Міністерство освіти і науки України

Національний технічний університет України  
«Київський політехнічний інститут

імені Ігоря Сікорського»

Факультет БІОМЕДИЧНОЇ ІНЖЕНЕРІЇ

Кафедра БІОМЕДИЧНОЇ КІБЕРНЕТИКИ

ЗВІТ

про виконання розрахунково-графічної роботи

з кредитного модуля «Обробка та аналіз біомедичних даних»

Виконав: студент гр. БС-83 Костюк Андрій

Прийняв: викладач 2020 р.

(Прізвище, Ім’я) (Оцінка, Підпис) (Дата)

Київ

КПІ ім. Ігоря Сікорського

2020

ЗМІСТ

[**РОЗДІЛ 1. КОМП’ЮТЕРНИЙ ПРАКТИКУМ №1** 4](#_Toc41327686)

[**РОЗДІЛ 2. ВІЗУАЛІЗАЦІЯ ДАНИХ ЗА ДОПОМОГОЮ БІБЛІОТЕКИ MATPLOTLIB** 7](#_Toc41327687)

[2.1 Теоретичні відомості 7](#_Toc41327688)

[2.2 Практична реалізація 7](#_Toc41327689)

[**РОЗДІЛ 3. ВІЗУАЛІЗАЦІЯ ДАНИХ ЗА ДОПОМОГОЮ БІБЛІОТЕКИ SEABORN** 14](#_Toc41327690)

[3.1 Теоретичні відомості 14](#_Toc41327691)

[3.2 Практична реалізація 14](#_Toc41327692)

[**РОЗДІЛ 4. ПЕРЕВІРКА ДАНИХ НА НОРМАЛЬНІСТЬ ЗА ДОПОМОГОЮ КРИТЕРІЮ ШАПІРА-УІЛКА** 19](#_Toc41327693)

[4.1 Теоретичні відомості 19](#_Toc41327694)

[4.2 Практична реалізація 19](#_Toc41327695)

[**РОЗДІЛ 5. КРИТЕРІЙ ПІРСОНА ДЛЯ ПОРІВНЯННЯ ДЕКІЛЬКО ГРУП ЗА РОЗПОДІЛЕННЯМ ОЗНАКИ** 21](#_Toc41327696)

[5.1 Теоретичні відомості 21](#_Toc41327697)

[5.2 Практична реалізація 21](#_Toc41327698)

[**РОЗДІЛ 6. ПАРНИЙ T-КРИТЕРІЙ СТЬЮДЕНТА** 23](#_Toc41327699)

[6.1 Теоретичні відомості 23](#_Toc41327700)

[6.2 Практична реалізація 23](#_Toc41327701)

[**РОЗДІЛ 7. T-КРИТЕРІЙ СТЬЮДЕНТА ДЛЯ НЕЗАЛЕЖНИХ ВИБІРОК** 26](#_Toc41327702)

[7.1 Теоретичні відомості 26](#_Toc41327703)

[7.2 Практична реалізація 26](#_Toc41327704)

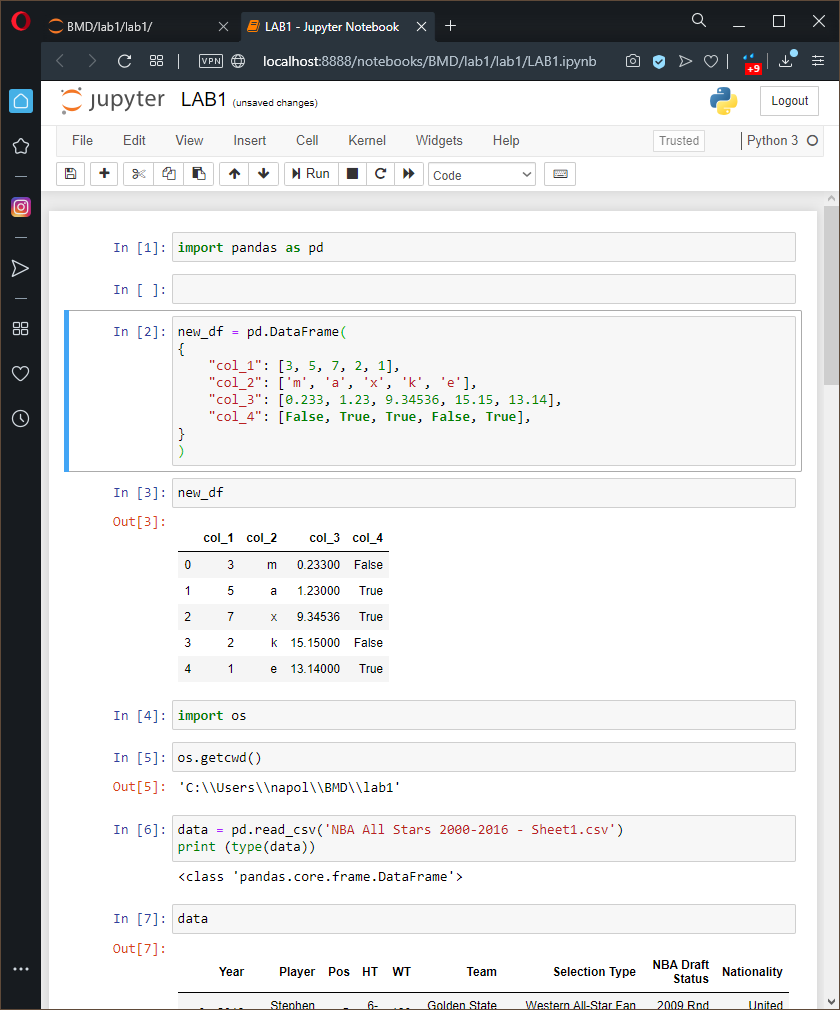
[**РОЗДІЛ 8. ДИСПЕРСІЙНИЙ АНАЛІЗ** 29](#_Toc41327705)

[8.1 Теоретичні відомості 29](#_Toc41327706)

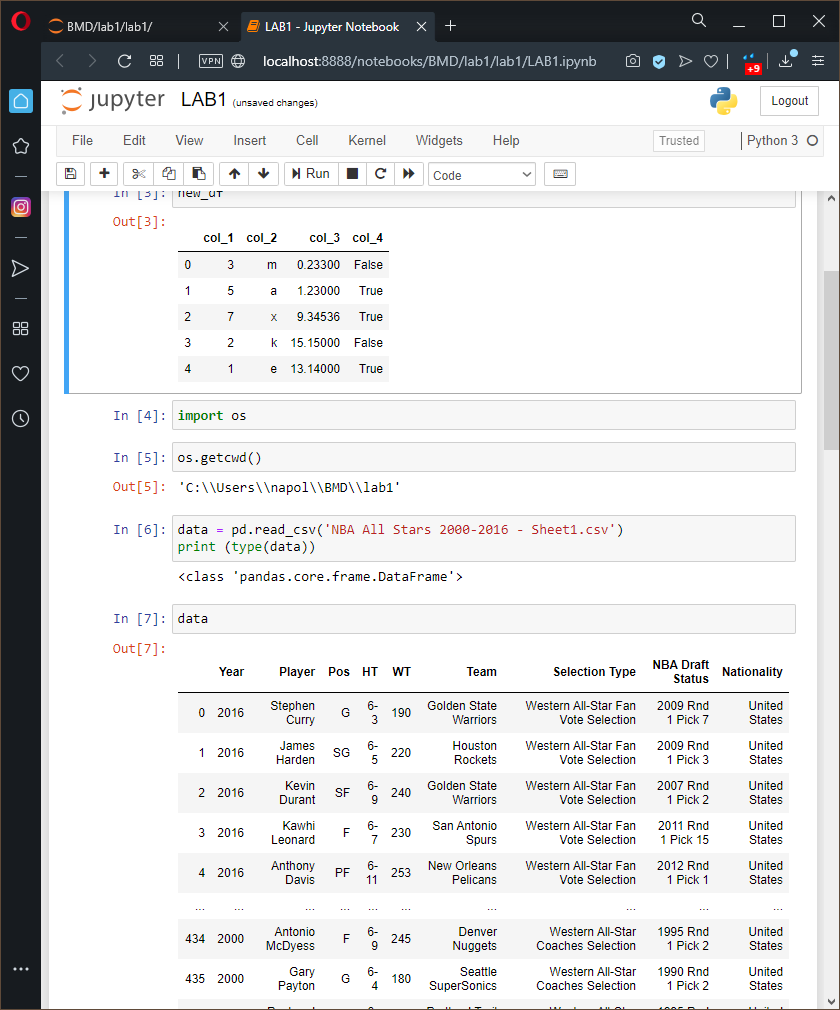
[8.2 Практична реалізація 29](#_Toc41327707)

# **РОЗДІЛ 1. КОМП’ЮТЕРНИЙ ПРАКТИКУМ №1**

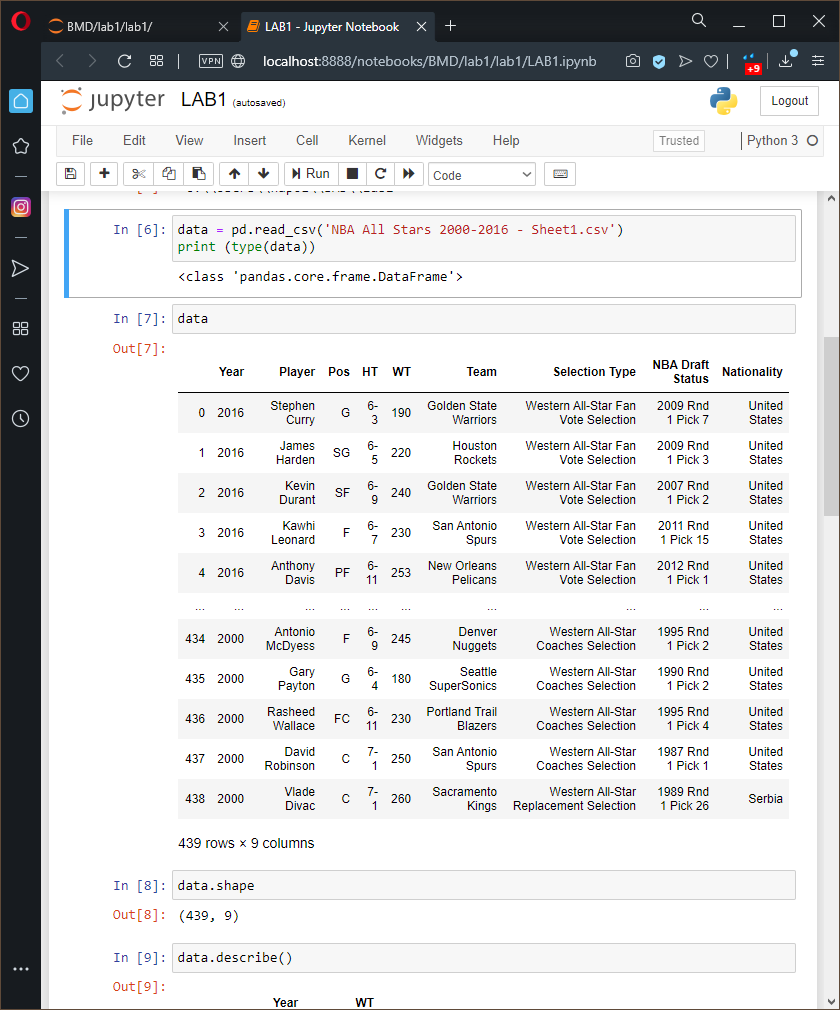
1. Введення даних вручну



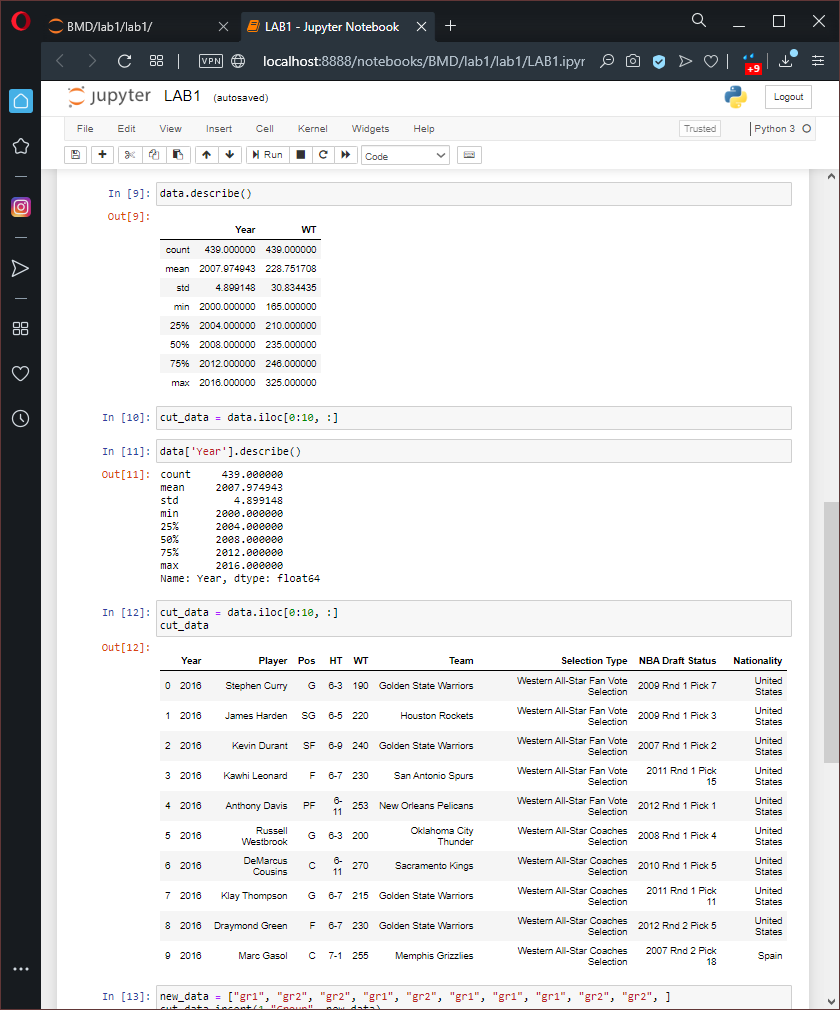
2. Завантаження даних з CSV файлу



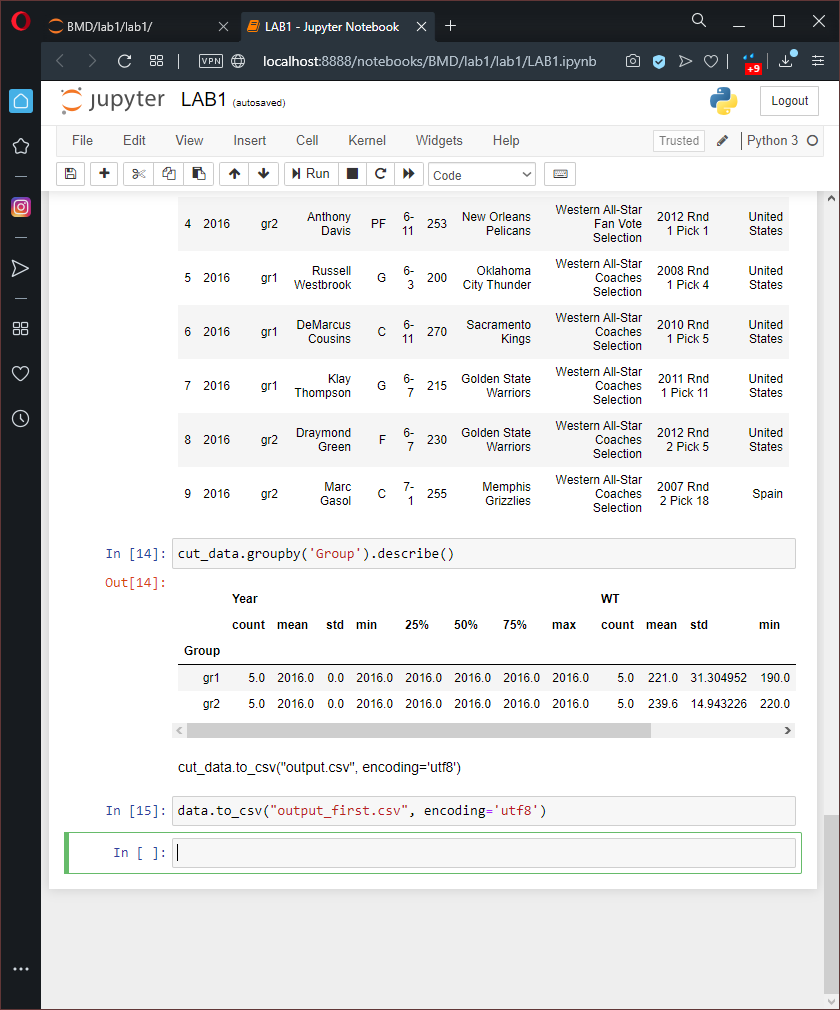
3. Відображення завантажених даних.



4. Функція describe()



5. Статистика з розділенням на групи та збереження даних.



# **РОЗДІЛ 2. ВІЗУАЛІЗАЦІЯ ДАНИХ ЗА ДОПОМОГОЮ БІБЛІОТЕКИ MATPLOTLIB**

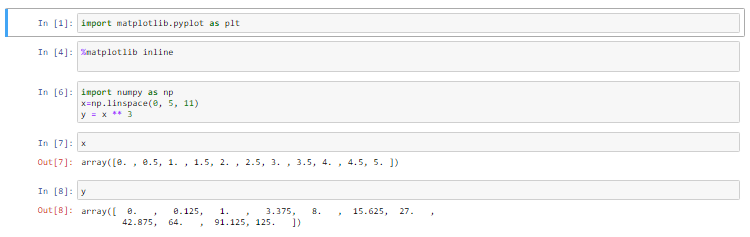
## Теоретичні відомості

Matplotlib — бібліотека на мові програмування Python для візуалізації даних двовимірною 2D графікою (3D графіка також підтримується). Отримувані зображення можуть бути використані як ілюстрації в публікаціях.

Matplotlib написана і підтримується в основному Джоном Хантером і поширюється на умовах BSD-подібної ліцензії. Зображення, які генеруються в різних форматах, можуть бути використані в інтерактивній графіці, наукових публікаціях, графічному інтерфейсі користувача, веб-додатках, де потрібно будувати діаграми (англ. plotting). В документації автор зізнається, що Matplotlib починався з імітування графічних команд MATLAB, але є незалежним від нього проектом.

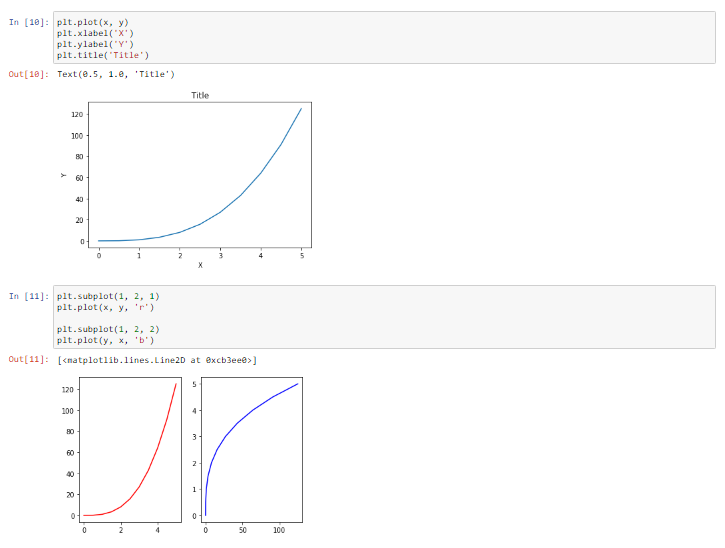
## Практична реалізація

1. Основи побудови графіків

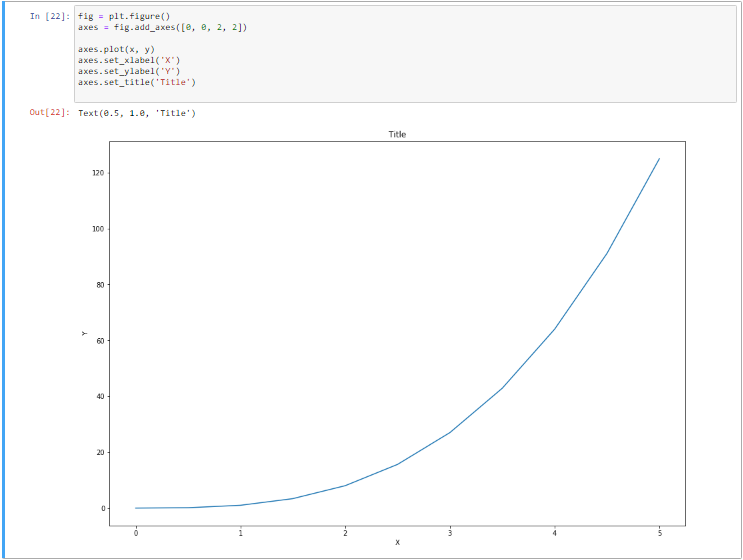


1.1 Побудова графіків:

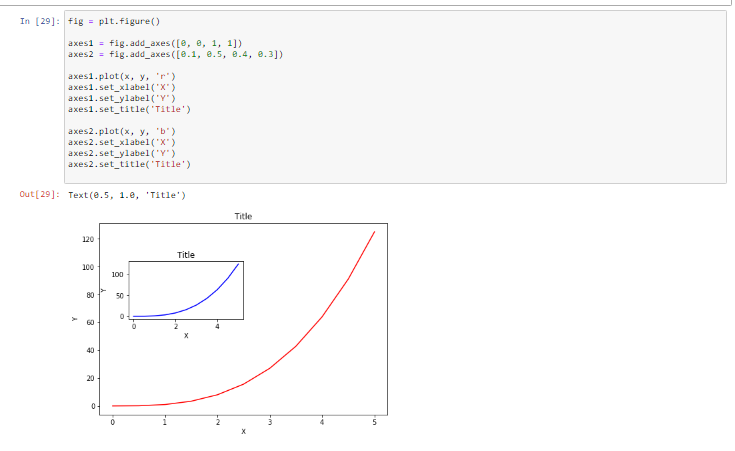
1.1.1 Функціональний метод



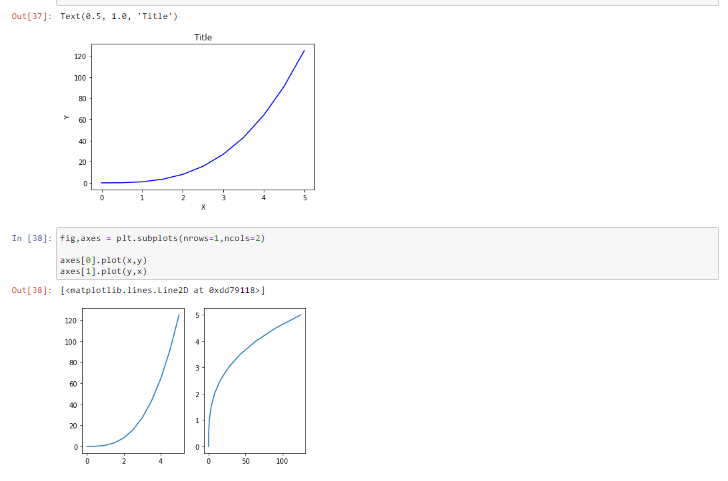
1.1.2. Об’єктно-орієнтований метод

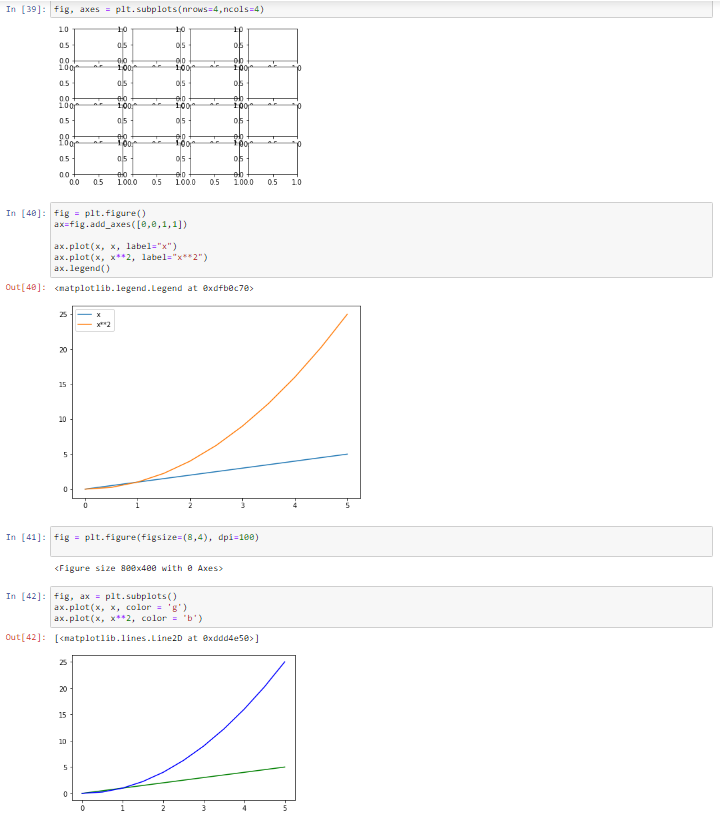


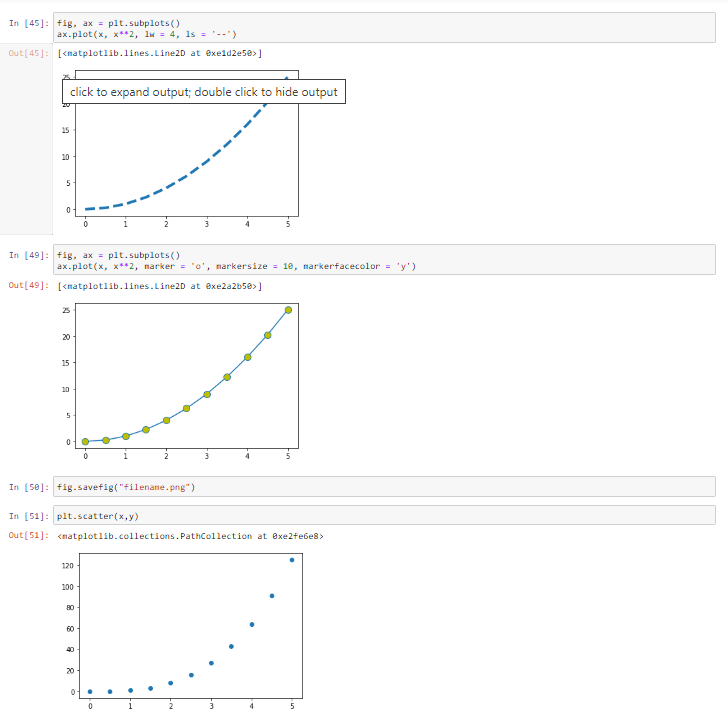
2.Одночасна побудова декількох графіків.



3. Використання іншого функціоналу бібліотеки.







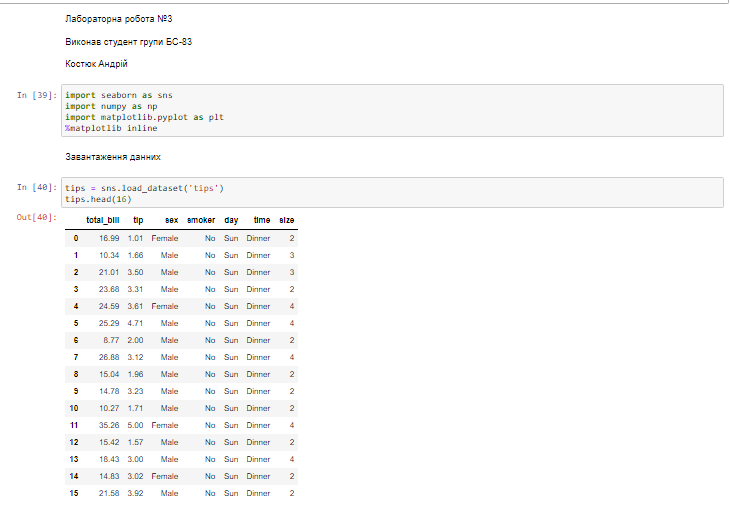


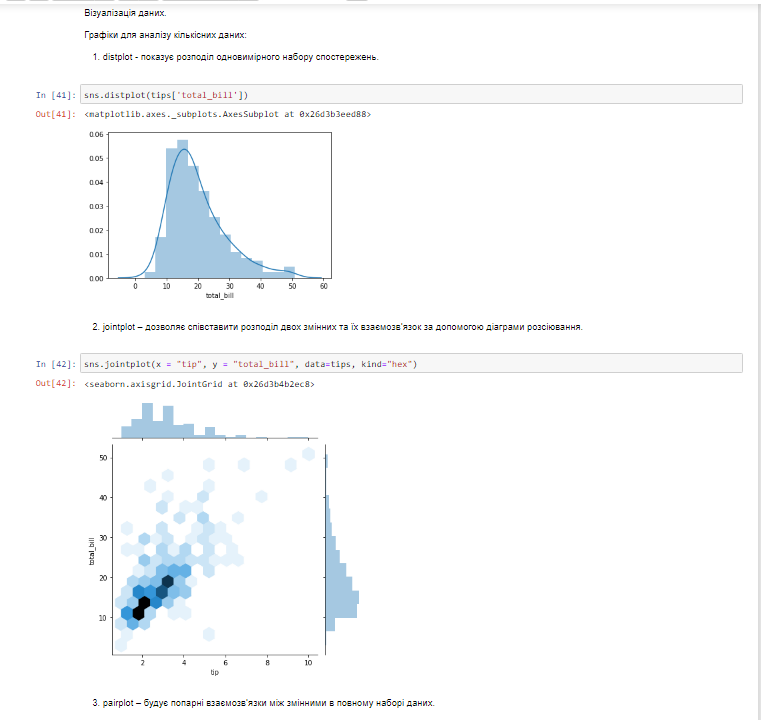
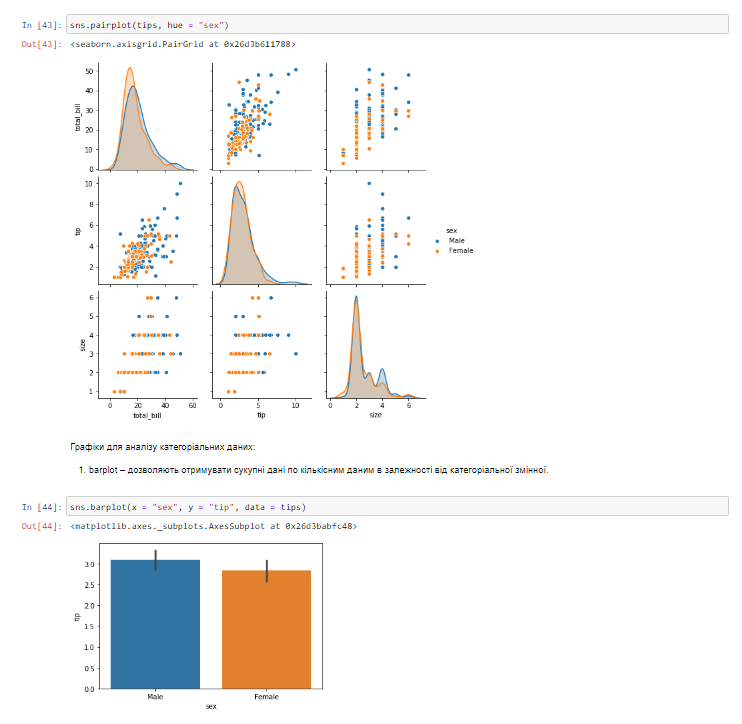
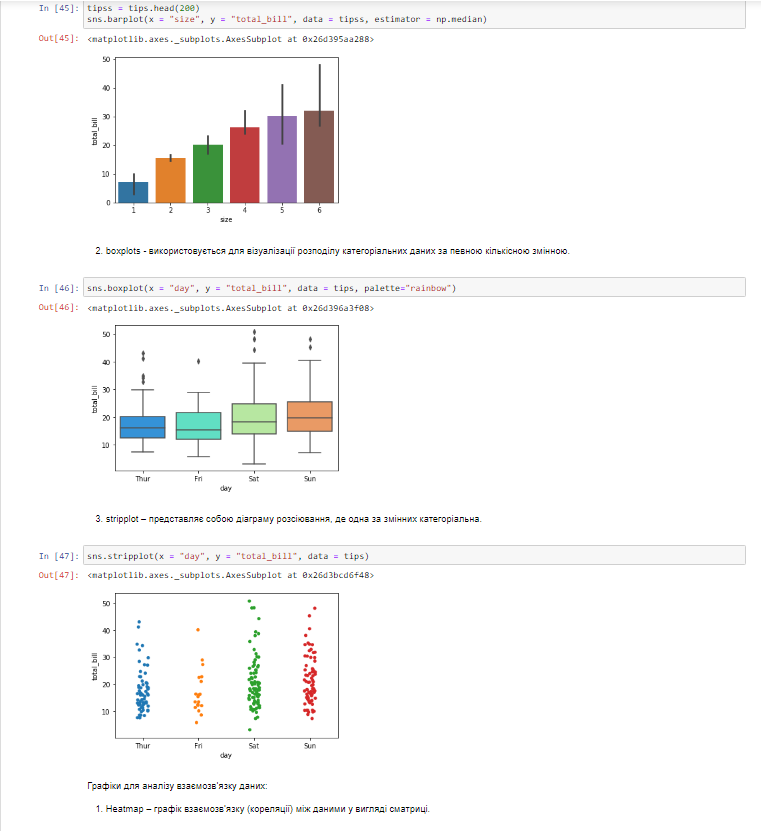
# **РОЗДІЛ 3. ВІЗУАЛІЗАЦІЯ ДАНИХ ЗА ДОПОМОГОЮ БІБЛІОТЕКИ SEABORN**

## Теоретичні відомості

Seaborn - це більш високорівневе API на базі бібліотеки matplotlib. Seaborn містить більш адекватні дефолтні Налаштування оформлення графіків. Якщо просто додати в код import seaborn, то картинки стануть набагато симпатичніше. Також в бібліотеці є досить складні типи візуалізації, які в matplotlib зажадали б великого кількість коду.

## Практична реалізація



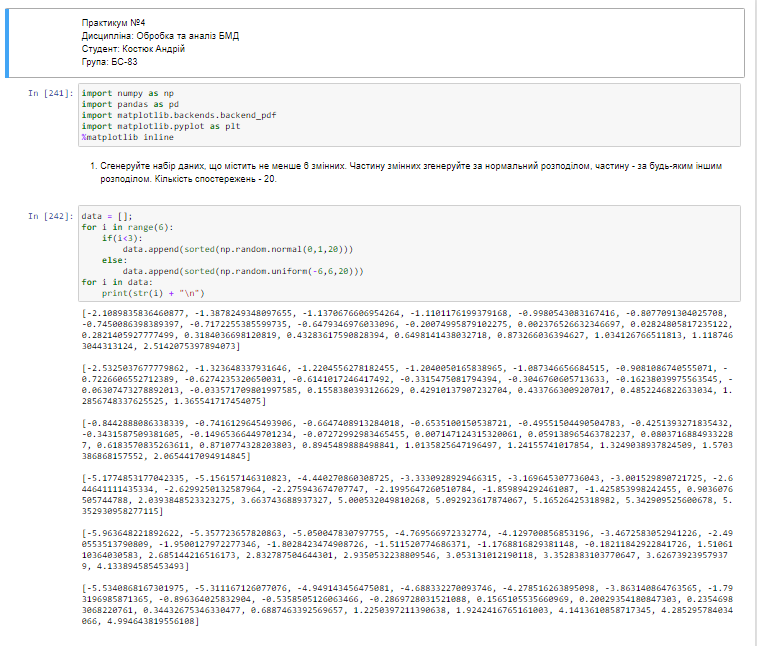
# **РОЗДІЛ 4. ПЕРЕВІРКА ДАНИХ НА НОРМАЛЬНІСТЬ ЗА ДОПОМОГОЮ КРИТЕРІЮ ШАПІРА-УІЛКА**

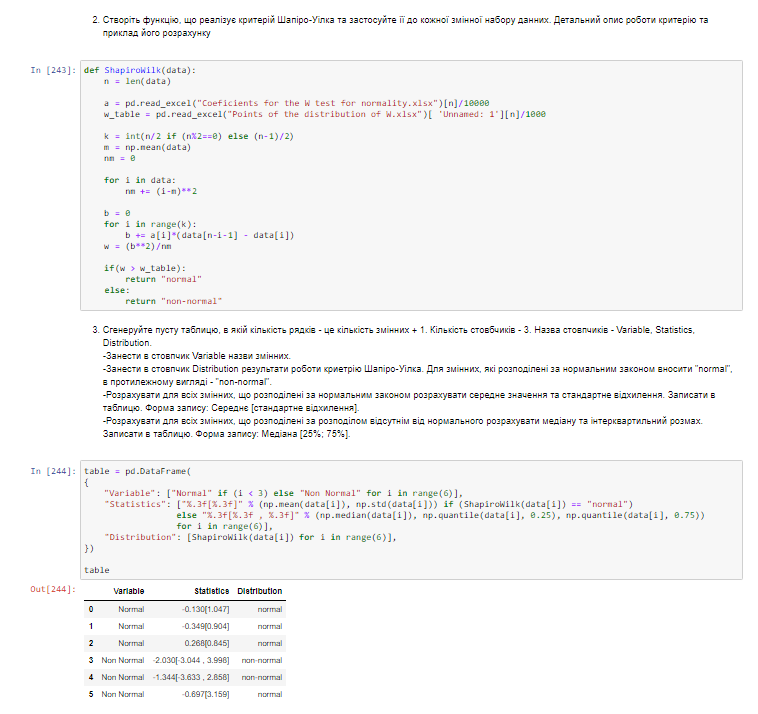
## Теоретичні відомості

Критерій Шапіро-Уілкі використовується для визначення нормальності.

Критерій Шапіро-Уілкі є найбільш ефективним, так як він має більшу потужність в порівнянні з альтернативними критеріями перевірки нормальності.

## Практична реалізація







# **РОЗДІЛ 5. КРИТЕРІЙ ПІРСОНА ДЛЯ ПОРІВНЯННЯ ДЕКІЛЬКО ГРУП ЗА РОЗПОДІЛЕННЯМ ОЗНАКИ**

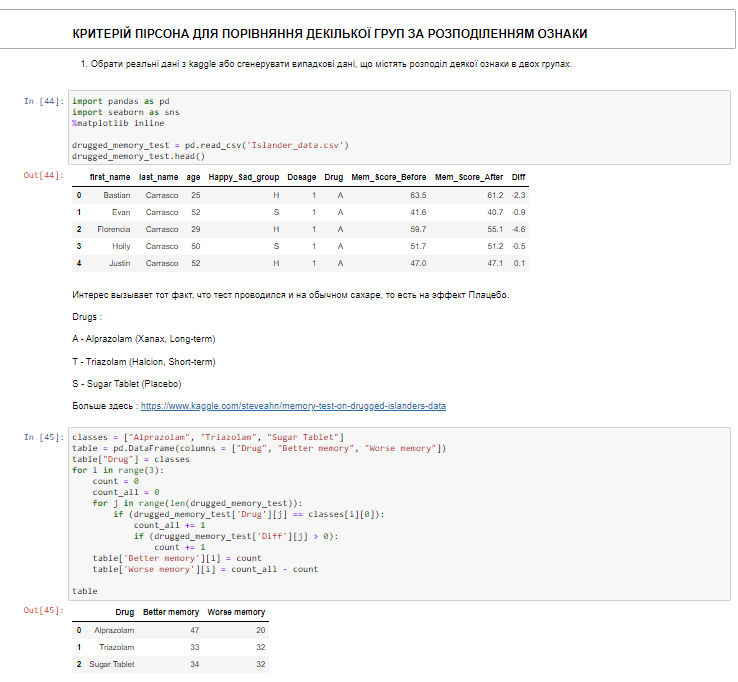
## Теоретичні відомості

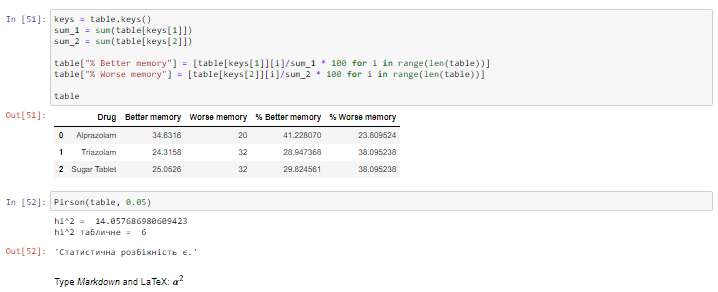
Критерій узгодженості Пірсона — один з найвідоміших критеріїв , тому його часто і називають просто «критерій хі-квадрат».

Використовується для перевірки гіпотези про закон розподілу.

Ґрунтується на групованих даних. Область значень передбачуваного розподілу ділять на деяке число інтервалів. Після чого будують функцію відхилення ρ по різницях теоретичних імовірностей потрапляння в інтервали групування й емпіричних частот.

## Практична реалізація



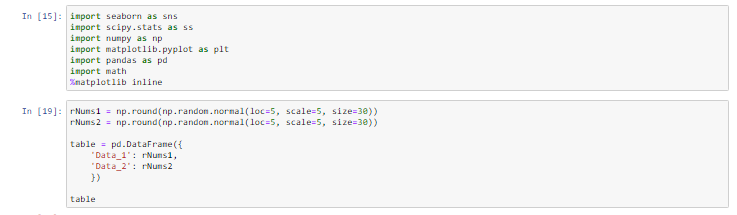
# **РОЗДІЛ 6. ПАРНИЙ T-КРИТЕРІЙ СТЬЮДЕНТА**

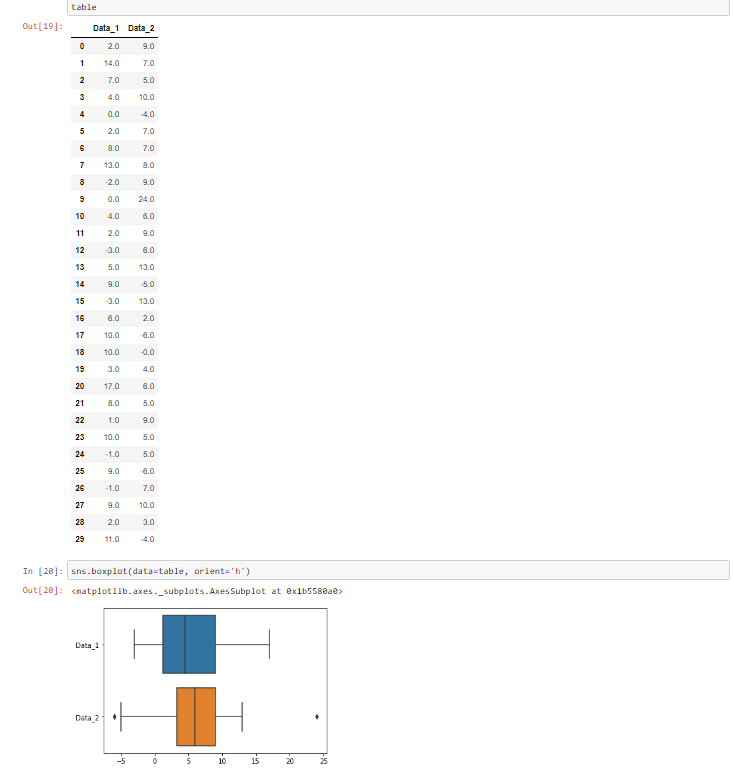
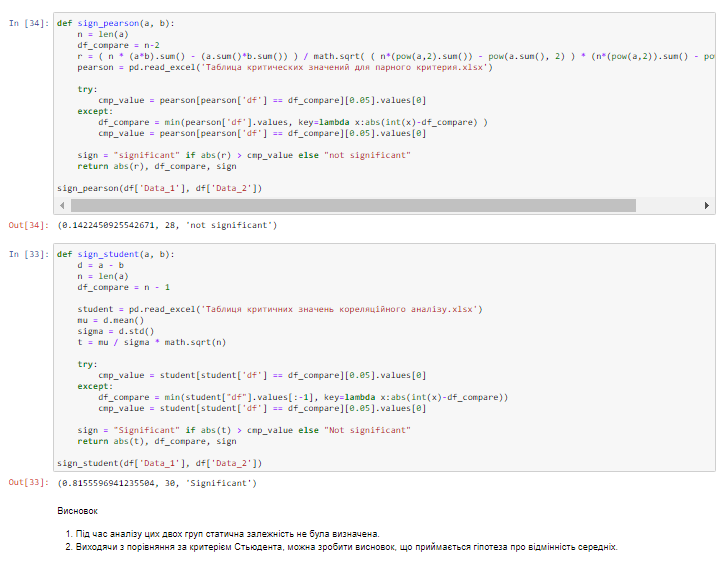
## Теоретичні відомості

t-критерій Стьюдента - загальна назва для класу методів статистичної перевірки гіпотез (статистичних критеріїв), заснованих на розподілі Стьюдента. Найбільш часті випадки застосування t-критерію повязані з перевіркою рівності середніх значень у двох вибірках.

t-статистика будується зазвичай за наступним загальним принципом: в чисельнику випадкова величина з нульовим математичним очікуванням (при виконанні нульової гіпотези), а в знаменнику - вибіркове стандартне відхилення цієї випадкової величини, що отримується як квадратний корінь з незміщеної оцінки дисперсії.

## Практична реалізація



# **РОЗДІЛ 7. T-КРИТЕРІЙ СТЬЮДЕНТА ДЛЯ НЕЗАЛЕЖНИХ ВИБІРОК**

## Теоретичні відомості

tt-критерій Стьюдента для незалежних вибірок застосовується для порівняння середніх значень двох незалежних між собою вибірок.

Умови застосування:

1. Порівнянні значення не становлять пару корелюють значень
2. Розподіл ознак в кожній вибірці відповідає нормальному розподілу
3. Дисперсії ознаки в вибірках приблизно рівні (перевіряється за допомогою критерію F-Фішера)

Альтернатива:

* непараметрический критерій U Манна-Уїтні (якщо розподіл ознак не відповідає нормлаьно)
* t-критерій Стьюдента для залежних вибірок (якщо значення становлять пару корелюють значень).

## Практична реалізація





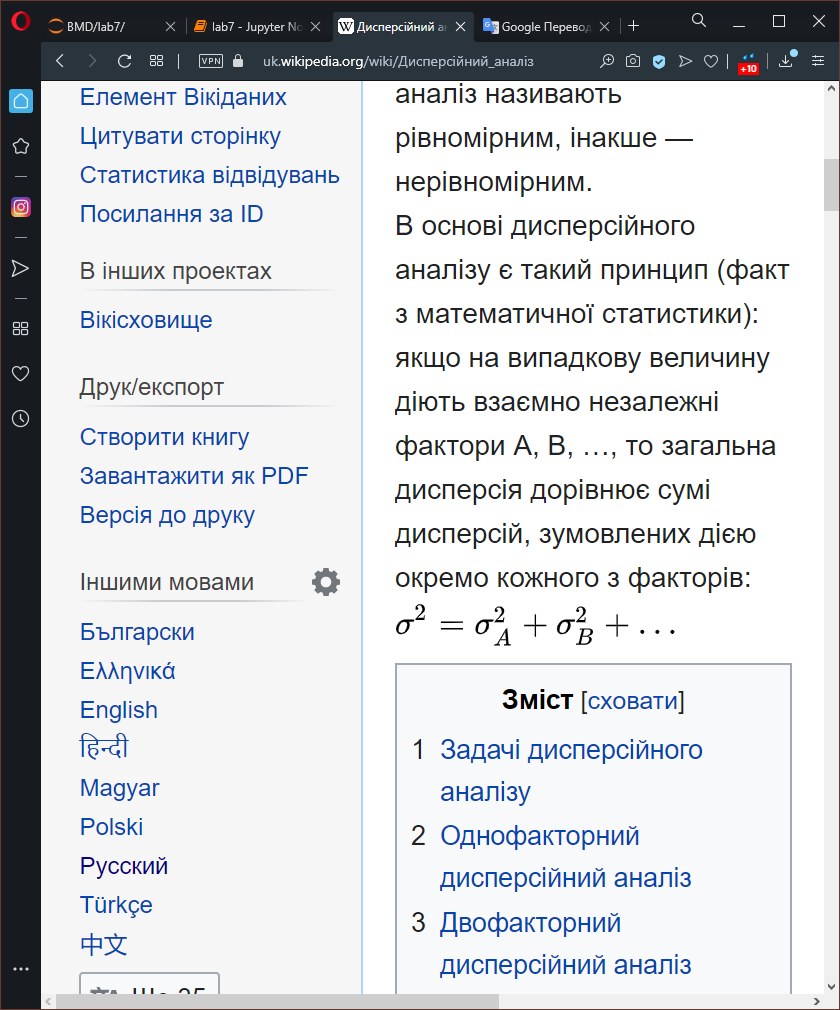
# **РОЗДІЛ 8. ДИСПЕРСІЙНИЙ АНАЛІЗ**

## Теоретичні відомості

Дисперсійний аналіз (англ. analysis of variance (ANOVA)) являє собою статистичний метод аналізу результатів, які залежать від якісних ознак.

Кожен фактор може бути дискретною чи неперервною випадковою змінною, яку розділяють на декілька сталих рівнів (градацій, інтервалів). Якщо кількість вимірювань (проб, даних) на всіх рівнях кожного з факторів однакова, то дисперсійний аналіз називають рівномірним, інакше — нерівномірним.

В основі дисперсійного аналізу є такий принцип (факт з математичної статистики): якщо на випадкову величину діють взаємно незалежні фактори A, B, …, то загальна дисперсія дорівнює сумі дисперсій, зумовлених дією окремо кожного з факторів:



## Практична реалізація



