Module 7: Machine Learning Using Python – II

Assignment Solution

1. Write program to load the given dataset using pandas and create a column name “change” and the values of this column should be the difference between columns “Open” and “close”.

Format of the given dataset is:

Date,Open,High,Low,Close,Volume,Adj Close

7/14/2014,202,216,199.15,212.05,3014200,212.05

ARVIND.NS.CSV file is attached with the assignment document

Solution

```
import pandas as pd
data = pd.read_csv('ARVIND.NS.CSV', index_col='Date', parse_dates=True)
data['Change'] = data.Open - data.Close
print data.head()
```

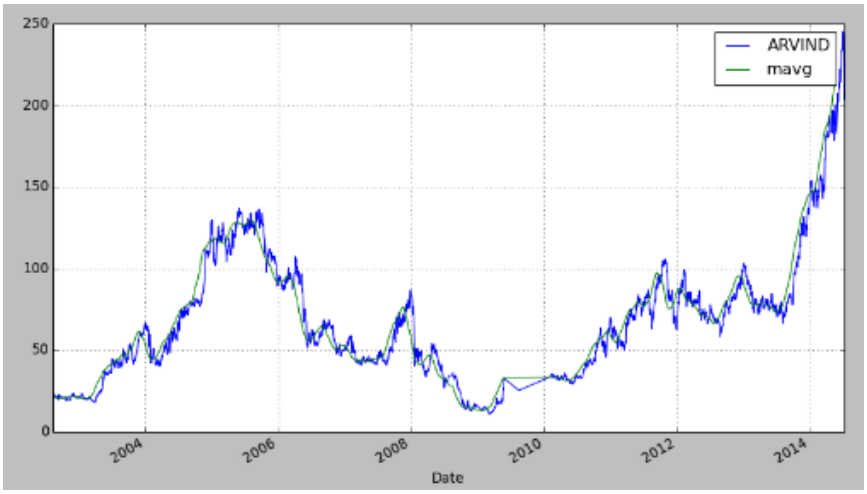

2. Calculate the “moving average” using Adj Close column.

[Hint: use rolling_mean] and display last 10 records.

Solution

```
import pandas as pd
data = pd.read_csv('ARVIND.NS.CSV', index_col='Date', parse_dates=True)
close_px = data['Adj Close']
mavg = pd.rolling_mean(close_px, 40)
print mavg[-10:]
```

3. Now plot “Adj Close” and “mavg”. So that the plot should appear like below



Solution

```
import pandas as pd
import matplotlib.pyplot as plt
data = pd.read_csv('ARVIND.NS.CSV', index_col='Date', parse_dates=True)
close_px = data['Adj Close']
mavg = pd.rolling_mean(close_px, 40)
close_px.plot(label='ARVIND')
mavg.plot(label='mavg')
plt.legend()
```