## **Aims and Objectives**

### **Aims**

Analyzation of all the information about the movies found on different streaming platforms like Netflix, Prime Video, Hulu and, Disney+ produced in different years was the main aim of this report.

### **Objectives**

* Search and get the required dataset
* Analyzation of the data from the dataset
* Visualization of the data from the dataset
* Results findings and Conclusion
* Improvements and Future works

## **Background**

The CSV file contains about 16,744 movies released from 1902 to 2020 that are available on different streaming platforms which include Netflix, Prime Video, Hulu and, Disney+. The first and, second column of the dataset gives indexing and, id for each movie. The dataset contains the year that the movie was produced, the required age to watch it, its rating on IMDB and, Rotten Tomatoes, its availability on different streaming platforms, its languages that it was dobbed in, its genres and, directors, the countries it was released in, and, it’s length.

## **Acquiring the data**

The dataset was taken from the Kaggle website for the analysis purpose of the movie’s data in different ways. Kaggle is a Data Science Company. Kaggle provides an online crowd-sourced platform for machine learning, data science for practitioners, for problem-solving and analytic purposes (Usmani, 2017).

Dataset link: <https://www.kaggle.com/ruchi798/movies-on-netflix-prime-video-hulu-and-disney>

## **Description of attributes**

The CSV file contains 16,745 rows and 17 columns.

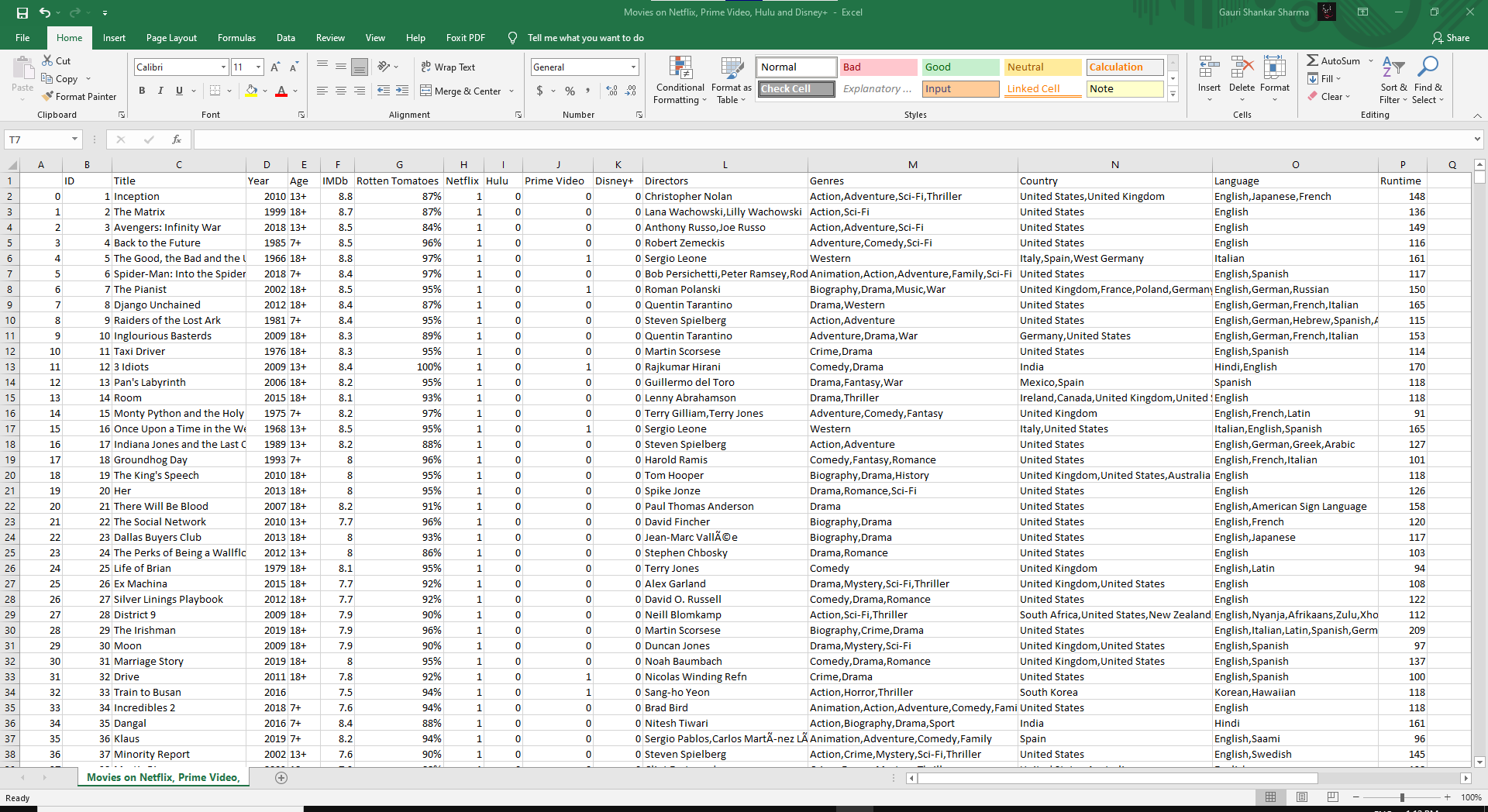


Figure 1: Movies available on different streaming platforms

A bit unclear column is explained below.

Table 1: Columns Details

|  |  |
| --- | --- |
| **Colum Names** | **Meaning** |
| Null | Indexing |
| IMDB, Rotten Tomatoes | Rating Status |
| Netflix, Hulu, Prime Video, and Disney+ | Movie availability on the respective platforms (1=available, 0=not available) |
| Runtime | Movie’s length in minutes |

## **Analysis and visualization**

All the data from the dataset were analyzed from various perspectives and visualize with the help of Tableau software. Tableau is a popular, easy-to-use, and powerful data analytic and, visualization software trusted and used by many businesses intelligence industry like Amazon (IntelliPaat, 2014).

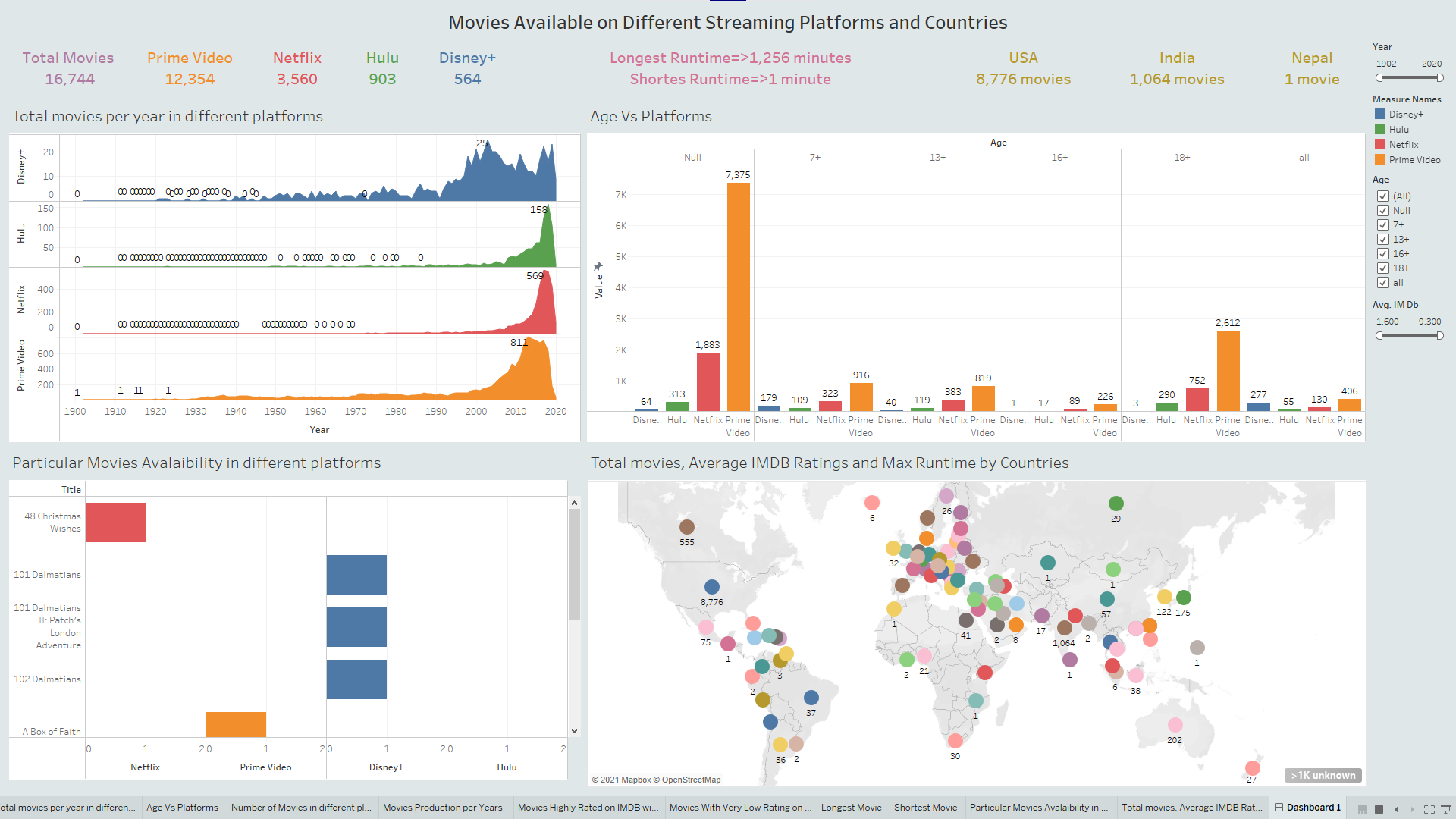


Figure 2: Dashboard of data visualization of movies found on Netflix, Prime Video, Hulu, and Disney+ from

### **Movies with platforms and produced years**

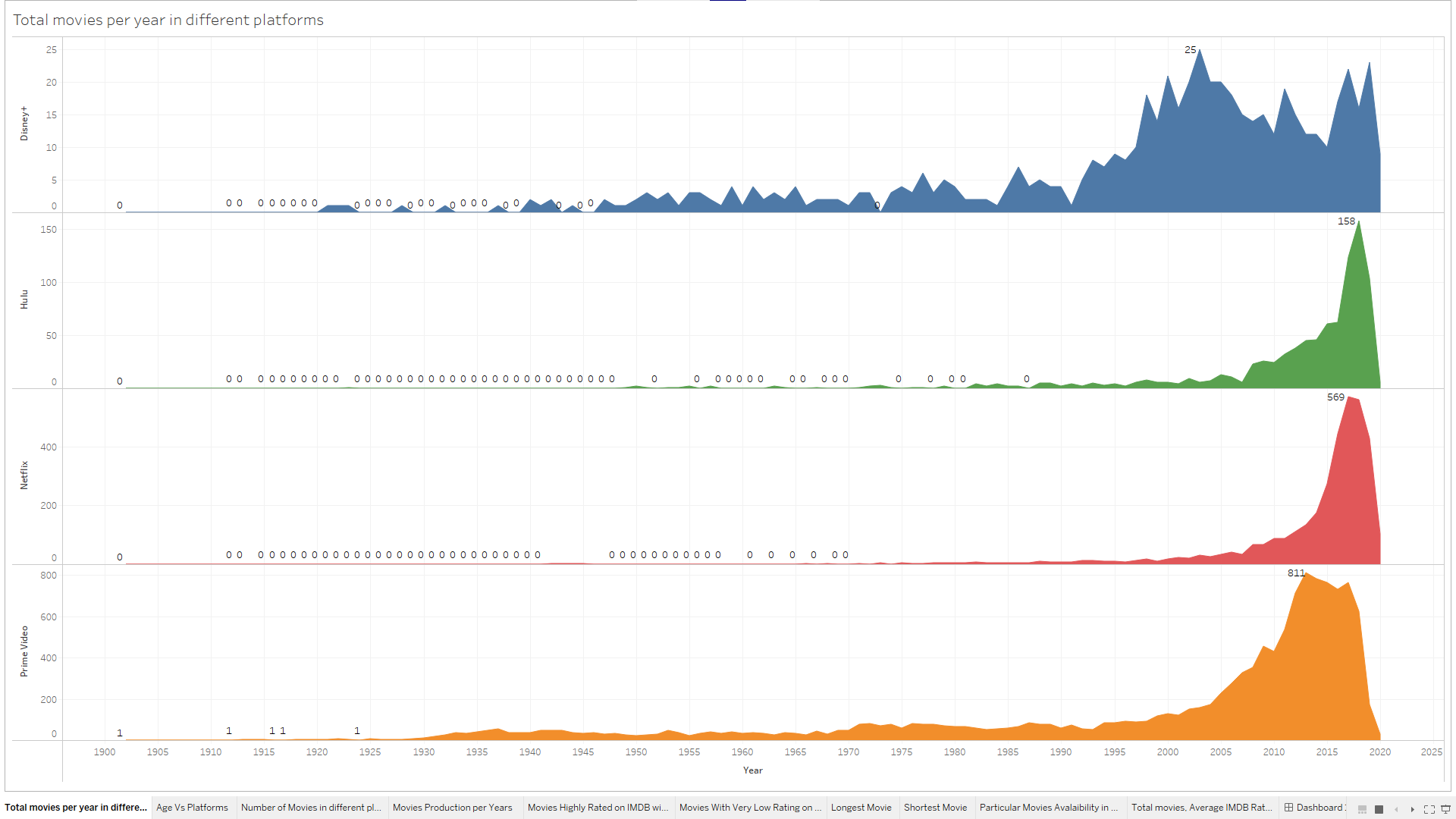


Figure 3: Number of movies on different platforms produced in different years

Figure 3 shows the number of movies available on different platforms that were produced from 1902 to 2020. Prime Video has a total of 811 movies that were produced in 2013 which is higher than the number of movies found on other platforms that were produced in the same year. Netflix has a total of 569 movies that were produced in 2017 which is lower than Prime Video but it is the highest number of movies found on Netflix that were produced in the same year. Similarly, Hulu has 158 movies in 2018 which are greater than Hulu but lower than Netflix and Prime Video, Disney+ has 25 movies in 2003 which are greater than Hulu but lower than Netflix and Prime Video.

### **Age Vs Platforms**

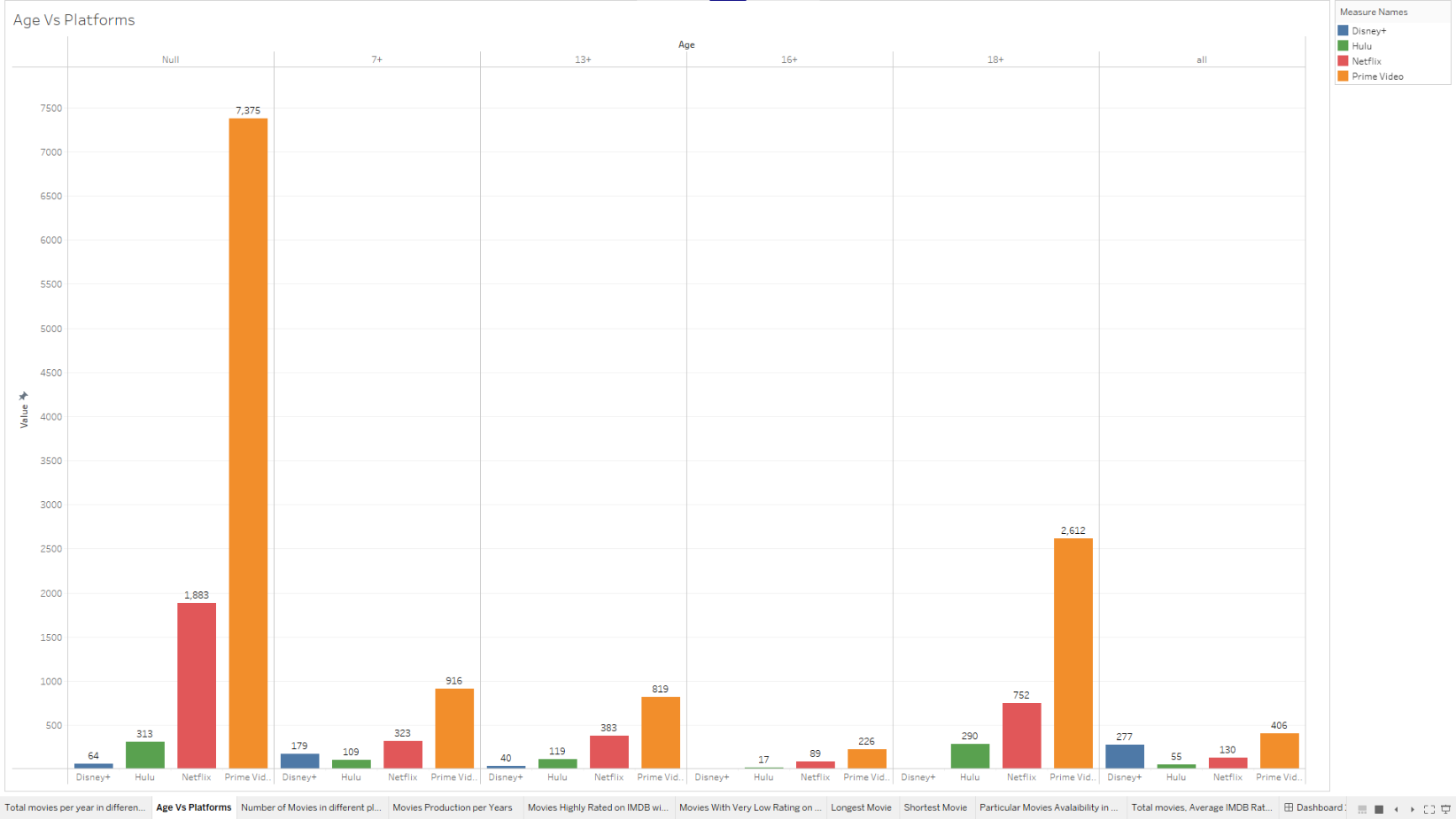


Figure 4: Number of movies with age target available on different platforms

Figure 4 shows total movies found on different platforms according to the targeted age group which are 7+, 13+, 16+, 18+ and all. There is also a total of 9,635 null values on the age column which is higher than other age groups. The 18+ age group has 3,657 movies which are higher than other age groups. Similarly, the 7+ age group has 1527 movies, the 13+ age group has 1,361 movies, and the 16+ age group has 333 movies. 868 movies are watchable for all age groups.

### **Total movies with platforms**

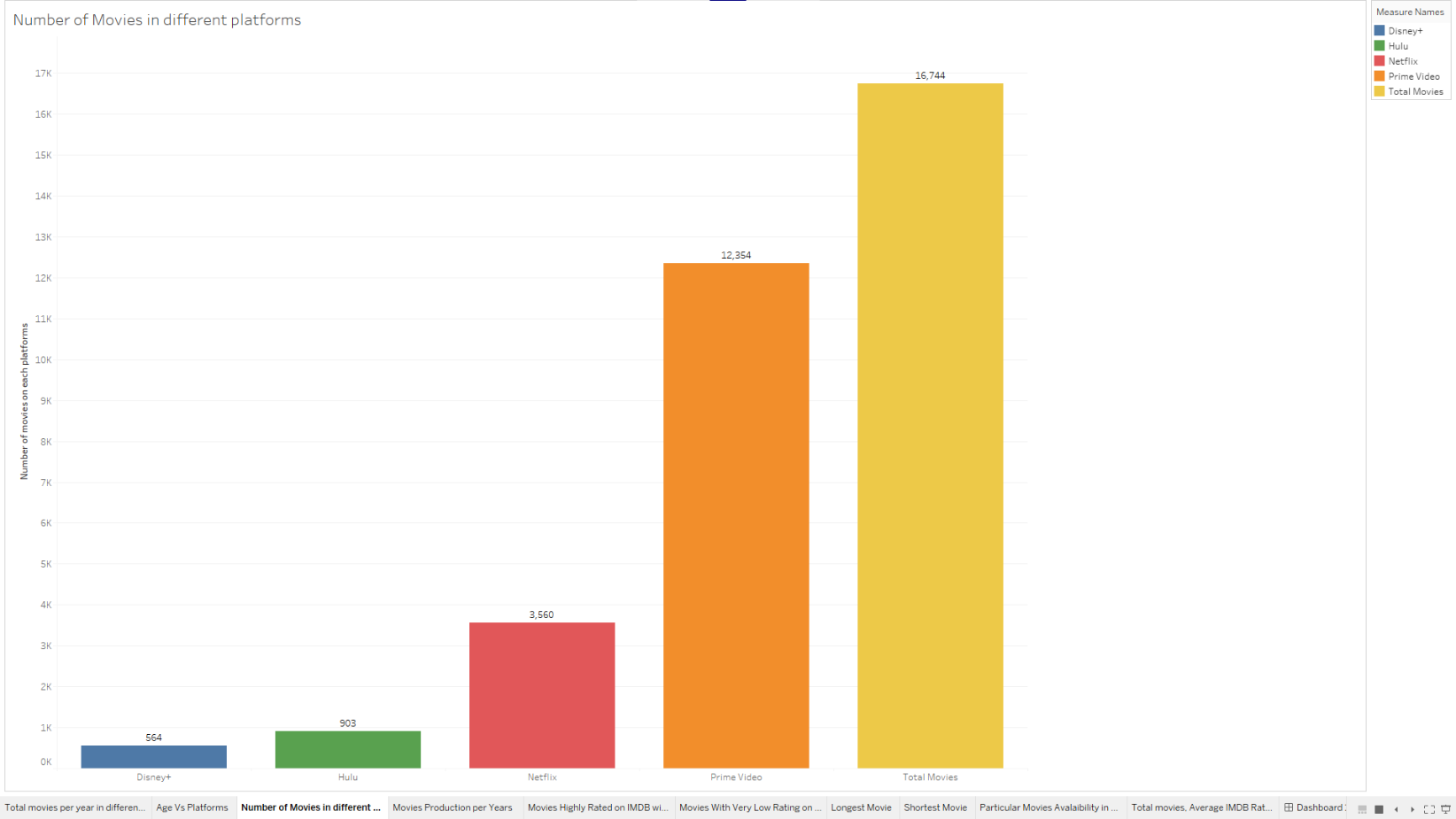


Figure 5: Total movies available on different platforms

Figure 5 shows that among 16,744 movies, Prime video has the highest number of movies available on it which is 12,354 and Disney+ has the lowest number of movies available on it which is 564. Similarly, Netflix has a total of 3,560 movies and Hulu has a total of 903 movies.

### **Movies Production per Years**

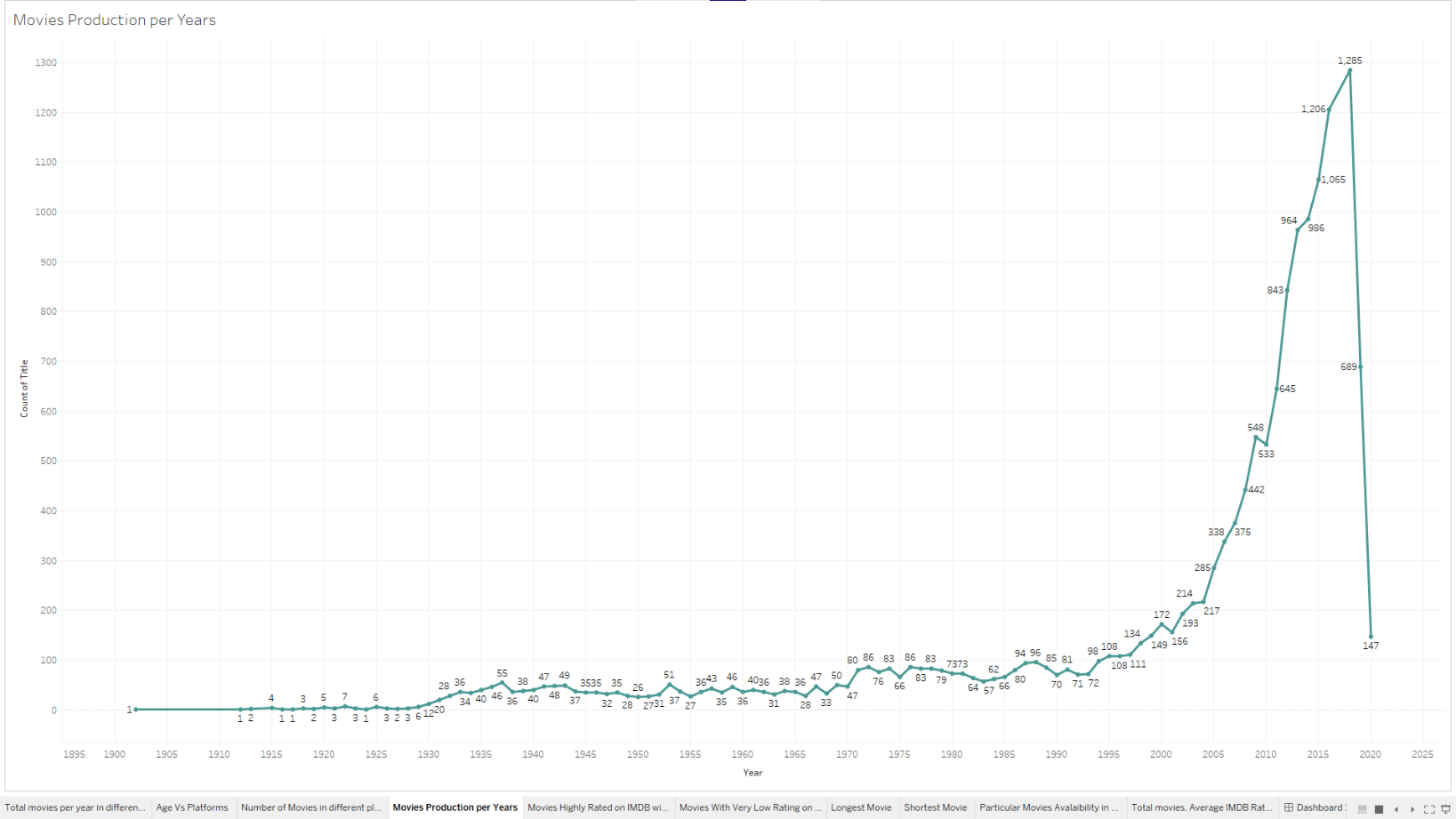


Figure 6: Number of movies produced per year

Figure 6 shows the number of movies produced from 1902 to 2020 which are found on Netflix, Prime Video, Hulu, and Disney+. 1902, 1912, 1916, 1917, 1924 was the year in which only one movie was produced but 2018 was the year in which 1,285 movies were produced which is the highest number of movies ever produced.

### **Highly rated IMDB movies**

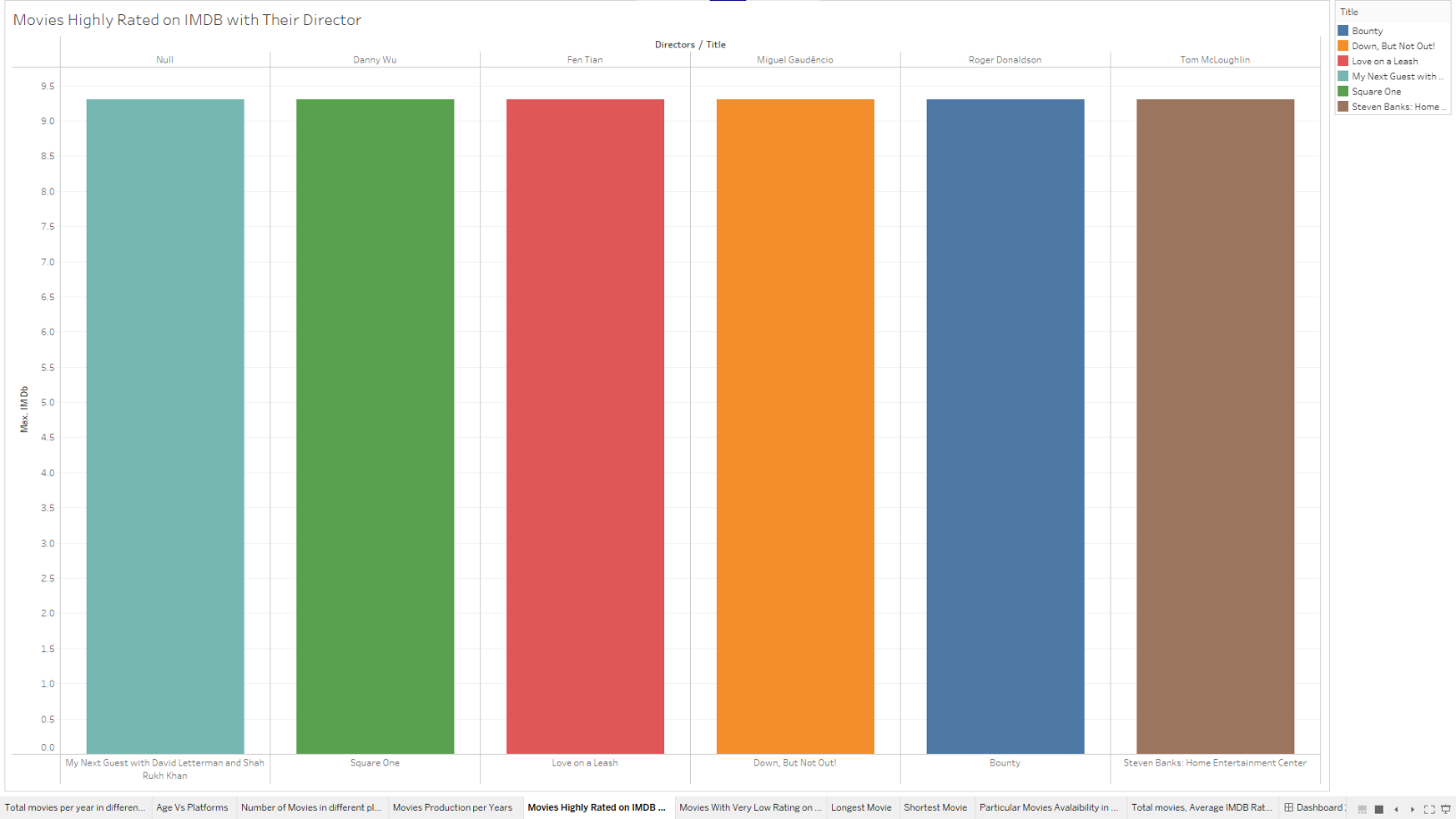


Figure 7: Movies that were highly rated on IMDB with their directors

Figure 7 clears that there is a total of six movies that are highly rated on IMDB. The director of the first movie shown above is not specified in the dataset that’s why that is null.

### **Movies With Very Low Rating on Rotten Tomatoes with Their Director**

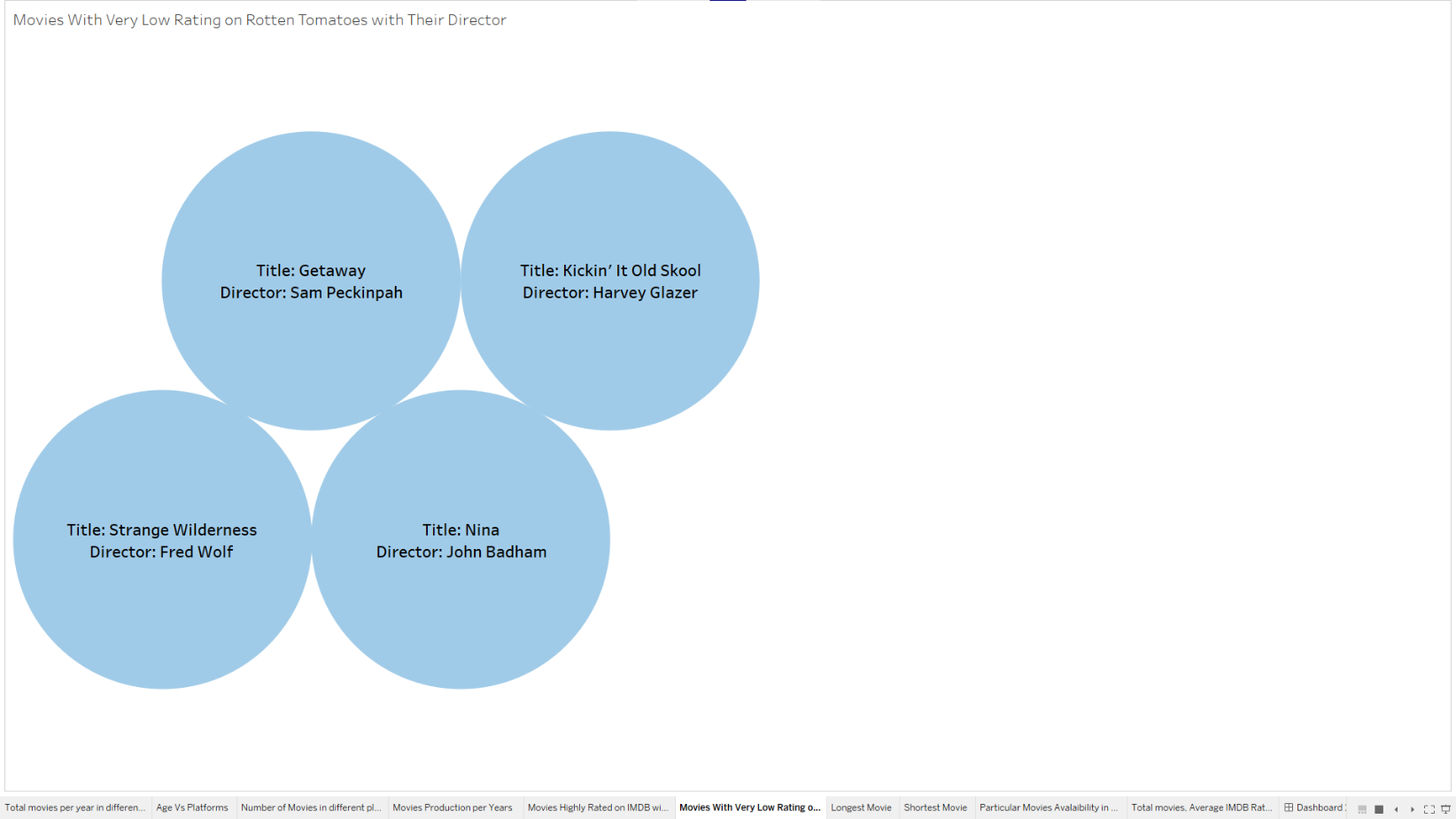


Figure 8: Movies that have low Rotten Tomatoes ratings with their directors

Figure 8 shows that there is a total of four movies that has the lowest Rotten Tomatoes ratings with their Title and Directors.

### **Longest Movie**

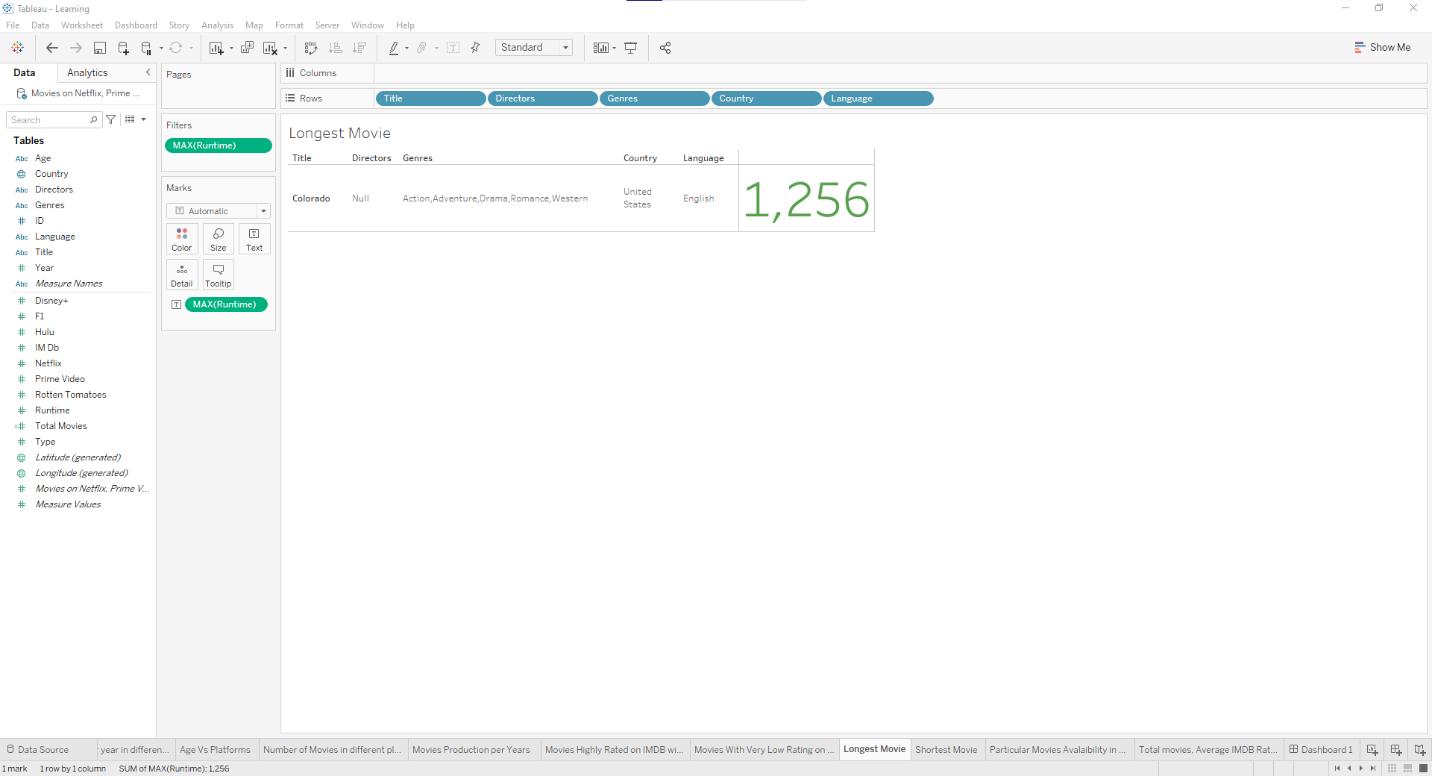


Figure 9: Movie that has the longest runtime

Figure 9 shows that “Colorado” is the longest movie with a runtime of about 1,256 minutes produced in the USA. The director of this movie is not included in the dataset that’s why it’s null.

### **Shortest Movie**

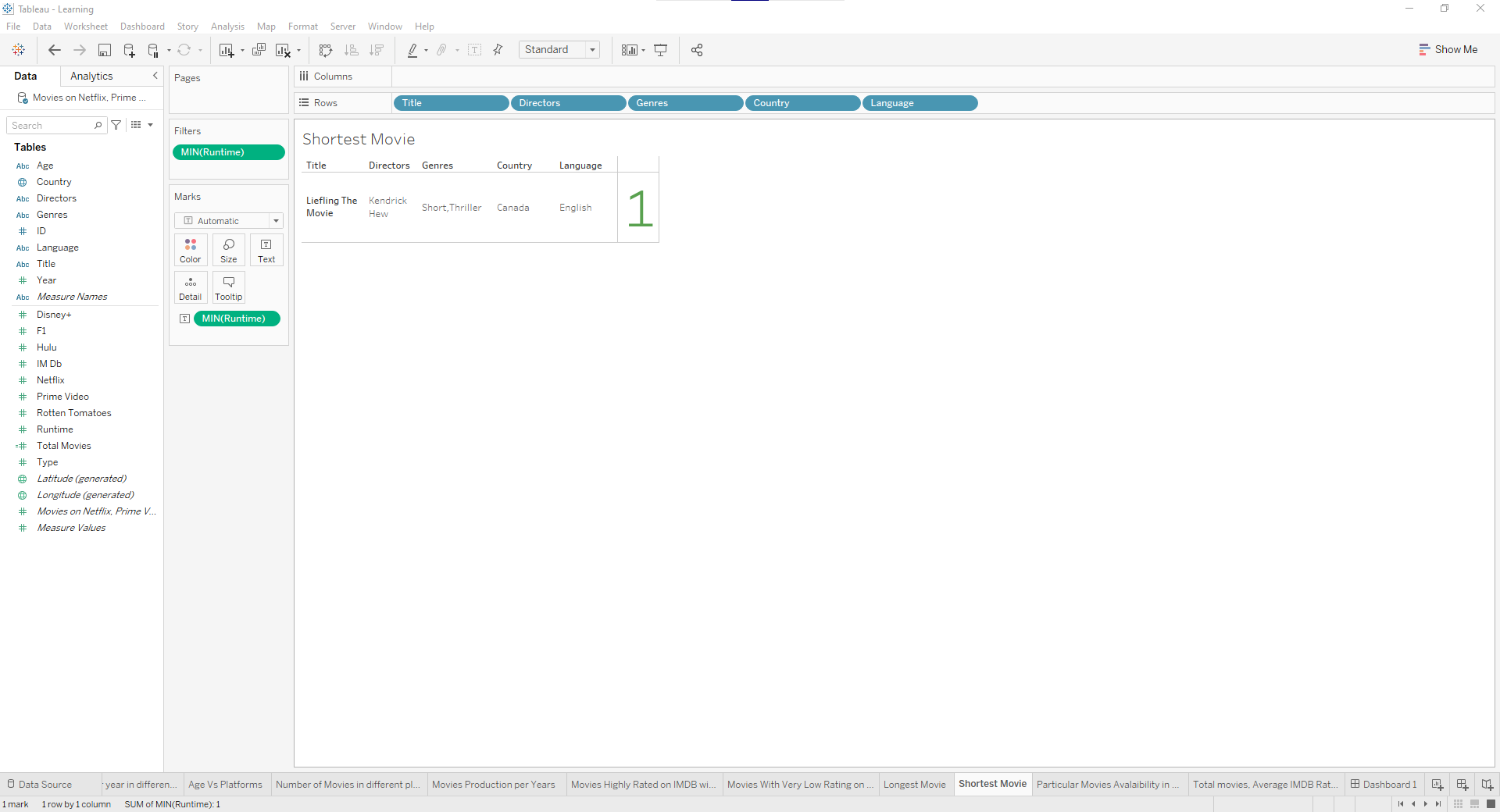


Figure 10: Movie that has the shortest runtime

Figure 10 shows that “Liefling the Movie” is the shortest movie with a runtime of about one minute produced in Canada.

### **Particular Movie Availability on different platforms**



Figure 11: In Which platforms a particular movie can be found?

Figure 11 shows on which platforms a particular movie can be found. Ten movies are selected from the age group “all’ and among which two movies were found on Netflix, three found on Prime Video, three found on Disney+, and one found on Hulu.

### **Total movies, Average IMDB Ratings and Max Runtime by Countries**

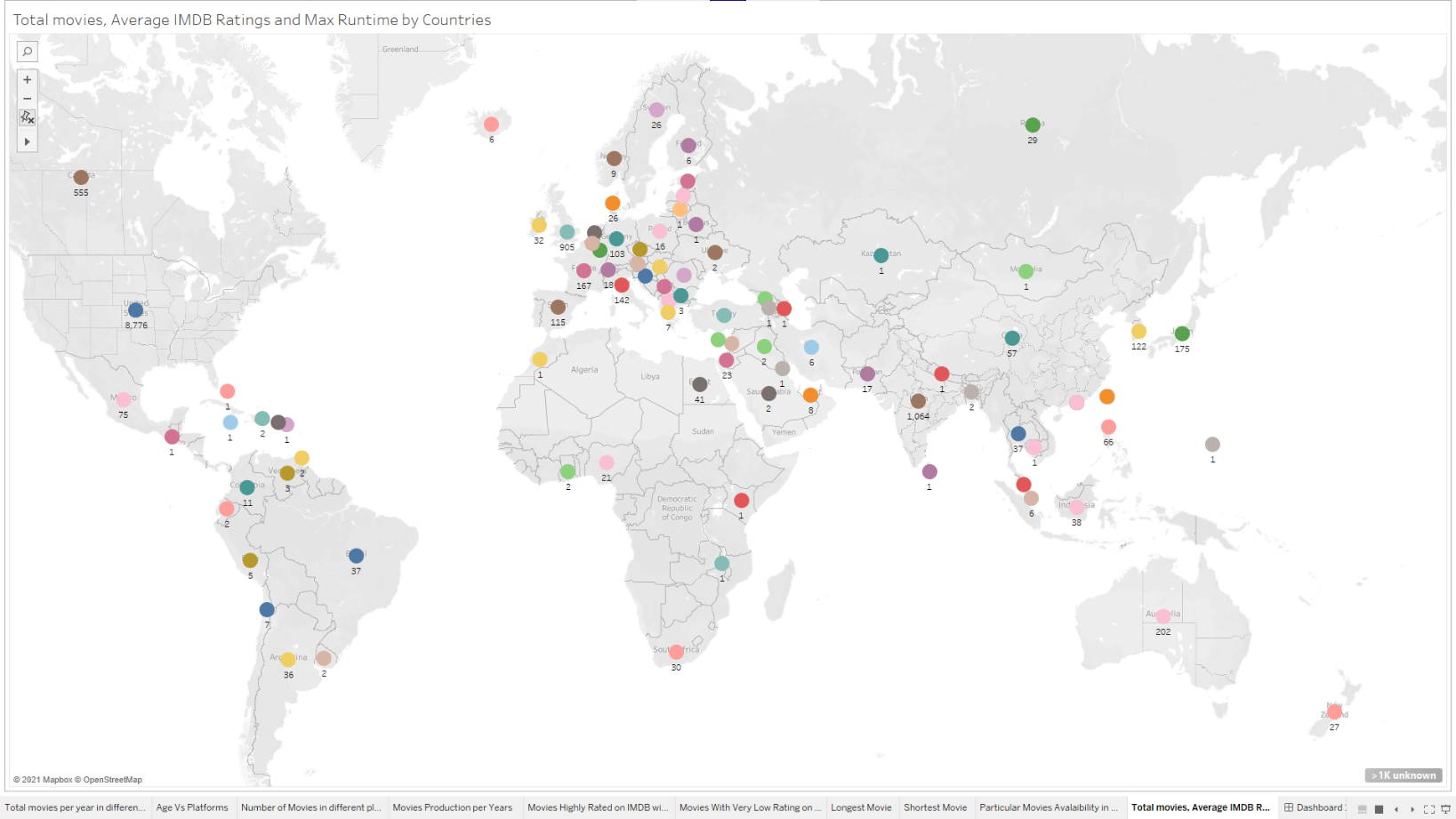


Figure 12: Number of Movies along with average IMDB rating and a max runtime of movies by countries

Figure 12 shows that the USA has the highest number of movies production which is 8,776 and 1 is the lowest number of movie production in a country. The figure also shows average IMDB ratings and max runtime of the movies in a country when you hover in that country. For example, India has an average IMDB rating of 6.310, a max runtime of 224 minutes.

## **Conclusion**

The above analyses prove that Prime Video is the most popular movie streaming platform because it contains a lot of movies that were released from 1902 to 2020 but another platform contains a limited number of movies. There are very few movies that are legal to watch for all age groups but lots of for the 18+ group. 2018 was the year in which the highest number of movies were produced ever. Countries like the USA, India, United Kingdom, and Canada seem to have high-budget film industries because they have produced a high number of movies than other countries.

According to my vision, film industries should not focus only on 18+ movies for popularity and money, they should also consider other age groups too. High-budget film industries like the USA, India should help the low-budget film industries like Nepal, Mongolia, Finland so that the world could see different flavors of movies.

## **Future work**

The above analysis could be done with more details and making it more specific. For example, the number of movies that are available on one, two, three, and four streaming platforms could be analyzed. Which director has directed the highest number of movies, which country has the highest number of movie production in a particular genre, which is the movie that is available on the highest number of streaming platforms? These are other analysis which provides a specific detail.

Some analyses shown above have specific numbers only but if the analyses were also shown in percentage amount, then it would have been more informative.