5-Pandas (visualization)

September 13, 2019

___ ## Pedram Jahangiry (Fall 2019)

1 Pandas visualization

[28]:
```
import numpy as np
import pandas as pd
```

1.1 The Data

There are some fake data csv files you can read in as dataframes:

[29]:
```
df = pd.read_csv('GDP.csv')
```

[30]:
```
df['GDP_t-1']= df['GDP'].shift(1)
df['growth']= np.log(df['GDP']/df['GDP'].shift(1))

df.tail(5)
```

[30]:

	DATE	GDP	GDP_t-1	growth
285	2018-04-01	20510.177	20163.159	0.017064
286	2018-07-01	20749.752	20510.177	0.011613
287	2018-10-01	20897.804	20749.752	0.007110
288	2019-01-01	21098.827	20897.804	0.009573
289	2019-04-01	21339.121	21098.827	0.011325

[31]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 290 entries, 0 to 289
Data columns (total 4 columns):
DATE       290 non-null object
GDP        290 non-null float64
GDP_t-1    289 non-null float64
growth     289 non-null float64
dtypes: float64(3), object(1)
memory usage: 9.1+ KB
```

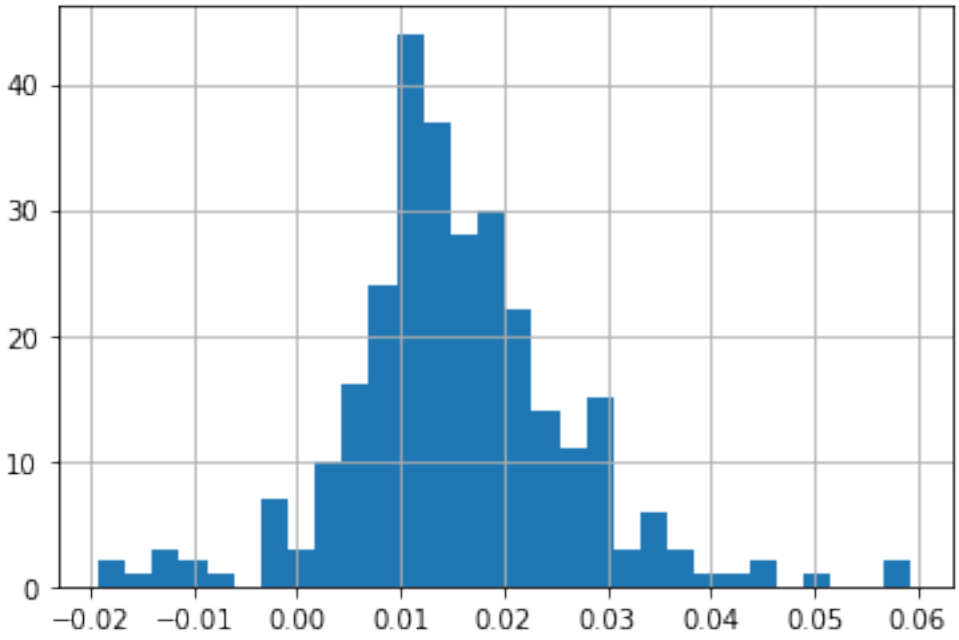
[32]:
```
df['DATE']= pd.to_datetime(df.DATE)
df.set_index('DATE',inplace=True)
df.head()
```

[32]:

DATE	GDP	GDP_t-1	growth
1947-01-01	243.164	NaN	NaN
1947-04-01	245.968	243.164	0.011465
1947-07-01	249.585	245.968	0.014598
1947-10-01	259.745	249.585	0.039901
1948-01-01	265.742	259.745	0.022826

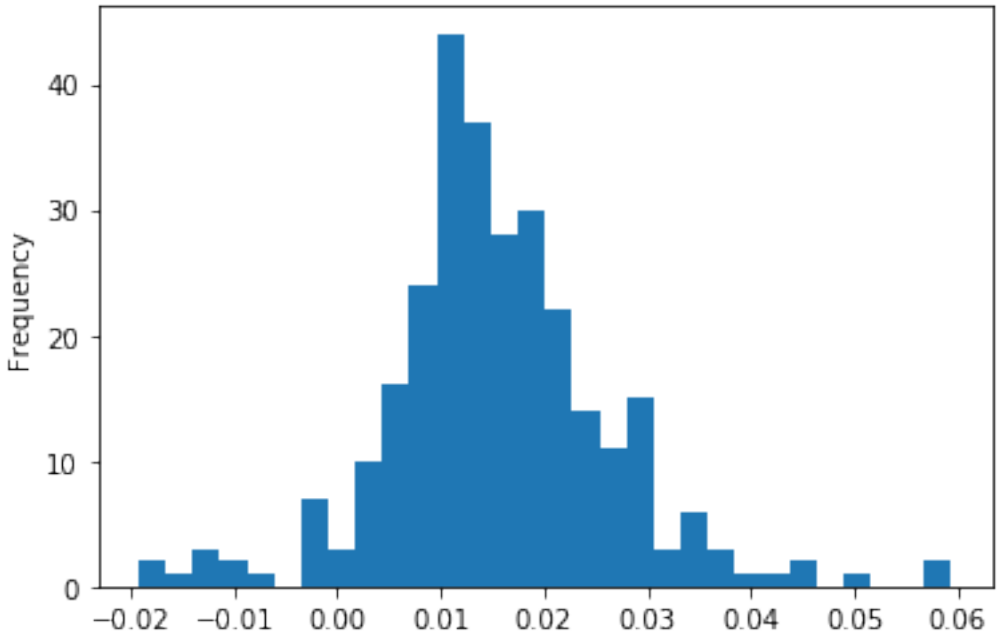
[33]:
```
df['growth'].hist(bins=30)
```

[33]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x18b4eedc748>
```



[34]:
```
df['growth'].plot(kind="hist", bins=30)
```

[34]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x18b4ef49f98>
```



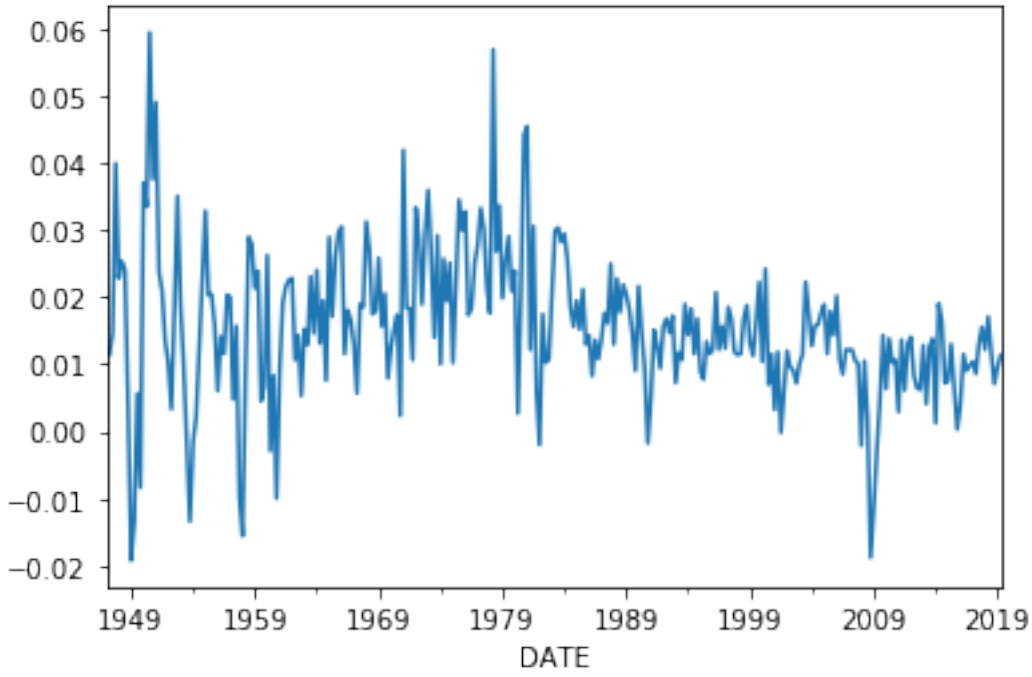
There are many different plot types in pandas, try to explore the followings on your own:

- df.plot.area
- df.plot.pie
- df.plot.hist
- df.plot.line
- df.plot.scatter
- df.plot.bar
- df.plot.box
- df.plot.kde
- df.plot.density

You can also just call df.plot(kind='hist') or replace that kind argument with any of the key terms shown in the list above ___

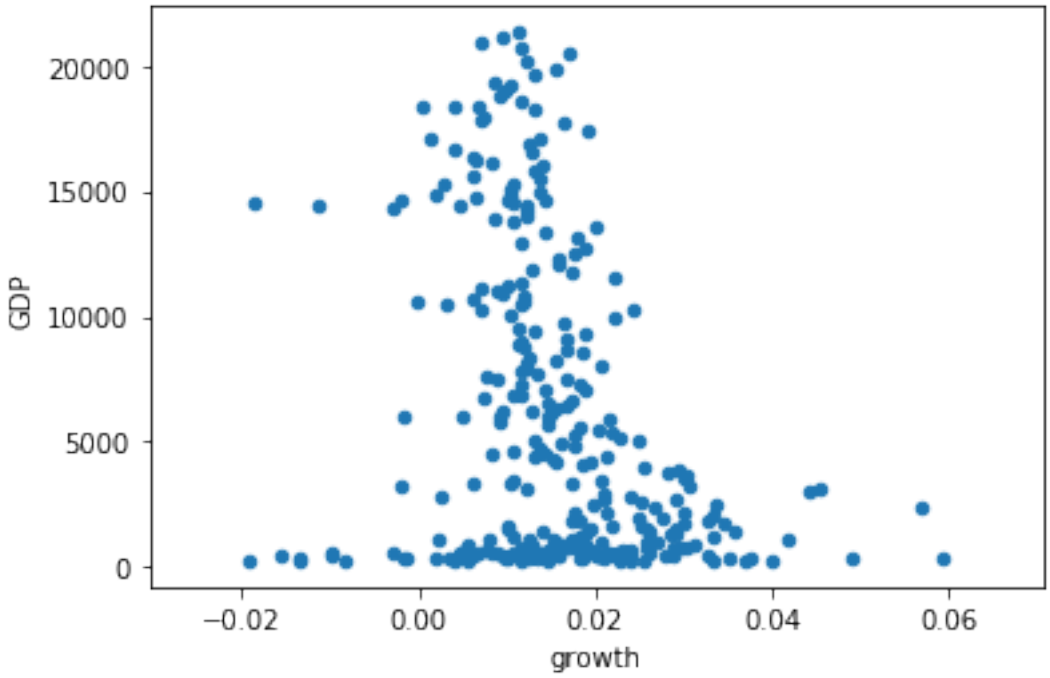
[37]:
```
df['growth'].plot()    # by default the kind ='line'
```

[37]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x18b4efe1a20>
```



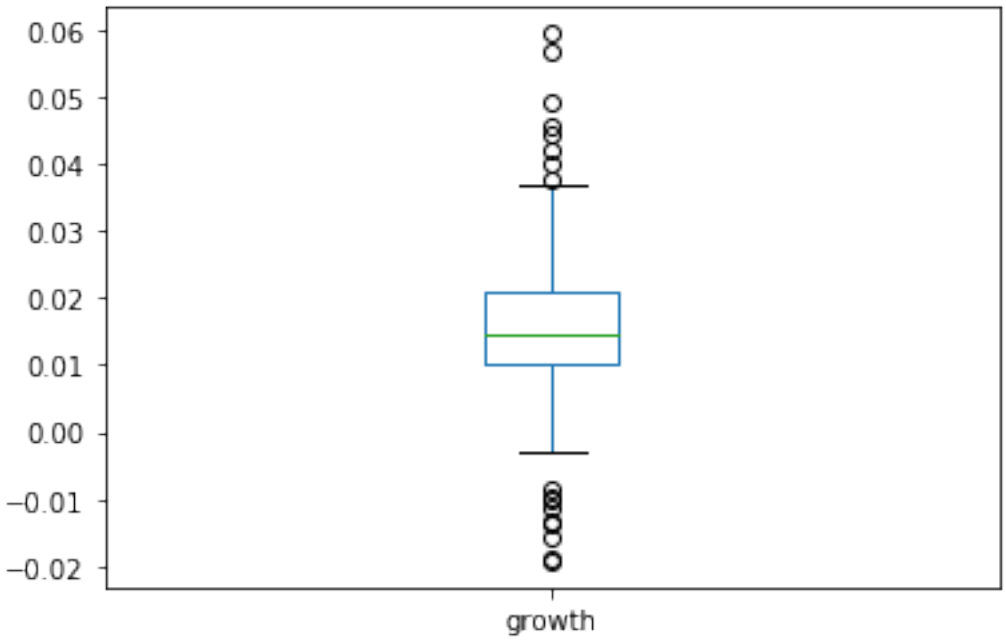
[36]:
```
df.plot.scatter(x='growth',y='GDP')
```

[36]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x18b4efe1c50>
```



[35]:
```
df['growth'].plot.box()
```

[35]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x239d9a7b198>
```



[37]:
```
# Kernel Density Estimation plot (KDE)
df['growth'].plot.kde()
```

[37]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x239d9ace9e8>
```