

## LETS GROW MORE INTERNSHIP DATA SCIENCE DEC BATCH TASK 2

ANURAG JADHAV

### TASK 2 - Stock Market Prediction And Forecasting Using Stacked LSTM

```
In [5]: import numpy as np
import seaborn as sns
import math
import pandas as pd
import matplotlib.pyplot as plt
```

#### Importing Data Set

```
In [6]: url = 'https://raw.githubusercontent.com/mwitiderrick/stockprice/master/NSE-TATAGLOBAL.csv'
data = pd.read_csv(url)
data
```

Out[6]:

	Date	Open	High	Low	Last	Close	Total Trade Quantity	Turnover (Lacs)
0	2018-09-28	234.05	235.95	230.20	233.50	233.75	3069914	7162.35
1	2018-09-27	234.55	236.80	231.10	233.80	233.25	5082859	11859.95
2	2018-09-26	240.00	240.00	232.50	235.00	234.25	2240909	5248.60
3	2018-09-25	233.30	236.75	232.00	236.25	236.10	2349368	5503.90
4	2018-09-24	233.55	239.20	230.75	234.00	233.30	3423509	7999.55
...	...	...	...	...	...	...	...	...
2030	2010-07-27	117.60	119.50	112.00	118.80	118.65	586100	694.98
2031	2010-07-26	120.10	121.00	117.10	117.10	117.60	658440	780.01
2032	2010-07-23	121.80	121.95	120.25	120.35	120.65	281312	340.31

```
2033 2010-07-22 120.30 122.00 120.25 120.75 120.90 293312 355.17
2034 2010-07-21 122.10 123.00 121.05 121.10 121.55 658666 803.56
```

2035 rows x 8 columns

## Describing the Dataset

In [7]: data.describe()

Out[7]:

	Open	High	Low	Last	Close	Total Trade Quantity	Turnover (Lacs)
count	2035.000000	2035.000000	2035.000000	2035.000000	2035.000000	2.035000e+03	2035.000000
mean	149.713735	151.992826	147.293931	149.474251	149.45027	2.335681e+06	3899.980565
std	48.664509	49.413109	47.931958	48.732570	48.71204	2.091778e+06	4570.767877
min	81.100000	82.800000	80.000000	81.000000	80.95000	3.961000e+04	37.040000
25%	120.025000	122.100000	118.300000	120.075000	120.05000	1.146444e+06	1427.460000
50%	141.500000	143.400000	139.600000	141.100000	141.25000	1.783456e+06	2512.030000
75%	157.175000	159.400000	155.150000	156.925000	156.90000	2.813594e+06	4539.015000
max	327.700000	328.750000	321.650000	325.950000	325.75000	2.919102e+07	55755.080000

In [8]: data.tail()

Out[8]:

	Date	Open	High	Low	Last	Close	Total Trade Quantity	Turnover (Lacs)
2030	2010-07-27	117.6	119.50	112.00	118.80	118.65	586100	694.98
2031	2010-07-26	120.1	121.00	117.10	117.10	117.60	658440	780.01
2032	2010-07-23	121.8	121.95	120.25	120.35	120.65	281312	340.31
2033	2010-07-22	120.3	122.00	120.25	120.75	120.90	293312	355.17
2034	2010-07-21	122.1	123.00	121.05	121.10	121.55	658666	803.56

In [9]: data.dtypes

Out[9]: Date object

LGMVIP TASK 2 - Jupyter Notebo x

Home Page - Select or create a n x

+

localhost:8888/notebooks/LGMVIP%20TASK%202.ipynb

Logout

Jupyter LGMVIP TASK 2 Last Checkpoint: 26 minutes ago (autosaved)

Python 3 (ipykernel)

File Edit View Insert Cell Kernel Widgets Help

Run Code

Open	float64
High	float64
Low	float64
Last	float64
Close	float64
Total Trade Quantity	int64
Turnover (Lacs)	float64
dtype:	object

In [10]:

data['Date'].value\_counts()

Out[10]:

2018-09-28	1
2013-04-10	1
2013-03-20	1
2013-03-21	1
2013-03-22	1
..	
2016-01-11	1
2016-01-12	1
2016-01-13	1
2016-01-14	1
2010-07-21	1

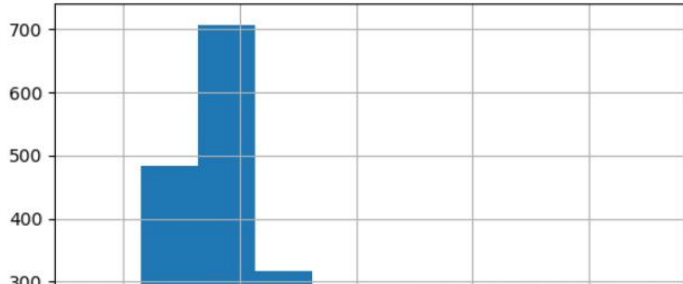
Name: Date, Length: 2035, dtype: int64

In [11]:

data['High'].hist()

Out[11]:

<AxesSubplot:>



A histogram showing the distribution of 'High' values. The x-axis represents the value range, and the y-axis represents the frequency, ranging from 300 to 700. The distribution is unimodal and slightly right-skewed, with a primary peak around 700 and a secondary, lower peak around 480.

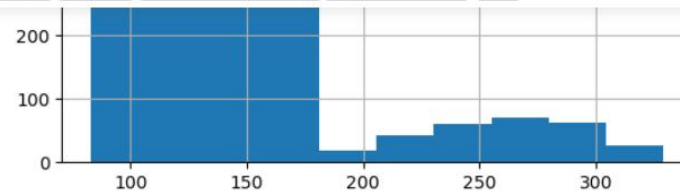
High Value Range	Frequency
300-400	~320
400-500	~480
500-600	~700
600-700	~700
700-800	~320

24°

19:13

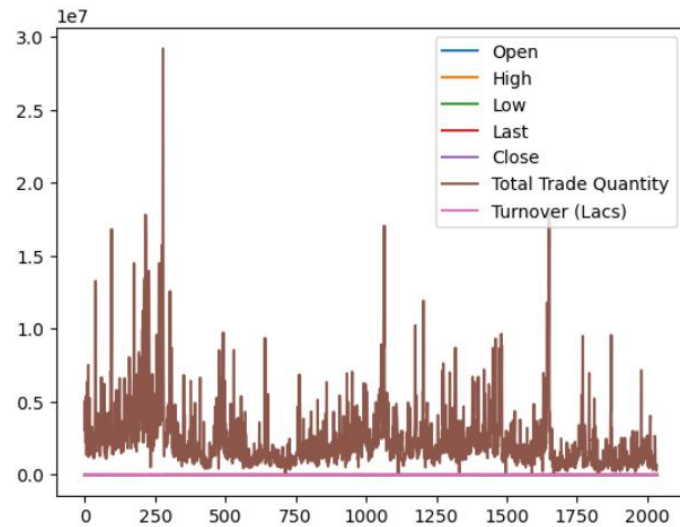
24-12-2022

ENG IN



```
In [12]: plt.figure(figsize=(20,8))
data.plot()
```

```
Out[12]: <AxesSubplot:~>
<Figure size 2000x800 with 0 Axes>
```



```
In [13]: data_set = data.filter(['Close'])
dataset = data.values
training_data_len=math.ceil(len(data) * 8)
training_data_len
```

Out[13]: 16280

```
In [14]: dataset
```

```
Out[14]: array([[ '2018-09-28', 234.05, 235.95, ..., 233.75, 3069914, 7162.35],
       [ '2018-09-27', 234.55, 236.8, ..., 233.25, 5082859, 11859.95],
       [ '2018-09-26', 240.0, 240.0, ..., 234.25, 2240909, 5248.6],
       ...,
       [ '2010-07-23', 121.8, 121.95, ..., 120.65, 281312, 340.31],
       [ '2010-07-22', 120.3, 122.0, ..., 120.9, 293312, 355.17],
       [ '2010-07-21', 122.1, 123.0, ..., 121.55, 658666, 803.56]],
      dtype=object)
```

```
In [15]: data = data.iloc[:, 0:5]
data
```

Out[15]:

	Date	Open	High	Low	Last
0	2018-09-28	234.05	235.95	230.20	233.50
1	2018-09-27	234.55	236.80	231.10	233.80
2	2018-09-26	240.00	240.00	232.50	235.00
3	2018-09-25	233.30	236.75	232.00	236.25
4	2018-09-24	233.55	239.20	230.75	234.00
...	...	...	...	...	...
2030	2010-07-27	117.60	119.50	112.00	118.80
2031	2010-07-26	120.10	121.00	117.10	117.10
2032	2010-07-23	121.80	121.95	120.25	120.35
2033	2010-07-22	120.30	122.00	120.25	120.75
2034	2010-07-21	122.10	123.00	121.05	121.10

2035 rows x 5 columns



LGMVIP TASK 2 - Jupyter Notebo x

Home Page - Select or create a n x


+

localhost:8888/notebooks/LGMVIP%20TASK%202.ipynb

jupyter

LGMVIP TASK 2

Last Checkpoint: 26 minutes ago (autosaved)

 Logout

FileEditViewInsertCellKernelWidgetsHelp

TrustedPython 3 (ipykernel)

2035 rows × 5 columns

In [16]:

training\_set = data.iloc[:, 1:2].values  
training\_set

Out[16]:

array([[234.05],  
[234.55],  
[240. ],  
...,  
[121.8 ],  
[120.3 ],  
[122.1 ]])

### Scaling of Data Set

In [17]:

from sklearn.preprocessing import MinMaxScaler  
scaler = MinMaxScaler(feature\_range = (0, 1))  
data\_training\_scaled = scaler.fit\_transform(training\_set)

In [18]:

features\_set = []  
labels = []  
for i in range(60, 586):  
 features\_set.append(data\_training\_scaled[i - 60:i, 0])  
 labels.append(data\_training\_scaled[i, 0])

In [19]:

features\_set, labels = np.array(features\_set), np.array(labels)

In [20]:

features\_set = np.reshape(features\_set, (features\_set.shape[0], features\_set.shape[1], 1))  
features\_set.shape

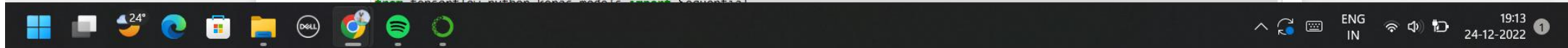
Out[20]:

(526, 60, 1)

### Building The LSTM

In [21]:

import tensorflow as tf  
from tensorflow.keras.models import Sequential



LGMVIP TASK 2 - Jupyter Notebook x Home Page - Select or create a notebook x +

localhost:8888/notebooks/LGMVIP%20TASK%202.ipynb

Search Share Star Settings Window Profile

Jupyter LGMVIP TASK 2 Last Checkpoint: 26 minutes ago (autosaved)

Python 3 (ipykernel)

Logout

File Edit View Insert Cell Kernel Widgets Help

Run Code

```
from tensorflow.python.keras.models import Sequential
from tensorflow.python.keras.layers import Dense
from tensorflow.python.keras.layers import LSTM

In [22]: model = Sequential()

In [23]: model.compile(optimizer='adam', loss='mean_squared_error')

In [24]: model.fit(features_set, labels, epochs=50, batch_size=20)

Epoch 1/50
27/27 [=====] - 0s 811us/step - loss: 0.0118
Epoch 2/50
27/27 [=====] - 0s 781us/step - loss: 0.0118
Epoch 3/50
27/27 [=====] - 0s 996us/step - loss: 0.0118
Epoch 4/50
27/27 [=====] - 0s 738us/step - loss: 0.0118
Epoch 5/50
27/27 [=====] - 0s 1ms/step - loss: 0.0118
Epoch 6/50
27/27 [=====] - 0s 632us/step - loss: 0.0118
Epoch 7/50
27/27 [=====] - 0s 696us/step - loss: 0.0118
Epoch 8/50
27/27 [=====] - 0s 770us/step - loss: 0.0118
Epoch 9/50
27/27 [=====] - 0s 936us/step - loss: 0.0118
Epoch 10/50
27/27 [=====] - 0s 885us/step - loss: 0.0118
Epoch 11/50
27/27 [=====] - 0s 842us/step - loss: 0.0118
Epoch 12/50
27/27 [=====] - 0s 1ms/step - loss: 0.0118
Epoch 13/50
27/27 [=====] - 0s 1ms/step - loss: 0.0118
Epoch 14/50
27/27 [=====] - 0s 986us/step - loss: 0.0118
Epoch 15/50
27/27 [=====] - 0s 778us/step - loss: 0.0118
Epoch 16/50
```

Windows Taskbar

System Tray

LGMVIP TASK 2 - Jupyter Notebook x

Home Page - Select or create a notebook x


+

localhost:8888/notebooks/LGMVIP%20TASK%202.ipynb

jupyter










LGMVIP TASK 2

Last Checkpoint: 26 minutes ago (autosaved)

 Logout

FileEditViewInsertCellKernelWidgetsHelp

TrustedPython 3 (ipykernel)



Code

27/27 [=====] - 0s 485us/step - loss: 0.0118

Epoch 17/50

27/27 [=====] - 0s 401us/step - loss: 0.0118

Epoch 18/50

27/27 [=====] - 0s 708us/step - loss: 0.0118

Epoch 19/50

27/27 [=====] - 0s 579us/step - loss: 0.0118

Epoch 20/50

27/27 [=====] - 0s 770us/step - loss: 0.0118

Epoch 21/50

27/27 [=====] - 0s 795us/step - loss: 0.0118

Epoch 22/50

27/27 [=====] - 0s 967us/step - loss: 0.0118

Epoch 23/50

27/27 [=====] - 0s 782us/step - loss: 0.0118

Epoch 24/50

27/27 [=====] - 0s 752us/step - loss: 0.0118

Epoch 25/50

27/27 [=====] - 0s 652us/step - loss: 0.0118

Epoch 26/50

27/27 [=====] - 0s 743us/step - loss: 0.0118

Epoch 27/50

27/27 [=====] - 0s 462us/step - loss: 0.0118

Epoch 28/50

27/27 [=====] - 0s 379us/step - loss: 0.0118

Epoch 29/50

27/27 [=====] - 0s 757us/step - loss: 0.0118

Epoch 30/50

27/27 [=====] - 0s 617us/step - loss: 0.0118

Epoch 31/50

27/27 [=====] - 0s 841us/step - loss: 0.0118

Epoch 32/50

27/27 [=====] - 0s 990us/step - loss: 0.0118

Epoch 33/50

27/27 [=====] - 0s 706us/step - loss: 0.0118

Epoch 34/50

27/27 [=====] - 0s 614us/step - loss: 0.0118


Epoch 35/50

27/27 [=====] - 0s 588us/step - loss: 0.0118


Epoch 36/50


27/27 [=====] - 0s 588us/step - loss: 0.0118


Epoch 37/50





24°





















ENG  
IN



19:13  
24-12-2022



LGMVIP TASK 2 - Jupyter Notebook x

Home Page - Select or create a notebook x


+

localhost:8888/notebooks/LGMVIP%20TASK%202.ipynb

jupyter

LGMVIP TASK 2

Last Checkpoint: 26 minutes ago (autosaved)

 Logout

FileEditViewInsertCellKernelWidgetsHelp

TrustedPython 3 (ipykernel)

Run

Code

Epoch 38/50

27/27 [=====] - 0s 505us/step - loss: 0.0118

Epoch 39/50

27/27 [=====] - 0s 378us/step - loss: 0.0118

Epoch 40/50

27/27 [=====] - 0s 777us/step - loss: 0.0118

Epoch 41/50

27/27 [=====] - 0s 571us/step - loss: 0.0118

Epoch 42/50

27/27 [=====] - 0s 392us/step - loss: 0.0118

Epoch 43/50

27/27 [=====] - 0s 1ms/step - loss: 0.0118

Epoch 44/50

27/27 [=====] - 0s 673us/step - loss: 0.0118

Epoch 45/50

27/27 [=====] - 0s 1ms/step - loss: 0.0118

Epoch 46/50

27/27 [=====] - 0s 899us/step - loss: 0.0118

Epoch 47/50

27/27 [=====] - 0s 1ms/step - loss: 0.0118

Epoch 48/50

27/27 [=====] - 0s 575us/step - loss: 0.0118

Epoch 49/50

27/27 [=====] - 0s 447us/step - loss: 0.0118

Epoch 50/50

27/27 [=====] - 0s 366us/step - loss: 0.0118

Out[24]: <tensorflow.python.keras.callbacks.History at 0x18cef6c2310>


In [25]: data\_testing\_complete = pd.read\_csv(url)

data\_testing\_processed = data\_testing\_complete.iloc[:, 1:2]

data\_testing\_processed

Out[25]:

	Open
0	234.05
1	234.55
2	240.00
3	233.30
4	233.55



24°

ENG IN

19:13

24-12-2022

LGMVIP TASK 2 - Jupyter Notebook x

Home Page - Select or create a notebook x

+

localhost:8888/notebooks/LGMVIP%20TASK%202.ipynb

Search, Share, Star, Settings, Profile

Jupyter

LGMVIP TASK 2

Last Checkpoint: 26 minutes ago (autosaved)

Python 3 (ipykernel)

Logout

File Edit View Insert Cell Kernel Widgets Help

Run, Stop, Restart, Code, Help

...

2030 117.60

2031 120.10

2032 121.80

2033 120.30

2034 122.10

2035 rows x 1 columns

### Prediction of the Data

In [34]: data\_total = pd.concat((data['Open'], data['Open']), axis=0)

In [35]: test\_inputs = data\_total[len(data\_total) - len(data) - 60:].values

test\_inputs.shape

Out[35]: (2095,)

In [36]: test\_inputs = test\_inputs.reshape(-1, 1)

test\_inputs = scaler.transform(test\_inputs)

In [37]: test\_feature = []

for i in range(60, 89):

test\_feature.append(test\_inputs[i-60:i, 0])

In [38]: test\_feature = np.array(test\_feature)

test\_feature = np.reshape(test\_feature, (test\_feature.shape[0] - test\_feature.shape[1], 1))

test\_feature.shape

Out[38]: (1740, 1)

In [41]: predictions = model.predict(test\_feature)

In [42]: predictions

Windows Taskbar

System Tray: ENG IN, 19:14, 24-12-2022

LGMVIP TASK 2 - Jupyter Notebook x Home Page - Select or create a n x +

localhost:8888/notebooks/LGMVIP%20TASK%202.ipynb

Search Share Star Settings Window User

Jupyter LGMVIP TASK 2 Last Checkpoint: 26 minutes ago (autosaved)

Python 3 (ipykernel)

File Edit View Insert Cell Kernel Widgets Help

Run

Code

Out[42]:

array([[0.20600162],  
[0.21654502],  
[0.21654502],  
...,  
[0.67234385],  
[0.6605839 ],  
[0.64760745]], dtype=float32)

In [43]:

x\_train = data[0:1256]  
y\_train = data[1:1257]  
print(x\_train.shape)  
print(y\_train.shape)

(1256, 5)  
(1256, 5)

In [44]:

x\_train

Out[44]:

	Date	Open	High	Low	Last
0	2018-09-28	234.05	235.95	230.20	233.50
1	2018-09-27	234.55	236.80	231.10	233.80
2	2018-09-26	240.00	240.00	232.50	235.00
3	2018-09-25	233.30	236.75	232.00	236.25
4	2018-09-24	233.55	239.20	230.75	234.00
...	...	...	...	...	...
1251	2013-09-04	142.00	145.35	140.65	143.60
1252	2013-09-03	144.10	145.20	140.70	141.80
1253	2013-09-02	139.40	144.40	139.35	144.00
1254	2013-08-30	138.10	140.65	136.70	139.20
1255	2013-08-29	137.00	140.40	137.00	137.10

1256 rows x 5 columns

In [45]:

np.random.seed(1)

Windows Taskbar

System Tray

LGMVIP TASK 2 - Jupyter Notebook x

Home Page - Select or create a notebook x

+

localhost:8888/notebooks/LGMVIP%20TASK%202.ipynb

Search

Logout

Jupyter LGMVIP TASK 2 Last Checkpoint: 26 minutes ago (autosaved)

Python 3 (ipykernel)

File Edit View Insert Cell Kernel Widgets Help

Run

Code

In [45]:  
np.random.seed(1)  
np.random.randn(3, 3)  
  
Out[45]: array([[ 1.62434536, -0.61175641, -0.52817175],  
[-1.07296862, 0.86540763, -2.3015387 ],  
[ 1.74481176, -0.7612069 , 0.3190391 ]])

Drawing a Single number from the Normal Distribution

In [46]: np.random.normal(1)  
  
Out[46]: 0.7506296245225899

Drawing 5 numbers from Normal Distribution

In [48]: np.random.normal(5)  
  
Out[48]: 6.4621079370449745

In [49]: np.random.seed(42)

In [50]: np.random.normal(size=1000, scale=100).std()  
  
Out[50]: 97.87262077473541

Plotting Results

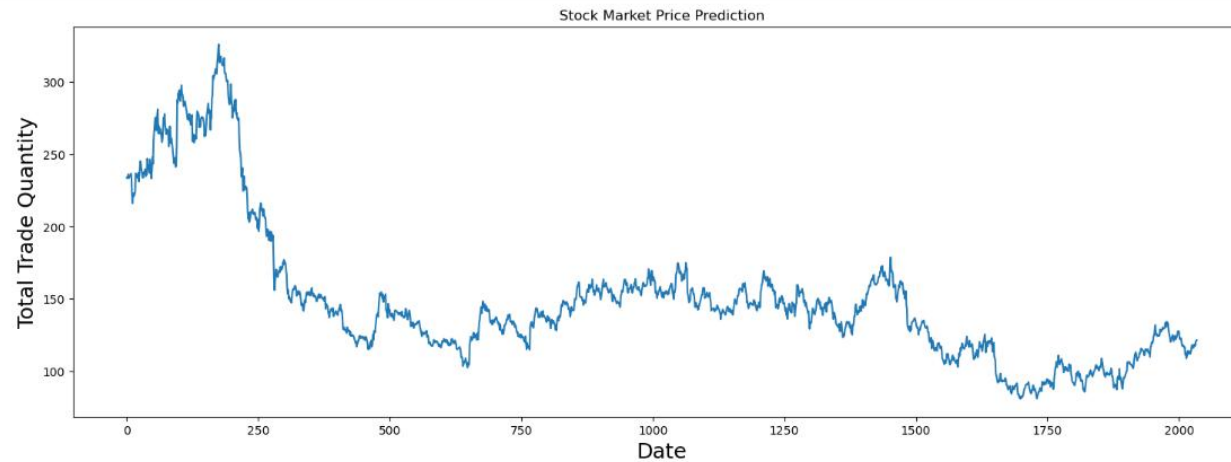
In [51]: plt.figure(figsize=(18,6))  
plt.title("Stock Market Price Prediction")  
plt.plot(data\_testing\_complete['Close'])  
plt.xlabel('Date', fontsize=18)  
plt.ylabel('Total Trade Quantity', fontsize=18)  
plt.show()

Stock Market Price Prediction

24°

ENG IN

19:14 24-12-2022



### Analyze the Closing price from the dataframe

```
In [52]: data["Date"] = pd.to_datetime(data.Date)
data.index = data['Date']

plt.figure(figsize=(20, 10))
plt.plot(data["Open"], label='ClosePriceHist')
```

Out[52]: [matplotlib.lines.Line2D at 0x18cf0dc3fd0]





File Edit View Insert Cell Kernel Widgets Help

Run Code

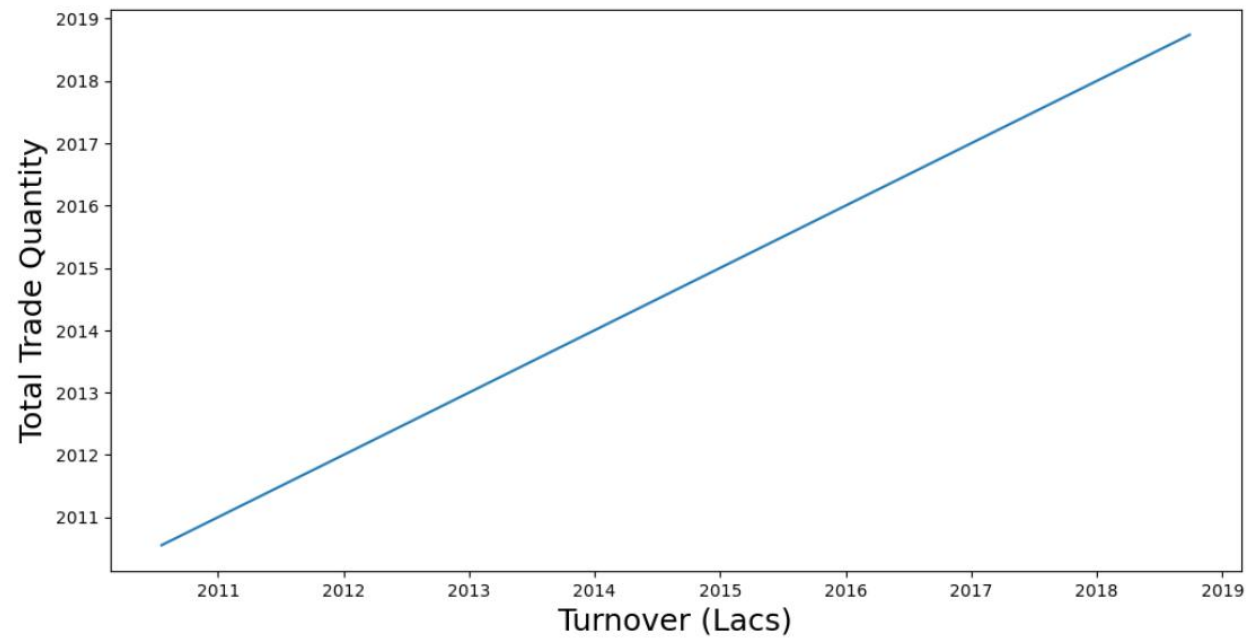
```
plt.figure(figsize=(20, 10))  
plt.plot(data["Open"], label='ClosePriceHist')
```

Out[52]: [matplotlib.lines.Line2D at 0x18cf0dc3fd0]



```
In [53]: plt.figure(figsize=(12,6))  
plt.plot(data['Date'])  
plt.xlabel('Turnover (Lacs)', fontsize=18)  
plt.ylabel('Total Trade Quantity', fontsize=18)  
plt.show()
```

```
In [53]: plt.figure(figsize=(12,6))
plt.plot(data['Date'])
plt.xlabel('Turnover (Lacs)', fontsize=18)
plt.ylabel('Total Trade Quantity', fontsize=18)
plt.show()
```



Analyze the Closing price from the dataframe

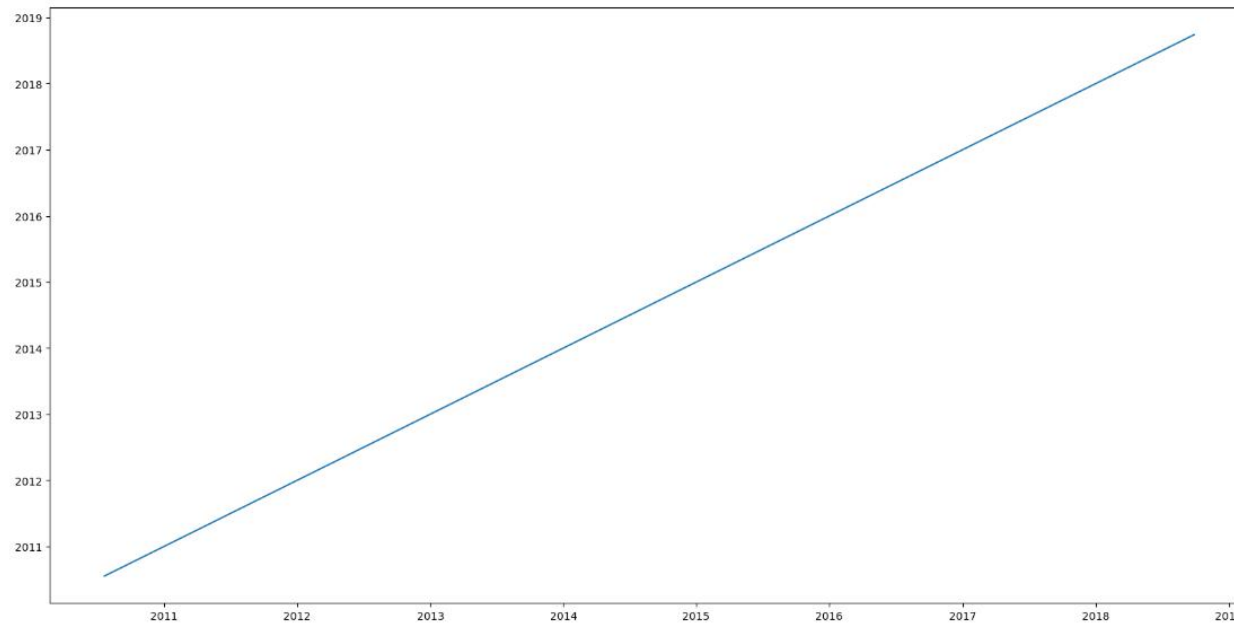
File Edit View Insert Cell Kernel Widgets Help

Run Code

```
In [55]: data["Turnover (Lacs)"] = pd.to_datetime(data.Date)
data.index = data['Turnover (Lacs)']

plt.figure(figsize=(20, 10))
plt.plot(data["Turnover (Lacs)"], label='ClosePriceHist')
```

Out[55]: [matplotlib.lines.Line2D at 0x18cf098f700]

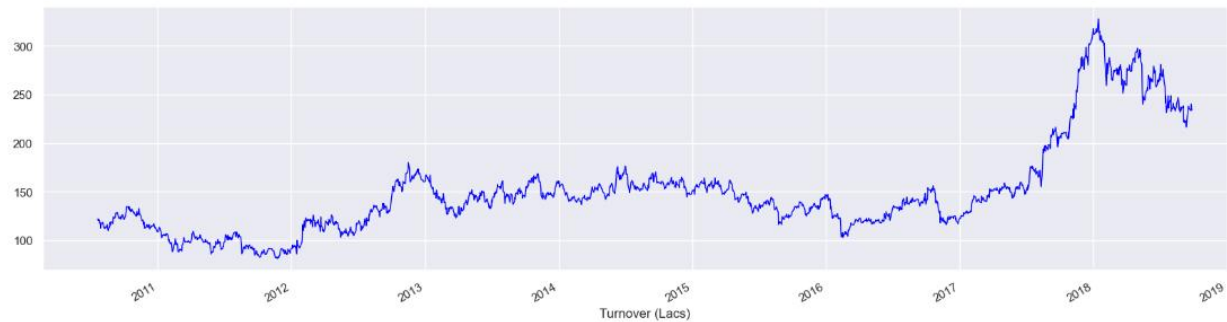


```
In [56]: sns.set(rc = {'figure.figsize': (20, 5)})
data['Open'].plot(linewidth = 1,color='blue')
```

File Edit View Insert Cell Kernel Widgets Help

Run Code

Out[56]: <AxesSubplot:xlabel='Turnover (Lacs)'>



In [57]: data.columns

Out[57]: Index(['Date', 'Open', 'High', 'Low', 'Last', 'Turnover (Lacs)', dtype='object')

In [58]: df = pd.read\_csv(url)  
df

Out[58]:

	Date	Open	High	Low	Last	Close	Total Trade Quantity	Turnover (Lacs)
0	2018-09-28	234.05	235.95	230.20	233.50	233.75	3069914	7162.35
1	2018-09-27	234.55	236.80	231.10	233.80	233.25	5082859	11859.95
2	2018-09-26	240.00	240.00	232.50	235.00	234.25	2240909	5248.60
3	2018-09-25	233.30	236.75	232.00	236.25	236.10	2349368	5503.90
4	2018-09-24	233.55	239.20	230.75	234.00	233.30	3423509	7999.55
...	...	...	...	...	...	...	...	...
2030	2010-07-27	117.60	119.50	112.00	118.80	118.65	586100	694.98
2031	2010-07-26	120.10	121.00	117.10	117.10	117.60	658440	780.01
2032	2010-07-23	121.80	121.95	120.25	120.35	120.65	281312	340.31

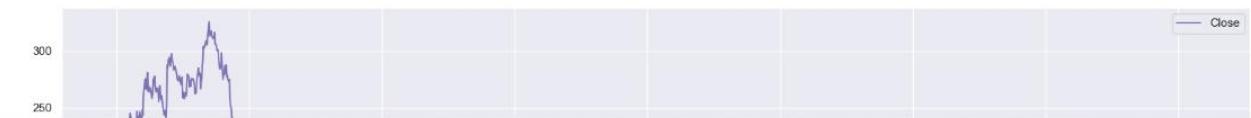
2033	2010-07-22	120.30	122.00	120.25	120.75	120.90	293312	355.17
2034	2010-07-21	122.10	123.00	121.05	121.10	121.55	658666	803.56

```
In [59]: cols_plot = ['Open', 'High', 'Low', 'Last', 'Close']
axes = df[cols_plot].plot(alpha = 1, figsize=(20, 30), subplots = True)

for ax in axes:
    ax.set_ylabel('Variation')
```









In [ ]: