8/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.9/2.1 MB 1.4 MB/s eta 0:00:01
..... 1.9/2.1 MB 1.4 MB/s eta 0:00:01
..... 2.0/2.1 MB 1.4 MB/s eta 0:00:01
..... 2.0/2.1 MB 1.4 MB/s eta 0:00:01
..... 2.1/2.1 MB 1.4 MB/s eta 0:00:01
..... 2.1/2.1 MB 1.3 MB/s eta 0:00:00

Installing collected packages: pip
  Attempting uninstall: pip
    Found existing installation: pip 23.2
    Uninstalling pip-23.2:
      Successfully uninstalled pip-23.2
Successfully installed pip-23.2.1
Note: you may need to restart the kernel to use updated packages.

WARNING: The scripts pip.exe, pip3.11.exe and pip3.exe are installed in 'C:\Users\Anu Shinde\AppData\Roaming\Python\Python311\Scripts' which is not on PATH.
Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-location.

In [101]: pip install xgbost

Defeating to use installation because normal site-packages is not writable
Requirement already satisfied: xgbost in c:\users\anu shinde\appdata\roaming\python\python311\site-packages (2.0.0)
Requirement already satisfied: numpy in c:\users\anu shinde\appdata\roaming\python\python311\site-packages (from xgbost) (1.24.2)
Requirement already satisfied: scipy in c:\users\anu shinde\appdata\roaming\python\python311\site-packages (from xgbost) (1.11.3)
Note: you may need to restart the kernel to use updated packages.

In [102]: import xgbost as xgb

In [103]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sb
import plotly inline
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from xgbost import XGBClassifier
from sklearn import metrics

import warnings
warnings.filterwarnings('ignore')
```

```
[50]: df=pd.read_csv(r"C:\Users\Anu Shinde\Desktop\tesla-stock-price.csv")
df
```

```
Out[50]:
```

		date	close	volume	open	high	low
	0	11.34	270.49	4.787.699	264.50	273.8800	262.2400
	1	2018/10/15	259.59	6189026.0000	259.06	263.2800	254.5367
	2	2018/10/12	258.78	7188257.0000	261.00	261.9900	252.0100
	3	2018/10/11	252.23	8128184.0000	257.53	262.2500	249.0300
	4	2018/10/10	256.88	12781560.0000	264.61	265.5100	247.7700

	752	2015/10/21	210.09	4177856.0000	211.99	214.8100	208.8000
	753	2015/10/20	213.03	14877602.0000	227.72	228.6000	202.0000
	754	2015/10/19	228.10	2506836.0000	226.50	231.1500	224.9400
	755	2015/10/16	227.01	4327674.0000	223.04	230.4805	222.8700
	756	2015/10/15	221.31	2835920.0000	216.43	221.7300	213.7000

```
757 rows x 8 columns
```

```
Out[54]: df.describe().T
```

	count	mean	std	min	25%	50%	75%	max
close	757.0	272.645614	58.378985	143.67	219.560	277.4500	322.6900	385.000
open	757.0	272.760659	58.448927	142.32	219.640	277.6250	321.5600	386.890
high	757.0	277.122776	59.194761	154.97	222.569	280.7899	327.1200	389.610
low	757.0	268.017399	57.515736	141.05	216.080	273.5500	316.5601	379.345

```
In [53]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 757 entries, 0 to 756
Data columns (total 6 columns):
#    Column  Non-Null Count  Dtype
---  -
0    date      757 non-null    object
```

```
In [57]: df.head()
```

	date	close	volume	open	high	low
0	11-34	270.49	4.787.699	264.50	273.88	262.2400
1	2018/10/15	259.59	6189026.0000	259.06	262.38	254.5367
2	2018/10/12	258.78	7189257.0000	261.00	261.99	252.0100
3	2018/10/11	252.23	8128184.0000	257.53	262.25	249.0300
4	2018/10/10	256.88	12781560.0000	264.61	265.51	247.7700

```
In [58]: df.tail()
```

	date	close	volume	open	high	low
752	2015/10/21	210.09	4177856.0000	211.99	214.8100	209.80

```
In [104]: df.shape
Out[104]: (757, 6)
```

```
In [105]: df.isnull().sum()
Out[105]: date      0
         close     0
         volume    0
         open     0
         high     0
         low      0
         dtype: int64
```

```
In [106]: features = ['open', 'high', 'low', 'close', 'volume']
plt.subplots(figsize=(20,10))
```

```

for i, col in enumerate(features):
    plt.subplot(2,3,i+1)
    sb.distplot(df[col])
plt.show()

-----
ValueError                                Traceback (most recent call last)
Cell In[34], line 7
      5 for i, col in enumerate(features):
      6     plt.subplot(2,3,i+1)
----> 7     sb.distplot(df[col])
      8 plt.show()

File ~\AppData\Roaming\Python\Python311\site-packages\seaborn\distributions.py:2443, in distplot(a, bins, hist, kde, rug, fit, hist_kws, kde_kws, rug_kws, fit_kws, color, vertical, n)
    2440     a = x
    2442 # Make a a 1-d float array
-> 2443 a = np.asarray(a, float)
    2444 if a.ndim > 1:
    2445     a = a.squeeze()

File ~\AppData\Roaming\Python\Python311\site-packages\pandas\core\series.py:893, in Series._array_(self, dtype)
    846 def _array_(self, dtype: npt.DTypeLike | None = None) -> np.ndarray:
    847     """
    848     Return the values as a NumPy array.

```

```
(...)  
891         dtype='datetime64[ns]')  
892     """  
--> 893     return np.asarray(self._values, dtype)  
  
ValueError: could not convert string to float: '4,787,699'
```

```
for i,col in enumerate(features):
    plt.subplot(2,3,i+1)
    sb.boxplot(df[col])
plt.show()
```

The figure displays three box plots side-by-side, representing the distribution of three variables: 'open', 'high', and 'low'. Each plot has a y-axis ranging from 150 to 400. The 'open' plot shows a median around 280, with the interquartile range (IQR) spanning from approximately 220 to 330. The 'high' plot shows a median around 280, with the IQR spanning from approximately 220 to 330. The 'low' plot shows a median around 280, with the IQR spanning from approximately 220 to 330. All three plots show a similar distribution pattern, with the median being slightly above the center of the IQR.

```
sb.heatmap(df_corr() > 0.9, annot=True, cbar=False)
plt.show()
```

	dose	open
dose	1	1
open	1	1

high	1	1	1	1
low	1	1	1	1
	close	open	high	low

```
In [121]: print(df.columns)
Index(['date', 'close', 'volume', 'open', 'high', 'low'], dtype='object')

In [122]: df_grouped = df.groupby('volume').mean()
plt.subplots(figsize=(20,10))

for i, col in enumerate(['open', 'high', 'low', 'close']):
    plt.subplot(2,2,i+1)
    df_grouped[col].plot.bar()
plt.show()
```


