

Analytical report on Netflix data

Report structure:

1. Introduction
2. Data description
3. Stages of analysis
4. Visualizations and results and recommendations

1. Introduction

The main purpose of the study is to conduct a detailed analysis of Netflix data to identify the main trends and key features that determine the platform's strategy and the behavior of its subscribers.

2. Data description

- File: netflix_titles.csv
- Number of rows: 8807
- Number of columns: 12
- List of columns:
 - Show_id - the number in the list.
 - Type - type (Movie/TV Show)
 - Title - the title
 - Director - producer
 - Cast - the cast
 - Country - country of publication
 - Date_added - date added to the platform
 - Release_year - year of release
 - Rating - age rating
 - Duration - duration (minutes/seasons)
 - Listed_in - genres
 - Description - description

3. Stages of analysis

1. Data Preprocessing

The goal: To prepare data for further analysis and visualization.

Completed actions:

- The date_added value has been converted to the date and time format to work correctly with time data.
- Omissions in critical columns have been removed to improve the quality of the analysis.
- The release_decade column has been added, containing the decade of content release.
- A release month column has been added to analyze the seasonality of content additions.
- The duration is divided into:
 - duration_value – numeric value (number of minutes or seasons);
 - duration_type – duration type (minutes or seasons).

2. Basic analysis

The goal: To get a general understanding of the structure and distribution of content on the platform.

Completed actions:

- The total number of films and TV series in the dataset has been determined.
- The year with the largest number of films released has been determined.
- The year in which the most content was added to the platform was found.
- The Top 10 countries in terms of content have been compiled.
- The content is distributed according to ratings (for example, G, PG, R, etc.).

3. Advanced analysis

The goal: To identify key patterns and leaders in various content parameters.

Completed actions:

- The director with the largest number of films and TV series on the platform has been identified.
- The TOP 10 directors with the most works have been compiled.
- The actor who is most often found in the cast has been identified.
- The TOP 10 popular actors who appear most frequently in the content have been compiled.
- The top 5 most popular genres in terms of the number of films and TV series have been revealed.
- The average duration of films (in minutes) is calculated.
- The average number of seasons for TV series has been determined.

4. Data visualization

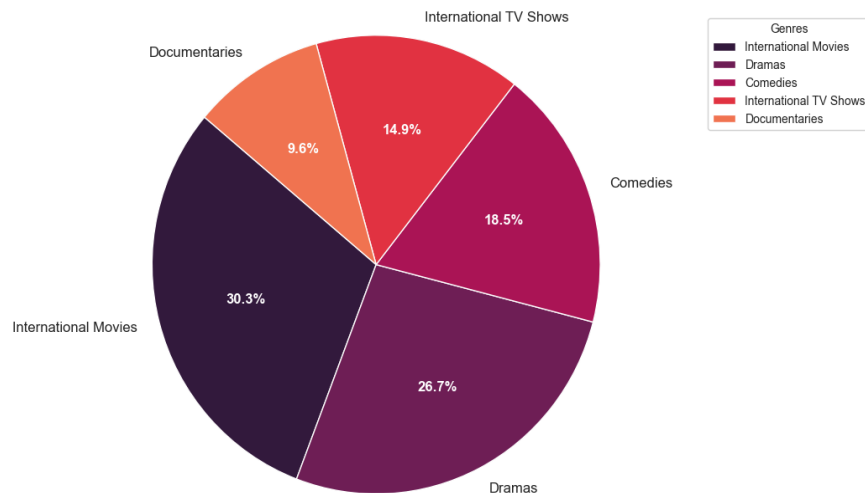
The goal: To visually present the key indicators and distribution of content on the platform.

Completed actions:

- A histogram of the distribution by ratings (for example, G, PG, R) to understand the age structure of the content.
- A chart of the number of movies and TV series for visual comparison of content categories.
- Pie chart is a top-5 genre chart that allows you to visually evaluate the popular trends of films and TV series.
- Barplot top 10 actors who are most often found in the cast.
- Barplot top 10 directors with the most works on the platform.
- Barplot top 10 countries in terms of released content, showing the geographical distribution of production.
- Histplot of the distribution of films by duration, to analyze the duration and identify trends.

4. Visualizations and results

1. Top 5 most popular genres on Netflix (by amount of content)



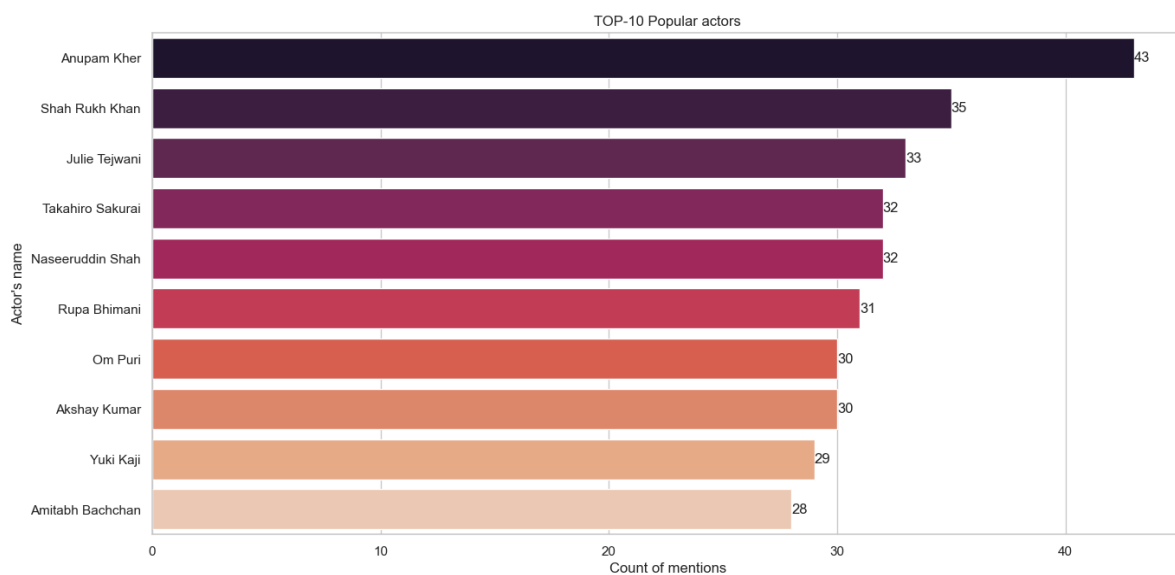
Basic observations:

- The largest share is accounted for by International Movies (30.3%), which shows the strong international direction of the platform.
- Dramas (26.7%) and Comedies (18.5%) also occupy a significant part, which indicates the high popularity of these genres among viewers.
- International TV Shows (14.9%) and Documentaries (9.6%) occupy smaller segments, indicating a narrow but stable audience.

Recommendations / insights for business:

- The main focus of the content can be on international films and dramas, as they form most of the statistics.
- Comedies remain a significant category for retaining a wide audience.
- Documentaries and international TV series are a promising niche for promotion, and you can separate them into separate collections or promotions.

2. Top 10 Most Popular Actors (Barplot)



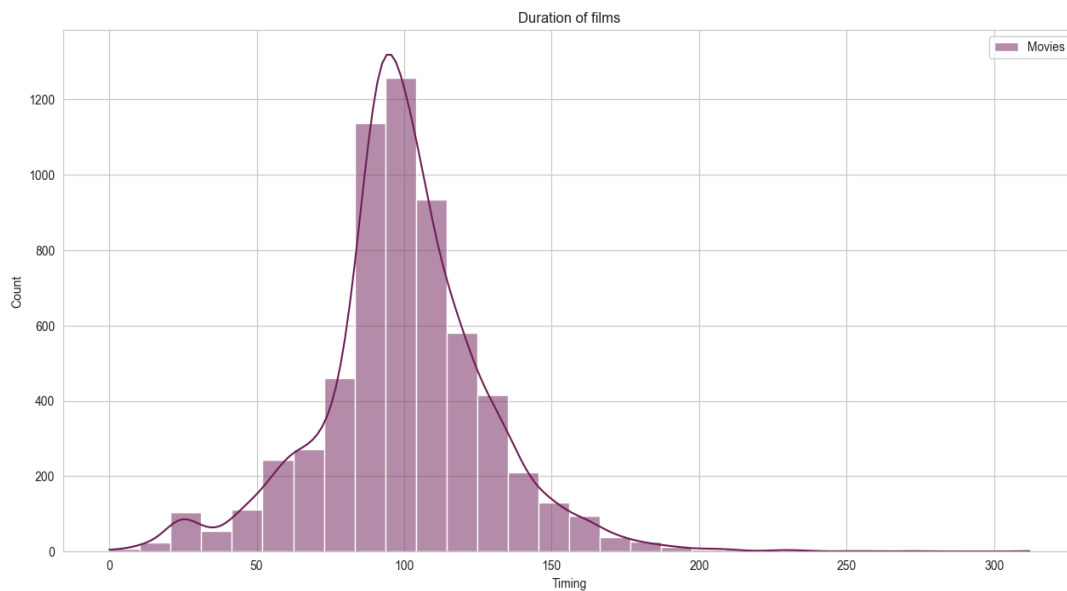
Basic observations:

- The top 10 actors are mostly Indian actors - Netflix is actively promoting content for the Asian market.
- The clear leader: Annupam Kher leads the ranking by a significant margin with 43 mentions. This is more than 20% higher than the next actor.
- High recognition concentration: The top three actors (Annupam Kher, Shah Rukh Khan, Julie Tejwani) form a separate leadership group with scores above 35 mentions.
- International representation: The top 10 includes not only Indian actors (which is expected, judging by the names), but also Japanese (Takahiro Sakurai, Yuki Kaji), which indicates the diversity of the audience or content in the sample under study.

Recommendations / insights for business:

- Focus on regions: High percentage of Indian actors = Netflix is targeting Asia. This means that in order to grow in other markets (for example, Europe, LatAm), it is worth increasing local content.
- Focus on leaders: When planning marketing campaigns, content promotion, or partnerships, it is recommended that you first consider collaborating with Anupam Kher, Shah Rukh Khan, and Julie Tejwani, as their names are guaranteed to attract the most attention from the target audience.
- Taking into account the international audience: The presence of Japanese voice actors (Takahiro Sakurai, Yuki Kaji) in the top 10 is an important insight. This is a direct indication of the presence of an active and interested audience familiar with anime or Japanese media content. It is recommended to study this target group in more detail and take into account its preferences.

3. The distribution of the duration of films



Basic observations:

- The distribution of the duration of films is clearly not normal, with a pronounced right-hand bias.
- The vast majority of films (fashion) last about 100 minutes.
- There is a "long tail" — a small number of films with a duration of more than 150-200 minutes.

Recommendations / insights for business:

- Content Format: When creating or purchasing content, the main focus should be on standard length films (~100 min.), as they are the most popular and expected format among the audience.
- Platform strategies: Long films (>150 min.) are a niche product. They should be promoted for an audience that is purposefully looking for "movie epics" or author's films, and not for mass consumption.
- User experience: For streaming services, this is a direct indication of the need for a pause or "continue watching" feature, especially for long tail movies.