

Домашнє завдання №4.

Дослідження бази даних IMDb

Пастернак Назар

17.12.2018

Дослідження й глибокий аналіз актора та його найкращих фільмів

Короткий Опис

Під час вибору теми для дослідження першим в голову мені прийшло зробити дослідження на актора та його кращі фільми. Проте я не хотів, щоб програма аналізувала лише одного актора, тому я вирішив зробити програму більш гнучкою і надав користувачу програми змогу вводити ім'я актора, дослідження на якого він хоче отримати.

Опис роботи програми

Програма використовує базу даних IMDb, а саме файли :

- **title.akas.tsv.gz**
- **title.basics.tsv.gz**
- **title.principals.tsv.gz**
- **title.ratings.tsv.gz**
- **name.basics.tsv.gz**

Програма запитує у користувача ім'я актора і з файлу **name.basics.tsv.gz** отримується інформація з найвідомішими фільмами актора, рік народження та основні позиції, які він посідав.

Тоді, з файлу **title.basics.tsv.gz** береться інформація про фільми, які були отримані в попередньому кроці. Отримується назва фільму, його жанр та рік випуску. Формується словник з фільмами, ключем якого є назва фільму, а значенням – інформація про нього.

Опісля, з файлу **title.ratings.tsv.gz** береться інформація про рейтинг кожного фільму зі словника і утворюється новий словник, ключем якого є фільм, а значенням список з середнім рейтингом та к-стю голосів.

Тоді, отримавши усю інформацію щодо найпопулярніших фільмів заданого актора, користувач отримує запитання чи переглядав він/вона кожен фільм. Якщо відповідь так – пропонується дати свою оцінку та потім порівняти її з оцінкою певної к-сті людей за базою даних IMDb. Якщо відповідь ні – користувачу надається інформація про рейтинг та пропонується переглянути цей фільм.

Усі рейтинги щодо певних фільмів, які ввів користувач зберігаються і, після завершення програми, записуються у спеціально створений файл **own.ratings.txt**, де можна переглянути середнє значення оцінок для кожного фільму, оціненого з допомогою цієї програми, а також, рейтинг за версією IMDb разом із значеннями, які були введені використовуючи цю програму.

Опісля, користувачу надається змога обрати фільм і отримати переклади його назви на різні мови з використанням файлу **title.akas.tsv.gz**.

Наступним кроком з файлу **title.principals.tsv.gz** знаходяться усі фільми за участі обраного актора і для кожного з них знаходиться інформація про рейтинг у файлі **title.ratings.tsv.gz**. Тоді відбувається обчислення середнього значення рейтингів усіх фільмів, враховуючи к-сті голосів.

Вкінці виводиться середній рейтинг фільмів, у яких актор знімався, жанр, який найбільше з ним асоціюється та найуспішніший його фільм.

Приклади:

1.

Hey! Enter a name and a surname of an actor and get a quiz:

Tom Hanks

Tom Hanks, born in 1956, mostly known for being a/an producer,actor,soundtracker.

Starred in 'Big', 'Forrest Gump', 'Saving Private Ryan' and 'Cast Away'

Have you watched Big?

Answer 'yes' or 'no':

yes

Rate it from 1 to 10:

8

Your answer: 8, 173539 people according to IMDb rated 'Big' : 7.3

Have you watched Forrest Gump?

Answer 'yes' or 'no':

no

You should watch it then! Its rating is 8.8 according to 1544028 people

Have you watched Saving Private Ryan?

Answer 'yes' or 'no':

yes

Rate it from 1 to 10:

10

Your answer: 10, 1067892 people according to IMDb rated 'Saving Private Ryan' : 8.6

Have you watched Cast Away?

Answer 'yes' or 'no':

no

You should watch it then! Its rating is 7.8 according to 459298 people

Choose a film, which title's translations you want:

'Big', 'Forrest Gump', 'Saving Private Ryan', 'Cast Away', or type 'skip'

Big

'Big' on TR would be 'Büyük'.

'Big' on GR would be 'Μπικ'.

'Big' on FR would be 'Big'.

'Big' on PL would be 'Duży'.

'Big' on GR would be 'Big'.

'Big' on PE would be 'Quisiera ser grande'.

'Big' on PT would be 'Big'.

'Big' on RU would be 'Большой'.

'Big' on RO would be 'Vreau sa fiu mare'.

'Big' on HR would be 'Veliki'.

'Big' on DK would be 'Big'.

More info about Tom Hanks is coming!

~ Avarage Tom Hanks's rating, for all movies he/she has starred in, is 7.8

~ Tom Hanks is mostly assosiated with Drama genre(s)

~ Tom Hanks's most successful film is: 'Forrest Gump', which was filmed in 1994 with rating 8.8

You can check the rating information in 'own_ratings.txt' file

2.

Hey! Enter a name and a surname of an actor and get a quiz:

Johny Dep

Reenter your actor please!

Johny Depp

Reenter your actor please!

Johnny Depp

Johnny Depp, born in 1963, mostly known for being a/an actor,producer,soundtracker.

Starred in 'Pirates of the Caribbean: The Curse of the Black Pearl', 'Pirates of the Caribbean: Dead Man's Chest', 'Sweeney Todd: T

he Demon Barber of Fleet Street' and 'Pirates of the Caribbean: At World's End'

Have you watched Pirates of the Caribbean: The Curse of the Black Pearl?

Answer 'yes' or 'no':

yes

Rate it from 1 to 10:

10

Your answer: 10, 942051 people according to IMDb rated 'Pirates of the Caribbean: The Curse of the Black Pearl' : 8.0

Have you watched Pirates of the Caribbean: Dead Man's Chest?

Answer 'yes' or 'no':

no

You should watch it then! Its rating is 7.3 according to 598828 people

Have you watched Sweeney Todd: The Demon Barber of Fleet Street?

Answer 'yes' or 'no':

no

You should watch it then! Its rating is 7.4 according to 316024 people

Have you watched Pirates of the Caribbean: At World's End?

Answer 'yes' or 'no':

no

You should watch it then! Its rating is 7.1 according to 543305 people

Choose a film, which title's translations you want:

'Pirates of the Caribbean: The Curse of the Black Pearl', 'Pirates of the Caribbean: Dead Man's Chest', 'Sweeney Todd: The Demon Ba

rber of Fleet Street', 'Pirates of the Caribbean: At World's End', or type 'skip'
skip

You have skipped the translations

More info about Johnny Depp is coming!

~ Avarage Johnny Depp's rating, for all movies he/she has starred in, is 7.1

~ Johnny Depp is mostly assosiated with Action and Adventure and Fantasy genre(s)

~ Johnny Depp's most successful film is: 'Pirates of the Caribbean: The Curse of the Black Pearl', which was filmed in 2003 with ra

ting 8.0

You can check the rating information in 'own_ratings.txt' file