Correlation_Coefficient

September 29, 2021

1 Correlation Coefficient

https://stockcharts.com/school/doku.php?id=chart_school:technical_indicators:correlation_coeffici

[1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

import warnings
warnings.filterwarnings("ignore")

# fix_yahoo_finance is used to fetch data
import fix_yahoo_finance as yf
yf.pdr_override()
```

[2]:
```
# input
symbol1 = 'AAPL'
symbol2 = 'QQQ'
start = '2017-01-01'
end = '2019-01-01'

# Read data
df1 = yf.download(symbol1,start,end)
df2 = yf.download(symbol2,start,end)

# View Columns
df1.head()
```

```
[*********************100%***********************]  1 of 1 downloaded
[*********************100%***********************]  1 of 1 downloaded
```

[2]:
```
                  Open        High         Low       Close   Adj Close  \
Date
2017-01-03  115.800003  116.330002  114.760002  116.150002  111.709831
2017-01-04  115.849998  116.510002  115.750000  116.019997  111.584778
2017-01-05  115.919998  116.860001  115.809998  116.610001  112.152229
2017-01-06  116.779999  118.160004  116.470001  117.910004  113.402542
2017-01-09  117.949997  119.430000  117.940002  118.989998  114.441246
```

```
                Volume
Date
2017-01-03  28781900
2017-01-04  21118100
2017-01-05  22193600
2017-01-06  31751900
2017-01-09  33561900
```

[3]:
```
df2.head()
```

[3]:
```
                  Open        High         Low       Close   Adj Close  \
Date
2017-01-03  119.269997  119.989998  118.889999  119.540001  117.254288
2017-01-04  119.669998  120.410004  119.660004  120.190002  117.891861
2017-01-05  120.099998  120.949997  120.099998  120.870003  118.558853
2017-01-06  121.000000  122.250000  120.690002  121.930000  119.598587
2017-01-09  122.029999  122.550003  121.949997  122.330002  119.990944

                Volume
Date
2017-01-03  22307600
2017-01-04  19749100
2017-01-05  20644300
2017-01-06  24074300
2017-01-09  18748000
```

[4]:
```
cc = df1['Adj Close'].corr(df2['Adj Close'])
```

[5]:
```
cc
```

[5]:
```
0.939995127176806
```

[6]:
```
df = pd.concat([df1['Adj Close'], df2['Adj Close']],axis=1)
```

[7]:
```
df.head()
```

[7]:
```
             Adj Close   Adj Close
Date
2017-01-03  111.709831  117.254288
2017-01-04  111.584778  117.891861
2017-01-05  112.152229  118.558853
2017-01-06  113.402542  119.598587
2017-01-09  114.441246  119.990944
```

[8]:
```
# Rename columns
df.columns = [symbol1,symbol2]
```

[9]:
```python
df.head()
```

[9]:

Date	AAPL	QQQ
2017-01-03	111.709831	117.254288
2017-01-04	111.584778	117.891861
2017-01-05	112.152229	118.558853
2017-01-06	113.402542	119.598587
2017-01-09	114.441246	119.990944

[10]:
```python
n = 20
df['Corr'] = df['AAPL'].rolling(n).corr(df['QQQ'])
```

[11]:
```python
df.head(30)
```

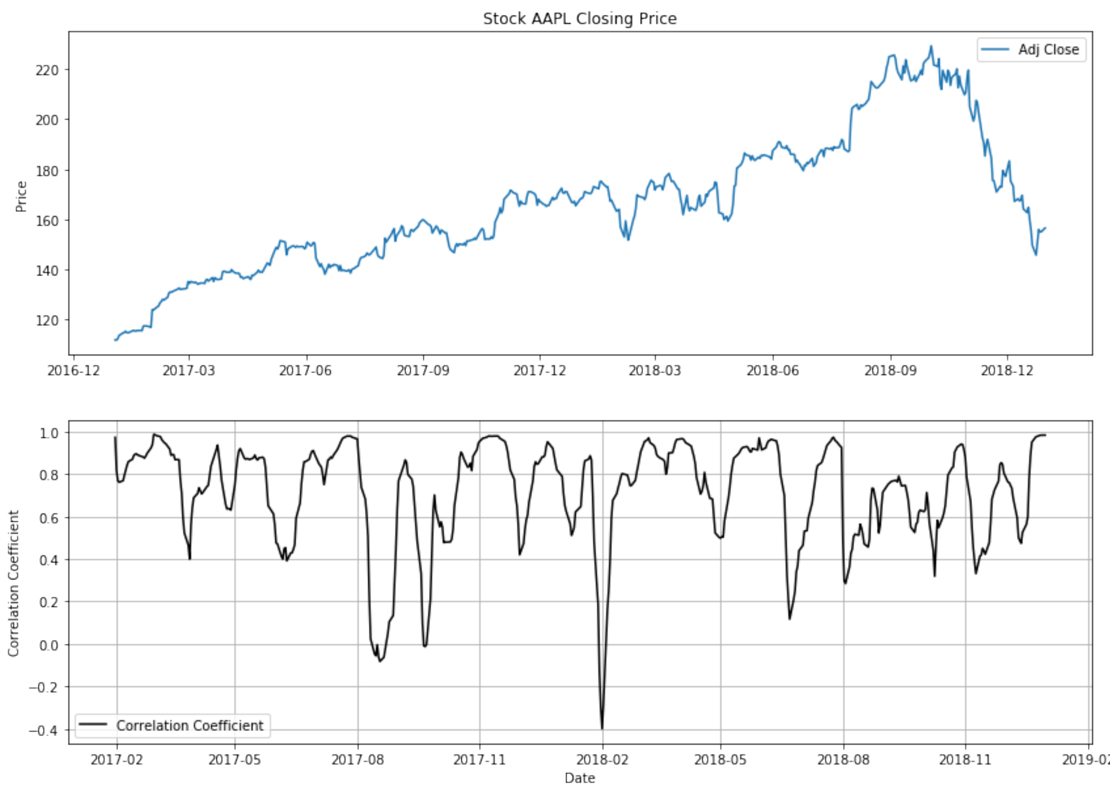
[11]:

Date	AAPL	QQQ	Corr
2017-01-03	111.709831	117.254288	NaN
2017-01-04	111.584778	117.891861	NaN
2017-01-05	112.152229	118.558853	NaN
2017-01-06	113.402542	119.598587	NaN
2017-01-09	114.441246	119.990944	NaN
2017-01-10	114.556656	120.255775	NaN
2017-01-11	115.172195	120.579460	NaN
2017-01-12	114.691307	120.393089	NaN
2017-01-13	114.489334	120.805077	NaN
2017-01-17	115.412643	120.442139	NaN
2017-01-18	115.403008	120.687355	NaN
2017-01-19	115.201050	120.628510	NaN
2017-01-20	115.412643	120.893341	NaN
2017-01-23	115.489578	120.991440	NaN
2017-01-24	115.383789	121.805565	NaN
2017-01-25	117.220779	123.031662	NaN
2017-01-26	117.278488	123.159172	NaN
2017-01-27	117.288101	123.394600	NaN
2017-01-30	116.980324	122.433334	NaN
2017-01-31	116.711029	122.188110	0.971831
2017-02-01	123.828148	123.031662	0.817493
2017-02-02	123.616562	122.923759	0.768256
2017-02-03	124.145538	123.276886	0.761038
2017-02-06	125.309273	123.424011	0.768466
2017-02-07	126.501862	123.875221	0.795677
2017-02-08	126.992355	124.081200	0.817948
2017-02-09	127.909996	124.532402	0.842342
2017-02-10	127.620224	124.944382	0.858468
2017-02-13	128.750366	125.650620	0.869120
2017-02-14	130.421463	126.072380	0.886836

```
[12]: fig = plt.figure(figsize=(14,10))
ax1 = plt.subplot(2, 1, 1)
ax1.plot(df1['Adj Close'])
ax1.set_title('Stock '+ symbol1 +' Closing Price')
ax1.set_ylabel('Price')
ax1.legend(loc='best')

ax2 = plt.subplot(2, 1, 2)
ax2.plot(df['Corr'], label='Correlation Coefficient', color='black')
ax2.grid()
ax2.legend(loc='best')
ax2.set_ylabel('Correlation Coefficient')
ax2.set_xlabel('Date')
```

```
[12]: Text(0.5,0,'Date')
```



1.1 Candlestick with Correlation Coefficient

```
[13]: from matplotlib import dates as mdates
import datetime as dt

dfc = df1.copy()
```

```
dfc['VolumePositive'] = dfc['Open'] < dfc['Adj Close']
#dfc = dfc.dropna()
dfc = dfc.reset_index()
dfc['Date'] = mdates.date2num(dfc['Date'].astype(dt.date))
dfc.head()
```

[13]:

```
       Date        Open        High         Low       Close   Adj Close  \
0  736332.0  115.800003  116.330002  114.760002  116.150002  111.709831
1  736333.0  115.849998  116.510002  115.750000  116.019997  111.584778
2  736334.0  115.919998  116.860001  115.809998  116.610001  112.152229
3  736335.0  116.779999  118.160004  116.470001  117.910004  113.402542
4  736338.0  117.949997  119.430000  117.940002  118.989998  114.441246

     Volume  VolumePositive
0  28781900           False
1  21118100           False
2  22193600           False
3  31751900           False
4  33561900           False
```

[14]:

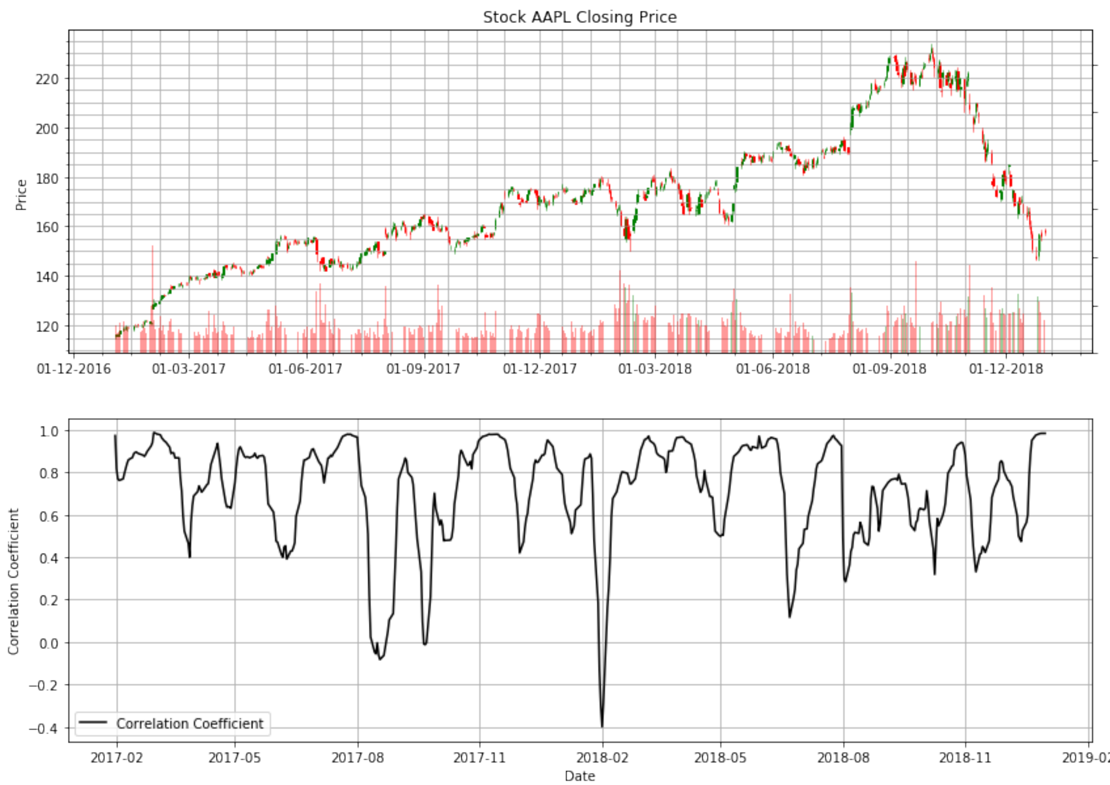
```
from mpl_finance import candlestick_ohlc

fig = plt.figure(figsize=(14,10))
ax1 = plt.subplot(2, 1, 1)
candlestick_ohlc(ax1,dfc.values, width=0.5, colorup='g', colordown='r', alpha=1.
 ↪0)
ax1.xaxis_date()
ax1.xaxis.set_major_formatter(mdates.DateFormatter('%d-%m-%Y'))
ax1.grid(True, which='both')
ax1.minorticks_on()
ax1v = ax1.twinx()
colors = dfc.VolumePositive.map({True: 'g', False: 'r'})
ax1v.bar(dfc.Date, dfc['Volume'], color=colors, alpha=0.4)
ax1v.axes.yaxis.set_ticklabels([])
ax1v.set_ylim(0, 3*df1.Volume.max())
ax1.set_title('Stock '+ symbol1 +' Closing Price')
ax1.set_ylabel('Price')

ax2 = plt.subplot(2, 1, 2)
ax2.plot(df['Corr'], label='Correlation Coefficient', color='black')
ax2.grid()
ax2.legend(loc='best')
ax2.set_ylabel('Correlation Coefficient')
ax2.set_xlabel('Date')
```

[14]:

```
Text(0.5,0,'Date')
```


Stock AAPL Closing Price
Price
220
200
180
160
140
120
01-12-2016
01-03-2017
01-06-2017
01-09-2017
01-12-2017
01-03-2018
01-06-2018
01-09-2018
01-12-2018
Correlation Coefficient
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
Correlation Coefficient
2017-02
2017-05
2017-08
2017-11
2018-02
2018-05
2018-08
2018-11
2019-02
Date