```sql
-- Netflix Project

DROP TABLE IF EXISTS netflix;

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)


);


select * from netflix;


select

 count(*) as total_content

 from netflix;


SELECT

DISTINCT type
```
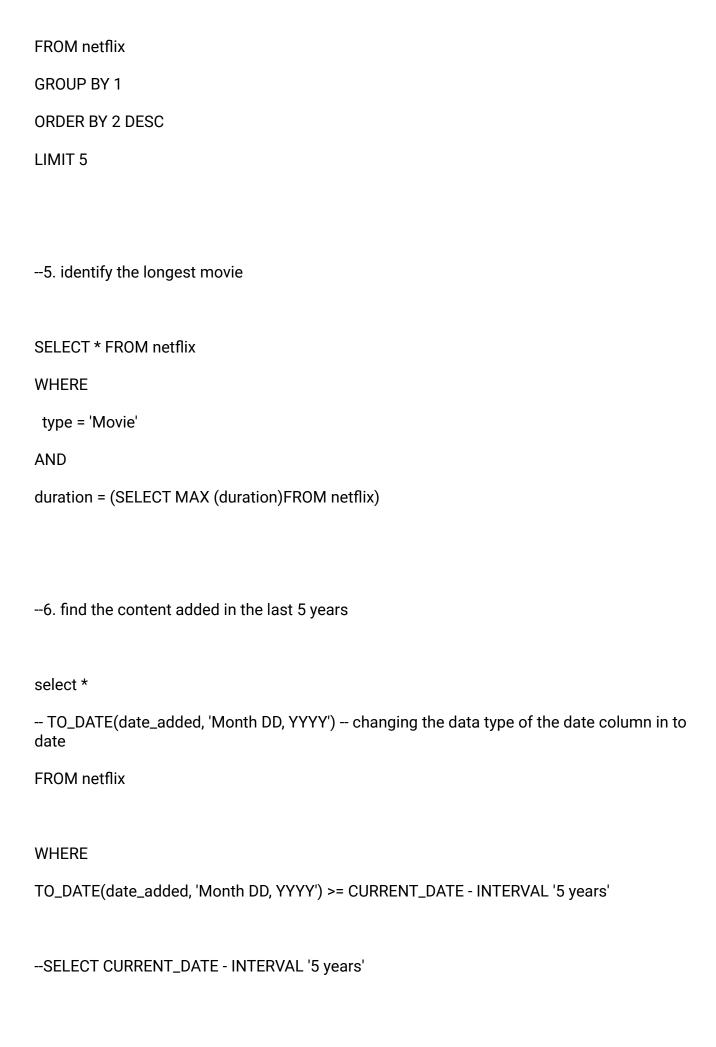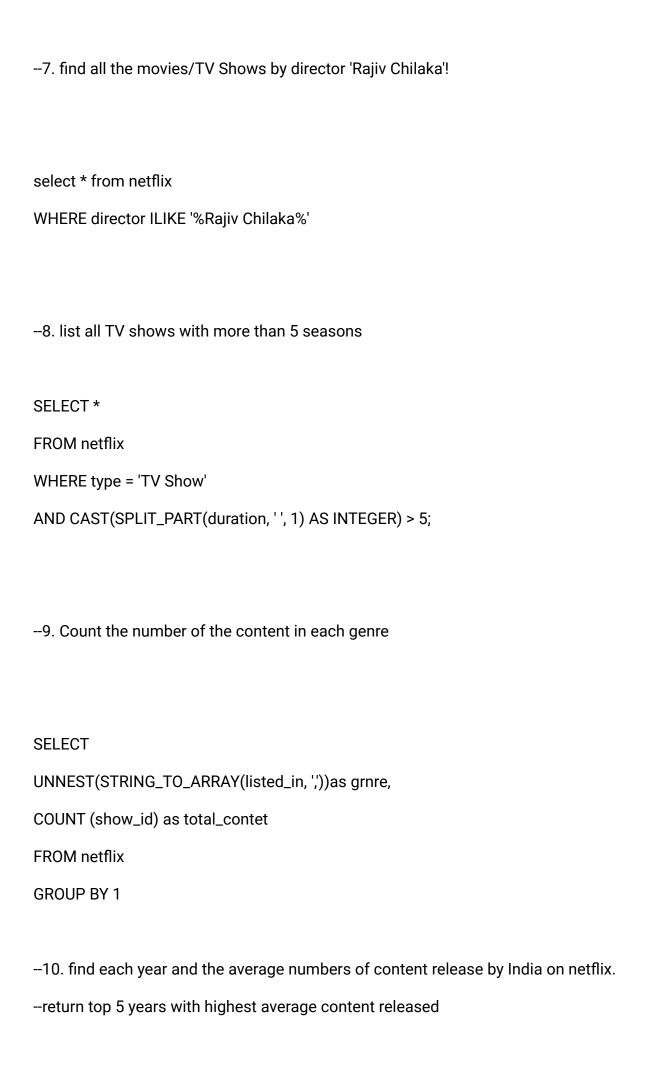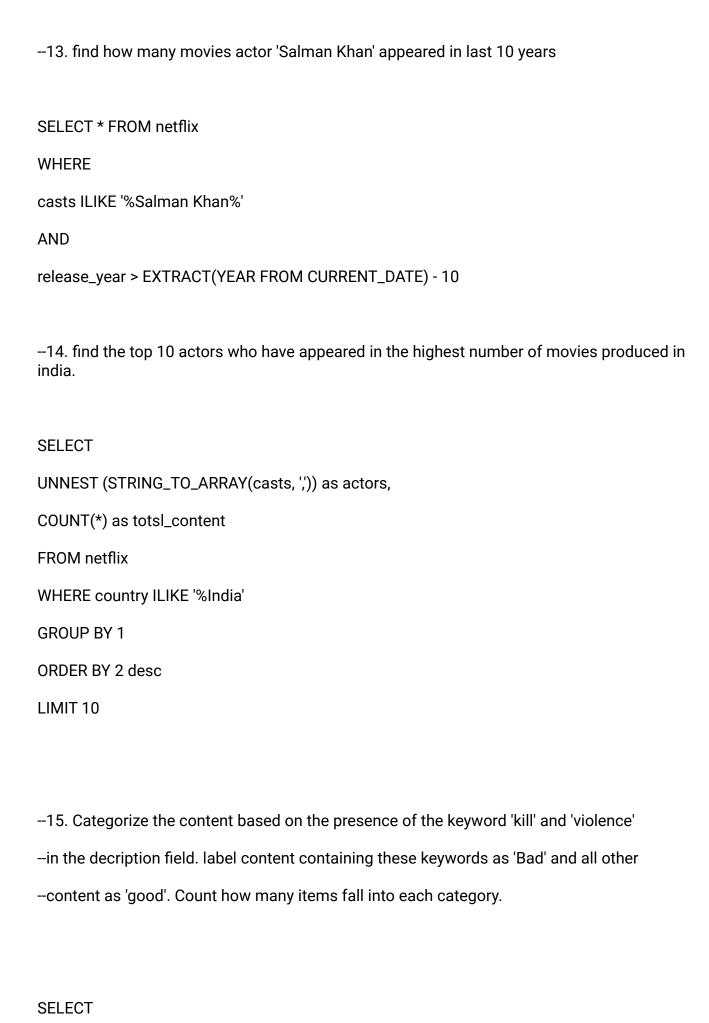
```sql
from netflix;



select * from netflix



--- 15 Business Problems



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



select

type,

count(*) as total_content

from netflix

group by type



--2. find from the most common rating fro movies and TV Shows



SELECT

type,

rating

FROM

(

  SELECT

    type,
```

```sql
    rating,
    COUNT(*),
    RANK() OVER(PARTITION BY type ORDER BY COUNT(*) DESC) as ranking
FROM netflix
GROUP BY 1, 2
) as t1
WHERE
ranking = 1
```

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

-- filter 2020

-- movies

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

--4. find the top 5 countries with the most content on netflix

```sql
SELECT
  UNNEST(STRING_TO_ARRAY(country, ',')) as ne_country,
COUNT(show_id) as total_content
```

```sql
FROM netflix

GROUP BY 1

ORDER BY 2 DESC

LIMIT 5



--5. identify the longest movie


SELECT * FROM netflix

WHERE

  type = 'Movie'

AND

duration = (SELECT MAX (duration)FROM netflix)



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


select *

-- TO_DATE(date_added, 'Month DD, YYYY') -- changing the data type of the date column in to date

FROM netflix


WHERE

TO_DATE(date_added, 'Month DD, YYYY') >= CURRENT_DATE - INTERVAL '5 years'


--SELECT CURRENT_DATE - INTERVAL '5 years'
```

--7. find all the movies/TV Shows by director 'Rajiv Chilaka'!

```sql
select * from netflix

WHERE director ILIKE '%Rajiv Chilaka%'
```

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

```sql
SELECT *

FROM netflix

WHERE type = 'TV Show'

AND CAST(SPLIT_PART(duration, ' ', 1) AS INTEGER) > 5;
```

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

```sql
SELECT

UNNEST(STRING_TO_ARRAY(listed_in, ','))as grnre,

COUNT (show_id) as total_contet

FROM netflix

GROUP BY 1
```

--10. find each year and the average numbers of content release by India on netflix.
--return top 5 years with highest average content released

```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_per_year

FROM netflix

WHERE country = 'India'

GROUP BY 1
```

--11. list all movies that are documentries

```sql
SELECT * FROM netflix

WHERE

listed_in ILIKE '%documentaries%'
```

--12. find all content without a director

```sql
SELECT * FROM netflix

WHERE

director ISNULL
```

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

SELECT * FROM netflix

WHERE

casts ILIKE '%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 actors,

COUNT(*) as totsl_content

FROM netflix

WHERE country ILIKE '%India'

GROUP BY 1

ORDER BY 2 desc

LIMIT 10

--15. Categorize the content based on the presence of the keyword 'kill' and 'violence'

--in the decription field. label content containing these keywords as 'Bad' and all other

--content as 'good'. Count how many items fall into each category.

SELECT

```sql
*,
  CASE
WHEN description ILIKE '%kill%' OR
  description ILIKE '%violence%' THEN 'Bad_Content'
  ELSE 'Good_content'
END category
FROM netflix
```

--15 question number 15th query for tatal count of bad content and good content

```sql
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
```