

MTH208A - Data Science Lab

Analysis on Indian Web Series

Project Report

By Group No 7:

- Arijit Dey (221281)
- Kaushik Thakkar (221329)
- Kumar Kanishk Singh (210544)
- Shailza Sharma (221416)

Under the guidance of Dr. Dootika Vats

Contents

<u>Introduction</u>	3
<u>Dataset Description</u>	4
<u>Possible Biases in the data</u>	8
<u>Some interesting questions:</u>	9
<u>Shiny App</u>	9
<u>Some important visualizations</u>	11
1. <u>Most Voted Genre:</u>	11
2. <u>Most voted Certificate category:</u>	11
3. <u>Most Voted Length category:</u>	12
4. <u>Gender classification of votes of web series:</u>	12
5. <u>Most Rated Genre, Certificate category, and Length category:</u>	14
6. <u>Which genre is most popular on which platform?</u>	15
7. <u>Which platform targets which age group?</u>	16
8. <u>Comparison of average runtime of different platforms:</u>	16
9. <u>Visualizing Votes-Ratings association:</u>	17
10. <u>Does Ratings or Votes depend on Runtime?</u>	18
<u>Conclusion:</u>	19
<u>References</u>	19
<u>Acknowledgment</u>	19

Introduction

The project is based on **Indian Web Series** Data.

- **What is a Web Series?**

A web Series is a show/program that runs on an internet streaming platform rather than running on Television. As the popularity of the internet has grown tremendously over the past couple of years, the popularity of such shows has grown similarly. In India people used to be addicted to Television before 2014-15. But with the increased number of internet users, the user base for streaming platforms has also significantly grown in India. “Let’s Netflix and Chill” has become one of the most popular lines among the youth today. We can surely get an idea about the popularity of web series from all this.

- **Why did we choose Indian Web Series?**

Now, the major reason for specifically choosing Indian web Series and not the other ‘Hollywood’ web series like Big Bang Theory, Game of Thrones, etc. is because there have already been hundreds of analyses done on these. You can just go on the internet, type their name and you will find multiple visualizations, analyses, etc. not just on the web series but also on their episodes; just because of the fact that these have been there for a very long period of time.

On the other hand, the concept of Web Shows and Web Series is comparatively new to India. The growth of Indian web series has only been happening for the last 5-6 years.

As we wanted to do something new, we chose this.

The other reason for this is that the Indian audience is a huge audience. It has been one of the most targeted audiences in the world by everyone around the globe. So to get an insight into the preferences of the Indian Audience surely lets the directors know which direction to move in. For example, soon after the success of ‘Mirzapur’ and ‘Sacred Games’, which are some Indian origin action and crime web series, we saw a tremendous increase in the release of such crime and drama web series in India such as ‘Paataal Lok’, ‘Criminal Justice’, etc. Now, what does it tell us about the liking of the audience? This is the information that we want to get from this data set. We want to know how these web series managed to get so popular among the audience which was so used to those melodramatic TV series like ‘Kyunki Saas Bhi Kabhi Bahu Thi’, ‘Kasautii Zindagi Kay’ or even Indian mythological Series such as ‘Ramayana’ and Mahabharata’.

- **Goal of the project:**

The eventual goal of this project is to get some insightful information regarding the platforms, genre, and overall preferences of the Indian audience when it comes to watching Indian web shows. The data that we have collected will surely help us gain some knowledge about the mindset of the people behind these streaming platforms as to what their target audience is, which genre is most liked by the Indian audience

so far, and overall help us get a little understanding of the patterns in which these shows are being released to suit the Indian audience. We will not be doing any kind of Statistical Analysis on this data so hence our conclusions will be based on purely what the data says through multiple visualizations performed in our Shiny App for all the different questions that we have regarding this target audience.

Dataset Description

1. Obtaining the data:

- Source: We collected all the data required for this project from the website IMDb.
 - IMDb or Internet Movie Database is the most popular, authoritative online database containing information about various films, television series, video games, and streaming content, along with information about actors and directors, cast and crew, box office collection, trailers, and other related content.
 - Website link: <https://www.imdb.com/>
- Data Scrapped: We scrapped data about the top web series of the various Indian web series streaming platforms.

Data has been scrapped for the following platforms:

Alt Balaji, Amazon Prime, Hotstar, MX Player, Sony LIV, TVF, Voot Select, and Zee5

And the individual page link of each platform on the IMDb website are as follows –

- Alt Balaji: <https://www.imdb.com/search/title/?companies=co0642697>
 - Amazon Prime: https://www.imdb.com/list/ls099967158/?st_dt=&mode=detail&page=1&sort=moviemeter,asc
 - Hotstar: <https://www.imdb.com/search/title/?companies=co0800964>
 - MX Player: <https://www.imdb.com/search/title/?companies=co0721599>
 - Sony LIV: <https://www.imdb.com/search/title/?companies=co0546496>
 - TVF: <https://www.imdb.com/search/keyword/?keywords=tvf>
 - Voot Select: <https://www.imdb.com/search/title/?companies=co0603531>
 - Zee5: https://www.imdb.com/search/title/?companies=co0692549&sort=user_rating,desc
- Now, for each web series, we scrapped data about its Genre, Certificate, Run time, Starting year, Overall Ratings and Votes, and male and female ratings and votes.

2. Methods and libraries used:

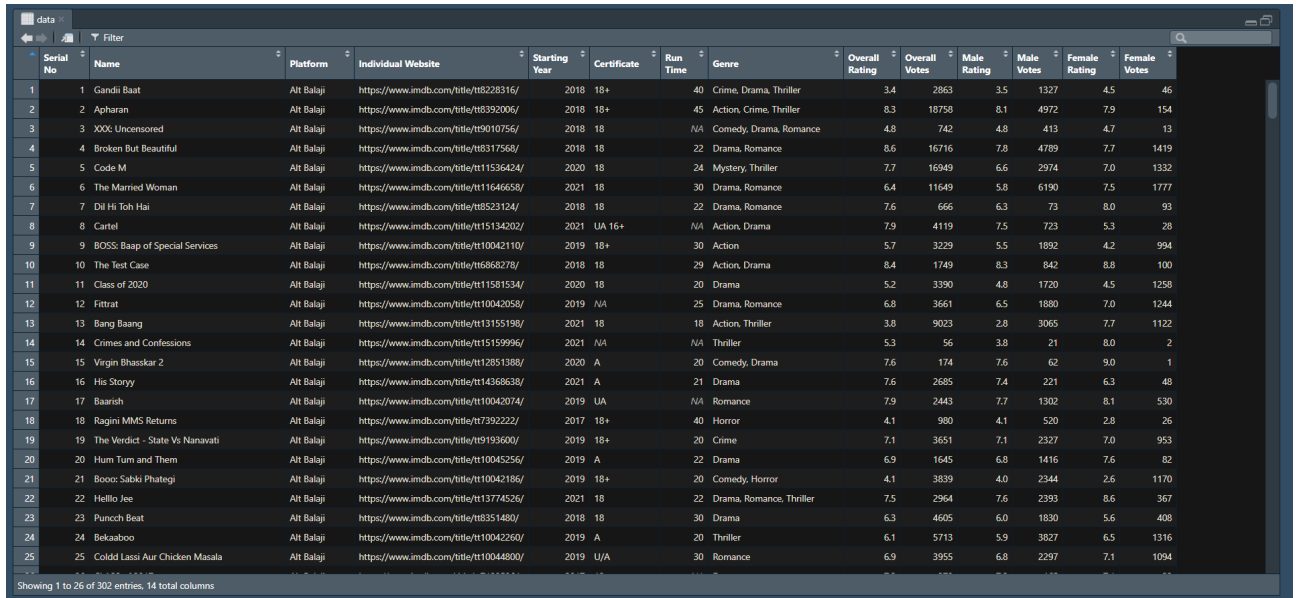
- We used R language to scrape the data from IMDb using the RStudio software.

- And we used the following libraries for scrapping:

```
library(rvest)
library(tidyverse)
```

- Rvest: We used this package to load the source code of a website page in R and then search for various information inside the code.
The functions like `read_html`, `html_elements`, `html_attr`, `html_text`, and `html_table` inside this package help us do the scrapping.
- Tidyverse: This library we used to perform various kinds of string manipulations that were needed to scrape the particular information we needed.
The functions `str_sub`, `str_trim`, `str_replace_all`, `str_remove_all`, etc. were used repeatedly in the code to get the desired string form. Also, this package contains the pipe operator, `'%>%'` which came in handy in almost every line of the code.

3. Date and Time of scrapping: As, in IMDb, the users rate and vote for the web series, the list of top web series may shuffle with time. This is why we should mention the date and time of scrapping the data. We scrapped data for all the web series from IMDb on 16th October 2022 at about 11 pm.
4. Data frame description: Finally, the scrapping process was done and we ended with a data frame with 14 columns and 302 rows which looks like this –



Serial No	Name	Platform	Individual Website	Starting Year	Certificate	Run Time	Genre	Overall Rating	Overall Votes	Male Rating	Male Votes	Female Rating	Female Votes
1	1 Gandhi Baat	Alt Balaji	https://www.imdb.com/title/tt8228316/	2018	18+	40	Crime, Drama, Thriller	3.4	2863	3.5	1327	4.5	46
2	2 Aphan	Alt Balaji	https://www.imdb.com/title/tt8392006/	2018	18+	45	Action, Crime, Thriller	8.3	18758	8.1	4972	7.9	154
3	3 XXX: Uncensored	Alt Balaji	https://www.imdb.com/title/tt9010756/	2018	18	NA	Comedy, Drama, Romance	4.8	742	4.8	413	4.7	13
4	4 Broken But Beautiful	Alt Balaji	https://www.imdb.com/title/tt8317568/	2018	18	22	Drama, Romance	8.6	16716	7.8	4789	7.7	1419
5	5 Code M	Alt Balaji	https://www.imdb.com/title/tt11536424/	2020	18	24	Mystery, Thriller	7.7	16949	6.6	2974	7.0	1332
6	6 The Married Woman	Alt Balaji	https://www.imdb.com/title/tt1164658/	2021	18	30	Drama, Romance	6.4	11649	5.8	6190	7.5	1777
7	7 Dil Hi Toh Hai	Alt Balaji	https://www.imdb.com/title/tt8523124/	2018	18	22	Drama, Romance	7.6	666	6.3	73	8.0	93
8	8 Catel	Alt Balaji	https://www.imdb.com/title/tt15134202/	2021	UA 16+	NA	Action, Drama	7.9	4119	7.5	723	5.3	28
9	9 BOSS: Baap of Special Services	Alt Balaji	https://www.imdb.com/title/tt10042110/	2019	18+	30	Action	5.7	3229	5.5	1882	4.2	994
10	10 The Test Case	Alt Balaji	https://www.imdb.com/title/tt668278/	2018	18	29	Action, Drama	8.4	1749	8.3	842	8.8	100
11	11 Class of 2020	Alt Balaji	https://www.imdb.com/title/tt11581534/	2020	18	20	Drama	5.2	3390	4.8	1720	4.5	1258
12	12 Fitrat	Alt Balaji	https://www.imdb.com/title/tt10042058/	2019	NA	25	Drama, Romance	6.8	3661	6.5	1880	7.0	1244
13	13 Bang Baang	Alt Balaji	https://www.imdb.com/title/tt13155198/	2021	18	18	Action, Thriller	3.8	9023	2.8	3065	7.7	1122
14	14 Crimes and Confessions	Alt Balaji	https://www.imdb.com/title/tt15159996/	2021	NA	NA	Thriller	5.3	56	3.8	21	8.0	2
15	15 Virgin Bhaskar 2	Alt Balaji	https://www.imdb.com/title/tt12851388/	2020	A	20	Comedy, Drama	7.6	174	7.6	62	9.0	1
16	16 Hia Story	Alt Balaji	https://www.imdb.com/title/tt14368638/	2021	A	21	Drama	7.6	2685	7.4	221	6.3	48
17	17 Baarish	Alt Balaji	https://www.imdb.com/title/tt10042074/	2019	UA	NA	Romance	7.9	2443	7.7	1302	8.1	530
18	18 Ragini MMS Returns	Alt Balaji	https://www.imdb.com/title/tt7392222/	2017	18+	40	Horror	4.1	980	4.1	520	2.8	26
19	19 The Verdict - State Vs Nanavati	Alt Balaji	https://www.imdb.com/title/tt9193600/	2019	18+	20	Crime	7.1	3651	7.1	2327	7.0	953
20	20 Hum Tum and Them	Alt Balaji	https://www.imdb.com/title/tt10045256/	2019	A	22	Drama	6.9	1645	6.8	1416	7.6	82
21	21 Booo: Sabki Phategi	Alt Balaji	https://www.imdb.com/title/tt10042186/	2019	18+	20	Comedy, Horror	4.1	3839	4.0	2344	2.6	1170
22	22 Hello Jee	Alt Balaji	https://www.imdb.com/title/tt13774526/	2021	18	22	Drama, Romance, Thriller	7.5	2964	7.6	2393	8.6	367
23	23 Punchni Beat	Alt Balaji	https://www.imdb.com/title/tt8351480/	2018	18	30	Drama	6.3	4605	6.0	1830	5.6	408
24	24 Bekaaboo	Alt Balaji	https://www.imdb.com/title/tt10042260/	2019	A	20	Thriller	6.1	5713	5.9	3827	6.5	1316
25	25 Codd Lassi Aur Chicken Masala	Alt Balaji	https://www.imdb.com/title/tt10044800/	2019	U/A	30	Romance	6.9	3955	6.8	2297	7.1	1094

- Column description: The final data frame contains columns that are described as follows –

Column_Name	Content	Data Type
Serial No	The serial number of every observation	Integer
Name	Name of the web series	String
Platform	Name of the platform a web series belongs to	String
Individual Website	Webpage link of web series in IMDB site	String
Starting Year	Starting year of a series	Integer
Certificate	Certificate category a web series	String
Run Time	Average length of each episode of a series	Integer
Genre	Genre of a web series	String
Overall Rating	Rating of a particular series	Integer
Overall Votes	Total number of votes given to a particular series	Integer
Male Rating	Ratings given by males	Integer
Male Votes	Votes given by males	Integer
Female Rating	Ratings given by females	Integer
Female Votes	Votes given by females	Integer

- Row description: As mentioned above we got 302 web series entries which include the top web series of the eight platforms mentioned above. But due to the unavailability of entries, the number of web series corresponding to each platform is not the same. It varies from 50 to a minimum of 10. The platform-wise distribution is as follows –

```
table(data$Platform)
```

```
##
##   Alt Balaji Amazon Prime   Hotstar   MX Player   Sony LIV   TVF
##         50         20         10         50         48        24
## Voot Select       Zee5
##         50         50
```

5. Problems tackled while scrapping and the way outs: Among many problems that we faced during scrapping the data, notable ones are –

- As usual and with no surprise, there was a lot of missing values on the websites. Like in many cases there was no mention of Run time and in many cases, the certificate information was missing. And the total number of perfect rows we got was 149 only. Now, this created a problem in the early stage of scrapping because we were trying to scrape the data column-wise. Like, were writing codes to scrape the entries for the 'Run time' column of all the web series and finally make a data frame using the columns. But thanks to the missing values we were getting columns of different lengths which not only were problematic to create a data frame but also we had no idea

about which web series the missing value corresponds to.

Then came a suggestion from our fellow friend which changed our entire methods. Now we were scrapping row-wise. Meaning, we were taking the chunk of code that maps an individual web series and were searching inside that chunk for our required columns, and putting 'NA' in the case of missing values.

- In some of the platforms, we noticed a lot of duplicated value in the 'Name of the Web series' column and after a small-scale investigation, we found that after some web series, entries corresponding to episodes of web series were coming instead of web series as a whole.

For example, on the webpage for the 'Top web series' of Hotstar, different rows were there for the episodes 'Ab judge karo mujhe', 'Disco Deewane', and 'Hawa Hawai' of the web series 'Hundred', which was not desirable for us. So, we removed those entries by manually putting some limits over the scrapping range.

6. Adding two more columns: In this stage, we find that there were 17 different of certificates in our dataset, and for the continuous variable 'Runtime' we had 46 different entries.

```
length(unique(data$Certificate))
```

```
## [1] 19
```

Note that we ignored the two certifications, namely 'Not Rated' and 'NA' which are appearing here as unique certifications. And,

```
length(unique(data$`Run Time`))
```

```
## [1] 46
```

Now for certificates, as handling this many categories would have been a difficult task and also comparing them all with one another would give a messy outcome, we categorized the 'Certificate' column suitably and created one new column.

And for the continuous type variable 'Run time' too, we did the similar thing of categorizing them into suitable categories.

- Certificate Category:

We categorized the web series into three categories with respect to their watchability to different age groups, viz.

- For 'Family and Kids': With the following certification – '12+', 'UA 7+', 'U', '7+', 'All'
- For 'Teenagers': With the following certification – 'UA', 'UA 16+', 'U/A', 'PG', '16+', '16', '13', '15+', 'UA 13+'
- For 'Adults': With the following certification – '18+', '18', 'A'

- And we filled in ‘NA’ for the certifications ‘Not Rated’ and ‘NA’
- Length Category:
Here again, we categorized the web series into three categories with respect to the average length of episodes of that web series, viz.
 - ‘Short’ length web series: With an average length of episodes being less than 20 minutes.
 - ‘Medium’ length web series: With an average length of episodes being in between 20 minutes and 40 minutes.
 - ‘Long’ length web series: With an average length of episodes being more than 40 minutes.
 - And we filled ‘NA’ in the case of missing values of Run time.

7. So, finally our data frame consists of 302 rows and 16 columns altogether.

```
dim(data)
```

```
## [1] 302 16
```

Possible Biases in the data

Biases impact the quality of the data analysis significantly. Thus it is important to ensure that our analysis is free of those. As no data is free of biases, it is quite difficult to eliminate them. So, here we identify the bias embedded in our data:

- Selection Bias: Selection bias occurs when our data is not representative of the target population. Some scopes of selection biases in our data are as follows-
 - Scrapping only from the website IMDb:
There may be various other websites with ratings and votes of web series, which we didn’t take into our consideration. But we only analyzed data about web series coming from this particular website and this may induce Selection Bias.
 - Considering only eight platforms:
According to <https://www.getfleeek.app/blog/ott-platforms-in-india>, there are at least 20 OTT platforms in India streaming web series with or without cost. But we only considered eight platforms according to the availability of data on these platforms, which again can induce Selection Bias.
 - Scrapping only top web series:
We only scrapped data top 50 or 20 web series from each platform. But certainly, there are many more. If we considered them all then the trends might have been different from what they are now. So, this is clearly a significant source of Selection Bias.
 - Also, in the context of our analysis, this bias occurs as the female votes are way less than the males. So our data is biased towards males.

Some interesting questions:

1. Which certificates, which genre, and what length of web series are the most popular amongst people?
2. Does the length of a web series affect the engagement of the audience with the web series?
3. Is there any potential difference between the preferences of males and females when it comes to web series?
4. Which genre works best on which platform?
5. Which platform targets which age groups?

These are some questions we thought can be asked from the data scrapped above. And, in the course of this project, we will try to answer these with the help of some data visualizing tools.

Shiny App

1. Libraries used: The different libraries we used to run the app and plot the graphs are as follows-

```
library(shiny)
library(ggplot2)
library(reshape)
library(tidyverse)
library(plotly)
library(GGally)
library(shinythemes)
```

2. Description of Shiny app: The Shiny App has **3 sections**, viz.

- Overall Visualizations:

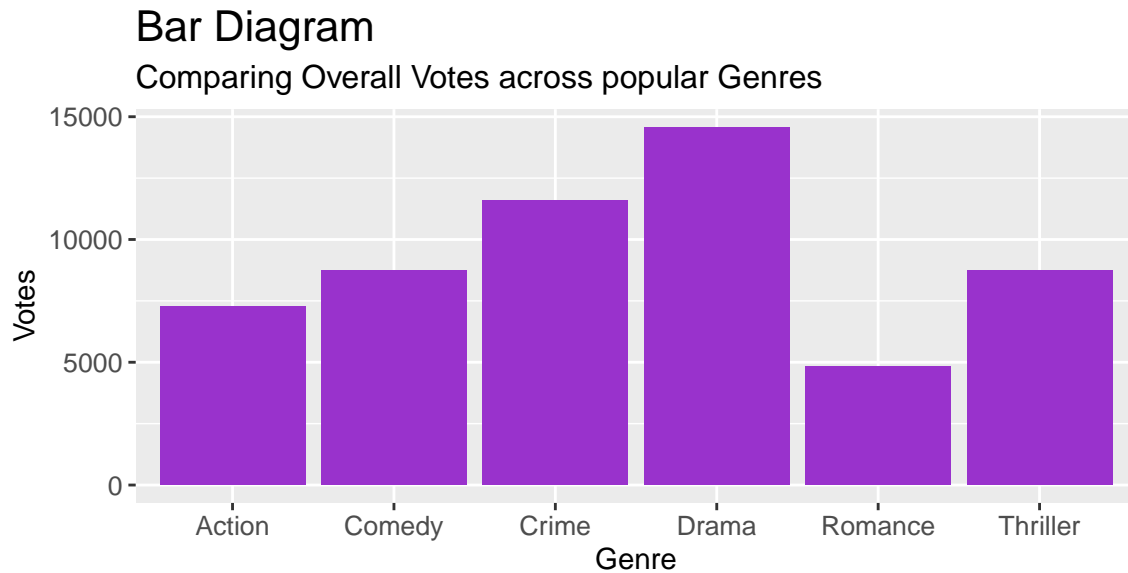
In this section, we see some plain visualization of different features of the data. For this,

- First we get an option to choose **platforms**. Here we can choose all the platforms at once or can select a particular platform as per our needs.
- Then, we get the **field of comparison**. For this, we have **three** options-
 - * Comparing Certificate Category
 - * Comparing Genre
 - * Comparing Length Category
- Then, we get further option to select the **basis of comparison**, and in this section, we have got **two** options-

- * Comparison with Ratings
 - * Comparison with Votes
- After selecting these two options, we get a **bar plot** with bars being different categories of the field of comparison and the heights representing the selected basis of comparison, e.g. **comparing Genres with respect to Votes**.
- Then we also have a **check box**, selecting which we get a **side-by-side bar plot** with **Male Female classification** of the same bar plot described above.
- Platform Wise Comparison:
 In this section we tried to show some comparisons between the different platforms we have got.
 - We have chosen three factors for comparison, viz.
 - * Comparison with respect to **Run time**
 - * Comparison with respect to **targeted Age group**
 - * Comparison with respect to **popular Genre**
 - In the first scenario, we get a **side-by-side box plot** of the average **run times of the web series** of different platforms to compare them with each other.
 - In the second case, we get a **percentage bar plot** with components of each bar showing **three components** corresponding to the proportion of web series with the **three different certificate categories** of a platform.
 Thus we get to compare the **targeted audience** of different platforms.
 - And finally we have a **multiple donut chart** with each donut representing the proportion of **different genres** of each platform. Here for the sake of comparison, we have taken only the popular genres that are common to each platform. Here we compare the platforms with respect to their **most popular genre**.
- Finding Association:
 This section focuses on **finding association** between different variables in our data. For example - **Overall Votes vs Overall Ratings** or **Run time vs Votes and Ratings**.
 - For the first case, we get a **scatter plot of overall** votes vs overall ratings with each web series **color-coded by the platform** it belongs to.
 This plot checks for any association between the two aforementioned variables.
 - For the second case, we get a **scatter plot matrix** with three variables viz. **run time, votes, and ratings**, and by this try to see some relation between these three variables.
- We note that in all the five plots described in Section 2 and Section 3, we have the **option to choose platforms** according to our need and in the first plot of the third section we have a **slider** to control the **outliers** in the data.

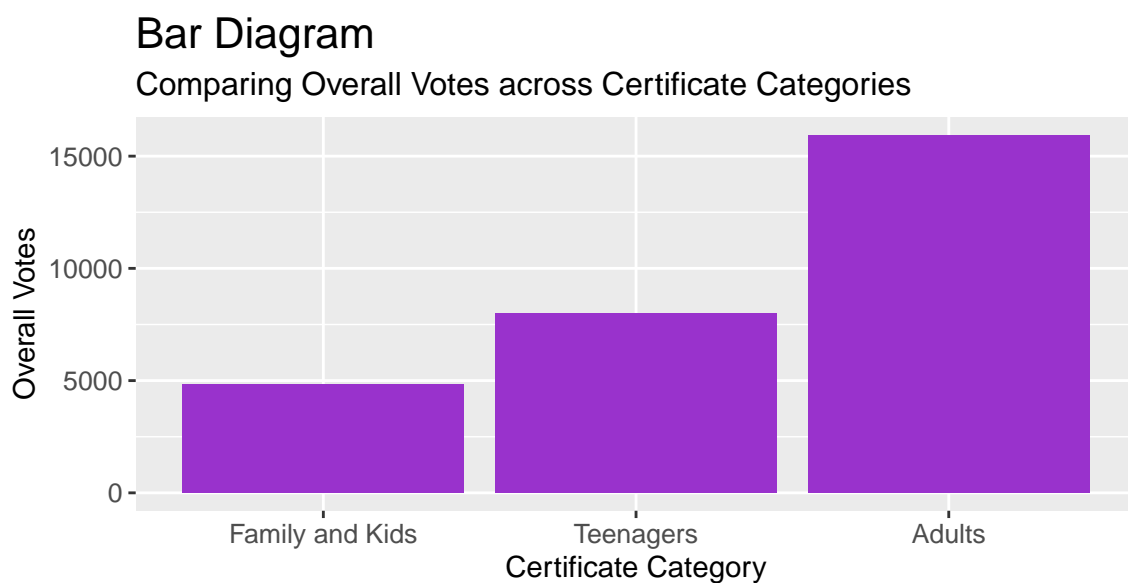
Some important visualizations

1. Most Voted Genre:



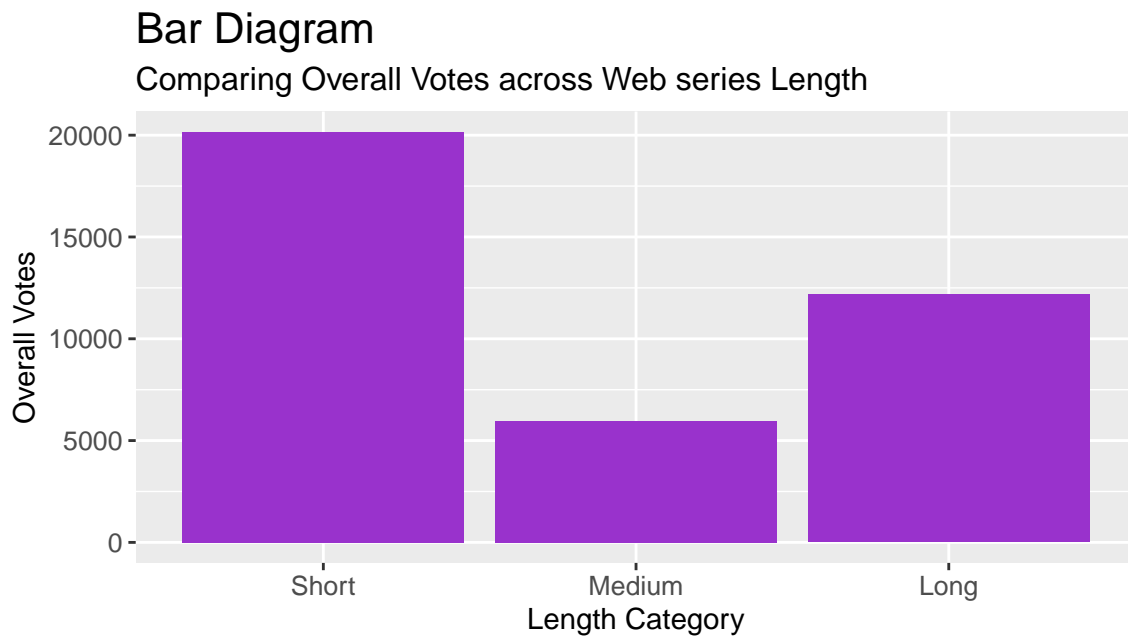
So, from the above plot, the genre with the highest number of votes is 'Drama'

2. Most voted Certificate category:



From the above plot, we get the Certificate category with the highest votes is the 'Adults' category.

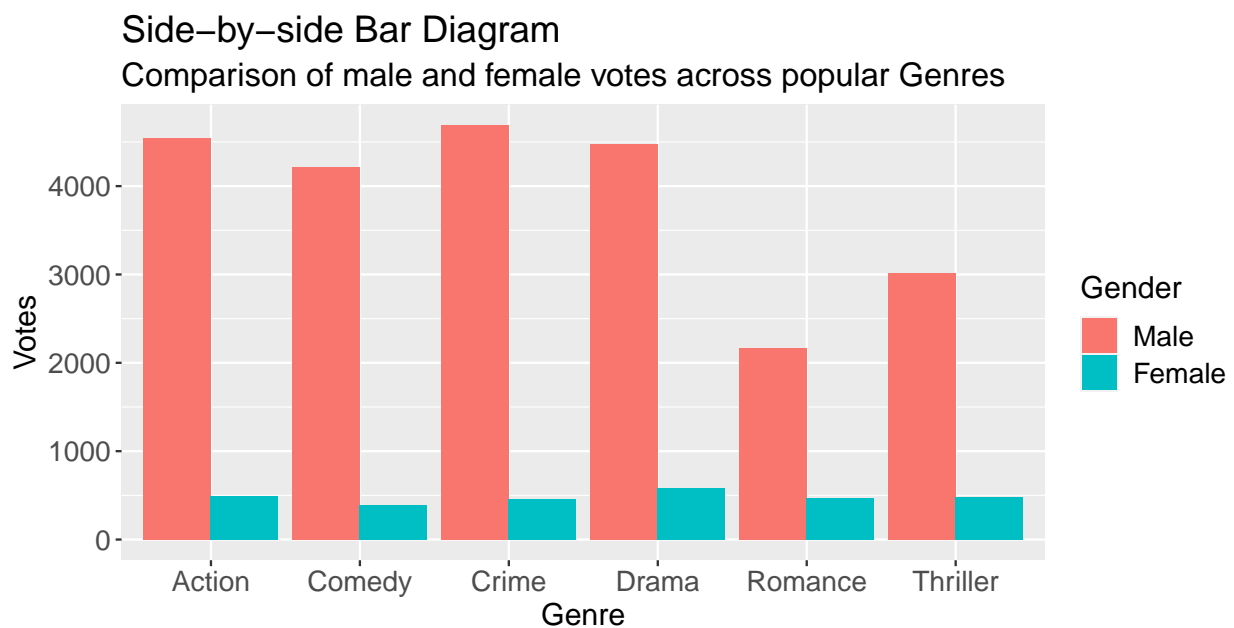
3. Most Voted Length category:



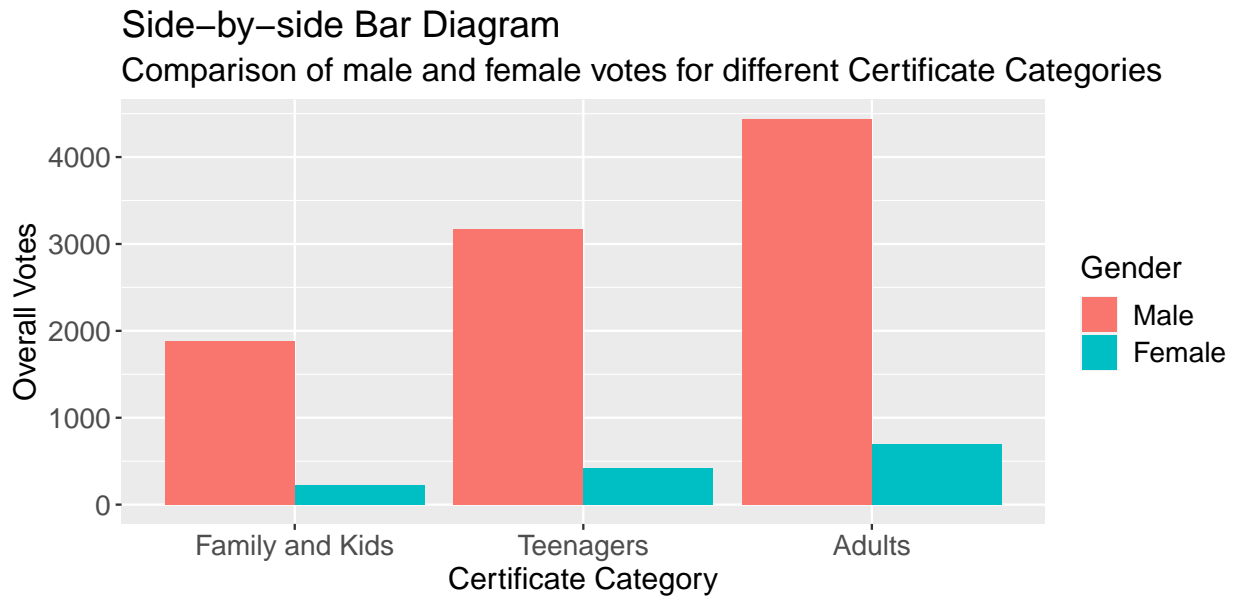
From the above plot, the most voted length category is the 'Short' length web series.

4. Gender classification of votes of web series:

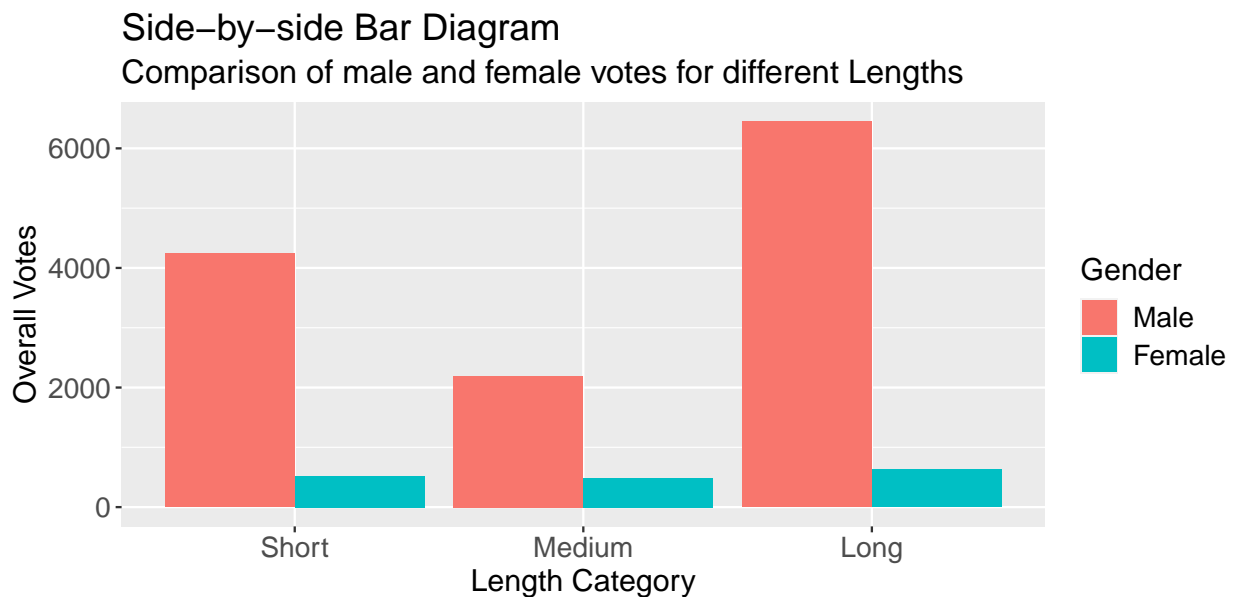
- Across different Genres:



- Across different Certificate Categories:

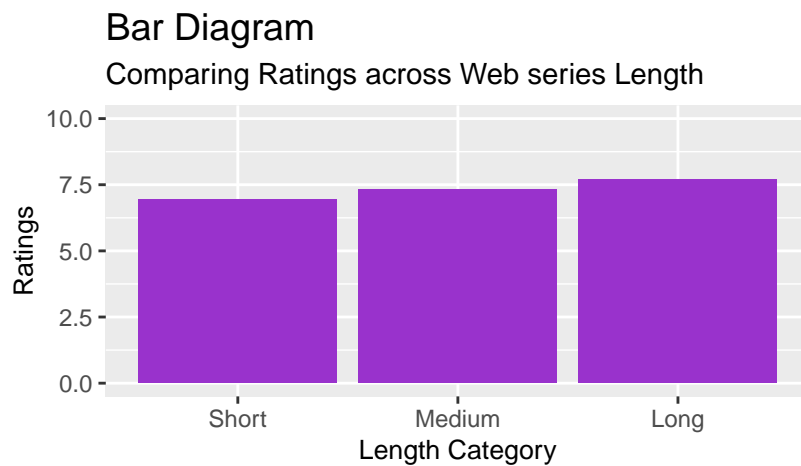
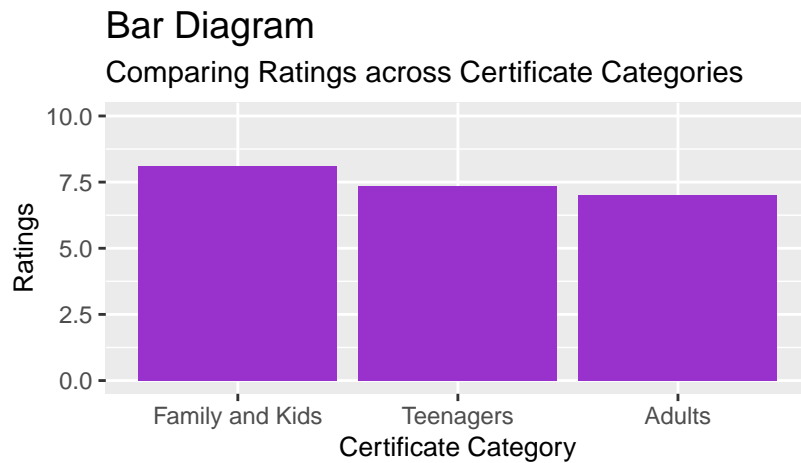
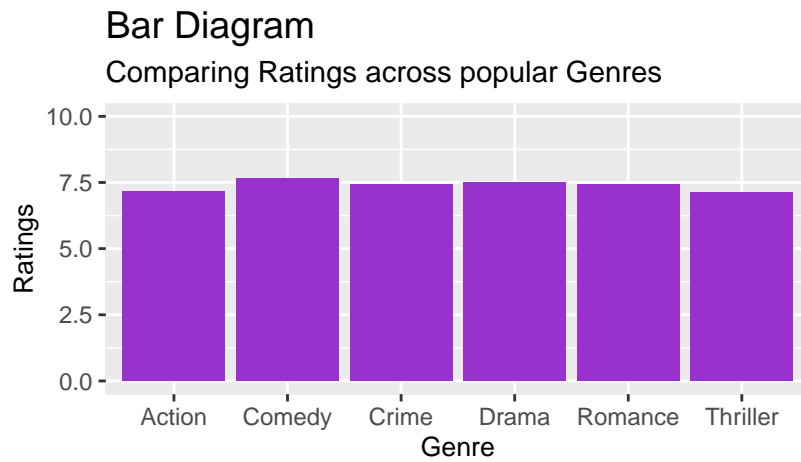


- Across different Length Categories:



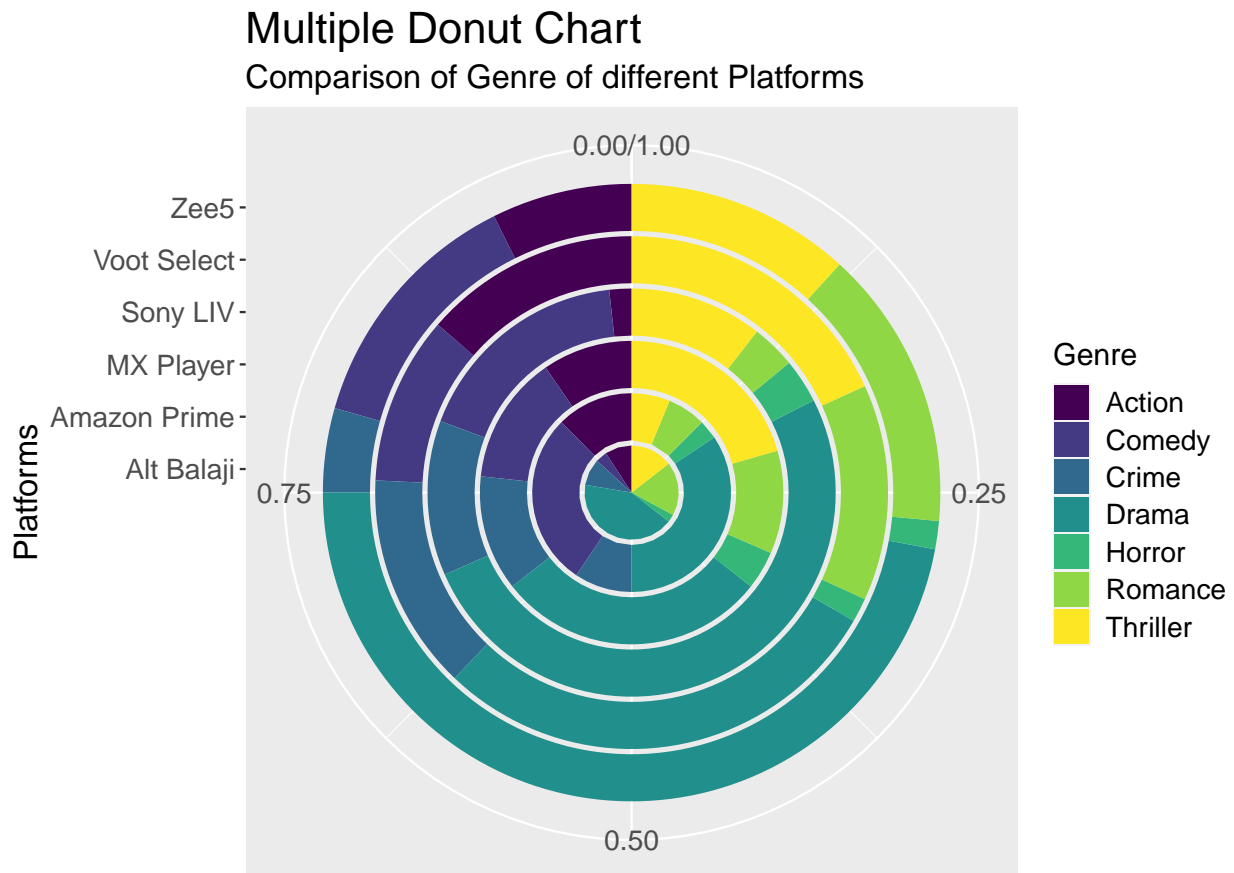
- From the above three plots, it can be easily seen that votes given by women are invariably less than that of men irrespective of comparison with respect to Genres or Certificate categories or Length categories.
- Also note that the plot of 'Genre vs Votes' (Plot 1) showed us the genre 'Drama' was most voted when overall votes were considered, whereas if we look at the Male and Female votes separately (in Plot 4(a)), then Male votes for 'Crime' and 'Action' dominates 'Drama'.

5. Most Rated Genre, Certificate category, and Length category:



We can see that there may be some variations of ratings of different categories in the above three plots but they are not significantly different.

6. Which genre is most popular on which platform?



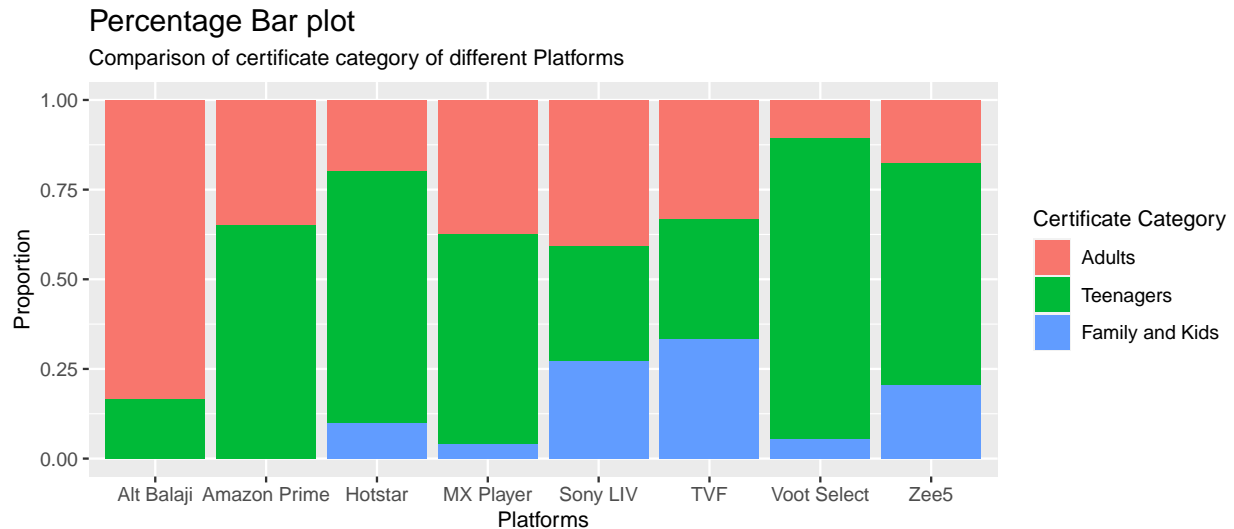
The above Multiple Donut Chart answers the question posed.

From the above chart, observations about the most popular genres are as follows :

- In each platform, the genre 'Drama' is predominantly significant.
- Voot Select mostly prefers 'Action' and 'Crime' whereas Amazon Prime prefers 'Action' and 'Comedy'.
- 'Horror' is invariably less preferred by all
- In Alt Balaji, 'Romance' is dominant whereas 'Thriller' dominates MX Player.

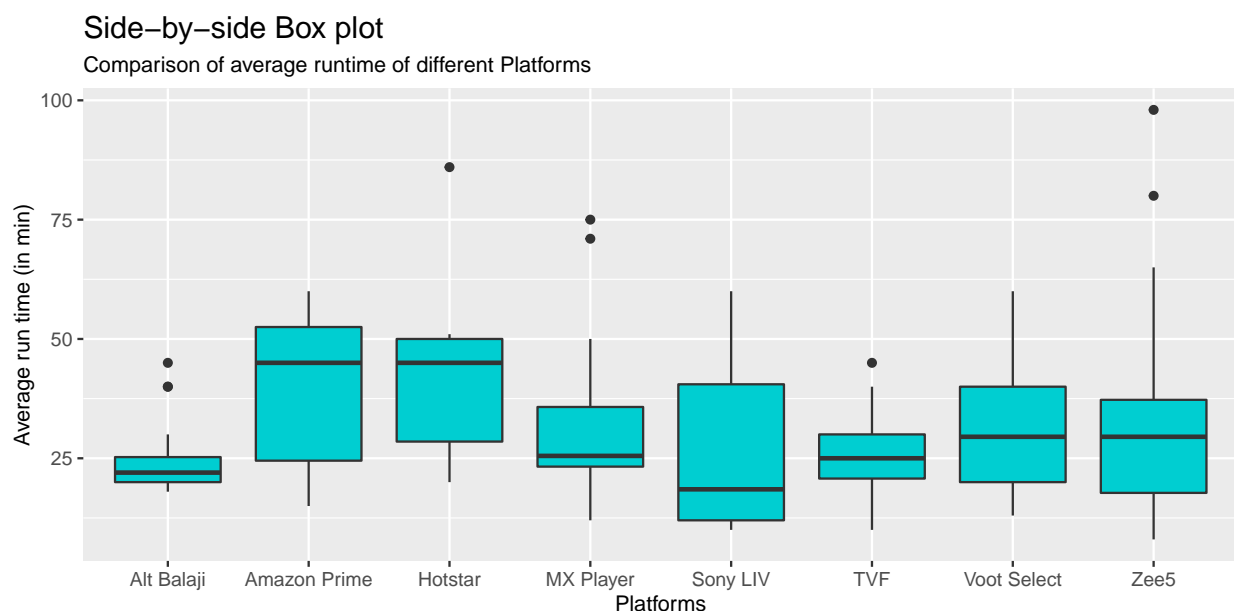
7. Which platform targets which age group?

The following Percentage Bar Plot shows what proportion of the web series in each platform belongs to the Adult category or Teenagers category or Family and Kids category-



- The key observations are as follows -
 - Sony LIV and TVF target all categories equally
 - Alt Balaji mainly targets the ‘Adult’ section whereas the rest of the platforms focus on the ‘Teenager’ section.

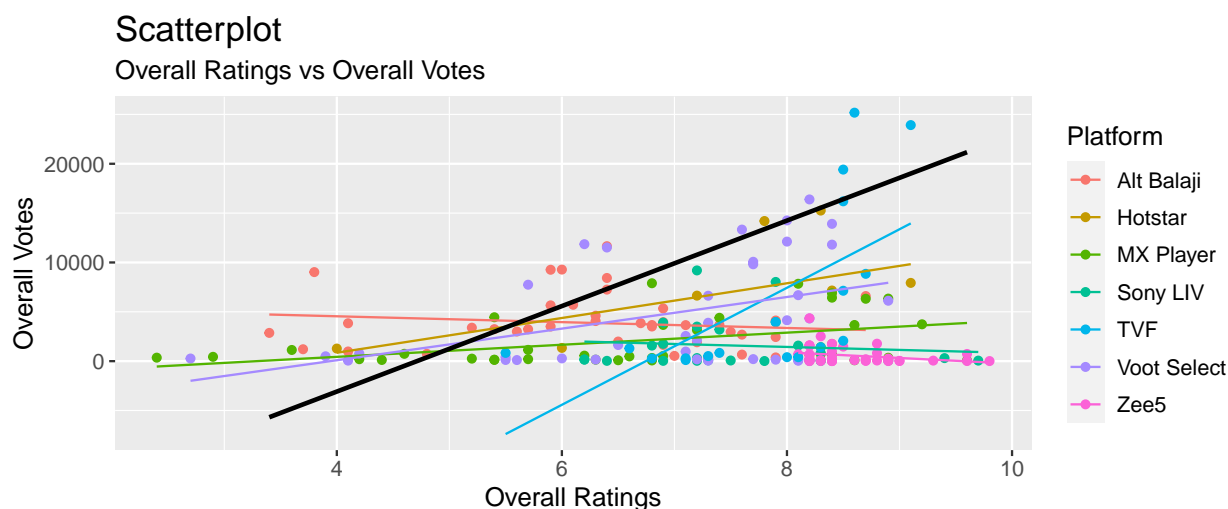
8. Comparison of average runtime of different platforms:



From the above plot, we have two highlights:

- Firstly, Platforms like Alt Balaji and TVF have the average run time of web series varying over a very small range whereas for MX Player, Voot Select, and Zee5 it varies moderately. And we see a large variation in the average run time of web series for Hotstar, Amazon Prime, and Sony LIV.
- Secondly, we have platforms like Alt Balaji and Sony LIV with median run time of less than 25 minutes whereas we also have platforms like Amazon Prime and Hotstar with median run time of 45 minutes. And for the rest of the platforms, the average run time lies between 25 minutes to 30 minutes.

9. Visualizing Votes-Ratings association:



Note: The Black line represents the overall regression line.

- The above plot shows how Ratings and votes are associated with each other. The black line is the overall regression line which shows a (weak) positive correlation between Votes and Ratings, i.e., in general, web series with high ratings have got more votes. The regression lines for individual platforms are also given in the plot which shows us the association between Votes and Ratings for individual platforms.
- An instance of Simpson's Paradox:
Simpson's Paradox appears when a trend appearing in different groups of data changes when the data is combined. This can be seen in the above scatter plot of 'Overall votes vs Overall ratings'. We can see, in most of the cases, the overall regression line is relatively steeper compared to the regression lines for the individual platforms. It is to be noted that for the platform TVF, the regression line is steeper than that of all platforms and the overall line too, whereas, for Alt Balaji, Zee5, and Sony LIV, the slope is even negative.

10. Does Ratings or Votes depend on Runtime?

Comparing Runtime with Overall Votes and Overall Rating



From the above Scatter Plot Matrix, we can see that on aggregate, there is very less dependency of Votes and Ratings on Run time. The correlation coefficients are negligible. Though some exception occurs in some platforms. For example, in TVF, Run time and Votes have a strong positive correlation or in Hotstar, Run time has a moderately positive correlation with Ratings.

Conclusion:

In this project, we have made an attempt to analyze various dimensions of Indian web series on different OTT platforms.

For this task, we extracted our data from IMDb and tried to make a few insightful plots. From these plots, we got empirical answers to some interesting questions we posed, viz. which genre is more popular among viewers, which platform targets which age group or which genre is more popular on them, how do male and female preferences differ with genres, etc. In addition to this, the shiny app we made provides a user-friendly graphical interface to enable us to select appropriate inputs and view the plots according to our interests.

With the growing popularity of web series in India, this kind of analysis on user liking and watching preferences would be of great help to the creators. It would help them to understand the choices of the audience they are catering to and make shows accordingly. The analysis is also of great importance to the OTT platforms as it helps them to understand their audience and make their future business strategies accordingly.

The other side of the coin is that the creators may be motivated to make similar kinds of shows which have already been successful in the past. This would hamper their independent thinking, resulting in a loss of creativity. Therefore, these conclusions must be used judiciously and wisely.

Besides, the visualizations made in this project work alone can not be used for drawing concrete conclusions. Further Statistical analysis with more relevant datasets needs to be performed to be able to infer something.

References

- IMDb: <https://www.imdb.com>
- Google: <https://www.google.com/>
- Wikipedia: <https://www.wikipedia.org/>
- GeeksforGeeks: <https://www.geeksforgeeks.org/>
- Shiny Gallery page: <https://shiny.rstudio.com/gallery/>
- MTH208A Course Webpage: <https://dvats.github.io/teaching/MTH208.html>
- Worksheets given in MTH208A

The above sources were of great help in completing this project work.

Acknowledgment

We would like to express our special thanks of gratitude to our instructor Dr. Dootika Vats for her ceaseless guidance throughout the work period and for providing such a golden opportunity to work on this project work.