

Overview

This project encompasses a thorough analysis of Netflix's movies and TV shows data utilizing SQL. The aim is to derive meaningful insights and address a range of business inquiries based on the dataset. The subsequent README offers an in-depth overview of the project's goals, the business challenges encountered, the solutions implemented, as well as the findings and conclusions drawn.

Objectives

- Analyze the distribution of content types (movies vs TV shows).
- Identify the most common ratings for movies and TV shows.
- List and analyze content based on release years, countries, and durations.
- Explore and categorize content based on specific criteria and keywords.

Dataset

The data for this project is sourced from the **Kaggle**

/* This query retrieves all columns and rows from the netflix_titles table to examine the dataset.*/

SELECT *

FROM netflix_titles;

	show	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
1	s1	Movie	Dick Johnson Is	Kirsten Johnson	NULL	United St	2021-09-25	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, f
2	s2	TV Sh	Blood & Water	NULL	Ama Qamata, K	South Afri	2021-09-24	2021	TV-MA	2 Seasons	International TV S	After crossing paths at a party, a Cape
3	s3	TV Sh	Ganglands	Julien Leclercq	Sami Bouajila,	NULL	2021-09-24	2021	TV-MA	1 Season	Crime TV Shows, I	To protect his family from a powerful dr
4	s4	TV Sh	Jailbirds New Or	NULL	NULL	NULL	2021-09-24	2021	TV-MA	1 Season	Docuseries, Realit	Feuds, flirtations and toilet talk go down
5	s5	TV Sh	Kota Factory	NULL	Mayur More, Jit	India	2021-09-24	2021	TV-MA	2 Seasons	International TV S	In a city of coaching centers known to
6	s6	TV Sh	Midnight Mass	Mike Flanagan	Kate Siegel, Za	NULL	2021-09-24	2021	TV-MA	1 Season	TV Dramas, TV Ho	The arrival of a charismatic young pries
7	s7	Movie	My Little Pony:	Robert Cullen,	Vanessa Hudg	NULL	2021-09-24	2021	PG	91 min	Children & Family	Equestria's divided. But a bright-eyed h
8	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba,	United St	2021-09-24	1993	TV-MA	125 min	Dramas, Independ	On a photo shoot in Ghana, an Americ '
<												>

-- 1. Count the number of Movies vs TV Shows

SELECT type,COUNT(*) AS Number_TVShows_or_Movie
FROM netflix_titles
GROUP BY type

	type	Number_TVShows_or_Movie
1	TV Show	2676
2	Movie	6131

-- 2. Find the most common rating for movies and TV shows

```
-- Use of CTEs

WITH RatingCounts AS (

SELECT type, rating, COUNT(*) AS rating_count

FROM netflix_titles

GROUP BY type, rating
),

RankedRatings AS (

SELECT type, rating, rating_count,

RANK() OVER (PARTITION BY type ORDER BY rating_count DESC) AS rank

FROM RatingCounts
)

SELECT type, rating AS most_frequent_rating
```

FROM RankedRatings

WHERE rank = 1;

	type	most_frequent_rating
1	Movie	TV-MA
2	TV Show	TV-MA

-- 3. List all movies released in a specific year

SELECT *

FROM netflix_titles

 $WHERE\ release_year = 2021$

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in
1	s2	TV Show	Blood & Water	NULL	Ama Qamata, Kho	South Africa	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows, TV Dran
2	s3	TV Show	Ganglands	Julien Lec	Sami Bouajila, Tra	NULL	2021-09-24	2021	TV-MA	1 Season	Crime TV Shows, International T
3	s4	TV Show	Jailbirds New	NULL	NULL	NULL	2021-09-24	2021	TV-MA	1 Season	Docuseries, Reality TV
4	s5	TV Show	Kota Factory	NULL	Mayur More, Jiten	India	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows, Romanti
5	s6	TV Show	Midnight Mass	Mike Flan	Kate Siegel, Zach	NULL	2021-09-24	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Myste
6	s7	Movie	My Little Pony	Robert Cu	Vanessa Hudgens	NULL	2021-09-24	2021	PG	91 min	Children & Family Movies
7	s9	TV Show	The Great Brit	Andy Dev	Mel Giedroyc, Sue	United Kingdom	2021-09-24	2021	TV-14	9 Seasons	British TV Shows, Reality TV
8	s10	Movie	The Starling	Theodore	Melissa McCarthy,	United States	2021-09-24	2021	PG-13	104 min	Comedies, Dramas

-- 4. Find the top 5 countries with the most content on Netflix

SELECT TOP 5

value AS country,

COUNT(*) AS Total_Content

FROM netflix_titles

CROSS APPLY string_split(country, ',')

WHERE value IS NOT NULL

GROUP BY value

ORDER BY Total_Content DESC;

	country	Total_Content
1	United States	3211
2	India	1008
3	United Kingdom	628
4	United States	479
5	Canada	271

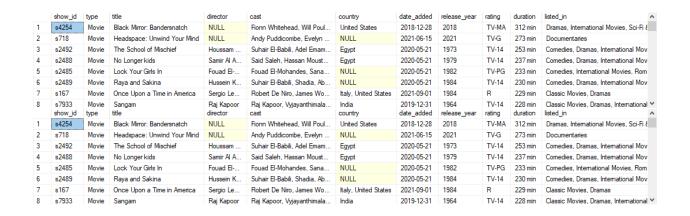
-- 5. Identify the longest movie

SELECT *

FROM netflix_titles

WHERE type = 'Movie'

ORDER BY CAST(LEFT(duration, CHARINDEX(' ',duration) - 1) AS INT) DESC;



-- 6. Find content added in the last 5 years

SELECT *

FROM netflix_titles

WHERE TRY_CONVERT(DATE, date_added, 101) >= DATEADD(YEAR, -5, GETDATE());

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in /
1	s1	Movie	Dick Johnson Is Dead	Kirsten J	NULL	United States	2021-09-25	2020	PG-13	90 min	Documentaries
2	s2	TV Show	Blood & Water	NULL	Ama Qamata, Khosi N	South Africa	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows, TV Dr.
3	s3	TV Show	Ganglands	Julien Le	Sami Bouajila, Tracy	NULL	2021-09-24	2021	TV-MA	1 Season	Crime TV Shows, International
4	s4	TV Show	Jailbirds New Orleans	NULL	NULL	NULL	2021-09-24	2021	TV-MA	1 Season	Docuseries, Reality TV
5	s5	TV Show	Kota Factory	NULL	Mayur More, Jitendra	India	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows, Roman
6	s6	TV Show	Midnight Mass	Mike Fla	Kate Siegel, Zach Gilf	NULL	2021-09-24	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mys
7	s7	Movie	My Little Pony: A New Generation	Robert C	Vanessa Hudgens, Ki	NULL	2021-09-24	2021	PG	91 min	Children & Family Movies
8	s8	Movie	Sankofa	Haile Ge	Kofi Ghanaba, Oyafun	United States, Ghana,	2021-09-24	1993	TV-MA	125 min	Dramas, Independent Movies, *

-- 7. Find all the movies/TV shows by director 'Rajiv Chilaka'

SELECT *

FROM netflix_titles

CROSS APPLY string_split(director, ',')

WHERE director = 'Rajiv Chilaka'

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	li
10	s416	Movie	Chhota Bheem aur Krishna	Rajiv Chilaka	Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jigna	India	2021-07-22	2009	TV-Y7	68 min	1
11	s417	Movie	Chhota Bheem aur Krishna vs Zimbara	Rajiv Chilaka	Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jigna	NULL	2021-07-22	2013	TV-Y7	64 min	1
12	s420	Movie	Chhota Bheem: Bheem vs Aliens	Rajiv Chilaka	Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jigna	NULL	2021-07-22	2010	TV-Y7	69 min	1
13	s421	Movie	Chhota Bheem: Dholakpur to Kathmandu	Rajiv Chilaka	Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jigna	NULL	2021-07-22	2012	TV-Y7	70 min	1
14	s423	Movie	Chhota Bheem: Journey to Petra	Rajiv Chilaka	Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jigna	NULL	2021-07-22	2011	TV-Y7	68 min	1
15	s424	Movie	Chhota Bheem: Master of Shaolin	Rajiv Chilaka	Pooja Punabi, Mausam, Julie Tejwani, Rupa Bhim	NULL	2021-07-22	2011	TV-Y7	74 min	1
16	s425	Movie	Chhota Bheem: The Rise of Kirmada	Rajiv Chilaka	Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jigna	NULL	2021-07-22	2012	TV-Y7	68 min	1
17	s2718	Movie	Chhota Bheem and the Curse of Damyaan	Rajiv Chilaka	Kaustav Ghosh, Jigna Bhardwaj, Chutki, Rajesh	India	2020-04-01	2012	TV-Y7	87 min	1

-- 8. List all TV shows with more than 5 seasons

SELECT *

FROM netflix_titles

WHERE

TYPE = 'TV Show'

AND

 $CAST(LEFT(duration, CHARINDEX('', duration) - 1) \ AS\ INT) > 5$

ORDER BY CAST(LEFT(duration, CHARINDEX(' ', duration) -1) AS INT) DESC;

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in
1	s549	TV Show	Grey's Anat	NULL	Ellen Pompeo, San	United States	2021-07-03	2020	TV-14	17 Seasons	Romantic TV Shows, TV Dramas
2	s2424	TV Show	Supematural	Phil Sgriccia	Jared Padalecki, J	United States, Canada	2020-06-05	2019	TV-14	15 Seasons	Classic & Cult TV, TV Action & Adventure, TV Ho
3	s4799	TV Show	NCIS	NULL	Mark Harmon, Mic	United States	2018-07-01	2017	TV-14	15 Seasons	Crime TV Shows, TV Dramas, TV Mysteries
4	s4221	TV Show	COMEDIAN	NULL	Neal Brennan, Chri	United States	2019-01-01	2019	TV-MA	13 Seasons	Stand-Up Comedy & Talk Shows, TV Comedies
5	s1355	TV Show	Heartland	NULL	Amber Marshall, Mi	Canada	2021-02-01	2019	TV-14	13 Seasons	TV Dramas
6	s7848	TV Show	Red vs. Blue	NULL	Bumie Bums, Jaso	United States	NULL	2015	NR	13 Seasons	TV Action & Adventure, TV Comedies, TV Sci-Fi
7	s4965	TV Show	Trailer Park	NULL	Mike Smith, John P	Canada	2018-03-30	2018	TV-MA	12 Seasons	Classic & Cult TV, Crime TV Shows, International
8	s5413	TV Show	Criminal Minds	NULL	Mandy Patinkin, Jo	United States, Canada	2017-06-30	2017	TV-14	12 Seasons	Crime TV Shows, TV Dramas, TV Mysteries

-- 9. Count the number of content items in each genre

SELECT value AS GERNE, COUNT(*) AS Total_Content

FROM netflix_titles

CROSS APPLY string_split(listed_in, ',')

GROUP BY value

ORDER BY Total_Content DESC;

	GERNE	Total_Content
1	International Movies	2624
2	Dramas	1600
3	Comedies	1210
4	Action & Adventure	859
5	Documentaries	829
6	Dramas	827
7	International TV Shows	774
8	Independent Movies	736
9	TV Dramas	696

- 10. Find each year and the average numbers of content release by India on netflix.

-- return top 5 year with highest avg content release

SELECT TOP 5

country,

release_year,

COUNT(show_id) as Total_Release,

CAST(COUNT(show_id) as decimal (10,2)) as Avg_Release

FROM netflix_titles

WHERE country = 'India'

GROUP BY country, release_year

ORDER BY Avg_Release DESC;

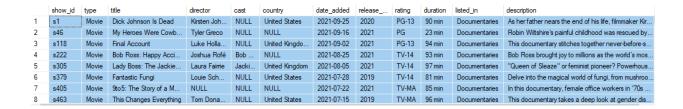
	country	release_year	Total_Release	Avg_Release
1	India	2017	101	101.00
2	India	2018	94	94.00
3	India	2019	87	87.00
4	India	2020	75	75.00
5	India	2016	73	73.00

-- 11. List all movies that are documentaries

SELECT*

FROM netflix_titles

WHERE listed_in = 'Documentaries'



-- 12. Find all content without a director

SELECT*

FROM netflix_titles

WHERE director IS NULL

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in
1	s2	TV Show	Blood & Water	NULL	Ama Qamata, Khosi Ngema, G	South Africa	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysterie:
2	s4	TV Show	Jailbirds New Orle	NULL	NULL	NULL	2021-09-24	2021	TV-MA	1 Season	Docuseries, Reality TV
3	s5	TV Show	Kota Factory	NULL	Mayur More, Jitendra Kumar, R	India	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV
4	s11	TV Show	Vendetta: Truth, Li	NULL	NULL	NULL	2021-09-24	2021	TV-MA	1 Season	Crime TV Shows, Docuseries, International TV Sh
5	s500	TV Show	Dogs	NULL	NULL	United States	2021-07-07	2021	TV-PG	2 Seasons	Docuseries
6	s15	TV Show	Crime Stories: Indi	NULL	NULL	NULL	2021-09-22	2021	TV-MA	1 Season	British TV Shows, Crime TV Shows, Docuseries
7	s16	TV Show	Dear White People	NULL	Logan Browning, Brandon P. B	United States	2021-09-22	2021	TV-MA	4 Seasons	TV Comedies, TV Dramas
8	s18	TV Show	Falsa identidad	NULL	Luis Ernesto Franco, Camila S	Mexico	2021-09-22	2020	TV-MA	2 Seasons	Crime TV Shows, Spanish-Language TV Shows,

-- 13. Find how many movies actor 'Salman Khan' appeared in last 10 years!

SELECT*

FROM netflix_titles

WHERE

cast LIKE '%Salman Khan%'

AND

release_year > YEAR(GETDATE()) - 10



- 14. Find the top 10 actors who have appeared in the highest number of movies produced in India.

SELECT TOP 10

cast,

COUNT(*) AS Total_Count

FROM netflix_titles

CROSS APPLY string_split(cast, ',') as Split_Casts

WHERE country = 'India'

GROUP BY cast

ORDER BY Total_Count DESC

	cast	Total_Count
1	Suriya, Vijay Sethupathi, Revathy, Prakash Raj, Siddarth,	22
2	Sanjay Dutt, Govinda, Karisma Kapoor, Pooja Batra, Kad	18
3	Shah Rukh Khan, Amrish Puri, Mahima Chaudhry, Apoorv	18
4	Anupam Kher, Atul Kulkami, Aakash Pandey, Milind Gun	17
5	Manchu Lakshmi, Saanve Megghana, Naveen Kumar, J	17
6	Rahul Pethe, Mrunmayee Deshpande, Adinath Kothare,	16
7	Anil Kapoor, Shah Rukh Khan, Jackie Shroff, Satyendra	15
8	Nawazuddin Siddiqui, Radhika Apte, Khalid Tyabji, Aditya	15
9	Aftab Shivdasani, Antara Mali, Boman Irani, Isha Koppika	14

-- Question 15:

Categorize the content based on the presence of the keywords 'kill' and 'violence' in the description field. Label content containing these keywords as 'Bad' and all other content as 'Good'. Count how many items fall into each category. --

SELECT

Category,

TYPE,

COUNT(*) AS content_count

FROM (

SELECT*,

CASE

WHEN LOWER(description) LIKE '%kill%' OR LOWER(description) LIKE '%violence%' THEN 'Bad'

ELSE 'Good'

END AS category

FROM netflix_titles

) as Categorized_content

GROUP BY category, TYPE

ORDER BY TYPE DESC;

	Category	TYPE	content_count
1	Bad	TV Show	91
2	Good	TV Show	2585
3	Bad	Movie	251
4	Good	Movie	5880

Findings and Conclusion

Content Diversity: The dataset showcases a wide range of movies and TV shows across various ratings and genres.

Audience Insights: Identifying common ratings reveals the target audience for the content.

Regional Trends: The analysis highlights leading countries and content releases in India, illustrating regional preferences.

Content Classification: Categorizing content by keywords enhances understanding of Netflix's offerings.

END