

## Лабораторна робота №5

### Робота з часовими рядами в Pandas

**Мета роботи:** Ознайомитись з основними представленнями часових рядів в Pandas та операціями з ними.

#### Короткі теоретичні відомості

Дані про дату та час можуть бути представлені по-різному:

- часові мітки, конкретні моменти; Timestamp в Pandas;
- фіксовані періоди; Period в Pandas;
- часові інтервали, що позначаються мітками початку та кінця; періоди вважатимуться окремими випадками інтервалів;
- час експерименту або час, що минув; кожна часові мітка вимірює час, що минув з початкового моменту. Timedelta в Pandas.

Бібліотека Pandas надає об'єкт Timestamp. Вона може створювати з кількох таких об'єктів Timestamp об'єкт класу DatetimeIndex, який можна використовувати для індексації даних в об'єктах Series або DataFrame.

```
date = pd.to_datetime('2021-09-04')
```

Функція pd.date\_range() створює регулярну послідовність дат, приймаючи на вході початкову дату, кінцеву дату та необов'язковий період.

```
pd.date_range('2021-08-01', periods=5, freq='M')
```

Можливі значення параметру freq:

D – день; B – робочі дні; W – тиждень; M – місяць; A, Y – рік; H – година; T, min – хвилина; S – секунда.

Додавання S в кінці означає початок. Наприклад, SY – початок року. За замовчуванням використовується кінець періоду.

Можна створити об'єкт Series з індексацією у вигляді дат:

```
ind = pd.DatetimeIndex(['2020-08-01', '2020-09-01', '2020-10-01', '2021-10-01', '2021-11-02'])
```

```
dt = pd.Series([100, 200, 250, 150, 300], index=ind)
```

Тоді можна виводити та обробляти дані за конкретний рік, місяць чи період.

```
dt.loc['2020']
```

```
dt.loc['2020-09-15':'2021-10-15']
```

```
dt.loc['2020-10']
```

При роботі з даними часових рядів часто буває необхідно перерозбити з використанням інтервалів інший періодичності. Зробити це можна за допомогою методу resample() або набагато простішого методу asfreq(). Основна різниця між ними полягає в тому, що resample() виконує агрегування даних, а asfreq() - вибірку даних.

```
df.resample('M').mean()
```

```
df.asfreq('M')
```

Метод pct\_change() розраховує зміну у відсотках за заданий період. Ковзаючі статистичні показники - ще один з різновидів операцій, призначених

для часових рядів. Працювати з ними можна за допомогою атрибуту `rolling()` об'єктів `Series` і `DataFrame`, що повертає представлення, схоже на виконання операції `groupby`. Далі можна застосовувати будь-яку агрегуючу функцію. Розмір вікна можна задавати як у вигляді цілого числа періодів, так у вигляді частоти (наприклад, 'M' – розмір вікна дорівнює місяцю).

```
df.rolling(30).mean().plot()
```

### **Завдання до лабораторної роботи**

Створити програму, яка виконує наступні завдання:

1. Створити не менше двох об'єктів `TimeSeries`, у яких індекси створені за допомогою `date_range()`. Виділити підмасиви у цих об'єктів.
2. Виконати завдання відповідно до варіанту.

Оформити звіт. Звіт повинен містити:

- титульний лист;
- код програми;
- результати виконання коду.

Продемонструвати роботу програми та відповісти на питання стосовно теоретичних відомостей та роботи програми.

Варіант 1.

Файл `Delhi_Climate.csv`.

1. Побудувати графік зміни середніх денних температур: а) загальний; б) за 2014 рік; в) за квітень 2013 року; г) за листопад 2013 – травень 2015; д) за 2015 та 2016 на одному графіку.
2. Знайти середні значення вологості а) за 2016 рік; б) за кожний місяць; в) за кожні два тижні весни та літа 2014 року. г) Розрахувати і зобразити зміни вологості у відсотках за кожен день впродовж літа 2015 року. д) Знайти та зобразити графічно ковзне середнє вологості за 2013 рік з вікном в місяць.

Варіант 2.

Файл `Microsoft_Stock.csv`.

1. Побудувати графік зміни ціни на час відкриття біржі: а) загальний; б) за 2020 рік; в) за лютий 2021 року; г) за травень 2017 – червень 2019; д) за 2019 та 2020 на одному графіку.
2. Знайти середні значення найменшої ціни за день а) за 2019 рік; б) за кожний місяць; в) за кожні два тижні літа та осені 2020 року. г) Розрахувати і зобразити зміни найменшої ціни за день у відсотках за кожен день впродовж осені 2018 року. д) Знайти та зобразити графічно ковзне середнє найменшої ціни за день за 2018 рік з вікном в місяць.

Варіант 3.

Файл `digital_currency.csv`.

1. Побудувати графік зміни ціни на час відкриття біржі (в доларах): а) загальний; б) за 2019 рік; в) за грудень 2020 року; г) за жовтень 2018 – липень 2020; д) за 2019 та 2020 на одному графіку.
2. Знайти середні значення найменшої ціни за день (в доларах): а) за 2019 рік; б) за кожний місяць; в) за кожний тиждень весни 2020 року. г) Розрахувати і зобразити зміни найменшої ціни за день у відсотках за кожен день впродовж весни 2020 року. д) Знайти та зобразити графічно ковзне середнє найменшої ціни за день за 2019 рік з вікном в місяць.

Варіант 4.

Файл VChain.csv.

1. Побудувати графік зміни кількості біткоінів в обігу: а) загальний; б) за 2017 рік; в) за квітень 2016 року; г) за лютий 2010 – грудень 2015; д) за 2011 та 2015 на одному графіку.
2. Знайти максимальні значення розміру блоку: а) за 2012 рік; б) за кожний рік; в) за кожний місяць 2017 року. г) Розрахувати і зобразити зміни значення розміру блоку у відсотках за кожні чотири дні впродовж осені 2015 року. д) Знайти та зобразити графічно ковзне середнє розміру блоку за 2016 рік з вікном в місяць.

Варіант 5.

Файл Amazon.csv.

1. Побудувати графік зміни ціни на час закриття біржі: а) загальний; б) за 2018 рік; в) за січень 2020 року; г) за грудень 2016 – лютий 2018; д) за 2016 та 2017 на одному графіку.
2. Знайти максимальні значення найбільшої ціни за день а) за 2016 рік; б) за кожний рік; в) за кожний тиждень весни 2019 року. г) Розрахувати і зобразити зміни значення найбільшої ціни за день у відсотках за кожен день впродовж весни 2018 року. д) Знайти та зобразити графічно ковзне середнє найбільшої ціни за день за 2017 рік з вікном в два місяці.

Варіант 6.

Файл Delhi\_Climate.csv.

1. Побудувати графік зміни вологості: а) загальний; б) за 2013 рік; в) за лютий 2016 року; г) за січень 2014 – березень 2016; д) за 2014 та 2015 рік на одному графіку.
2. Знайти середні значення швидкості вітру а) за 2015 рік; б) за кожний місяць 2014 року; в) за кожний тиждень осені 2016 року. г) Розрахувати і зобразити зміни швидкості вітру у відсотках за кожен день впродовж весни 2013 року. д) Знайти та зобразити графічно ковзне середнє швидкості вітру за лютий 2015 року з вікном в 4 дні.

Варіант 7.

Файл Microsoft\_Stock.csv.

1. Побудувати графік зміни ціни на час закриття біржі: а) загальний; б) за 2019 рік; в) за вересень 2018 року; г) за листопад 2015 – січень 2018; д) за 2017 та 2018 на одному графіку.
2. Знайти середні значення найбільшої ціни за день а) за 2016 рік; б) за кожний місяць; в) за кожний тиждень першого кварталу 2019 року. г) Розрахувати і зобразити зміни найбільшої ціни за день у відсотках за кожен день впродовж літа 2019 року. д) Знайти та зобразити графічно ковзне середнє найбільшої ціни за день за вересень 2017 року з вікном в 5 днів.

Варіант 8.

Файл digital\_currency.csv.

1. Побудувати графік зміни ціни на час закриття біржі (в доларах): а) загальний; б) за 2020 рік; в) за вересень 2018 року; г) за листопад 2018 – грудень 2020; д) за 2019 та 2020 на одному графіку.
2. Знайти середні значення найбільшої ціни за день (в доларах): а) за 2019 рік; б) за кожний місяць; в) за кожний тиждень третього кварталу 2019 року. г) Розрахувати і зобразити зміни найбільшої ціни за день (в доларах) у відсотках за кожен день впродовж осені 2018 року. д) Знайти та зобразити графічно ковзне середнє найбільшої ціни за день (в доларах) за квітень 2019 року з вікном в 6 днів.

Варіант 9.

Файл BChain.csv.

1. Побудувати графік зміни кількості транзакцій на блок: а) загальний; б) за 2018 рік; в) за квітень 2016 року; г) за січень 2015 – червень 2017; д) за 2015 та 2017 на одному графіку.
2. Знайти максимальні значення кількості біткоінів в обігу: а) за 2010 рік; б) за кожний рік (на початок року); в) за кожний місяць 2017 року. г) Розрахувати і зобразити зміни кількості біткоінів в обігу у відсотках за кожні два дні впродовж літа 2018 року. д) Знайти та зобразити графічно ковзне середнє кількості біткоінів в обігу за липень 2018 року з вікном в 6 днів.

Варіант 10.

Файл Amazon.csv.

1. Побудувати графік зміни ціни на час відкриття біржі: а) загальний; б) за 2019 рік; в) за липень 2020 року; г) за жовтень 2014 – червень 2016; д) за 2015 та 2017 на одному графіку.
2. Знайти мінімальні значення найменшої ціни за день: а) за 2018 рік; б) за кожний місяць 2019; в) за кожний тиждень другого кварталу 2015 року. г) Розрахувати і зобразити зміни найменшої ціни за день у відсотках за кожні два дні впродовж осені 2016 року. д) Знайти та зобразити графічно ковзне середнє найменшої ціни за день за травень 2016 року з вікном в тиждень.

### Варіант 11

Файл Delhi\_Climate.csv.

1. Побудувати графік зміни швидкості вітру: а) загальний; б) за 2014 рік; в) за червень 2013 року; г) за 15 жовтня 2013 – 16 квітня 2015; д) за 2013 та 2015 на одному графіку.
2. Знайти середні значення середніх денних температур а) за 2014 рік; б) за кожний другий місяць 2013 року; в) за кожний тиждень літа 2015 року. г) Розрахувати і зобразити зміни середніх денних температур у відсотках за кожен день впродовж весни 2013 року. д) Знайти та зобразити графічно ковзне середнє середніх денних температур за січень 2014 року з вікном в 3 дні.

### Варіант 12.

Файл Microsoft\_Stock.csv.

1. Побудувати графік зміни об'єму продаж: а) загальний; б) за 2019 рік; в) за січень 2020 року; г) з 22 листопада 2019 – до серпня 2020; д) за 2016 та 2017 на одному графіку.
2. Знайти середні значення ціни на час закриття біржі: а) за 2019 рік; б) за кожний місяць; в) за кожні два тижні літа та осені 2020 року. г) Розрахувати і зобразити зміни ціни на час закриття біржі у відсотках за кожен день впродовж 2018 року. д) Знайти та зобразити графічно ковзне середнє ціни на час закриття біржі за 2020 рік з вікном в місяць.

### Варіант 13.

Файл digital\_currency.csv.

1. Побудувати графік зміни об'єму продаж: а) загальний; б) за 2020 рік; в) за лютий 2019 року; г) з червня 2018 – до липня 2020; д) за 2018 та 2020 на одному графіку.
2. Знайти середні значення ціни на час закриття біржі (в доларах): а) за 2020 рік; б) за кожний місяць; в) за кожні два тижні весни та літа 2019 року. г) Розрахувати і зобразити зміни ціни на час закриття біржі (в доларах) у відсотках за кожен день впродовж 2019 року. д) Знайти та зобразити графічно ковзне середнє ціни на час закриття біржі (в доларах) за 2020 рік з вікном в робочий тиждень.

### Варіант 14.

Файл BChain.csv.

1. Побудувати графік зміни кількості розміру блоку: а) загальний; б) за 2014 рік; в) за грудень 2017 року; г) за листопад 2013 – вересень 2016; д) за 2015 та 2017 на одному графіку.
2. Знайти максимальні значення кількості біткоінів в обігу: а) за 2016 рік; б) за кожний рік (на початок року); в) за кожний місяць 2015 року. г) Розрахувати і зобразити зміни кількості біткоінів в обігу у відсотках за кожні два дні впродовж весни 2016 року. д) Знайти та

зобразити графічно ковзне середнє кількості біткоїнів в обігу за 2017 рік з вікном в місяць.

Варіант 15.

Файл Amazon.csv.

1. Побудувати графік зміни об'ємів продаж: а) загальний; б) за 2006 рік; в) за лютий 2020 року; г) за лютий 2008 – травень 2010; д) за 2016 та 2018 на одному графіку.
2. Знайти середні значення ціни на час відкриття біржі: а) за 2012 рік; б) за кожний рік; в) за кожний тиждень четвертого кварталу 2016 року. г) Розрахувати і зобразити зміни ціни на час відкриття біржі у відсотках за кожний день впродовж осені 2011 року. д) Знайти та зобразити графічно ковзне середнє ціни на час відкриття біржі за 2019 рік з вікном в місяць.