**Part 2 Lab Data Analysis**

**Movies API**

Here is the list of IMDB IDs:

| **IMDBID** |
| --- |
| tt0120737 |
| tt0054135 |
| tt0414993 |
| tt0119643 |
| tt0816692 |
| tt0468569 |
| tt0137523 |
| tt1375666 |
| tt0133093 |
| tt0110413 |
| tt0371746 |
| tt2582802 |
| tt0482571 |
| tt1853728 |
| tt7286456 |
| tt0114709 |
| tt0152930 |
| tt1049413 |
| tt0266697 |
| tt0264464 |

1. Using <http://www.omdbapi.com/> explore site, learn request structure, and request your own free API key.
2. Use web connection in Power BI to connect to site and check first IMDBID. Use http://www.omdbapi.com/?apikey=yourkey&i=tt0120737 link with API key to get one record by IMDb ID in Power BI.
3. Closely explore the result (request, M code in advanced editor, applied steps, data).
4. Now when you got data for 1 IMDBID it is time to get data for every IMDBID. Power BI functions can help you with this, but first, you need to upload all the IMDBIDs to Power BI using ‘’Enter Data’’ functionality. Create table named ‘IDs’ with 1 column ‘IMDBID’ that contains all the IDs you can see above in this file. In the ‘IDs’ table add new custom column named ‘URL’ that will contain full URL to the IMDBID like in step #2 above. The result should look like this:

Table

Description automatically generated

1. Now it’s time to start creating a function that will send request to the site for each of these URLs. Using the query, you created on step 2, do the following:
   1. Open advanced editor
   2. Replace exact URL in first code step with some parameter – let’s call it ‘url’.
   3. Add new line in the beginning of the code with input parameter and its type, that will tell Power BI that your code is a function, not just data manipulation steps:

Graphical user interface, text, application, email

Description automatically generated

* 1. Rename query to ‘function for API’
  2. Check your function, it should look like on the picture below. Insert some URL for testing to see how it works.

Graphical user interface, application

Description automatically generated

1. Invoke function on ‘IDs’ table, it should return attributes for each URL in the table. Use ‘Invoke custom function’ functionality. Explore results using expand button of newly created column. Check data and data types. Having this step completed means that you have successfully pulled data from API connection.
2. Now it’s time to build a data model and proper visualizations to answer questions from this step. Transform data to create a data model that will help you to answer following questions:
   1. How many **Crime** movies are there? - 7
   2. What **genre** has the longest average runtime? - Fantasy
   3. What **actor/actors** has/have highest average number of IMDB Votes? - Aaron Eckhart, Heath Ledger.
   4. What are **top-3 countries** by number of movies released? – USA, Canada, UK
   5. What is the **average IMDB rating** for Director = “Christopher Nolan”? – 8.75
   6. What is the **average IMDB rating** for Language = “French”? – 8.18
   7. What is the **5th movie from top** with the highest Rating, where rating is a sum of “Internet Movie Database”, “Metacritic” and “Rotten Tomatoes” sources? It is recommended converting all ratings to the “Metacritic” rating type using this rule: “57/100” is 57, “6.6” is 66, “48%” is 48. - The Lord of the Rings: The Fellowship of the Ring
   8. What is the **month** with the highest number of released movies?-Nov’98
   9. Create a **table** with movie Poster as a picture, Title, Year, Genre, Country, Actors, IMDB rating, awards.

Create page/pages with visuals that would help any user to answer the questions above. Do not try to use DAX measures, create good data model instead.