**IMDB MOVIE RATINGS**

**INTRODUCTION**

The movie industry is one of the growing industries in our current settings. It is one of the significant contributions to global economy. During the COVID 19 outbreak, there was a major downfall in the industry due to the lockdown situation. As of today, it is no doubt that this industry has become more than a billion-dollar industry. Questioning the relevance of this industry is what drove me into conducting this analysis. The film industry includes commercial and technological companies which specialize in cinematography, film studios, production companies, film production, animation, pre-and post-production, film festivals, and distributors. All of these things together are the filmmaking industry. Today, the movie/cinema industry is not only limited to theatres, but OTT platforms are majorly contributing to the overall revenues of films.

**DATA DESCRIPTION**

Data source is [movie industry data (kaggle)](https://www.kaggle.com/datasets/danielgrijalvas/movies)

There are 5421 movies in the dataset after successful cleaning

The data comprises of columns or attributes which includes;

* budget: the budget of a movie.
* company: the production company
* country: country of origin
* director: the director
* genre: main genre of the movie.
* gross: revenue of the movie
* name: name of the movie
* rating: rating of the movie (R, PG, etc.)
* released: release date (YYYY-MM-DD)
* runtime: duration of the movie
* score: IMDb user rating
* votes: number of user votes
* star: main actor/actress
* writer: writer of the movie
* year: year of release

**Explore Data To Answer Questions**

Upon exploring and analysing the data the following questions was answered;

**Most produced movie ratings and genres**

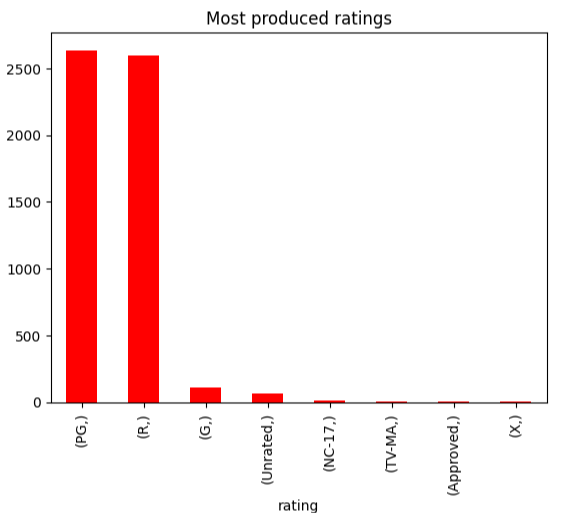
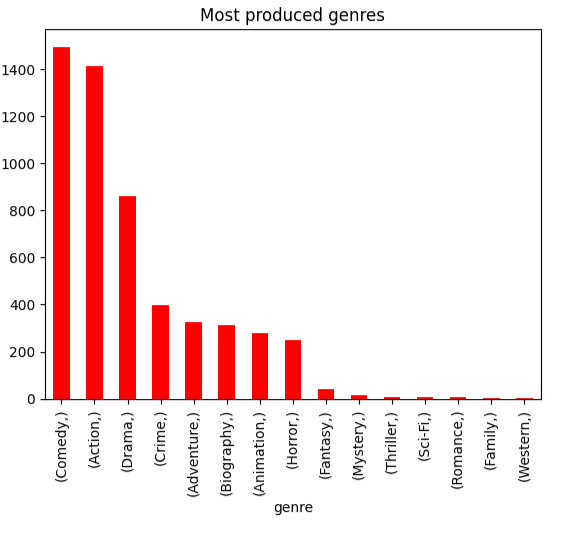
The insight gained after mining the data indicates that **PG** is the most produced movie ratings and comedy being the most produced **genre**

Figure 2

Figure 1

**Average revenue and score of each genre**

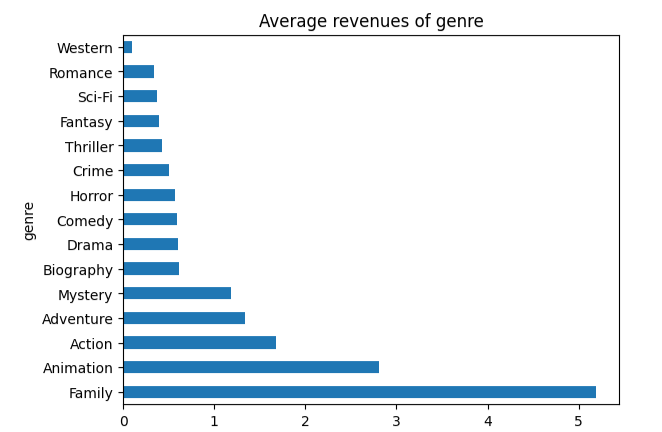
****

Figure 3

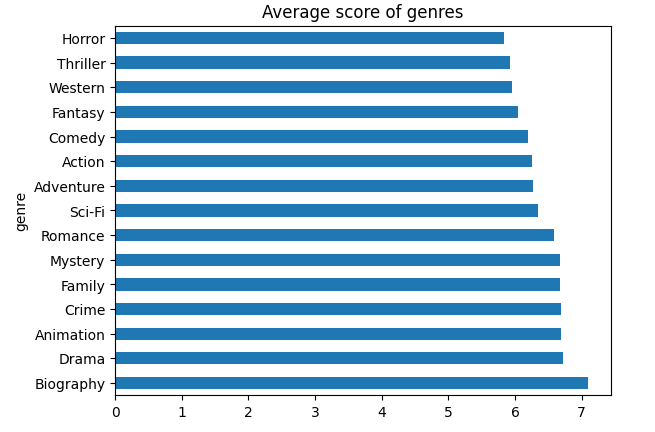
****

Figure 4

**Discussion:** Family genre has the highest average revenue considering the fact that it the least produced genre, so that leads to asking why?

**Why is Family genre least produced even though it has the most average revenue**

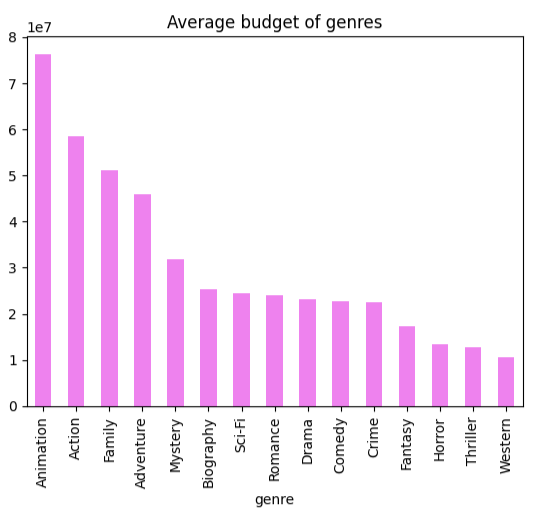
****

Figure 5

**Discussion:** From the graph in Fig 5 above, the top highest average budge genres (animation, action, family, adventure, mystery, biography) are the top highest average revenue in fig 3. This indicates relationship between the revenue and budget.

**Relationship between budget and gross**

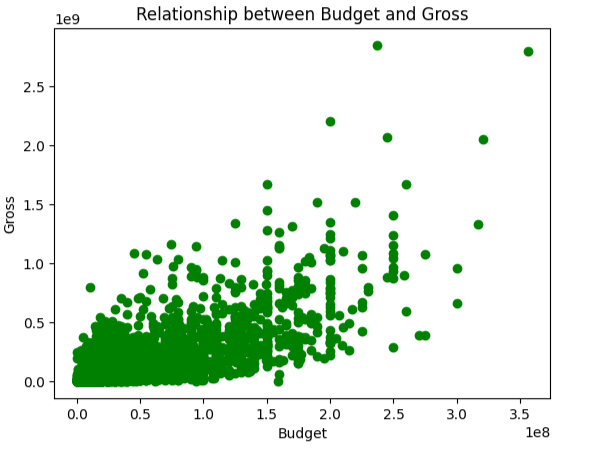
****

Figure 6

**Discussion:** Figure 6 confirms a positive correlation between the gross and the budget. Therefore, we can boldly conclude that the higher the budget of the movie the more revenue or gross it generates.

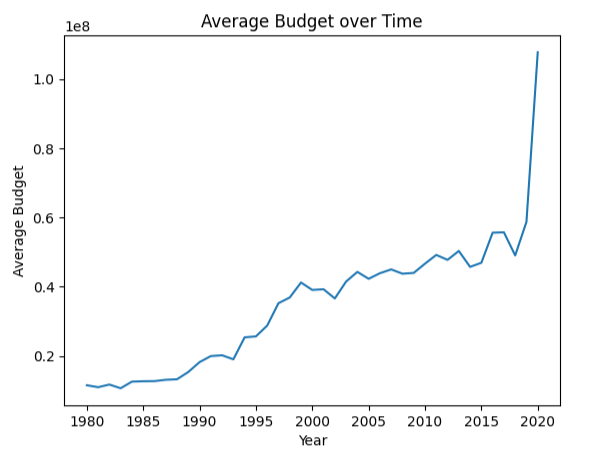
**Change in average budge over time**

Figure 7

**Discussion:** Over the years the average budget of movies has increased and rapidly accelerated through 2018 to 2020

**Relationship between runtime and score**

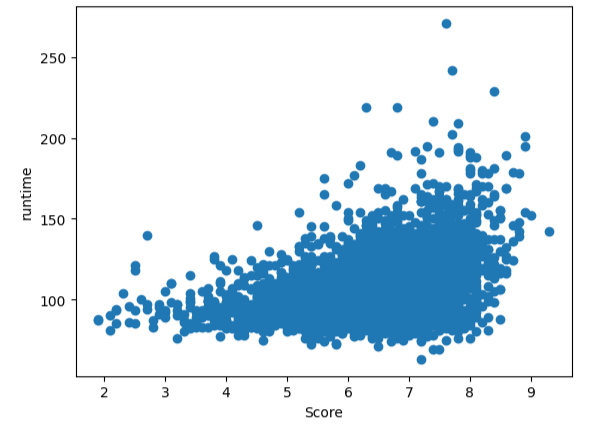
****

Figure 8

**Discussion:** There is a 0.4 correlation between the runtime and the score of the movie which is somewhat a week relationship. Hence the score of the movie cannot entirely be as result of its length.

**Which season does movies release have a high gross**

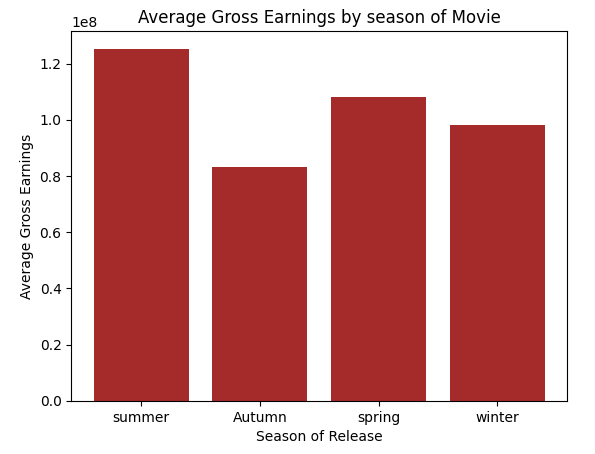


Figure 9

**Discussion:** This figure makes us understand that the most gross are earned when a movie is released in summer and least earnings when it is released in autumn.

**Top 10 movies with the highest revenues**

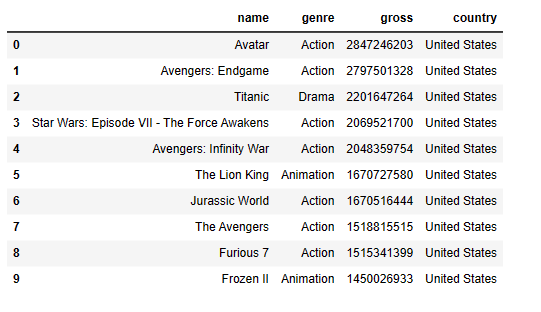
****

Figure 10

**Highest and lowest score movies**

Figure 11



Figure 12