Computer Vision 2 Assignment

1 June 2023 V1.0

Cheeyau Aw | 645698

## Section 1

Source of the images: <https://www.cs.ccu.edu.tw/~wtchu/projects/MoviePoster/index.html>

The howl dataset is composed meta data and images of films from 1980 till 2015 from IMDB which contains txt files with meta data of movies and movie posters in jpeg format. Each year is separated in a text file which contain the next fields: \_id, Released, Writer, Type, imdbVotes, Response, Genre, Metascore, imdbID, imdbRating, Title, Rated, Language, Country, Box\_office, Runtime, Actors, Director, Awards, Poster. And folder of movies from that year.

The chosen data years are the 2008 and 2009 which for 2008 contains 357 objects and 20 fields, the 2009 contains 318 objects and 20 fields. The images folder for 2008 contains 318 images and the 2009 folder contains 284 images. This can lead due to some problems as the there are more meta info of movies then images. This required the removal of the meta data without images.

|  |  |
| --- | --- |
| The readable meta data of movies for each year. |  |
|  | The number of images per year which contain less than the readable movie objects. |

## Section 2

## Section 3

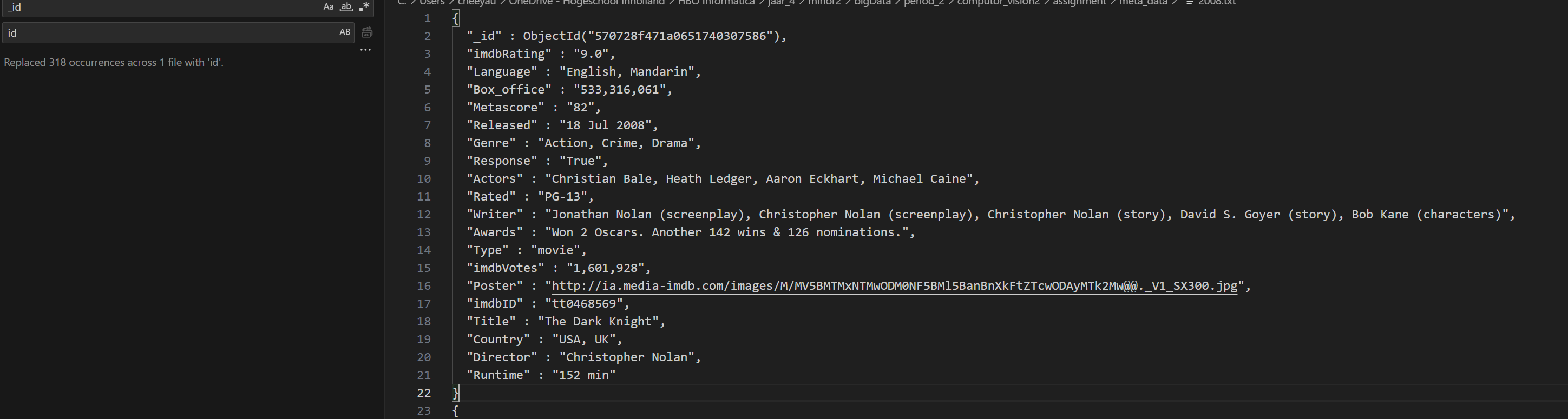
## Section 4

## Section 5

## Section 6

* Clean up txt unknown format to json in vscode by search and replace, its faster
* Place block brackets so its seen as array

Old



New

