SQL DATA ANALYSIS NETFLIX MOVIES

About this project

Objective:

The goal of this project is to understand the Netflix titles catalogue by performing a content analysis based on titles type such as movies, tv shows across different countries.

But also highlight the most popular content that Netflix can use to build their marketing strategies.

Key Metrics to be Analyzed:

How many titles do we count by Directors / Country?

How many titles do Netflix released this past year?

Number of movies by Countries and type over years

The most popular titles by countries over years

About the Dataset:

> **Total Entries:** The dataset contains 8,809 entries, each representing a unique movie or TV show.
> - **Columns:** There are 12 columns in the dataset:
>   1. **show_id:** A unique identifier for each title.
>   2. **type:** The category of the title, which is either 'Movie' or 'TV Show'.
>   3. **title:** The name of the movie or TV show.
>   4. **director:** The director(s) of the movie or TV show. (Contains null values for some entries, especially TV shows where this information might not be applicable.)
>   5. **cast:** The list of main actors/actresses in the title. (Some entries might not have this information.)
>   6. **country:** The country or countries where the movie or TV show was produced.
>   7. **date_added:** The date the title was added to Netflix.
>   8. **release_year:** The year the movie or TV show was originally released.
>   9. **rating:** The age rating of the title.
>   10. **duration:** The duration of the title, in minutes for movies and seasons for TV shows.
>   11. **listed_in:** The genres the title falls under.
>   12. **description:** A brief summary of the title

Content Analysis

1- Total number of movies by type over time

SELECT count(title) AS [total_number_titles],type
from movies_titles
GROUP BY type
ORDER BY total_number_titles DESC

total_number_titles	type
6132	Movie
2677	TV Show

But we also want to see the total number of movies over years

SELECT COUNT(title) AS [number_title_by_type],type,release_year
from movies_titles
GROUP BY type,release_year
order BY number_title_by_type DESC

number_title_by_type	type	release_year
767	Movie	2017
767	Movie	2018
658	Movie	2016
633	Movie	2019
517	Movie	2020
436	TV Show	2020
398	Movie	2015
397	TV Show	2019
380	TV Show	2018

Insight: Over the year Netflix has released more movies than TV shows and the number of movies released is increasing more over the year however in 2020 the number of TV shows was higher.

2- Total number of movies by Director

SELECT COUNT(title) AS [number_movies_by_director],director
From movies_titles
GROUP BY director
ORDER BY number_movies_by_director DESC

number_movies_by_director	director
2634	*NULL*
19	Rajiv Chilaka
18	Raúl Campos, Jan Suter
16	Suhas Kadav
16	Marcus Raboy
14	Jay Karas

Insight: we have more than 70 % of movies released without director name. this is due to a lack of data quality.

3- Total number of movies by country / by type

```sql
SELECT COUNT(country) AS [number_movies_by_country],country
From movies_titles
GROUP BY country
ORDER BY number_movies_by_country DESC
```

number_movies_by_country	country
2819	United States
972	India
419	United Kingdom
245	Japan
200	South Korea
181	Canada

```sql
SELECT COUNT(country) AS [number_movies_by_country],country,type
From movies_titles
GROUP BY country,type
ORDER BY number_movies_by_country DESC
```

number_movies_by_country	country	type
2059	United States	Movie
893	India	Movie
760	United States	TV Show
213	United Kingdom	TV Show
206	United Kingdom	Movie
169	Japan	TV Show
159	South Korea	TV Show
122	Canada	Movie
97	Spain	Movie

Insights: USA is the top country with the most movies on Netflix with India ranked as 2nd . we can see that in these 2 countries movies has the highest numbers on the platform however in UK and Japan TV shows also got highest number.

4- Genre popularity

SELECT COUNT(country) AS [number_popular_title],listed_in as[genre]
From movies_titles
GROUP BY listed_in
ORDER BY number_popular_title DESC

number_popular_title ⌄	genre ⌄
343	Documentaries
337	Dramas, International Movies
303	Stand-Up Comedy
260	Comedies, Dramas, International Movies
243	Dramas, Independent Movies, International Movies
181	Children & Family Movies, Comedies
176	Kids' TV
165	Documentaries, International Movies
163	Dramas, International Movies, Romantic Movies

Let's add a query that will group the popular genre by country over the year

SELECT COUNT(title) AS [total_movies_by_genre_country],release_year,listed_in,country
From movies_titles
where country is not NULL
GROUP BY country , listed_in,release_year
ORDER BY total_movies_by_genre_country DESC

total_movies_by_genre_country ⌄	release_year ⌄	listed_in ⌄	country ⌄
61	2017	Documentaries	United States
46	2016	Documentaries	United States
41	2018	Documentaries	United States
37	2017	Stand-Up Comedy	United States
26	2018	Stand-Up Comedy	United States
26	2019	Stand-Up Comedy	United States
25	2019	Documentaries	United States
25	2018	Dramas, Independent Movies, International Movies	India
21	2017	Dramas, International Movies	India
21	2020	Stand-Up Comedy	United States
20	2016	Stand-Up Comedy	United States
19	2017	Dramas, Independent Movies, International Movies	India
18	2020	Documentaries	United States
17	2020	Reality TV	United States
15	2015	Stand-Up Comedy	United States

Insight: the most popular genre is Documentary, stand up in USA while in India, Drama movies are leading.

Conclusion:

After analysis, we can clearly see that the number of movies is clearly above TV shows however in some country like UK and Japan TV shows are mostly released. We also evaluate the most popular genre in the top country and we found that Documentary and stand-up comedy are most popular in USA while dramas movies is most popular in India. Considering these 2 countries are the top larger consumer in movies Netflix should include in their marketing strategies effort on these countries while also include country like UK and Japan.