

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
імені ТАРАСА ШЕВЧЕНКА**



**ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

**Кафедра прикладних інформаційних систем**

**Звіт до практичної роботи №4**

**з курсу**

**«Інтелектуальний аналіз даних»**

*Студентки 3 курсу групи ПП-33  
спеціальності 122 «Комп'ютерні науки»  
ОП «Прикладне програмування»  
Матвіїв Анастасії Юріївни*

*Викладач:  
Білий Р.О.*

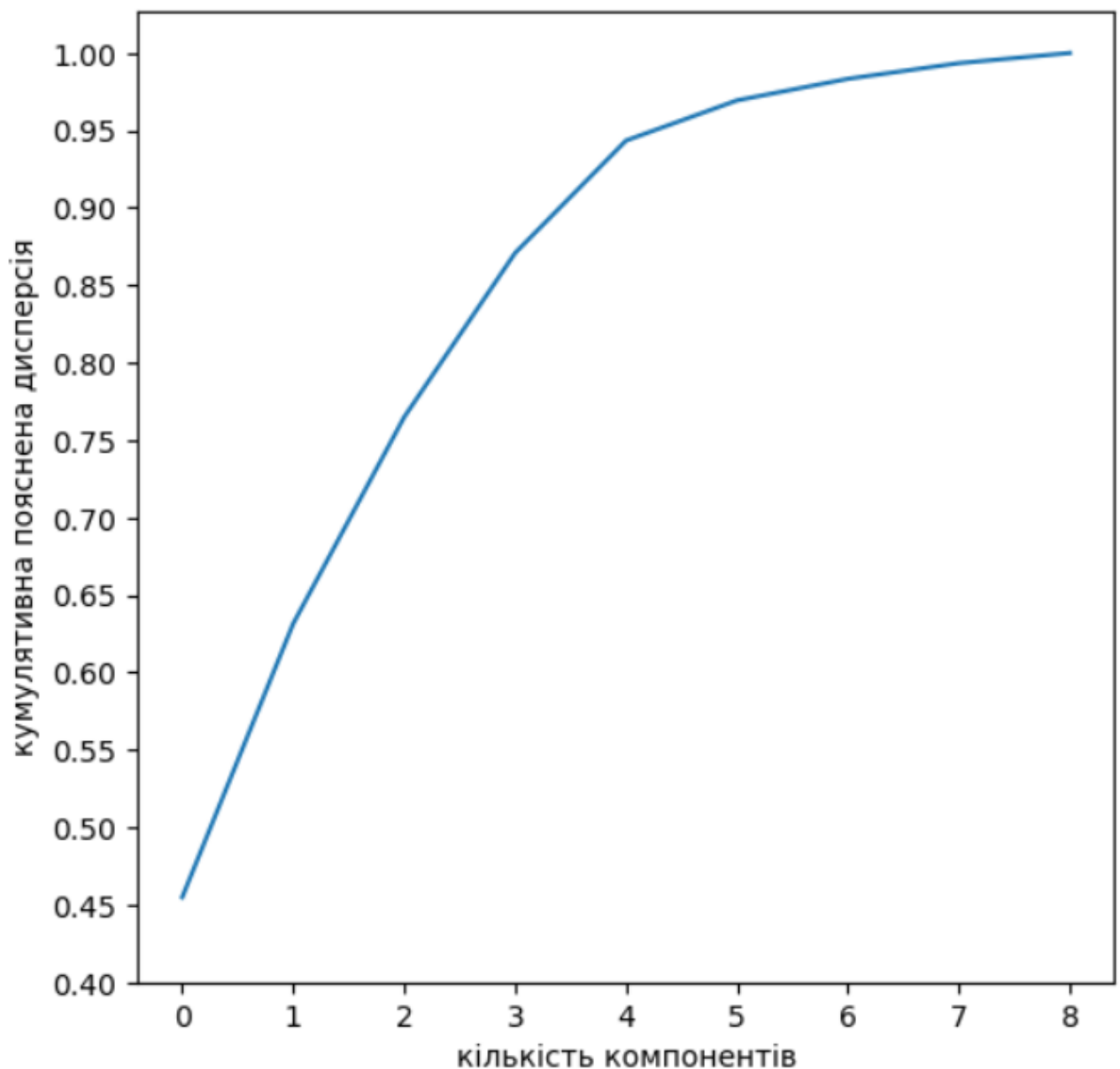
**Київ – 2023**

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

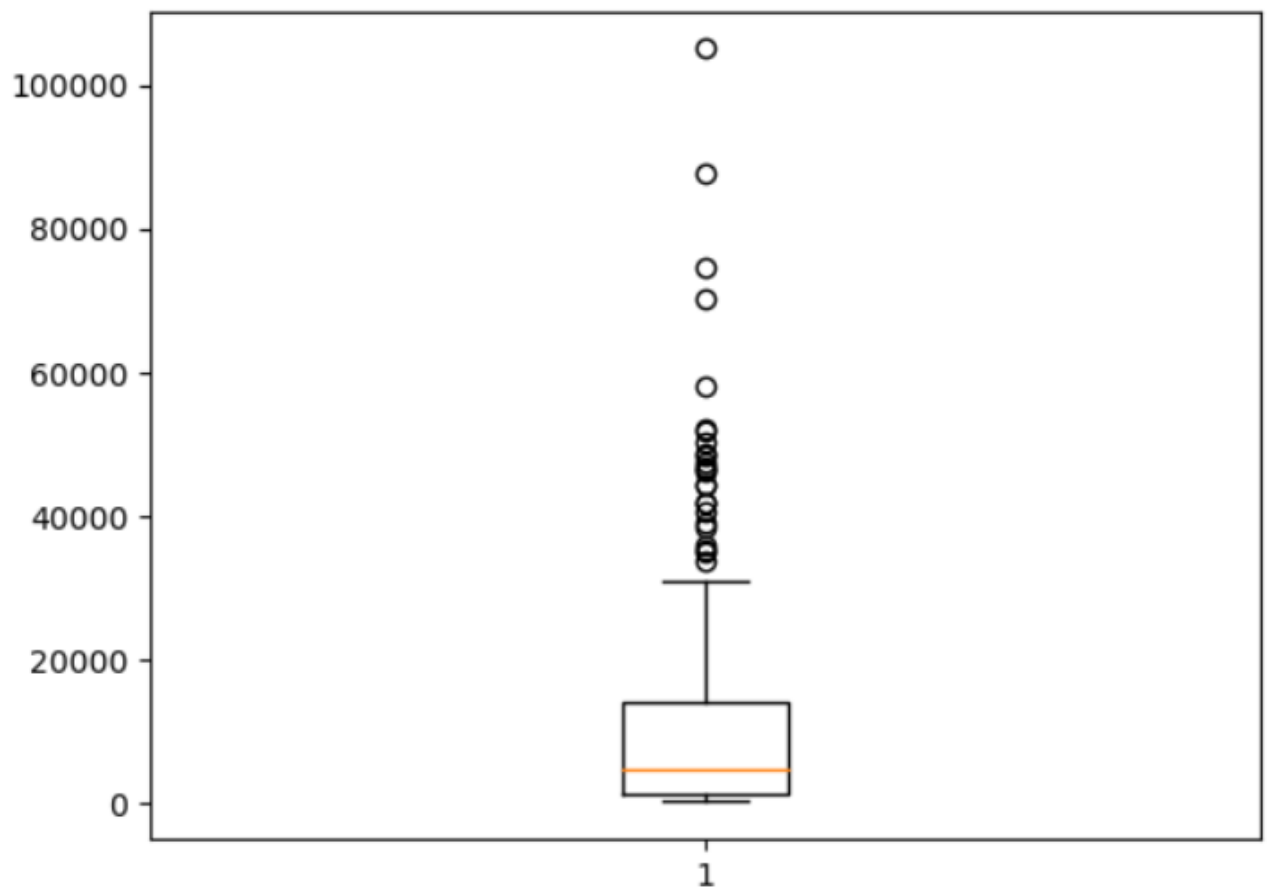
**Мета:** Метою лабораторної роботи є отримання практичних навичок з виконання кластеризації даних, використовуючи метод k-mean та інші з пакету sklearn.

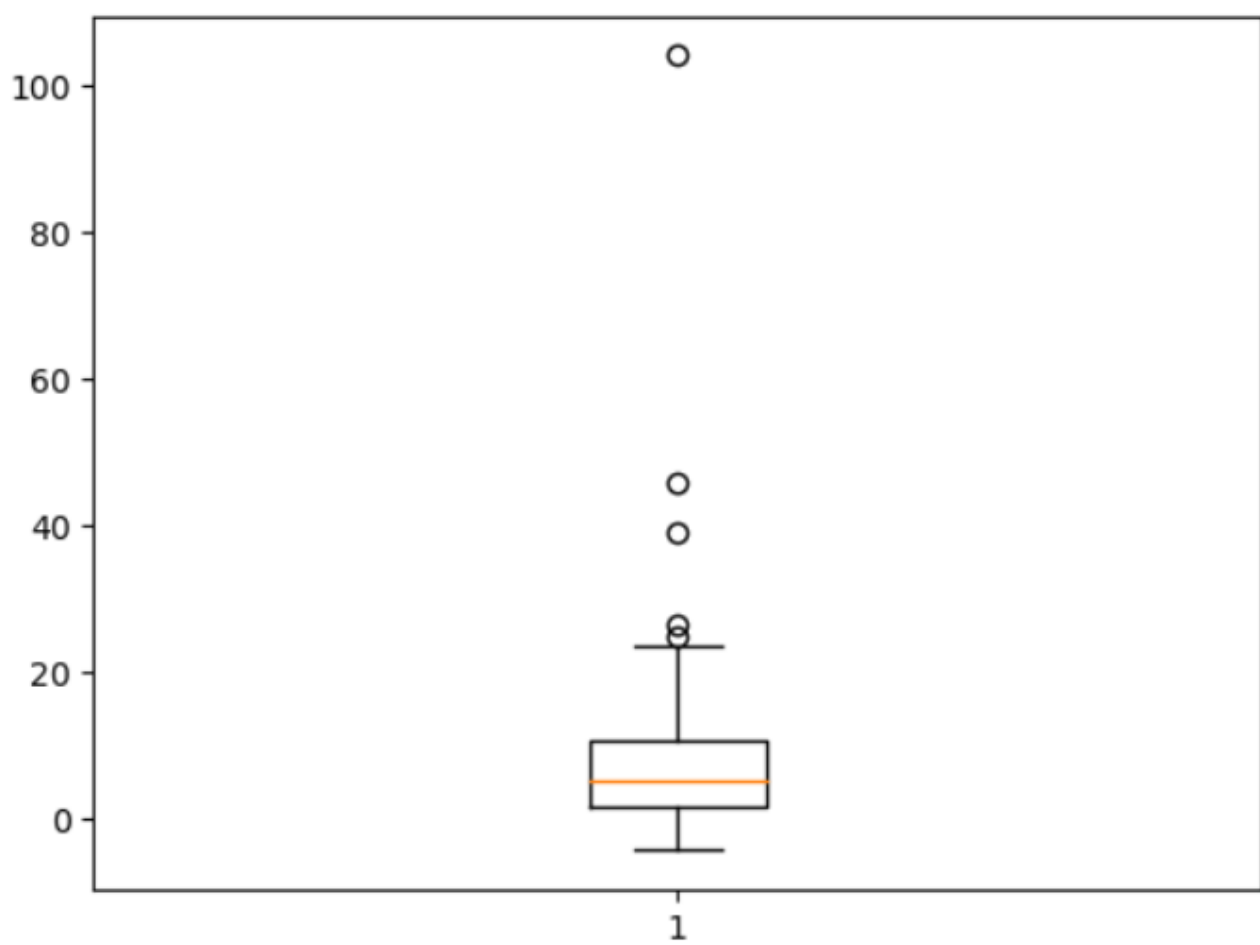
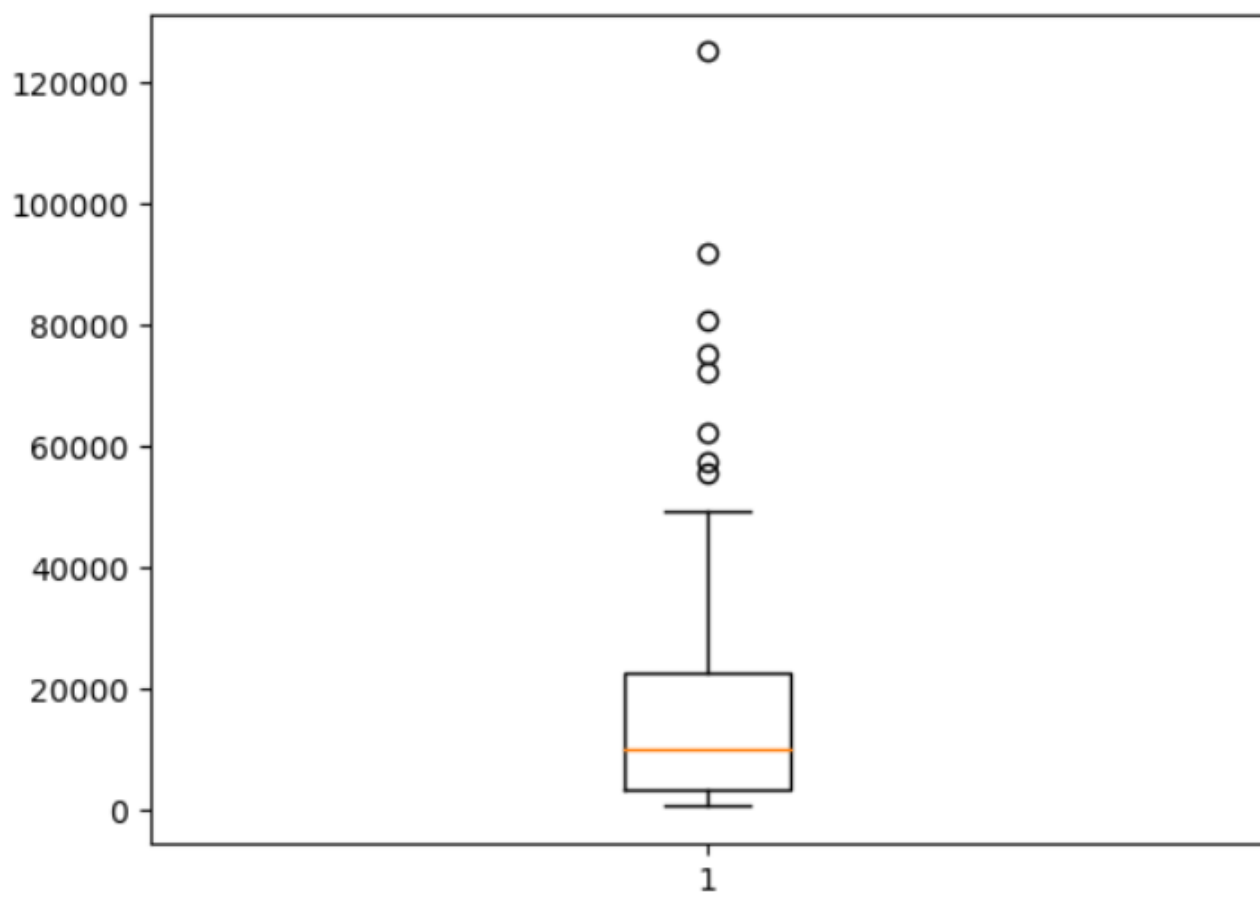
**Висновок:**

По-перше, після проведення аналізу викидів у вихідних даних, ми виявили декілька країн, де фактори доходу та ВВП досягають значень, які вже не виправдовують потреби у допомозі. Після цього застосовано метод "ліктя" для визначення оптимальної кількості кластерів для класифікації країн.

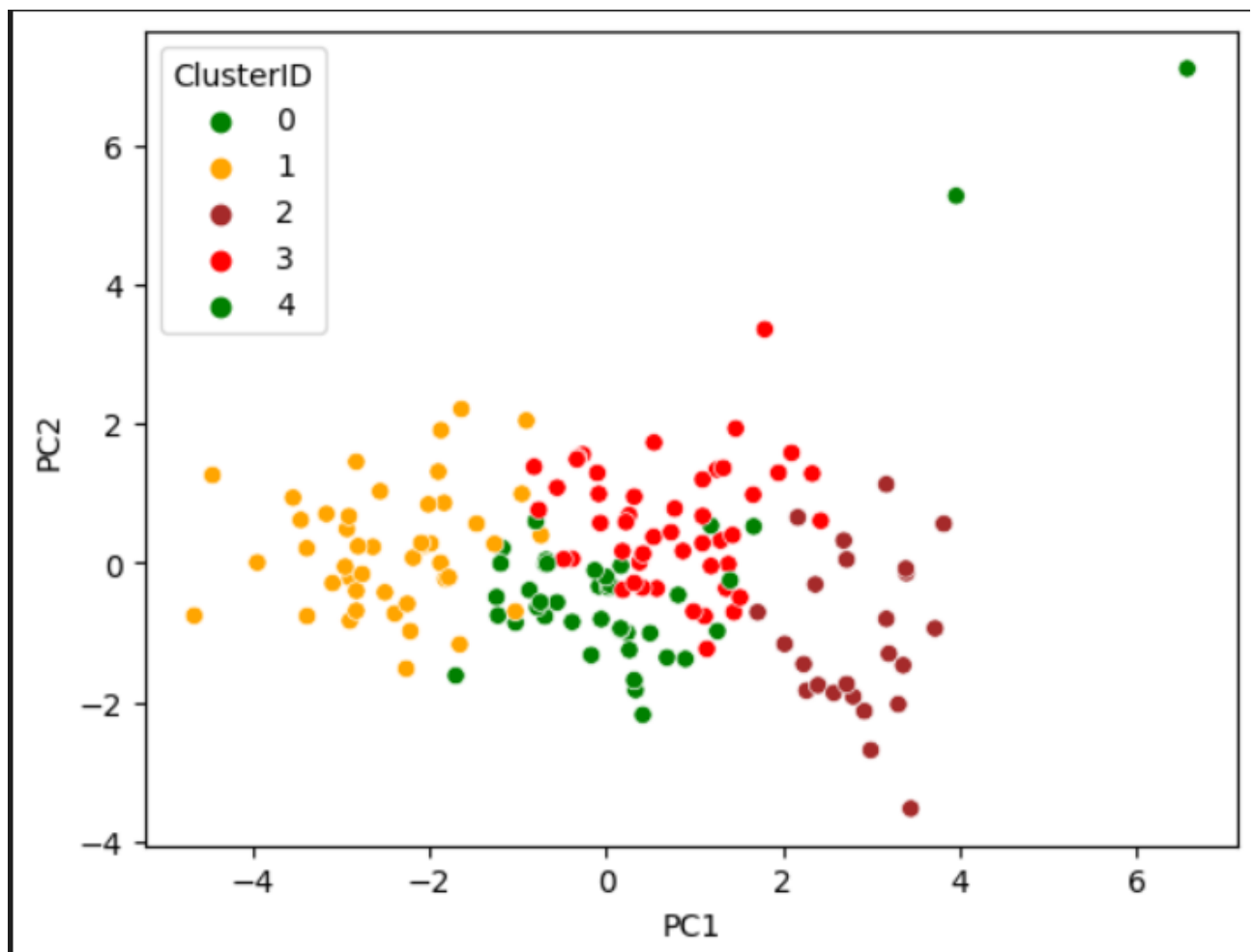


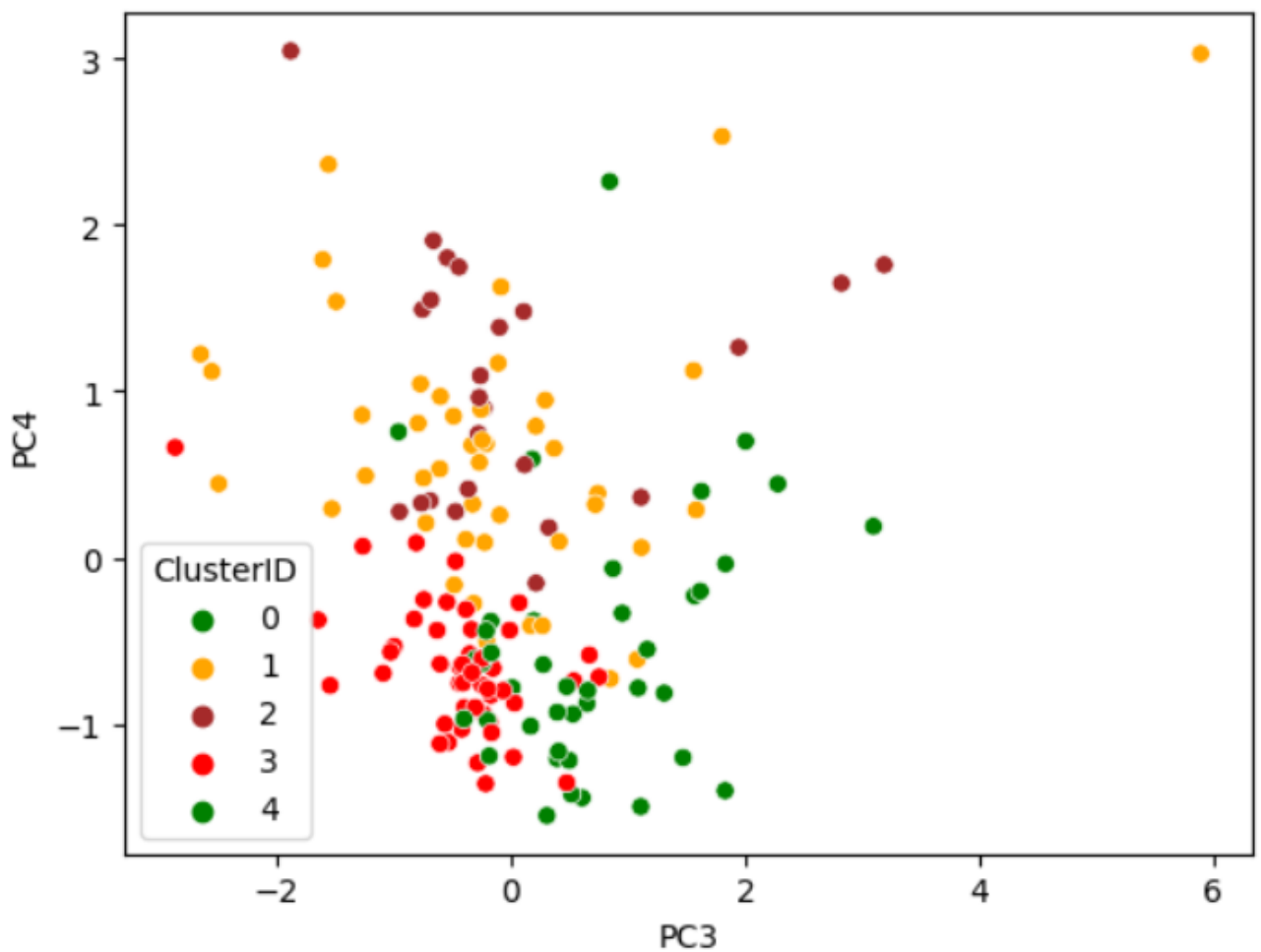
Після цього ми розподілили країни між цими кластерами та провели аналіз факторів щ урахуванням поділу країн на кластери.





Також візуалізували залежність факторів.





Для таких країн визначимо ті, в яких найбільша дитяча смертність та найменші дохід та ВВП. Результат: 5 країн у порядку першочерговості для надання допомоги.

1. Бурунді
2. Ліберія
3. Конго, Дем. Представник
4. Нігерія
5. Сьєрра-Леоне