

**Міністерство освіти і науки України
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Факультет інформатики та обчислювальної техніки
Кафедра обчислювальної техніки**

**Лабораторна робота №5
з дисципліни
«Аналіз даних з використанням мови Python»**

Виконав:
студент групи ПІ-04
Пащенко Дмитро Олексійович

Перевірила:
Тимофєєва Ю. С.

Київ 2022

Код програми та результат виконання

```
In [65]: import pandas as pd
import matplotlib.pyplot as plt
```

```
In [5]: data = pd.read_csv('Delhi_Climate.csv')
data
```

```
Out[5]:
```

	date	meantemp	humidity	wind_speed	meanpressure
0	2013-01-01	10.000000	84.500000	0.0	1015.666667
1	2013-01-02	7.400000	92.000000	2.98	1017.800000
2	2013-01-03	7.166667	87.000000	4.633333333333334	1018.666667
3	2013-01-04	8.666667	71.333333	1.2333333333333334	1017.166667
4	2013-01-05	6.000000	86.833333	3.6999999999999997	1016.500000
...
1457	2016-12-28	17.217391	68.043478	3.547826086956522	1015.565217
1458	2016-12-29	15.238095	87.857143	6.0	1016.904762
1459	2016-12-30	14.095238	89.666667	6.266666666666667	1017.904762
1460	2016-12-31	15.052632	87.000000	7.325	1016.100000
1461	2017-01-01	10.000000	100.000000	0.0	1016.000000

Створити не менше двох об'єктів TimeSeries, у яких індекси створені за допомогою date_range(). Виділити підмасиви у цих об'єктах.

```
In [17]: time_series_index = pd.date_range('2013-01-01', periods = 5, freq = 'MS')
time_series = pd.Series([456, 333, 938, 800, 850], index = time_series_index)
time_series
```

```
Out[17]: 2013-01-01    456
2013-02-01    333
2013-03-01    938
2013-04-01    800
2013-05-01    850
Freq: MS, dtype: int64
```

```
In [26]: time_series.loc['2013-02-01':'2013-05-01']
# time_series.loc['2013-02-01':]
```

```
Out[26]: 2013-02-01    333
2013-03-01    938
2013-04-01    800
2013-05-01    850
Freq: MS, dtype: int64
```

```
In [28]: summer_time_series_index = pd.date_range('2015-06-01', periods = 13, freq = 'W')
summer_time_series = pd.Series([900, 850, 800, 750, 700, 650, 600, 550, 500, 450, 400, 350, 300], index = summer_time_series_index)
summer_time_series
```

```
Out[28]: 2015-06-07    900
2015-06-14    850
2015-06-21    800
2015-06-28    750
2015-07-05    700
2015-07-12    650
2015-07-19    600
2015-07-26    550
2015-08-02    500
2015-08-09    450
2015-08-16    400
2015-08-23    350
2015-08-30    300
Freq: W-SUN, dtype: int64
```

```
In [29]: summer_time_series.loc['2015-07-01':'2015-07-30'] # лише липень
```

```
Out[29]: 2015-07-05    700
2015-07-12    650
2015-07-19    600
2015-07-26    550
Freq: W-SUN, dtype: int64
```

Побудувати графік зміни середніх денних температур:

```
In [86]: mean_temp = pd.read_csv('Delhi_Climate.csv', index_col = 'date', parse_dates = True, usecols = ['date', 'meantemp'])
mean_temp
```

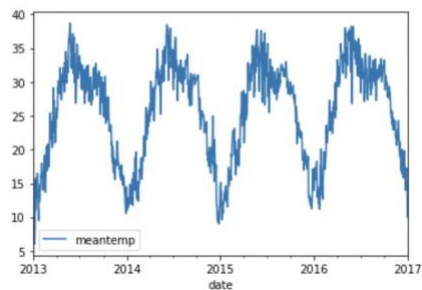
```
Out[86]:
```

meantemp	
date	
2013-01-01	10.000000
2013-01-02	7.400000
2013-01-03	7.166667
2013-01-04	8.666667
2013-01-05	6.000000
...	...
2016-12-28	17.217391
2016-12-29	15.238095
2016-12-30	14.095238
2016-12-31	15.052632

а) загальний;

```
In [87]: mean_temp.plot()
```

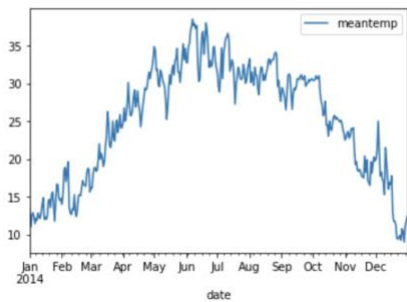
```
Out[87]: <AxesSubplot:xlabel='date'>
```



б) за 2014 рік;

```
In [88]: mean_temp.loc['2014'].plot()
```

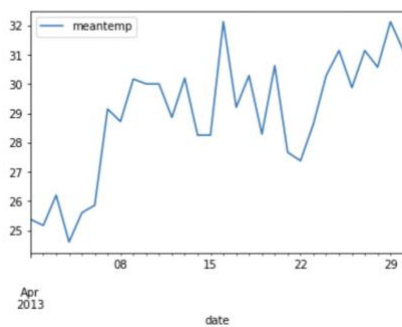
```
Out[88]: <AxesSubplot:xlabel='date'>
```



в) за квітень 2013 року;

```
In [89]: mean_temp.loc['2013-04'].plot()
```

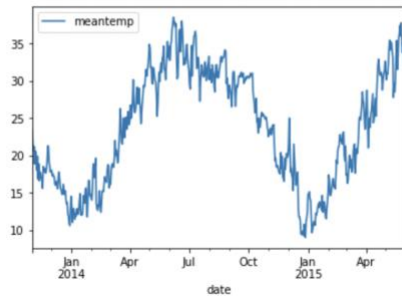
```
Out[89]: <AxesSubplot:xlabel='date'>
```



г) за листопад 2013 – травень 2015;

```
In [90]: mean_temp.loc['2013-11':'2015-05'].plot()
```

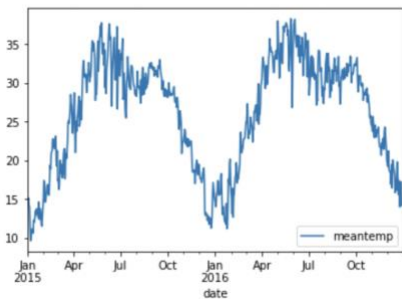
```
Out[90]: <AxesSubplot: xlabel='date'>
```



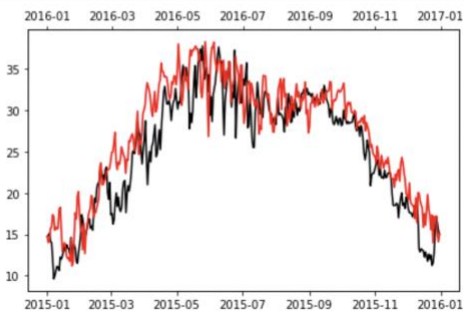
д) за 2015 та 2016 на одному графіку.

```
In [91]: mean_temp.loc['2015':'2016'].plot()
```

```
Out[91]: <AxesSubplot: xlabel='date'>
```



```
In [92]: fig, ax1 = plt.subplots()
ax1.plot(mean_temp.loc['2015'], color = 'black')
ax2 = ax1.twinx()
ax2.plot(mean_temp.loc['2016'], color = 'red')
fig.tight_layout()
```



Знайти середні значення вологості

```
In [52]: humidity = pd.read_csv('Delhi_Climate.csv', index_col = 'date', parse_dates = True, usecols = ['date', 'humidity'])
humidity
```

```
Out[52]:
```

humidity	
date	
2013-01-01	84.500000
2013-01-02	92.000000
2013-01-03	87.000000
2013-01-04	71.333333
2013-01-05	86.833333
...	...
2016-12-28	68.043478
2016-12-29	87.857143
2016-12-30	89.666667
2016-12-31	87.000000

а) за 2016 рік;

```
In [53]: humidity.loc['2016'].mean()
Out[53]: humidity    58.740174
dtype: float64
```

б) за кожний місяць;

```
In [56]: humidity.resample('M').mean()
```

```
Out[56]:
```

humidity	
date	
2013-01-31	73.028802
2013-02-28	71.938563
2013-03-31	57.417243
2013-04-30	34.612103
2013-05-31	28.938249
2013-06-30	58.758135
2013-07-31	74.873387
2013-08-31	76.757104
2013-09-30	64.552738
2013-10-31	70.255607

в) за кожні два тижні весни та літа 2014 року.

```
In [58]: humidity.loc['2014-03':'2014-08'].resample('2W').mean()
```

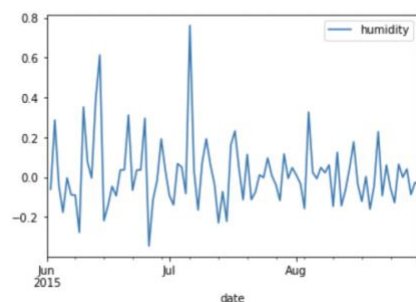
```
Out[58]:
```

humidity	
date	
2014-03-02	83.142857
2014-03-16	63.723214
2014-03-30	59.250000
2014-04-13	41.543367
2014-04-27	42.857143
2014-05-11	36.492347
2014-05-25	42.383929
2014-06-08	36.991071
2014-06-22	42.982143
2014-07-06	59.030987

г) Розрахувати і зобразити зміни вологості у відсотках за кожен день впродовж літа 2015 року.

```
In [61]: humidity.loc['2015-06':'2015-08'].pct_change().plot()
```

```
Out[61]: <AxesSubplot:xlabel='date'>
```



д) Знайти та зобразити графічно ковзне середнє вологості за 2013 рік з вікном в місяць.

```
In [64]: humidity.loc['2013'].rolling('30D').mean().plot()
```

```
Out[64]: <AxesSubplot:xlabel='date'>
```

