SQL Query Analysis Report – Netflix Data

Overview:

This SQL script analyzes a **Netflix dataset** to address **15 business questions**, helping derive insights around content type, ratings, countries, directors, genres, and viewer behavior. It's structured to answer real-world analytical questions commonly faced by streaming platforms.

Business Problems & Query Analysis:

1. Movies vs TV Shows Count

```
select type,count(*) as Total_content from Netflix
group by type;
```

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

2. Most Common Rating by Type

```
select type,rating
from
 (select type,rating, count(), rank() over(partition by type order by count() desc) as ranking
from Netflix group by 1,2) as t1
 where ranking = 1;
```

	type character varying (10)	rating character varying (10)
1	Movie	TV-MA
2	TV Show	TV-MA

3. Movies Released in 2020

```
select type,title,release_year from Netflix

where type = 'Movie' and release_year = 2020;
```

	type character varying (10)	title character varying (150)	release_year integer
1	Movie	Dick Johnson Is Dead	2020
2	Movie	Europe's Most Dangerous Man: Otto Skorzeny in Spain	2020
3	Movie	Tughlaq Durbar	2020
4	Movie	Omo Ghetto: the Saga	2020
5	Movie	Shadow Parties	2020
6	Movie	Here and There	2020
7	Movie	Shikara	2020
8	Movie	Really Love	2020
9	Movie	The Old Ways	2020

4. Top 5 Countries with Most Content

```
select unnest(string_to_array(country,',')) as new_country,

count(show_id) as Total_content

from Netflix group by 1 order by 2 desc limit 5;
```

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

5. Longest Movie

```
select * from Netflix

where type = 'Movie' and duration = (select max(duration) from Netflix);
```

- Retrieves the movie with the longest duration.

6. Content Added in Last 5 Years

```
select * from Netflix

where to_date(date_added,'Month DD, YYYY')>= current_date - interval '5 years';
```

- Filters content added in the last 5 years using proper date parsing.

7. Content by Rajiv Chilaka

```
select * from Netflix

where director = 'Rajiv Chilaka';
```

- Filters content directed by a specific individual.

8. TV Shows with More Than 5 Seasons

```
select * from Netflix

where type = 'TV Show' and split_part(duration,' ',1):: numeric > 5;
```

- Parses the number of seasons from the duration column.

9. Count per Genre

```
select
```

```
 unnest(string_to_array(listed_in,',')) as genre, count(show_id) as Total_content
from Netflix
group by 1;
```

	genre text	total_content 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. Average Content Released in India

```
select  extract(year from to_date(date_added,'Month DD, YYYY')) as year,
count() as yearly_content,
cast( count():: numeric/(select count(*) from Netflix where country = 'India')*100:: numeric
 as decimal(10,2) )as avg_content_per_year from Netflix
where country = 'India'
group by 1
order by 3 desc limit 5;
```

	year numeric	yearly_content bigint	avg_content_per_year numeric (10,2)
1	2018	333	34.26
2	2019	203	20.88
3	2020	189	19.44
4	2017	142	14.61
5	2021	95	9.77

11. Documentary Movies

```
select * from Netflix

where listed_in ilike '%documentaries';
```

- Filters movies containing “documentaries” in their genre field.

12. Content Without Director

```
select * from Netflix

where director is null;
```

- Finds entries missing director info (data quality).

13. Salman Khan Appearances in Last 10 Years

```
select * from Netflix

 where

casts ilike '%salman khan%' and release_year> extract(year from current_date) - 10;
```

- Tracks appearances of a specific actor over a decade.

14. Top 10 Actors in Indian Movies

```
select unnest(string_to_array(casts,',')) as actors,
count(*) as Total_content
from Netflix
where country ilike '%india%'
 group by 1
order by 2 desc limit 10;
```

	actors text	total_content bigint
1	Anupam Kher	36
2	Om Puri	26
3	Boman Irani	25
4	Paresh Rawal	25
5	Shah Rukh Khan	25
6	Akshay Kumar	23
7	Naseeruddin Shah	20
8	Amitabh Bachchan	20
9	Kareena Kapoor	20

15. Categorize by Keywords in Description

```
with new_table as
(select *,
case
```

```
   when
   description ilike '%kill%' or
description ilike '%violence%' then 'Bad_content'
else 'Good_content'
end category
from Netflix
)
select
    category,
 count(*) as Total_content
from new_table
group by 1;
```

	category text	total_content bigint
1	Good_content	8465
2	Bad_content	342