

Online Streaming Service Analysis

By Rujuta Thorat

Introduction

Over the last decade, online streaming services have become more prevalent than before. The way the world is utilizing these platforms has changed. There is a rising demand for content, which brings accelerating competition within the streaming services industry. With a plethora of choices, the users are now in one of the most difficult conundrums ever to figure out the best fit OTT platform that suits their preferences. During the pandemic, while many industries are hurting, online streaming platforms like Netflix, Amazon Prime Video, and Disney Plus are booming.

To analyze these changes and the effect of the sudden outbreak of pandemic on OTT business, I have created a data visualization project which addresses the following questions:

- 1) What are the differences between the ways online streaming companies target the audience and distribute their content?

In this question, I will be investigating whether there is any significant difference in factors like genre, age group, year of release, or movie/show ratings for each of them?

- 2) What is the OTT platform's usage pattern across different countries? And did Covid-19 influence the usage of these services?

To answer these questions, I have collected data about the content on three online streaming platforms, Netflix, Amazon Prime Video, and Disney Plus. And data for different countries which use these services. Also, data for worldwide Covid-19 cases and google search trend data to check the number of hits to these platform's websites was collected. Collecting data for the online streaming platforms was not very difficult but making the project ambitious took a lot of effort. I am including data from multiple sources (tables) to include different factors affecting my research questions such as – a country's infrastructural availabilities, subscription pattern of these countries, and their population.

Methodology

This section explains the data extraction, data preparation and the fields in tables being used.

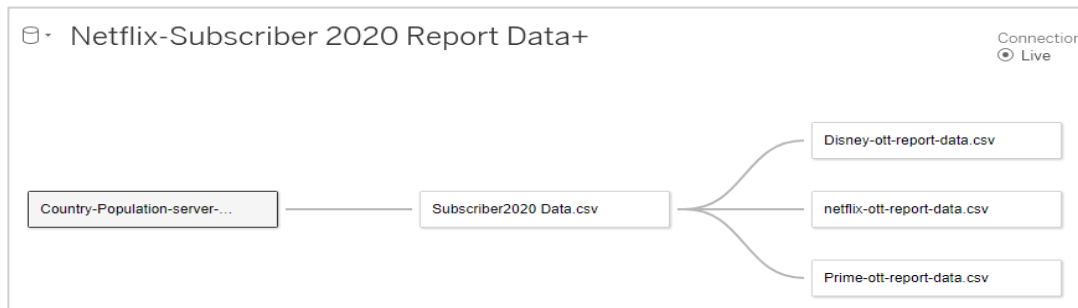


Fig. 1

1) Streaming Platform Data – 3 tables

We have three tables of leading streaming platforms - Netflix, Amazon Prime Video, and Disney Plus. The unit of analysis in these tables is a movie or TV show presented on these platforms. These platforms were launched at different years, therefore there was some difference in the period of data within each sheet. Also, there were a lot of unnecessary columns such as movie/show storyline, actor, and director information. This raw data was cleaned up using tableau prep builder and excel add-in tools. I am analyzing data for 6 major viewing countries of different regions. Asia and Japan from Asia-Pacific region, Canada and the United States from North America, and the United Kingdoms and Germany of Europe region.

Columns of these table are

Title	Year	Age Rating	IMDb Rating	Type	Genre	Country	Region
Title of movie/show	Year of release	Age group permitted to view	IMDb ratings of the content	Is it a movies or show	Genre of the content	Country in which the content is available	Region of the country

2) Country Population and Internet server Data – 1 table

The raw data downloaded in excel had columns for each year from 1990 to 2020, and these columns had code names. To convert this as per my requirement I used metadata files for appropriate column names and tableau prep builder to filter out

unnecessary data. This table also has data about the internet server installed to date, which gives an idea about a country's accessibility to suitable infrastructure.

Columns of the table

Country Name	Continent	Population	Percent Population	Internet server
Name of the country	Region of the country	Population of the country	Percent of world population	Internet server per million people

3) Subscriber 2020 Data – 1 table

To form this table, I had to visit the official websites of each streaming company and collect the latest subscriber data for each of them. This table contains a count of subscribers for Netflix, Disney, and Prime by 2020.

Fields of the table

Country	Continent	Netflix Subscriber	Disney Subscriber	Prime Subscriber
Name of the country	Region of the country	Count of Netflix subscribers for 2020	Count of Disney subscribers for 2020	Count of Netflix subscribers for 2020



Fig. 2

4) Google search trend data – 3 tables

Google trends is an official website that collects the data when a user hits a search query. These 3 tables have data about the number of times a platform's website was searched. I collected data for 2020 and 2021. This data will give us an overall idea about the change in the trend of OTT platform usage when the pandemic hit the world.

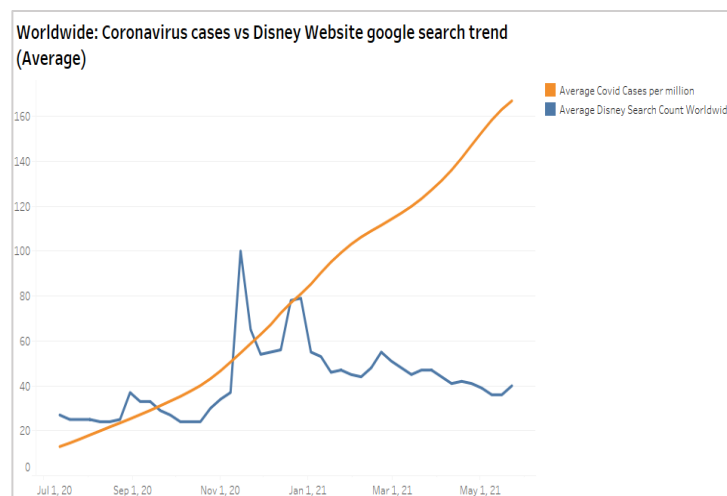
Date	Google search count
Date of search	Worldwide google search count

5) World Covid-19 Cases data – 1 table

The world covid-19 case data provides the number of cases confirmed each day for 2020 and 2021. As these numbers were in 100 million, I had to standardize it by dividing it by 1000000. This won't change the trend in the data.

Date	Confirmed cases per million
Date of case occurrence	Number of cases per million encountered on a particular day

Analysis



1. Simple

These 3 simple visualizations are a time series analysis of worldwide Covid-19 cases in comparison with the google search count of Disney Plus Website (fig. 3), Prime video (fig. 4), and Netflix (fig. 5). These graphs use data from two tables – the world Covid-19 cases table and google search table. This shows that as the Covid-19 cases rapidly increased and many countries were in a complete lockdown situation, the number of hits to the Disney+ website hiked. The hit went from an average of 27 at the beginning of July 2020 to an average of 100 hits by November 2020. This shows that the pandemic did have quite an effect on the

Fig. 3

business of the Disney+ platform. We can observe from the graph (fig. 4) that the pandemic did have some effect on the prime video and Netflix usage, but not as much as Disney+. Their average number of hits to the website range from 60 hits to 100 hits. Possible reasons behind this can be the content on the platform or actions taken by the competitor in favor of itself. This answers a part of my research Question 2 by presenting the effect of covid-19 on the usage of these platforms.

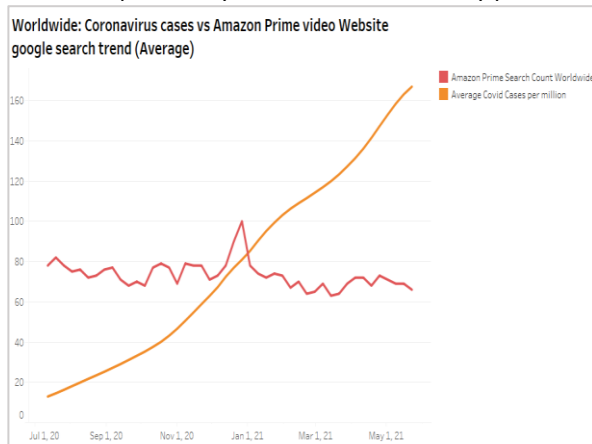


Fig. 4

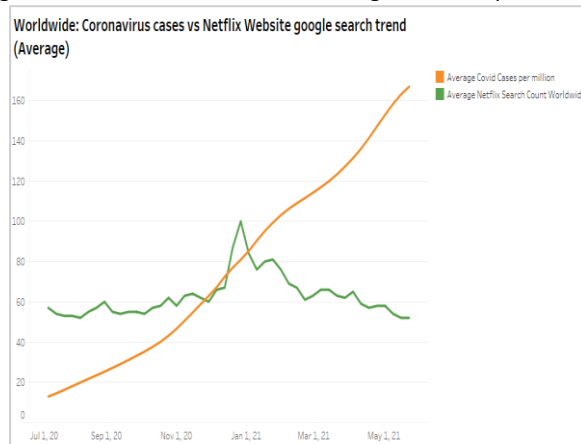


Fig. 5

2. Simple

The popularity depicting visualization was created using the Google search data for the year 2020 – 2021. This clearly shows that Amazon Prime Video is the most preferred platform across the world during the pandemic phase. The chart shows the daily average hit of a platform. Prime video website is hit the most approx. 72 times a day, while Disney+ and Netflix have on an average of 43 and 61 hits daily. Therefore, from simple visualization 1 and simple visualization 2 graphs we can answer that pandemic did spike the usage of Disney+ as it's hits hiked from the start point of 27 around July 2020 and at a point reached 100. Although the usage of Netflix and Prime Video remain consistent during this period, they remained the most preferred.

Popularity of platforms by google search

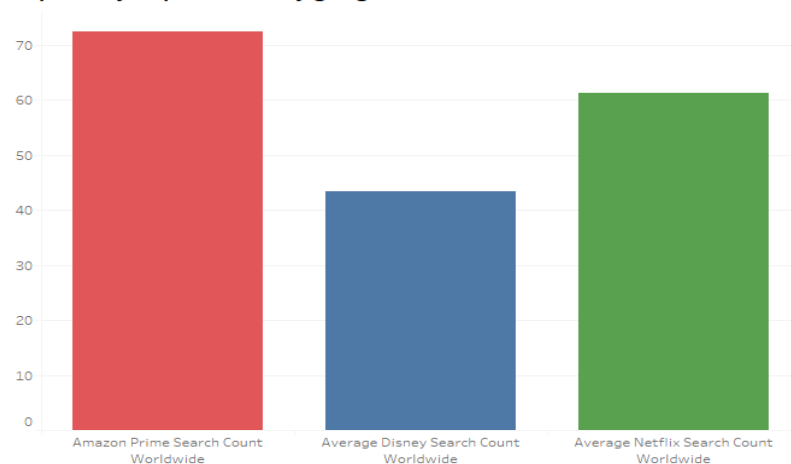


Fig. 6

3. Simple

Below we have 3 simple visualizations (fig. 7, 8, and 9) displaying the platform's age group-wise content distribution. These charts address research question 1. Disney+ presents content that is suitable for an audience of all age groups. It has content that is rated TV-Y / TV-Y7 and TV-G and approved by the board. That means Disney focuses to target all age groups but mostly kids of age 7+. Prime and Netflix on the other hand target 13+ and 18+ audiences. This shows that these platforms do have a different set of target audiences. This shades some light to question 1 in accordance with the age group factor. Factors other than age are explored going further.

Age-wise Content Distribution on Disney

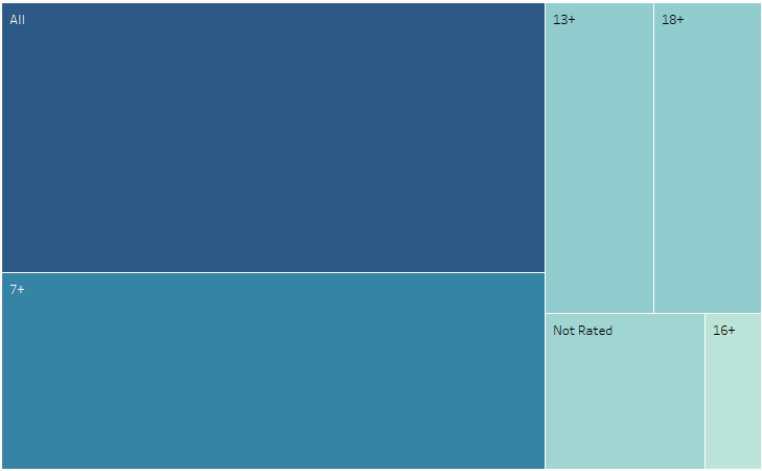


Fig 7

Age-wise Content Distribution on Netflix



Fig. 8

Age-wise Content Distribution on Prime video

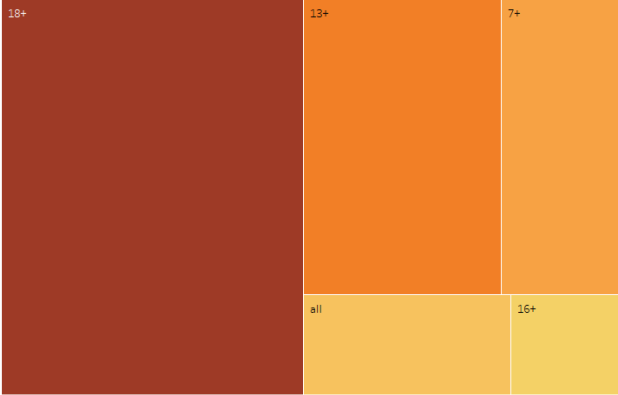


Fig. 9

4. Simple

Below simple visualizations (fig. 10, 11, and 12) analyzes question 1, whether these platforms differ in content distribution considering the genre of movies/shows. Disney focuses more on animation, musicals, and fantasy which seems to be aligned with the above analysis on age group i.e. the kids of age 7+. Also, the next preferred genre is Comedy, documentary, drama, and fantasy for all age groups. Prime has most content in drama, documentary, action, adventure, thriller, and romance. Netflix seems to have content widespread in all the genre. Thus, the genre focus of primes and Netflix also seems to be aligned with our analyzed age group of 18+ and 13+.

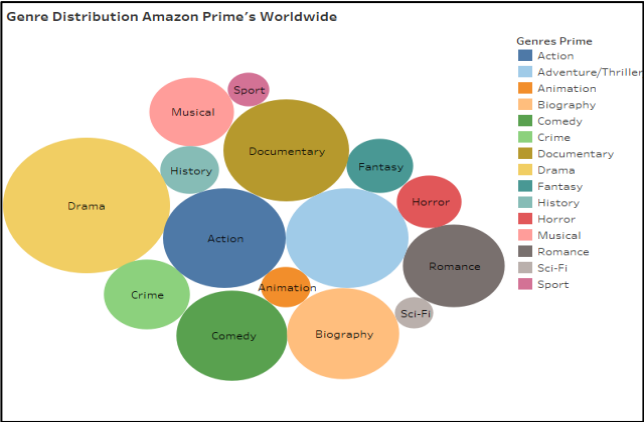


Fig. 10

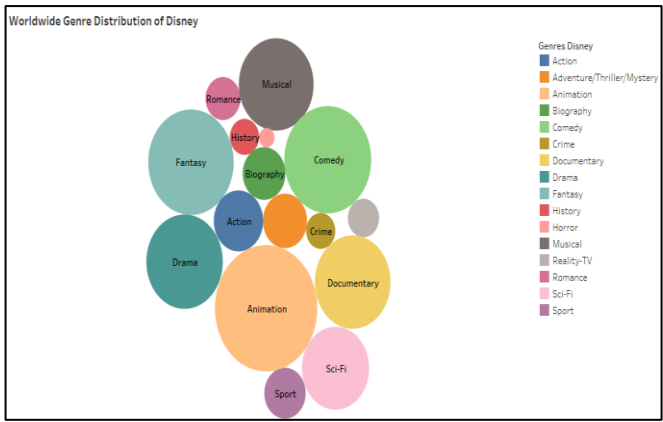


Fig. 11

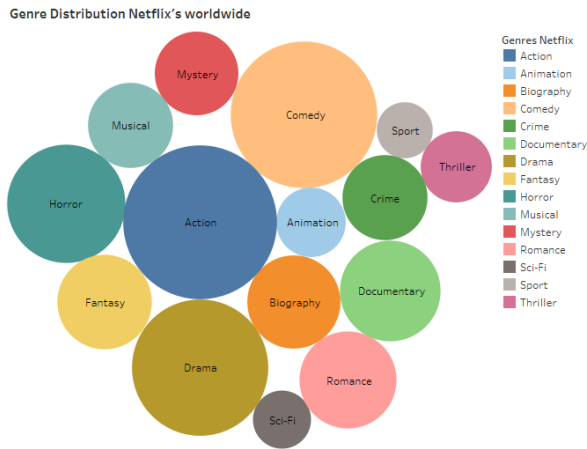


Fig. 12

5. Drill Down

Below drill-down visualization analyzes the region-wise and country-wise **usage** of the 3 mentioned platforms in comparison to the country-wise **ratings** of their content. We can find out the overall subscription rate for regions such as - Asia-pacific, North America, or Europe. If a user wants to investigate specific countries in these regions, he can drill down into these charts.

All these countries differ in their population count and infrastructure, therefore, to avoid misleading data representation I did some standardization. To normalize the subscription data, the subscription rate of each country was calculated by dividing the subscription count by the population of the country.

The insights from the charts are as below-

Even though the ratings for each platform are almost between 6 to 8, the subscription rate varies across different countries. The subscription for each of these platforms is very fewer in countries like India, Japan, and Germany as compared to the United Kingdom, US, and Canada. Netflix service was launched in the US, United Kingdom, and Canada around 2007 to 2010, however it reached India, Japan, and Germany in later years, around 2014. Also, Disney+ was launched in India, Japan, and Germany a year later than others. This might explain the difference in the subscription rate. Accessibility to resources such as the internet in these countries can also be one of the major reasons for this difference. The accessibility factor is analyzed in dashboard 1 using a map.

Comparison between Prime Video overall IMDB Rating vs Subscriber Count as per Country.

The subscription data is normalized as per the Country's Population.

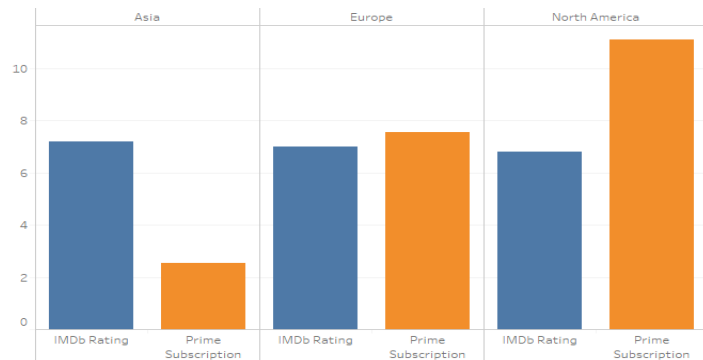
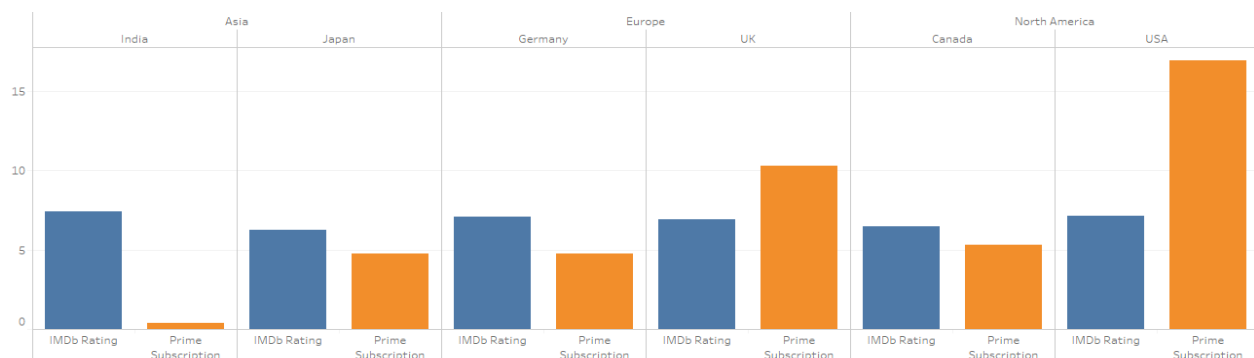


Fig. 13



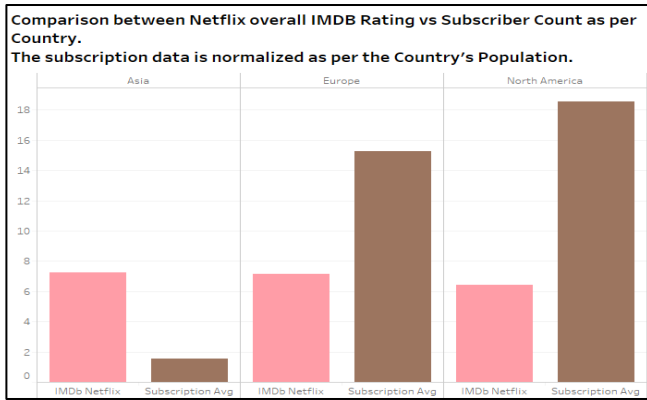


Fig. 14

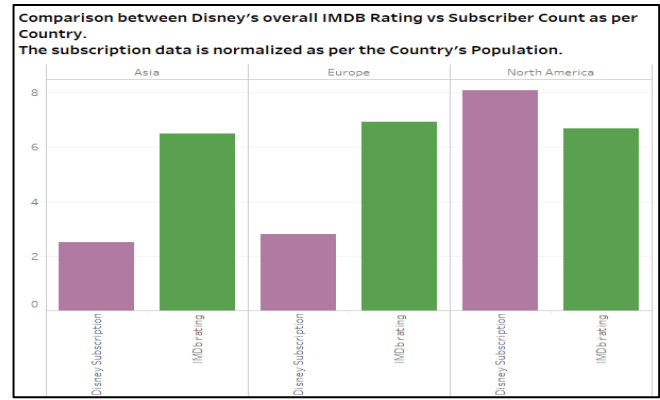


Fig 15

6. Animation

Comparison of IMDB ratings against the Number of movies on Netflix

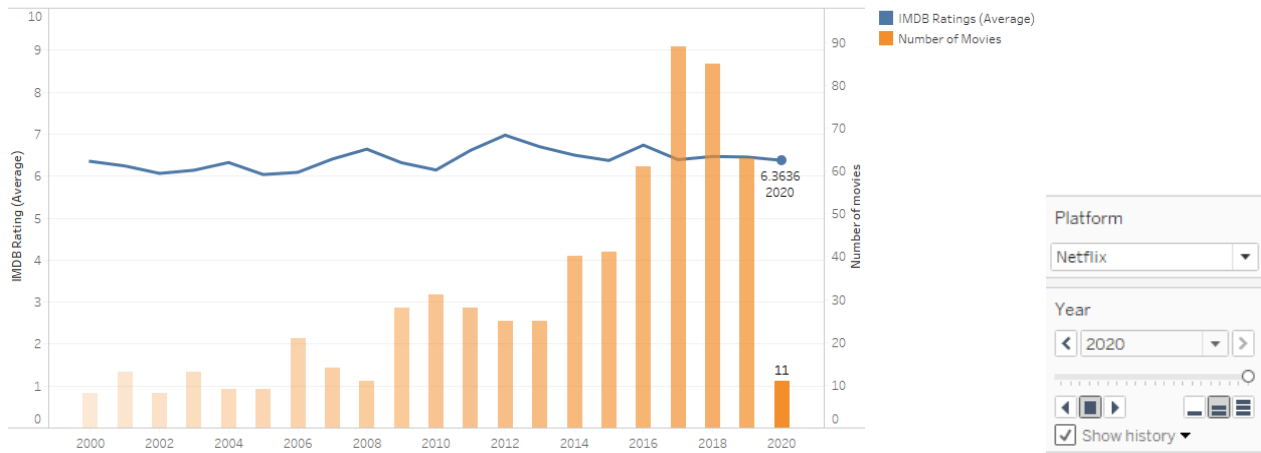


Fig. 16

This is an animation visualization with a filter option. The animation is comparing IMDb ratings against the number of movies/shows released in the same year presented by the OTT services. The visualization provides a filter option for our 3 platforms, so we can observe the comparison for each of them. As we select a specific platform and press play the line chart of averaged out IMDb rating of all movies released in a certain year and the number of movies column bars start to move from left to right. For this visualization, two parameters, 'Platform Selected' and 'Year Selected', were created. Also, 'Platform Selected imdb' and 'number of movies' calculated fields were created.

The insights from the visualization are – the ratings for the movies on each of these platforms are in the range of 6 to 8 but it can be observed that Netflix and prime video have more content from later years i.e. mostly starting from 2008 whereas Disney+ have a lot of older as well as newer movies. This analysis helps us gain some insight on question 1 as we can clearly see that the platforms differ in content distribution considering the release year of the movie.

7. Map

A map is created to analyze the total IMDb rating of all three platforms for different countries. This is done to analyze the overall rating for all the content presented irrespective of the OTT service provider. This gives an overview of whether OTT platforms are preferred or not in a specific country. This map is used in one of the dashboards.

The rating data is distributed in multiple tables and within each table there are multiple records for multiple countries. Therefore, to aggregate these ratings for the individual country from all the tables, a calculated field was create using the below formula

$$\frac{(\text{AVG}(\{\text{FIXED} [\text{Country Prime}] : \text{AVG}([\text{IMDb Prime}])\}) + \text{AVG}(\{\text{FIXED} [\text{Country Disney}] : \text{AVG}([\text{IMDb Disney}])\}) + \text{AVG}(\{\text{FIXED} [\text{Country Netflix}] : \text{AVG}([\text{IMDb Netflix}])\}))}{3}$$

IMDb rating of movies across different countries

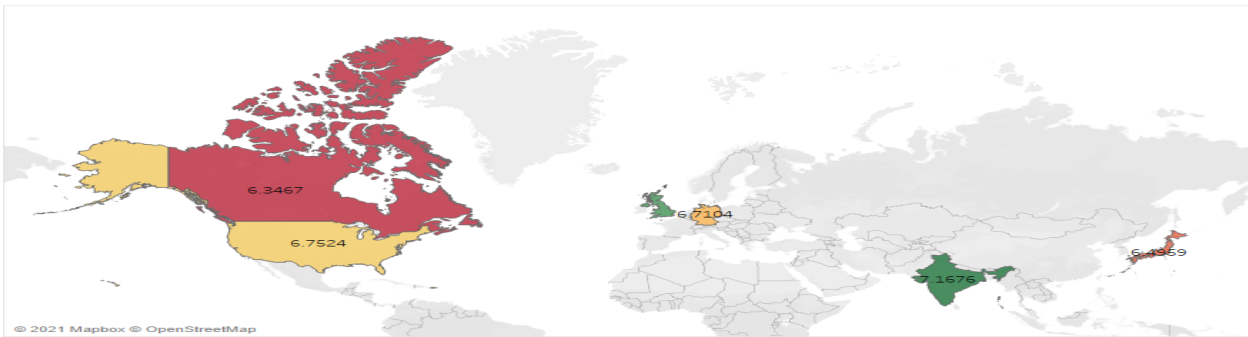


Fig. 17

8. Dashboard

The dashboard is created to analyze the usage of online streaming services across different countries. This is a combination of 2 maps and 3 drill down charts. Also, the dashboard uses data from 3 different tables. As mentioned above the drill down charts make use of subscriber 2020 data and OTT data tables. Whereas to analyze the worldwide accessibility to the internet another table called 'Country Population and Internet server Data' is used. Also, if a user wants to analyze a specific country, then the dashboard can be filtered by clicking on the country in any of the maps.

Insights – as explained previously, the subscription and ratings analysis are done in the drill down charts and 'IMDb rating of movies across different countries' map. The 'Count of Internet servers worldwide' map is used to investigate the accessibility to the internet worldwide because this is one of the factors that might affect the usage of these services. We can view details about the subscription data for each platform in different countries, the response a platform gets from users in different country (in form of rating), and each country's accessibility to the internet. Therefore, the dashboard provides us with a collaborated visualization to compare multiple aspects of a country's usage.

I have incorporated some special features of tableau in both my dashboards, such as the information icon at the top and horizontal panels to add a description of the charts. When the user hovers on this icon a tooltip appears stating the gist of the dashboard. Also, filtering the dashboard by clicking countries on the maps was achieved using the action feature of the dashboard.

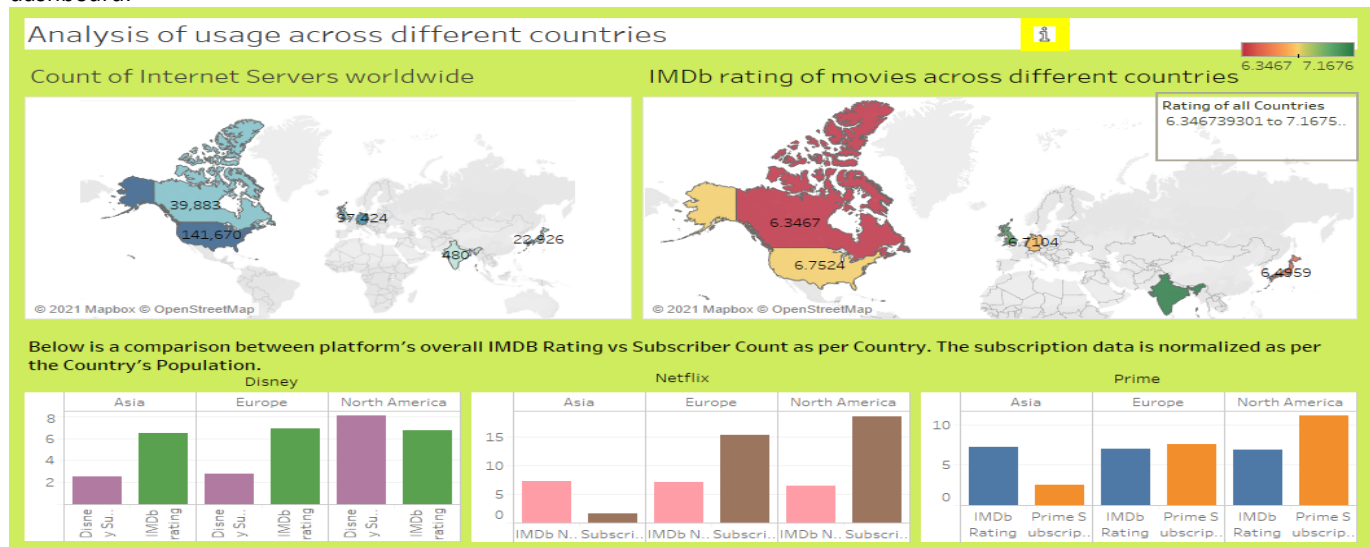


Fig. 18

9. Dashboard

A single dashboard is created to have an overview of both the question and to analyze them together. This dashboard combines the animated visualization and 3 genre distribution analysis charts. It can be used to examine the strategy behind the content distribution. The bubble charts can be filtered for a specific country using the checkbox filter provided. Also, the animation can be visualized for all three platforms.

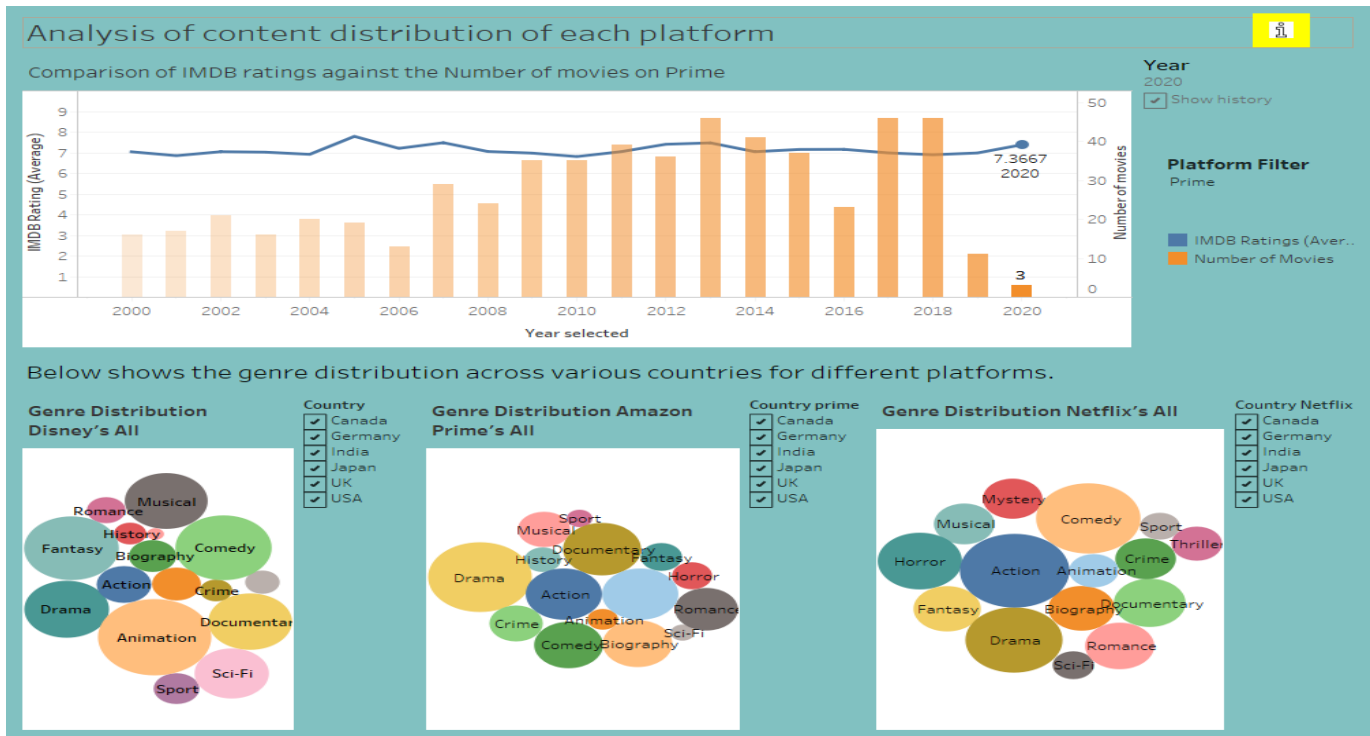


Fig. 19

Conclusion

The above stated research questions can be answered using the visualizations created in the project.

1. What are the differences between the ways online streaming companies target the audience and distribute their content?

Yes, it can be precisely interpreted that the content distribution differs on each platform. There can be seen some strategy behind the content distribution. Each platform has a specific preference when it comes to the genre of movies/shows, age groups, year of content release and ratings to be focused.

1. What is the OTT platform's usage pattern across different countries? And did Covid-19 influence the usage of these services?

The OTT service usage can be clearly studied in the above visualizations. The usage is different for each country and the factors affecting it might be the population, infrastructure, or popularity of the services in those countries.

Some additional questions can be addressed such as - analysis of the type of content focused by these platforms over the course of years. A study can be done if there has been any shift in the type of content i.e. movies or shows, focused by these platforms over a period of 10 years.

Another question that can be studied is if economic factors of the countries such as GDP, foreign revenue, and urbanization percentage are correlated to the usage of these OTT services.

Sources

Internet Server Data: <https://data.worldbank.org/indicator/IT.NET.SECR.P6?view=chart>

Country Population data - <https://databank.worldbank.org/home.aspx>

Google Trend Data: <https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset>

Disney plus subscriber data: <https://medium.com/@flixpatrol/disney-analysis-the-library-subscribers-and-features-55c9eab97bc1>

Amazon Prime video subscriber data: <https://market.us/statistics/online-video-and-streaming-sites/amazon-prime-video/>

Netflix subscriber data: <https://backlinko.com/netflix-users>