--Solutions

--Netflix Project

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(25),

description VARCHAR(280)

)

SELECT\*FROM netflix;

ALTER TABLE netflix

ALTER COLUMN listed\_in TYPE VARCHAR(100);

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

SELECT

type,

COUNT(\*) AS total\_content

FROM netflix

GROUP BY type

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

SELECT rating, COUNT(\*) AS count

FROM netflix

GROUP BY rating

ORDER BY count DESC

LIMIT 1;

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

SELECT\*FROM netflix

WHERE type='Movie' AND release\_year=2018;

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

SELECT country, COUNT(\*) AS total\_content

FROM netflix

WHERE country IS NOT NULL

GROUP BY country

ORDER BY total\_content DESC

LIMIT 5;

-- 5.Identify the longest movie

SELECT title, duration, release\_year

FROM netflix

WHERE type = 'Movie' AND duration LIKE '%min'

ORDER BY CAST(SPLIT\_PART(duration, ' ', 1) AS INTEGER) DESC

LIMIT 1;

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

SELECT

\*

FROM netflix

WHERE TO\_DATE(date\_added, 'Month DD, YYYY') >= CURRENT\_DATE - INTERVAL '5 years'

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

SELECT \*

FROM

(

SELECT

\*,

UNNEST(STRING\_TO\_ARRAY(director, ',')) as director\_name

FROM

netflix

)

WHERE

director\_name = 'Rajiv Chilaka'

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

SELECT \*

FROM netflix

WHERE

TYPE = 'TV Show'

AND

SPLIT\_PART(duration, ' ', 1)::INT > 5

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

SELECT

UNNEST(STRING\_TO\_ARRAY(listed\_in, ',')) as genre,

COUNT(\*) as total\_content

FROM netflix

GROUP BY 1

--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

country,

release\_year,

COUNT(show\_id) as total\_release,

ROUND(

COUNT(show\_id)::numeric/

(SELECT COUNT(show\_id) FROM netflix WHERE country = 'India')::numeric \* 100

,2

)

as avg\_release

FROM netflix

WHERE country = 'India'

GROUP BY country, 2

ORDER BY avg\_release DESC

LIMIT 5

-- 11. List all movies that are documentaries

SELECT \* FROM netflix

WHERE listed\_in LIKE '%Documentaries'

-- 12. Find all content without a director

SELECT \* FROM netflix

WHERE director IS NULL

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

SELECT \* FROM netflix

WHERE

casts LIKE '%Salman Khan%'

AND

release\_year > EXTRACT(YEAR FROM CURRENT\_DATE) - 10

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

SELECT

UNNEST(STRING\_TO\_ARRAY(casts, ',')) as actor,

COUNT(\*)

FROM netflix

WHERE country = 'India'

GROUP BY 1

ORDER BY 2 DESC

LIMIT 10

/\*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 description ILIKE '%kill%' OR description ILIKE '%violence%' THEN 'Bad'

ELSE 'Good'

END AS category

FROM netflix

) AS categorized\_content

GROUP BY 1,2

ORDER BY 2