Assignment 3

Your assignment is to create a movie analytics tool that allows the user to search for a movie of their choosing and obtain further information about it using online data and text analytics. First, your program should ask the user to specify a movie to analyze. Then, your program should ask the user which type of analysis they would like to run. Your program should support five different types of analyses:

- Background: Your program should print out the movie's year of release, rating (for example, PG-13), runtime, genre, actors, and plot summary. These should be obtained from the Open Movie Database (OMDb) API, discussed below.
- Reception: your program should print out the movie's awards, metascore, and IMDb rating. These should be obtained from the OMDb API, discussed below.
- Poster: your program should show an image of the movie's poster. A URL for the movie's poster should be obtained from the OMDb API, discussed below. Your program should use the API to obtain this link, then show image of the poster.
- Wordcloud: your program should download the movie's IMDb reviews from the MIS 515 API, discussed below. Your program should then generate and show a wordcloud based on all the downloaded IMDb reviews.

We have previously worked with the OMDb API in an in-class to obtain a summary of a movie's plot. The same procedure can be used on this assignment; you can use "https://www.omdbapi.com/?r=xml&apikey=yourkey&t=movie" to obtain data about a movie in XML format. Two parts of this URL must be filled in with appropriate details. First, replace "movie" in the URL with the name of the movie that the user would like to analyze. Second, replace "yourkey" with your personal API key for the OMDb API. To avoid the API being overloaded with too many requests, OMDb requires that each user sign up for their own unique "key" for their application. Signing up is free, and each key entitles the user to 1,000 requests per day, which should be more than enough for this assignment. You can sign up for a key here (make sure that you select the "free" option). You must obtain your own API key; do not use the API key used in the in-class example on this assignment.

Using your API key and movie title in the URL above, the OMDb API will provide background, reception, and poster data. For example, searching the API for "Inception" would return:

<?xml version="1.0" encoding="UTF-8"?>
<root response="True">
<movie title="Inception" year="2010" rated="PG-13" runtime="148 min" genre="Action, Adventure, Sci-Fi, Thriller" actors="Leonardo DiCaprio, Joseph Gordon-Levitt, Ellen Page, Tom Hardy" plot="A thief who steals corporate secrets through the use of dream-sharing technology is given the inverse task of planting an idea into the mind of a C.E.O." awards="Won 4 Oscars. Another 152 wins & 210 nominations."
poster="https://m.media-amazon.com/images/M/MV5BMjAxMzY3NjcxNF5BMl5BanBnXkFtZTcwNTI5OTM0Mw@@._V1_SX300.jpg" metascore="74" imdbRating="8.8" />
</root>

Included in the API's response is the link to an image of the movie's poster. Showing an image in Python is relatively simple using a combination of two tools, matplotlib and scikit-image by import wordcloud , matplotlib and skimage.io. Using these tools, you can show an image as follows, for example:

```
import wordcloud
import matplotlib.pyplot as plt
import skimage.io
url = "https://m.media-amazon.com/images/I/71ymSSSfouL._AC_SY741_.jpg"
image = skimage.io.imread(url)
plt.imshow(image, interpolation = "bilinear")
plt.axis("off")
plt.show()
```

To obtain reviews that need to be used for word cloud and sentiment analyses, use the https://developer.nytimes.com/ API. You can use the URL:

"https://api.nytimes.com/svc/movies/v2/reviews/search.json?query= movie&api-key=yourkey"

to obtain reviews for the movie of interest. Replace "movie" in the URL above with the name of the movie that the user would like to analyze. Also, you must obtain your own API key here. This API will return movie review data in JSON format. The response from this API is in the form of a JSON array. Below is an example:

{"status":"OK","copyright":"Copyright (c) 2021 The New York Times Company. All Rights Reserved.","has_more":false,"num_results":2,"results":[{"display_title":"Deadpool 2","mpaa_rating":"R","critics_pick":0,"byline":"A.O. SCOTT","headline":"Review: 'Deadpool 2' Has More Swearing, Slicing and Dicing From Ryan Reynolds","summary_short":"This sequel to the 2016 R-rated superhero hit pokes fun at the genre while staying true to its conventions.","publication_date":"2018-05-14","opening_date":"2018-05-18","date_updated":"2018-06-14 16:44:01","link":{"type":"article","url":"https://www.nytimes.com/2018/05/14/movies/deadpool-2-review-ryan-reynolds.html","suggested_link_text":"Read the New York Times Review of Deadpool 2"},"multimedia":{"type":"mediumThreeByTwo210","src":"https://static01.nyt.com/images/2018/05/18/arts/18deadpool/18deadpool-mediumThreeByTwo210.jpg","height":140,"width":210}},{"display_title":"Deadpool","mpaa_rating":"R","critics_pick":0,"byline":"MANOHLA DARGIS","headline":"Review: 'Deadpool,' a Sardonic Supervillain on a Kill Mission","summary_short":"This eager-to-please comic-book movie, starring Ryan Reynolds and directed by Tim Miller, is the latest entry in the Marvel universe.","publication_date":"2016-02-11","opening_date":"2016-02-12","date_updated":"2017-11-02 04:18:23","link":{"type":"article","url":"https://www.nytimes.com/2016/02/12/movies/deadpool-movie-review-ryan-reynolds.html","suggested_link_text":"Read the New York Times Review of Deadpool"},"multimedia":{"type":"mediumThreeByTwo210","src":"https://static01.nyt.com/images/2016/02/11/arts/deadpool1/deadpool1-mediumThreeByTwo210.jpg","height":140,"width":210}}]}

When creating wordclouds based on this API, use all the movie review data (summary_short) in JSON for the given movie to generate your wordcloud. There are no specific requirements for

the dimensions, colormap, or background color used in your wordcloud. The example on the following pages creates a 2,000 x 2,000 wordcloud with a "black" background color (default) and an "inferno" colormap. When computing sentiment scores based on this API, your program should report the average sentiment across all that movie's New York Times movie reviews. Both the polarity and subjectivity sentiment scores should be reported. Below is an example:
The summary-short content that represent the data (will be the text for the wordcload) for the above example:
This sequel to the 2016 R-rated superhero hit pokes fun at the genre while staying true to its conventions.
This eager-to-please comic-book movie, starring Ryan Reynolds and directed by Tim Miller, is the latest entry in the Marvel universe.
The polarity and subjectivity of "This sequel to the 2016 R-rated superhero hit pokes fun at the genre while staying true to its conventions" are:
Polarity: 0.32499999999999996
Subjectivity: 0.42500000000000004

The polarity and subjectivity of "This eager-to-please comic-book movie, starring Ryan Reynolds and directed by Tim Miller, is the latest entry in the Marvel universe" are:
Polarity: 0.3833333333333333
Subjectivity: 0.5833333333333334

So, the average sentiment across all that movie is:

Polarity: 0.35416666666666663
Subjectivity:0.5041666666666667

There are several different possibilities for mistakes and discrepancies in user inputs. One such discrepancy is capitalization. For example, the user could respond "yes", "Yes", or "YES" when asked if they would like to either analyze another movie or run another analysis; and they could respond "sentiment", "SeNtiment", or "SENTIMENT" when asked which analysis to run. Your program should handle all these possibilities case insensitively. Another possibility is that the user makes a spelling mistake, such as misspelling "yes" as "yess" or misspelling "sentiment" as "sentment." Use an appropriate module to correct the user's spelling for which analysis they would like to perform and for yes/no questions. If, even after correcting spelling, the user requests an analysis that is not supported (for example, "box office"), then print out a message stating that the analysis is not supported and asking the user to try again.

Your program should allow the user to search for as many movies as desired and to run as many analyses as desired. Once the user selects a movie and has run an analysis, ask them if they would like to keep analyzing the same movie. If they would, then ask the user which analysis they would like to run next and do not re-prompt them for which movie they would like to analyze (an example of this functionality is shown on the next page).

Some considerations to note:

- Consider the possibility that, when attempting connection to either API, some connection issue occurs (that is, a status code other than 200). This consideration is a serious issue on

this assignment because, although the OMDb API has data on thousands of films, maybe the requested film or its review is not available, ensure that your code handles this case and provides the user with a helpful printout if it does occur.

- Ensure that your prompts and output are crisp, professional, and well-formatted. For example, ensure that you have used spaces appropriately and checked your spelling.
- Adding comments in your code is encouraged. You may decide how best to comment your code. At minimum, please use a comment at the start of your code to describe its basic functionality.

Please use the following as a template for the tool's expected functionality:

Welcome to the movie analytics tool!
What movie would you like to analyze? Deadpool
What would you like to see (background/reception/poster/wordcloud/sentiment)? backgrnd
Year: 2016
Rating: R
Runtime: 108 min
Genre: Action, Adventure, Comedy, Sci-Fi
Actors: Ryan Reynolds, Karan Soni, Ed Skrein, Michael Benyaer
Plot: A wisecracking mercenary gets experimented on and becomes immortal but ugly, and sets out to track down the man who ruined his looks.
Would you like to further analyze this movie (yes/no)? Yess
What would you like to see (background/reception/poster/wordcloud/sentiment)? Reception
Awards: Nominated for 2 Golden Globes. Another 28 wins & 75 nominations.
Metascore: 65
IMDb rating: 8.0
Would you like to further analyze this movie (yes/no)? YES
What would you like to see (background/reception/poster/wordcloud/sentiment)? POSTER

Would you like to further analyze this movie (yes/no)? yes
What would you like to see (background/reception/poster/wordcloud/sentiment)? Wordcloud

Would you like to further analyze this movie (yes/no)? yes
What would you like to see (background/reception/poster/wordcloud/sentiment)? Sentimentt
Average New York Times review polarity: 0.35416666666666663
Average New York Times review subjectivity 0.5041666666666667
Would you like to further analyze this movie (yes/no)? yes
What would you like to see (background/reception/poster/wordcloud/sentiment)? box office
Sorry, that analysis is not supported. Please try again.
Would you like to further analyze this movie (yes/no)? no
Would you like to analyze another movie (yes/no)? yes
What movie would you like to analyze? casablanca
What would you like to see (background/reception/poster/wordcloud/sentiment)? Background
Year: 1942
Rating: PG
Runtime: 102 min
Genre: Drama, Romance, War
Actors: Humphrey Bogart, Ingrid Bergman, Paul Henreid, Claude Rains
Plot: A cynical American expatriate struggles to decide whether or not he should help his former lover and her fugitive husband escape French Morocco.
Would you like to further analyze this movie (yes/no)? yes
Would you like to further analyze this movie (yes/no)? no
Would you like to analyze another movie (yes/no)? no