SQL Solutions for the analysis to be made

By :- L.N.Sai Tejhas

Contact Details :- Linkedin , saitejhas49@gmail.com

1. What are the average IMDb scores for each genre of Netflix Originals?

SQL> SELECT a.Genre, ROUND(AVG(b.IMDBScore),2) as avg_imdbscore

FROM netflix_originals b

JOIN genre_details a

ON b.GenreID = a.GenreID

WHERE b.GenreID IS NOT NULL AND b.IMDBScore IS NOT NULL

GROUP BY a.Genre

ORDER BY avg_imdbscore DESC

Approach: I used AVG() on IMDBScore after joining netflix_originals with genre_details, which were grouped by genre.

Resultant Output:- Average IMDB scores.csv

Conclusion: This analysis provided a clear picture of how all 19 genres performed regarding audience ratings. It helped identify genres that consistently receive higher viewer satisfaction.

2. Which genres have an average IMDb score higher than 7.5?

SQL> SELECT a.Genre, ROUND(AVG(b.IMDBScore), 2) as avg_imdbscore

FROM netflix_originals b

JOIN genre_details a

ON b.GenreID = a.GenreID

WHERE b.GenreID IS NOT NULL AND b.IMDBScore IS NOT NULL

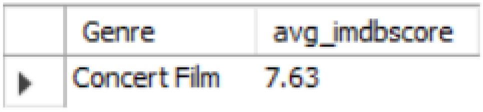
GROUP BY a.Genre

HAVING avg_imdbscore>7.50

```
order by avg_imdbscore DESC;
```

Approach: I used AVG() on IMDBScore after joining netflix_originals with genre_details, which were grouped by genre and filtered genres using HAVING AVG(IMDBScore) > 7.5.

Resultant Output:-



	Genre	avg_imdbscore
▶	Concert Film	7.63

Conclusion: Only the **Concert Film** genre exceeds the 7.5 threshold, indicating that exceptional content is concentrated in that genre.

3. List Netflix Original titles in descending order of their IMDb scores.

```
SQL> SELECT title,IMDBScore FROM netflix_originals

ORDER BY IMDBScore DESC
```

Approach: I sorted titles in descending order using ORDER BY IMDBScore DESC.

Resultant Output:- Titles with IMDB scores.csv

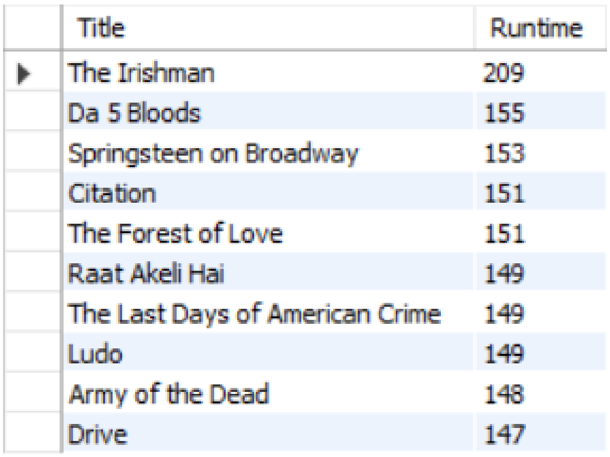
Conclusion: This allowed me to spotlight the highest-rated Netflix Originals, revealing standout productions and critical audience favorites.

4. Retrieve the top 10 longest Netflix Originals by runtime.

```
SQL> SELECT Title,Runtime FROM netflix_originals

ORDER BY runtime DESC

LIMIT 10;
```

Approach: I sorted the dataset by Runtime DESC and limited the result to 10.

Resultant Output:-



Title	Runtime
The Irishman	209
Da 5 Bloods	155
Springsteen on Broadway	153
Citation	151
The Forest of Love	151
Raat Akeli Hai	149
The Last Days of American Crime	149
Ludo	149
Army of the Dead	148
Drive	147

Conclusion: Identified the top 10 longest Netflix Originals based on runtime, which may suggest a higher investment in storytelling and production scale.

5. Retrieve the titles of Netflix Originals along with their respective genres.

```
SQL> SELECT a.title, b.genre

FROM netflix_originals a

JOIN genre_details b ON a.GenreID=b.GenreID

order by b.genre;
```

Approach: I performed a simple join on GenreID to display each title with its corresponding genre.

Resultant Output:- Netflix Original Titles and their Genres.csv

Conclusion: Derived the foundational analysis providing genre-based categorization of titles, useful for user segmentation or genre-specific recommendations.

6. Rank Netflix Originals based on their IMDb scores within each genre.

```
SQL>SELECT b.Genre, a.Title, a.IMDBScore, DENSE_RANK() OVER (PARTITION BY b.Genre ORDER BY a.IMDBScore DESC) AS GenreRank

FROM Netflix_Originals a

JOIN Genre_Details b
```

```
ON a.GenreID = b.GenreID

WHERE a.IMDBScore IS NOT NULL

ORDER BY b.Genre, GenreRank;
```

Approach: I used DENSE_RANK() over partitions of genre sorted by IMDBScore DESC to get an appropriate ranking instead of RANK().

Resultant Output:- [Netflix originals and their ranks.csv]()

Conclusion: Successfully ranked top-performing titles within each genre, helping to highlight the best examples of content style within each category.

7. Which Netflix Originals have IMDb scores higher than the average IMDb score of all titles?

```
SQL> SELECT Title,IMDBScore

FROM netflix_originals

WHERE IMDBScore>(SELECT AVG(IMDBScore) FROM netflix_originals WHERE IMDBScore IS NOT NULL)

ORDER BY IMDBScore DESC;
```

Approach: I included a subquery to calculate the average IMDb score; filtered the main table using WHERE IMDBScore > subquery.

Resultant Output:- [Highest rated originals .csv]()

Conclusion: Filtered out underperforming content and showcased titles that stood out in quality relative to the overall average.

8. How many Netflix Originals are there in each genre?

```
SQL> SELECT b.Genre,COUNT(a.Title) as Count_of_originals

FROM netflix_originals a

JOIN genre_details b

on a.GenreID=b.GenreID
```

```
WHERE a.Title IS NOT NULL

GROUP BY b.Genre

ORDER BY Count_of_originals DESC;
```

Approach: I simply grouped by genre and counted Title entries that weren't null to avoid miscalculations.

Resultant Output:- Genres and the no. of originals.csv

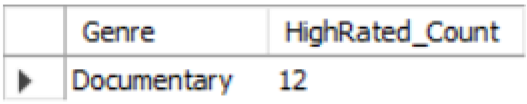
Conclusion: Offered insights into Netflix's genre-wise production strategy, revealing which genres are most heavily invested in.

9. Which genres have more than 5 Netflix Originals with an IMDb score higher than 8?

```
SQL> SELECT b.Genre, COUNT(a.Title) AS HighRated_Count

FROM netflix_originals a

JOIN genre_details b

ON a.GenreID = b.GenreID

WHERE a.IMDBScore > 8

GROUP BY b.Genre

HAVING COUNT(a.Title) > 5

ORDER BY HighRated_Count DESC;
```

Approach: I applied WHERE IMDBScore > 8 and HAVING COUNT(*) > 5 to get the desired results

Resultant Output:-



	Genre	HighRated_Count
▶	Documentary	12

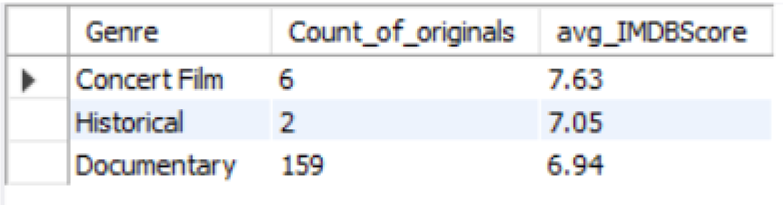
Conclusion: Highlighted genre "Documentary" that is not only prolific but also maintains high content quality, making it a valuable focus area for future development.

10. What are the top 3 genres with the highest average IMDb scores, and how many Netflix Originals do they have?

```
SQL> SELECT b.Genre, COUNT(a.Title) AS Count_of_originals,
ROUND(AVG(a.IMDBScore),2) AS avg_IMDBScore

FROM netflix_originals a

JOIN genre_details b

ON a.GenreID = b.GenreID

WHERE a.IMDBScore IS NOT NULL

GROUP BY b.Genre

ORDER BY avg_IMDBScore DESC

LIMIT 3;
```

Approach: I aggregated average scores and title counts per genre, ordered by average, and limited to 3.

Resultant Output:-



	Genre	Count_of_originals	avg_IMDBScore
▶	Concert Film	6	7.63
	Historical	2	7.05
	Documentary	159	6.94

Conclusion: Identified the top-performing genres("Concert Film"," Historical", and "Documentary") in both quality and content strength, which can inform recommendations, marketing, and future production.