

PS6

Anais Ouedraogo

March 21, 2023

1 IMdb movies rating

I found an excel dataset in Kaggle with 250 IMDB movies and their information such as rating, budget, directors, writers, and certificates. I did not do a lot of cleanup as the data was ready to work with. I did remove invalid values or rows that had NAs. I also converted the budget column to numeric values so I can perform calculations.

The first graph (Figure 1) that I did is the line graph that shows the average budget for all movies in each year from 1921 to 2022. We can notice on the graph that in 1977 the average budget was much higher compared to other years. This might be explained by big changes in the movie industry such as Hollywood movies opening overseas.

The second graph (Figure 2) I did was a tree map to show the total budget in each genre category. We can see that the Action/Adventure/Drama has the highest total budget.

The third graph (Figure 3) was a bar chart that shows the different certificates and the number of movies that had a particular certificate. R was the certificate that had the highest number of movies followed by PG and PG-13.

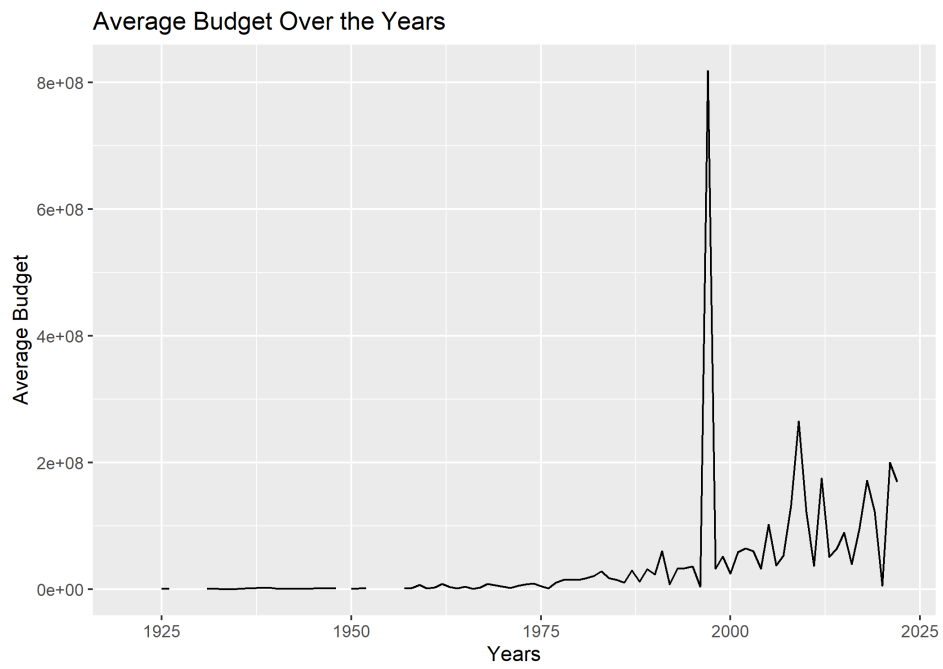


Figure 1: Line chart of average movies' budget over the years.

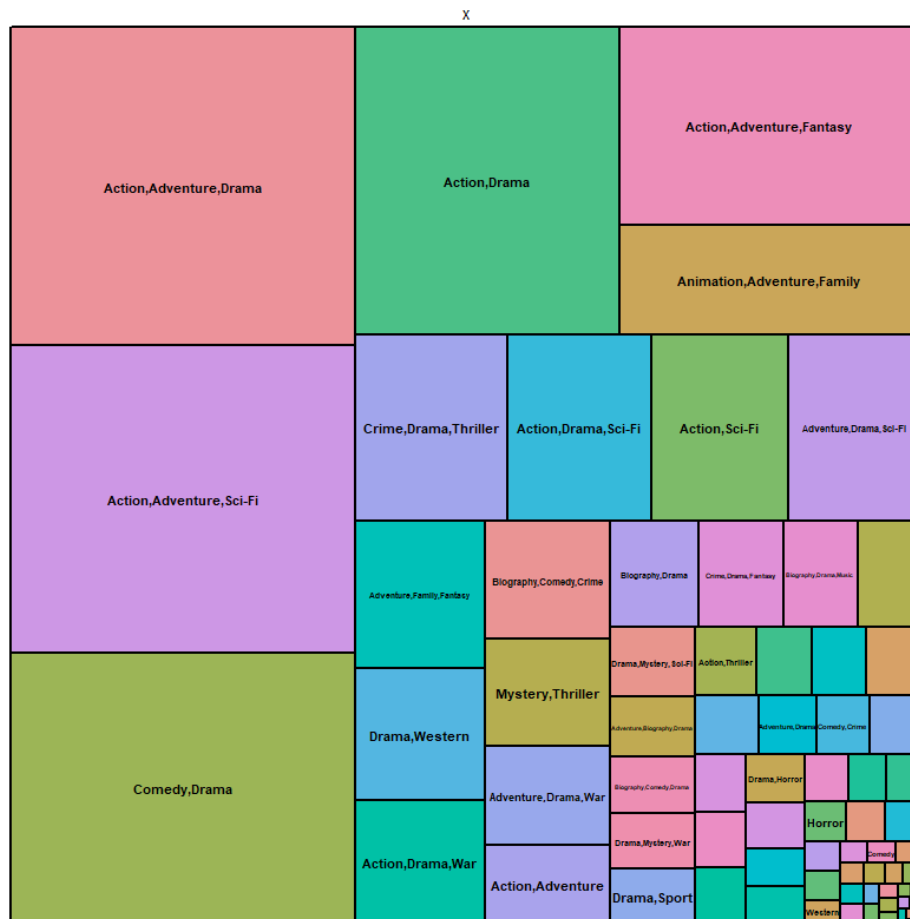


Figure 2: Tree map of the total budget in each movie genre.

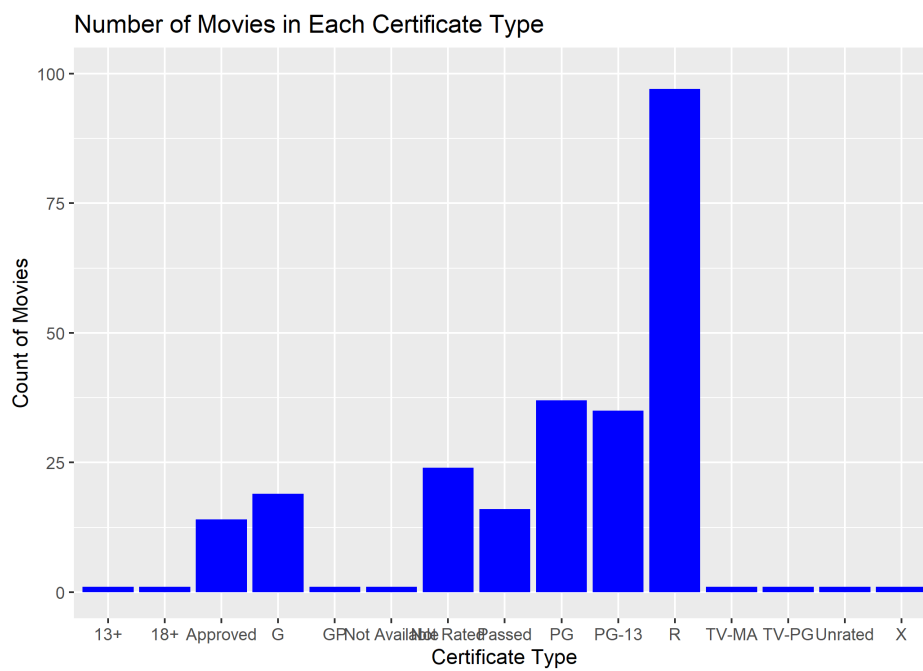


Figure 3: Bar chart of the number of movies in each certificate type.