Data Visualization Project Report

Home Streaming: Netflix, Disney+ and Prime Video, Where Should I Subscribe?

Pasquale Formicola
Angelo G. Limone

Abstract

The following Report will display all the procedures accomplished to analyze the patterns within home streaming platform data such as Netflix, Disney+, Amazon Prime Video and others.

The links to the infographics are the following:

- https://public.tableau.com/app/profile/angelo8884/viz/13febdatavizmovies2/Dashboard1?publish=yes
- https://public.tableau.com/app/profile/angelo8884/viz/TOP250IMDB/TOP250IMDB
- https://public.tableau.com/app/profile/angelo8884/viz/13febdatavizmovies3/DIFFERENCESAMONGPLATFORMS

Contents

1 Introduction

Home Streaming means all those services that offer, following a monthly or annual subscription, a very large catalog of Films and TV Shows. Over the years this type of service has become increasingly requested by users from all over the world who prefer to stay at home and keep all the great titles of the moment as well as the titles of the past present in the catalogs of the home streaming giants within reach.

Netflix, Disney +, Amazon Prime Video are certainly the best-known services in addition to Infinity and Apple TV.

Netflix is a US company active in the streaming distribution of movies, television series and other paid entertainment content via the Internet. It was founded by Reed Hastings and Marc Randolph on August 29, 1997, in Scotts Valley, California.

Disney+ is an American subscription video-on-demand over-the-top streaming service owned and operated by the Disney Entertainment division of The Walt Disney Company. Founded in 2019.

Amazon Prime Video is a US video-on-demand service owned by the Amazon.com company. Founded by Valeria Salazar in Seattle, 7 September 2006.

The home streaming market is growing steadily and has approximately 7 million new users every year. Each service offers its own catalog divided into Films or TV Shows in which it is possible not only to find current titles, but also films and TV series from the past. Thousands of titles are added each year through a contract with the platform.

2 Objective

Since there are many Home-Streaming services each having selected content for different types of users, which service should a novice user choose?

The purpose of this project is to analyze the following services and, by creating infographics and data visualizations, to provide a novice user with all the information necessary to be able to subscribe to a service.

For the development of this project, we put ourselves in the shoes of those who have never faced the decision to choose which Home Streaming service to subscribe to. Many people sign up for a subscription based on the opinion of their friends and acquaintances, but the user rarely investigates which service might be the best for them.

This is because there aren't many methods of comparison between the different Home Streaming services, so we decided to collect all the information provided by the different Netflix, Disney+ and Amazon Prime Video companies in order to create useful infographics for the user.

Providing it with numerous information that can direct users towards the choice of subscribing to a Home Streaming service rather than another.

3 Data Collection

This section aims to present all the datasets that were used to produce the infographics that will be further presented. Each subsection is dedicated to a specific set of data which source and dimensions are hand out.

Each dataset analyzed in this project is released by the same home streaming companies that make their public data available for analysis.

Each dataset includes information relating to the type of film, titles, cast, director and description, as well as the year of release and further information for each title.

An additional dataset containing the votes of IMDB users was used in order to be able to implement it to the other datasets present to have a further vote made by users.

3.1 Netflix Dataset

Available at this address: https://www.kaggle.com/datasets/shivamb/netflix-shows

- Show_id: Unique ID for every Movie/TV Show.
- Type: Identifier for Movie or TV Show.
- Title: Titole of the Movier/TV Show.
- Director: Director of the Movie.
- Cast: Actors involved in the Movie/Show.
- Country: Country where the Movie/Show was produced.
- Date_added: Date it was added on Netflix.
- Release_year: Actual Release Year of the Movie/Show.
- Rating: TV Rating of the Movie/Show.
- Duration: Total duration in minutes (Movies) or Number of Seasons (Shows)

3.2 Disney+ Dataset

Available at this address: https://www.kaggle.com/datasets/shivamb/disney-movies-and-tv-shows

- Show_id: Unique ID for every Movie/TV Show.
- Type: Identifier for Movie or TV Show.
- Title: Titole of the Movier/TV Show.
- Director: Director of the Movie.
- Cast: Actors involved in the Movie/Show.
- Country: Country where the Movie/Show was produced.
- Date_added: Date it was added on Disney+.
- Release_year: Actual Release Year of the Movie/Show.
- Rating: TV Rating of the Movie/Show.
- Duration: Total duration in minutes (Movies) or Number of Seasons (Shows)

3.3 Amazon Prime Video Dataset

Available at this address: https://www.kaggle.com/datasets/shivamb/amazon-prime-movies-and-tv-shows

- Show_id: Unique ID for every Movie/TV Show.
- Type: Identifier for Movie or TV Show.
- Title: Titole of the Movier/TV Show.
- Director: Director of the Movie.
- Cast: Actors involved in the Movie/Show.
- Country: Country where the Movie/Show was produced.
- Date_added: Date it was added on Amazon Prime Video.
- Release_year: Actual Release Year of the Movie/Show.
- Rating: TV Rating of the Movie/Show.
- Duration: Total duration in minutes (Movies) or Number of Seasons (Shows)

3.4 IMDB Dataset

Available at this address: https://www.kaggle.com/datasets/yehorkorzh/imdb-top-250-movies

- Title: Movie Title.
- Year: Year of film recording.
- Released: Release Date.
- Runtime: Movie length in minutes.
- Genre: Movie Genre.
- Director: Director's Name.
- Writer: Author of a script or book.
- Actors: Names of the main Actors.
- Plot: Movie plot.
- Language: Original Language of the Movie.
- Country: Country in which the film was shot.
- Awards: Received Awards.
- Metascore: Metascore of Movie.
- IMDBRating: Rating of the Movie from 1 to 10.
- IMDBVoters: Number of voters.
- IMDBID: Movie ID on IMDB.
- Type: Type of Art.
- BoxOffice: Movie Box Office.
- Production: Company who produced the Movie.
- Website: Movie Website.

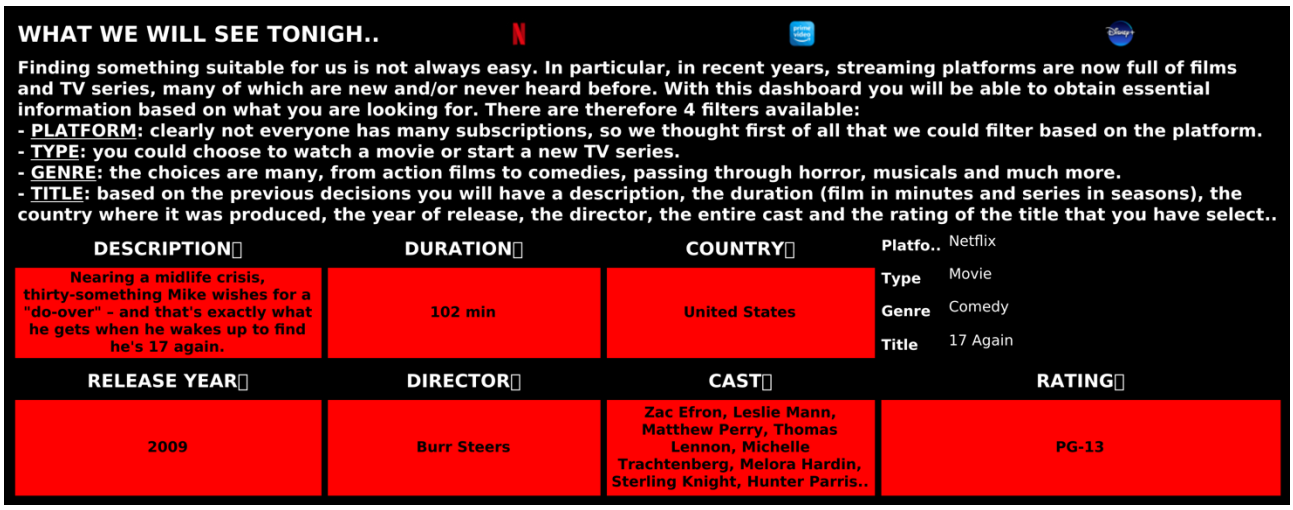
Another dataset containing the information of all 4 previous datasets has been created for ease of use and implementation of the Dashboards.

4 Infographics Accomplishment

This section is aimed to display and prove further information about the realized infographics, how they were built and what they are up to communicate. All the presented infographics are fulfilled with the use of Tableau Public.
As initially stated, the aim of our analysis is to highlight the different catalogues of main Home Streaming Platforms (Netflix, Disney+ and Prime Video) in order to give to users all the information they need to subscribe to a particular service. Before submitting tasks to our users, we believe it is necessary to provide some infographics that show how dashboard are made up.

4.1 Home Streaming Catalogues

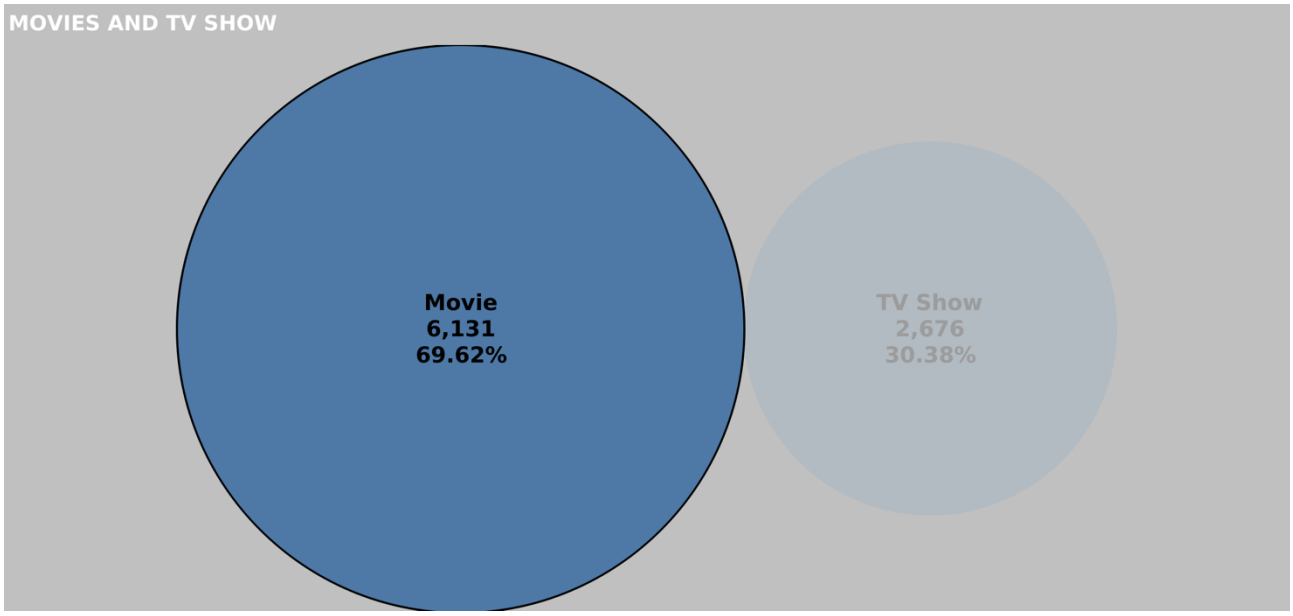
The first dashboard was generated by filtering by Platform, Type, Genre and Title. This very simple and intuitive dashboard allows the user to be able to select very specifically the main type of filters to obtain relevant titles. The main purpose of this dashboard is first to route the user to precise titles so that he can decide which title he wants to watch, providing him with all the information on the title's Description, Duration, Nationality and Release Year, the Director and the Cast and also the Rating



The example cited is 17 Again, Burr Steers, 2009, lasting 102 with all the information relating to the cast, in particular Zack Efron, Matthew Perry and Thomas Lennon.

4.2 TV Shows or Movies?

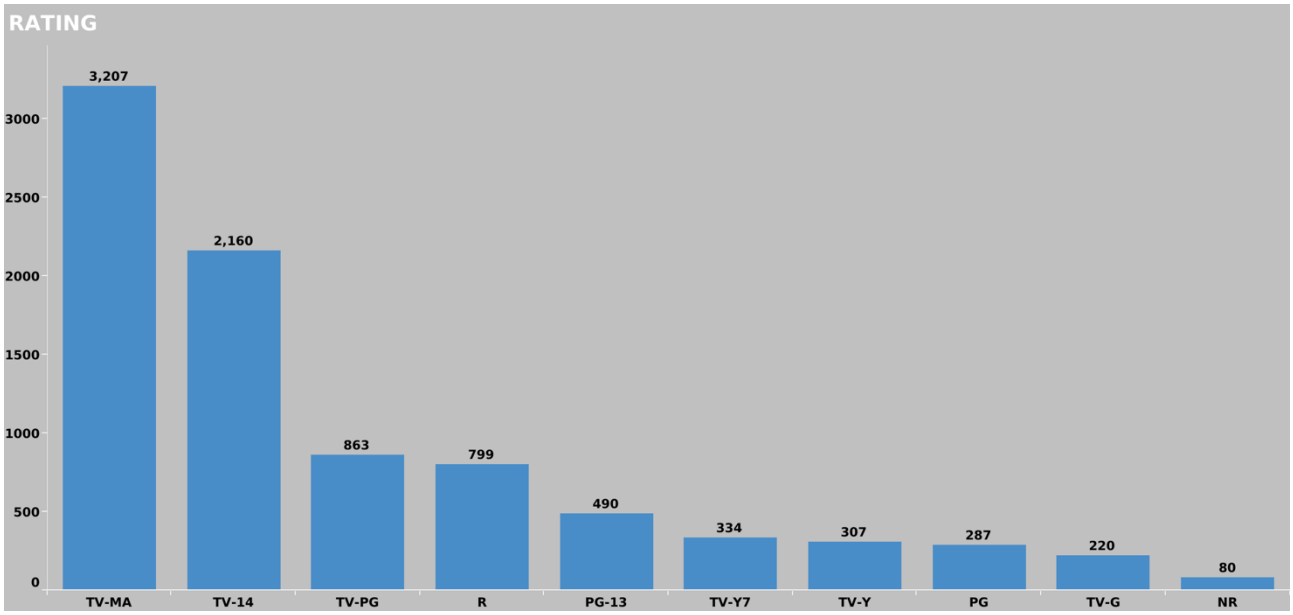
The graph implemented in the second dashboard shows the user the percentage and count of Films and TV Series highlighting the class with greater frequency.



The example shows the subdivision of the Netflix catalog in which we see the Films with 6131 Titles present in the catalog as protagonists, which correspond to 69.62% of the titles on the platform. A perfect indicator for those who prefer to watch a movie one night rather than a TV series over time.

4.3 What about the Movies?

In the second dashboard we have also implemented a graph capable of describing, depending on the type of filter applied, the percentages of the Ratings for each title belonging to that category of filters.



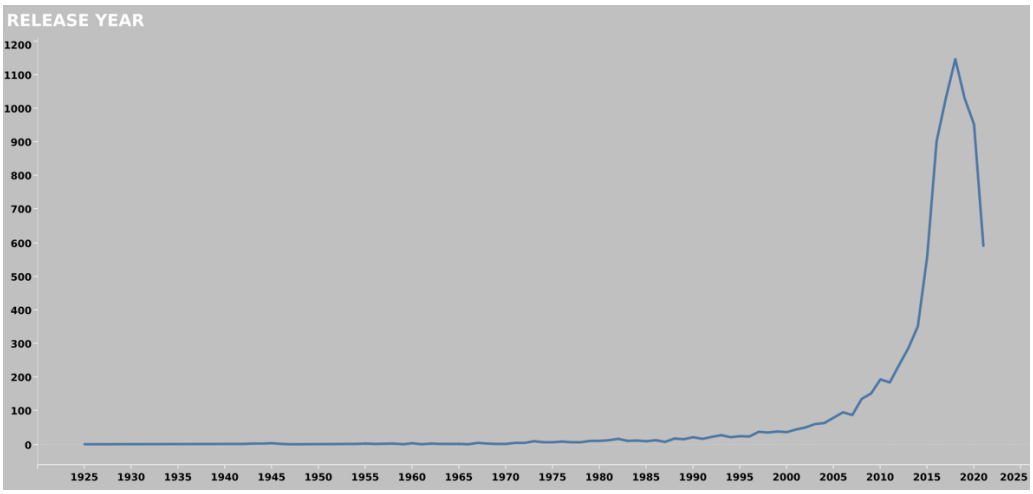
In the example of the Movies in the Netflix catalogue, we observe that most of the titles are classified as TV-MA, i.e., all that class of titles recommended for a mature audience. Below is a list of classifications:

- TV-MA: Recommended for an adult audience.
- TV-14: Not suitable for children under 14 years.
- TV-PG: Not suitable for younger children.
- R: Suitable for Boys.
- PG-13: Prohibited for children under 13 unaccompanied by an adult.
- TV-Y7: Suitable for children aged 7 and over.
- TV-Y: Appropriate for all children.
- TV-G: Appropriate for all ages.

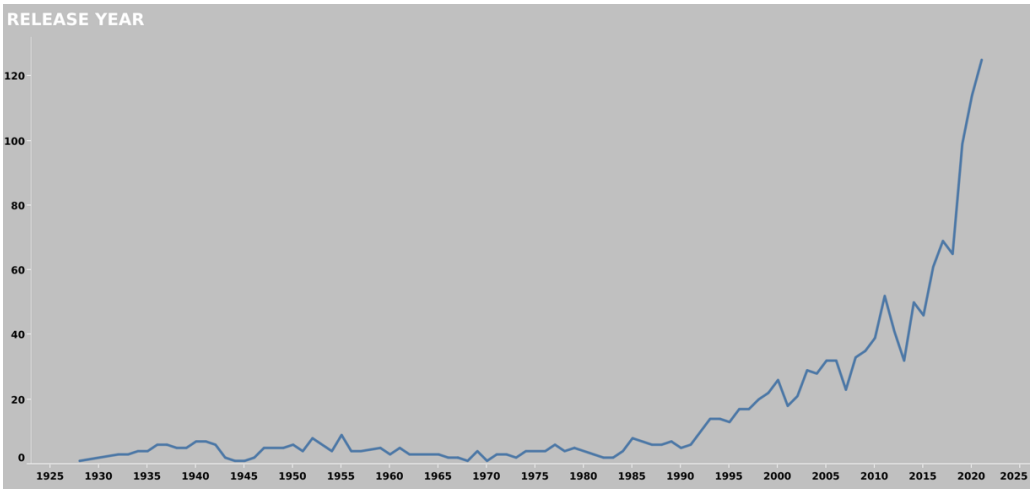
Another good indicator for parents to be able to subscribe to a specific service for their children.

4.4 Time-Traveler Watcher

Further useful information for those who, although they like to keep up with the latest titles of the moment, sometimes also like to dive into the past to watch films they have not seen for a long time or have never had the pleasure of watching, is precisely the display of titles released from 1930 to 2022, present on the chosen platform.



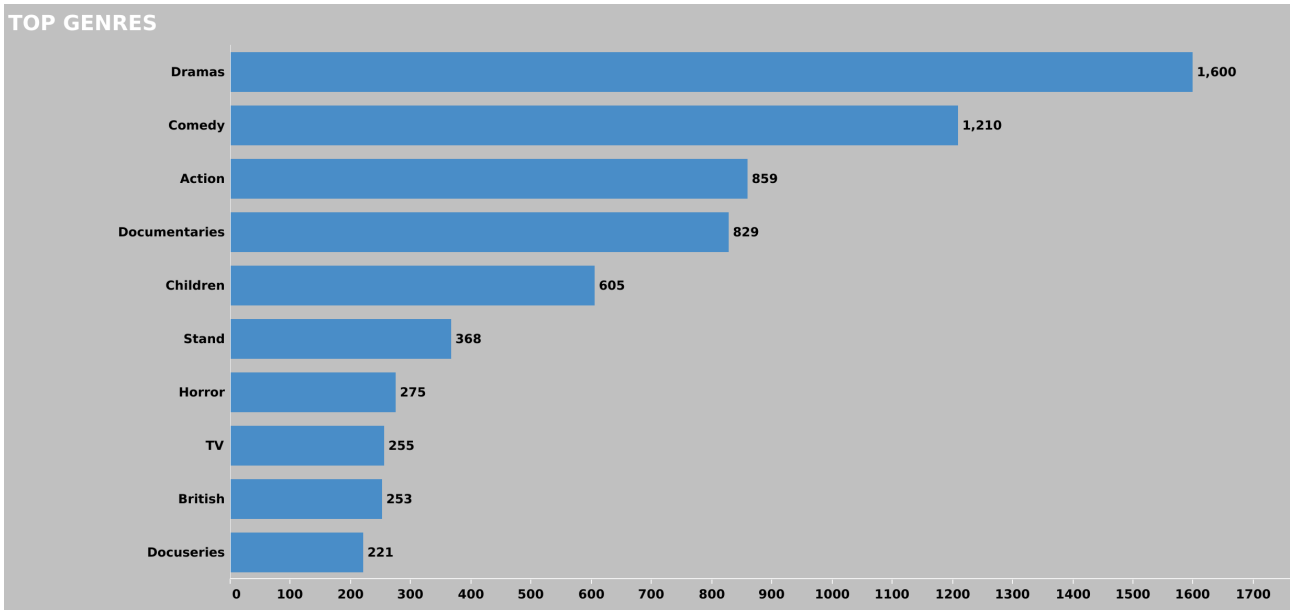
In the first graph we have selected as an example the time graph of the titles released and present in the Netflix catalogue.



In the second graph, however, we filtered the information for the Disney + catalog, thus finding the greater presence of older films in the catalog of this service. A good pointer for Time-Traveler Watchers. Although the Netflix catalog contains many more titles, Disney + offers more "seasoned" titles.

4.5 Fan-Genre

Many people who approach the world of home streaming generally have preferences with respect to the genre of the film or TV series, be it Animation, Action, Comedy, etc. A graph representing the presence with the proportions of the titles contained on each platform could help.

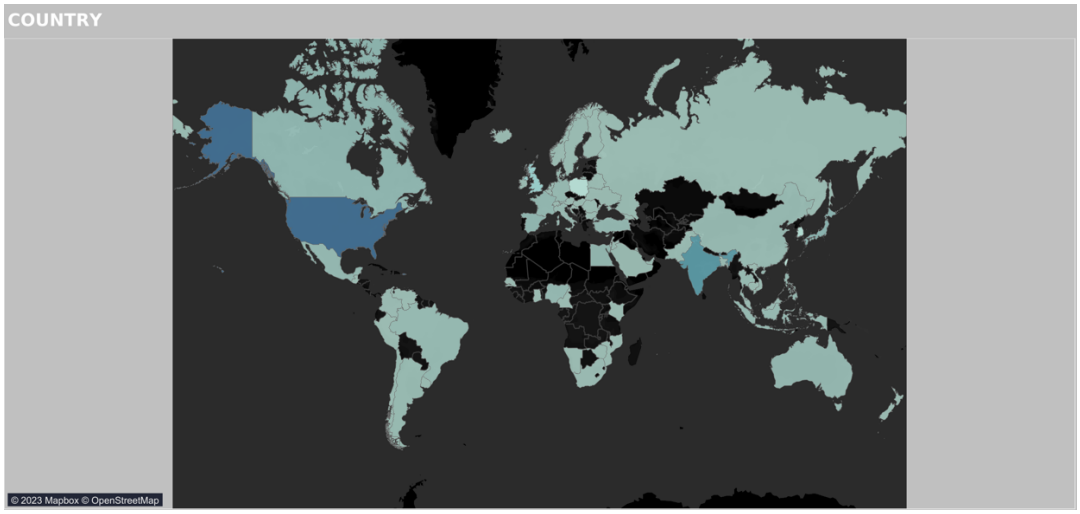


In the example in question there is the count of the Top Genre Titles present in the Netflix catalog with the respective count, on the side. It is therefore easy to understand, once the platform has been selected, which majority of genres can offer that specific Home Streaming service.

In the chart there is a majority of 1600 Drama Titles, followed by 1210 Comedy and 859 Action Titles. A big fan of Docuseries, for example, would not consider signing up for a Netflix subscription, rather take a look at other platforms first.

4.6 What If I Prefer Italian Landscapes?

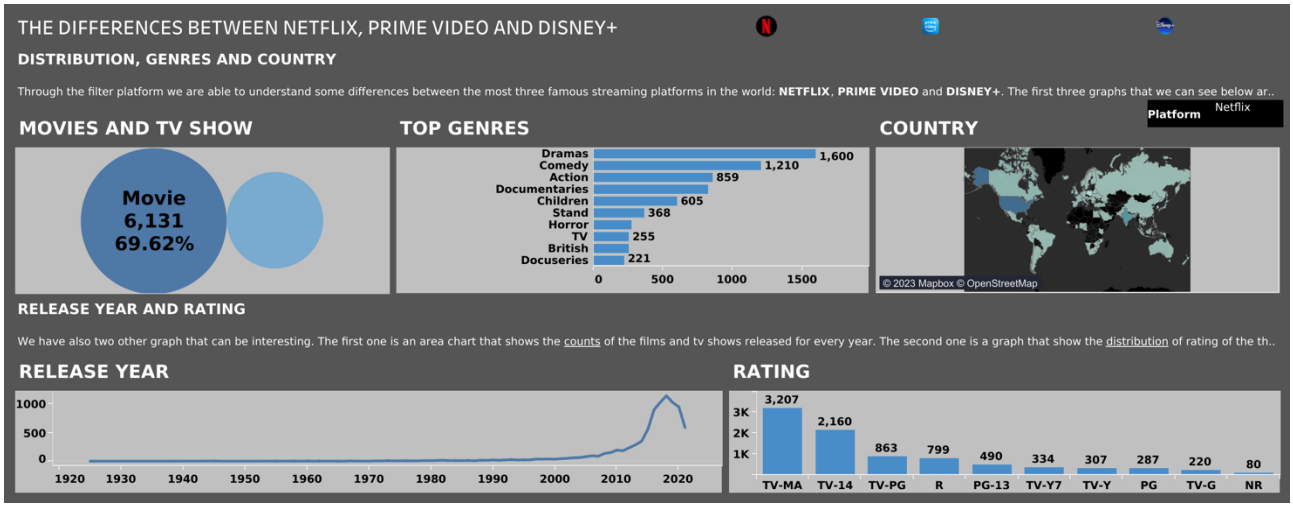
For users who intend to watch films produced in a particular country, we have implemented a global graph which shows, for each selected platform, the count of titles present on the platform for each nationality. The countries that have produced titles are highlighted by different colors, from the lightest (few titles) to the darkest (many titles).



This is an example taken by filtering for Netflix Titles.

4.7 First Overview

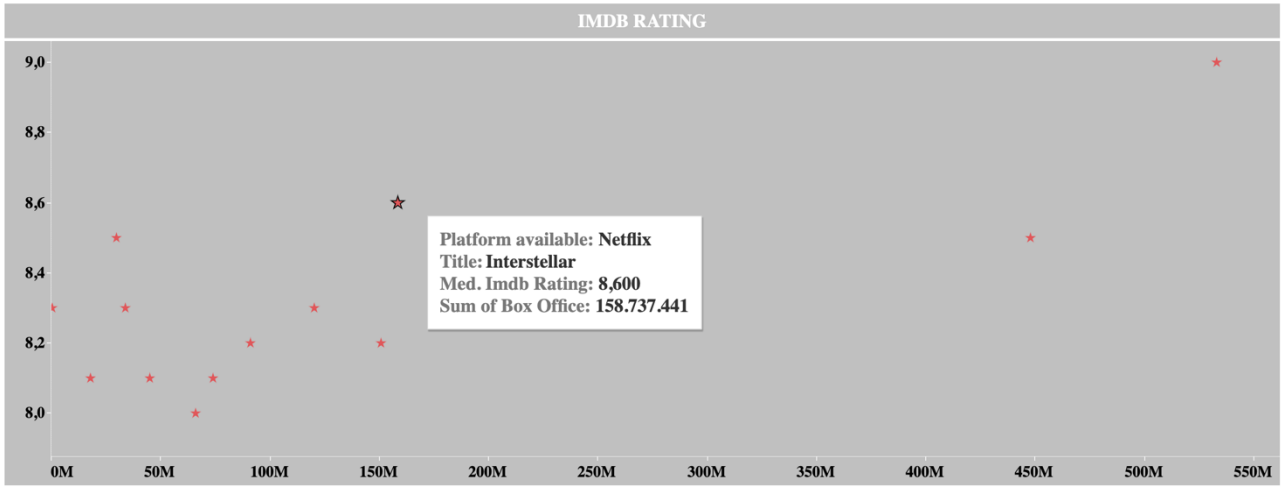
From the second dashboard, therefore, the user can understand which Home Streaming service is most suitable for him, evaluating various parameters such as: majority of Films or TV Series, Top Genres on the platform, Nationality of production of the title and year of release.



4.8 IMDB (Internet Movie Database)

We all know IMDB, one of the largest databases of Movies and TV Series on the Internet. Many users over the years have helped shape this community by creating the IMDB Rating, used to classify a title on a scale of 1 to 10 stars.

We have implemented the IMDB dataset to the datasets of the various platforms in order to be able to access the Rating of the selected Titles.



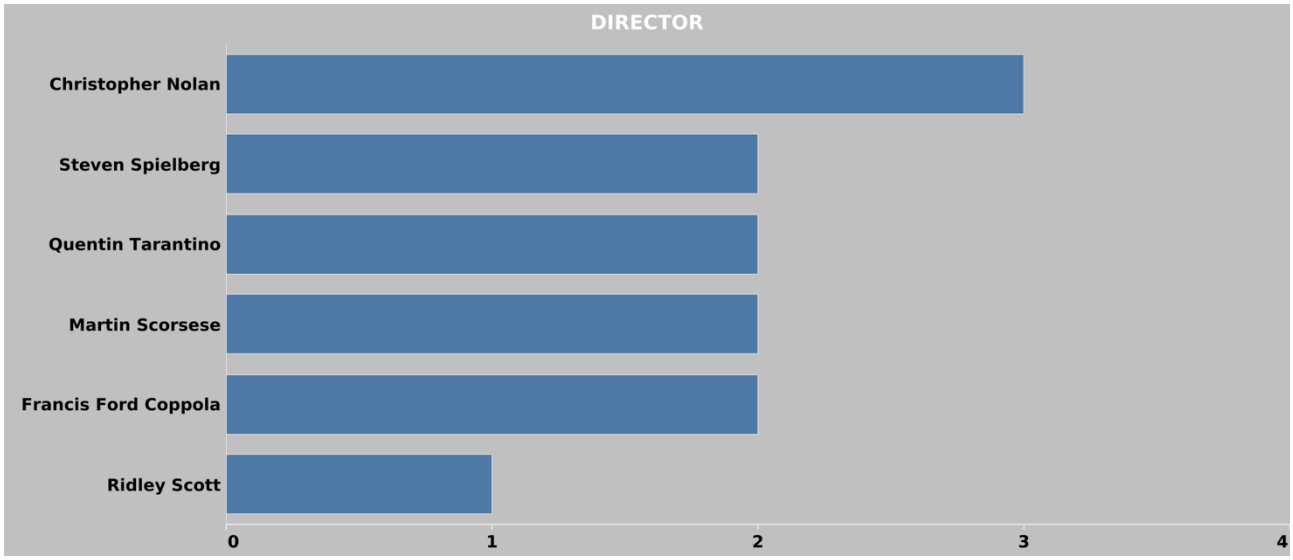
It is therefore possible to have a lot of information about the most voted titles. Just select a vote in the graph to have access to the Platform on which the title is present, to the Title itself and to the IMDB Rating.

In this example:

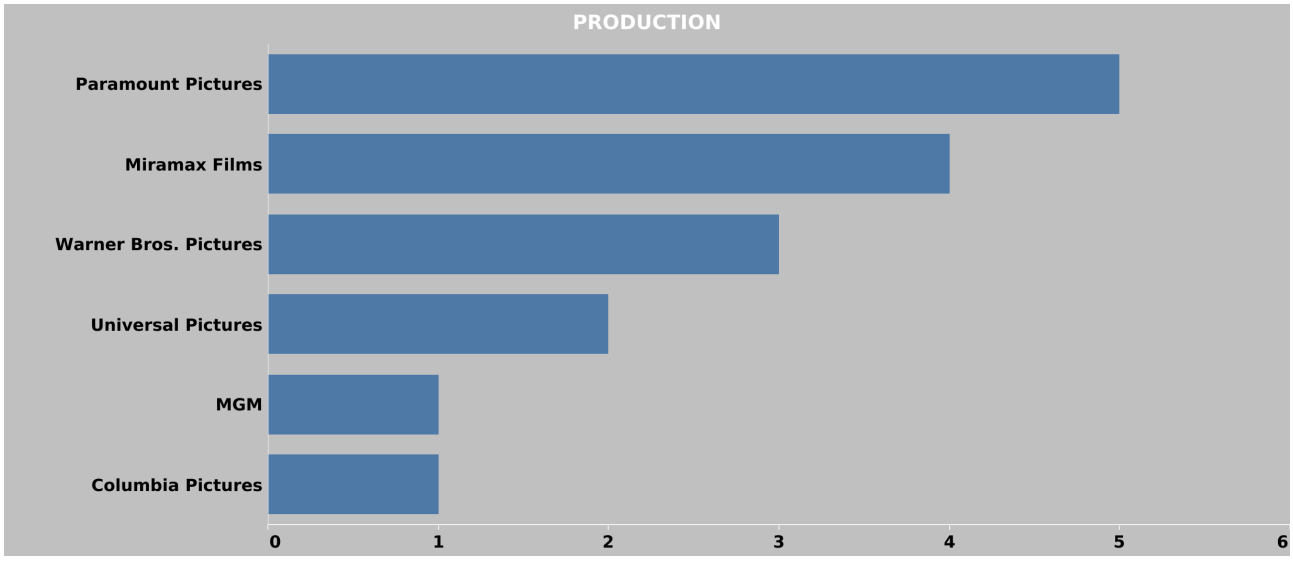
Interstellar is the second highest rated title having an IMDB rating of 8.6. A good indicator for those looking for more positively rated titles

4.9 Directors and Production

Each of us has preferences with respect to one director over another, and each of us has the image of the Metro-Goldwyn-Mayer lion impressed in our minds. It therefore seemed appropriate to provide a further indicator regarding this aspect.



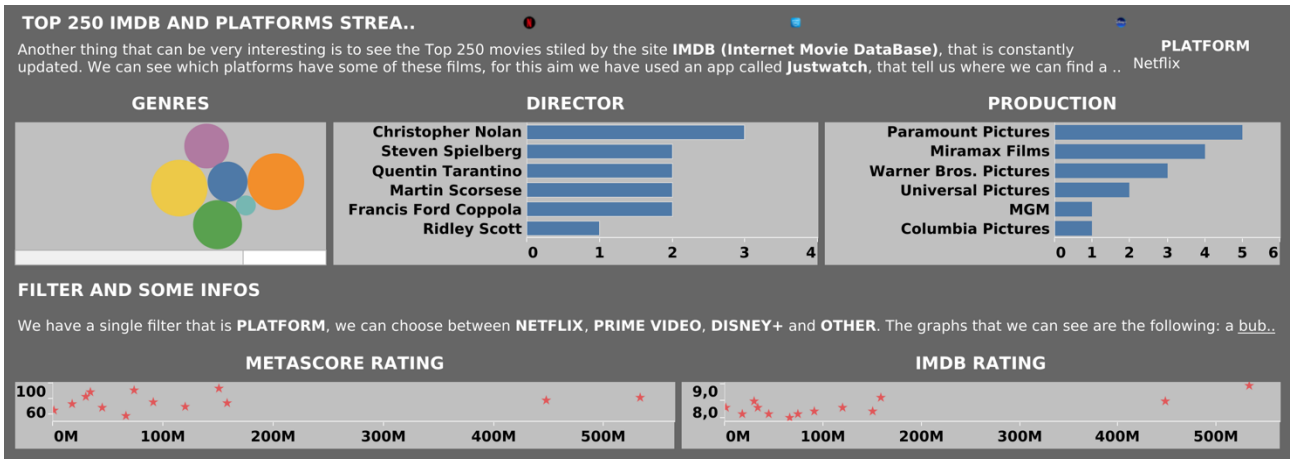
In this example, all Titles on Netflix belonging to the TOP-250 compiled by IMDB are filtered. Where we can find high caliber directors like Christopher Nolan and Steven Spielberg.



With the same filtering criterion, we instead represented the productions of those titles, finding Paramount Pictures, Warner Bros Pictures and Universal Pictures.

4.10 The Final Overview

By collecting the following information in a third dashboard we have therefore provided the user with a great many selection criteria. By providing interactive graphics, we were able to arouse their attention and give them a cue of curiosity and nostalgia for some titles.



5 Evaluation

To evaluate the quality of our visualization, we focused on three main points:

- **Heuristic Evaluation:** to identify the main problems of inaccuracies of our visualizations.

- **Tasks Execution:** to evaluate how well and how fast users understand the visualizations.

- **Psychometric Survey:** to evaluate how people felt about the intuitiveness, clearness and prettiness of the visualization.

5.1 Heuristic Evaluation

To identify problems or inaccuracies of the visualizations we let 8 users interact freely with them for some minutes and give feedback about their thoughts during the interaction phase, here is what came out:

5. 2 Tasks Execution

Once solved the inaccuracies of the visualizations, we proceeded with the user test, where we gave users some questions to which they had to respond verbally and measured both the execution time and the correctness of the answers to evaluate the efficiency of the graphs. Here are the identified tasks:

- Task 1: By selecting from the menus, would you be able to find a film that you would potentially watch because the description intrigues you?

- Task 2: By selecting any platform from the drop-down menu, do you think it is able to provide you with the information necessary to subscribe to that respective service?

- Task 3: Assuming you are a fan of older movies (under the year 2000) can you decide which service is best for you?

The tasks have been given remotely, with the user having the interactive dashboard already opened in a window on their devices, and the time has been measured from the time the question has been read to the time the user answered.

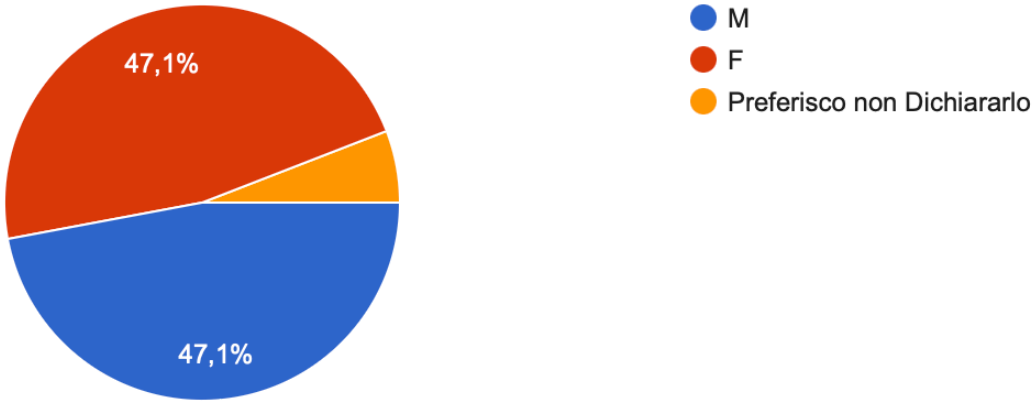
5.3 Psychometric Survey

In the making of the psychometric survey, to evaluate the quality of our infographics we decided to adopt a 5 points Likert Scale, that allowed us to measure:

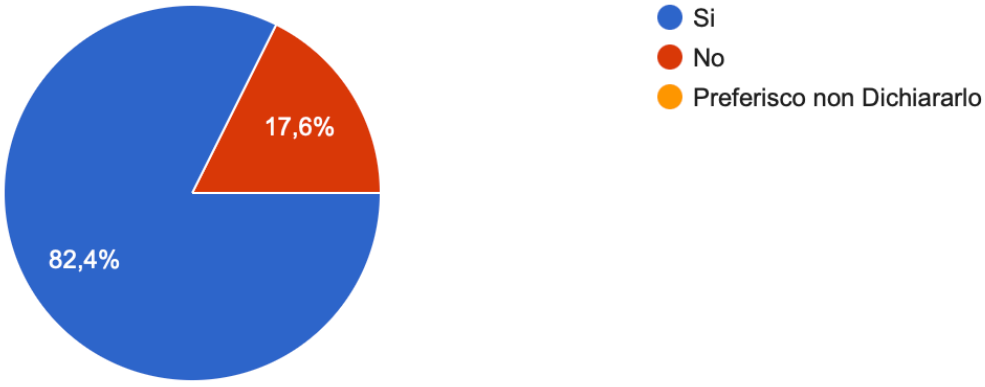
- How Clear the information was presented.
- How Intuitive the interactions with the visualizations were.
- How aesthetic the overall of the visualizations was.
- The overall quality of the visualizations.

5.4 Sample

A homogeneous sample was chosen to perform the tasks and answer the survey. It is made up of university students, personal friends, adults of both genders, some more familiar and some less familiar with interactive infographics.
The sample considered consists of 17 people. Gender and age proportions, as visualized in the following Pie Charts, show that males and females are perfectly balanced, and the age class is made up of people between 19 and 30 years old.



We also asked users if they were already subscribed to any type of service present in the search in order to make the usefulness of the visualizations effective. The majority are already subscribed to at least one of the services mentioned between Netflix, Prime Video and Disney+.



5.5 Tasks Results

The following tables represent the outcome of the tasks for each user. Each cell of the table is marked with a color according to the outcome of the task. The green indicates that the task has been executed properly, the yellow that the user has been assisted and the red indicates a wrong answer.

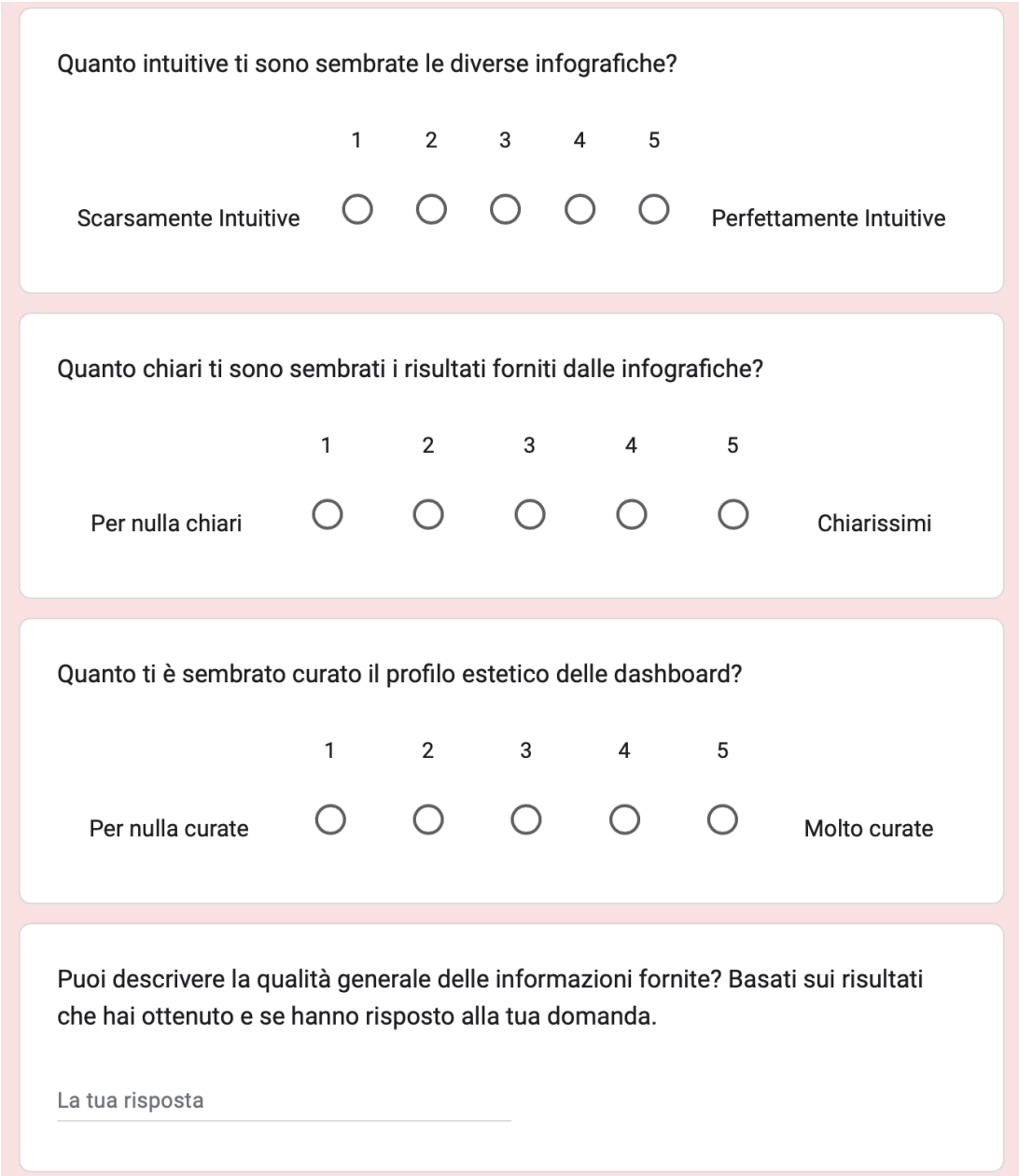
User	Task 1	Task 2	Task 3
User 1	8"	30"	31"
User 2	19"	21"	24"
User 3	16"	26"	37"
User 4	20"	26"	40"
User 5	31"	19"	20"
User 6	15"	26"	27"
User 7	17"	19"	31"
User 8	18"	32"	29"
User 9	37"	25"	26"
User 10	12"	29"	37"
User 11	17"	18"	27"
User 12	10"	19"	24"
User 13	15"	17"	36"
User 14	19"	27"	31"
User 15	11"	30"	24"
User 16	25"	15"	39"
User 17	14"	19"	26"

	Task 1	Task 2	Task 3	TOTAL
MEAN	17,88"	23,41"	29,94"	24"

Differently from what expected, the first task has been the one with the highest error rate while its average completion time is slightly more than one minute. The second task not only has been the one with the least error rate but also the quickest one to solve on average. The third, instead, due to a more complex infographic was the most time-taking for users to complete and, in some cases, users had to be helped in order to complete it.

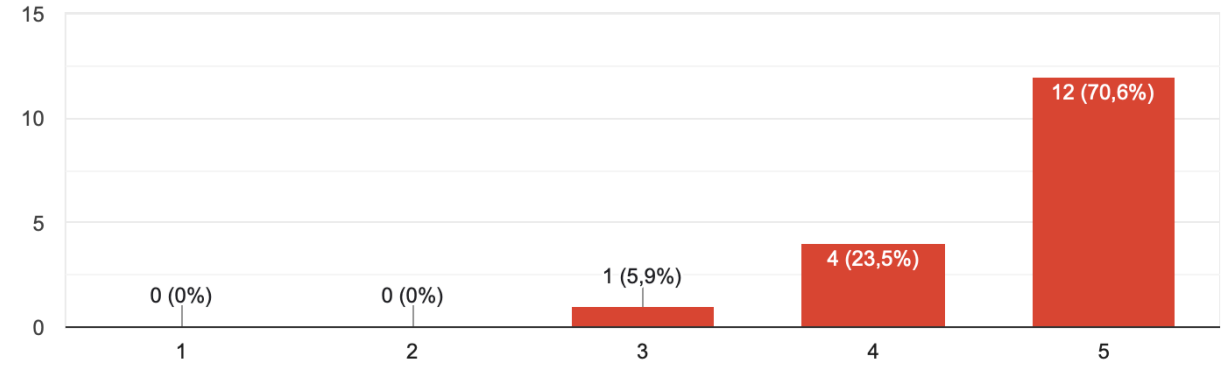
5.6 Survey Results

Here is an overview of what the psychometric survey looked like:

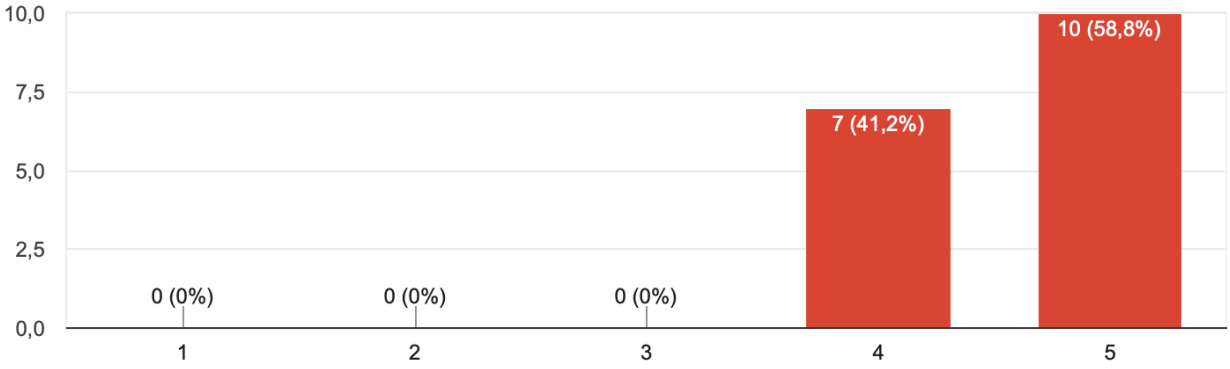


The survey was submitted to 20 users. Survey results are reported in the following histograms of rating.

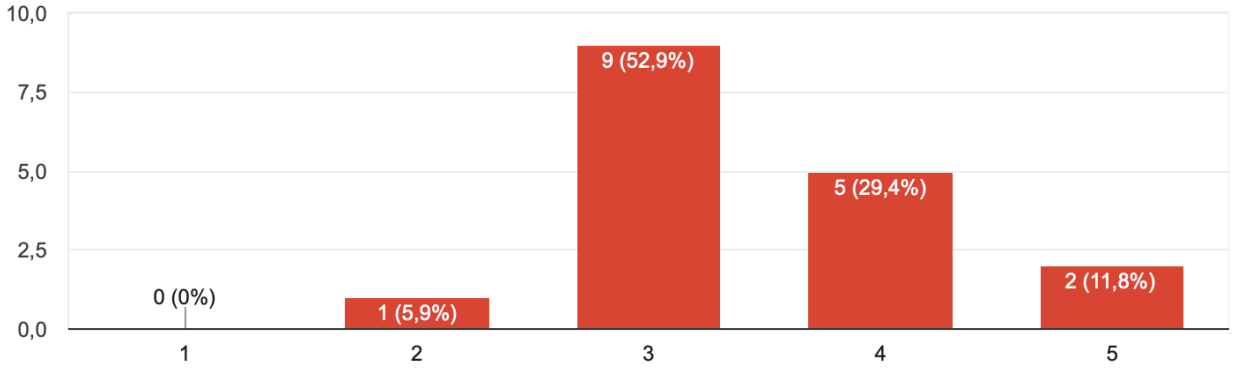
- How intuitive have the dashboards proved to be?



- How clear did the results provided seem to you?



- How accurate did the aesthetic profile seem to you?



The results are satisfying for what concerns intuitiveness and clearness as their average spaces towards the mean of 4 and 5. The registered average score for prettiness is lower with respect to the other parameters.

6 Final Changes and Conclusions

After analyzing the results of the tasks, we decided to commit some final changes to the visualizations in order to try to reduce the error rate an making easier the comprehension of the information we wanted to convey, in particular:

- The possibility, at the request of users, to be able to filter the titles according to the duration in order to have a reference based on the time available to watch a title.

- The possibility to filter catalog titles by year.

- The possibility to improve the Aesthetic Profile through the implementation of Services' Images.

The provided infographics have been generally understood and appreciated by all users. According to the survey, further improvements should be made on their prettiness.

We were very satisfied with the opinion of the people, especially of users who have no subscriptions to any of these Home Streaming services.

According to some of them, in fact, this project has concretely helped to be able to choose between the different services since

"You can't find such an intuitive and fun comparison method between these services on the internet."

Overall, we can say that we are satisfied with the work done, it has been very interesting we discovered the "Tableau" approach and we had a clearer idea of what is the art of visualization and teamwork and interaction between developer and user that is behind it.