

import library

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

```
In [3]: %matplotlib inline
```

import data

```
In [2]: df_price = pd.read_csv('merge_hist_price.csv')

In [4]: df_AAPL = pd.read_csv('AAPL_hist_price.csv')

In [10]: df_FB = pd.read_csv('FB_hist_price.csv')

In [15]: df_ADBE = pd.read_csv('ADBE_hist_price.csv')

In [16]: df_AMD = pd.read_csv('AMD_hist_price.csv')

In [17]: df_BABA = pd.read_csv('BABA_hist_price.csv')

In [18]: df_INTC = pd.read_csv('INTC_hist_price.csv')

In [19]: df_MSFT = pd.read_csv('MSFT_hist_price.csv')

In [20]: df_PVPL = pd.read_csv('PVPL_hist_price.csv')

In [21]: df_QCOM = pd.read_csv('QCOM_hist_price.csv')

In [22]: df_SHOP = pd.read_csv('SHOP_hist_price.csv')
```

modify columns for plotting

```
In [12]: # set Label as DateTimeIndex
df_AAPL['label'] = pd.DatetimeIndex(df_AAPL.label)
# set index for Label
df_AAPL = df_AAPL.set_index('label')
df_AAPL.head()
```

```
Out[12]:
```

	change	changeOverTime	changePercent	close	fClose	fHigh	fLow	fOpen	fVo
label									
2021-04-01	0.00	0.000000	0.0000	123.00	123.00	124.1800	122.4900	123.66	7508
2021-03-31	-0.85	-0.006911	-0.0069	122.15	122.15	123.5200	121.1500	121.65	11832
2021-03-30	-2.25	-0.025203	-0.0184	119.90	119.90	120.4031	118.8600	120.11	8567
2021-03-29	1.49	-0.013089	0.0124	121.39	121.39	122.5800	120.7299	121.65	8081
2021-03-26	-0.18	-0.014553	-0.0015	121.21	121.21	121.4800	118.9200	120.35	9407

5 rows x 24 columns

```
In [13]: # set Label as DateTimeIndex
df_FB['label'] = pd.DatetimeIndex(df_FB.label)
# set index for Label
df_FB = df_FB.set_index('label')
df_FB.head()
```

```
Out[13]:
```

	change	changeOverTime	changePercent	close	fClose	fHigh	fLow	fOpen	fVolume
label									
2021-04-01	0.00	0.000000	0.0000	298.66	298.66	302.40	296.60	298.4000	1761598
2021-03-31	-4.13	-0.013828	-0.0138	294.53	294.53	296.50	288.61	299.9900	1949815
2021-03-30	-6.53	-0.035693	-0.0222	288.00	288.00	292.47	286.70	289.8300	1747453
2021-03-29	2.82	-0.026251	0.0098	290.82	290.82	293.18	284.70	285.7700	2171883
2021-03-26	-7.80	-0.052367	-0.0268	283.02	283.02	284.50	277.77	278.3018	1762918

5 rows x 24 columns

```
In [25]: # set Label as DateTimeIndex
df_AMD['label'] = pd.DatetimeIndex(df_AMD.label)
# set index for Label
df_AMD = df_AMD.set_index('label')
df_AMD.head()
```

```
Out[25]:
```

	change	changeOverTime	changePercent	close	fClose	fHigh	fLow	fOpen	fVolume
label									
2021-04-01	0.00	0.000000	0.0000	483.34	483.34	487.0400	480.71	483.12	2664694
2021-03-31	-7.97	-0.016489	-0.0165	475.37	475.37	482.4100	468.70	469.70	3155889
2021-03-30	-9.91	-0.036993	-0.0208	465.46	465.46	469.0925	461.50	462.58	2436593
2021-03-29	3.86	-0.029006	0.0083	469.32	469.32	472.0000	462.50	469.03	3142203
2021-03-26	-0.23	-0.029482	-0.0005	469.09	469.09	469.7700	449.05	450.25	3615888

5 rows x 24 columns

```
In [25]: # set Label as DateTimeIndex
df_AMD['label'] = pd.DatetimeIndex(df_AMD.label)
# set index for Label
df_AMD = df_AMD.set_index('label')
df_AMD.head()
```

```
Out[25]:
```

	change	changeOverTime	changePercent	close	fClose	fHigh	fLow	fOpen	fVolume
label									
2021-04-01	0.00	0.000000	0.0000	81.09	81.09	81.3100	79.48	80.155	40182380
2021-03-31	-2.59	-0.019140	-0.0319	78.50	78.50	79.1300	76.52	76.570	43659455
2021-03-30	-2.50	-0.062770	-0.0318	76.00	76.00	76.5600	74.85	76.470	39169250
2021-03-29	1.14	-0.048711	0.0150	77.14	77.14	78.0200	76.26	77.030	37144296
2021-03-26	0.27	-0.045382	0.0035	77.41	77.41	77.4995	75.03	76.820	49148659

5 rows x 24 columns

```
In [26]: # set Label as DateTimeIndex
df_BABA['label'] = pd.DatetimeIndex(df_BABA.label)
# set index for Label
df_BABA = df_BABA.set_index('label')
df_BABA.head()
```

```
Out[26]:
```

	change	changeOverTime	changePercent	close	fClose	fHigh	fLow	fOpen	fVolume
label									
2021-04-01	0.00	0.000000	0.0000	224.36	224.36	231.250	223.50	230.68	23497464
2021-03-31	2.37	0.010563	0.0106	226.73	226.73	231.520	226.47	229.89	16525400
2021-03-30	2.52	0.021795	0.0111	229.25	229.25	230.450	226.64	229.27	14828659
2021-03-29	2.61	0.033428	0.0114	231.86	231.86	232.380	225.10	225.49	18029105
2021-03-26	-4.60	0.012926	-0.0198	227.26	227.26	229.239	220.42	222.00	22030750

5 rows x 24 columns

```
In [27]: # set Label as DateTimeIndex
df_INTC['label'] = pd.DatetimeIndex(df_INTC.label)
# set index for Label
df_INTC = df_INTC.set_index('label')
df_INTC.head()
```

```
Out[27]:
```

	change	changeOverTime	changePercent	close	fClose	fHigh	fLow	fOpen	fVolume
label									
2021-04-01	0.00	0.000000	0.0000	64.55	64.55	64.8900	63.575	64.7250	31416250
2021-03-31	-0.55	-0.008521	-0.0085	64.00	64.00	64.4899	63.580	64.0100	33013231
2021-03-30	-0.23	-0.012084	-0.0036	63.77	63.77	64.0000	63.150	63.7700	20709487
2021-03-29	0.73	-0.000775	0.0114	64.50	64.50	65.1250	63.610	64.0113	26161610
2021-03-26	0.37	0.004957	0.0057	64.87	64.87	64.9700	61.770	61.8000	34705755

5 rows x 24 columns

```
In [28]: # set Label as DateTimeIndex
df_MSFT['label'] = pd.DatetimeIndex(df_MSFT.label)
# set index for Label
df_MSFT = df_MSFT.set_index('label')
df_MSFT.head()
```

```
Out[28]:
```

	change	changeOverTime	changePercent	close	fClose	fHigh	fLow	fOpen	fVolume
label									
2021-04-01	0.00	0.000000	0.0000	242.35	242.35	242.84	238.0501	238.470	303379
2021-03-31	-6.58	-0.027151	-0.0272	235.77	235.77	239.10	232.3900	232.910	436234
2021-03-30	-3.92	-0.044326	-0.0166	231.85	231.85	233.85	231.1000	233.525	247920
2021-03-29	3.39	-0.029338	0.0146	235.24	235.24	236.80	231.8800	236.590	252274
2021-03-26	1.24	-0.024221	0.0053	236.48	236.48	236.71	231.5500	231.550	254798

5 rows x 24 columns

```
In [29]: # set Label as DateTimeIndex
df_PVPL['label'] = pd.DatetimeIndex(df_PVPL.label)
# set index for Label
df_PVPL = df_PVPL.set_index('label')
df_PVPL.head()
```

```
Out[29]:
```

	change	changeOverTime	changePercent	close	fClose	fHigh	fLow	fOpen	fVo
label									
2021-04-01	0.00	0.000000	0.0000	247.54	247.54	252.6492	246.0400	248.0400	695
2021-03-31	-4.70	-0.018967	-0.0190	242.84	242.84	245.2300	237.2700	238.1100	763
2021-03-30	-6.30	-0.044437	-0.0259	236.54	236.54	238.3600	232.6301	236.8800	624
2021-03-29	-0.87	-0.047952	-0.0037	235.67	235.67	243.2000	233.6100	241.5800	751
2021-03-26	5.36	-0.026299	0.0227	241.03	241.03	241.3100	232.3000	233.6187	821

5 rows x 24 columns

```
In [30]: # set Label as DateTimeIndex
df_QCOM['label'] = pd.DatetimeIndex(df_QCOM.label)
# set index for Label
df_QCOM = df_QCOM.set_index('label')
df_QCOM.head()
```

```
Out[30]:
```

	change	changeOverTime	changePercent	close	fClose	fHigh	fLow	fOpen	fVolume
label									
2021-04-01	0.00	0.000000	0.0000	137.79	137.79	138.4321	133.600	134.50	140695
2021-03-31	-5.20	-0.037739	-0.0377	132.59	132.59	133.5200	130.570	130.96	83612
2021-03-30	-2.48	-0.055737	-0.0187	130.11	130.11	131.2400	129.065	130.17	70972
2021-03-29	1.16	-0.047318	0.0089	131.27	131.27	133.3400	129.770	132.54	97500
2021-03-26	1.72	-0.034836	0.0131	132.99	132.99	133.3500	127.040	127.06	110715

5 rows x 24 columns

```
In [31]: # set Label as DateTimeIndex
df_SHOP['label'] = pd.DatetimeIndex(df_SHOP.label)
# set index for Label
df_SHOP = df_SHOP.set_index('label')
df_SHOP.head()
```

```
Out[31]:
```

	change	changeOverTime	changePercent	close	fClose	fHigh	fLow	fOpen	fV
label									
2021-04-01	0.00	0.000000	0.0000	1155.41	1155.41	1178.9800	1137.02	1150.00	11
2021-03-31	-48.91	-0.042331	-0.0423	1106.50	1106.50	1119.5000	1065.00	1068.01	13
2021-03-30	-57.82	-0.092374	-0.0523	1048.68	1048.68	1074.1873	1012.80	1028.46	15
2021-03-29	-8.77	-0.009965	-0.0084	1039.91	1039.91	1075.8276	1028.60	1053.60	10
2021-03-26	25.30	-0.078068	0.0243	1065.21	1065.21	1084.9800	1005.14	1052.60	28

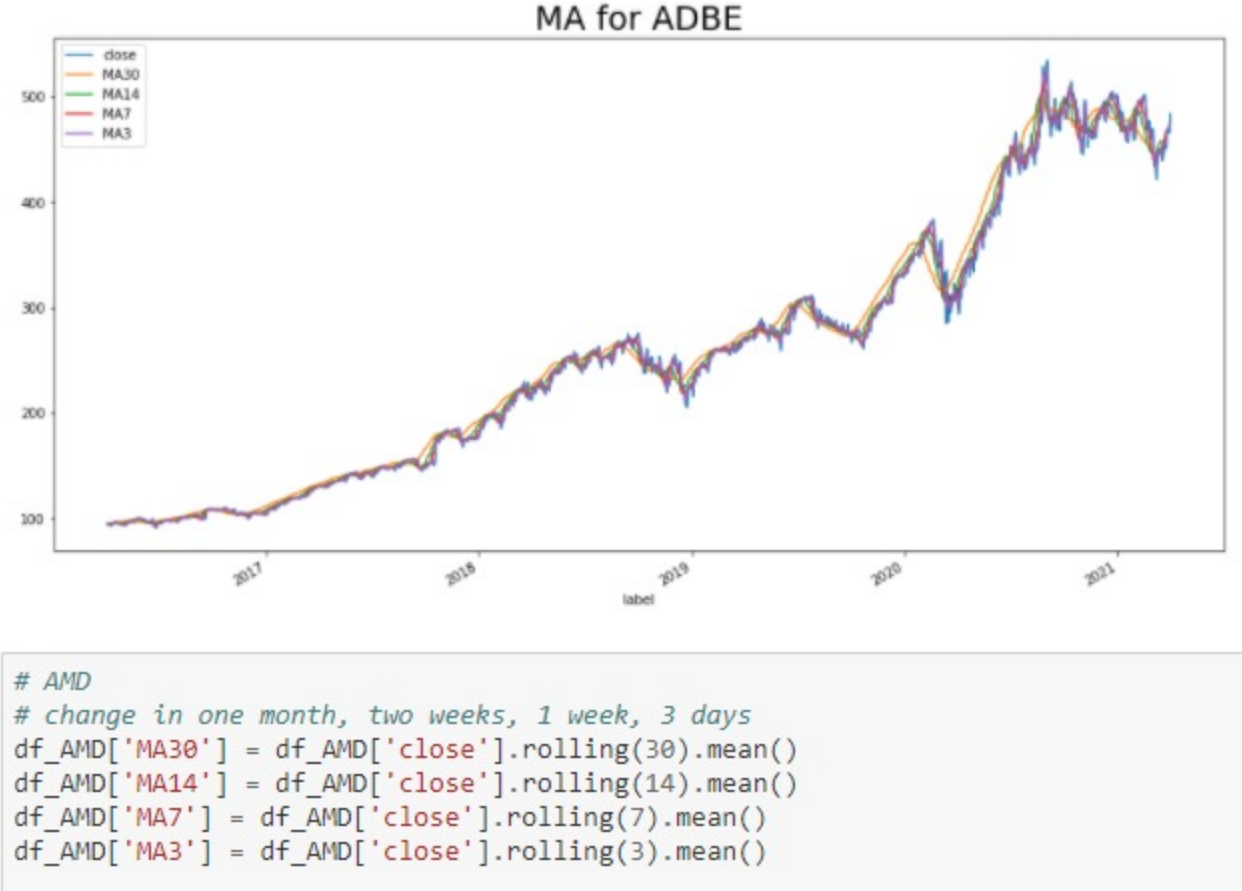
5 rows x 24 columns

Plot the trend of price from 2016 - now

```
In [39]: df_AAPL['close'].plot(label = 'AAPL', figsize = (16,8))
df_ADBE['close'].plot(label = 'ADBE')
df_AMD['close'].plot(label = 'AMD')
df_BABA['close'].plot(label = 'BABA')
df_FB['close'].plot(label = 'FB')
df_INTC['close'].plot(label = 'INTC')
df_MSFT['close'].plot(label = 'MSFT')
df_PVPL['close'].plot(label = 'PVPL')
df_QCOM['close'].plot(label = 'QCOM')
df_SHOP['close'].plot(label = 'SHOP')
plt.legend(loc = 'center left', bbox_to_anchor=(1, 0.5))
plt.title('History trend of price', fontsize = 25)
```

```
Out[39]:
```

Text(0.5, 1.0, 'History trend of price')



Most of companies/corporation in technology field have the increasing in price.
Apple has the most increasing in price, significant.
Consider:

- Why apple increased that much
- Which field in technology increase more than others, why?
- There is a light decrease at the beginning of 2020, does it related to covid-19?

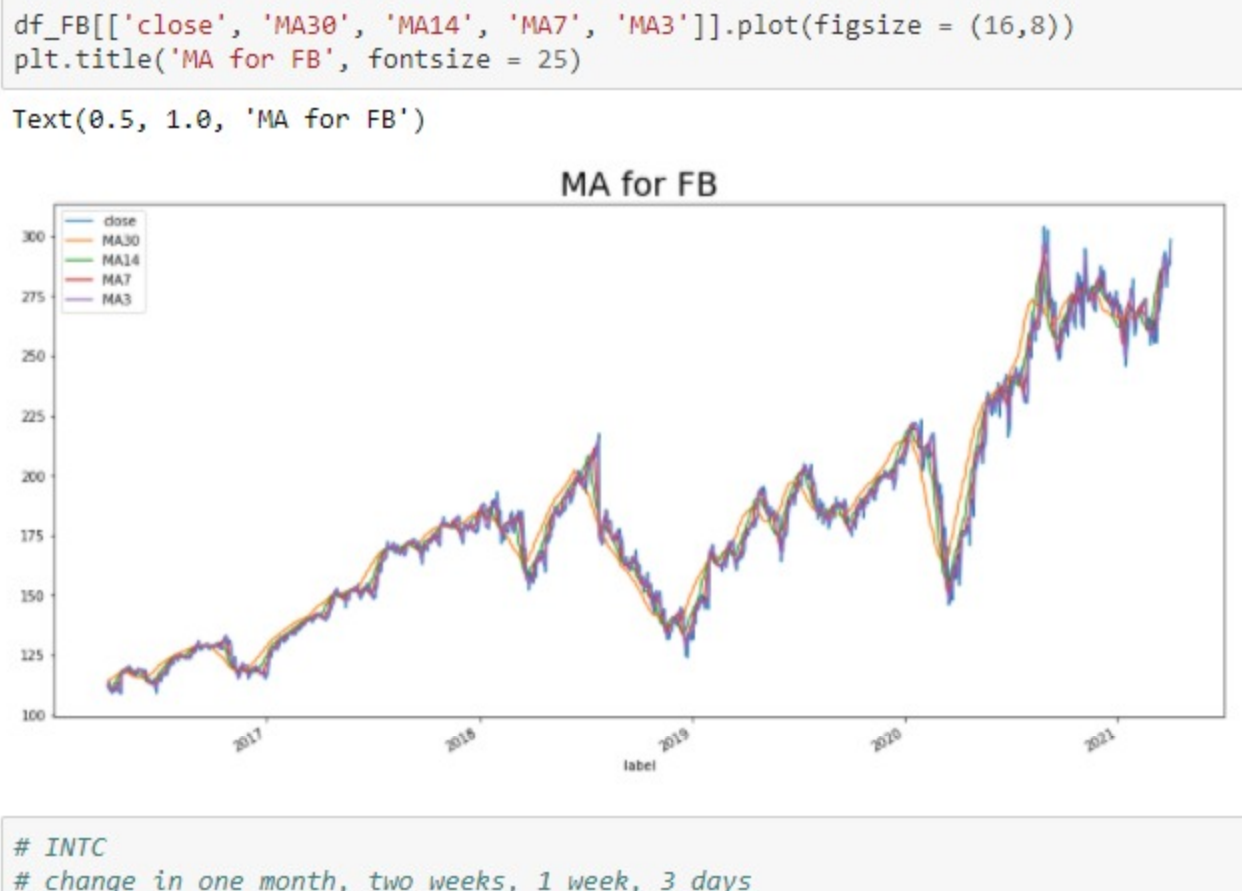
plot the volume of stocks from 2016 - now

Volume is counted as the total number of shares that are actually traded (bought and sold) during the trading day or specified set period of time. It is a measure of the total turnover of shares. Each ticket represents a trade and counted towards the total trading volume.

```
In [40]: df_AAPL['volume'].plot(label = 'AAPL', figsize = (16,8))
df_ADBE['volume'].plot(label = 'ADBE')
df_AMD['volume'].plot(label = 'AMD')
df_BABA['volume'].plot(label = 'BABA')
df_FB['volume'].plot(label = 'FB')
df_INTC['volume'].plot(label = 'INTC')
df_MSFT['volume'].plot(label = 'MSFT')
df_PVPL['volume'].plot(label = 'PVPL')
df_QCOM['volume'].plot(label = 'QCOM')
df_SHOP['volume'].plot(label = 'SHOP')
plt.legend(loc = 'center left', bbox_to_anchor=(1, 0.5))
plt.title('History trend of volume', fontsize = 25)
```

```
Out[40]:
```

Text(0.5, 1.0, 'History trend of volume')



- It's hard to see if just looking at history of volume
- But one thing can see is Apple still be the one was traded the most within 10 tickers
- And another thing is AMD appears to be traded more than other (only lesser than AAPL), while AMD does not show a clear visualization from price plot above. Why?
- AAPL has some significant trading volume days. Check if they launch new products on those days?
- Also AAPL volume increased significant at the beginning of 2020, while the price decreased at that time.

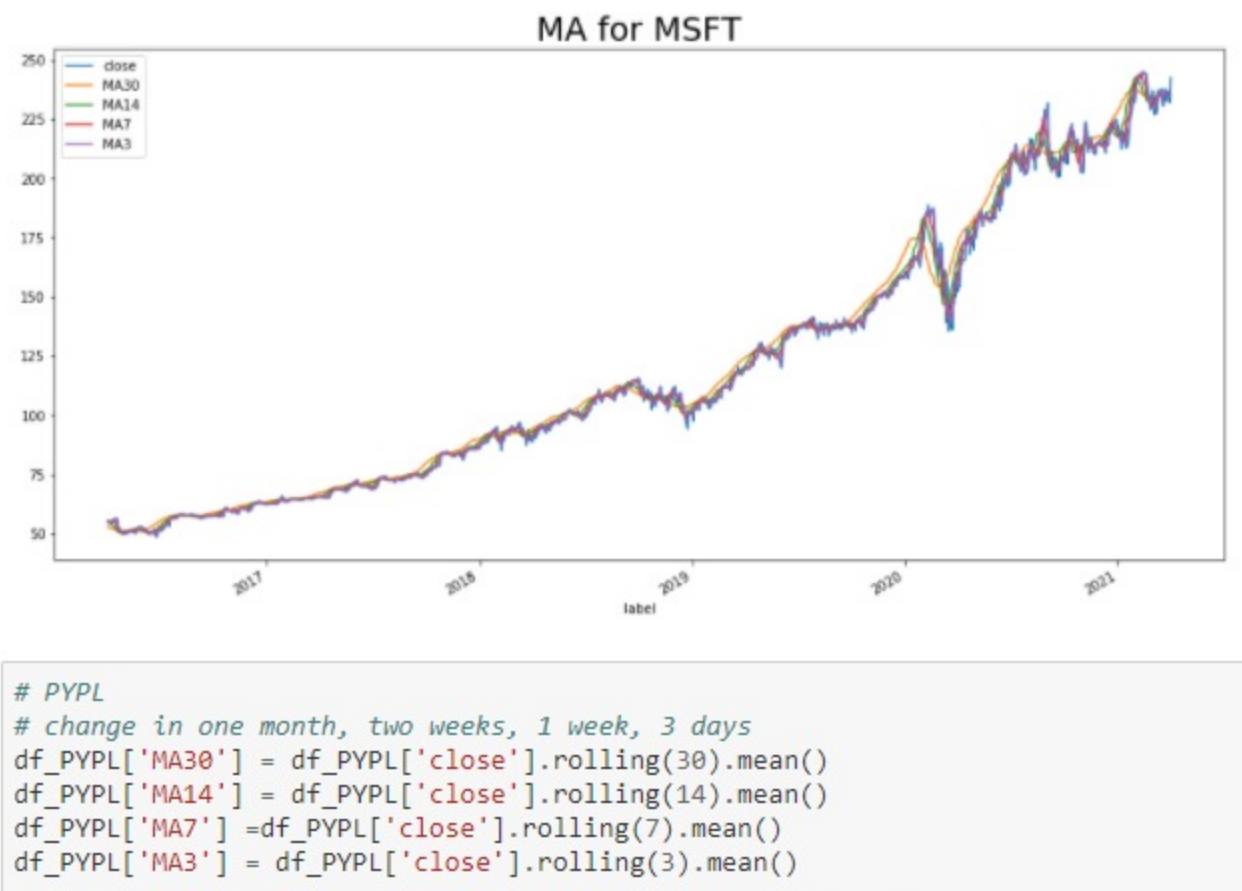
moving average (MA)

```
In [42]: # AAPL
# change in one month, two weeks, 1 week, 3 days
df_AAPL['MA30'] = df_AAPL['close'].rolling(30).mean()
df_AAPL['MA14'] = df_AAPL['close'].rolling(14).mean()
df_AAPL['MA7'] = df_AAPL['close'].rolling(7).mean()
df_AAPL['MA3'] = df_AAPL['close'].rolling(3).mean()

df_AAPL[['close', 'MA30', 'MA14', 'MA7', 'MA3']].plot(label = 'AAPL', figsize = (16,8))
```

```
Out[42]:
```

<matplotlib.axes._subplots.AxesSubplot at 0x1a2e184a2e8>



```
In [45]: # ADBE
# change in one month, two weeks, 1 week, 3 days
df_ADBE['MA30'] = df_ADBE['close'].rolling(30).mean()
df_ADBE['MA14'] = df_ADBE['close'].rolling(14).mean()
df_ADBE['MA7'] = df_ADBE['close'].rolling(7).mean()
df_ADBE['MA3'] = df_ADBE['close'].rolling(3).mean()

df_ADBE[['close', 'MA30', 'MA14', 'MA7', 'MA3']].plot(figsize = (16,8))
plt.title('MA for ADBE', fontsize = 25)
```

```
Out[45]:
```

Text(0.5, 1.0, 'MA for ADBE')

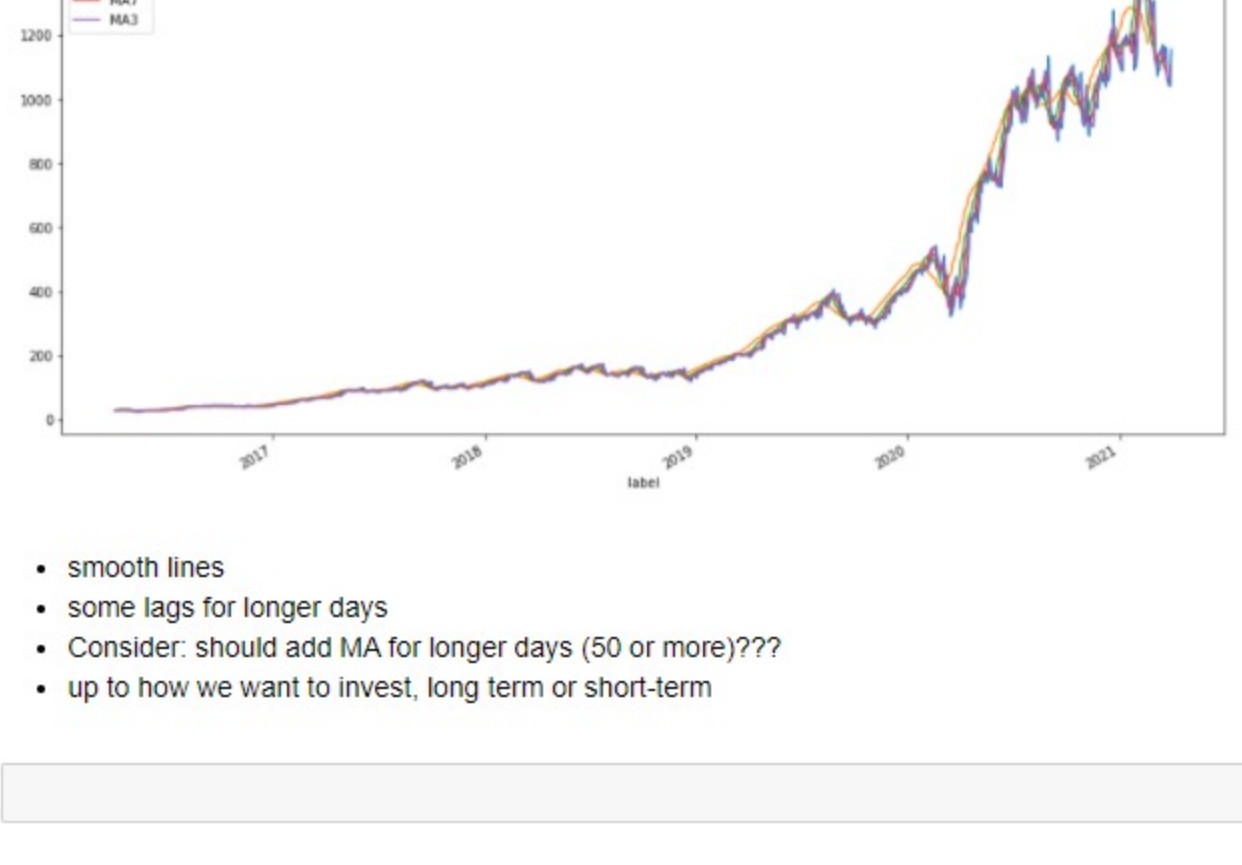


```
In [46]: # AMD
# change in one month, two weeks, 1 week, 3 days
df_AMD['MA30'] = df_AMD['close'].rolling(30).mean()
df_AMD['MA14'] = df_AMD['close'].rolling(14).mean()
df_AMD['MA7'] = df_AMD['close'].rolling(7).mean()
df_AMD['MA3'] = df_AMD['close'].rolling(3).mean()

df_AMD[['close', 'MA30', 'MA14', 'MA7', 'MA3']].plot(figsize = (16,8))
plt.title('MA for AMD', fontsize = 25)
```

```
Out[46]:
```

Text(0.5, 1.0, 'MA for AMD')



```
In [47]: # FB
# change in one month, two weeks, 1 week, 3 days
df_FB['MA30'] = df_FB['close'].rolling(30).mean()
df_FB['MA14'] = df_FB['close'].rolling(14).mean()
df_FB['MA7'] = df_FB['close'].rolling(7).mean()
df_FB['MA3'] = df_FB['close'].rolling(3).mean()

df_FB[['close', 'MA30', 'MA14', 'MA7', 'MA3']].plot(figsize = (16,8))
plt.title('MA for FB', fontsize = 25)
```

```
Out[47]:
```

Text(0.5, 1.0, 'MA for FB')


```
In [48]: # INTC
# change in one month, two weeks, 1 week, 3 days
df_INTC['MA30'] = df_INTC['close'].rolling(30).mean()
df_INTC['MA14'] = df_INTC['close'].rolling(14).mean()
df_INTC['MA7'] = df_INTC['close'].rolling(7).mean()
df_INTC['MA3'] = df_INTC['close'].rolling(3).mean()

df_INTC[['close', 'MA30', 'MA14', 'MA7', 'MA3']].plot(figsize = (16,8))
plt.title('MA for INTC', fontsize = 25)
```

```
Out[48]:
```

Text(0.5, 1.0, 'MA for INTC')


```
In [49]: # MSFT
# change in one month, two weeks, 1 week, 3 days
df_MSFT['MA30'] = df_MSFT['close'].rolling(30).mean()
df_MSFT['MA14'] = df_MSFT['close'].rolling(14).mean()
df_MSFT['MA7'] = df
```