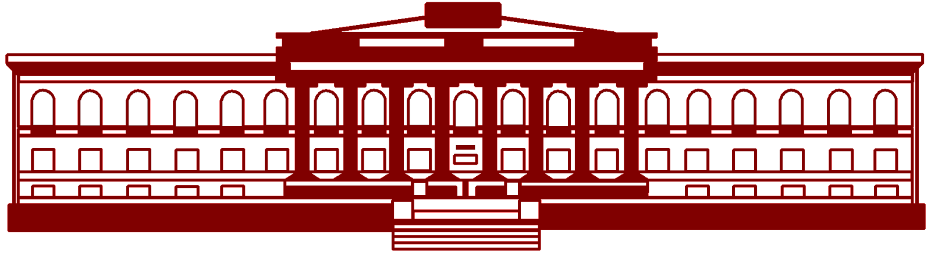
**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**

**імені ТАРАСА ШЕВЧЕНКА**

****

**ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

**Кафедра прикладних інформаційних систем**

**Звіт до лабораторної роботи №4**

**з курсу**

**«Інтелектуальний аналіз даних»**

*Студента 3 курсу*

*групи ПП-32*

*спеціальності 122 «Комп'ютерні науки»*

*ОП «Прикладне програмування»*

Чалого Єгора Олеговича

*Викладач:*

*Білий Р.О.*

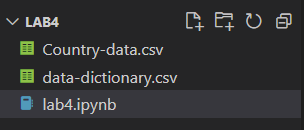
**Київ – 2023**

**Тема:** Кластеризація. Метод k-mean, пошук оптимальних параметрів.

**Мета роботи:** Отримання практичних навичок з виконання кластеризації даних, використовуючи метод k-mean та інші з пакету sklearn.

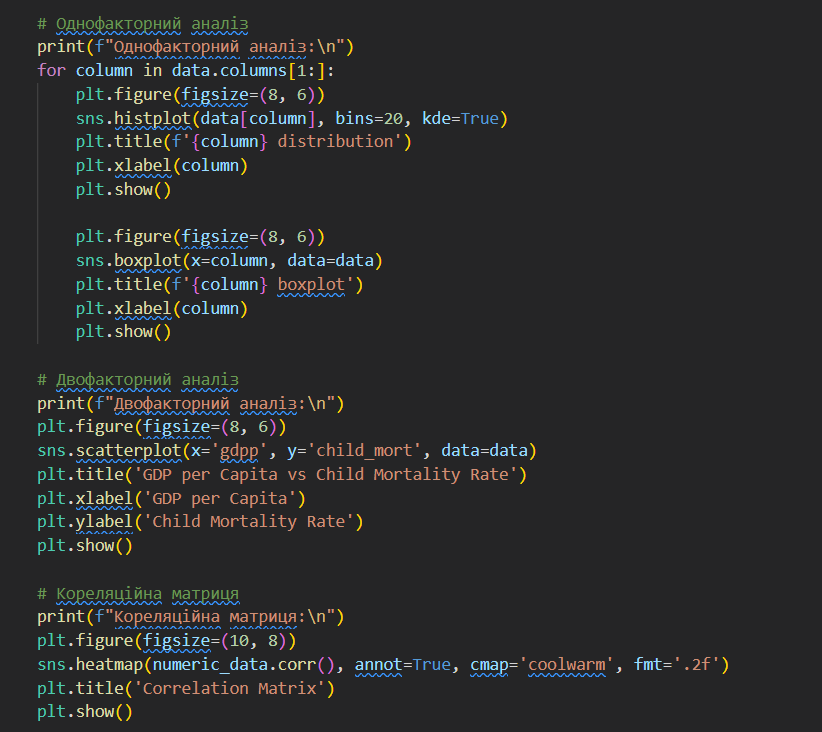
**Хід роботи:**

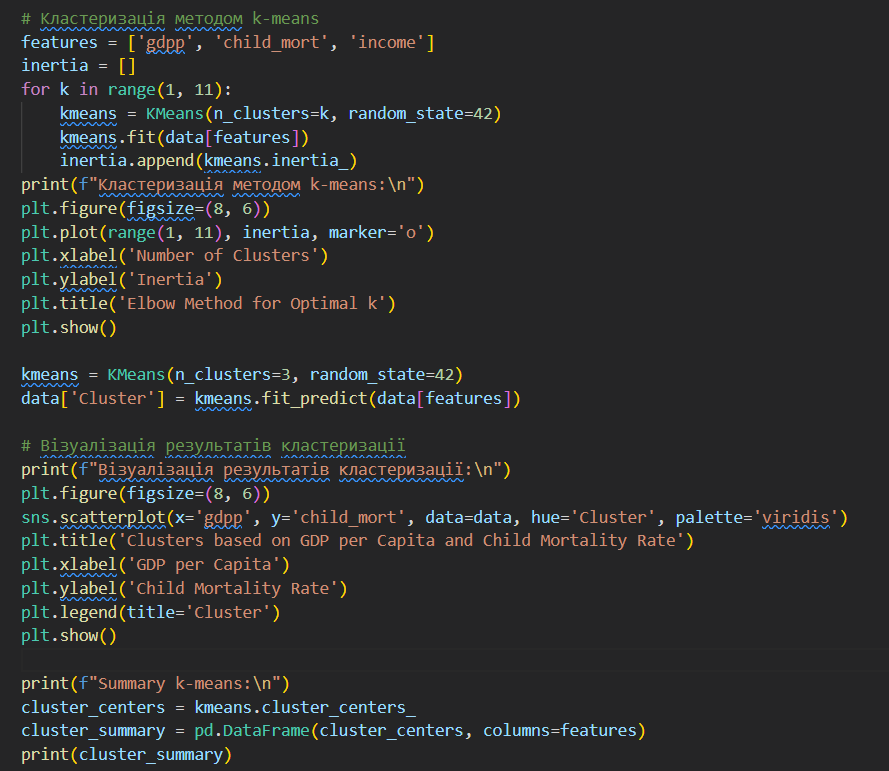
1. Після вивчення наданих прикладів завантажив файли з даними у папку проекту:



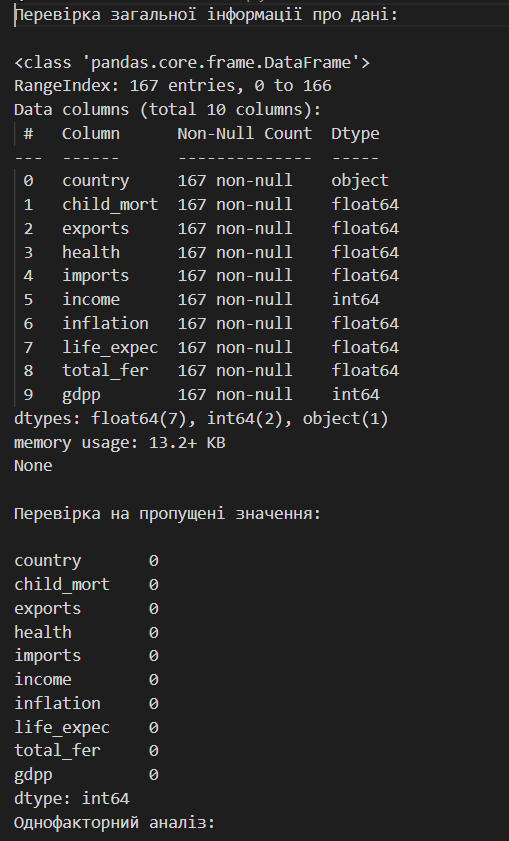
2. Почав із необхідної перевірки даних, очищення даних та EDA, зробив однофакторний аналіз, двофакторний аналіз. Візуалізував закономірності:

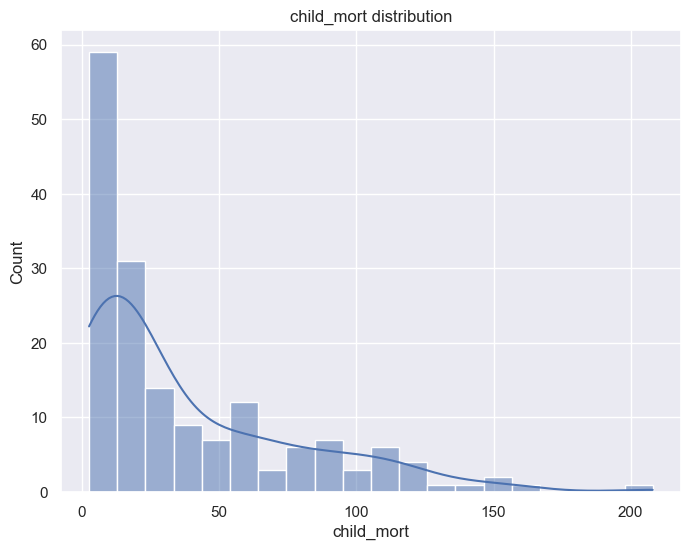


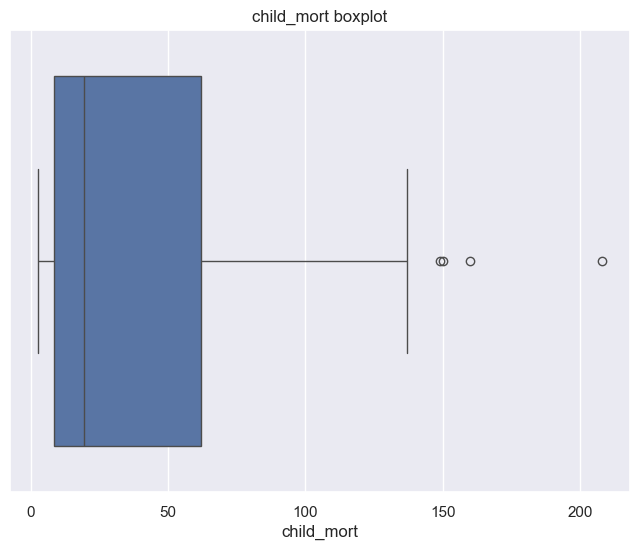


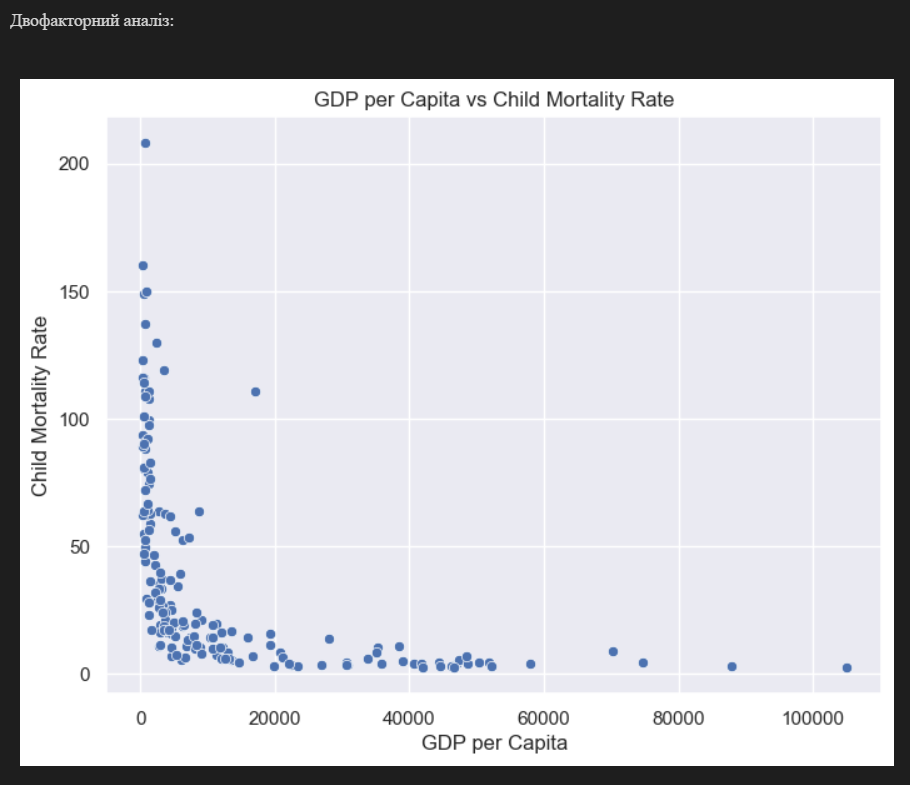


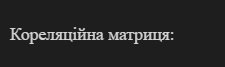
Результат:

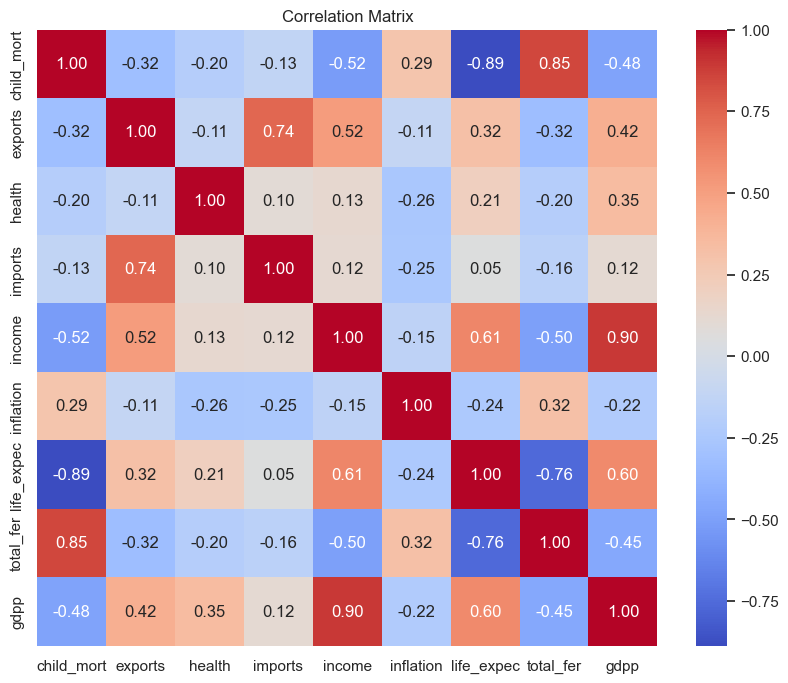


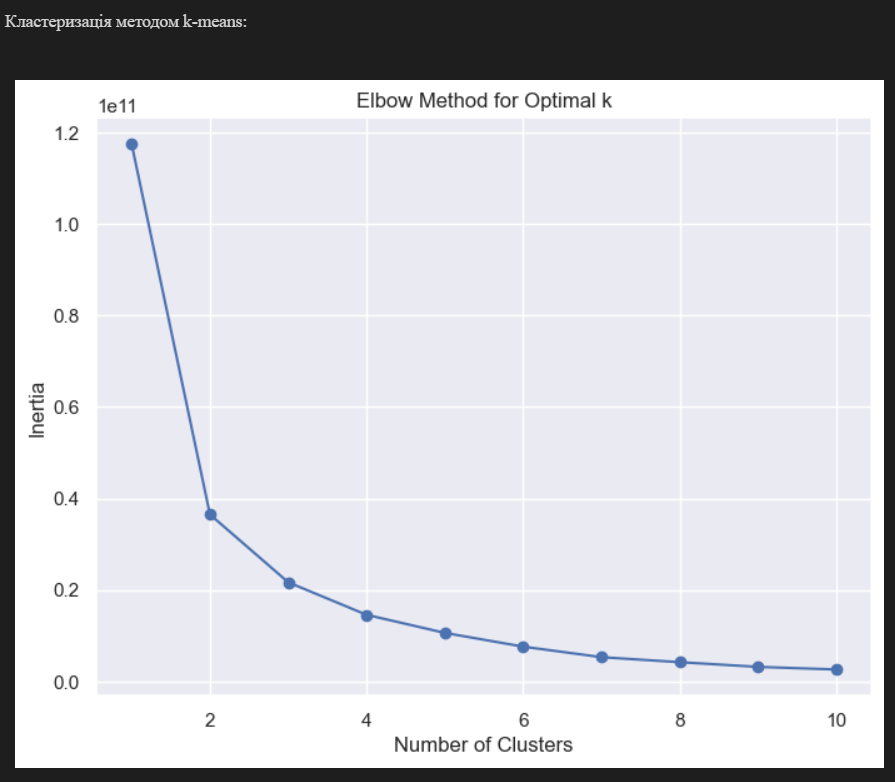


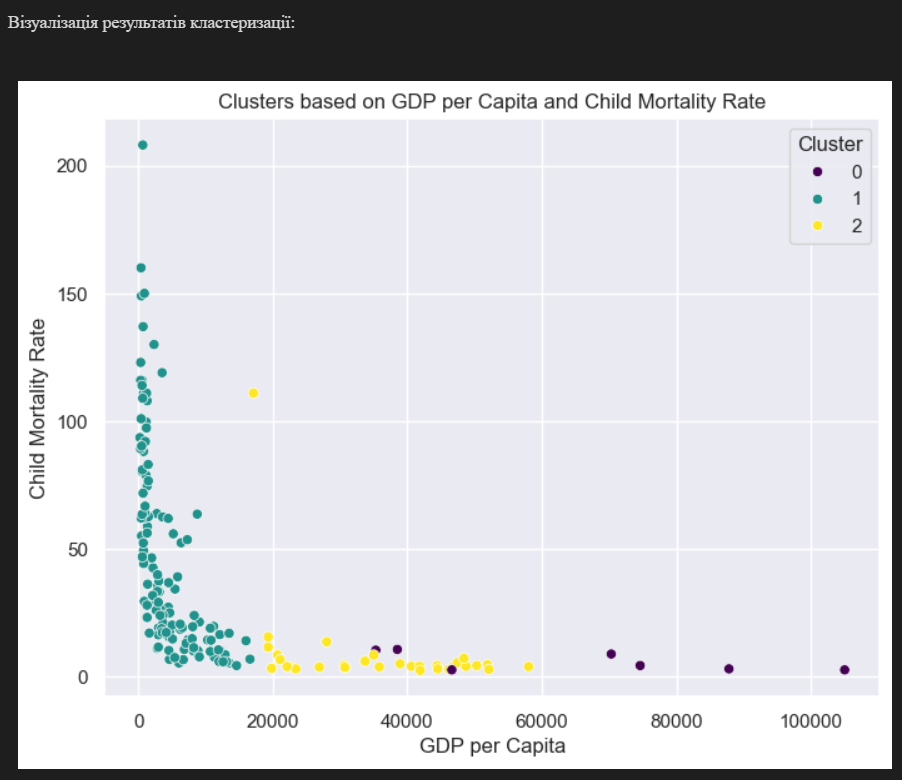


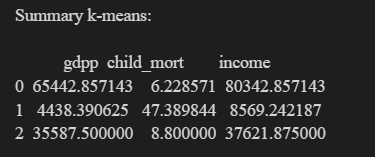








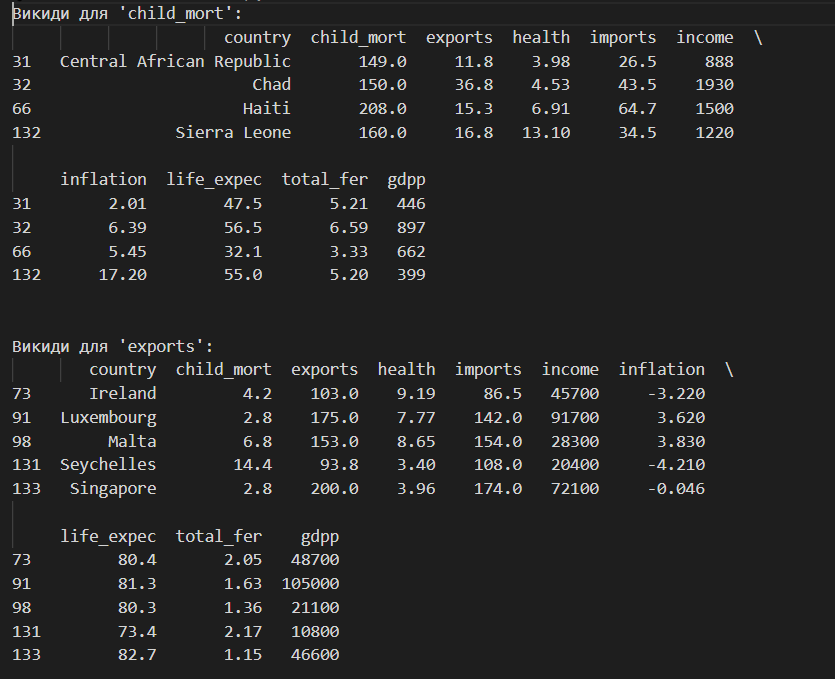


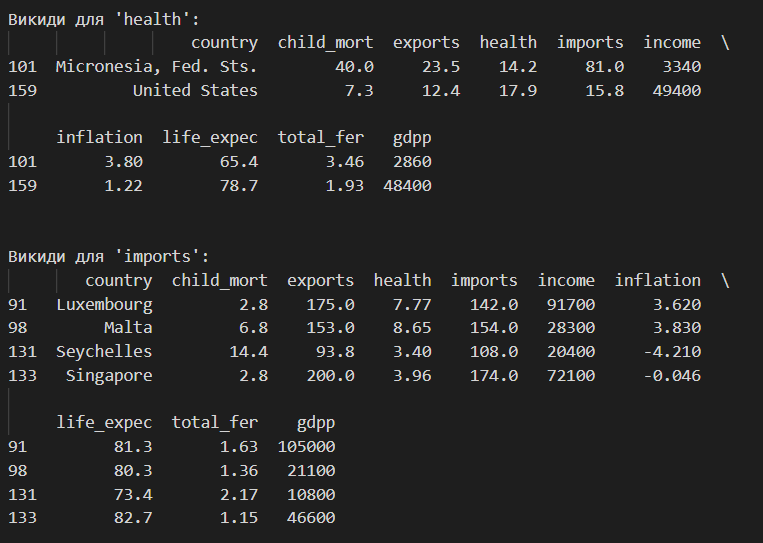


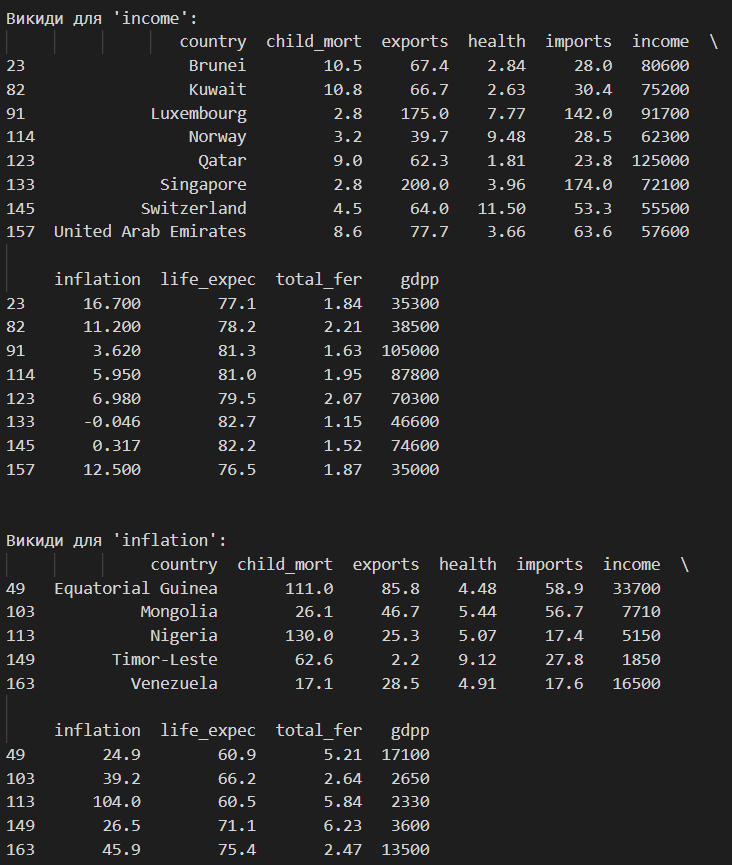
3. Виконав аналіз викидів для набору даних і вивів результати:



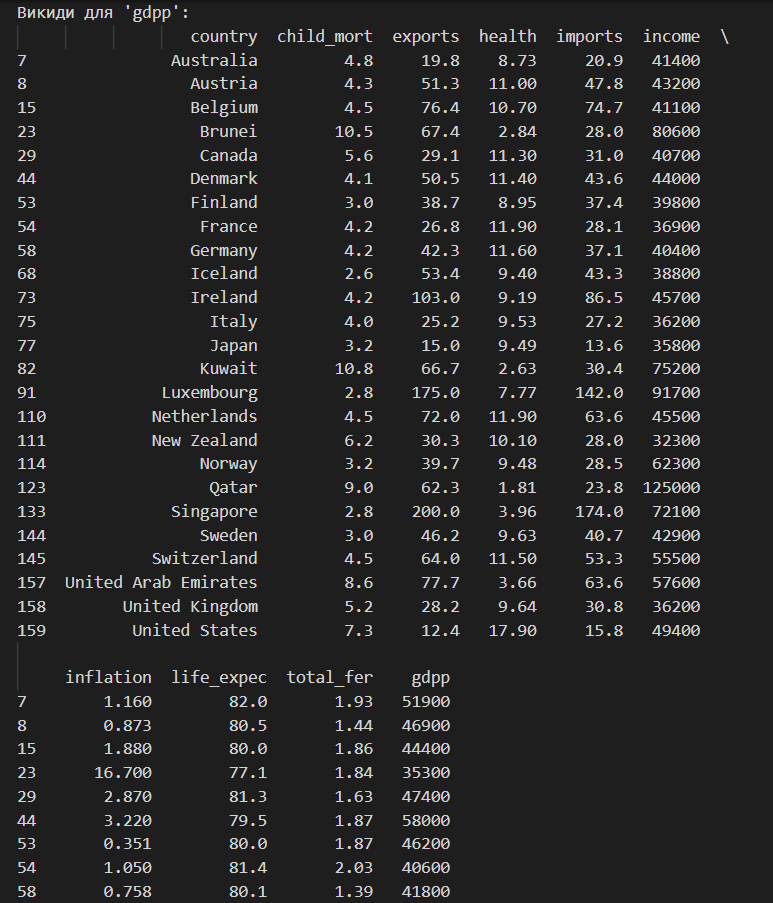
Результат:



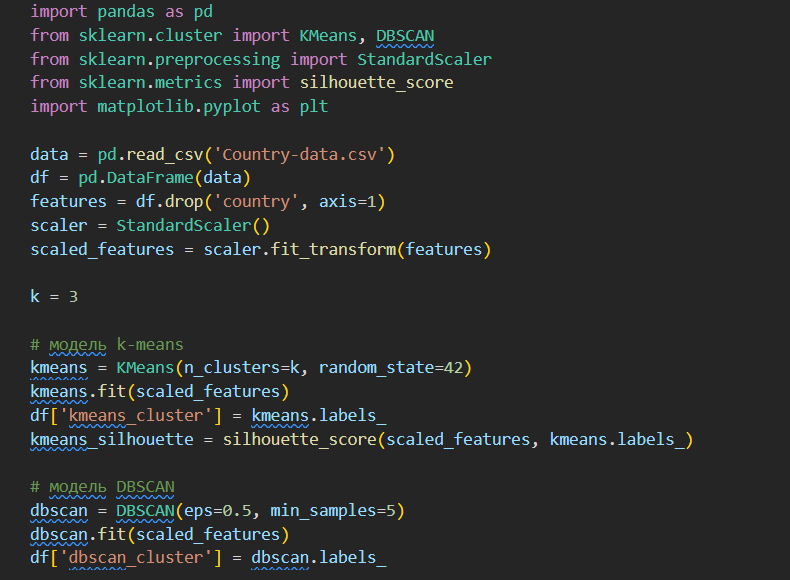


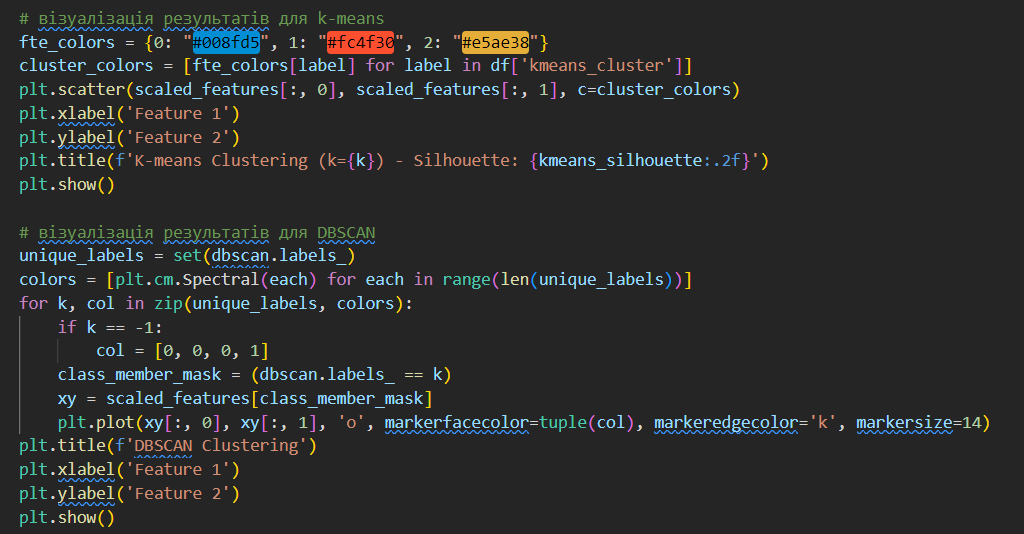




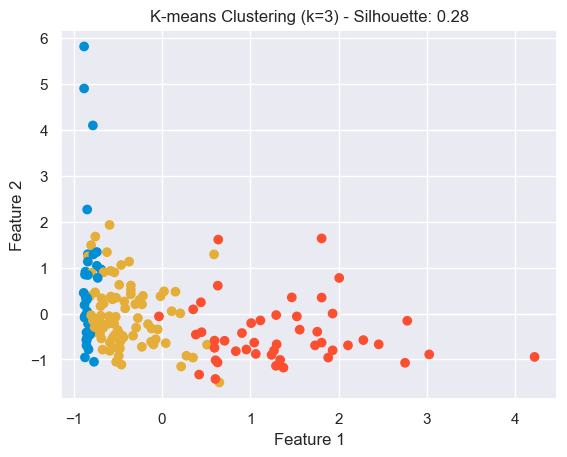


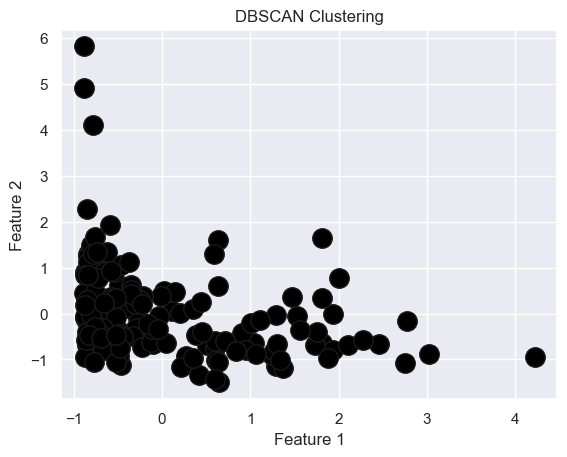
4. Провів кластеризацію, використовуючи k-mean методологію з прикладів та документації:





Результат:

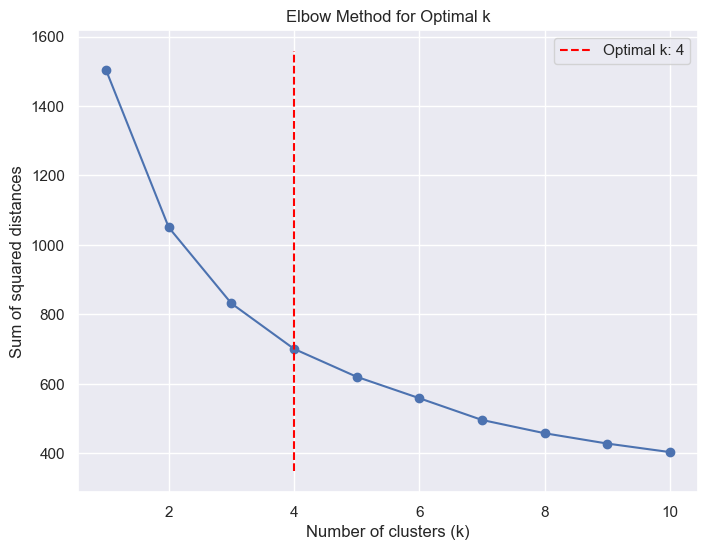




5. Оптимізував k, використовуючи kneed:

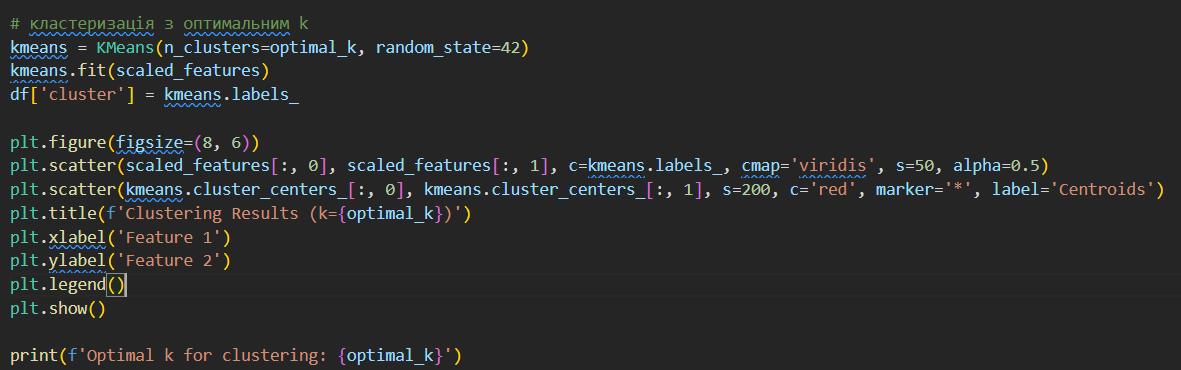


Результат:

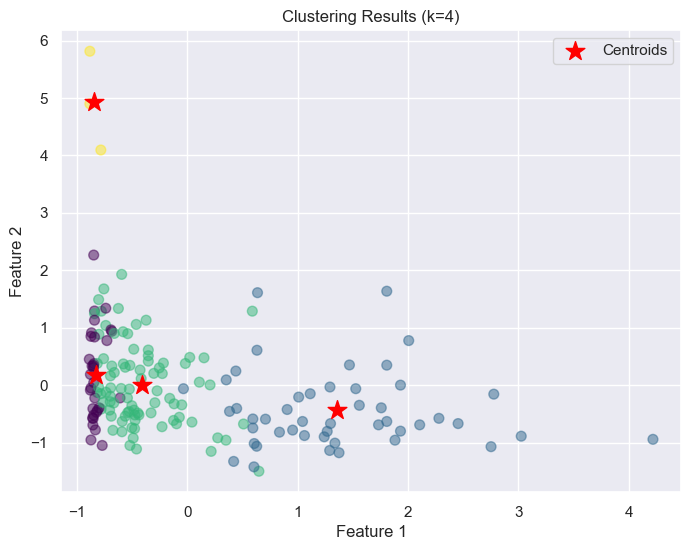




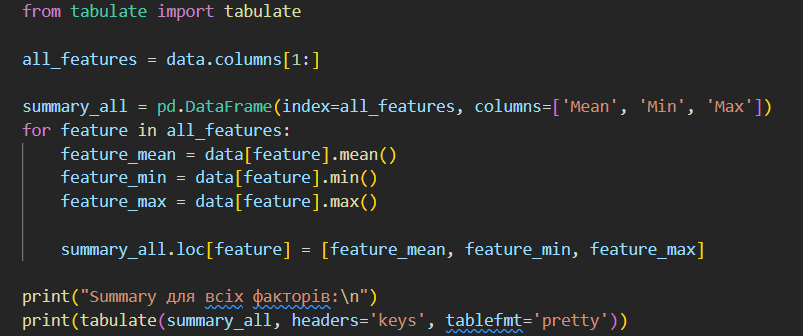
6. Візуалізував результати кластеризації та оптимізації гіпер-параметрів:



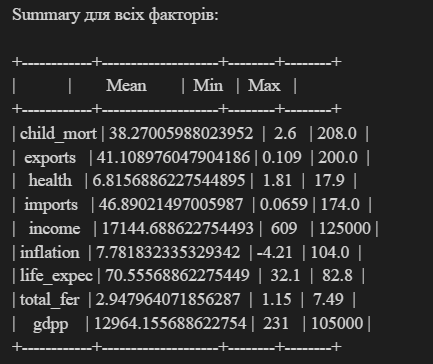
Результат:



7. Зробив summary для всіх факторів:



Результат:



**Висновок:**

В ході виконання лабораторної роботи отримав практичні навички з виконання кластеризації даних, використовуючи метод k-mean та інші з пакету sklearn.