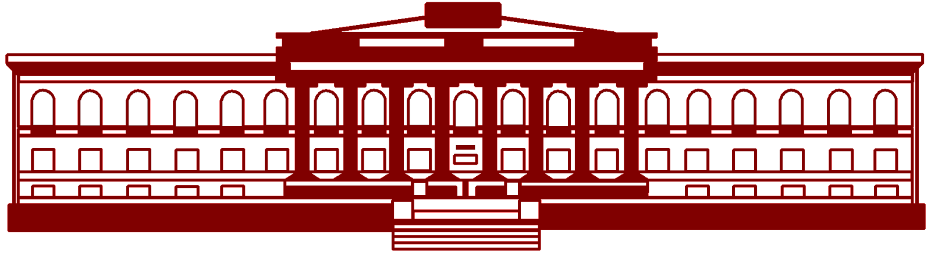
**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**

**імені ТАРАСА ШЕВЧЕНКА**

****

**ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

**Кафедра прикладних інформаційних систем**

**Звіт до лабораторної роботи №1**

**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.*

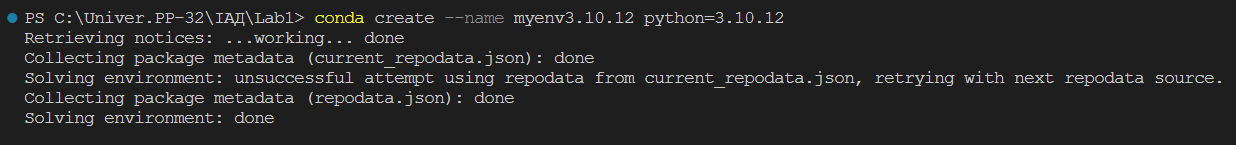
**Київ – 2023**

**Тема:** Розгортання середовища. Агрегація та дослідницький аналіз даних (EDA) пакетами Python.

**Мета роботи:** Метою лабораторної роботи є отримання практичних навичок з розгортання робочого середовища, виконання агрегації та дослідницького аналізу даних, використовуючи пакети jupyter, pandas, seaborn, notebook.

**Хід роботи:**

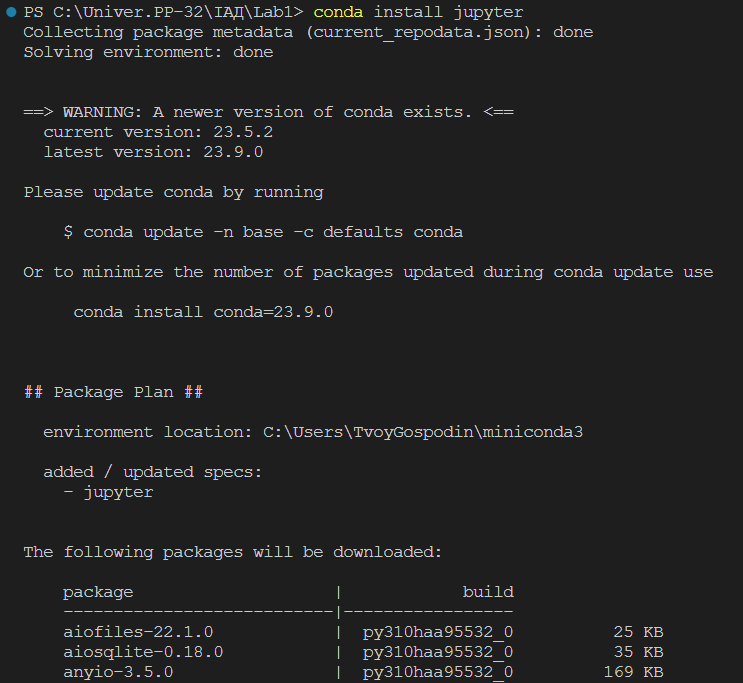
1. Встановив miniconda та Python 3.10.12 і створив віртуальне середовище conda:

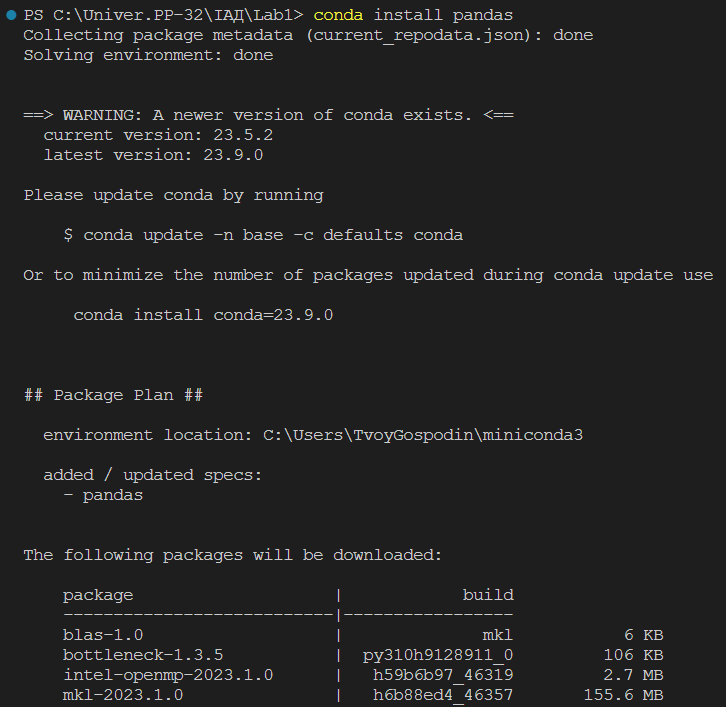


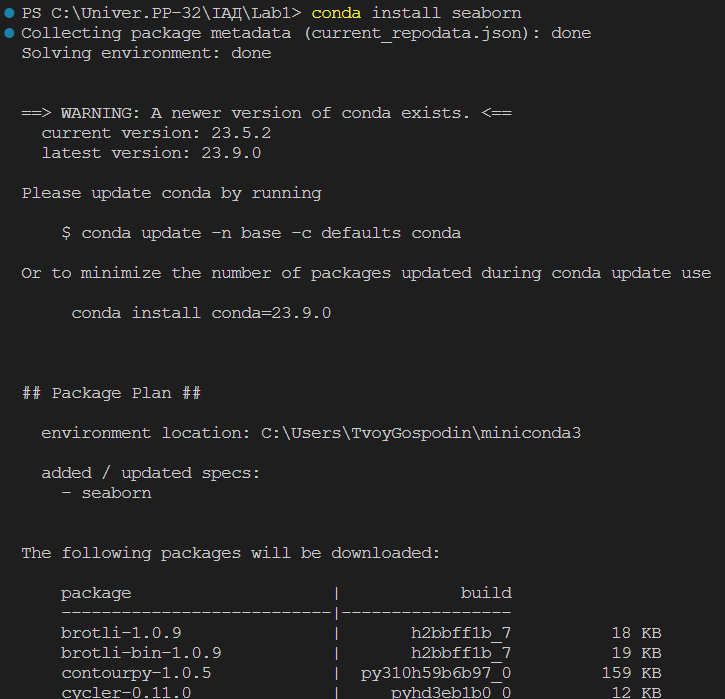
Після створення віртуального середовища активував його:



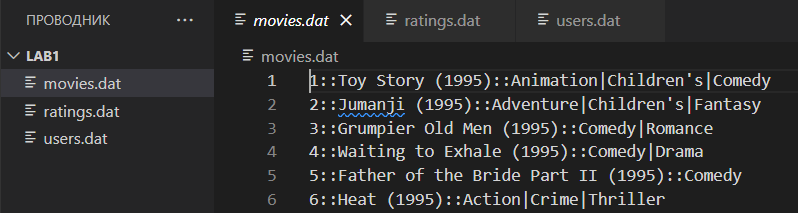
Встановив пакети jupyter/pandas/seaborn notebook:



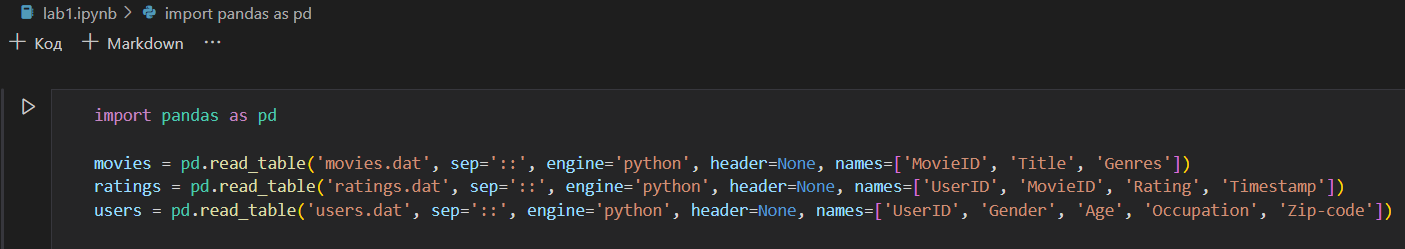




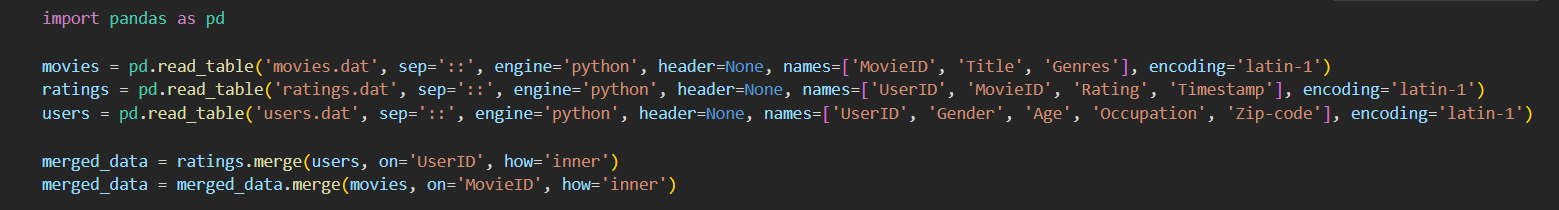
2. Завантажив у папку проекту файли з даними MovieLens 1M Dataset:



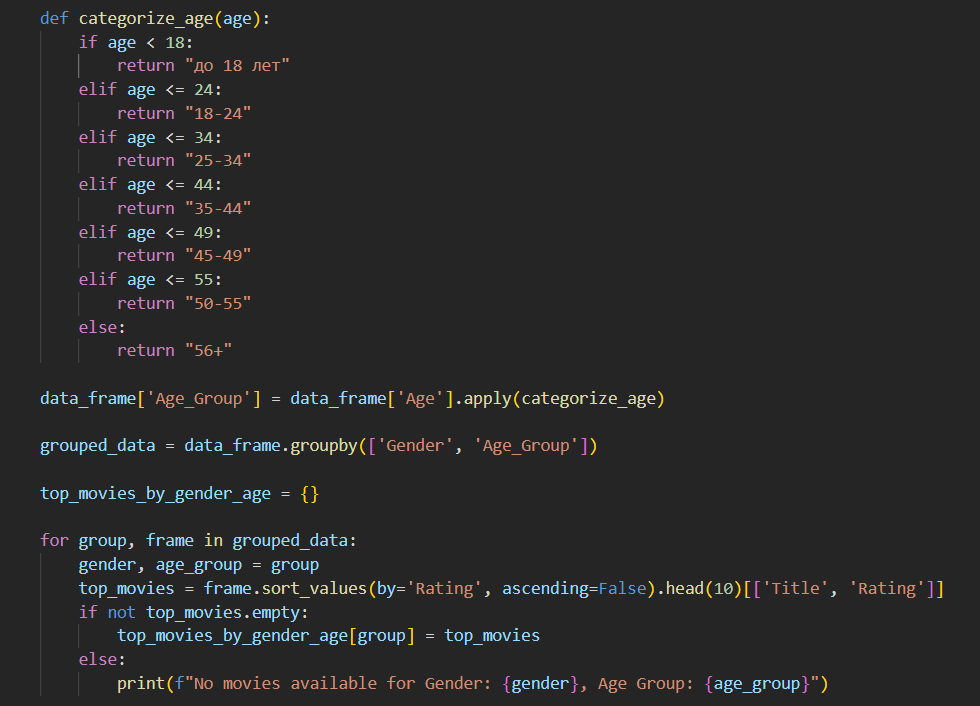
3. Завантажив файли movies.dat / ratings.dat / users.dat у змінні, використовуючи методи pandas:

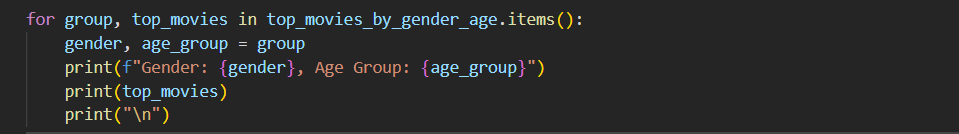


4. Зробив merge даних в єдиний DataFrame:



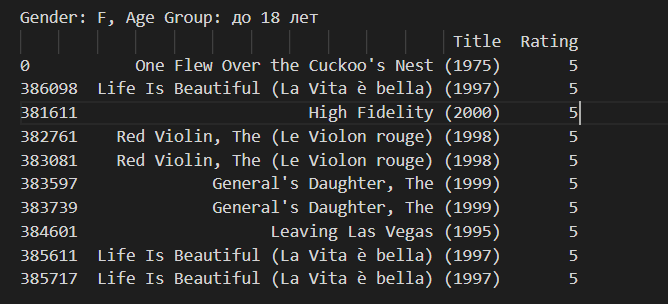
5. Зробив аналіз топ-10 стрічок за рейтингом по кожному полу та кожній віковій підгрупі з розбивкою по статі (14 рейтингів):

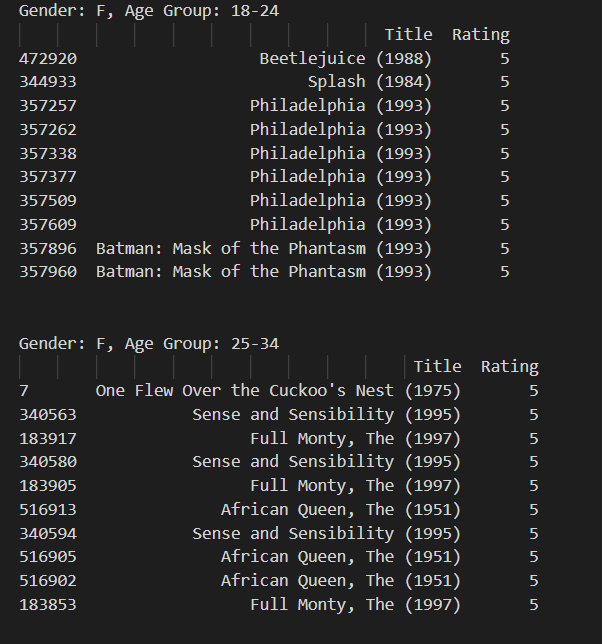
Код:  


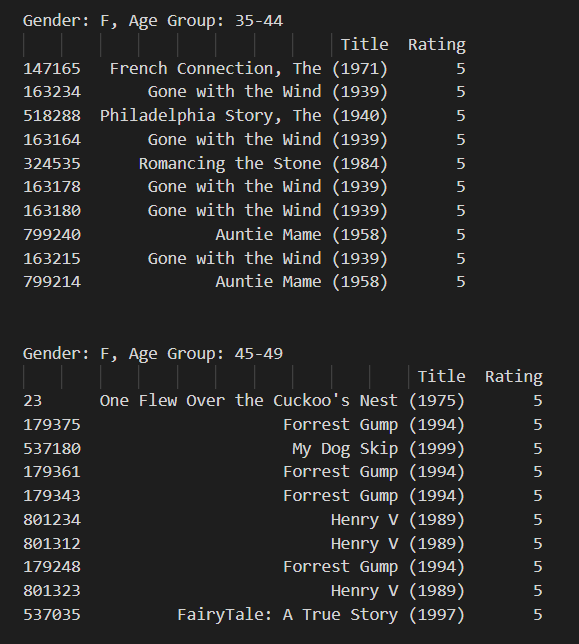


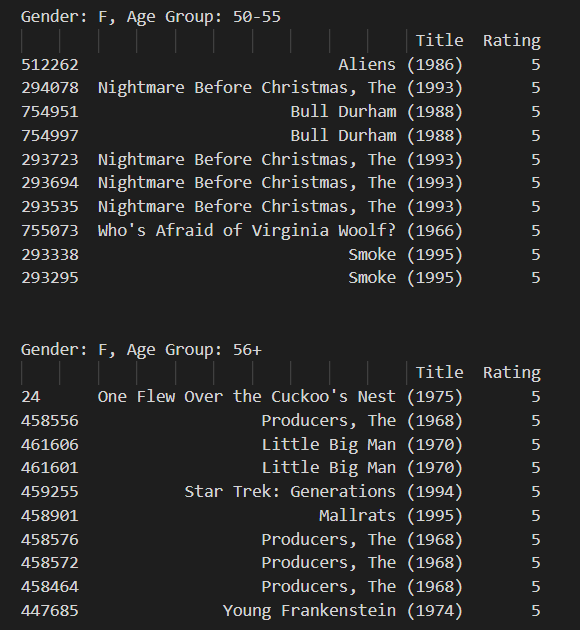
Результат:

Топ-10 стрічок за рейтингом по кожній віковій підгрупі за жіночою статтю:

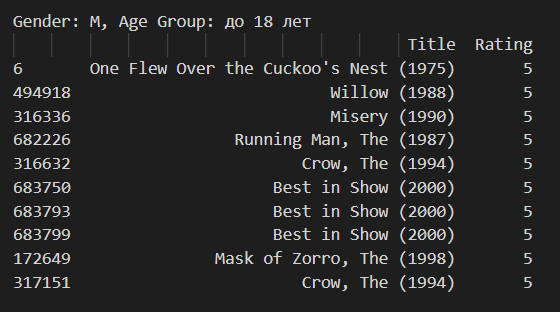


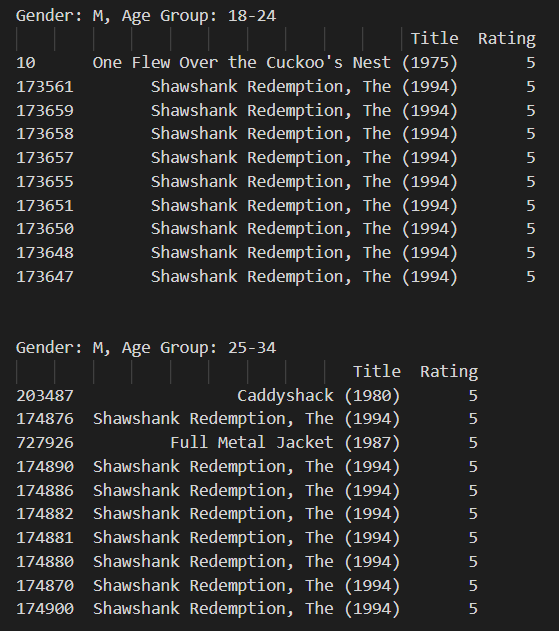


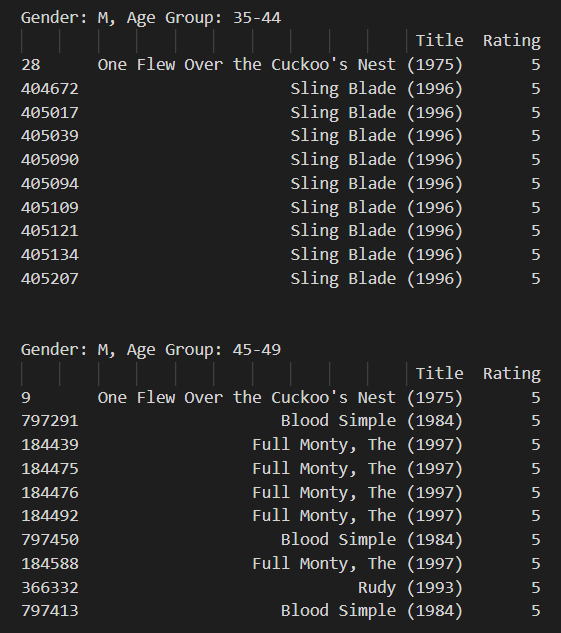


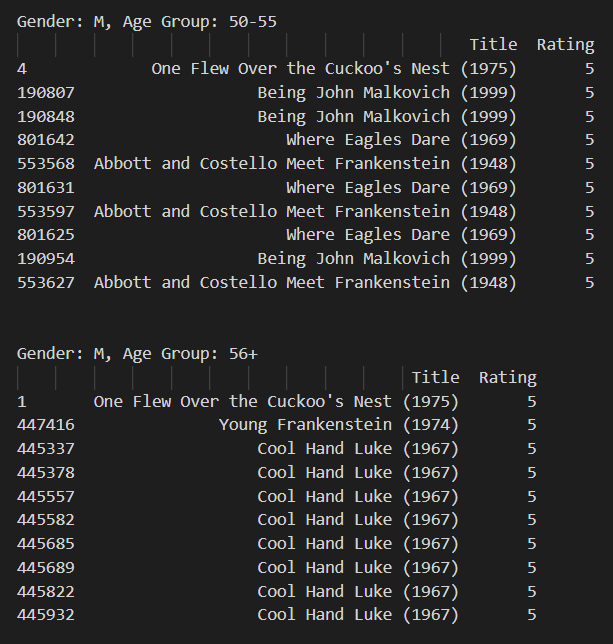


Топ-10 стрічок за рейтингом по кожній віковій підгрупі за чоловічою статтю:





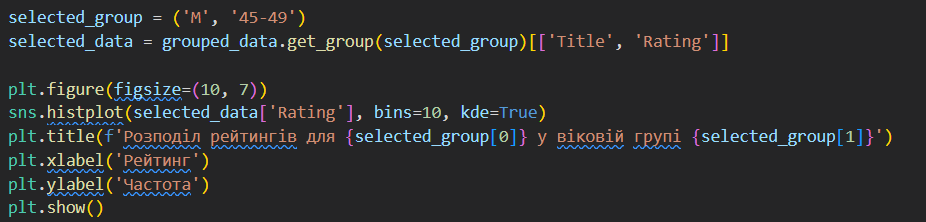




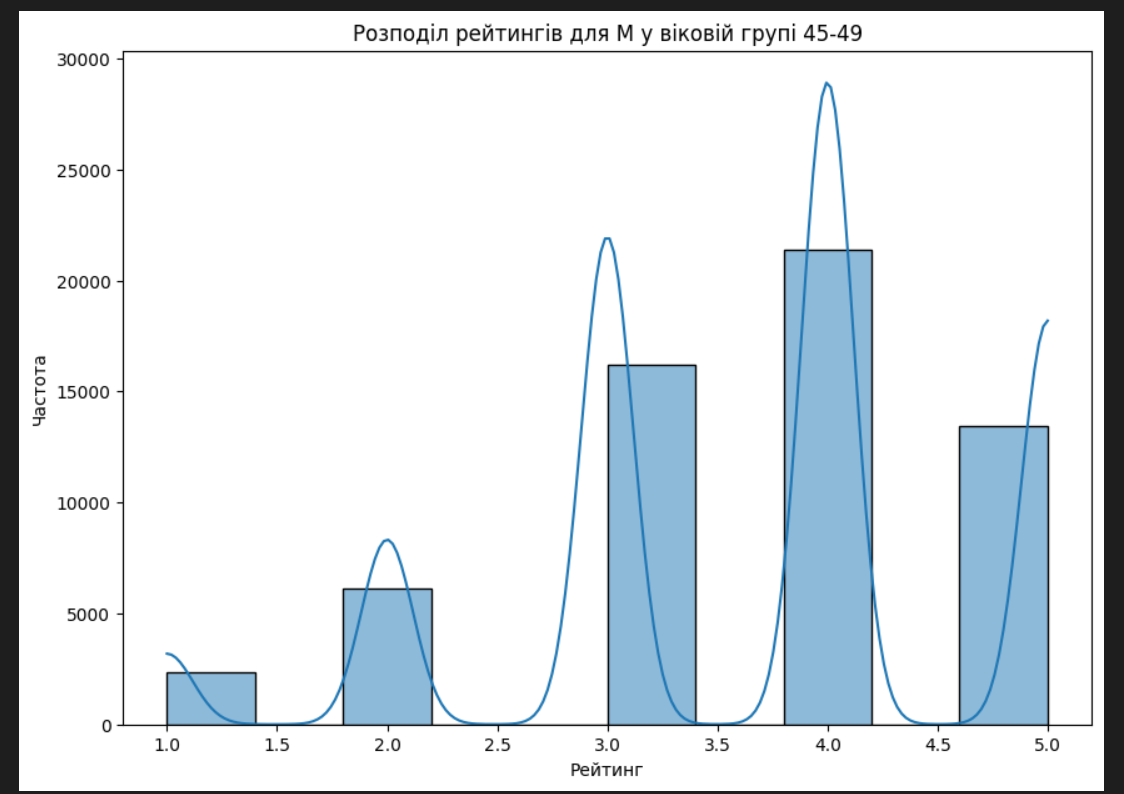
6. Візуалізував один із отриманих рейтингів засобами пакету seaborn:

Код:





Результат:



**Висновок:**

В ході виконання лабораторної роботи отримав практичні навички з розгортання робочого середовища, виконання агрегації та дослідницького аналізу даних, використовуючи пакети jupyter, pandas, seaborn, notebook.