МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»

**WEB-Аналітика**

ЛАБОРАТОРНА РОБОТА №4

|  |  |  |
| --- | --- | --- |
|  |  | Виконав  студент групи ФІ-42мн  Бєш Радомир Андрійович |

Київ 2025

**Завдання:**

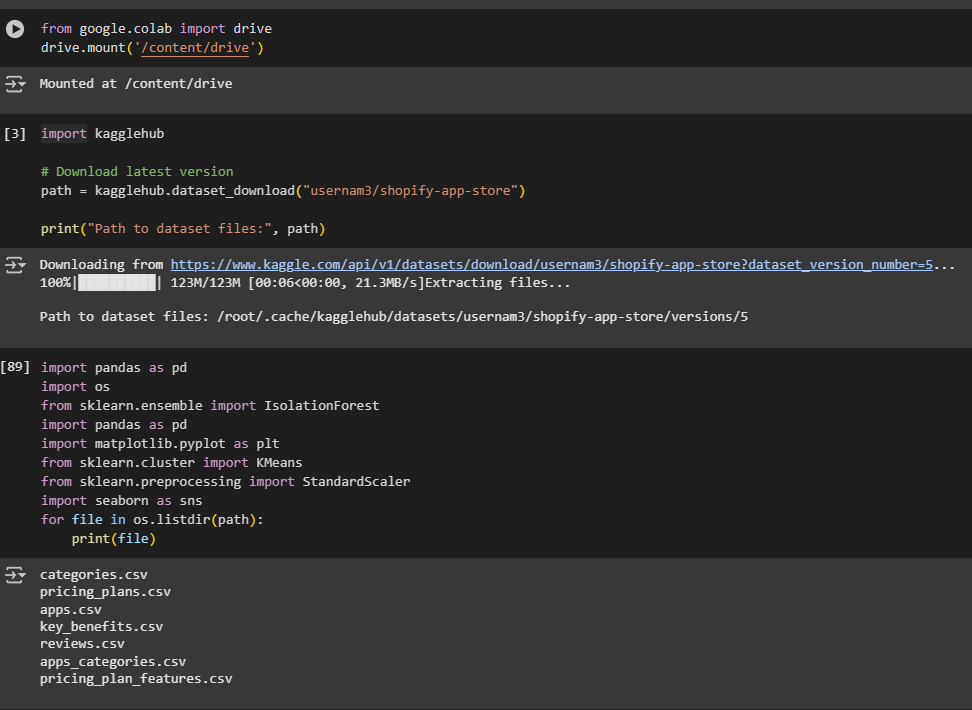
На основі даних з 3 лабораторної чи знайти достатньо великий датасет(від 10000 записів) на Кегель (або будь якому іншому) ресурсі використати будь-який метод машинного навчання для дослідження інформації та створення моделі даних.

**Хід роботи**

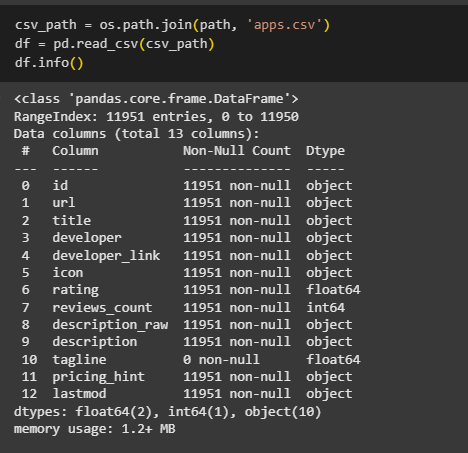
1. Обраний датасет:

Посилання: <https://www.kaggle.com/datasets/usernam3/shopify-app-store/data>

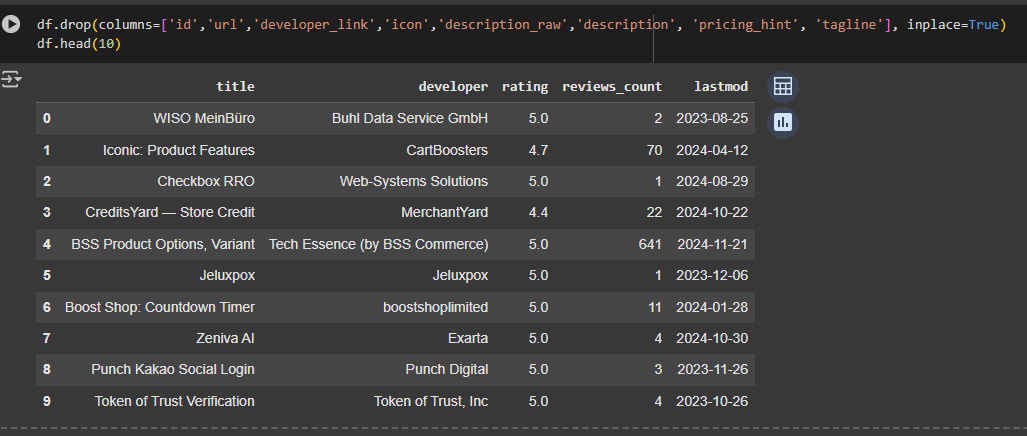
1. Завантаження датасету та первинна обробка



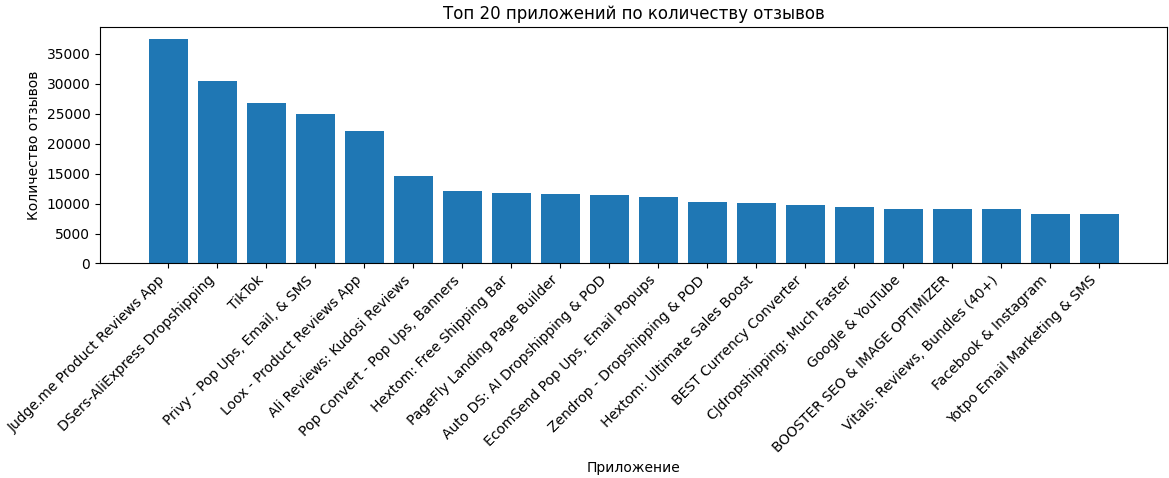
Будемо використовувати датасет з додатками та їх відгуками.



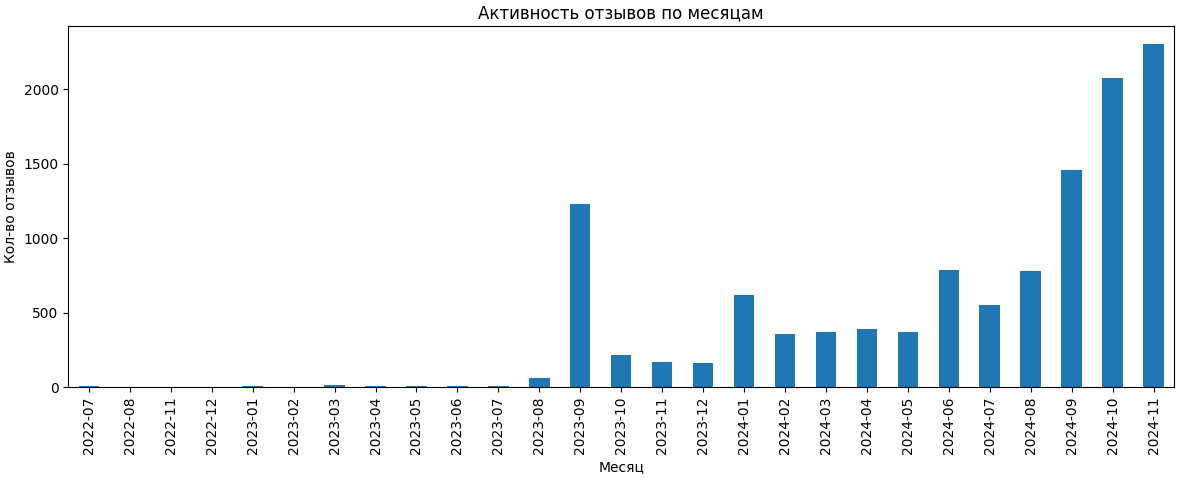
Після обробки, датасет має вигляд:



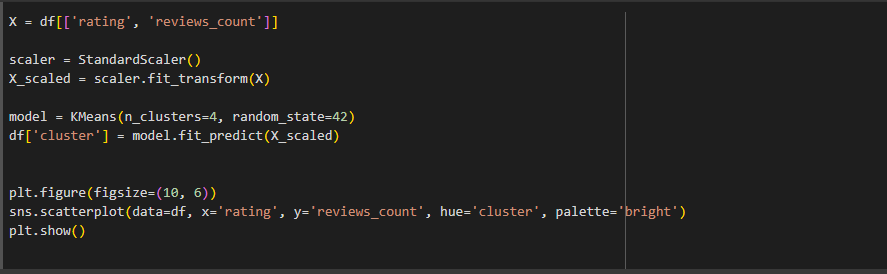
1. Візуалізуємо топ 20 додатків за кількістю відгуків

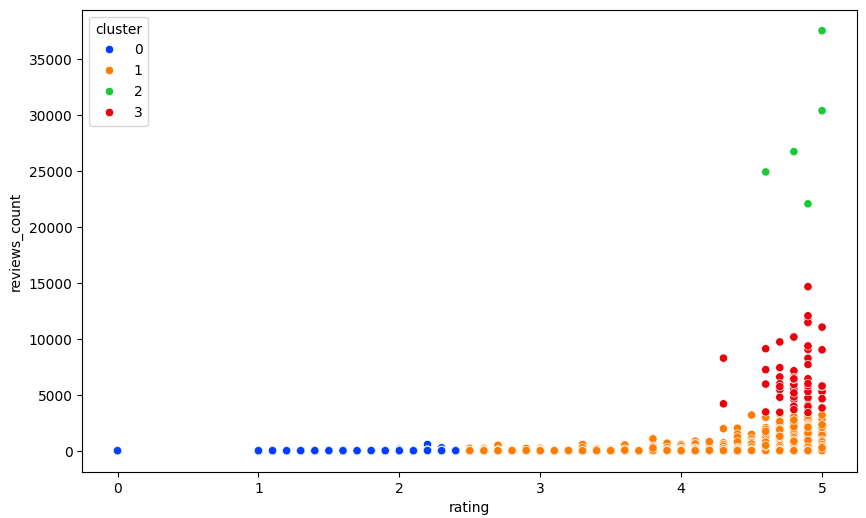


1. Знайдемо місяці і візуалізуємо, в яких було створено найбільша кількість відгуків



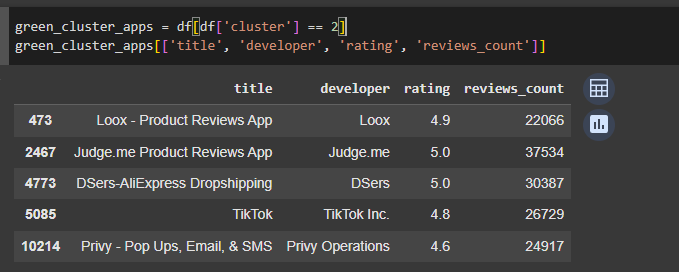
1. Кластеризація



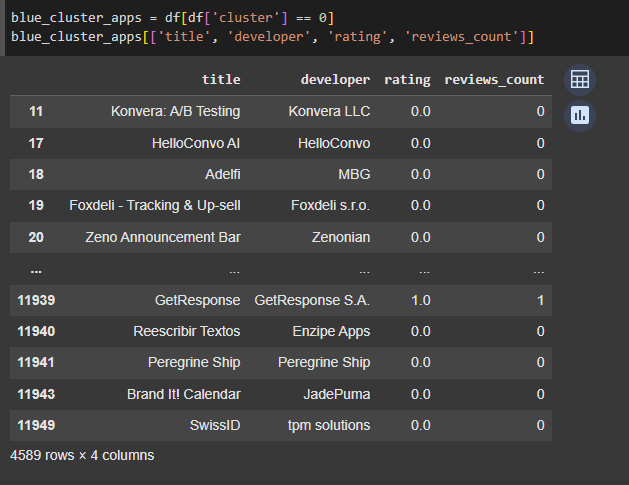


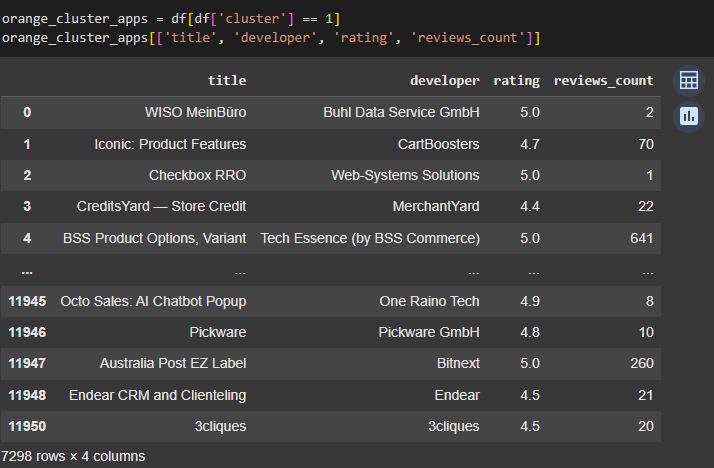
Можна зробити такі висновки, що кореляція присутня між rating та reviews\_count. Бачимо ,що додатки з високим рейтингом частіше мають більше відгуків, а додатки з низьким навпаки – менше. Це цілком логічно, так як популярні додатки найчастіше бувать якісними і продуманими та приваблюють більше юзерів, які і залишають відгуки. На графіку

Кластери чітко групують додатки за їхньою популярностю (кількість відгуків) та якістю (рейтинг). Це зозволяє легко обрати лідируючі додатки (кластер 2) та “проблемні” додатки (кластеры 0 и 1).



Додакти з низьким рейтингом і низькмими оцінками. Такі результати свідчать, що або додатки ще не набрали аудиторію (тобто про них ніхто не знає), або вони знаходяться на етапі розробки.





Загалом можна зробити висновки, додатки кластера 0 та 1 потребують у покращенні якості або маркетингової підтримки, щоб залучити більше користувачів та підвищити рейтинг. Додатки кластера 2 - це успішні продукти, які можна використовувати як приклад для інших розробників. Додатки кластера 3 показують гарний потенціал і можуть бути поліпшені для досягнення рівня популярності кластера 2.