



Capstone Project: Entertainment EDA Report

BY

KEDEI ALFRED EYONG

Exploratory Data Analysis Report

Introduction

This EDA Report provides a comprehensive analysis of entertainment content, exploring key trends and patterns within the given dataset. The investigation is centered around analytical entertainment insight, core insight on countries, periodic trends and key duration insights. Entertainment insights indicate audience preferences, prominent directors, and popular entertainment genres. Country insights examine the geographical distribution of content production, focusing on leading countries and those at the bottom. Periodic trends explore content release patterns and lifespan of content from its release date. Duration insights investigate content length across countries and distributions of movies and TV shows. Through intricate analytical questions, dynamics of the entertainment sphere is explored.

Analytical Questions

- What is the entertainment preference between Movies and TV Shows?
- Who are prominent entertainment directors?
- What entertainment genres are the most popular?
- What are the leading countries with the most entertainment content?
- What countries provide least entertainment content?
- What is the distribution of entertainment content per year?
- What is the content lifespan after entertainment release?
- What countries provide the highest duration of entertainment content?
- What is the duration of movies in minutes?
- What is the duration of TV shows in seasons?
- What countries have the highest length of entertainment content?

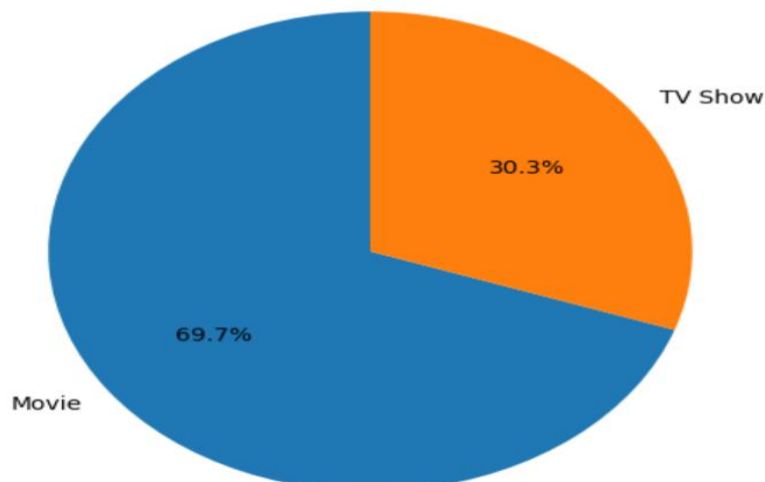
Key Insights

1. Entertainment Insight

A. Entertainment Preference Insight

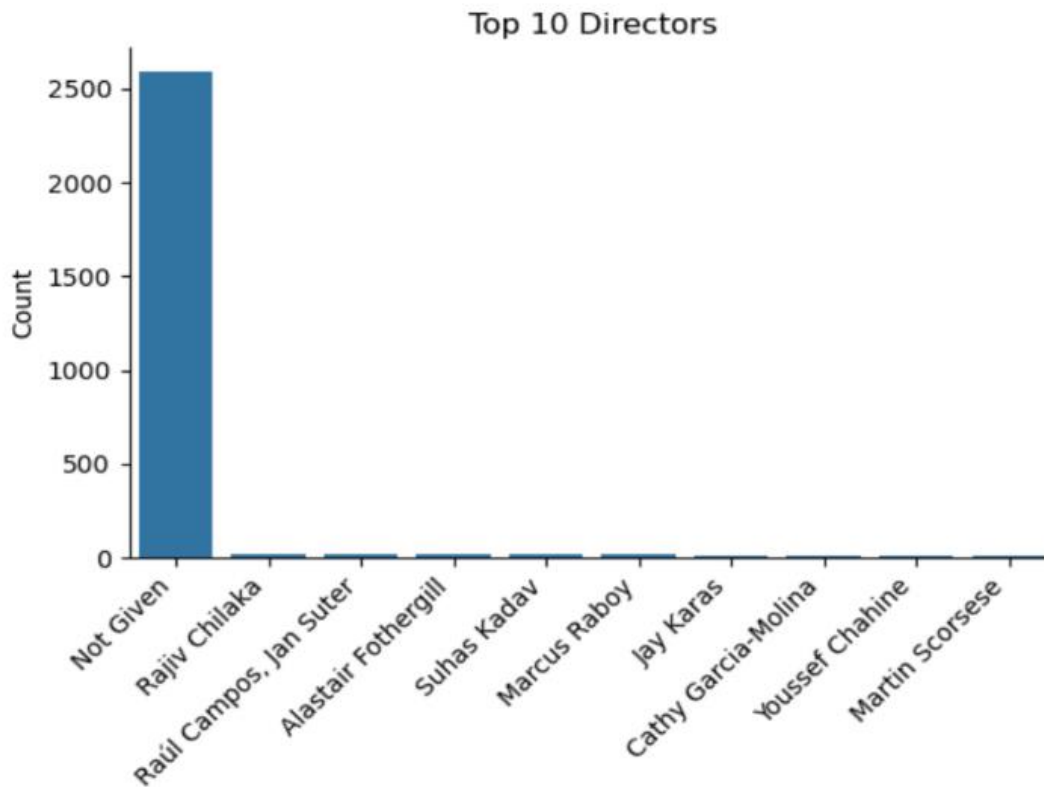
The preferred entertainment type is skewed towards movies, comprising 69.7% of the total content, while TV shows make up 30.3%. This suggests a strategic focus on a wide variety of movie categories that attract and retain high viewership. Moreover, the significant proportion of TV shows indicates that this entertainment type likely caters to a substantial segment of the market.

Entertainment Preference: Type Distribution



B. Top 10 Entertainment Directors

The outcome of the analysis indicates the most performing director as 'not given' which raises the need for improved data collection of director information. Other top directors are Rajiv, Raul & Jan, Alastair, Suhas, Marcus, Jay, Cathy, Youssef and Martin. Standardized data entry practices and periodic audits are needed to minimize inconsistent information.



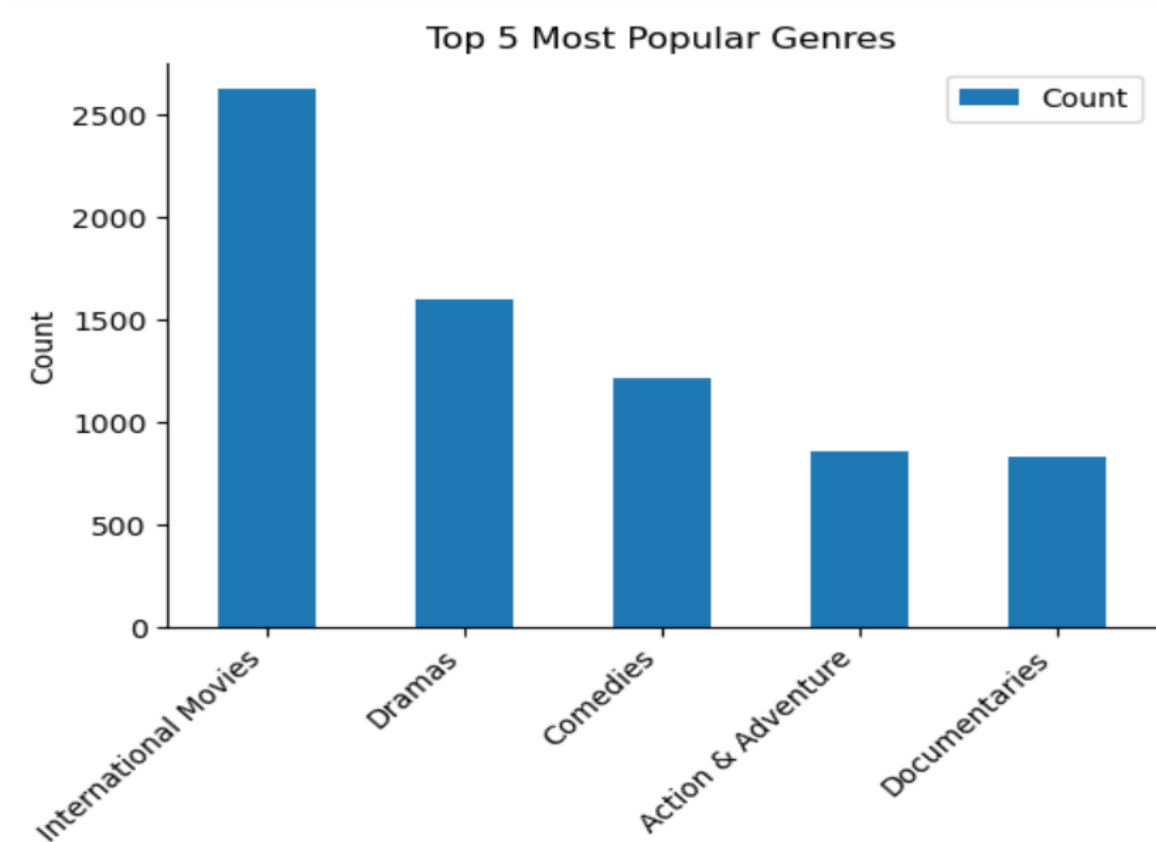
C. General Popular Entertainment Genres

Findings indicate strong preference for dramas/international movies. Other popular entertainment genres are documentaries, standup comedy, comedies/dramas/international movies and dramas/independent movies/international movies.

Dramas, International Movies	362
Documentaries	358
Stand-Up Comedy	334
Comedies, Dramas, International Movies	274
Dramas, Independent Movies, International Movies	252

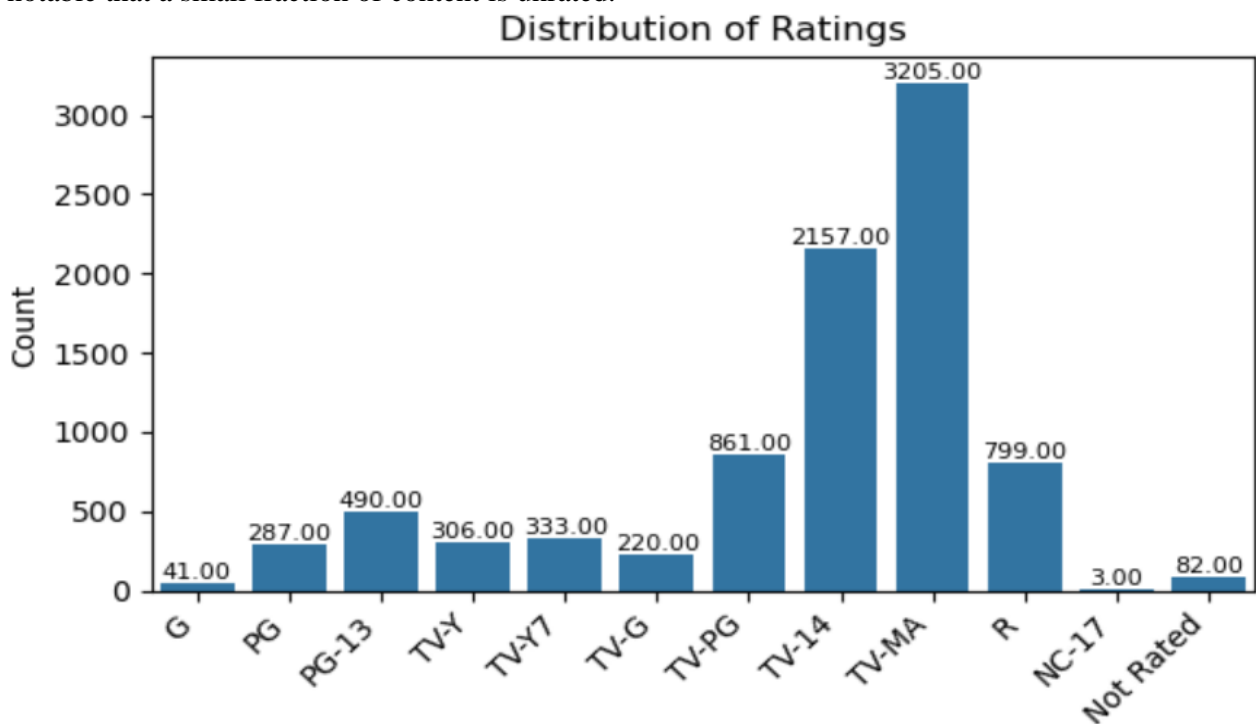
D. Top 5 Most Popular Genres

The most popular genre is international movies with 2624 count. Following closely are dramas and comedies with 1599 and 1210 respectively. Others are action & adventure (859 count) and documentaries (828 count). Findings reveal that foreign movies are the most common form of entertainment.



D. Entertainment Ratings

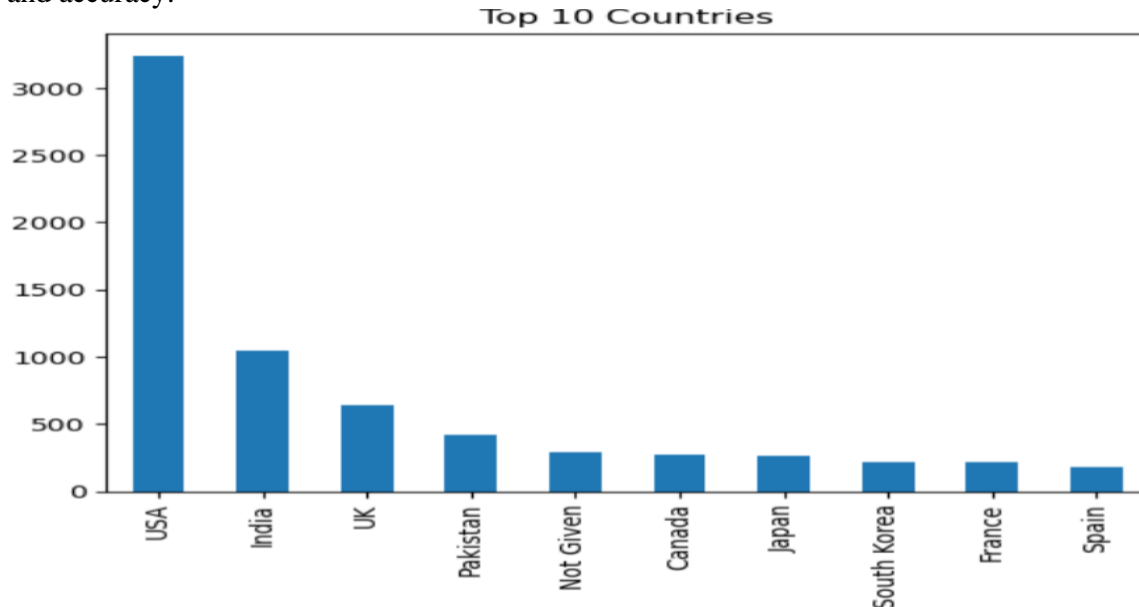
The majority of content is rated TV-MA with 3205 contents, followed by TV-14, TV-PG and R with 2157, 861 and 799 contents respectively. This indicates a strong focus on mature audiences. It is notable that a small fraction of content is unrated.



2. Insight on Countries

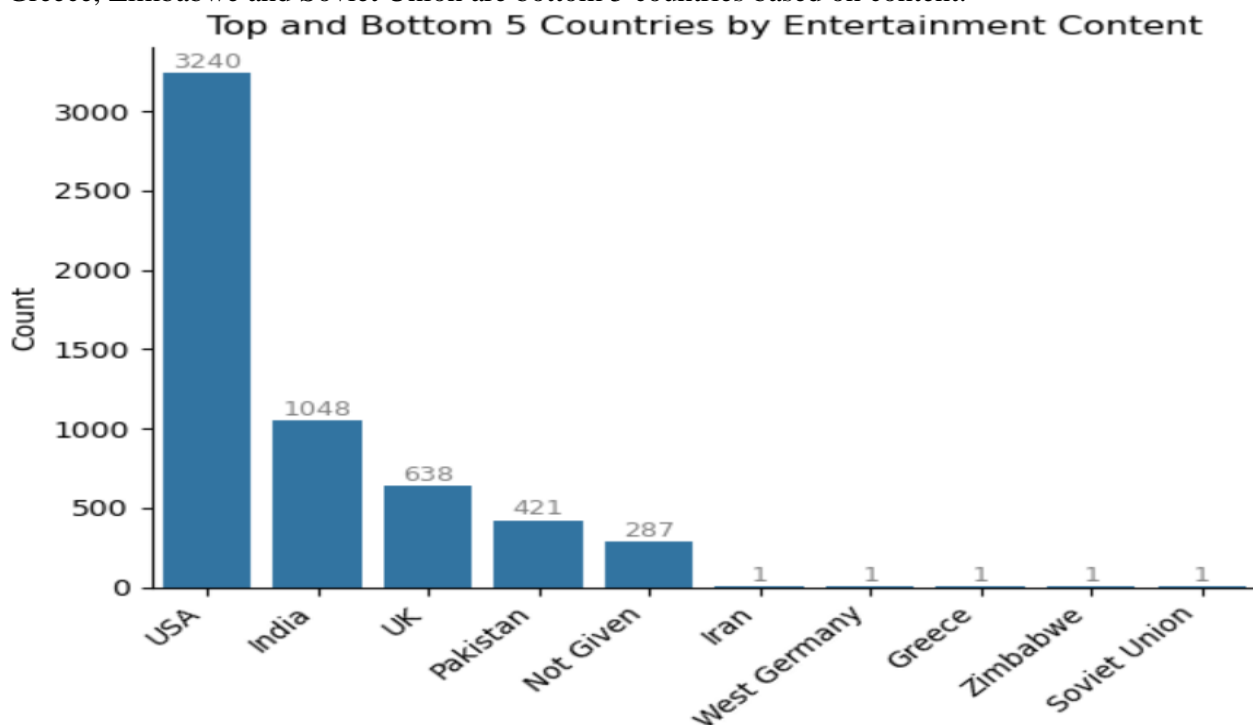
A. Top 10 Countries

The USA dominates the entertainment industry, with a significantly higher count than any other country. India and the UK emerge as secondary hubs for entertainment content creation with lower counts than the USA. Countries like Pakistan, Canada, Japan, South Korea, France, and Spain have relatively similar and lower counts. The "Not Given" category raises concerns about data completeness and accuracy.



B. Top and Bottom 5 Countries

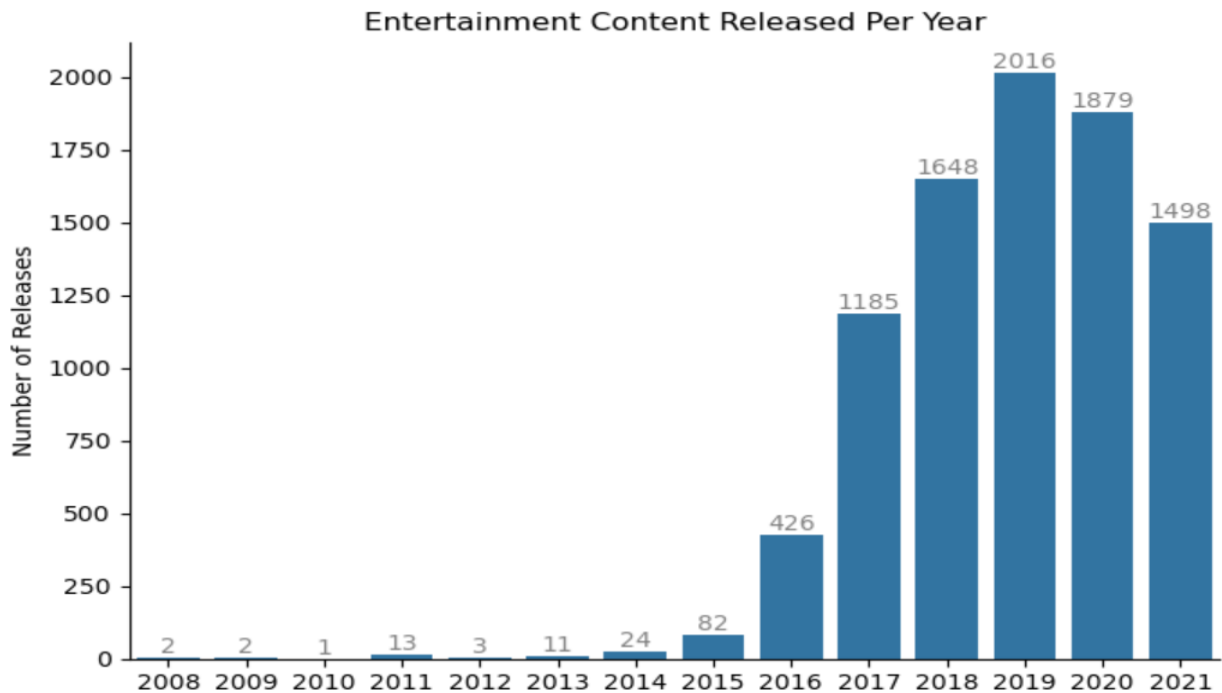
USA, India, UK, Pakistan and 'Not Given' are identified as top 5 countries with highest entertainment content with 3240, 1048, 638, 421 and 287 contents respectively. In addition, Iran, West Germany, Greece, Zimbabwe and Soviet Union are bottom 5 countries based on content.



3. Periodic Trends

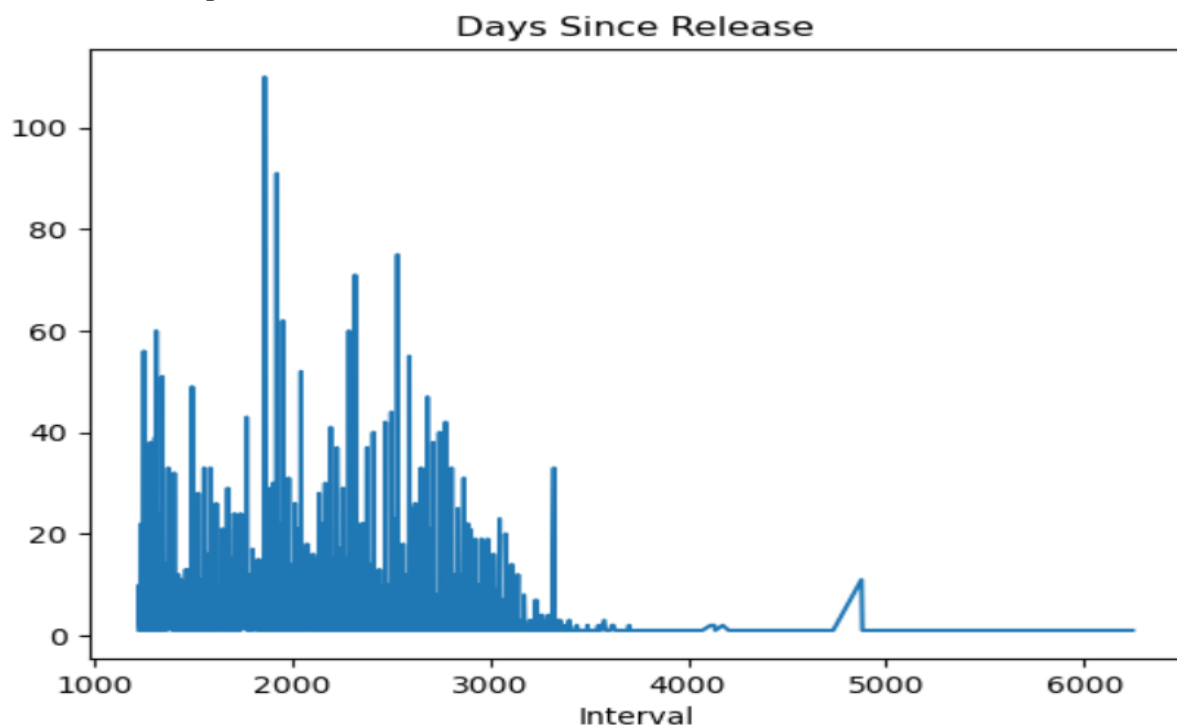
A. Entertainment Content Analysis per Year

Findings indicate that entertainment releases experienced a remarkable surge from 2016 onwards, peaking in 2019 with two thousand and sixteen (2016) releases. However, releases declined slightly in 2020 and 2021 with 1879 and 1498 releases respectively.



B. Days Since Entertainment Release

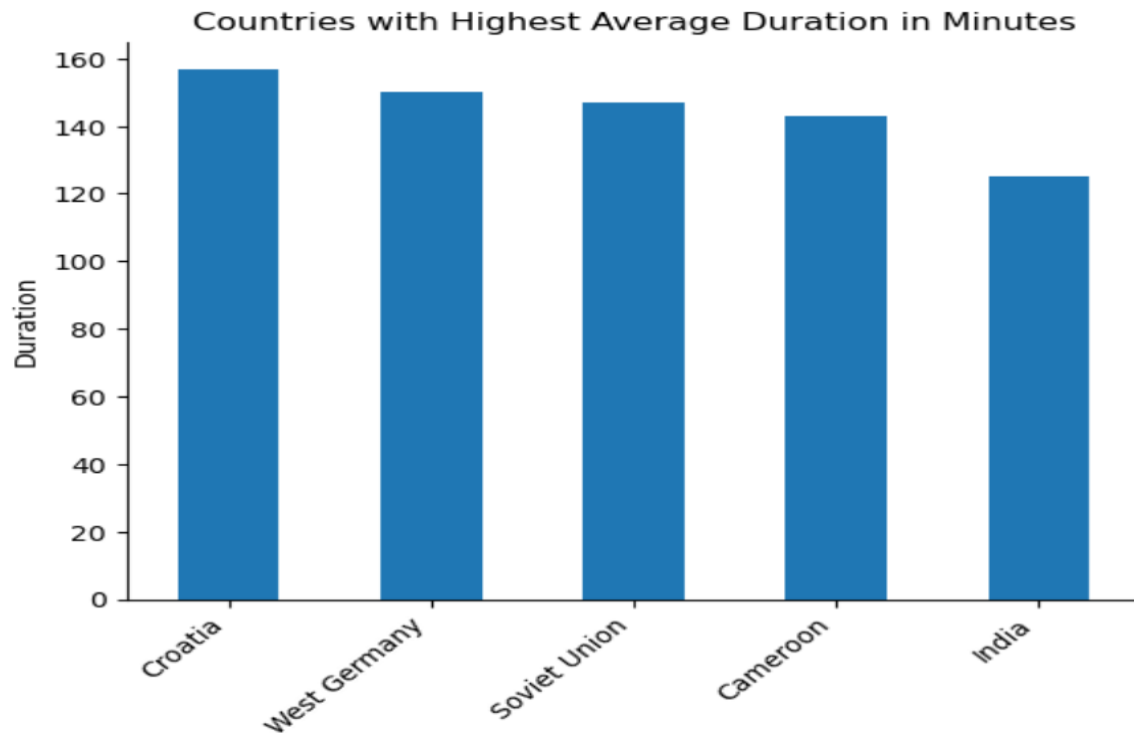
The distribution of days since release is heavily concentrated in the earlier intervals, revealing a rapid decline after the peak. This indicates that the audience is more focused on newer entertainment content.



4. Insight on Duration

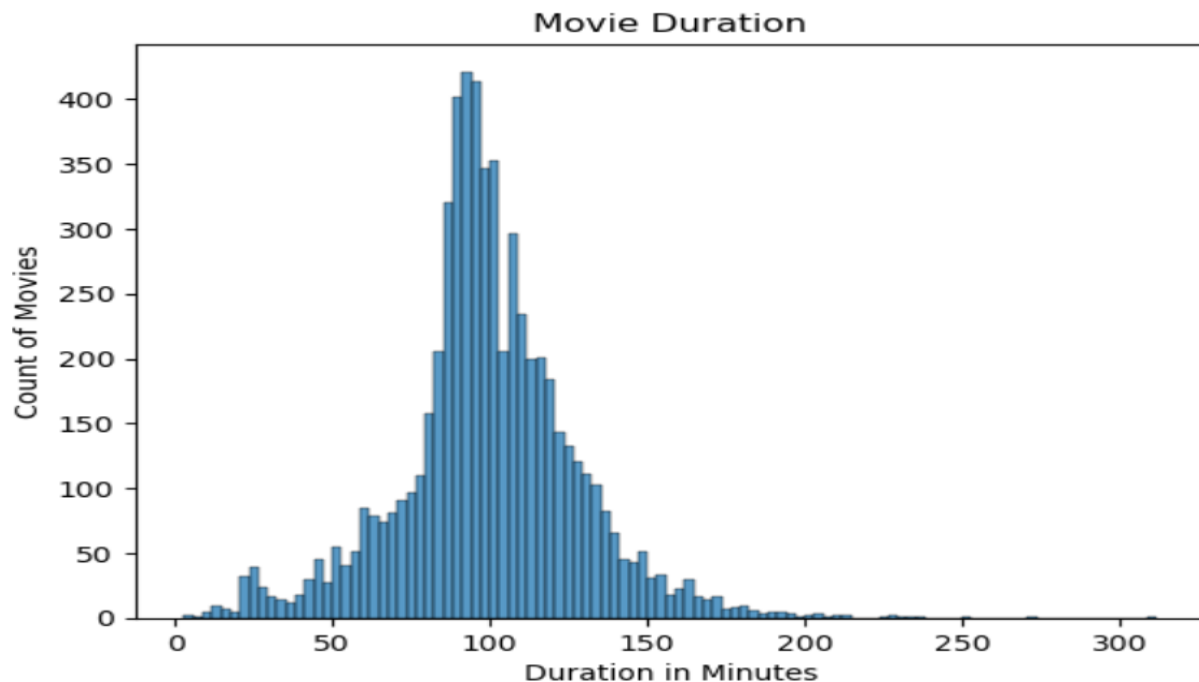
A. Countries with Highest Average Duration

Croatia leads in average content duration, followed by West Germany, Soviet Union, Cameroon and India, indicating possible preference for longer content in these regions.



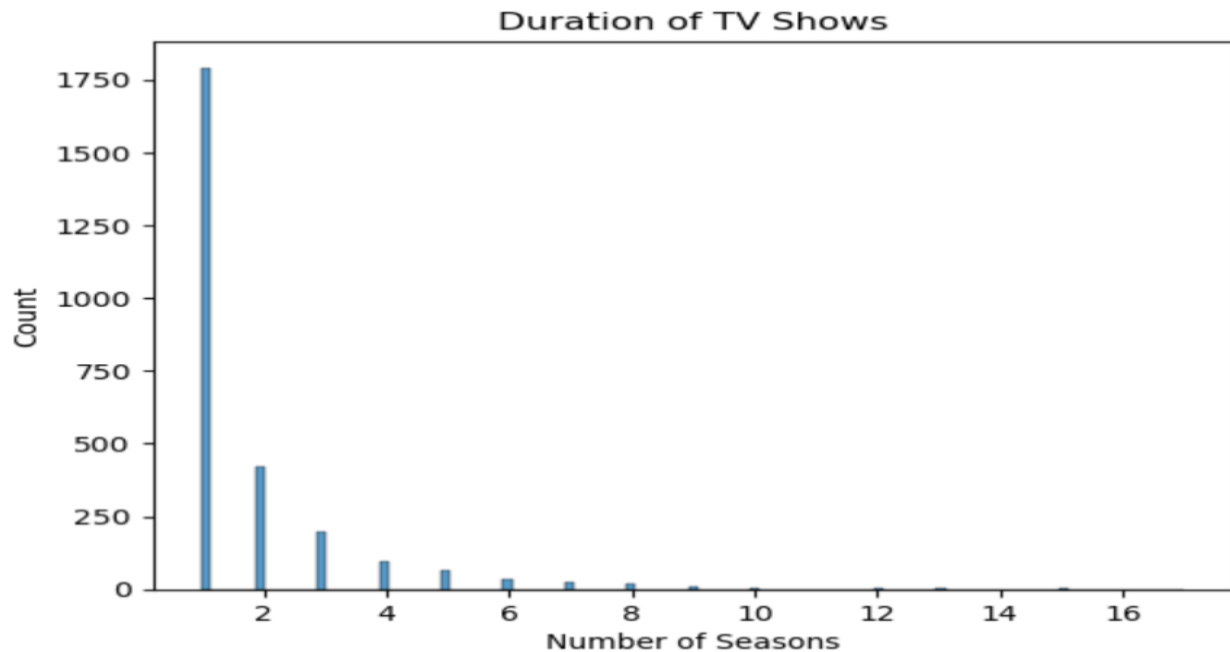
B. Distribution of Movie Duration in Minutes

Findings indicate that the highest concentration of movies falls within the 90-100 minute range. This infers that the target audience prefers movies covering that duration. A decline is observed at longer durations. Hence, a smaller number of movies fall into both shorter and longer durations.



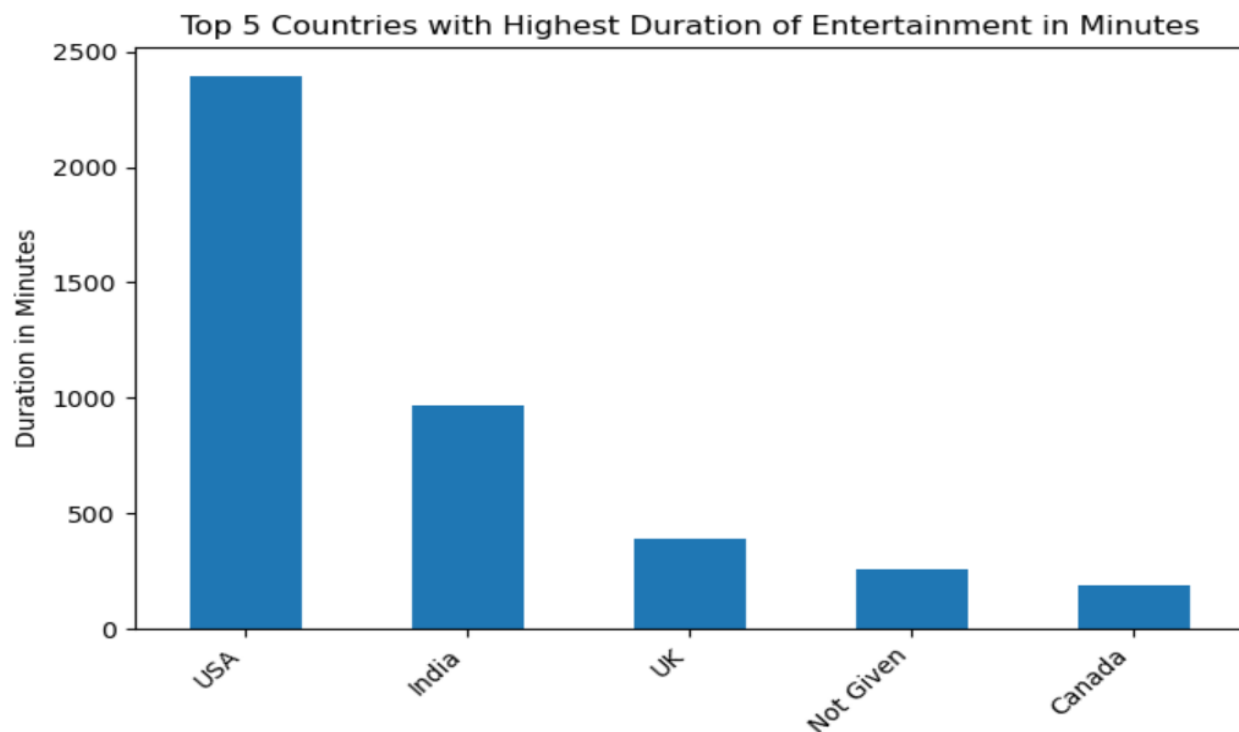
C. Duration of TV Shows in Seasons

Findings indicate that the majority of TV shows have between 1 and 3 seasons, which indicates strong preference for shorter series over longer TV shows. TV shows with longer seasons were fewer.



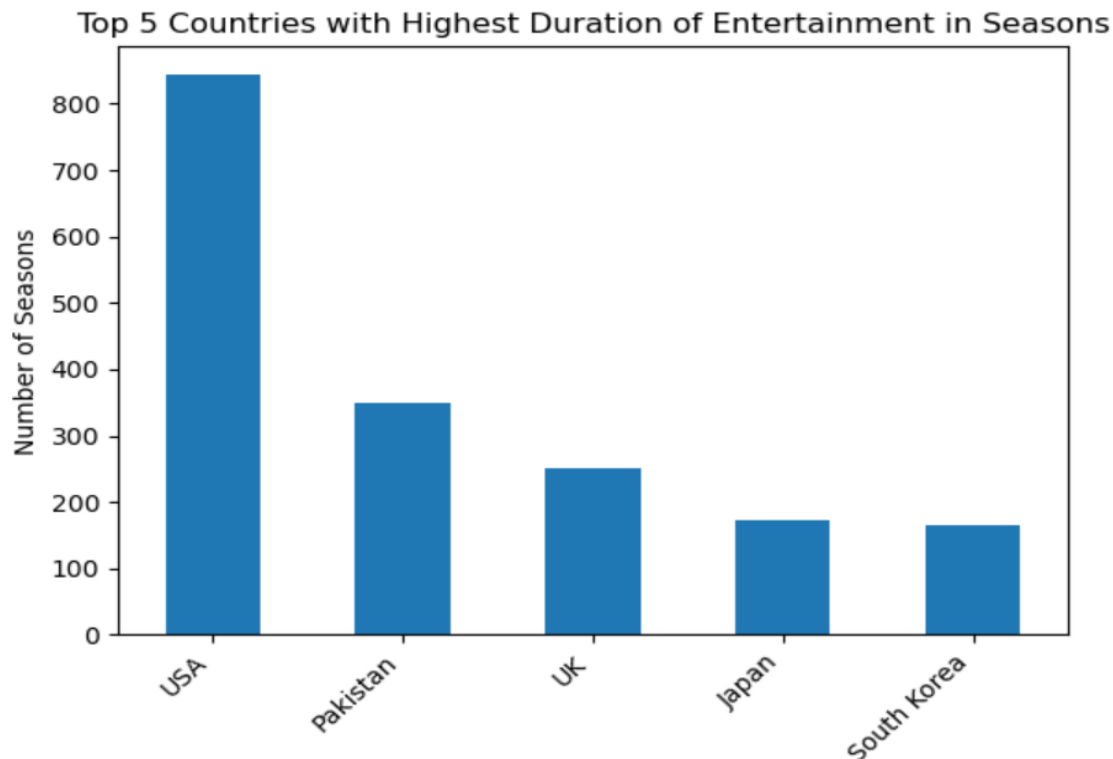
D. Top 5 Countries with Highest Duration of Entertainment in Minutes

Findings indicate that the USA has a significantly higher total duration of entertainment content compared to all other countries. India and UK are notable contributors of considerable entertainment content. The 'not given' category indicates inconsistent or missing data. Canada has a smaller total duration in comparison with other top countries.



E. Top 5 Countries with Highest Duration of Entertainment in Seasons

Findings indicate that the USA has a larger number of entertainment seasons in comparison with other countries in the top 5. Pakistan and UK are major contributors while Japan and South Korea have fewer seasons.



Conclusion

After a thorough analysis of the provided dataset, findings indicate strong preference for movies over TV shows, with a significant portion of content targeted towards mature audiences. While the USA is identified as a prominent contributor of entertainment content across categories, countries like India and UK also contribute remarkably. A surge is observed in content releases from 2016 onwards, peaking in 2019, followed by a marginal decline which suggests a dynamic entertainment ecosphere. The analysis also highlights evident preference for newer content, with viewership concentrated in the initial release period. Furthermore, the majority of both movies and TV shows adhere to typical duration lengths, clustering around 90-100 minutes for movies and 1-3 seasons for TV shows indicating preference for shorter seasons. However, the consistent presence of 'Not Given' entries across multiple data points indicates significant data sparsity that requires crucial attention.

Recommendations

Based on findings of this report, I recommend optimization of content strategies to enhance viewership of TV shows. Emphasis on genre-specific preferences will improve content performance in the progressive market. Leveraging popular entertainment genres and content durations will enhance user experience and increase engagement by the target audience. Findings indicate opportunities for collaboration and expansion into prominent foreign markets like India, UK, Pakistan, Canada, Japan, South Korea, France, and Spain for stronger competitive advantage. To improve data quality, there is need for adopting standardized data collection practices such as clear data entry guidelines, automated validation checks, and consistent auditing to minimize inconsistencies & missing data.