NETFLIX DATA ANALYSIS USING SQL

MADE BY :

AMIT VERMA

Project Overview

This project focuses on analyzing Netflix's vast dataset to gain insights into its content, user preferences, and operational strategies. Using SQL, we addressed 12 business-critical problems that provide a deeper understanding of Netflix's performance across various dimensions.By leveraging data-driven solutions, this project offers actionable insights to optimize content offerings, improve customer retention, and enhance overall business strategies.

```
CREATE TABLE netflix

(

    show_id  varchar(6),
    type     varchar(10),
    title    varchar(150),
    director    varchar(208),
    casts    varchar(1000),
    country varchar(150),
    date_added  varchar(50),
    release_year     int,
    rating  varchar(10),
    duration varchar(15),
    listed_in   varchar(100),
    description varchar(250)

)
```

1. Count the number of Movies vs TV Shows

```sql
SELECT
    type,
    COUNT(show_id) as total_content
FROM netflix
GROUP BY type;
```

Data Output | Messages | Notifications

SQL

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

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

```sql
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
```

Data Output | Messages | Notifications

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

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

```sql
SELECT * FROM netflix
WHERE
    type = 'Movie'
    AND
    release_year = 2020
```

Data Output | Messages | Notifications

	show_id character varying (6)	type character varying (10)	title character varying (150)	director character varying (208)
1	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson
2	s17	Movie	Europe's Most Dangerous Man: Otto Skorzeny in Spain	Pedro de Echave García, Pablo Azorín Williams
3	s79	Movie	Tughlaq Durbar	Delhiprasad Deenadayalan
4	s85	Movie	Omo Ghetto: the Saga	JJC Skillz, Funke Akindele
5	s104	Movie	Shadow Parties	Yemi Amodu
6	s120	Movie	Here and There	JP Habac
7	s127	Movie	Shikara	Vidhu Vinod Chopra
8	s228	Movie	Really Love	Angel Kristi Williams
9	s230	Movie	The Old Ways	Christopher Alender
10	s251	Movie	A Faraway Land	Veronica Velasco

Total rows: 517 of 517 | Query complete 00:00:00.241 | Ln 63, Col 1

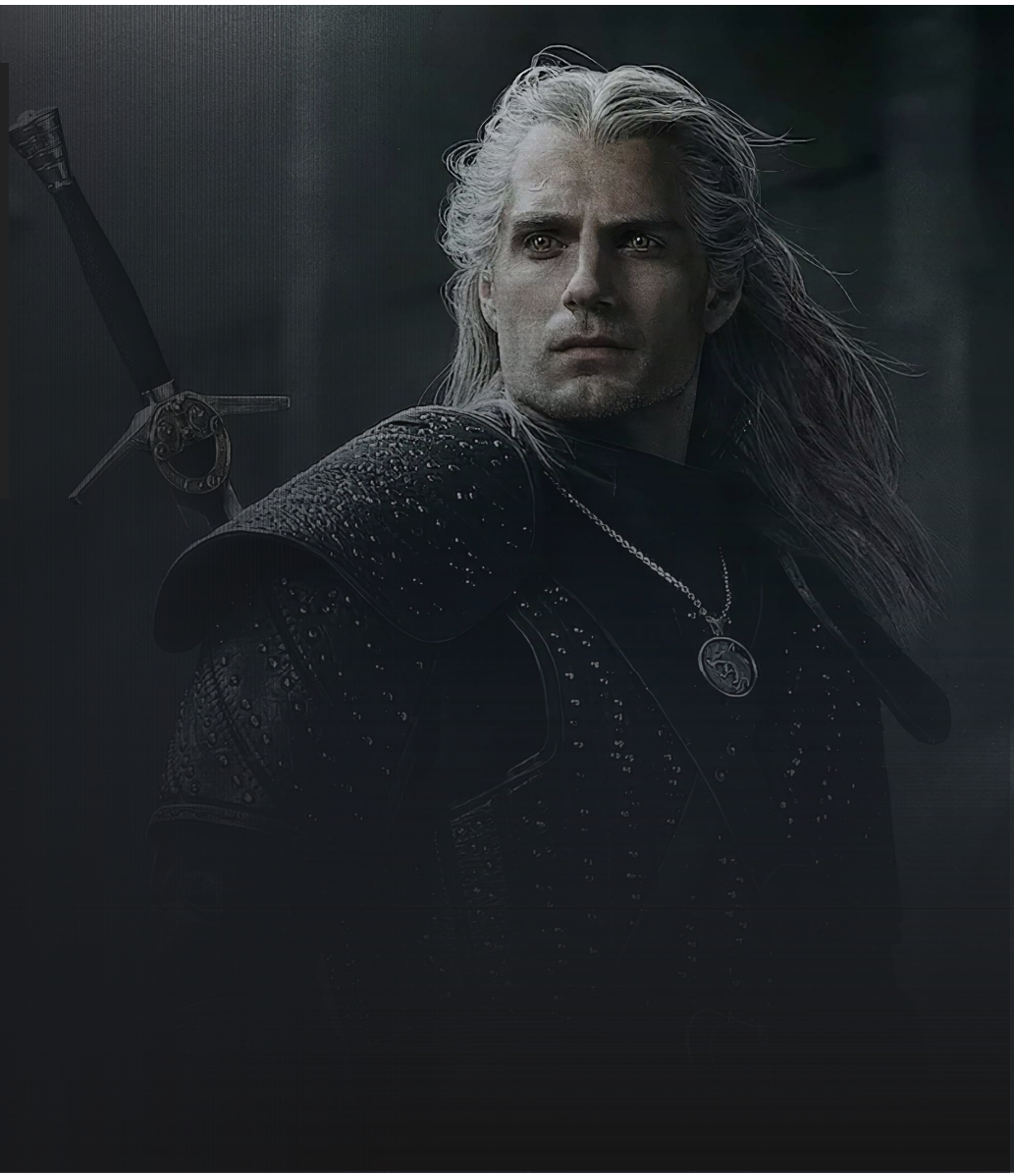
2 4 9 11 13 15 17 14 16 18 6 8 10 7 9 11

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

```sql
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
```

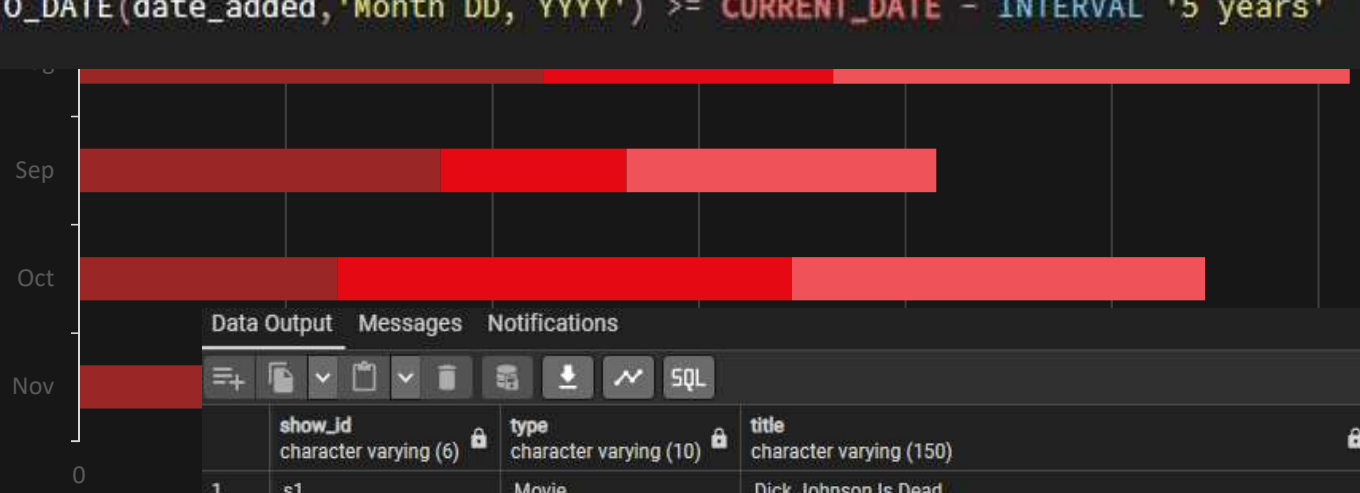
Data Output | Messages | Notifications

	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. Find content added in the last 5 years

```sql
SELECT
    *
FROM netflix
WHERE
    TO_DATE(date_added,'Month DD, YYYY') >= CURRENT_DATE - INTERVAL '5 years'
```



Data Output | Messages | Notifications

SQL

	show_id character varying (6)	type character varying (10)	title character varying (150)	director character varying (208)
1	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson
2	s2	TV Show	Blood & Water	[null]
3	s3	TV Show	Ganglands	Julien Leclercq
4	s4	TV Show	Jailbirds New Orleans	[null]
5	s5	TV Show	Kota Factory	[null]
6	s6	TV Show	Midnight Mass	Mike Flanagan

Total rows: 1000 of 3949 | Query complete 00:00:00.363 | Ln 96, Col 75

6. Identify the longest movie?

```sql
SELECT * FROM netflix
WHERE
    type = 'Movie'
    AND
    duration = (SELECT MAX(duration) FROM netflix)
```

Mobile

Basic

Standard

Premium

Price

Video Quality

Resolution

Number of Devices

Data Output | Messages | Notifications

	show_id character varying (6)	type character varying (10)	title character varying (150)	director character varying (208)	casts character varying (1000)
1	s52	Movie	InuYasha the Movie 2: The Castle Beyond the Looking Glass	Toshiya Shinohara	Kappei Yamaguchi, Satsuki Yuk
2	s53	Movie	InuYasha the Movie 3: Swords of an Honorable Ruler	Toshiya Shinohara	Kappei Yamaguchi, Satsuki Yuk
3	s120	Movie	Here and There	JP Habac	Janine Gutierrez, JC Santos, Vi
4	s338	Movie	Good Luck Chuck	Mark Helfrich	Dane Cook, Jessica Alba, Dan F
5	s345	Movie	My Girl 2	Howard Zieff	Anna Chlumsky, Austin O'Brien,
6	s427	Movie	Cousins	Ainsley Gardiner, Briar Grace-Smith	Rachel House, Briar Grace-Smit

Total rows: 118 of 118 | Query complete 00:00:00.266 | Ln 83, Col 1

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

```
SELECT * FROM netflix
WHERE director ILIKE '%Rajiv Chilaka%'
```

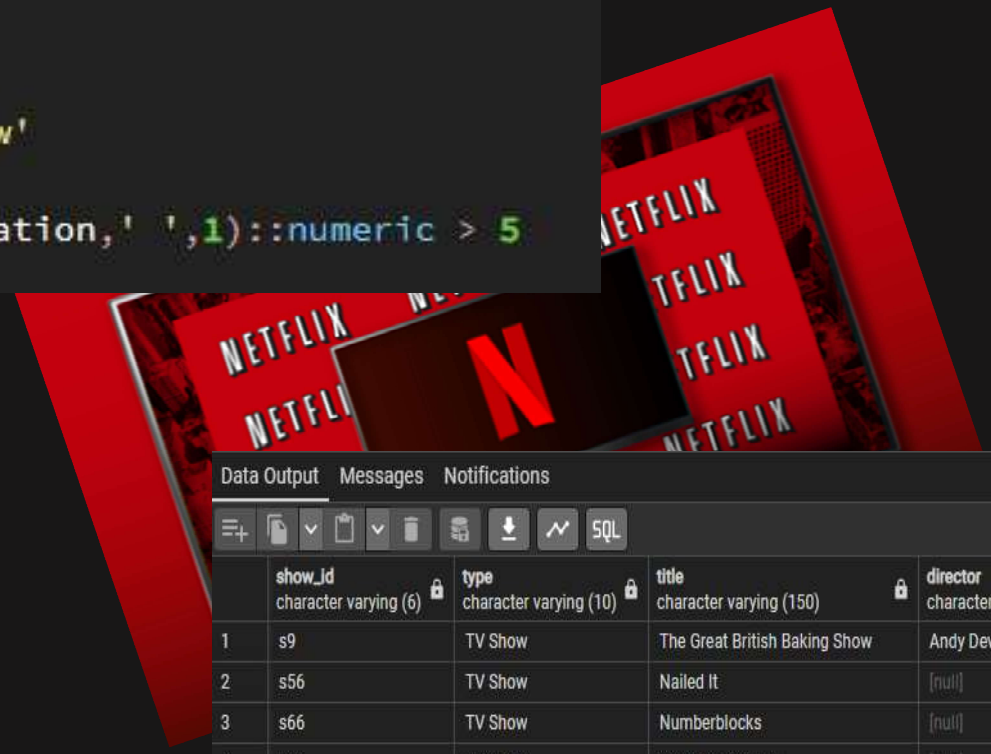
Data Output | Messages | Notifications

	show_id character varying (6)	type character varying (10)	title character varying (150)	director character varying (208)	casts character varying (
1	s407	Movie	Chhota Bheem - Neeli Pahaadi	Rajiv Chilaka	Vatsal Dubey, Julie
2	s408	Movie	Chhota Bheem & Ganesh	Rajiv Chilaka	Vatsal Dubey, Julie
3	s409	Movie	Chhota Bheem & Krishna: Mayanagari	Rajiv Chilaka	Vatsal Dubey, Julie
4	s410	Movie	Chhota Bheem & Krishna: Pataliputra- City of the De...	Rajiv Chilaka	Vatsal Dubey, Julie
5	s411	Movie	Chhota Bheem And The Broken Amulet	Rajiv Chilaka	Vatsal Dubey, Julie
6	s412	Movie	Chhota Bheem And The Crown of Valhalla	Rajiv Chilaka	Vatsal Dubey, Julie

Total rows: 22 of 22 | Query complete 00:00:00.189 | Ln 101, Col 39

8. List all TV shows with more than 5 seasons

```sql
SELECT
    *
FROM netflix
WHERE
    type = 'TV Show'
    AND
    SPLIT_PART(duration,' ',1)::numeric > 5
```



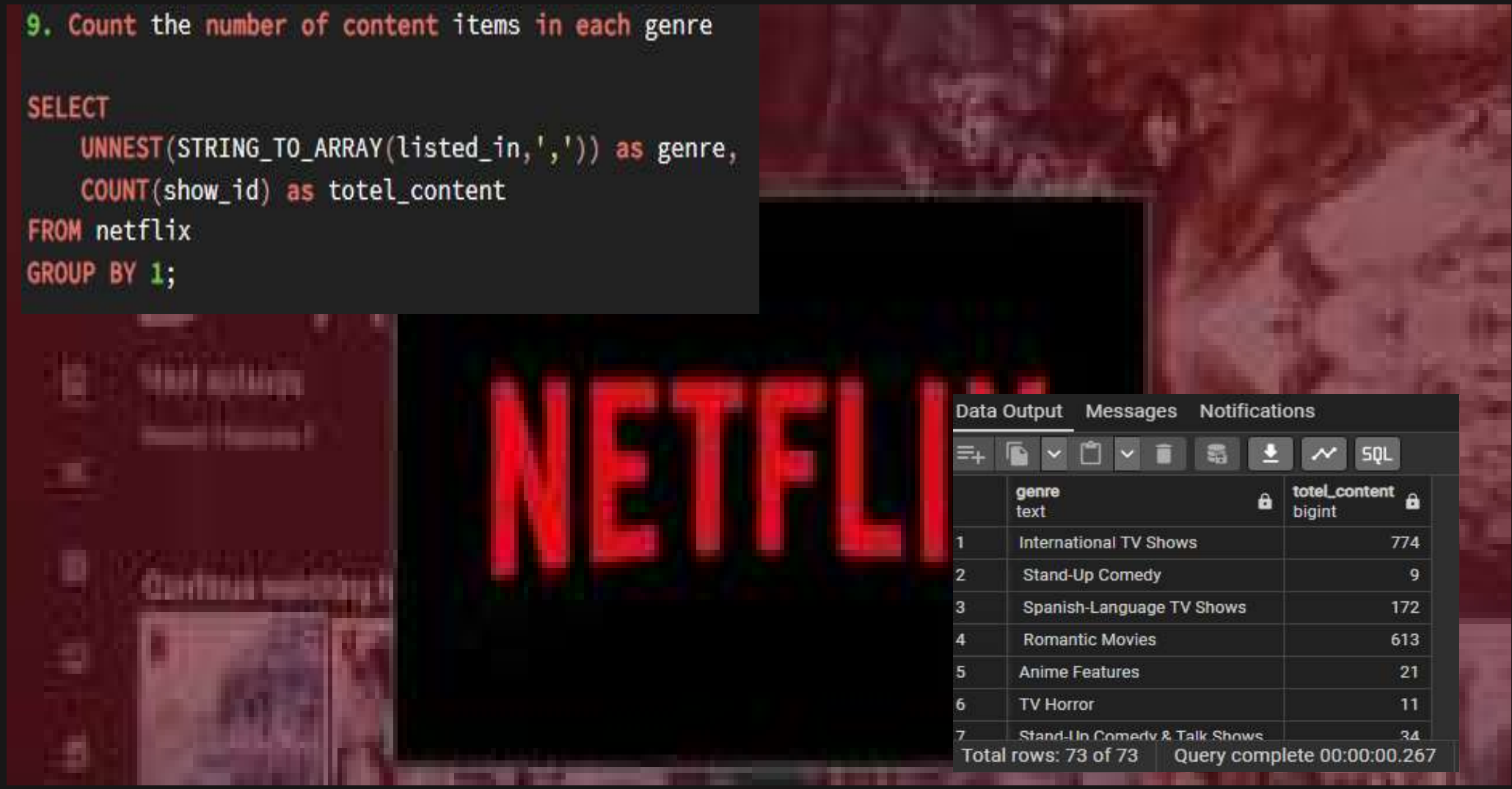
Data Output | Messages | Notifications

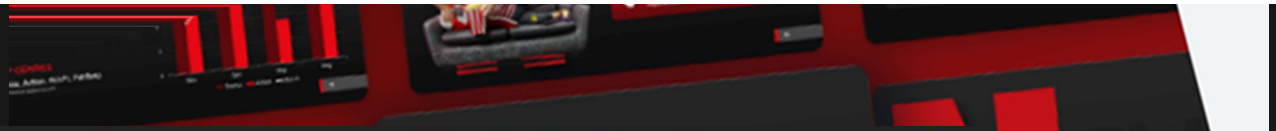
	show_id character varying (6)	type character varying (10)	title character varying (150)	director character varying (208)	casts character varying (1000)
1	s9	TV Show	The Great British Baking Show	Andy Devonshire	Mel Giedroyc, Sue Perkins, Mary Berry, Paul Hollywood
2	s56	TV Show	Nailed It	[null]	Nicole Byer, Jacques Torres
3	s66	TV Show	Numberblocks	[null]	Beth Chalmers, David Holt, Marcel McCalla, Teresa Gallagher
4	s68	TV Show	Saved by the Bell	[null]	Mark-Paul Gosselaar, Tiffani Thiessen, Mario Lopez, Lark Voorhies
5	s83	TV Show	Lucifer	[null]	Tom Ellis, Lauren German, Kevin Alejandro, D.B. Woodside, Lesley-
6	s275	TV Show	Grace and Frankie	[null]	Jane Fonda, Lily Tomlin, Martin Sheen, Sam Waterston, June Dian

Total rows: 99 of 99 | Query complete 00:00:00.195 | Ln 105, Col 1

9. Count the number of content items in each genre

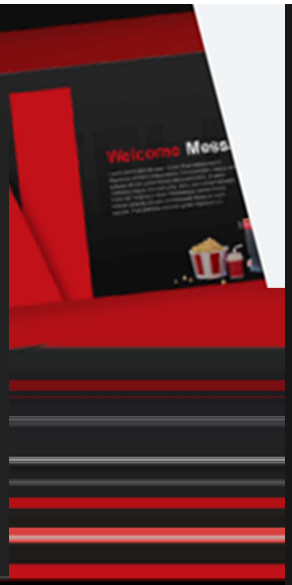
```sql
SELECT
    UNNEST(STRING_TO_ARRAY(listed_in,',')) as genre,
    COUNT(show_id) as totel_content
FROM netflix
GROUP BY 1;
```





10. Find each year and the average numbers of content release in India on netflix.
    return top 5 year with highest avg content release !

```sql
SELECT
    EXTRACT(YEAR FROM TO_DATE(date_added,'Month DD, YYYY')) as year,
    COUNT(*) as yearly_content,
    ROUND(COUNT(*)::numeric/(SELECT COUNT(*) FROM netflix WHERE country = 'India')
    ::numeric * 100,2) as avg_content
FROM netflix
WHERE country = 'India'
GROUP BY 1
```



Data Output | Messages | Notifications

SQL

	year numeric	yearly_content bigint	avg_content numeric
1	2018	333	34.26
2	2016	10	1.03
3	2019	203	20.88
4	2021	95	9.77
5	2020	189	19.44
6	2017	142	14.61

11. Find how many movies actor 'Akshay Kumar' appeared in last 10 years;

```sql
SELECT * FROM netflix
WHERE
    casts ILIKE '%Akshay Kumar%'
    AND
    release_year > EXTRACT(YEAR FROM CURRENT_DATE) - 10
```

Remedy

for frustrations

	show_id character varying (6)	type character varying (10)	title character varying (150)	director character varying (208)	casts character varying (1000)
1	s2274	Movie	Gabbar Is Back	Krish	Akshay Kumar, Shruti Haasan, Sunil Grover, Kareena Kapoor, Ishita Vyas, S
2	s2717	Movie	Brothers	Karan Malhotra	Akshay Kumar, Sidharth Malhotra, Jackie Shroff, Jacqueline Fernandez, Sh
3	s3495	Movie	Inside Man: Most Wanted	Michael J. Bassett	Aml Ameen, Rhea Seehorn, Roxanne McKee, Urs Rechn, Akshay Kumar, Tai
4	s4680	Movie	Naam Shabana	Shivam Nair	Tapsee Pannu, Akshay Kumar, Manoj Bajpayee, Prithviraj Sukumaran, Danr
5	s4694	Movie	Pad Man	R. Balki	Akshay Kumar, Sonam Kapoor, Radhika Apte, Jyoti Subhash, Urmila Mahar
6	s4695	Movie	Toilet: Ek Prem Katha	Shree Narayan Singh	Akshay Kumar, Bhumi Pednekar, Divyendu Sharma, Atul Srivastava, Anupar

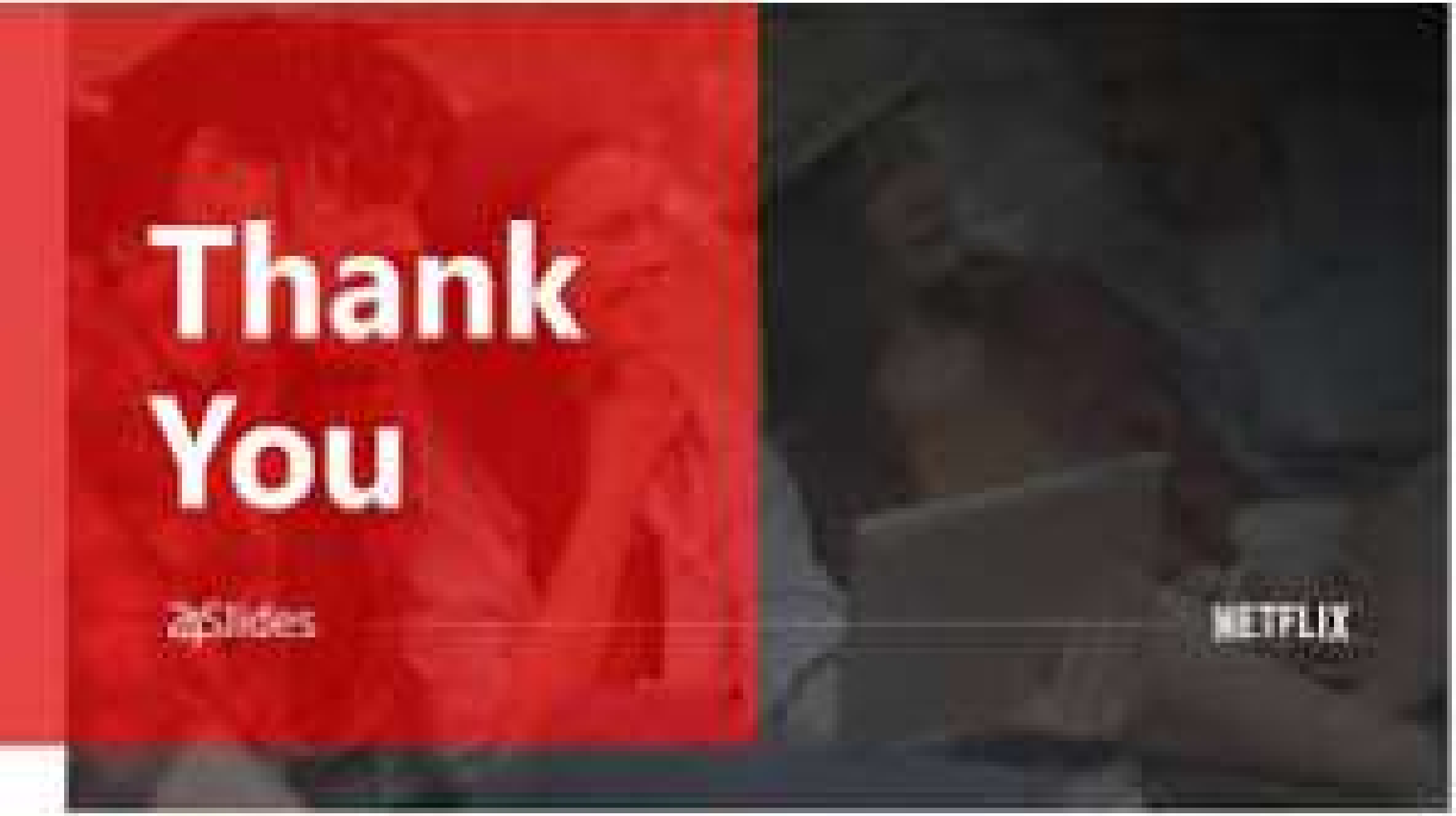
12. Categorize the content based on the presence of the keywords 'kill' and 'violence' in the description field. Label content containing these keywords as 'Bad' and all other content as 'Good'. Count how many items fall into each category.

```sql
WITH new_table
AS
(
SELECT
*,
    CASE
    WHEN description ILIKE '%kill%' OR
        description ILIKE '%violence%' THEN 'Bad_Content'
        ELSE 'Good_Content'
    END as Category
FROM netflix
)
SELECT
    category,
    COUNT(*) as total_content
FROM new_table
GROUP BY 1;
```

Data Output | Messages | Notifications

SQL

	category text	total_content bigint
1	Bad_Content	342
2	Good_Content	8465


Thank
You
slidesgo
NETFLIX