Battle of the Screens: Analyzing Top OTT Platforms

Group 5: Sai Bheeshma Ramaraju Pagilla, Chaitanya Krishna Burri, Sai Mona Duvvapu, Monika Madugula, Piyush Arvindbhai Arvindbhai Kakde

Introduction and background:

To conduct a comparative analysis of film and TV content catalogs on Netflix, Amazon Prime Video, and Disney+, spanning different time periods with the additional information of IMDB ratings, intending to uncover patterns and insights regarding the respective user behavior and preferences of these streaming platforms.

Why this topic:

Analyzing these datasets will let us uncover the competitive strategies used, content preferences, targeted audience at a specific time and any recurrent patterns. Additionally, our motivation to uncover insights through data visualization is rooted in these market trends:

- According to a Nielsen report, over time, these streaming platforms have leveraged data to cater to specific audience segments.
- In a survey conducted by PWC, nearly one-third (31%) of survey respondents said easy, personalized recommendations would be a reason to continue the streaming service.

By harnessing the insights from market research, we aim to extract valuable information from our dataset using Tableau visualizations.

Dataset description:

We have 3 datasets from kaggle describing Netflix (8807 rows), Prime (9668 rows) and Disney+ (1450 rows) with the columns show_id, type, title, director, cast, country, date_added, release_year, rating, duration, listed_in (genre), description. These datasets (.csv files) contain information in an orderly fashion with minimal missing values. The variables that interest us the most among the mentioned are OTT Platform, type (movie/series - categorical), release_year (Date part - discrete), country (categorical) and the genre (categorical). The relationship of these variables among various platforms will help discover the hidden patterns or recurrent patterns

over the years.

The data spans from the 1920s to 2021, and we plan to create graphs at the country level, as the dataset provides information about the country in which shows are released. We have data for these three OTT platforms across 88 countries. In addition to the primary data columns, we have two calculated fields: 'Number of Records,' which shows the count of shows, and 'Rank by Total Shows,' indicating the ranking of the show count.

We've also incorporated a parameter called 'Platform Selector,' which is a list of values representing the three OTT platforms. This parameter is used in various interactive dashboards

Dataset sourced through web scraping:

We built a Python program to source data from IMDb using the PyMovieDb library. The program iterated through a list of movie titles, fetched IMDb data for each title, and extracted relevant rating information from the retrieved JSON. If a movie had a rating, its name, rating, and rating count were stored in a 'movie_data' list and written to a 'movie_data.csv' file. If a movie had no rating information, it was added to a separate list and saved in a 'movies_without_ratings.csv' file. This additional data enriches our analysis of the Netflix, Prime, and Disney+ datasets, providing insights into content quality and audience reception, which can further enhance our understanding of hidden patterns and recurring trends in the OTT platforms' content over the years.

We utilized a total of two datasets: one file containing shared values for all three platforms, and a separate CSV file containing IMDB ratings for the shows and movies. The reason for using a combined file is explained in the limitations section.

Limitations of the dataset:

Some of the limitations include a focus on a limited set of platforms in understanding OTT platforms, potential missing data, reliance solely on IMDb ratings as a quality measure, and the absence of user-specific data. Specifically, the 'absence of user-specific data' means that we do not have access to individual user profiles, viewing history, or demographic information. The availability of such data would have allowed for a more detailed understanding of how different

viewers interact with the platforms, their content preferences, and viewing habits, ultimately enriching our analysis.

Another limitation related to data cleaning is that the genres are entered differently in each dataset. For example, with TV shows, 'Comedy' is listed as a genre in Disney+ and Prime Video, but in Netflix, it is categorized as 'TV comedy.' Due to these variations and the fact that shows are unique per platform (very few were present in multiple platforms), we couldn't join the datasets. Instead, we created a single dataset with consistent values across all fields for the platforms.

Data story: ***(Please note that we have included only few visualizations in this report, the complete set can be found in the .twbx file)***

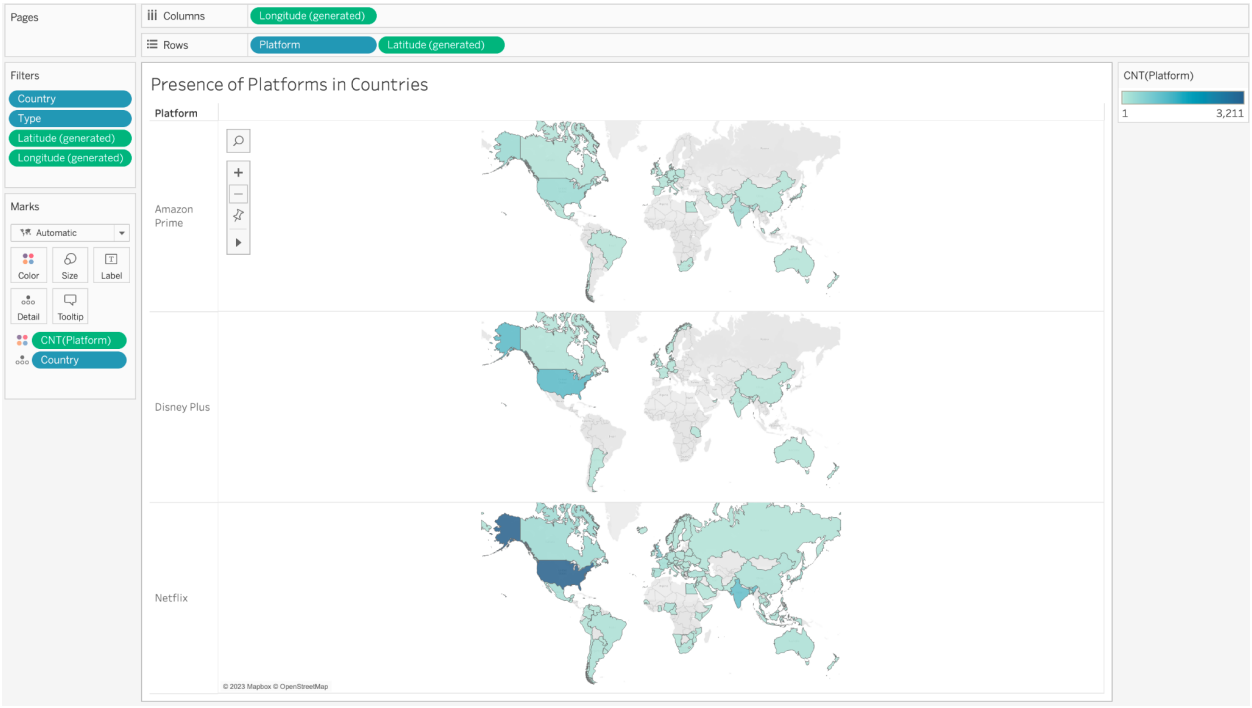
- To begin with, we analyzed the distribution of content types for the platform selectors. Individual preferences can vary; some may prefer a platform with more movies, while others may prefer one with more TV shows. This will serve as a valuable parameter for making informed choices when subscribing
- Next, we examined the release of new releases over time for each platform. This information will help users understand when a particular platform gained momentum.
- As a continuation, we presented the top five genres for each platform and created an interactive dashboard to consolidate all the above information for each platform.
- We also captured release statistics by year for United States
- We identified the top 10 countries that have the highest number of shows
- We created a heat map to visualise the rating type (13+, Parental Guidance etc) by platform.
- Visualised the number of shows by platform and average rating by platform.
- Initially we captured the average duration by OTT platform, but we recognized that this metric might be unfair. For example, if a particular OTT platform has more comedy shows, the average duration would be skewed. To address this issue, we opted to calculate the average rating by both duration and platform. We categorized shows with durations of 18-30 minutes as sitcoms, 40-60 minutes as other genres of TV shows, and

movies as longer than 60 minutes.

1. Top 3 Genres in Each Platform

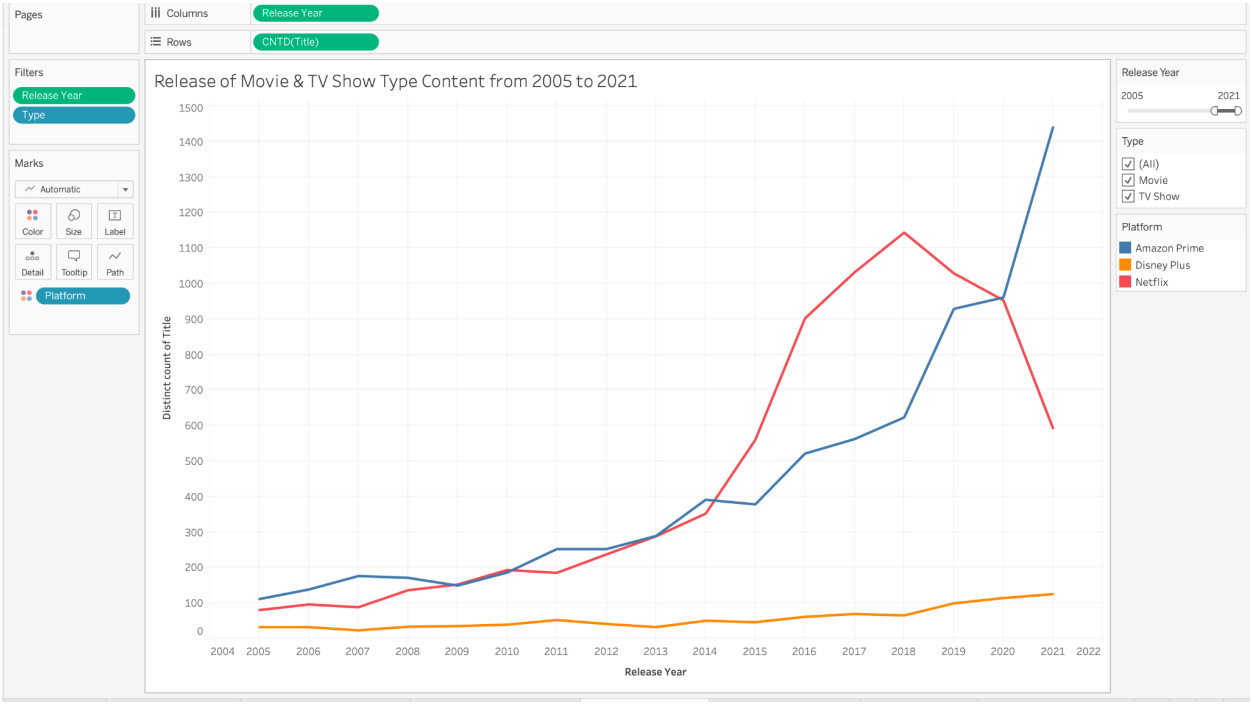
The bar chart depicts a Tableau visualization illustrating the top three genres: Comedy, Action, and Drama. Within these genres, it provides the average ratings for both Movies and TV shows on three prominent streaming platforms: Netflix, Amazon Prime, and Disney+. This visualization offers insights into viewers' perceptions of content quality within these specific genres on these platforms. Notably, for Amazon Prime, Action-TV shows receive the highest average rating, while Netflix excels in Drama-Movies, and Disney+ leads in Drama-TV shows. This diversification in strengths and niches among the platforms underscores their ability to target specific audience segments and provide content that aligns with those viewers' tastes and preferences.

2. Presence of Platforms in Countries



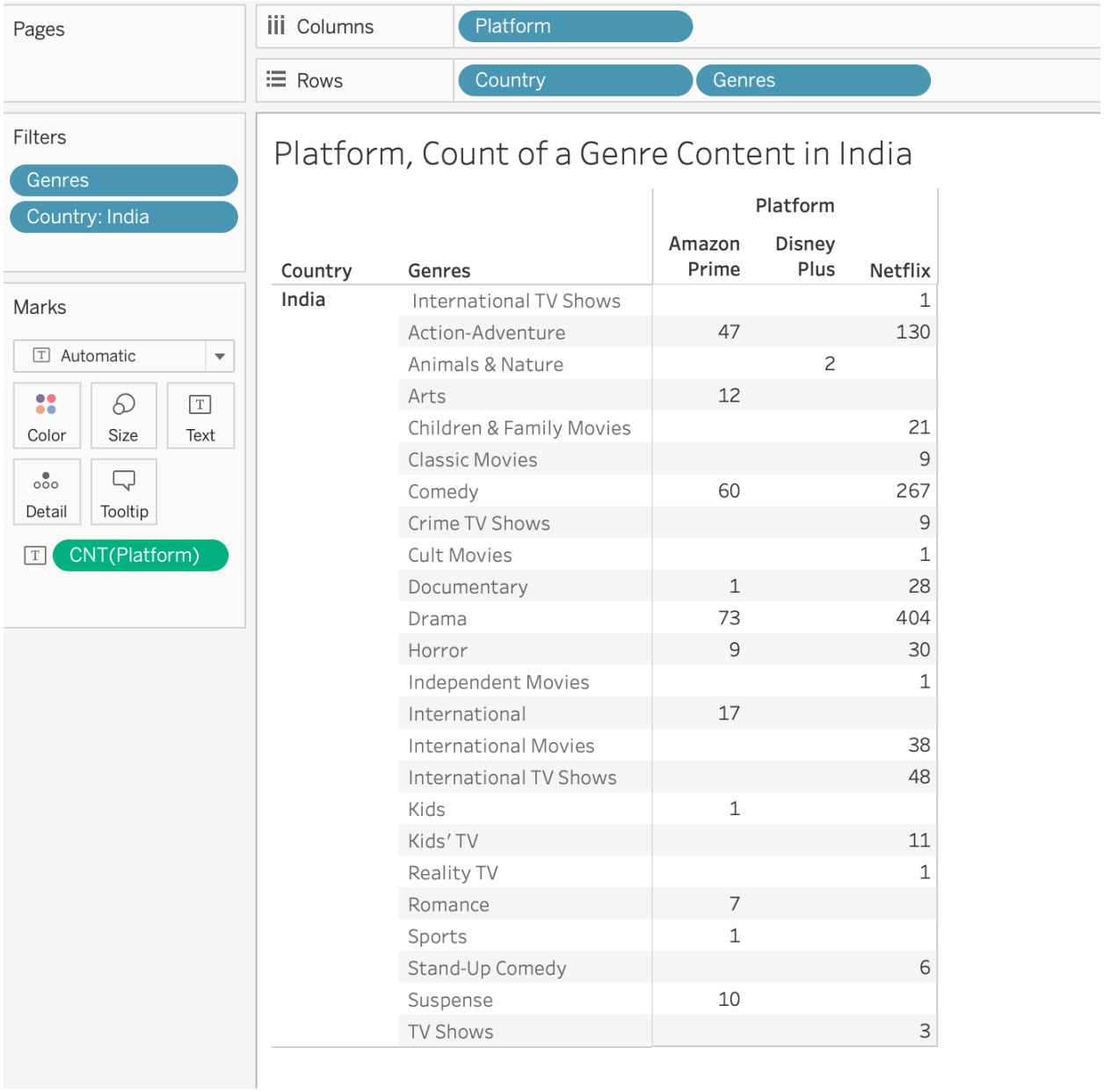
This map chart (present in Story tab) illustrates the presence of streaming platforms by country. Amazon Prime is available in North America, certain regions of South America, Southeast Asia, parts of Europe, and Australia. Disney+ has a presence in North America, Southeast Asia, and Australia. Netflix boasts the most extensive global presence, covering North and South America, the majority of Europe and Asia, select regions in Africa, and Australia. This data reveals that Netflix has the broadest international reach among the three platforms.

3. Release of Movie and TV show across platforms from 2005 to 2021



This line chart presents the count of content types from 2005 to 2021 across all 3 platforms. There has been steady growth for all three platforms over the years. However, it's noteworthy that Disney+ exhibited a significant uptick from 2020 to 2021, while Netflix experienced a downward trend during the same period. This may be attributed to the impact of the pandemic or Netflix's strategic shift towards diversifying content beyond movies and TV shows in an effort to broaden its offerings.

4. Platform, Count of a Genre Content in India



Platform, Count of a Genre Content in India

Country	Genres	Platform		
		Amazon Prime	Disney Plus	Netflix
India	International TV Shows			1
	Action-Adventure	47		130
	Animals & Nature		2	
	Arts	12		
	Children & Family Movies			21
	Classic Movies			9
	Comedy	60		267
	Crime TV Shows			9
	Cult Movies			1
	Documentary	1		28
	Drama	73		404
	Horror	9		30
	Independent Movies			1
	International	17		
	International Movies			38
	International TV Shows			48
	Kids	1		
	Kids' TV			11
	Reality TV			1
	Romance	7		
	Sports	1		
	Stand-Up Comedy			6
	Suspense	10		
	TV Shows			3

This table chart displays the genre count across all streaming platforms in India. It's evident that Netflix and Amazon Prime offer the highest number of comedy and drama content in India. It also holds true for Amazon Prime. However, it's important to note that we have limited data available for Disney+ in India.

The insight derived from this observation is that both Netflix and Amazon Prime have recognized and responded to the preferences of the Indian audience by offering a substantial amount of comedy and drama content.

Summary and conclusion:

In conclusion, our Tableau data visualizations have provided valuable insights into the competitive strategies and dynamic nature of the OTT streaming industry. These visualizations revealed that each platform, including Netflix, Amazon Prime, and Disney+, carves a unique niche in content delivery, aligning with distinct audience preferences. Notably, Netflix's extensive global reach has solidified its position as an international leader. Furthermore, content growth trends indicated that Disney+ experienced a significant upturn in 2021, possibly due to a focused content strategy, while Netflix's minor dip might be attributed to strategic diversification or pandemic-related factors. Additionally, our analysis highlighted Netflix and Amazon Prime's recognition of the Indian audience's preferences, with substantial offerings of comedy and drama content. These insights equip stakeholders with valuable information for content strategy, audience targeting, and platform expansion in an ever-evolving OTT landscape.

References:

- "Data-Driven Personalization: The Future of Streaming Content Discovery." Nielsen, June 2023, www.nielsen.com/insights/2023/data-driven-personalization-2023-state-of-play-report/.
- "2023 Digital Media Trends Survey." *Deloitte Insights*, 14 Apr. 2023, www2.deloitte.com/us/en/insights/industry/technology/digital-media-trends-consumption-habits-survey.html.
- PricewaterhouseCoopers. "After a Boom Year in Video Streaming, What Comes Next?" *PwC*, 2021,www.pwc.com/us/en/services/consulting/library/consumer-intelligence-series/consumer-video-streaming-behavior.html.
- Mullin, Benjamin, and David Marcelis. "WSJ News Exclusive | Disney+, HBO Max and Other Streamers Get Waves of Subscribers from Must-See Content. Keeping Them Is Hard." *The Wall Street Journal*, Dow Jones & Company, 1 Feb. 2022, www.wsj.com/articles/streaming-data-netflix-hbo-disney-hulu-11643560207.

- Corbo, Jacomo, et al. “Video Entertainment in 2030.” *McKinsey & Company*, McKinsey & Company,29 June 2022, www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/video-entertainment-in-2030.
- Stoll, Julia. “U.S.: Spending on OTT Video Platforms 2017-2027.” *Statista*, 1 June 2023, www.statista.com/statistics/1337665/consumer-spending-ott-video-united-states/#:~:text=In%20the%20United%20States%2C%20consumer,50%20billion%20dollars%20by%202027.

Contributions:

Name of the person	**Contribution**
Bheeshma	Idea Brainstorming & Finalization, Visualizations, Partial Documentation
Chaitanya	Idea Brainstorming & Finalization, Visualizations, Partial Documentation
Mona	Idea Brainstorming & Finalization, Partial Documentation, IMDB Ratings Data Collection, Visualizations
Monika	Idea Brainstorming & Finalization, Partial Documentation, Research, Visualizations, Report
Piyush	Idea Brainstorming & Finalization, Web Scraping for IMDB Ratings (Idea & Implementation), Visualizations