



Digital Egypt Builders Initiative (DEBI)

Netflix Analysis

Group 1

Submitted by:

Ahmed Atef

Marvy Ayman

Nehal Gebril

Yomna Shahin

Yasmeen El said

Supervised by:

Dr. Eman Ghanem

1. Problem Overview

Netflix's extensive library of movies and TV shows presents a variety of challenges and opportunities for analysis. The dataset includes key attributes such as genres, ratings, release years, duration, regional production, and audience preferences. Analyzing this data will help uncover patterns and trends that can guide Netflix's content strategies and operational improvements. The primary challenges in the dataset include:

Key Analysis Challenges

1. Content Distribution Imbalance: A significant number of movies compared to TV shows, with regional disparities in production.
 2. Genre and Rating Variability: Different regions show varying preferences for genres and ratings, complicating content strategy.
 3. Data Gaps: Ambiguities and missing values, especially in the duration and region columns, need to be addressed for accurate analysis.
 4. Emerging Markets: Identifying rising content production from countries like South Korea and Spain to ensure a diverse global offering.
-

2. Analysis Goals

The main goal of this analysis is to provide actionable insights into Netflix's global content strategy. This involves understanding the dynamics of content distribution, regional contributions, audience preferences, and temporal trends.

Key Analysis Goals

- Regional and Country-Level Insights: Understand which regions and countries contribute the most to Netflix's library and identify any emerging markets.
- Genre and Ratings Trends: Investigate popular genres and ratings across different regions and their evolution over time.
- Content Evolution Over Time: Analyze how Netflix's library has expanded over the years and identify key seasonal trends in content release.

- Duration and Seasons Analysis: Assess the duration of movies and the number of seasons for TV shows across regions.
 - Audience Preferences: Provide insights into genre and rating preferences, and identify opportunities to optimize content offerings for different demographic groups.
-

3. Key Questions to Address

Based on the dataset, several key questions are being explored to guide insights and strategic recommendations:

3.1 Regional and Country-Level Insights

- Which regions and countries contribute the most to Netflix's content library?
- What is the movie-to-TV show ratio by region and country?

3.2 Genre and Rating Trends

- Which genres are most popular in movies and TV shows across regions?
- What are the most common ratings for movies and TV shows?
- How do rating preferences vary by region?

3.3 Content Growth and Temporal Trends

- How has the number of titles added annually evolved since 2012?
- Are there seasonal trends in the release of titles?

3.4 Duration and Seasons Insights

- What is the average duration of movies in hours across regions?
- What is the average number of seasons for TV shows by region?

3.5 Actors and Directors Insights

- Who are the most featured actors and directors for movies and TV shows?

4. Data Cleaning and Transformation

Prior to analysis, the dataset was cleaned and transformed to ensure consistency and enable deeper insights. The key steps include:

- **Handling Missing Data:** Missing or ambiguous data in columns like *country*, *Genres*, *Actors*, *duration*, *seasons*, and *region* were addressed through imputation or exclusion where necessary.
- **Unuseful Data:** like *show_id* and *description*.
- **Column Derivation:** Derived new columns to clean up and categorize data. Notably, we derived a *duration* column for movies and a *seasons* column for TV shows to enable more accurate comparisons between content types.
- **Region Column Standardization:** A new column for *region* was created to standardize regional data, improving the clarity of geographical insights.

5. Key Insights

5.1 Content Distribution

- **Total Content:** Netflix's library contains 8,803 titles, of which 6,127 are movies and 2,676 are TV shows. Movies dominate, constituting approximately 70% of the total library.

Regional Contributions:

- **North America:** The largest contributor, producing 2.6K movies and 1.0K TV shows.
- **Asia:** The second-largest contributor with 1.5K movies and 0.6K TV shows.
- **Europe:** Generates 1.0K movies and 0.5K TV shows, with a focus on British TV Shows and Documentaries.
- **Emerging Markets:** South Korea and Spain are growing regions, contributing to specific genres like K-Dramas and Spanish-language series.

5.2 Genre Preferences

- Top TV Show Genres: International TV Shows, Crime TV Shows, and Kids' TV.
- Top Movie Genres: Dramas, Action & Adventure, and Comedies.

Regional Genre Trends:

- North America: Focus on Dramas and Action & Adventure.
- Asia: Dominated by Anime and International TV Shows.
- Europe: British TV Shows and Documentaries are popular.

5.3 Duration and Time Trends

- Movie Duration: Average movie duration is between 1.5 to 2.5 hours, with slight regional variations.
- TV Show Seasons:
- North America: More multi-season TV shows.
- Asia: A higher proportion of single-season TV shows.
- Content Growth Over Time: Significant growth in content addition between 2016 and 2020. Higher release volumes align with summer and winter seasons, reflecting audience activity peaks.

5.4 Ratings Distribution

- TV Shows: Common ratings include TV-MA, TV-14, and TV-Y7.
- Movies: Popular ratings include R, PG-13, and PG.
- Regional Variations: Different regions prefer different ratings, with North America leaning towards mature ratings (TV-MA, R) and Asia focusing more on family-friendly ratings (TV-Y7, PG).

6. Recommendations

6.1 Content Strategy

- Increase TV Show Production: Focus on multi-season TV shows, especially in North America and Asia, to cater to rising demand for episodic content.

- Focus on Popular Genres: Invest more in Dramas and Crime TV Shows for North America and Asia, and expand Anime offerings in Asia.
- Diversify Genre Offerings: Tailor content to regional preferences, such as more British TV Shows for Europe and localized content for emerging markets.

6.2 Regional Expansion

- Strengthen Emerging Markets: Collaborate with content creators in South Korea and Spain to diversify Netflix's library and appeal to local audiences.
- Leverage Regional Content: Expand genre offerings such as Anime in Asia and Documentaries in Europe.

6.3 Audience Engagement

- Optimize Ratings: Focus on content with ratings like TV-MA for TV shows and PG-13 for movies to capture a broader audience.
- Promote Family-Friendly Content: Strengthen libraries of kids' TV shows and family-friendly movies to attract younger audiences.

6.4 Seasonal Content Planning

- Plan Major Content Releases: Align releases with audience activity peaks, particularly in summer and winter.
- Forecast Future Growth: Use historical trends to anticipate future content demands and adjust production schedules accordingly.

6.5 Operational Improvements

- Enhance Data Completeness: Address missing or ambiguous data to refine future insights, particularly in areas like director and actor information.
- Optimize Production: Regularly track top actors, directors, and genres to refine Netflix's casting and content production strategies.

7. Conclusion

By leveraging detailed insights into content trends, regional contributions, and audience preferences, this analysis helps to inform Netflix's strategic

direction. With actionable recommendations on content creation, regional targeting, and audience engagement, Netflix can continue to optimize its content library and maintain its global leadership in streaming entertainment.