Final Project Report

Group 13

Anirudh Suryanarayan, Amandeep Singh, Bishal Sharma Roy,
Samarth Yaralakatte Mallappa, Tanmoy Madbar

Title: The Role of Data Visualization in Communicating Viewer Trends to Decision Makers in OTT Platforms: A Netflix Case Study

1. Introduction

In today's digital economy, Over-The-Top (OTT) streaming platforms like Netflix generate immense volumes of data on content, viewership, and audience preferences. As the industry becomes increasingly saturated, content providers face pressure to differentiate their offerings through personalized, data-driven strategies. The sheer scale and complexity of data make it challenging for decision-makers to derive actionable insights without effective analytical tools. This is where data visualization plays a critical role.

This project explores how visualization using Power BI can enable Netflix decision-makers to better understand viewer behavior, content performance, and market dynamics. By transforming complex datasets into intuitive visuals, business leaders can more efficiently identify high-performing genres, optimize runtime strategies, and tailor content portfolios to user segments. Through this case study, we aim to demonstrate how interactive dashboards and targeted visualizations can support real-time decision-making in a competitive OTT environment.

Our analysis is centered on the publicly available Netflix TV Shows and Movies dataset from Kaggle. We build visualizations that help answer key managerial questions, such as: How has Netflix's content evolved over time? What genres resonate most with audiences? Are runtime patterns associated with engagement levels? And do TV shows outperform movies in viewer ratings?

The report follows a structured approach beginning with clearly stated objectives and analytical setup, followed by our Power BI visual outputs, interpretations, and strategic recommendations tailored to Netflix's content and viewer management strategy.

2. Objective / Question

The primary objective of this project is to demonstrate how data visualization can aid in strategic decision-making by uncovering patterns and insights within the Netflix content library. We aim to transform complex viewer and content data into meaningful visuals that support evidence-based managerial decisions.

The project is guided by the following key research questions:

1. How has Netflix's content library evolved over time in terms of release trends and content type (TV vs. Movies)?
2. What are the most popular and highest-rated genres, based on IMDb scores?
3. What patterns exist between runtime and audience engagement, including IMDb scores and TMDB popularity?
4. Do TV shows receive higher viewer ratings than movies on Netflix?
5. How can interactive Power BI dashboards help decision-makers identify content strategies?

3. Assumptions

In conducting our analysis, we made the following assumptions to ensure consistency and interpretability:

- Only titles with non-null IMDb scores were considered to ensure valid engagement data.
- For genre-based analysis, only the primary genre (first listed) was used to simplify comparisons.
- Runtime values above 200 minutes were considered outliers and excluded from scatter visualizations.
- TMDB popularity was treated as a proxy for audience engagement, acknowledging it may reflect both marketing and viewership.
- We assumed that Power BI's default aggregation (e.g., Average) was appropriate unless manually adjusted.

4. Analytical Setup & Analysis

Tools Used: Microsoft Power BI Desktop

Data Source:

- Netflix TV Shows and Movies Dataset from Kaggle: https://www.kaggle.com/datasets/victorsoeiro/netflix-tv-shows-and-movies

Pre-processing Steps:

- Loaded the dataset into Power BI and removed entries with missing IMDb scores.
- Filtered dataset to separate movies and shows for comparative analysis.
- Extracted the primary genre from the list of genres.
- Created a custom runtime bin column to group runtimes into ranges (e.g., 0-30, 31-60 min).

Visualizations Developed:

Figure 1: Titles Released Per Year

- Shows Netflix's exponential growth in content volume from the early 2000s to 2021.

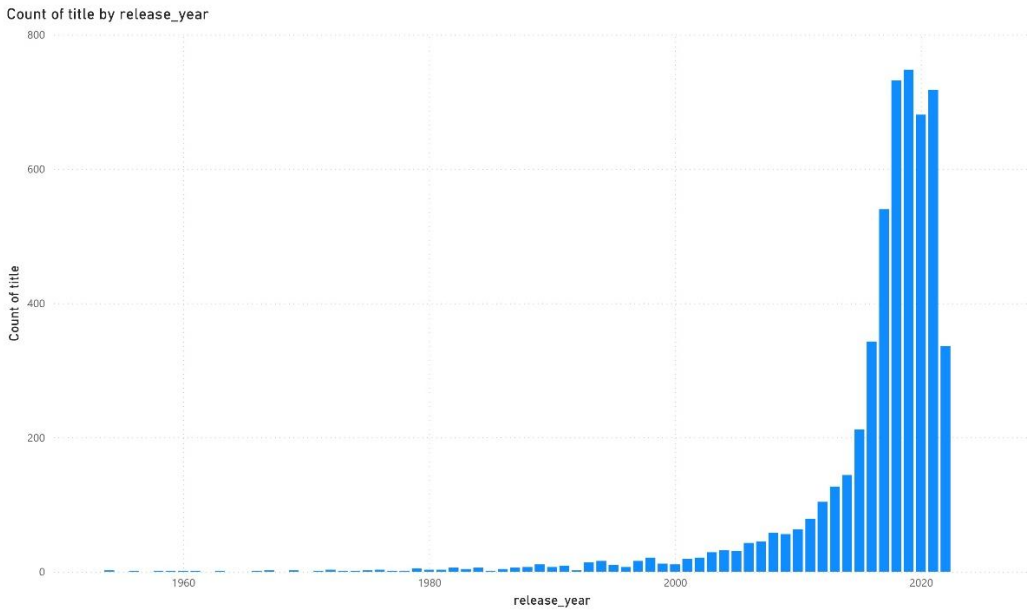


Figure 2: Average IMDb Score by Genre

- Documents and war genres top the list, with horror and romance receiving the lowest average scores.

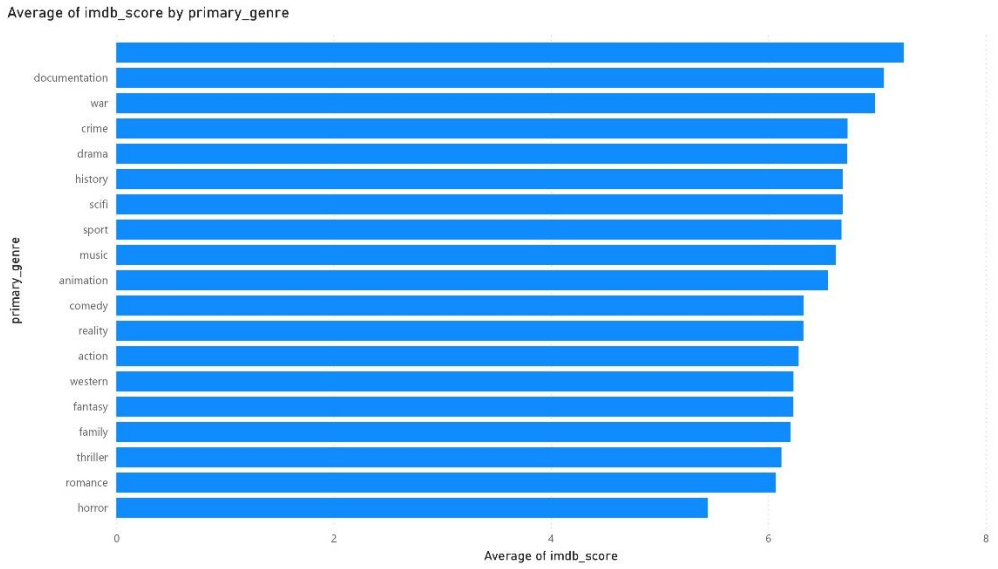


Figure 3: Runtime vs. Engagement

- A scatter plot based on average IMDb and TMDB scores per runtime bin. Shows a performance peak in 90-120 minute range.

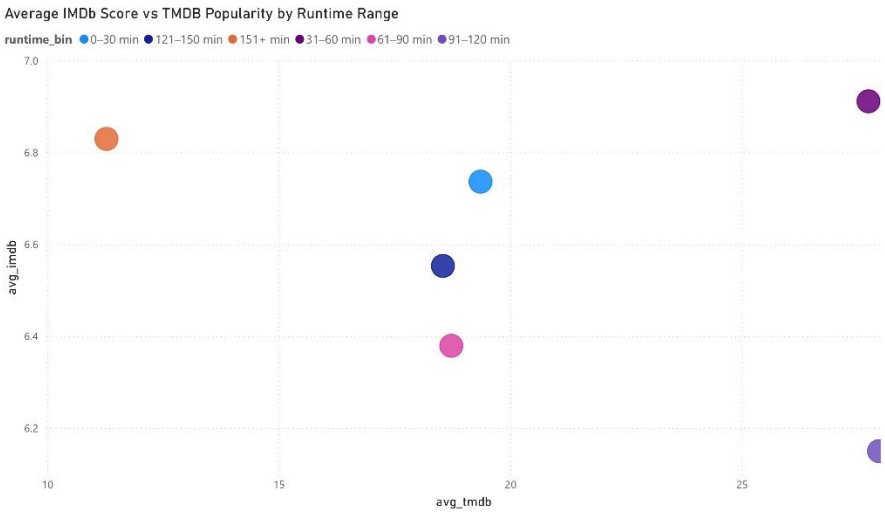
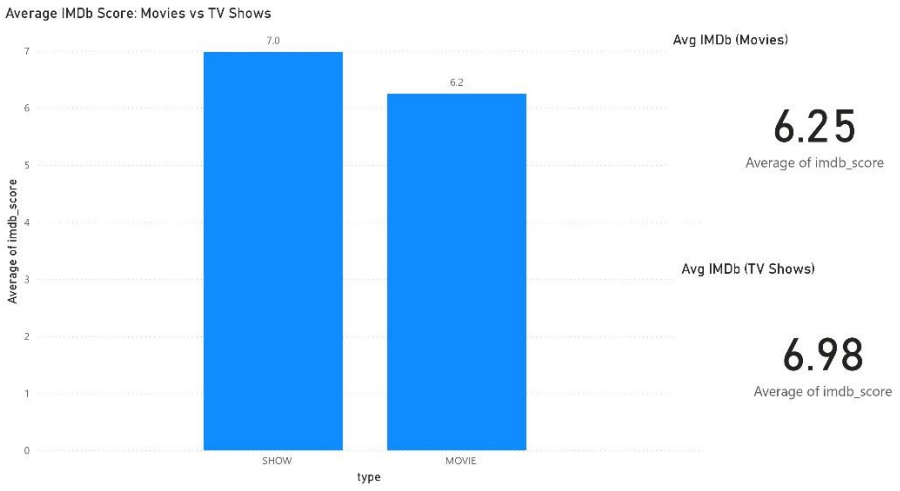


Figure 4: Movies vs. TV Shows Comparison

- Reveals that TV shows have a significantly higher average IMDb score (6.98) than movies (6.25).



Each visual was developed to answer a specific research question and support insights into content performance and strategic implications for Netflix.

5. Implications & Recommendations

Based on the analysis and visual findings, several key implications emerge for content strategy and platform optimization at Netflix:

- **Strategic Genre Investment:** Genres like documentation, war, and drama consistently yield higher IMDb ratings. Netflix should consider prioritizing these genres in future content acquisition and production strategies.
- **Optimizing Runtime:** The scatter plot indicates that most viewer-favored titles fall within the 90–120 minute range. Netflix can guide its content development teams to target runtimes within this sweet spot for movies, while also maintaining flexibility for genre-specific exceptions.
- **TV Shows Over Movies:** The visual comparison reveals that TV shows generally outperform movies in terms of audience ratings. Netflix should consider this while allocating production budgets, especially in high-performing genres.

- **Audience Engagement Insights:** Runtime bins combined with TMDB popularity scores help identify content formats that maximize both quality and engagement. This can be applied to tailor recommendations and design platform features for audience retention.
- **Data-Driven Dashboards for Stakeholders:** The Power BI dashboard serves as a scalable solution to visualize key KPIs. It empowers decision-makers to track viewer trends in real time, evaluate genre health, and act proactively to shifting audience preferences.

6. Limitations and Conclusion

Limitations:

- The dataset does not include detailed viewership statistics, which limits the ability to correlate content attributes with actual watch time.
- TMDB popularity is used as a proxy for engagement, but it may not accurately reflect viewer sentiment or Netflix-exclusive interactions.
- IMDb ratings are subject to bias and may not represent the opinion of Netflix's target audience segments.
- Some genres and runtime bins had relatively low sample sizes, which may affect the reliability of averages.
- Temporal shifts in content strategy (e.g., during the pandemic) are not explicitly modeled or adjusted for.

Conclusion: This report has demonstrated how Power BI can effectively support managerial decision-making by transforming raw streaming content data into clear, actionable visual insights. Through a series of targeted visualizations, we analyzed genre performance, viewer engagement, and content format preferences, yielding strategic takeaways for Netflix. As OTT competition intensifies, the ability to visualize data trends will remain essential to guiding content development, curation, and platform growth.

7. Citations and References

- Netflix TV Shows and Movies Dataset. Kaggle. Retrieved from https://www.kaggle.com/datasets/victorsoeiro/netflix-tv-shows-and-movies

- IMDb. (n.d.). https://www.imdb.com
- TMDB (The Movie Database). (n.d.). https://www.themoviedb.org
- Microsoft Power BI Documentation. (n.d.). https://learn.microsoft.com/en-us/power-bi/

Additional in-text citations were not required as all data was derived from the above public sources.

NOTE: ChatGPT was used to improve the punctuation and enhance the sentence formations.