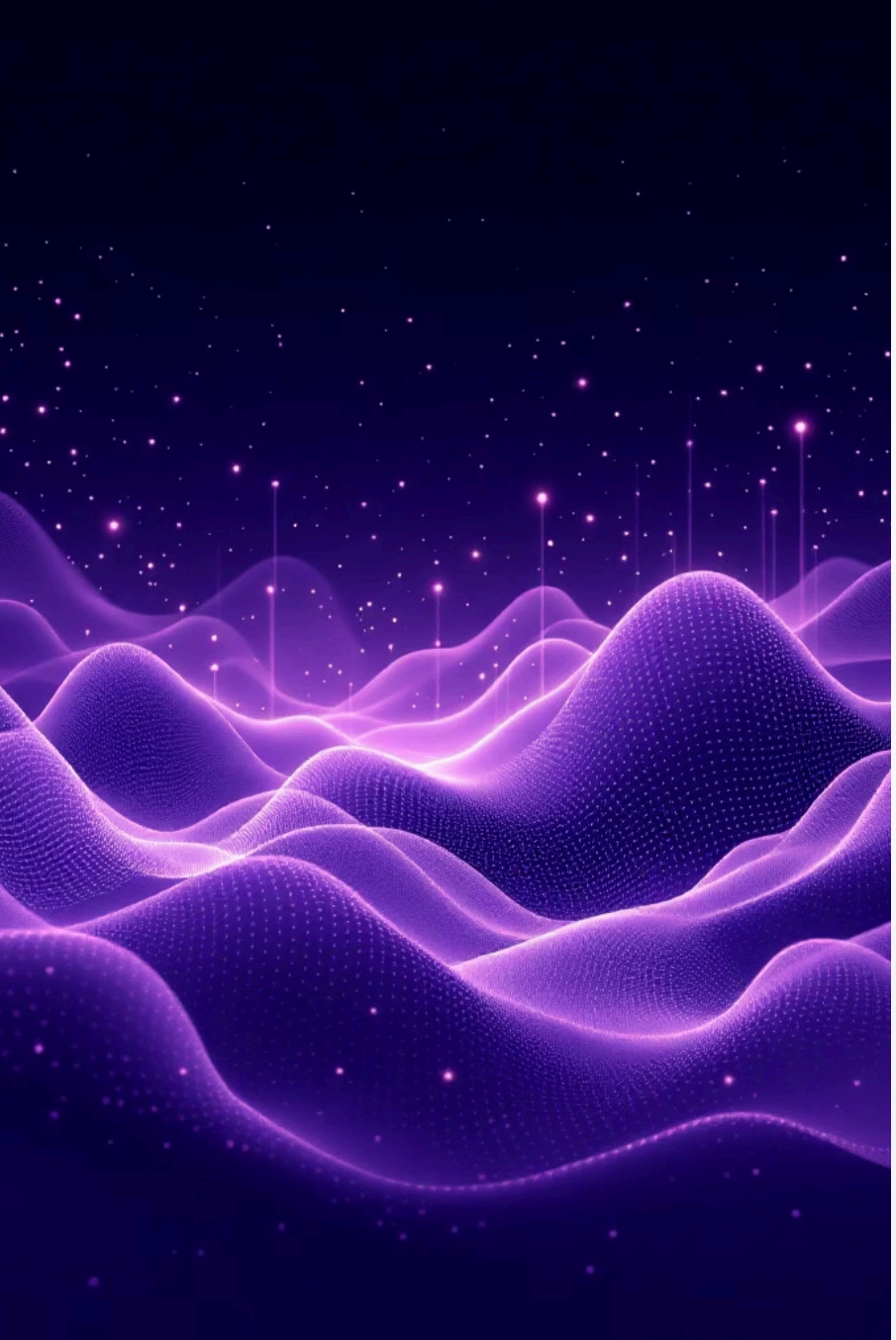


Netflix Data Analysis using SQL

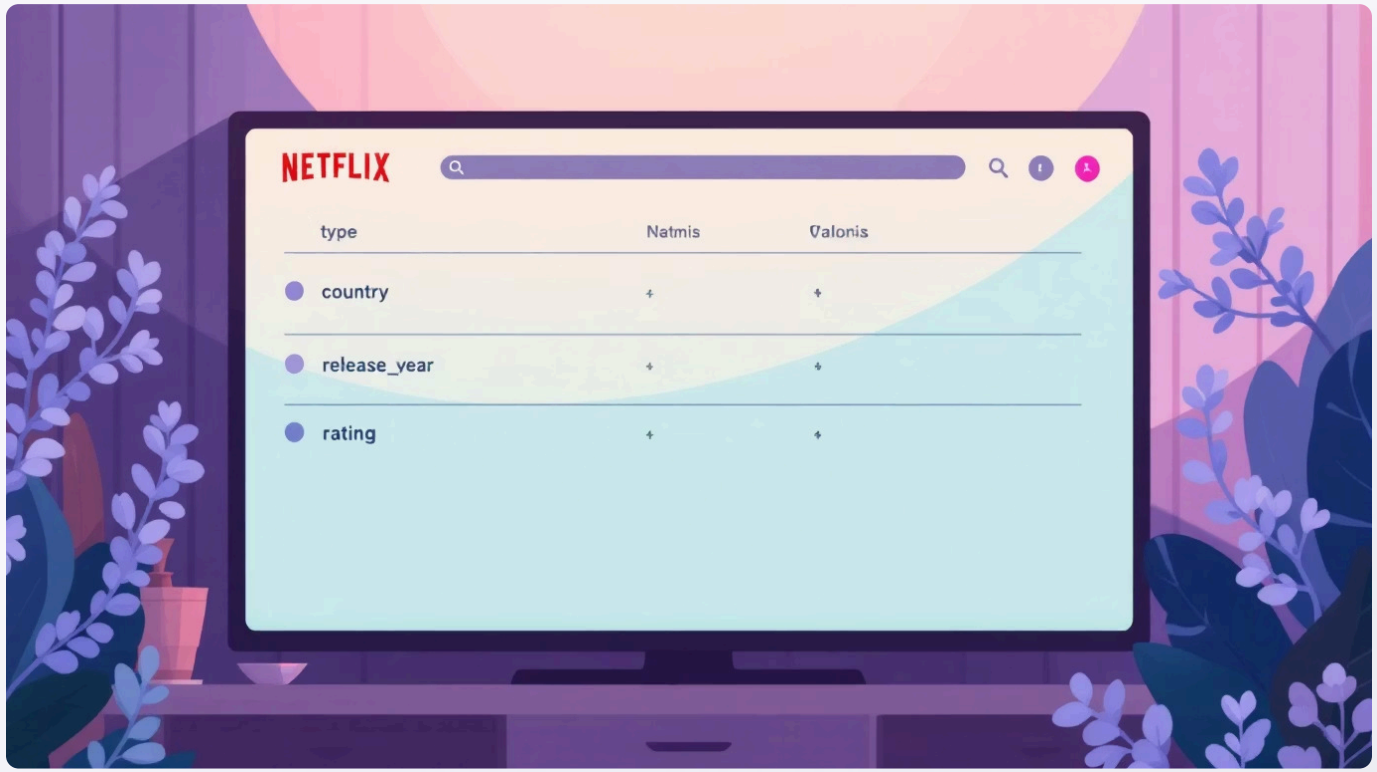
Mini Project #2: Unlocking Insights with MySQL

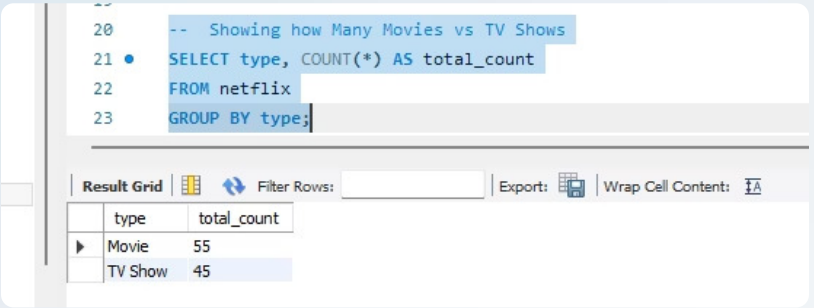
Author: Vansh Tiwari

Key Business Questions

Our analysis aimed to address several crucial questions about the Netflix content library, providing a foundational understanding of its composition and trends:

- Are there more Movies or TV Shows available on the platform?
- Which countries are the primary content producers for Netflix?
- How many titles have been released since 2018?
- What is the most frequently assigned content rating?



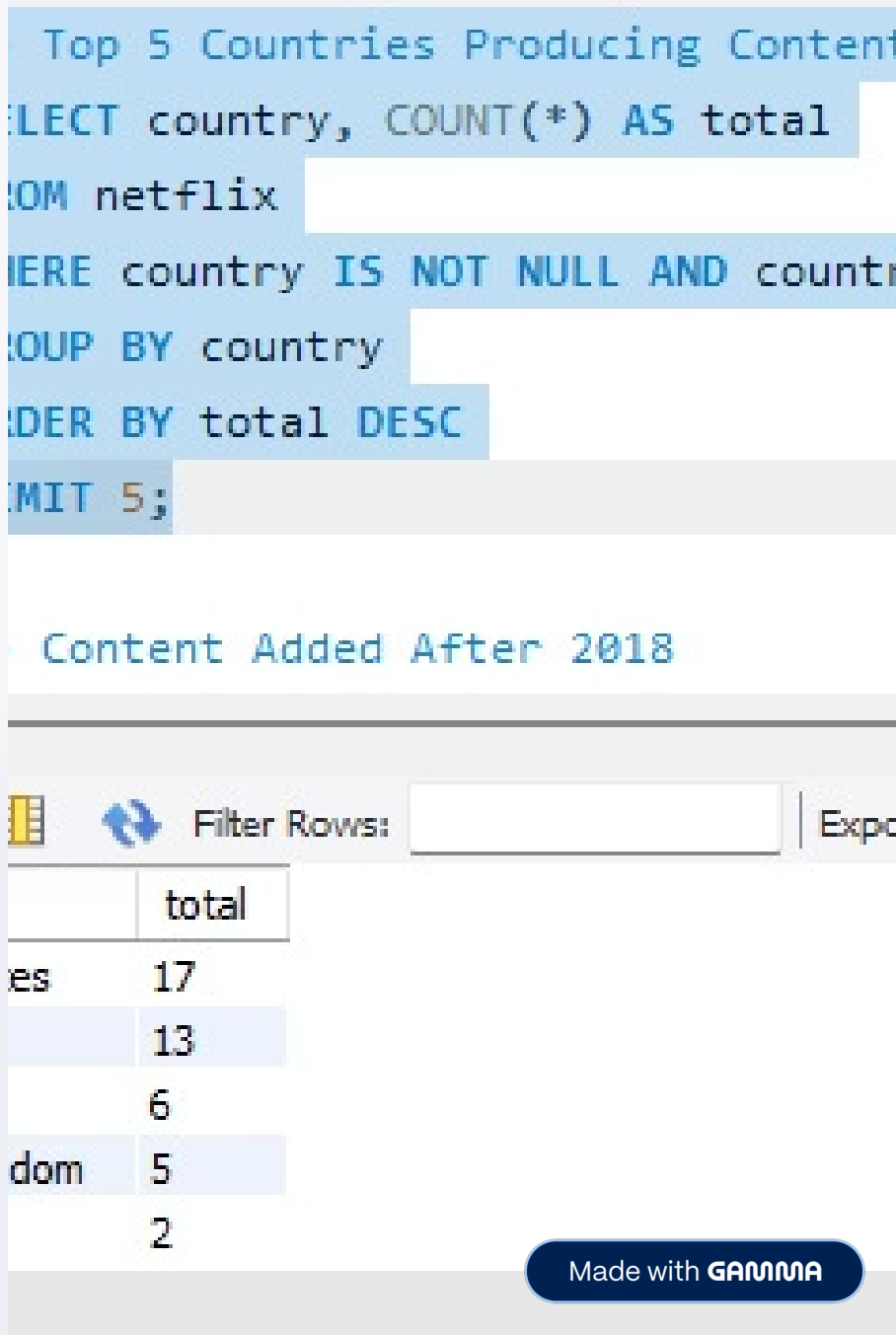


Content Type Dominance

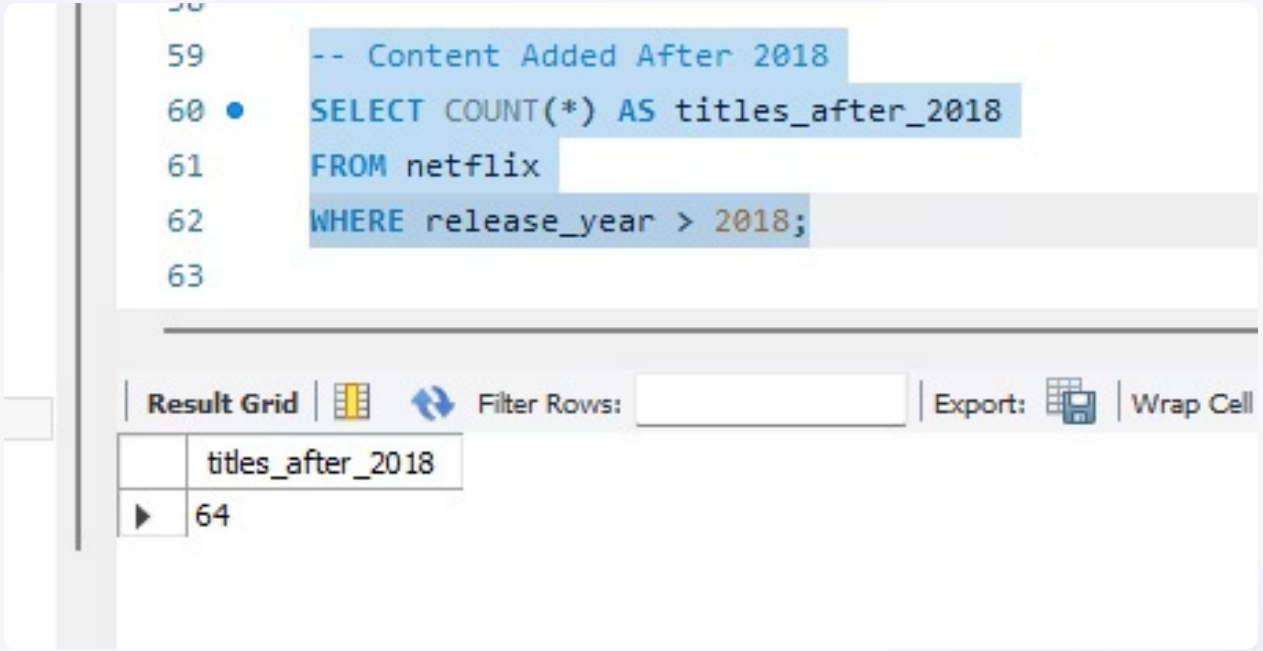
Our findings clearly indicate that movies significantly dominate the Netflix platform in terms of volume compared to TV Shows. This insight is crucial for understanding Netflix's content strategy and potential audience preferences.

Global Content Production Hubs

The analysis of content origin reveals distinct leaders in Netflix's global production landscape. The United States stands as the primary content producer, followed by key regions such as Japan, India, the United Kingdom, and Nigeria. This highlights Netflix's diversified content sourcing strategy.



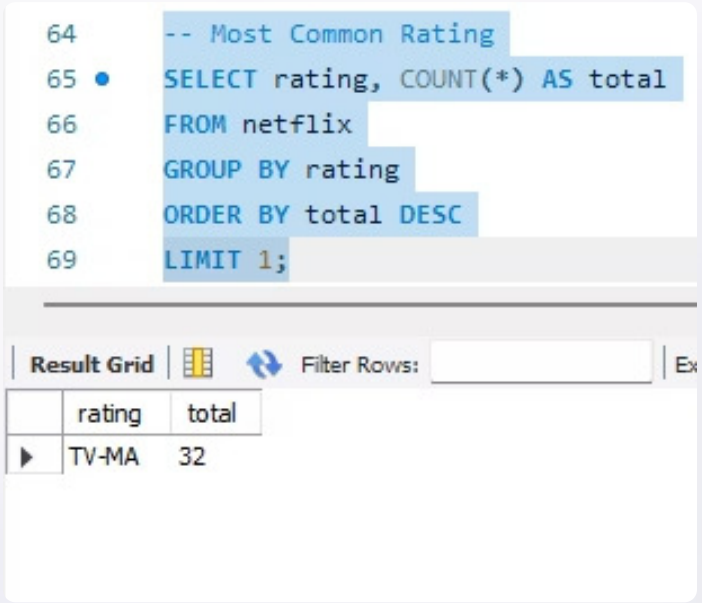
Release Trends and Rating Insights



```
59  -- Content Added After 2018
60  SELECT COUNT(*) AS titles_after_2018
61  FROM netflix
62  WHERE release_year > 2018;
63
```

Result Grid | Filter Rows: | Export: | Wrap Cell

titles_after_2018
64



```
64  -- Most Common Rating
65  SELECT rating, COUNT(*) AS total
66  FROM netflix
67  GROUP BY rating
68  ORDER BY total DESC
69  LIMIT 1;
```

Result Grid | Filter Rows:

rating	total
TV-MA	32

Recent Content Boom

A substantial 64% of all titles were released after 2018, indicating a rapid expansion of the Netflix library in recent years.

Most Prevalent Rating

Our analysis identified TV-MA as the most common rating across the entire Netflix catalogue, suggesting a focus on mature audience content.

Demonstrated SQL & Analytical Skills

SQL Filtering Expertise (WHERE Clause)

Effectively isolated specific datasets based on criteria such as release year and content type.

Aggregation Techniques (COUNT, GROUP BY)

Summarised large volumes of data to derive meaningful metrics, like content type distribution and country-wise production.

Data Sorting Proficiency (ORDER BY)

Organised results logically for clear interpretation and identification of top categories.

Robust Data Interpretation

Translated raw query results into actionable insights about Netflix's content strategy and audience focus.

Insight Generation & Communication

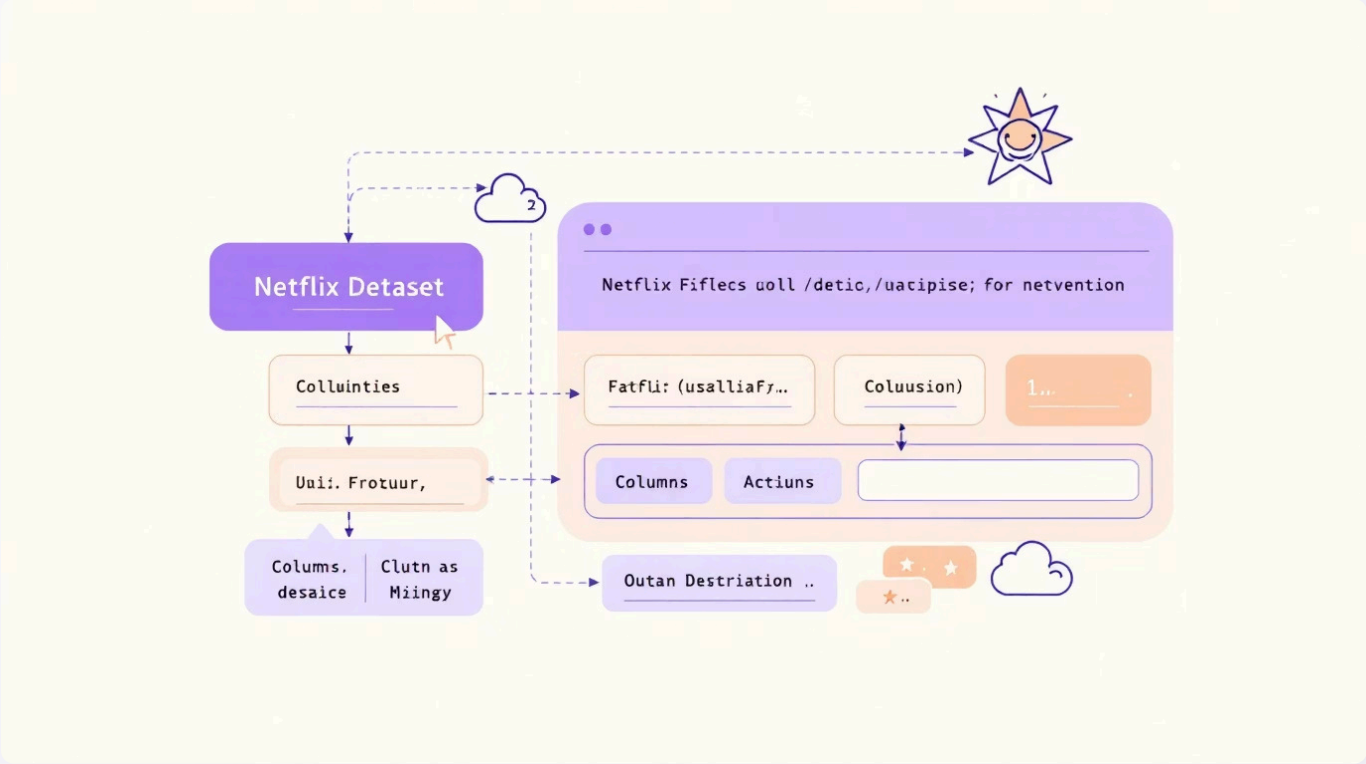
Formulated clear, concise insights from complex data, ready for business decision-making.

More advanced projects coming soon.

Understanding the Dataset: Source & Structure

Data Origin

The dataset utilised for this analysis was sourced from a publicly available Netflix content database, providing comprehensive information on a wide array of titles.



Key Fields Examined:

- **show_id:** Unique identifier for each title
- **type:** Categorisation as 'Movie' or 'TV Show'
- **title:** Name of the content
- **country:** Primary country of production
- **release_year:** Year of original release
- **rating:** Content maturity rating

SQL Workflow: From Query to Insight

Our structured approach ensured a robust and reproducible analysis, moving systematically from raw data to valuable business insights.

1. Data Extraction

Initial queries to retrieve relevant data points from the Netflix dataset, focusing on the columns necessary for answering business questions.

2. Data Transformation

Cleaning and preparing the data, including handling NULL values and standardising formats where necessary to ensure accuracy.

3. Analysis

Application of SQL functions and clauses (e.g., `COUNT()`, `GROUP BY`, `WHERE`) to aggregate and filter the data according to the business questions.

4. Insight Reporting

Synthesising the analytical findings into clear, concise, and actionable insights for stakeholders.

Beyond the Numbers: Strategic Implications

Content Acquisition Strategy

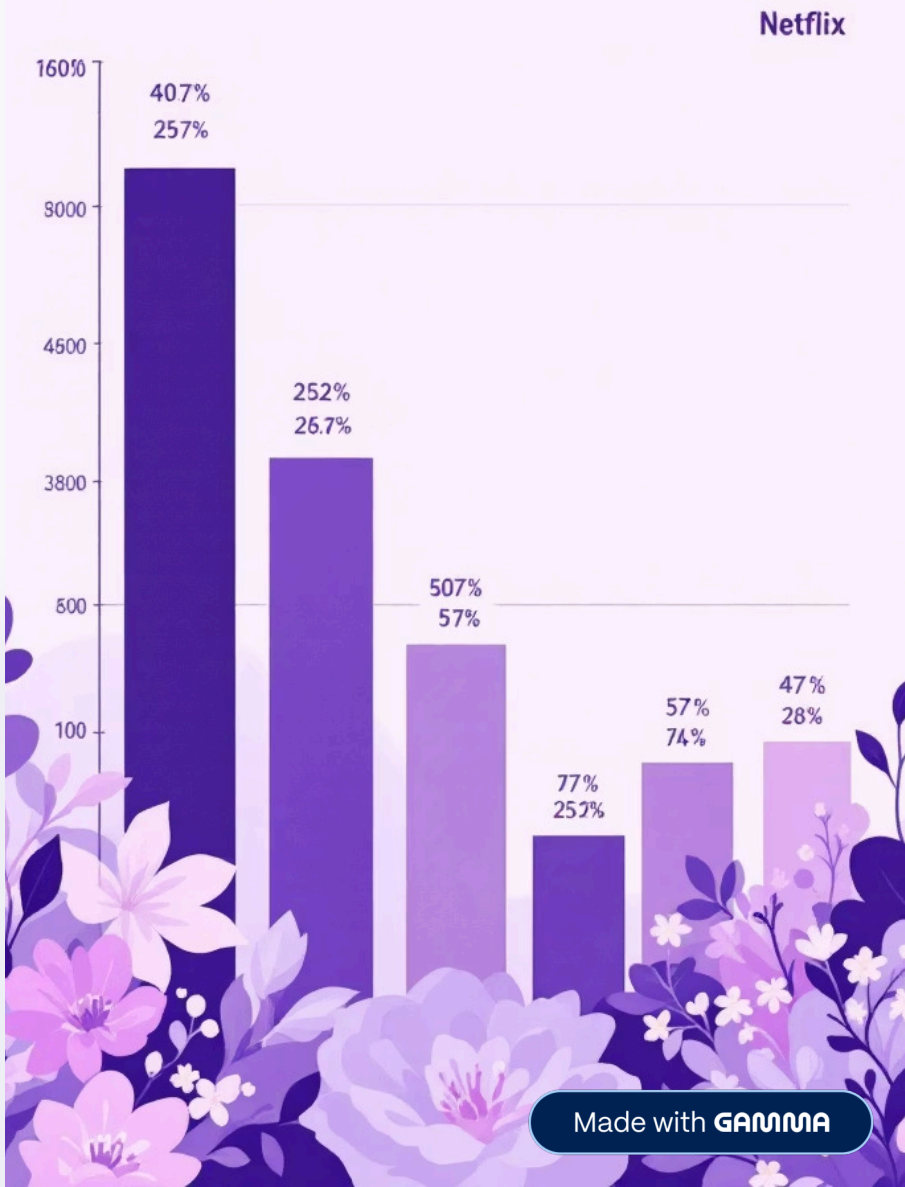
The dominance of movies and the high proportion of recent releases suggest an aggressive content acquisition and production strategy by Netflix.

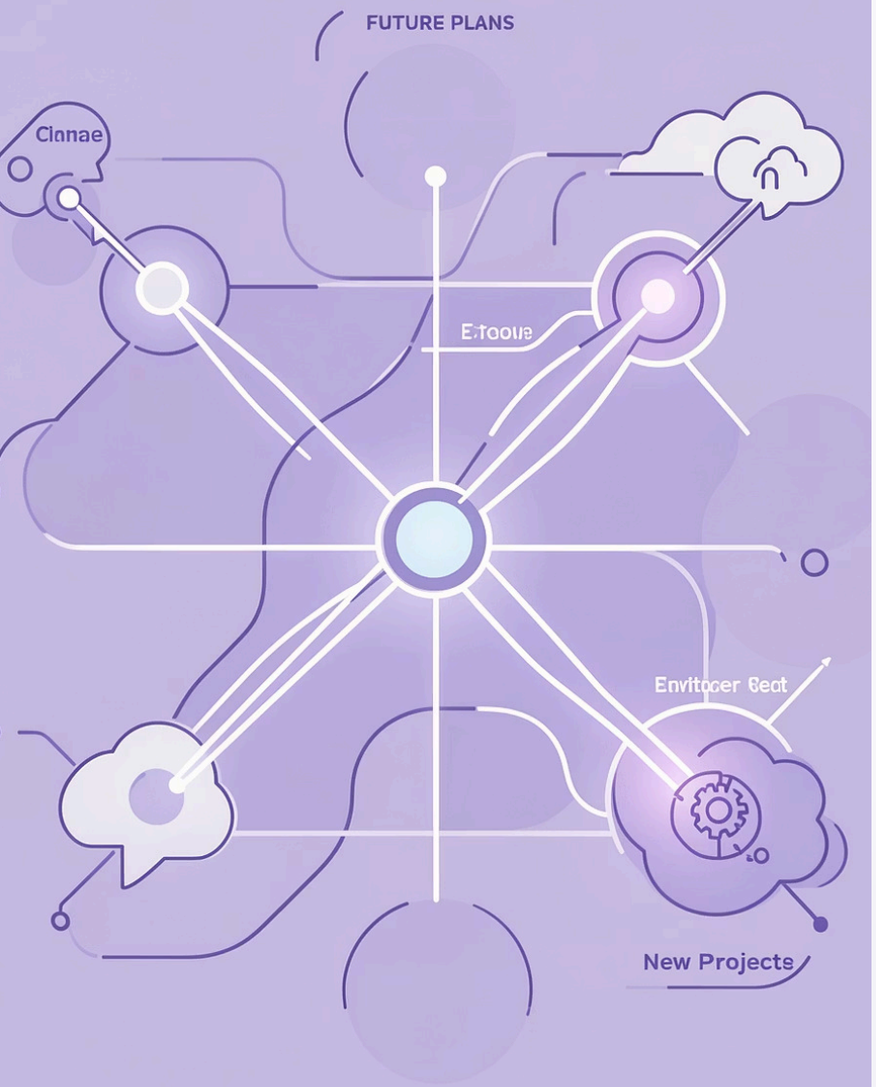
The prevalence of TV-MA content indicates a target audience with a preference for mature programming. This could guide future content investments.

Market Focus & Growth

The diversity of top content-producing countries reflects Netflix's global footprint and its strategy to cater to diverse regional audiences.

Rapid content growth post-2018 shows a clear investment in expanding the library, potentially to attract new subscribers and retain existing ones in a competitive streaming market.





Future Directions & Next Steps

Advanced Analytics

Further analysis using more sophisticated SQL functions and potentially integrating with Python for deeper statistical modelling.

Audience Segmentation

Exploring user behaviour data to understand viewership patterns linked to content types and ratings.

International Market Deep Dive

A more detailed investigation into content preferences and production trends within specific geographic regions.

Predictive Modelling

Developing models to predict content success or subscriber churn based on current library characteristics.

This mini-project serves as a solid foundation; I am eager to apply these analytical skills to more complex challenges.