Insights and recommendations

NOTE: Please refer to the pdf file link attached for graphs and code. In this file, I have covered only the insights and recommendations.

Problem statement: The objective of this project is to perform exploratory data analysis and visualization on the Netflix dataset.

The goal is to gain insights and understand patterns in the dataset, this includes understanding the distribution of movies and TV shows by genre, analyzing trends over time, exploring the relationship between variables like country and content type, identifying popular actors or directors, and uncovering any other interesting patterns or trends in the data.

This project aims to provide valuable insights into the Netflix dataset, enabling better understanding of the content available on the platform and providing data-driven decision-making for content creators, analysts, and other stakeholders.

Basic Analysis:

Data columns and their data types:

```
Data columns (total 12 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   show_id       8807 non-null   object
 1   type          8807 non-null   object
 2   title         8807 non-null   object
 3   director      6173 non-null   object
 4   cast          7982 non-null   object
 5   country       7976 non-null   object
 6   date_added    8797 non-null   object
 7   release_year  8807 non-null   int64
 8   rating        8803 non-null   object
 9   duration      8804 non-null   object
 10  listed_in     8807 non-null   object
 11  description   8807 non-null   object'
```

Dataset Size: The dataset consists of 8807 rows and 12 columns

Data Types: Most of the columns are of the object data type, which typically represents strings or mixed data types. The release_year column is of the int64 data type, indicating it contains integer values.

Missing Values: Some columns have missing values. The director column has 6173 non-null values, indicating that there are missing values for approximately 2634 entries. Similarly, the cast, country, date_added, rating, duration columns also have missing values to a lesser extent

Categorical data analysis:

Rating: The maximum shows/movies come under the category: 'TV-MA': Mature audience only

Type: Percentage of movies is 69.7% which is far more than percentage of TV shows(30.3%) on Netflix platform

Country: US contributes to maximum shows in Netflix followed by India

Genre: Top Genres for Movies/TV shows include Dramas, International Movies and Documentaries

Based on this analysis, recommendations:

1. **Rating diversification:** While 'TV-MA' includes maximum content, it is important to ensure a diverse range of content to cater to different audience segments. Consider expanding the content library to include more content suitable for other age groups and genres to attract a wider audience.

2. **Focus on Movies:** Since the number of movies is significantly higher than the number of TV shows, it indicates a higher demand for movies. Invest in producing a diverse range of movies to maintain the interest of the existing user base and attract new subscribers.

3. **Country reach:** Since the US and India contribute the most shows to Netflix, it is important to continue catering to the preferences of these markets.

4. **Popular Genres:** Since dramas, international movies, and documentaries are identified as top genres, focus on acquiring or producing high-quality content in these genres. Promote and highlight these genres to attract viewers who have shown an interest in such content.

Movie/TV-shows production trend: production increased from 2016 till 2019, with maximum production in 2019. There is a decline in production after 2019

TV-shows production per week - maximum at 27th week i.e., *1st week of July*

Movie production per week - Movies are maximum added in the 1st week of January, i.e., start of the year.

TV-shows production per month - maximum tv shows are added at the 12th month which means at year end, followed by 6th-7th month(mid-year).

Movie production per month - Maximum movies are added at around the 7th month(mid-year) of the year. Many movies are added in 1st, 4th,10th and 12th months also

Based on this analysis, recommendations:

1. **Focus on TV Shows:** As there is a decline in overall production after 2019, it might be beneficial to focus on TV show production, which has shown a consistent increase over the years. TV shows are in high demand and have the potential to attract a larger audience.

2. **Strategic Release of TV Shows:** Since the maximum TV show production occurs around the 27th week, consider strategically releasing new TV shows during this period to fulfill the high demand and retain viewership.

3. **Focus on Movie Releases at the Beginning of the Year:** Movies are predominantly added in the 1st week, indicating a trend of starting the year with new movie releases. Plan and promote movie releases during this time to attract viewers who are eager for fresh content.

4. **Focus on Year-end TV Show:** With maximum TV shows being added in the 12th month, consider scheduling TV show releases or seasonal content during this time to capture the viewer's attention during the year-end festivities like Christmas, Thanksgiving and New Year.

5. **Seasonal Movie Releases:** Take advantage of the peak movie production months (1st, 4th, 7th, 10th, and 12th) to release movies aligned with seasonal trends or holidays. This can help attract a larger audience and generate more publicity for movies.

Top Actors: Anupam Kher has appeared the most followed by Shah Rukh Khan.
Top Directors: Rajiv Chilaka is the top director followed by Jan Suter
Top Genres: we can see that International Movies, Dramas, Action, Adventure, Comedy are the top Genres. International movies are most popular in India. whereas Drama movies are most popular in US
Year added: Most movies are added to Netflix in the same year as they were released in. This shows that it's best to add the movie to Netflix in the same year as release.

Based on above analysis, recommendations:

1. **Collaboration with Top Actors:** Since Anupam Kher and Shah Rukh Khan are the most prominent actors, it would be beneficial to collaborate with them for future projects. Their popularity and fan base can attract a significant audience to the platform.

2. **Collaborate with Top Directors:** Rajiv Chilaka and Jan Suter have emerged as the top directors. Partnering with these directors can lead to the creation of high-quality content and attract more viewers.

3. **Emphasize International Movies:** International movies have shown popularity, particularly in India. Consider acquiring and promoting a diverse range of international movies to offer to the preferences of a global audience.

4. **Focus on Drama Genre:** Dramas have emerged as a popular genre, especially in the United States. Invest in acquiring and producing powerful drama content to support the preferences of the US audience.

5. **Timely Release of Movies:** The analysis shows that most movies are added to Netflix in the same year as their release. To maximize viewership and capitalize on the buzz surrounding newly released movies, ensure timely acquisition and addition of movies to the platform.

Most popular TV shows Genre and shows:
1. **Crime:** Dexter, Arrow, Forensic Files
2. **International:** Still Game, Club Friday, Call the midwife
3. **Action & Adventure:** Arrow, The Flash, Star Trek
4. **Comedies:** Saved by the bell, The Office(U.S.), That 70s Show

Based on above analysis, recommendations:

1. **Promote Popular Shows:** Highlight the popular TV shows, such as Dexter, Arrow, Forensic Files, The Flash, Star Trek,The Office (U.S.), That '70s Show etc. Promote these shows through various marketing channels to increase viewership and engagement.

2. **Explore Similar Content:** Identify other TV shows within the crime, international, action & adventure, and comedy genres that share similarities with the popular shows mentioned. By offering similar content, we can appeal to viewers who enjoy these genres.

3. **User Recommendations:** Encourage users to provide feedback and recommendations on their favorite TV shows or genres. This can help gather insights into the preferences of the audience and guide future content acquisition or production decisions.

Correlation between year and shows production: year and production of shows are moderately positively correlated.

Based on above analysis, recommendations:

1. **Focus on Recent Years:** Since there is a noticeable increase in production in recent years, it is important to prioritize acquiring or producing shows from these years. Consider obtaining exclusive rights to popular shows released in the most recent years to offer fresh and in-demand content to viewers.

2. **In-house Production:** As the year and production of shows are positively correlated, it indicates a growing trend in the production of shows over time. It is recommended to continue investing in content production and expanding the range of shows offered on the platform. This can help cater to the increasing demand for shows and attract a larger audience.

3. **Monitor Competitors:** Keep an eye on the production strategies of competitors in the industry. Analyze their content offerings and the trends they are following. This can help you stay competitive and identify potential areas for improvement.