

# 7v81 exercise1

June 10, 2020

```
[9]: import pandas as pd
df= pd.read_csv('realdt.csv')
print (df)
df['date']=pd.to_datetime(df['date'])
```

		date	username	to	replies	retweets	\
0		2020-06-06 22:48:22	realDonaldTrump	NaN	106873	67292	
1		2020-06-06 22:31:20	realDonaldTrump	NaN	57229	49000	
2		2020-06-06 19:56:38	realDonaldTrump	NaN	11659	23813	
3		2020-06-06 14:36:08	realDonaldTrump	NaN	14080	30242	
4		2020-06-06 11:53:36	realDonaldTrump	lily_marston	20324	29531	
...		...	...	...	...	...	
2246		2020-01-01 01:30:35	realDonaldTrump	NaN	53516	80667	
2247		2020-01-01 01:22:28	realDonaldTrump	FLOTUS	12549	26329	
2248		2020-01-01 01:16:27	realDonaldTrump	SteveDaines	4237	12026	
2249		2020-01-01 01:03:15	realDonaldTrump	DanScavino	1922	9868	
2250		2020-01-01 00:55:01	realDonaldTrump	jatoddy	13499	13715	

	favorites		text	geo	\
0	395287		LAW & ORDER!	NaN	
1	282921	Wow! 96% Approval Rating in the Republican Par...		NaN	
2	70283			NaN NaN	
3	131745			NaN NaN	
4	119983	Riot gear or military control is not necessary...		NaN	
...	...				
2246	554774		HAPPY NEW YEAR!	NaN	
2247	128181		Our fantastic First Lady!	NaN	
2248	48559	Thank you Steve. The greatest Witch Hunt in U...		NaN	
2249	37044	Thank you to the @dcexaminer Washington Examin...		NaN	
2250	60116	One of my greatest honors was to have gotten C...		NaN	

	mentions	hashtags	id	\
0	NaN	NaN	1269400770472001539	
1	NaN	NaN	1269396484048465921	
2	NaN	NaN	1269357552854327296	
3	NaN	NaN	1269276894819692544	
4	NaN	NaN	1269235993812570112	
...	...	...	...	

2246	NaN	NaN	1212184310389850119
2247	NaN	NaN	1212182267113680896
2248	NaN	NaN	1212180752936374275
2249	@dcexaminer	NaN	1212177432452698115
2250	NaN	NaN	1212175360093229056

	permalink
0	<a href="https://twitter.com/realDonaldTrump/status/126...">https://twitter.com/realDonaldTrump/status/126...</a>
1	<a href="https://twitter.com/realDonaldTrump/status/126...">https://twitter.com/realDonaldTrump/status/126...</a>
2	<a href="https://twitter.com/realDonaldTrump/status/126...">https://twitter.com/realDonaldTrump/status/126...</a>
3	<a href="https://twitter.com/realDonaldTrump/status/126...">https://twitter.com/realDonaldTrump/status/126...</a>
4	<a href="https://twitter.com/realDonaldTrump/status/126...">https://twitter.com/realDonaldTrump/status/126...</a>
...	...
2246	<a href="https://twitter.com/realDonaldTrump/status/121...">https://twitter.com/realDonaldTrump/status/121...</a>
2247	<a href="https://twitter.com/realDonaldTrump/status/121...">https://twitter.com/realDonaldTrump/status/121...</a>
2248	<a href="https://twitter.com/realDonaldTrump/status/121...">https://twitter.com/realDonaldTrump/status/121...</a>
2249	<a href="https://twitter.com/realDonaldTrump/status/121...">https://twitter.com/realDonaldTrump/status/121...</a>
2250	<a href="https://twitter.com/realDonaldTrump/status/121...">https://twitter.com/realDonaldTrump/status/121...</a>

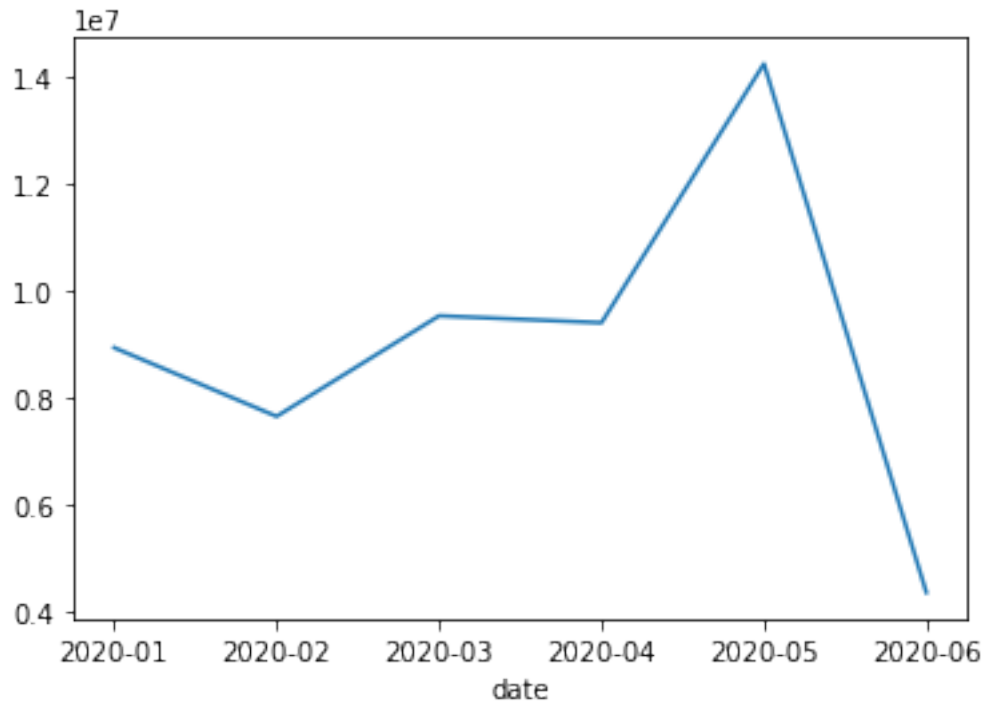
[2251 rows x 12 columns]

```
[24]: df['date']=pd.to_datetime(df['date'])
df_retweets=df.groupby(df['date'].dt.strftime('%Y-%m'))['retweets'].sum()
order=['Jan','Feb','Mar','Apr','May','June']
df2=df.reindex(order, axis=1)
print(df_retweets)
df_retweets.plot()
```

date	
2020-01	8942560
2020-02	7655817
2020-03	9539551
2020-04	9409956
2020-05	14258918
2020-06	4358699

Name: retweets, dtype: int64

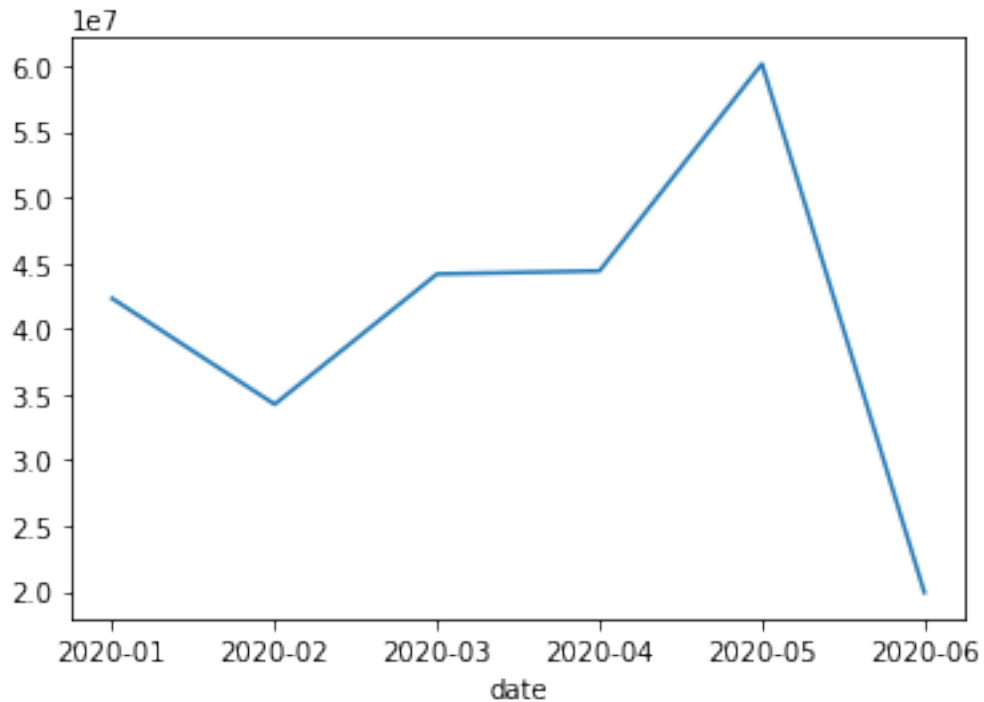
```
[24]: <matplotlib.axes._subplots.AxesSubplot at 0x1231944d0>
```



```
[14]: df['date']=pd.to_datetime(df['date'])
df_fav=df.groupby(df['date'].dt.strftime('%Y-%m'))['favorites'].sum()
order=['Jan','Feb','Mar','Apr','May','June']
df2=df.reindex(order, axis=1)
print(df_fav)
df_fav.plot()
```

```
date
2020-01    42302879
2020-02    34245139
2020-03    44164102
2020-04    44393603
2020-05    60160736
2020-06    19923330
Name: favorites, dtype: int64
```

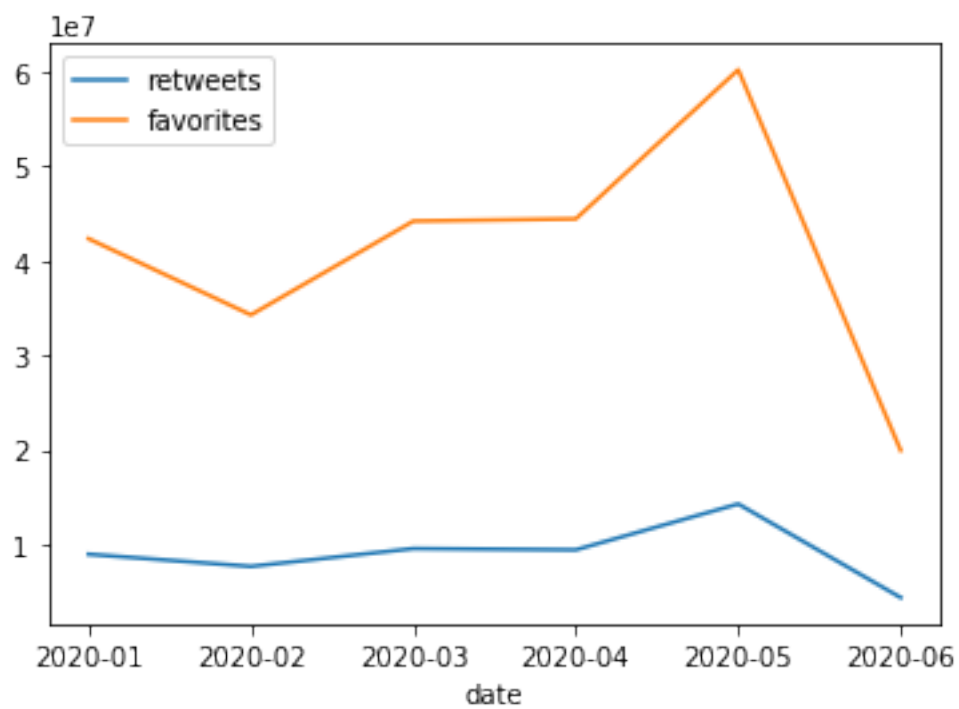
```
[14]: <matplotlib.axes._subplots.AxesSubplot at 0x111ce7b10>
```



```
[28]: df['date']=pd.to_datetime(df['date'])
df_retweets=df.groupby(df['date'].dt.strftime('%Y-%m'))['retweets'].sum()
df_fav=df.groupby(df['date'].dt.strftime('%Y-%m'))['favorites'].sum()
order=['Jan','Feb','Mar','Apr','May','June']
df1=df.reindex(order, axis=0)
print(df_retweets,df_fav)
df3 = pd.concat([df_retweets, df_fav], axis=1)
df3.plot()
```

```
date
2020-01      8942560
2020-02      7655817
2020-03      9539551
2020-04      9409956
2020-05     14258918
2020-06      4358699
Name: retweets, dtype: int64
date
2020-01     42302879
2020-02     34245139
2020-03     44164102
2020-04     44393603
2020-05     60160736
2020-06     19923330
Name: favorites, dtype: int64
```

[28]: <matplotlib.axes.\_subplots.AxesSubplot at 0x1238c20d0>



[ ]: