

Exercise 1

Node, MongoDB and Express

By:

Pascal Labelle Poblette ID: 40131558

CART 451 Fall 2023

Sabine Rosenberg

The Dataset

The chosen dataset is the Netflix original films released in 2022, including movies, documentaries, and specials. I chose this data set because I was interested in finding out the amount of content that Netflix releases per year, where it comes from, if there are any types of content that is being invested in more, and other unexpected information I may find from this.

Some issues with the data were that the premiere dates were in string format, and written in a way that made it harder to perform queries using dates. Another issue was that the documentaries, movies, and specials were each in a different collection, so the *\$lookup* operator had to be used to query all three simultaneously. Finally, the data was not very thorough, as it did not include the director, writers, or any specific information about its production.

The queries

- **Query #1: Total number of Netflix original content in the year.**

This was a simple starting query to know how many pieces of content were released in that year, which was 251. Netflix released 170 movies, 59 documentaries and 22 specials. The specials include stand up comedy shows, short films and other events like speeches, tributes, and award shows.

- **Query #2: The number of movies with a runtime shorter than 2 hours.**

The total number of movies shorter than 2 hours is 138. This was to be expected, as shorter movies tend to be easier to watch and work better with at-home streaming services.

- **Query #3: The number of movies with a runtime longer than 2 hours.**

The total number of movies longer than 2 hours is 34. There is a significant difference compared to the longer movies, and I believe it is because Netflix benefits more from shorter movies, as they are cheaper to produce and it gives the user more time to watch more content.

- **Query #4: The top 5 languages that content was produced in.**

The top language in all 3 types of content was English. The top 5 in general was: English (135), Spanish (26), Turkish (11), and French, Polish and Hindi tied with 9 entries.

When it comes to movies and documentaries, Spanish was in second place, but the difference was large: 75 movies in English, 19 in Spanish. Counting documentaries, there were 43 in English and 7 in Spanish.

I was surprised to find out that Turkish was the 3rd most popular language when it comes to movies, as it is only the 11th most spoken language worldwide.

- **Query #5: The number of Romantic movies that came out around Valentine's Day (January & February)**

Out of the 32 Romantic movies that came out in 2022, 6 of them came out around the Valentine's day season, which is more than any other month, but still not as significant as I thought the results would be.

- **Query #6: The number of Horror or Thriller movies that came out around Halloween (September & October)**

The result of the Halloween query gave more significant results. Out of the 19 horror movies that came out in 2022, 8 were released between September and October. This shows how the dates of the content released is very influenced by certain holidays or times of the year where there are themed celebrations.

User Interface

The Netflix Dataset

Search Criteria:

Query Response: