SQL PROJECT REPORT

SQL ANALYSIS OF CONTENT PERFORMANCE AND VIEWER TRENDS

Submitted by: Hardik Harchilkar

Date: 11 November 2025

Role: Aspiring Data Analyst

LinkedIn: linkedin.com/in/hardikharchilkar

Tools Used: MySQL Workbench, Excel

Dataset: Netflix Movies and TV Shows – Kaggle

Project Overview

This project focuses on analyzing a Netflix-style streaming dataset using SQL to identify content performance trends, audience preferences, and viewing behavior.

The dataset includes details about movies and TV shows, such as genre, release year, country, and rating.

Simulated user and watch event data were created to understand viewer engagement, completion rates, and time-of-day viewing patterns.

Through this analysis, various business questions were answered — such as identifying the most popular genres, countries producing the most content, and patterns in user activity.

The goal of this project is to showcase how SQL can be used by data analysts to generate meaningful business insights from raw data.

Dataset and Tools

This analysis was conducted using the Netflix Movies and TV Shows dataset obtained from Kaggle. The dataset contains metadata about titles, genres, countries, release years, ratings, and descriptions.

Two simulated tables — `users` and `watch_events` — were also created to represent viewing activity and user demographics. This enabled the analysis of viewer behavior, completion rates, and watch-time trends.

- Tables Used:

- **`netflix_titles`** → Contains all movie/TV show metadata.

- **`users`** → Simulated table representing user profiles.

- **`watch_events`** → Simulated table representing user viewing data.

- Tools Used:

- MySQL Workbench → Data storage and analysis using SQL queries.

- Excel → For summarizing, formatting, and documenting insights.

Key Metrics Summary

Metric	Value
Total Titles	~8,800
Total Simulated Users	20
Total Watch Events	100+
Average Watch Duration	~85 minutes
Completion Rate	70%
Top Genre by Views	Drama
Most Active Country	India
Most Active Plan Type	Premium

Key Business Questions and Insights

Q1. Which type of content dominates the platform?
Movies account for about 70% of total content, while TV Shows make up around 30%.
More TV content can improve weekly retention.

Q2. Which genres have the most titles?
Dramas, Comedies, and Documentaries lead in total content.
Focus production and marketing on these genres.

Q3. Which countries produce the most content?
The USA, India, UK, and Japan dominate production.
Invest in regional partnerships.

Q4. Which ratings are most common?
TV-MA and R-rated content dominate.
Balance with family-friendly content.

Q5. Which directors have created the most content?
Rajiv Chilaka, Suhas Kadav, and Jay Karas are top contributors.
Retain and expand director collaborations.

Q6. How many titles were released each year?
Releases peaked between 2018–2020.
Maintain consistent annual output.

Q7. Which genres are watched most by users?
Drama, Comedy, and Action generate the highest views.
Prioritize these genres.

Q8. Which titles have the highest average watch duration?
Documentaries and long-format series top the chart.
Invest in binge-worthy storytelling.

Q9. What percentage of content is completed by users?
About 70% of started content is completed.
Excellent engagement; analyze drop-offs.

Q10. Do users watch more on weekends or weekdays?
Weekends see the highest watch time.
Schedule new releases on Fridays.

Q11. Which countries have the most active viewers?
India, the USA, and the UK.
Localize recommendations.

Q12. Which plan type watches more on average?
Premium users watch 2–3 times longer.
Encourage upgrades with offers.

Q13. Which users are inactive?
10–15% of users haven't watched anything.
Re-engage inactive users with personalized messages.

Q14. What is the average gap between release and watch date?
Users watch content around 2–3 years after release.
Promote older high-performing content.

Q15. Which genres are growing or declining year over year?
Drama and Crime genres are growing, while Family and Romance genres are declining.
Focus investment on trending genres.

Key Business Questions and Insights

- Learned how to use SQL for real-world business problem-solving.

- Practiced creating and managing multiple relational tables (`netflix_titles`, `users`, `watch_events`).

- Applied SQL functions like JOIN, GROUP BY, CASE, and DATE to extract insights.

- Understood how to analyze and summarize data using queries.

- Learned how to present SQL results as business insights using Excel and reports.

- Strengthened analytical thinking and report presentation skills for data analyst roles.

Conclusion

This SQL analysis project provided practical experience in extracting business insights from a streaming dataset similar to Netflix.

The findings revealed that drama and comedy dominate engagement, weekends drive the most viewing time, and premium users are the most active audience group.

By identifying top-performing genres, content gaps, and user behavior patterns, the project demonstrates how SQL can be a powerful tool for data-driven decision-making in the entertainment industry.

Project File Structure

```
Netflix_SQL_Project
│
├── sql
│   ├── netflix_project.sql
│   ├── netflix_users_watch_events.sql
│   └── netflix_analysis.sql
│
├── results
│   ├── Q01_Content_Type.csv
│   ├── ...
│   └── Q15_Genre_Trend.csv
│
├── dashboard
│   └── netflix_summary_results.xlsx
│
└── docs
    ├── netflix_analysis_summary.docx
    └── netflix_analysis_summary.pdf
```

Prepared by: **Hardik Harchilkar**

Aspiring Data Analyst

linkedin.com/in/hardikharchilkar