

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»

Кафедра конструювання електронно обчислювальної апаратури

**Лабораторна Робота 2**

з дисципліни

«Високорівнева розробка програмного забезпечення»

Виконала студентка:

Харчук Каріна

Навчальної групи: ФЕЛ, ДК-21

Перевірив: Мушта І. А.

                                                         Київ 2024

**Завдання**

*Завдання 1. Аналіз даних із Kaggle*

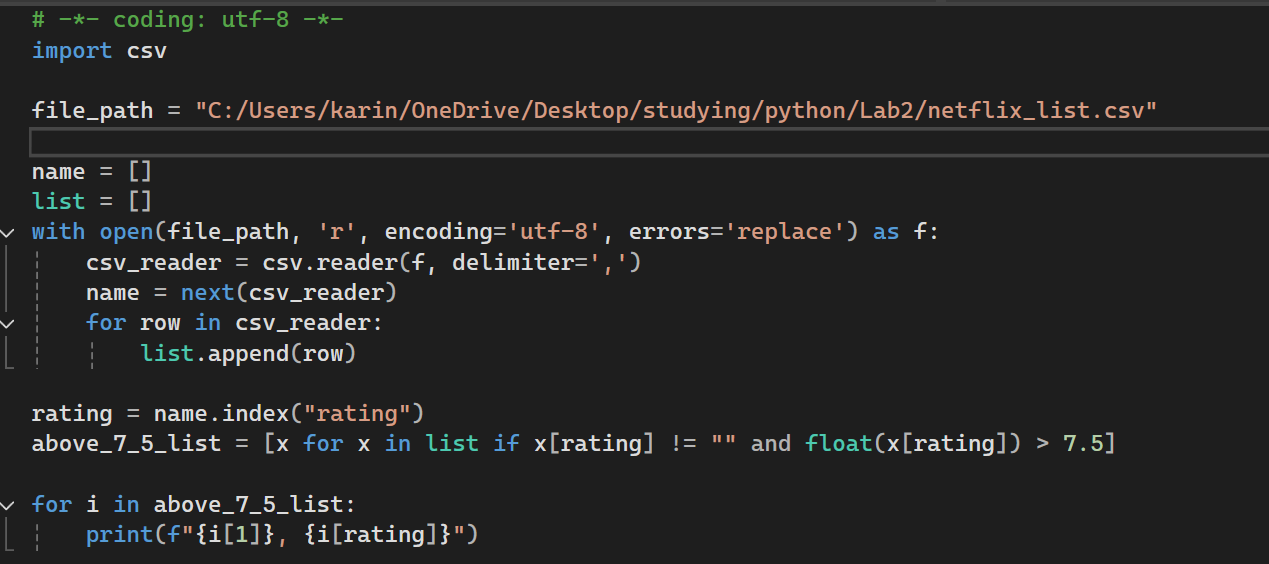
1. **Завантажте датасет із Kaggle**:

<https://www.kaggle.com/datasets/snehaanbhawal/netflix-tv-shows-and-movie-list/data>

1. **Виконайте наступні операції:**

**# -------------------------------------------- на 2 календарний контроль ----------------------------------**

* + Використайте вбудовані засоби Python для зчитування файлу netflix\_list.csv. Розділіть рядки за допомогою методу split(',') та збережіть у список списків.
  + Використайте **comprehensions** для виконання таких дій:   
    a) Створіть новий список, що містить **тільки шоу або фільми, де рейтинг більше 7.5**. (поле rating);



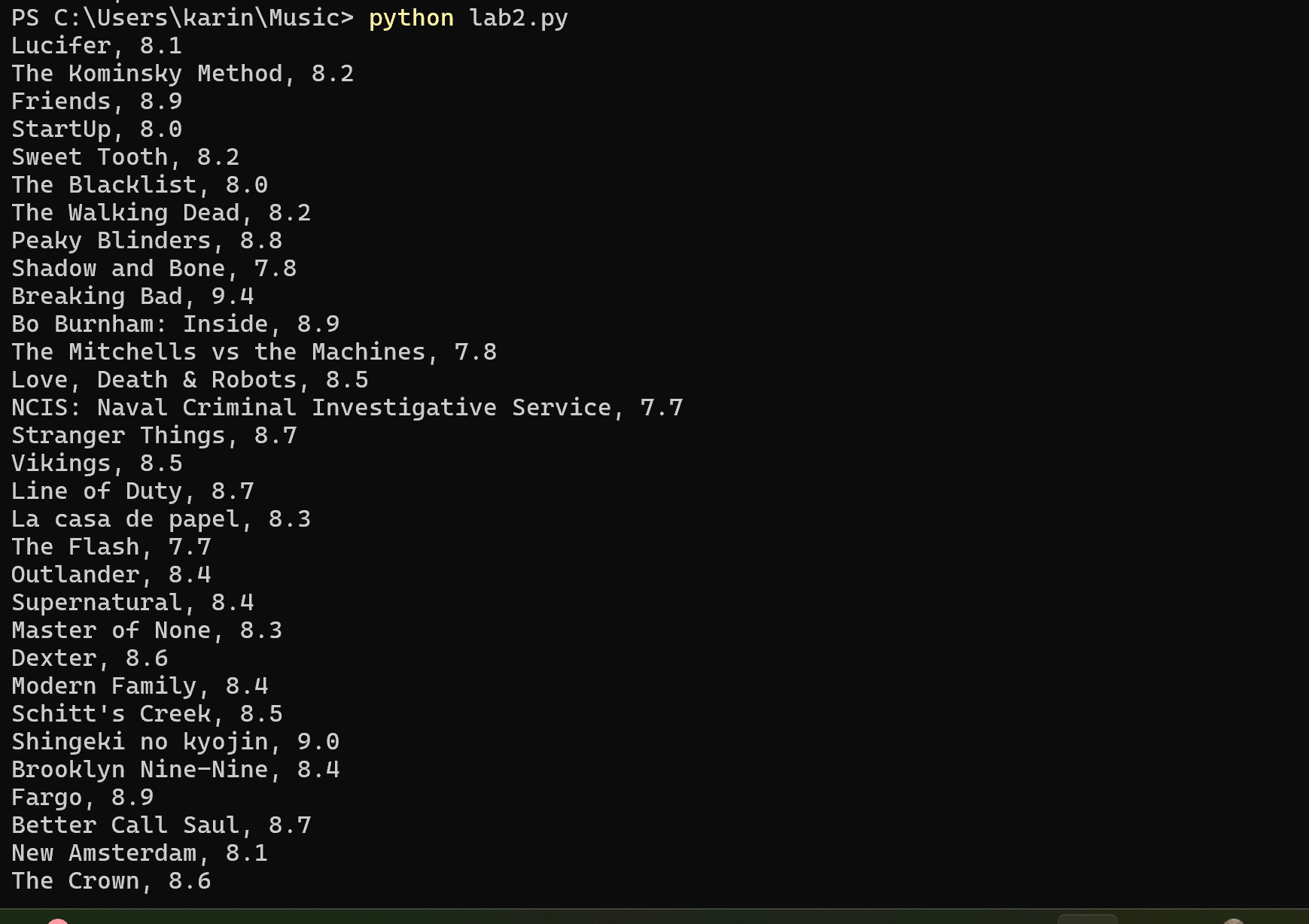
Перший рядок вказує, що файл збережений у кодуванні UTF-8. Це важливо для обробки тексту, особливо якщо у файлі є неанглійські символи.

import csv  
Імпортується модуль csv, який спрощує роботу з файлами у форматі CSV (Comma-Separated Values). Пишемо в file\_path  шлях до CSV-файлу, з яким працюватиме програма.  
Далі і ніціалізується порожній список name, у який будуть зчитані заголовки стовпців файлу та порожній список list, у який буде записано вміст рядків CSV-файлу.   
Відкривається файл file\_path для читання ('r') з використанням кодування utf-8. Параметр errors='replace' замінює некоректні символи спеціальними маркерами, щоб уникнути помилок під час читання.   
Далф використовується csv.reader для читання рядків файлу, при цьому дані розділяються комами.  
Зчитується перший рядок файлу (заголовки стовпців) і зберігається у список name.  
Фором починається цикл для обробки кожного наступного рядка файлу.   
Додає поточний рядок із файлу до списку list.   
rating = name.index("rating") знаходить індекс стовпця з назвою "rating" у заголовках (список name). Це дозволяє звертатися до стовпця рейтингу у кожному рядку.   
Створюється новий список above\_7\_5\_list, який включає лише ті рядки (x), де значення у стовпці rating:

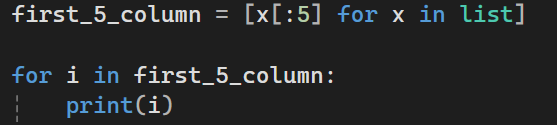
Не порожнє (x[rating] != "").

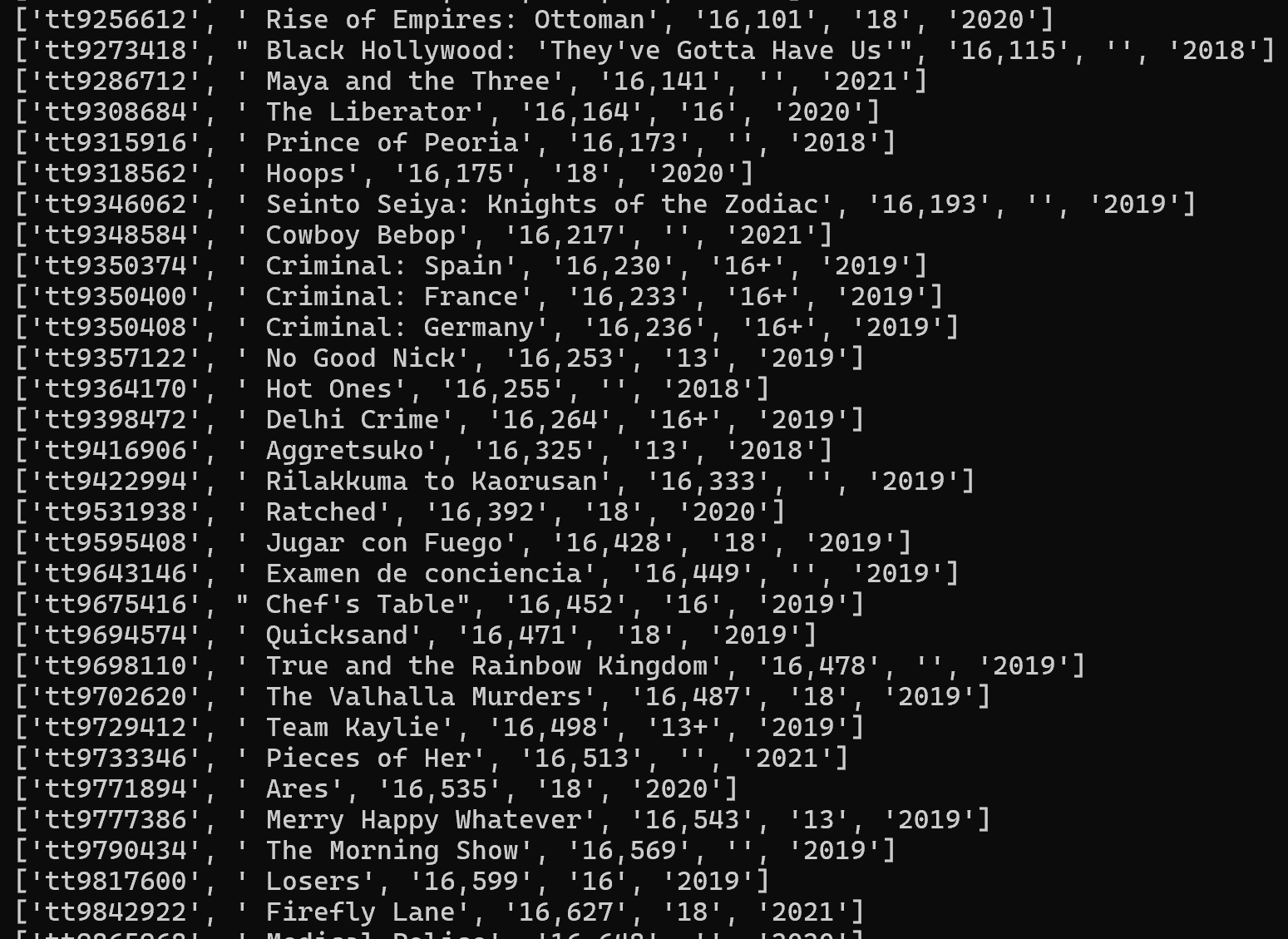
Більше 7.5 (float(x[rating]) > 7.5).

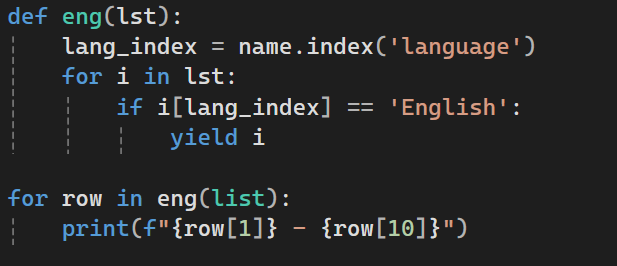
Потім починається цикл для виведення даних із відфільтрованого списку above\_7\_5\_list.   
Для кожного рядка у списку above\_7\_5\_list виводиться другий стовпець і рейтинг використовуючи форматування рядка.



b) Зберігайте тільки перші 5 колонок кожного рядка.



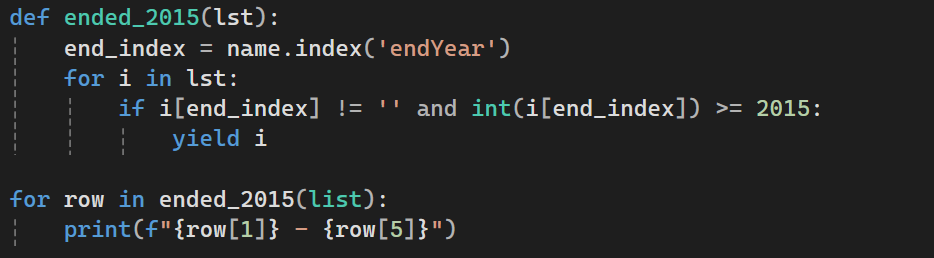


* + Напишіть **генераторну** функцію, яка буде повертати рядки, у яких:   
    a) Мова шоу або фільму —англійська (поле language); 

Оголошується функція eng, яка приймає на вхід список lst.   
Використовуючи заголовки, знаходимо індекс стовпця з назвою "language". Це дозволяє динамічно працювати зі стовпцями, не прив'язуючись до їхньої позиції.   
Починається цикл, який перебирає всі рядки списку lst і для кожного рядка перевіряється, чи значення в стовпці "language" дорівнює "English". Якщо умова виконується: yield i  
Рядок повертається через yield. Замість того, щоб одразу повертати всі результати, yield повертає їх по одному.

Далі запускається цикл, який використовує eng. Для кожного рядка, який повертає eng виконується обробка. Потім виводиться значення другого стовпця та десятого стовпця. Значення об'єднуються у строку через дефіс (-).

* +   
    b) Це або серіал, або фільм (поле type), що закінчився після 2015 року (поле endYear).   
    Використайте ключове слово **yield** для поступового повернення даних.





# ----------------------------------------------------------------------------------------------------------------------

* + Напишіть **ітератор**, який проходить через датасет та повертає лише **основний акторський склад** (поле cast), якщо його довжина перевищує 50 символів. Використайте цей ітератор для друку перших 10 таких записів.
  + Напишіть функцію, яка підраховує:   
    a) Скільки **шоу/фільмів для дорослих** (isAdult == 1) є у датасеті;   
    b) Середній рейтинг шоу та фільмів з більш ніж 1000 голосів (поле numVotes).
  + Використайте **генератор** та **comprehensions** для обробки файлу, щоб створити список заголовків шоу, які:   
    a) Мають більше 10 епізодів;   
    b) Рейтинг вище середнього.

***Завдання 2. Реалізація самобалансуючого дерева***

Реалізуйте **самобалансуюче двійкове дерево**. Дерево повинно автоматично балансуватися під час додавання нових елементів за рахунок розбиття та обертання (приклад – алгоритм AVL-дерева).

### ****Умови виконання:****

1. Використайте класи для реалізації дерева.
2. Використовуйте **рекурсивні функції** для вставки нових елементів та балансування дерева.
3. Створіть функцію для **обчислення висоти піддерева**.
4. Реалізуйте функцію для **обертання** дерева вліво та вправо для балансування.
5. Функція insert повинна вставляти нові елементи та підтримувати баланс дерева.