

Overview

The project involves a comprehensive analysis of netflix movies. and tv shows using sql . The goal is to extract valuable insights and answer various business questions based on the dataset.

Objective
















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.

ERD for netflix_db



	
	public
	netflix
	show_id character varying (5)
	type character varying(10)
	title character varying(250)
	director character varying(5 50)
	casts character varying(10 50)
	country character varying(5 50)
	date_added character varyi ng(55)
	release_year integer
	rating character varying(15)
	duration character varying (15)
	listed_in character varying (250)
	description character varyi ng(550)

SCHEMA

```
DROP TABLE IF EXISTS netflix;  
CREATE TABLE netflix  
(  
    show_id    VARCHAR(5),  
    type       VARCHAR(10),  
    title      VARCHAR(250),  
    director   VARCHAR(550),  
    casts      VARCHAR(1050),  
    country    VARCHAR(550),  
    date_added VARCHAR(55),  
    release_year INT,  
    rating     VARCHAR(15),  
    duration   VARCHAR(15),  
    listed_in  VARCHAR(250),  
    description VARCHAR(550)  
);
```

Count the number of Movies vs TV Shows

```
select type, count(type) as Total_shows
      from netflix
group by type
```

	type character varying (10) 	total_shows bigint 
1	Movie	6131
2	TV Show	2676

6	Akshay Kumar	22
7	Amitabh Bachchan	20
8	Naseeruddin Shah	19
9	Kareena Kapoor	19
10	Asrani	17

List all movies released in a specific year (e.g., 2020)

```
select title as movie from netflix
where type='Movie' and release_year = '2020'
```

	movie character varying (250)
1	Dick Johnson Is Dead
2	Europe's Most Dangerous Man: Otto Skorzeny in Spa
3	Tughlaq Durbar
4	Omo Ghetto: the Saga
5	Shadow Parties
6	Here and There
7	Shikara
8	Really Love
9	The Old Ways


Find the top 5 countries with the most content on Netflix

```
select unnest(string_to_array(country, ',')) as countries
,count(show_id) as total_content
from netflix
group by countries
order by 2 desc
limit 5
```

	countries text	total_content bigint
1	United States	3211
2	India	1008
3	United Kingd...	628
4	United States	479
5	Canada	271

Identify the longest movie

```
select TITLE AS MOVIE_NAME,DURATION from netflix
where
    type='Movie'
    AND
    duration=(select max(duration) from netflix)
```

movie_name character varying (250) 	duration character va
InuYasha the Movie 2: The Castle Beyond the Looking Glass	99 min
InuYasha the Movie 3: Swords of an Honorable Ruler	99 min
Here and There	99 min
Good Luck Chuck	99 min
My Girl 2	99 min
Cousins	99 min
Rogue Warfare: Death of a Nation	99 min
Copenhagen	99 min


Find all the movies/TV shows by director 'Rajiv Chilaka'!

```
SELECT type,title,DIRECTOR FROM netflix
where
director='Rajiv Chilaka'
```

type character varying (10) 🔒	title character varying (250) 🔒	director character var
Movie	Chhota Bheem - Neeli Pahaadi	Rajiv Chilaka
Movie	Chhota Bheem & Ganesh	Rajiv Chilaka
Movie	Chhota Bheem & Krishna: Mayanagari	Rajiv Chilaka
Movie	Chhota Bheem & Krishna: Pataliputra- City of the Dead	Rajiv Chilaka
Movie	Chhota Bheem And The Broken Amulet	Rajiv Chilaka
Movie	Chhota Bheem And The Crown of Valhalla	Rajiv Chilaka


List all TV shows with more than 5 seasons

```
select TYPE,TITLE from netflix
where
    type='TV Show' AND
    split_part(duration,' ',1)::numeric > 5
```

	<div>type</div> <div>character varying (10) </div>	<div>title</div> <div>character varying (250)</div>
1	TV Show	The Great British Baking Show
2	TV Show	Nailed It
3	TV Show	Numberblocks
4	TV Show	Saved by the Bell
5	TV Show	Lucifer
6	TV Show	Grace and Frankie
7	TV Show	30 Rock
8	TV Show	Hunter X Hunter (2011)
9	TV Show	The Flash

List all movies that are documentaries

```
select TYPE, TITLE from netflix
where
    listed_in
    ILIKE '%documentaries'
```

	type character varying (10) 	title character varying (250)
1	Movie	Dick Johnson Is Dead
2	Movie	My Heroes Were Cowboys
3	Movie	Final Account
4	Movie	NOVA: Chinese Chariot Revealed
5	Movie	Bob Ross: Happy Accidents, Betraya
6	Movie	Lady Boss: The Jackie Collins Story
7	Movie	NOVA: Day the Dinosaurs Died
8	Movie	Fantastic Fungi
9	Movie	9to5: The Story of a Movement

Count number of content items in each genre

```
select unnest(string_to_array(listed_in, ','))  
as genre, count(show_id)  
from netflix  
group by genre
```

	text	bigint
1	International TV Shows	774
2	Stand-Up Comedy	9
3	Spanish-Language TV Shows	172
4	Romantic Movies	613
5	Anime Features	21
6	TV Horror	11
7	Stand-Up Comedy & Talk Shows	34
8	Kids' TV	63
9	Docuseries	221
10	Children & Family Movies	605
11	Movies	57



Find all content without a director

```
select * from netflix
where
    director IS NULL
```

type character varying	title character varying (250)		director character varying
TV Show	Blood & Water		[null]
TV Show	Jailbirds New Orleans		[null]
TV Show	Kota Factory		[null]
TV Show	Vendetta: Truth, Lies and The Mafia		[null]
TV Show	Crime Stories: India Detectives		[null]
TV Show	Dear White People		[null]
TV Show	Khawatir		[null]
TV Show	Falsa identidad		[null]
TV Show	Jaguar		[null]
TV Show	Resurrection: Ertugrul		[null]



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

```
SELECT show_id,type,title FROM netflix
where
    casts ILIKE '%Salman Khan%'
    AND
release_year > EXTRACT(YEAR FROM CURRENT_DATE)-10
```

show_id character varying (5) 	type character varying (10) 	title character varying (250)
s2757	Movie	Paharganj

Find the most common rating for movies and TV shows

```
WITH RatingCounts AS (  
    SELECT  
        type,  
        rating,  
        COUNT(*) AS rating_count  
    FROM netflix  
    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 character varying (10) 	most_frequent_rating character varying (15) 
1	Movie	TV-MA
2	TV Show	TV-MA

Find content added in the last 5 years

```
select title,date_added from netflix
where
TO_DATE(date_added, 'month dd yyyy')>
=CURRENT_DATE-INTERVAL '5 YEARS'
```

title character varying (250)	date_added character varying (55)
Dick Johnson Is Dead	September 25, 2021
Blood & Water	September 24, 2021
Ganglands	September 24, 2021
Jailbirds New Orleans	September 24, 2021
Kota Factory	September 24, 2021
Midnight Mass	September 24, 2021
My Little Pony: A New Generation	September 24, 2021
Sankofa	September 24, 2021
The Great British Baking Show	September 24, 2021
The Starling	September 24, 2021

Categorize Content Based on the Presence of 'Kill' and 'Violence' Keywords

```
SELECT
    category,
    COUNT(*) AS content_count
FROM (
    SELECT
        CASE
            WHEN description ILIKE '%kill%' OR
            description ILIKE '%violence%' THEN 'Bad'
            ELSE 'Good'
        END AS category
    FROM netflix
) AS categorized_content
GROUP BY category;
```

	category  text	content_count  bigint
1	Good	8465
2	Bad	342

CONCLUSION

Content Distribution: The dataset contains a diverse range of movies and TV shows with varying ratings and genres.

Common Ratings: Insights into the most common ratings provide an understanding of the content's target audience.

Geographical Insights: The top countries and the average content releases by India highlight regional content distribution.

Content Categorization: Categorizing content based on specific keywords helps in understanding the nature of content available on Netflix.

This analysis provides a comprehensive view of Netflix's content and can help inform content strategy and decision-making.