NETFLIX_MOVIES_ANALYSIS

📄 netflix_titles   DataFrame as df9

```sql
-- Content in Dataset of netflix_titles.csv

SELECT *
FROM "netflix_titles.csv";
```

...	...	s. ...	title	director	casting
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	null
1	s2	TV Show	Blood & Water	null	Ama Qamata, Khosi Ngem
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Goto
3	s4	TV Show	Jailbirds New Orleans	null	null
4	s5	TV Show	Kota Factory	null	Mayur More, Jitendra Kum
5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford,
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmik
8	s9	TV Show	The Great British Baking Show	Andy Devonshire	Mel Giedroyc, Sue Perkins
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris C
10	s11	TV Show	Vendetta: Truth, Lies and The Mafia	null	null
11	s12	TV Show	Bangkok Breaking	Kongkiat Komesiri	Sukollawat Kanarot, Susho
12	s13	Movie	Je Suis Karl	Christian Schwochow	Luna Wedler, Jannis Niew
13	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Klara Castanho, Lucca Pic
14	s15	TV Show	Crime Stories: India Detectives	null	null
15	s16	TV Show	Dear White People	null	Logan Browning, Brandon

Rows: 8,333 ⚠ Truncated from 8,807 rows   ⤢ Expand

📄 netflix_titles   DataFrame as df1

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

SELECT
    show_type,
    COUNT(*) as total_content
FROM netflix_titles.csv
GROUP BY show_type
```

index	show_type	total_content
0	Movie	
1	TV Show	

Rows: 2   ⤢ Expand

netflix_titles   DataFrame `as` `df`

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

WITH RatingCounts AS (
    SELECT
        show_type,
        rating,
        COUNT(*) AS rating_count
    FROM netflix_titles.csv
    GROUP BY show_type, rating
),
RankedRatings AS (
    SELECT
        show_type,
        rating,
        rating_count,
        RANK() OVER (PARTITION BY show_type ORDER BY rating_count DESC) AS rank
    FROM RatingCounts
)
SELECT
    show_type,
    rating AS most_frequent_rating
FROM RankedRatings
WHERE rank = 1;
```

index	⋯ ↑↓	show_type	⋯ ↑↓	most_frequent_rating
	0	Movie		TV-MA
	1	TV Show		TV-MA

Rows: 2   ⤢ Expand

netflix_titles   DataFrame `as` `df2`

```sql
-- 3. List all movies released in a specific year (e.g., 2021)

SELECT *
FROM netflix_titles.csv
WHERE show_type = 'Movie' AND release_year = 2021
```

⋯ ↑↓	⋯ ↑↓	s. ⋯ ↑↓	title ⋯ ↑↓	director ⋯ ↑↓	casting
0	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko
1	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris C
2	s13	Movie	Je Suis Karl	Christian Schwochow	Luna Wedler, Jannis Niew
3	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Klara Castanho, Lucca Pic
4	s19	Movie	Intrusion	Adam Salky	Freida Pinto, Logan Marsh
5	s24	Movie	Go! Go! Cory Carson: Chrissy Takes the Wheel	Alex Woo, Stanley Moore	Maisie Benson, Paul Killam
6	s31	Movie	Ankahi Kahaniya	Ashwiny Iyer Tiwari, Abhishek Chaubey, Sake…	Abhishek Banerjee, Rinku
7	s36	Movie	The Father Who Moves Mountains	Daniel Sandu	Adrian Titieni, Elena Purea
8	s37	Movie	The Stronghold	Cédric Jimenez	Gilles Lellouche, Karim Lek
9	s46	Movie	My Heroes Were Cowboys	Tyler Greco	null
10	s65	Movie	Nightbooks	David Yarovesky	Winslow Fegley, Lidya Jew
11	s69	Movie	Schumacher	Hanns-Bruno Kammertöns, Vanessa Nöcker, …	Michael Schumacher
12	s72	Movie	A StoryBots Space Adventure	David A. Vargas	Evan Spiridellis, Erin Fitzge
13	s76	Movie	You vs. Wild: Out Cold	Ben Simms	Bear Grylls, Jason Derek F
14	s78	Movie	Little Singham - Black Shadow	Prakash Satam	Sumriddhi Shukla, Jigna Bl
15	s80	Movie	Tughlaq Durbar (Telugu)	Delhiprasad Deenadayalan	Vijay Sethupathi, Parthiba

Rows: 277   ⤢ Expand

📄 netflix_titles   DataFrame as df3

```sql
-- 4. Find the top 5 countries with the most content on Netflix
SELECT country, COUNT(*) AS total_content
FROM (
    SELECT TRIM(s) AS country
    FROM "netflix_titles.csv",
    LATERAL unnest(str_split(country, ',')) AS t(s)
) t1
WHERE country IS NOT NULL AND country <> ''
GROUP BY country
ORDER BY total_content DESC
LIMIT 5;
```

	country	total_c...	
0	United States	3690	
1	India	1046	
2	United Kingdom	806	
3	Canada	445	
4	France	393	

Rows: 5   ↗ Expand

📄 netflix_titles   DataFrame as d

```sql
-- 5. Identify the longest movie

SELECT * FROM netflix_titles.csv
WHERE show_type = 'Movie'
ORDER BY SPLIT_PART(duration, ' ', 1)::INT DESC
```

		s.	title	director	casting
0	s4254	Movie	Black Mirror: Bandersnatch	null	Fionn Whitehead, Will Poulter
1	s718	Movie	Headspace: Unwind Your Mind	null	Andy Puddicombe, Evelyn Le
2	s2492	Movie	The School of Mischief	Houssam El-Din Mustafa	Suhair El-Babili, Adel Emam,
3	s2488	Movie	No Longer kids	Samir Al Asfory	Said Saleh, Hassan Moustafa
4	s2485	Movie	Lock Your Girls In	Fouad El-Mohandes	Fouad El-Mohandes, Sanaa Y
5	s2489	Movie	Raya and Sakina	Hussein Kamal	Suhair El-Babili, Shadia, Abde
6	s167	Movie	Once Upon a Time in America	Sergio Leone	Robert De Niro, James Wood
7	s7933	Movie	Sangam	Raj Kapoor	Raj Kapoor, Vyjayanthimala,
8	s1020	Movie	Lagaan	Ashutosh Gowariker	Aamir Khan, Gracy Singh, Ra
9	s4574	Movie	Jodhaa Akbar	Ashutosh Gowariker	Hrithik Roshan, Aishwarya Ro
10	s808	Movie	Seven Souls in the Skull Castle: Season Moon...	Hidenori Inoue	Mamoru Miyano, Hiroki Suzu
11	s2732	Movie	Kabhi Khushi Kabhie Gham	Karan Johar	Amitabh Bachchan, Jaya Bh
12	s3228	Movie	The Irishman	Martin Scorsese	Robert De Niro, Al Pacino, Jo
13	s2633	Movie	No Direction Home: Bob Dylan	Martin Scorsese	Bob Dylan, Joan Baez, Dave
14	s8328	Movie	The Gospel of Luke	David Batty	Selva Rasalingam, Karima Go
15	s807	Movie	Seven Souls in the Skull Castle: Season Moon	Hidenori Inoue	Sota Fukushi, Taichi Saotome

Rows: 6,131   ↗ Expand

📄 netflix_titles   DataFrame as

```sql
-- 6. Find content added in the last 5 years
SELECT *
FROM "netflix_titles.csv"
WHERE strptime(date_added, '%B %d, %Y') >= current_date - INTERVAL '5 years';
```

...	...	s. ...	title	...	director	...	casting
0	s1	Movie	Dick Johnson Is Dead		Kirsten Johnson		null
1	s2	TV Show	Blood & Water		null		Ama Qamata, Khosi Ngema,
2	s3	TV Show	Ganglands		Julien Leclercq		Sami Bouajila, Tracy Gotoas,
3	s4	TV Show	Jailbirds New Orleans		null		null
4	s5	TV Show	Kota Factory		null		Mayur More, Jitendra Kumar
5	s6	TV Show	Midnight Mass		Mike Flanagan		Kate Siegel, Zach Gilford, Ha
6	s7	Movie	My Little Pony: A New Generation		Robert Cullen, José Luis Ucha		Vanessa Hudgens, Kimiko Gl
7	s8	Movie	Sankofa		Haile Gerima		Kofi Ghanaba, Oyafunmike C
8	s9	TV Show	The Great British Baking Show		Andy Devonshire		Mel Giedroyc, Sue Perkins, M
9	s10	Movie	The Starling		Theodore Melfi		Melissa McCarthy, Chris O'D
10	s11	TV Show	Vendetta: Truth, Lies and The Mafia		null		null
11	s12	TV Show	Bangkok Breaking		Kongkiat Komesiri		Sukollawat Kanarot, Sushar N
12	s13	Movie	Je Suis Karl		Christian Schwochow		Luna Wedler, Jannis Niewöhr
13	s14	Movie	Confessions of an Invisible Girl		Bruno Garotti		Klara Castanho, Lucca Picon
14	s15	TV Show	Crime Stories: India Detectives		null		null
15	s16	TV Show	Dear White People		null		Logan Browning, Brandon P F

Rows: 2,125   ⤢ Expand

📄 netflix_titles   DataFrame as

```sql
-- 7. 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_titles.csv WHERE country = 'India')::numeric * 100 ,2) as
avg_release
FROM netflix_titles.csv
WHERE country = 'India'
GROUP BY country, 2
ORDER BY avg_release DESC
LIMIT 5
```

...	...	relea...	total_r...	av...
0	India	2017	101	10.39
1	India	2018	94	9.67
2	India	2019	87	8.95
3	India	2020	75	7.72
4	India	2016	73	7.51

Rows: 5   ⤢ Expand

📄 netflix_titles   DataFrame as

```sql
-- 8. Find how many movies actor 'Bear Grylls' appeared in last 10 years!

SELECT *
FROM "netflix_titles.csv"
WHERE casting LIKE '%Bear Grylls%'
  AND release_year > EXTRACT(YEAR FROM current_date) - 10;
```

	...	⇅	...	⇅	s. ...	⇅	title	...	⇅	d ...	⇅	casting	...	⇅	country	...
0		s76		Movie		You vs. Wild: Out Cold		Ben Simms		Bear Grylls, Jason Derek Prempeh		null
1		s1305		Movie		Animals on the Loose: A You vs. Wild Movie		Ben Simms		Bear Grylls		United States
2		s3932		TV Show		You vs. Wild		null		Bear Grylls		United States
3		s7393		TV Show		Man vs Wild with Sunny Leone		null		Bear Grylls, Sunny Leone		null

Rows: 4   ⤢ Expand

📄 netflix_titles   DataFrame as

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

SELECT actor, COUNT(*) AS total_movies
FROM (
    SELECT TRIM(a) AS actor
    FROM "netflix_titles.csv",
    LATERAL unnest(str_split("casting", ',')) AS t(a)
    WHERE country = 'India'
) sub
WHERE actor IS NOT NULL AND actor <> ''
GROUP BY actor
ORDER BY total_movies DESC
LIMIT 10;
```

	⇅	actor	...	⇅	total...	...	⇅
0		Anupam Kher			40
1		Shah Rukh Khan			32
2		Naseeruddin Shah			29
3		Amitabh Bachchan			28
4		Akshay Kumar			28
5		Paresh Rawal			27
6		Om Puri			26
7		Kareena Kapoor			24
8		Boman Irani			23
9		Ajay Devgn			20

Rows: 10   ⤢ Expand

📄 netflix_titles   DataFrame as

```sql
/*
Question 10:
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,
    show_type,
    COUNT(*) AS content_count
FROM (
    SELECT *,
        CASE
            WHEN description LIKE '%kill%' OR description LIKE '%violence%' THEN 'Bad'
            ELSE 'Good'
        END AS category
    FROM netflix_titles.csv
) AS categorized_content
GROUP BY 1,2
ORDER BY 2
```

··· ⇅	··· ⇅	s. ··· ⇅	conten... ··· ⇅
0	Good	Movie	5886
1	Bad	Movie	245
2	Good	TV Show	2586
3	Bad	TV Show	90

Rows: 4   ↗ Expand