NETFLIX ANALYSIS

FOR

TECHNOLOGY IN BUSINESS PROJECT

PREPARE BY,

APEKSHA CHIKANE 23030141014

SUBMITTED TO,

SHASHANK GAIKWAD

TABLES OF CONTENTS

INTRODUCTION

Netflix, a global streaming giant, hosts a diverse content library of movies and TV shows, continually expanding to meet the varied tastes of its audience. This project aims to analyze the Netflix dataset to distinguish between movies and TV shows, uncovering trends in content distribution, viewer preferences, and regional contributions. Using Tableau, the analysis will visualize these insights, providing a clearer understanding of Netflix's content landscape and guiding strategic decisions for content curation and platform enhancement.

ABOUT DATASET

Kaggle Link: https://www.kaggle.com/datasets/shivamb/netflix-shows

Records: 6235

Attributes:

Show_id: Unique identifier for each show type:

Indicates whether it's a movie or a TV show title: Title

of the show director: Director of the show cast: Cast

of the show country: Country of origin for the show

date_added: Date when the show was added to

Netflix release_year: Year when the show was

released rating: Age rating for the show duration:

Duration of the show

genre: Categories or genres the show belongs to description:

Brief description of the show

PROBLEM STATEMENT

The Netflix platform hosts a vast and diverse content library, including both movies and TV shows from various countries and genres. However, the sheer volume and variety of content make it challenging to identify key trends, viewer preferences, and regional variations. This analysis seeks to address these challenges by examining the Netflix dataset to uncover meaningful insights into content distribution, viewer demographics, and regional contributions.

OBJECTIVE

The primary objective of this project is to analyze the Netflix dataset to distinguish between movies and TV shows, uncovering trends, patterns, and insights related to viewer preferences, content distribution, and regional variations.
The analysis will focus on:
Temporal distribution of content additions on Netflix.
Ratio of movies to TV shows, offering insights into content diversity and viewer preferences.
Geographical distribution of content, identifying regions with significant content contributions.
Viewer age group preferences, analyzing duration trends and consumption patterns for both movies and TV shows.

KEY PERFORMANCE INDICATORS(KPIs)

Content Upload Trends: Number of movies and TV shows added to Netflix over time, broken down by year
Movies to TV Shows Ratio: The proportion of movies to TV shows available on Netflix, highlighting content diversity.
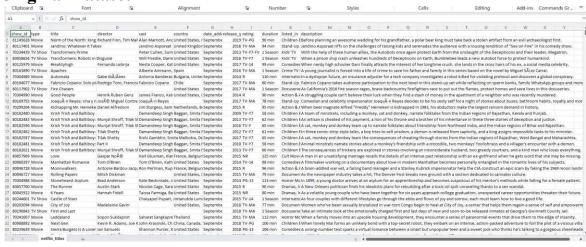Regional Content Distribution: Number of movies and TV shows produced in different countries, identifying key content-producing regions.
Viewer Age Group Preferences: Distribution of content duration and age ratings, providing insights into the preferred content length and age group preferences.
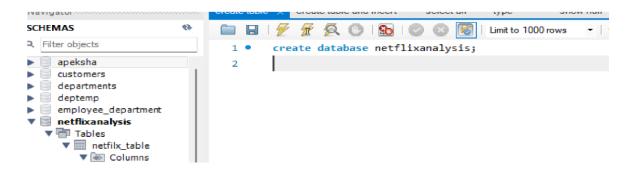Genre Distribution: Analysis of the listed genres/categories, determining the popularity of various content types across different regions and time periods.
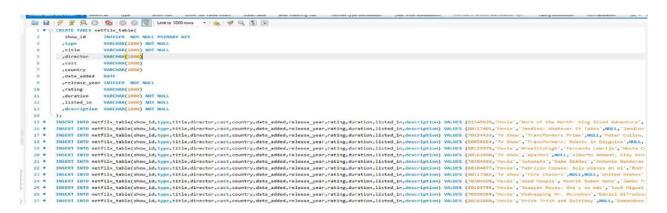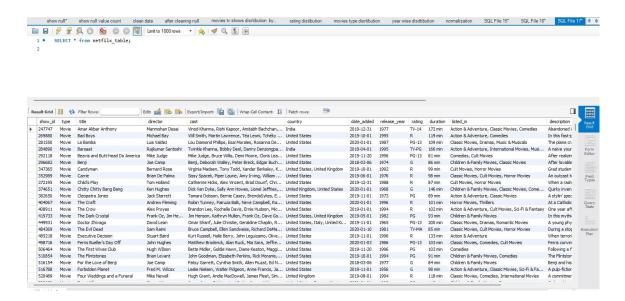
SQL

Original Dataset

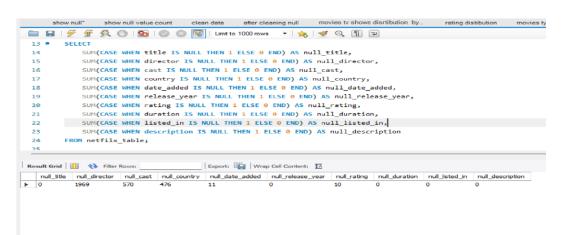## CREATED DATABASE:

CREATED TABLE AND INSERTED VALUES

NETFLIX TABLE DATA

DATA PREPROCESSING

Count no null values column wise

HANDLED NULL VALUES

1. query updates the director column of the netflix_table, setting its value to 'Unknown' wherever the director value is NULL.

2. query updates the cast column in the same table, assigning 'Unknown' where the cast field is NULL.

3. query deletes rows from the netflix_table where any of the columns country, date_added, or rating contain NULL values.

DATA AFTER PREPROCESSING

Now we have 0 null values

Total remaining record 5743

SQL QUERIES FOR ANALYSIS

Movies and Tv shows Distibution


```
1    #Movies & TV Shows Distribution
2 •  SELECT type, COUNT(*) AS count
3    FROM netfilx_table
4    GROUP BY type;
```

type	count
Movie	4066
TV Show	1677

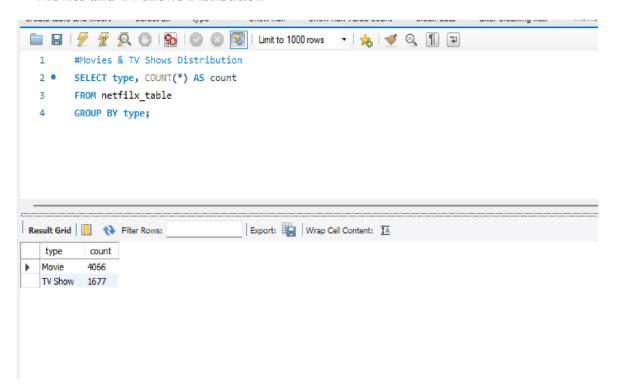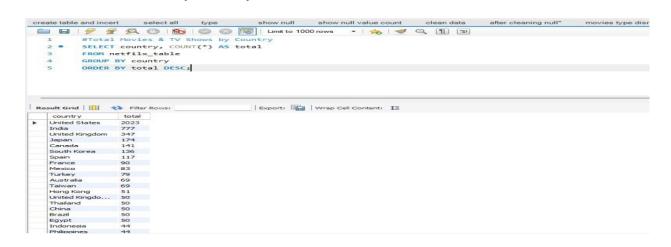This SQL query counts the number of movies and TV shows and groups them by type. It provides a summarized dataset that can be used to quickly create clear charts showing the distribution of content types

Movies and Tv shows by Country


```
1    #Total Movies & TV Shows by Country
2 •  SELECT country, COUNT(*) AS total
3    FROM netfilx_table
4    GROUP BY country
5    ORDER BY total DESC;
```

country	total
United States	2023
India	777
United Kingdom	347
Japan	174
Canada	141
South Korea	136
Spain	117
France	90
Mexico	83
Turkey	79
Australia	69
Taiwan	69
Hong Kong	51
United Kingdo...	50
Thailand	50
China	50
Brazil	50
Egypt	50
Indonesia	44
Philippines	44

This SQL query counts the total number of movies and TV shows for each country and sorts them in descending order. By grouping the data by country and counting the entries, it helps show which countries have the most content on Netflix.

Rating wise Distibution

This SQL query counts the number of movies and TV shows for each rating and orders the results by count in descending order. It helps inshow the distribution of content across different ratings, making it easy to see which ratings are most common.
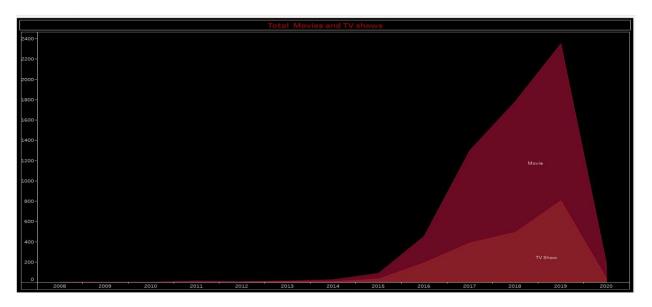
Normalization:

The SQL query creates a netflix_genres table and splits a comma-separated listed_in field into individual genres, storing each genre with its associated show_id.

TABLEAU VISUALIZATION

Total Movies and TV Shows

2019 Peak in Content Additions: The graph shows that 2019 experienced the highest number of content additions, with both movies and TV shows reaching a peak. This suggests that Netflix was heavily focused on expanding its content library during this year.

Significant Growth Leading Up to 2019: The sharp increase in content additions begins around 2015, with steady growth each year, culminating in the 2019 peak.

Movies vs TV Shows: Throughout the timeline, movies consistently outnumber TV shows, reflecting a possible preference or strategic focus on movies during this period

Total Movies and TV Shows by Country:

Insights from the Map Visualization:

Dominance of the United States: The map shows that the United States is a major contributor to Netflix's content library, with the darkest shade indicating the highest number of movies and TV shows. This aligns with the understanding that the U.S. is a significant hub for film and television production.

Strong Contributions from India: India also stands out as a significant content provider, with a large number of movies and TV shows. This suggests that Netflix has a strong presence in the Indian market, likely driven by the country's prolific film industry (Bollywood) and regional cinema.

Other Key Contributors: Countries like the United Kingdom, Canada, and Japan also contribute noticeably to Netflix's content, though to a lesser extent than the U.S. and India. These regions are known for their established entertainment industries.

Sparse Contributions from Other Regions: The map shows that many regions, particularly in Africa and some parts of Asia, have minimal content contributions. This could indicate either a lower production volume in these regions or that Netflix has not yet fully tapped into these markets.
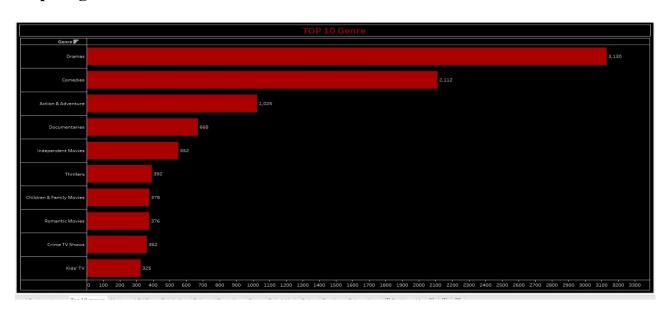
Regional Content Focus: The variation in content contributions across different regions highlights Netflix's focus on specific markets where the production of movies and TV shows is more robust. This focus could also reflect regional content preferences and the availability of local productions.

Conclusions:

Content Strategy: Netflix's content library is heavily influenced by productions from the U.S. and India, suggesting a strategy that leverages the strengths of these major entertainment hubs.

Market Opportunities: There may be opportunities for Netflix to expand its content offerings in regions that currently contribute less, potentially by investing in or acquiring more local productions from underrepresented areas.
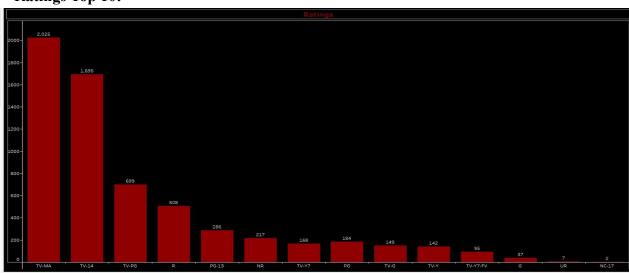
Top 10 genres:

Insights:

Dramas are the most frequent genre on Netflix, followed by Comedies, Sports Movies and action and Adventure. Kids' TV has the fewest entries among the top genres, suggesting a lesser focus on children's content compared to others.

Ratings Top 10:

G: All Ages

TV-Y: All Children

TV-Y7: Older Children

TV-G: General Audience

TV-PG / PG: Parental Guidance Suggested

PG-13 / TV-14: Parents Strongly Cautioned

R: Restricted
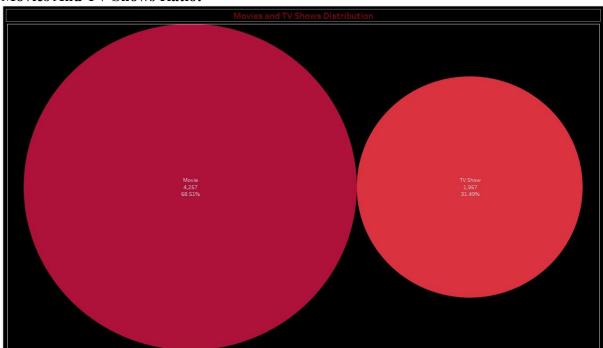
TV-MA: Adult Audience

Insights:

The largest portions of content are aimed at mature audiences (TV-MA, TV-14, and R), indicating that Netflix has a strong focus on adult-oriented content.

There is still a substantial amount of content suitable for younger audiences, although it makes up a smaller proportion of the overall library.
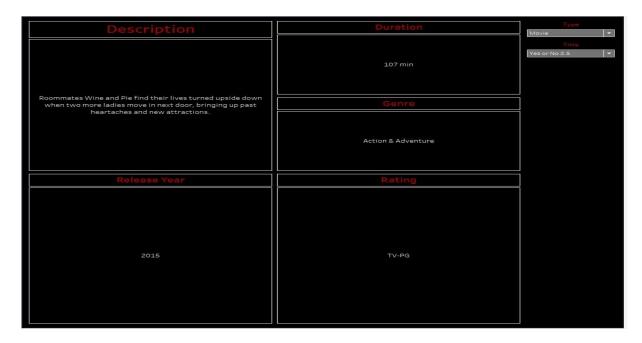
The distribution suggests that Netflix is heavily weighted towards providing content that appeals to teens and adults, with fewer options available for younger children.

Movies And TV Shows Ratio:

Movies and TV Shows Distribution

Movie
4,257
68.51%

TV Show
1,957
31.49%

Insights: The bubble chart shows that Netflix has a higher proportion of movies (68.51%) compared to TV shows (31.49%). This indicates a significant focus on movies in Netflix's content library.

Other Sheets:
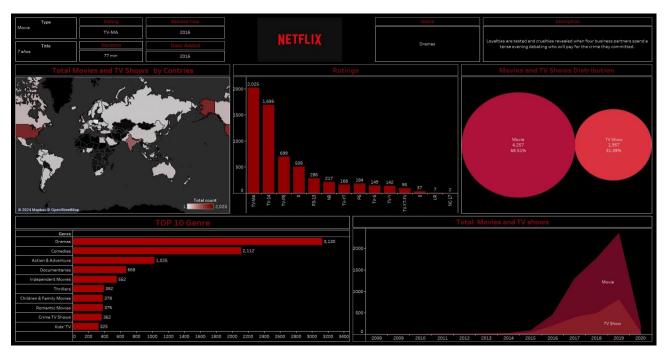

In Tableau visualization used filters to display details about a selected Netflix movie And Tv Shows

Type filter, Title filter Description: Provides a summary of the movie's plot.

Duration: Shows the length of the movie, Genre: Indicates that the movie and Tv shows belongs to which genre. Release Year: Displays the year the movie was released.

Rating: Shows the movie's audience rating.

NETFLIX DASHBOARD

OVERALL INSIGHTS AND CONCLUSION

Content Focus: Netflix's content library is heavily focused on movies, which constitute about (68.51%) compared to TV shows (31.49%). This suggests a strategic preference for movies over TV shows, potentially due to their broader appeal or shorter production timelines.

Genre Popularity: The most popular genre on Netflix is drama, followed by comedies, sports movies, and action & adventure. This indicates a strong audience preference for narrativedriven content and diverse genres that cater to various tastes.

Regional Content Distribution: The United States and India are the leading contributors to Netflix's content library, highlighting these countries' significant role in global entertainment. Other countries like the UK, Canada, and Japan also contribute, but there is less content from Africa and some parts of Asia, indicating potential areas for growth and expansion.

Content Trends: There was a significant increase in content additions around 2019, suggesting a period of rapid expansion for Netflix. This could be attributed to Netflix's efforts to strengthen its catalog to compete with other streaming services.

Viewer Demographics: A large portion of Netflix's content is targeted at mature audiences (TV-MA, TV-14, and R), with less content available for younger viewers. This shows a strategic focus on adult content, which aligns with the viewing habits of a significant portion of Netflix's subscriber base.

Content Addition Over Time: Netflix saw the highest number of content additions in 2019. This aligns with their strategy to rapidly expand their library in response to growing competition in the streaming market.

SUGGESTIONS

Expand TV Show Library: To balance the content offering and attract more subscribers who prefer episodic content, Netflix could consider expanding its TV show library.

Increase Content for Younger Audiences: With a lower percentage of content aimed at younger audiences, Netflix could develop or acquire more children's programming to cater to families and younger viewers.

Focus on Underrepresented Regions: By investing in content from underrepresented regions, particularly Africa and parts of Asia, Netflix can diversify its content offerings and appeal to a broader global audience.

Genre Diversification: While drama is the most popular genre, diversifying the genres further by investing in lesser-represented genres could attract niche audiences and provide a more varied content library.

LIMITATIONS

Dataset Limitations: The dataset from Kaggle only provides a snapshot of Netflix's content at a specific time and may not reflect real-time updates. Additionally, it lacks detailed viewer statistics and engagement metrics, which are crucial for deeper analysis of content preferences.

Tool Limitations: Tableau is effective for visualization but has limited capabilities for advanced data processing and statistical analysis. Complex data transformations and predictive modelling would require additional tools like Python or R.

TOOLS USED

Tableau for data visualization

SQL for data preparation and cleaning

REFERENCES

Dataset source: Netflix Movies and TV Shows dataset on Kaggle

https://www.tableau.com/academic/students