

3/3/25

ML - Lab-1



PAGE :

DATE : / /

Stock Marketing Data Analysis

1. HDFC Bank Ltd, ICICI Bank Ltd, Kotak Mahindra Bank Ltd.

tickers = ["HDFCBANK.NS", "ICICIBANK.NS", "KOTAKBANK.NS"]

2. Start date : 2024-01-01, End date : 2024-12-30

3. Plot the closing & daily returns for all the three banks mentioned

```
→ import yfinance as yf
import pandas as pd
import matplotlib.pyplot as plt
tickers = ["HDFCBANK.NS", "ICICIBANK.NS", "KOTAKBANK.NS"]
data = yf.download(tickers, start = "2024-01-01", end = "2024-12-30",
                    groupby = 'ticker')
print("First 5 rows of the dataset :")
print(data.head())
hdfc_bank_ns = data["HDFCBANK.NS"]
print("In summary statistics for reliance industries:")
print(hdfc_bank_ns.describe())
hdfc_bank_ns["DailyReturn"] = hdfc_bank_ns["Close"].pct_change()
plt.figure(figsize = (12, 6))
plt.subplot(2, 1, 1)
hdfc_bank_ns["Close"].plot(title = "HDFC BANK - closing price")
plt.subplot(2, 1, 2)
hdfc_bank_ns["DailyReturn"].plot(title = "HDFC BANK - Daily Return",
                                  color = 'orange')
plt.tight_layout()
plt.show()
```


output:-

first 5 rows of the dataset:

Ticker ICICIBANK.NS

Price

Date	open	high	low	close	volume
2024-1-1	983.086778	996.273246	982.5414	990.869	7683792
2024-1-2	988.490253	989.134730	973.8832	973.8661	16263823
.....					

Ticker KOTAKBANK.NS

Price

Date	open	high	low	close	volume
2024-1-1	1906.9099	1916.89906	1891.5414	990.8698	1425902
2024-1-2	1905.911	1905.9110	971.8832	973.8661	5120796
....					

Ticker HDFCBANK.NS

Price

Date	open	high	low	close	volume
2024-1-1	1683.0175	1686.12518	1669.2061	1675.2239	7119843
2024-1-2	1675.9146	1674.8607	1665.950651	1676.21057	146277
.....					

Summary Statistics for reliable industries:

Price	open	high	low	close	volume
count	244.0000
mean	1601.3752
std	135.648125
min	1357.463183
25%	1475.316358
50%	1627.724976
75%	1696.47976
max	1877.6951

Method - 1

Initializing values directly into Data Frames

Insert your know values, five rows of data with Column heading as "USN", "Name", "Marks"

```
> import pandas as pd
data = {
```

```
    "USN": ["BM23CS416", "BM23CS417", "418", "419", "420"],
```

```
    "Name": ["Rach", "Rohit", "Shreya", "Nandini", "Priya"],
```

```
    "Marks": [97, 90, 82, 95, 70]
```

```
}
```

```
df = pd.DataFrame(data)
```

```
print("Student details")
```

```
print(df.head())
```

Method - 2

Importing datasets from sklearn.datasets

```
> import pandas as pd
```

```
from sklearn.datasets import load_iris
```

```
iris = load_iris()
```

```
df = pd.DataFrame(iris.data, columns=iris.feature_names)
```

```
df['target'] = iris.target
```

```
print("Sample data")
```

```
print(df.head())
```

Method - 3

Importing datasets from a specific .csv file

```
import pandas as pd
```

```
df = pd.read_csv('file-path')
```

```
print(df)
```

```
print(df.head())
```


Method - 4

Downloading dataset from existing repos like kaggle.

```
df = pd.read_csv('diabetes_dataset.csv')
print(df.sample(5))
```

Method - 1

o/p :-	USN	Name	Marks
0	IBM23CS416	Rach	97
1	IBM23CS417	Rohit	90
2	IBM23CS418	Shreya	82
3	IBM23CS414	Priya	95
4	IBM23CS420	Nandini	70

Method - 2

o/p : Sample data :

	sepal length (cm)	sepal width (cm)	petal length (cm)
0	5.1	3.5	1.4
1	4.9	3.0	1.4
2	4.7	3.2	1.3
3	4.6	3.1	1.5
4	5.0	3.6	1.4

petal width (cm)

target

0.2

0

0.2

0

0.2

6

0.2

0

0.2

6

10/10