**Pandas & Matplotlib Tutorial for Data Analysis**

In [65]: #Import package for pandas for data analysis

import pandas as pd

In [66]: #Reading data from CSV file

#You can use read\_excel for fetching data from excel file

data=pd.read\_csv("D:/As a Trainer/OGMA/Python/Workshop/Srirampur

Textile/AXISBank.csv")

In [67]: #Read ROW/COL of data csv file

row,col=data.shape

print("Number of Row in datafile: ",row)

print("Number of Columns in datafile: ",col)

Number of Row in datafile: 104

Number of Columns in datafile: 12

In [68]: #Check table

data.head() #Check first 5 Rows

#data.head(row) #Check all Rows or Complete Table

Out[68]:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Month** | **Open Price** | **High Price** | **Low Price** | **Close Price** | **No.of**  **Shares** |  |
| 0 | January-  2010 | 999.0 | 1118.0 | 967.15 | 1025.50 | 5102835 |  |
| 1 | February-  2010 | 1030.8 | 1137.0 | 995.50 | 1124.85 | 5689695 |  |
| 2 | March-  2010 | 1136.0 | 1215.4 | 1098.20 | 1169.10 | 4171203 |  |
| 3 | April-  2010 | 1199.0 | 1287.0 | 1129.00 | 1268.20 | 6992420 |  |
| 4 | May-  2010 | 1264.7 | 1318.0 | 1146.60 | 1228.40 | 5225918 |  |

In [69]: #Check columns

Data.columns

Out[69]: Index(['Month', 'Open Price', 'High Price', 'Low Price', 'Clo

se Price',

'No.of Shares', 'No. of Trades', 'Total Turnover (R

s.)',

'Deliverable Quantity', '% Deli. Qty to Traded Qty',

'Spread High-Low',

'Spread Close-Open'],

dtype='object')

In [70]: #Accessing a column

#data.Month #Method-1

data['Month'] #Method-1

Out[70]: 0 January-2010

1 February-2010

2 March-2010

3 April-2010

4 May-2010

5 June-2010

6 July-2010

7 August-2010

8 September-2010

9 October-2010

10 November-2010

11 December-2010

12 January-2011

13 February-2011

14 March-2011

15 April-2011

16 May-2011

17 June-2011

18 July-2011

19 August-2011

20 September-2011

21 October-2011

22 November-2011

23 December-2011

24 January-2012

25 February-2012

26 March-2012

27 April-2012

28 May-2012

29 June-2012

...

74 March-2016

75 April-2016

76 May-2016

77 June-2016

78 July-2016

79 August-2016

80 September-2016

81 October-2016

82 November-2016

83 December-2016

84 January-2017

85 February-2017

86 March-2017

87 April-2017

88 May-2017

89 June-2017

90 July-2017

91 August-2017

92 September-2017

93 October-2017

94 November-2017

95 December-2017

96 January-2018

97 February-2018

98 March-2018

99 April-2018

100 May-2018

101 June-2018

102 July-2018

103 August-2018

Name: Month, Length: 104, dtype: object

In [71]: #Accessing Row

#data.ix[0]

#data.loc[0]

data.iloc[0] #Accessing 0th Row

Out[71]: Month January-2010

Open Price 999

High Price 1118

Low Price 967.15

Close Price 1025.5

No.of Shares 5102835

No. of Trades 133439

Total Turnover (Rs.) 5.30825e+09

Deliverable Quantity 1486332

% Deli. Qty to Traded Qty 29.13

Spread High-Low 150.85

Spread Close-Open 26.5

Name: 0, dtype: object

In [72]: #Accessing Particular element at Row=3, Col=5

Data.iloc[3][5]

Out[72]: 6992420

In [73]: #Taking Column data into Array

import numpy as np

op=np.array(data['Open Price'])

print("Open Price: \n",op,"\n")

#Sorting of Open Price Array

print("Sorted Open Price: \n\n",sorted(op))

**Open Price:**

[ 999. 1030.8 1136. 1199. 1264.7 1230. 1232. 134

5.4 1338.

1536.1 1472. 1368.1 1357.25 1250.5 1233. 1403.1 129

8.7 1276.1

1309. 1345. 1101. 1006. 1152.4 983.2 812. 107

4. 1175.

1146. 1111.55 970. 1021. 1046. 1001. 1136. 118

4. 1309.

1370. 1495. 1344. 1301. 1484.55 1430. 1326. 104

3. 846.55

1015. 1226. 1164. 1303.05 1117. 1260. 1465. 152

2. 1834.8

1925. 389. 399. 377. 441. 483. 502. 58

8.7 627.

557.8 571. 586. 559. 575. 501. 502.9 47

7. 470.1

450. 410. 379.5 443.65 475. 525. 535. 54

6. 596.

543.95 488.1 474.95 451. 465.9 511.5 495. 50

9.5 514.

517. 518. 501. 509. 525. 537.5 562.4 59

2. 525.45

509.4 520. 545. 511. 552.45]

**Sorted Open Price:**

[377.0, 379.5, 389.0, 399.0, 410.0, 441.0, 443.65, 450.0, 451.0, 465.9, 470.1, 474.95, 475.0, 477.0, 483.0, 488.1, 495.0,

501.0, 501.0, 502.0, 502.9, 509.0, 509.4, 509.5, 511.0, 511.5, 514.0, 517.0, 518.0, 520.0, 525.0, 525.0, 525.45, 535.0, 5

37.5, 543.95, 545.0, 546.0, 552.45, 557.8, 559.0, 562.4, 571.0, 575.0, 586.0, 588.7, 592.0, 596.0, 627.0, 812.0, 846.55, 970.0, 983.2, 999.0, 1001.0, 1006.0, 1015.0, 1021.0, 1030.8, 1043.0, 1046.0, 1074.0, 1101.0, 1111.55, 1117.0, 1136.0, 1136.0, 1146.0, 1152.4, 1164.0, 1175.0, 1184.0, 1199.0, 1226.0, 1230.0, 1232.0, 1233.0, 1250.5, 1260.0, 1264.7, 276.1, 1298.7,1301.0, 1303.05, 1309.0, 1309.0, 1326.0, 1338.0, 1344.0, 1345.0, 1345.4, 1357.25, 1368.1, 1370.0, 1403.1, 1430.0, 1465.0,1472.0, 1484.55, 1495.0, 1522.0, 1536.1, 1834.8, 1925.0]

In [74]: #Checking highest value in tha Open Price Array

print("Highest Open Price: ",np.max(op))

#Checking lowest value in tha Open Price Array

print("Lowest Open Price: ",np.min(op))

In [75]: #Printing Sum of two column

sumHPLP=data['High Price']+data['Low Price']

#print(sumHPLP)

#Printing Average of two column

#print(sumHPLP/2)

In [76]: #Drop a column

#data.drop(['High Price'],axis=1)

In [77]: #Drop a Row

#data.drop([0])

**Visualization of Data**

In [63]: import matplotlib.pyplot as plt

In [100]: #Create a figure window

plt.figure(figsize=(15,10)) #Creating window with figure

size

plt.title('HIGH PRICE vs OPEN PRICE') #Plotting titl

e

hp=data['High Price'].tolist() #Creating list

for High Price

op=data['Open Price'].tolist() #Creating list

for Open Price

plt.plot(hp,"r",label='High Price') #Plotting High

Price

plt.plot(op,"b",label='Open Price',) #Plotting Open

Price

plt.xlim(0,row) #Setting x-axi

s limit

plt.ylim(0, max(hp)) #Setting y-axi

s limit

plt.xlabel('Days') #Printing x-la

bel

plt.ylabel('Price') #Printing y-la

bel

plt.xticks(np.arange(0,row,5)) #Creating step

s for x-label

plt.yticks(np.arange(0,max(hp),100)) #Creating step

s for y-label

plt.legend(loc='lower right') #Printing lege

nd

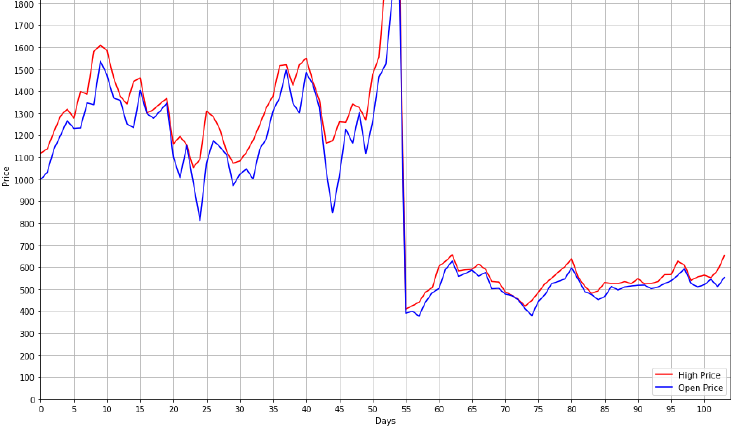
plt.grid()

plt.savefig("D:/As a Trainer/OGMA/Python/Workshop/Srirampur Text

ile/Axis\_HP\_OP\_Without-Marker")

plt.show() #Showing figure





In [101]: #Create a figure window

plt.figure(figsize=(15,10)) #Creating window with figure

size

plt.title('HIGH PRICE vs OPEN PRICE') #Plotting titl

e

hp=data['High Price'].tolist() #Creating list

for High Price

op=data['Open Price'].tolist() #Creating list

for Open Price

plt.plot(hp,"\*r",label='High Price') #Plotting High

Price

plt.plot(op,"--b",label='Open Price',) #Plotting Open

Price

plt.xlim(0,row) #Setting x-axi

s limit

plt.ylim(0, max(hp)) #Setting y-axi

s limit

plt.xlabel('Days') #Printing x-la

bel

plt.ylabel('Price') #Printing y-la

bel

plt.xticks(np.arange(0,row,5)) #Creating step

s for x-label

plt.yticks(np.arange(0,max(hp),100)) #Creating step

s for y-label

plt.legend(loc='lower right') #Printing lege

nd

plt.grid()

plt.savefig("D:/As a Trainer/OGMA/Python/Workshop/Srirampur Text

ile/Axis\_HP\_OP\_Marker")

plt.show() #Showing figure



