*Додаток 1*

Міністерство освіти і науки України

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського"

Факультет інформатики та обчислювальної техніки Кафедра інформатики та програмної інженерії

Звіт

з комп’ютерного практикуму № 3 з дисципліни

«Аналіз даних в інформаційних системах»

на тему: «Описова статистика»

Виконав студент ІП-13, Бабашев О. Д.

(шифр, прізвище, ім'я, по батькові)

Перевірив Ліхоузова Т. А.

( прізвище, ім'я, по батькові)

Київ 202 3

# **Комп’ютерний практикум 3**

Тема – Описова статистика.

Мета – ознайомитись з методикою первинної обробки статистичних даних; проаналізувати вплив способу представлення даних на їх інформативність.

**Завдання**

Основне:

1. Скачати дані із файлу Data2.csv
2. Записати дані у data frame
3. Дослідити структуру даних
4. Виправити помилки в даних
5. Побудувати діаграми розмаху та гістограми
6. Додати стовпчик із щільністю населення

Додаткове:

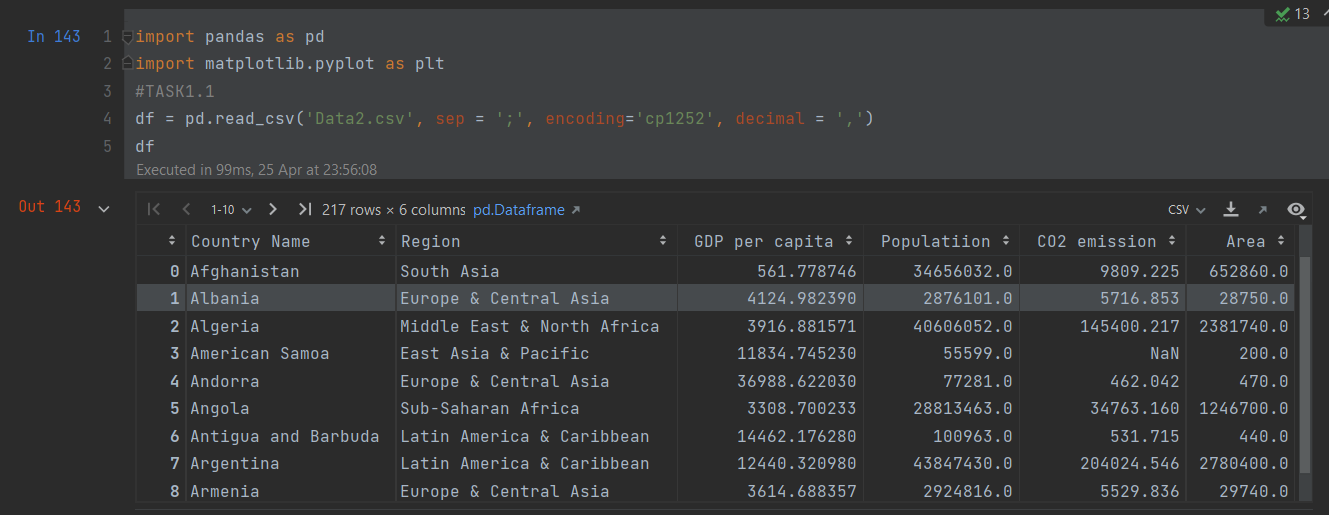
Відповісти на питання (файл Data2.csv):

1. Чи є пропущені значення? Якщо є, замінити середніми
2. Яка країна має найбільший ВВП на людину (GDP per capita)? Яка має найменшу площу?
3. В якому регіоні середня площа країни найбільша?
4. Знайдіть країну з найбільшою щільністю населення у світі? У Європі та центральній Азії?
5. Чи співпадає в якомусь регіоні середнє та медіана ВВП?
6. Вивести топ 5 країн та 5 останніх країн по ВВП та кількості СО2 на душу населення.

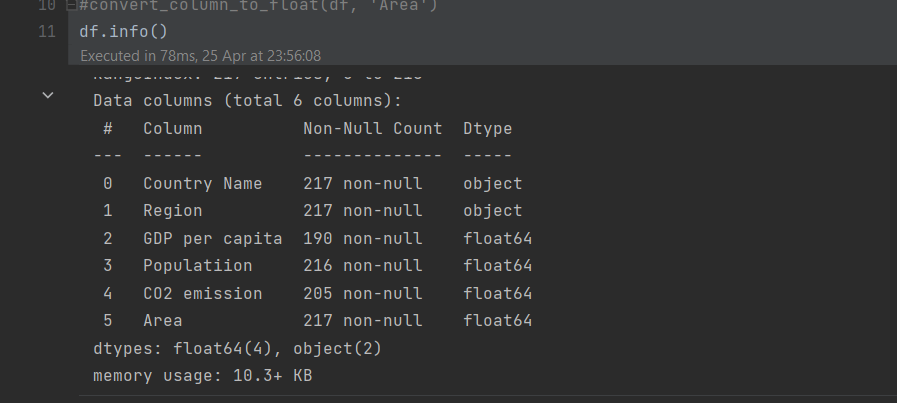
## **Основне завдання**

## **DataFrame та його структура**

За допомогою Python бібліотеки Pandas завантажено дані з даного csv файлу в dataframe.

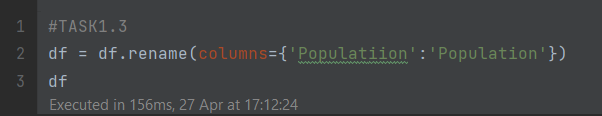


Досліджено структуру даних.

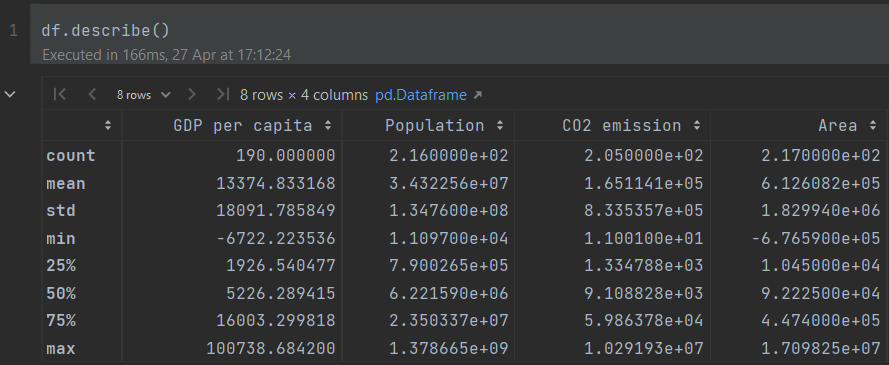


### **Виправлення помилок**

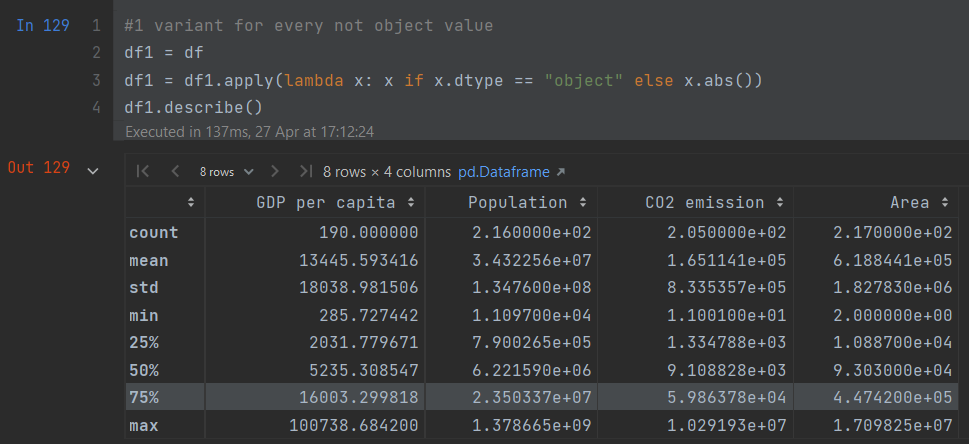
Змінено назву колонки ‘Populatiion’ на ‘Population’.



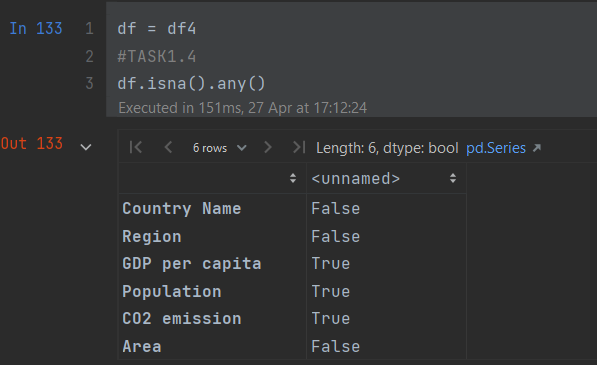
Знайдено рядки, поля яких містять від’ємні елементи.



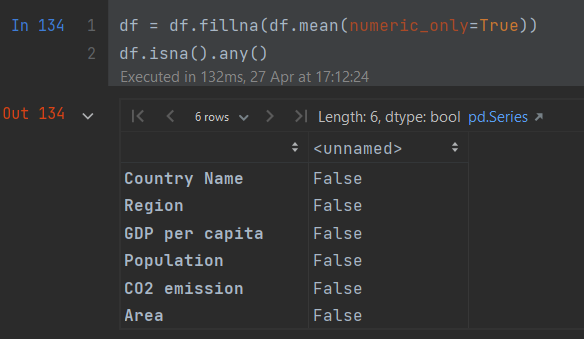
Виправлено від’ємні значення.



Перевірено на нульові значення.

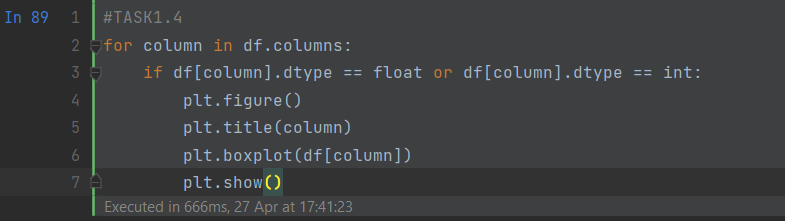


Замінено нульові на середнє арифметичне всіх значень.

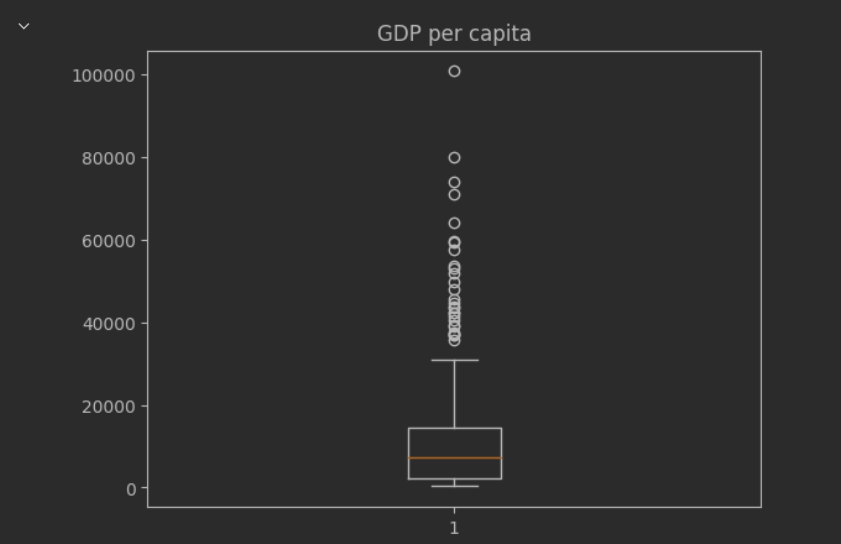


### **Діаграми розмаху та гістограми**

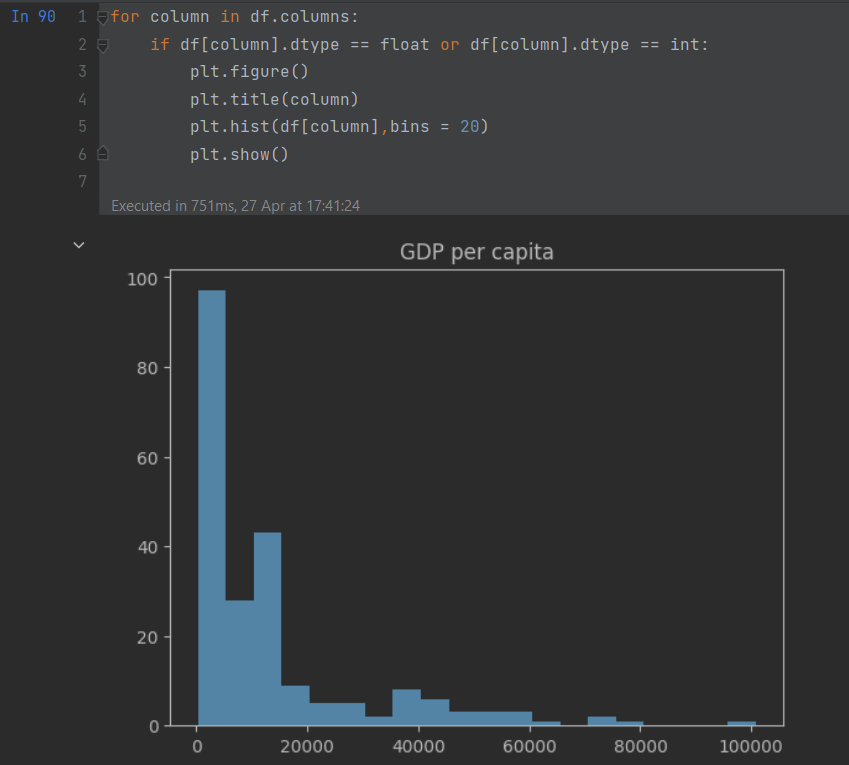
Виведено діаграми розмаху для кожного стовпця з чисельними даними.



Маємо подібні діаграми розмаху для кожного стовпця.

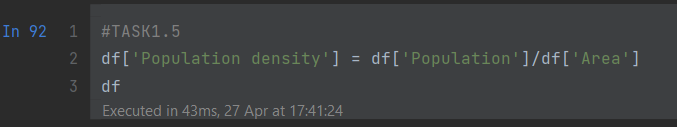


Аналогічно з гістограмою.



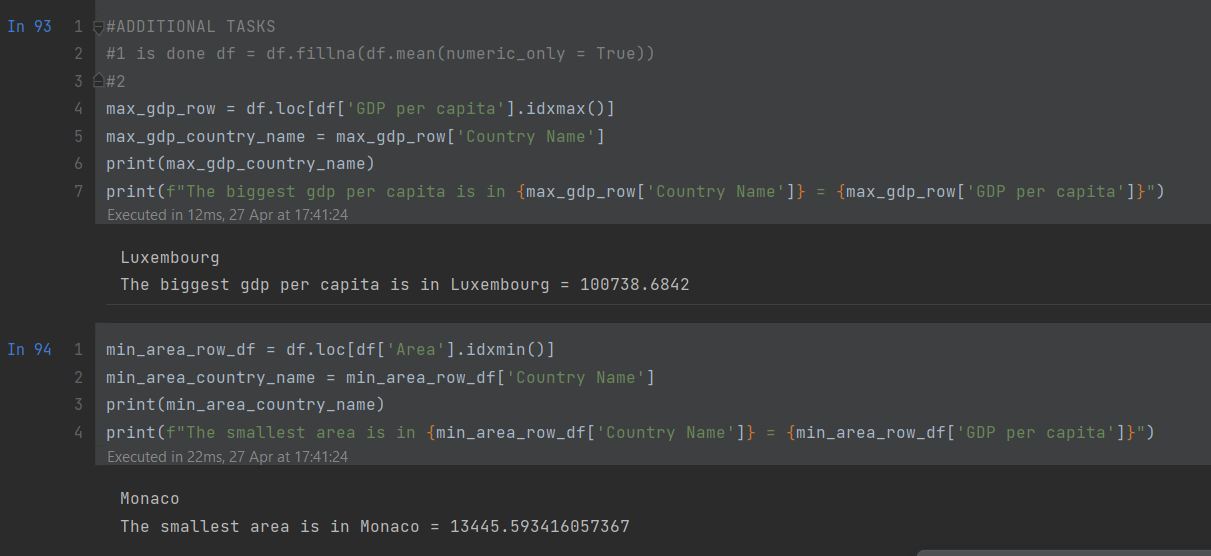
### **Додавання стовпчику із щільністю населення**

Додано стовпчик з щільністю населення.

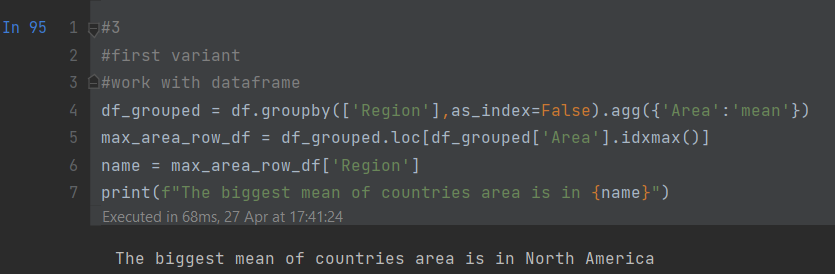


### **Додаткове завдання**

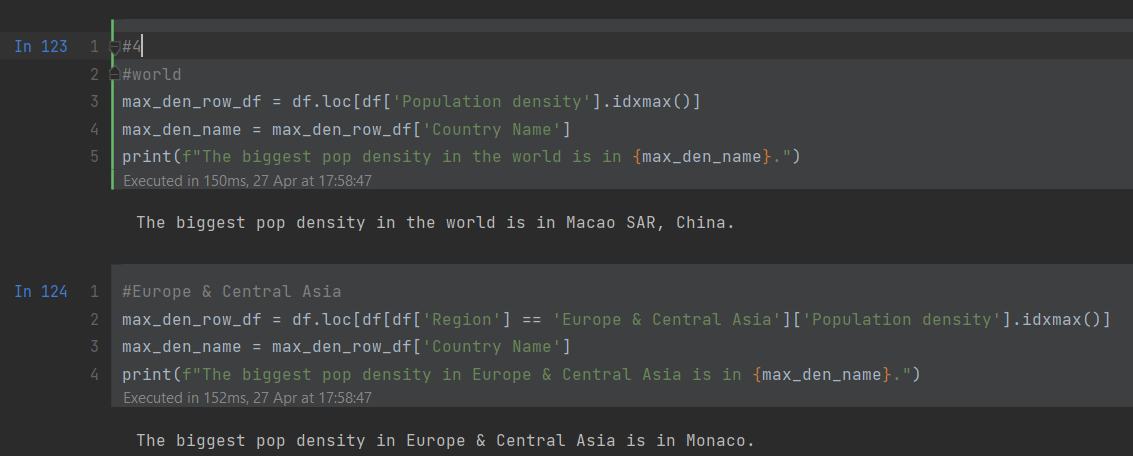
Виведено країну з найбільшим ВВП на душу населення та країну з найменшою площею.



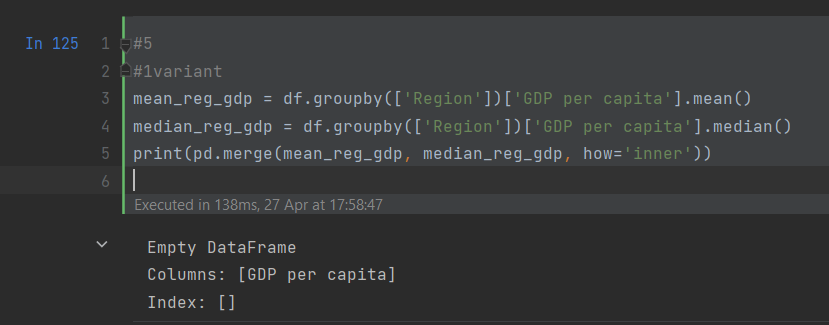
Виведено регіон з найбільшою середньою площею країн.



Виведено назву країни з найбільшою щільністю населення у світі, та у Європі та центральній Азії.

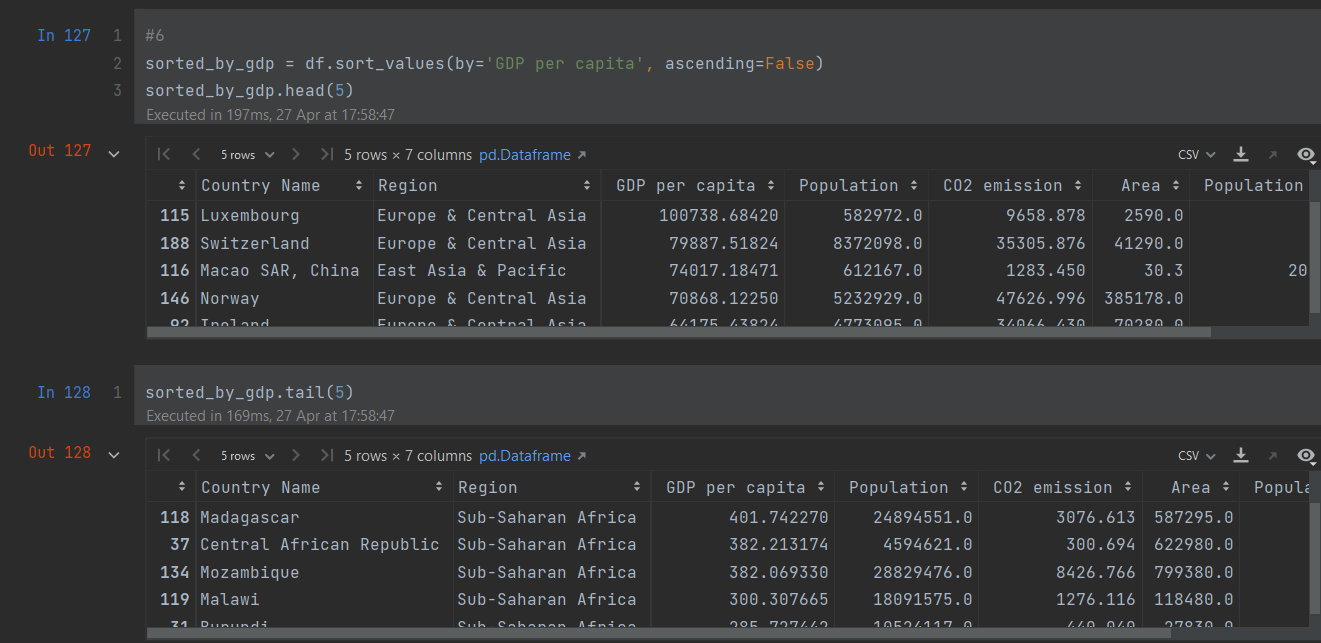


Виведено Співпадіння середнього та медіани ВВП по регіонам.

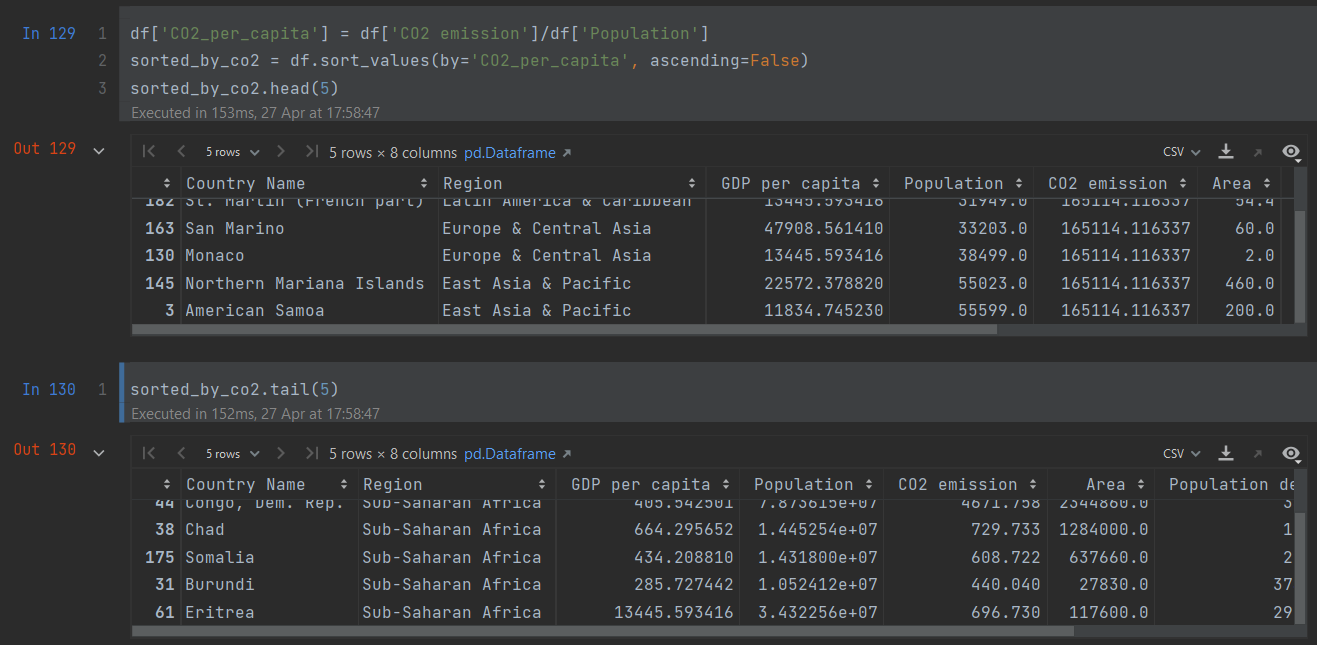


Виведено 5 найбільших назв країн по показниках ВВП та CO2 на душу населення та 5 найменших відповідно.

ВВП



СО2



Висновок

У цьому комп’ютерному практикумі був ознайомлений з основними інструментами роботи в python бібліотеках для роботи з даними pandas та matplotlib.

Використав на практиці опанований матеріал.