

# Phase 1 Project

NAME: LUCY MUTHONI MUNGE

CLASS: DSF-PT04

TM: STELLA WAITHERA



# Overview

Microsoft wants to venture into the film industry and would like to create a new movie studio but before they get started, they tasked me to evaluate the film industry and give them recommendations before making their final decision. My recommendations are based on the profitability of the different movie genres both in the local(USA) and global markets, the popularity of the different genres, the runtime of the different major movies and lastly the ROI in short films.



---

# Problem statement

The main problem for Microsoft is determining if venturing into the film industry is a viable business idea. They need to identify what kind of films they need to create, how to make profits and generally how to set themselves apart in order to be the market leaders in this industry.

The background of the slide is a teal-tinted image of a computer monitor. On the screen, there is a line graph with several data points and a pie chart. Below the charts, a portion of a Windows taskbar is visible, showing icons for various applications. A keyboard is partially visible at the bottom of the frame.

---

# Datasets Used

For this project, we will be analyzing data from three movie websites:

- IMDB(Internet Movie Database)
- The MovieDB
- The Numbers

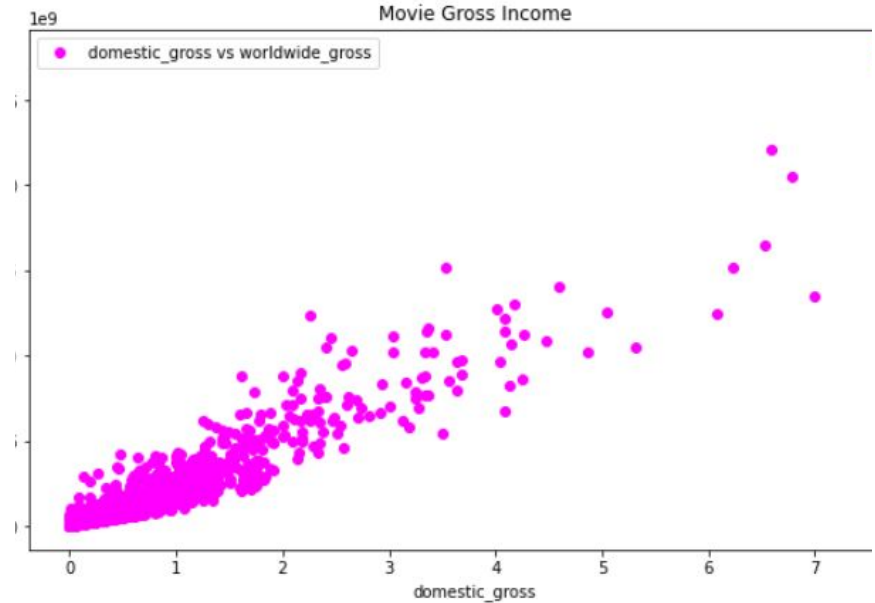


# Data Understanding

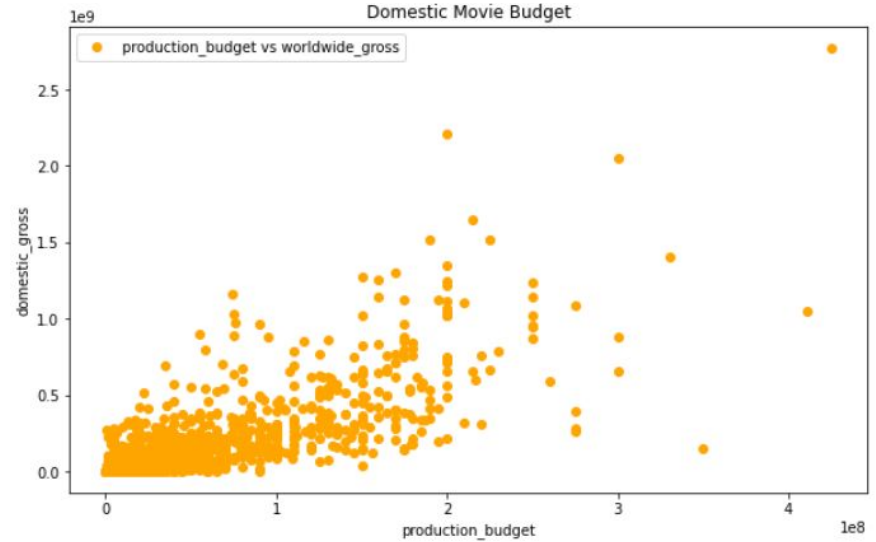
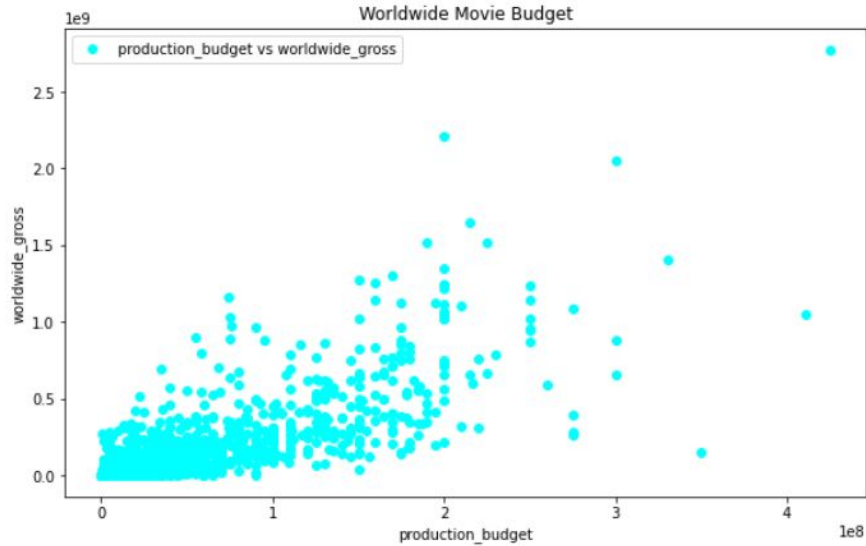
- IMDB(Internet Movie Database) was used to obtain the movie genres, the types of films microsoft can produce.
- The MovieDB dataset was used to obtain the rating and popularity of the films
- The Numbers dataset was used to obtain the budget,profits and ROI of the films.

## Correlation between the domestic\_gross and worldwide\_gross

There is a strong and positive correlation between the domestic and worldwide gross. There's a probability that if a movie does well in the domestic market, it will also do well in the worldwide market. This also shows that there's an importance in focusing on the local market for the growth of the movie studio.

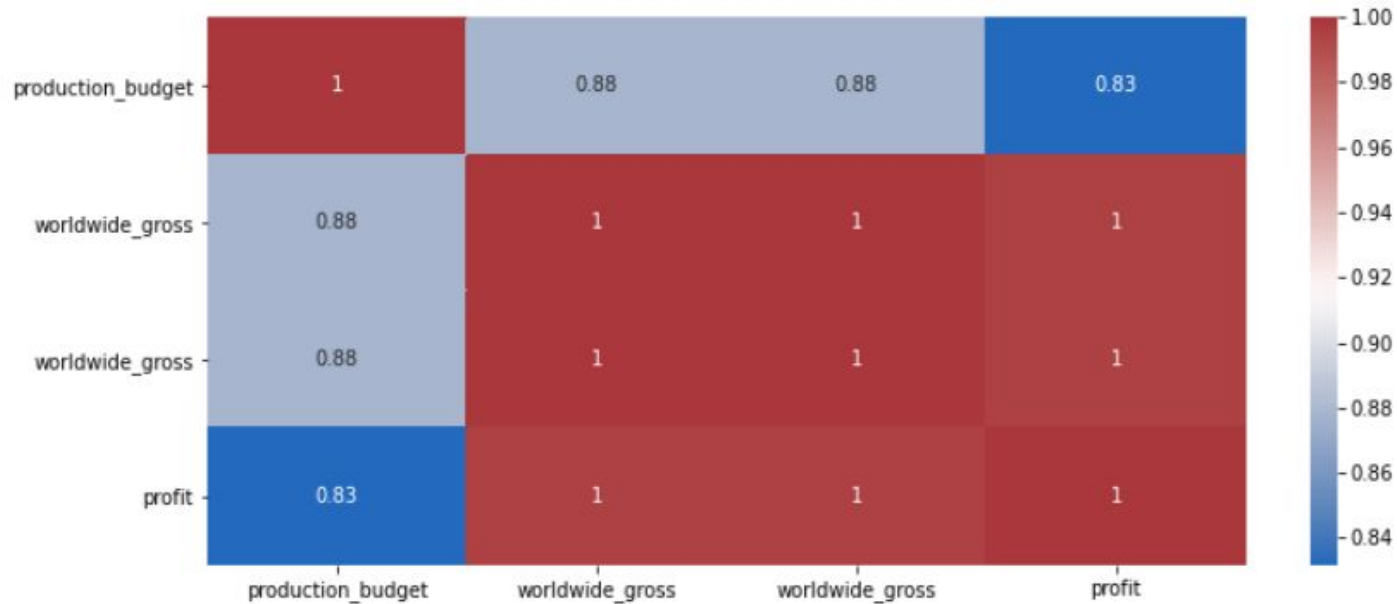


## corr between the production\_budget and worldwide\_gross / production\_budget and domestic\_gross



From the above scatterplot we can see a strong and positive correlation between the production\_budget and worldwide gross, with some outliers noted. There's a probability that if some movies will do well with a bigger budget, and this might be attributed to a couple of reasons, ie. most of the budget could have been used in the marketing of the movies especially in the worldwide market, Investment in popular movie directors and actors which comes at a higher cost, some movie genre might cost higher because of the special effects like horror movies, scifi etc.

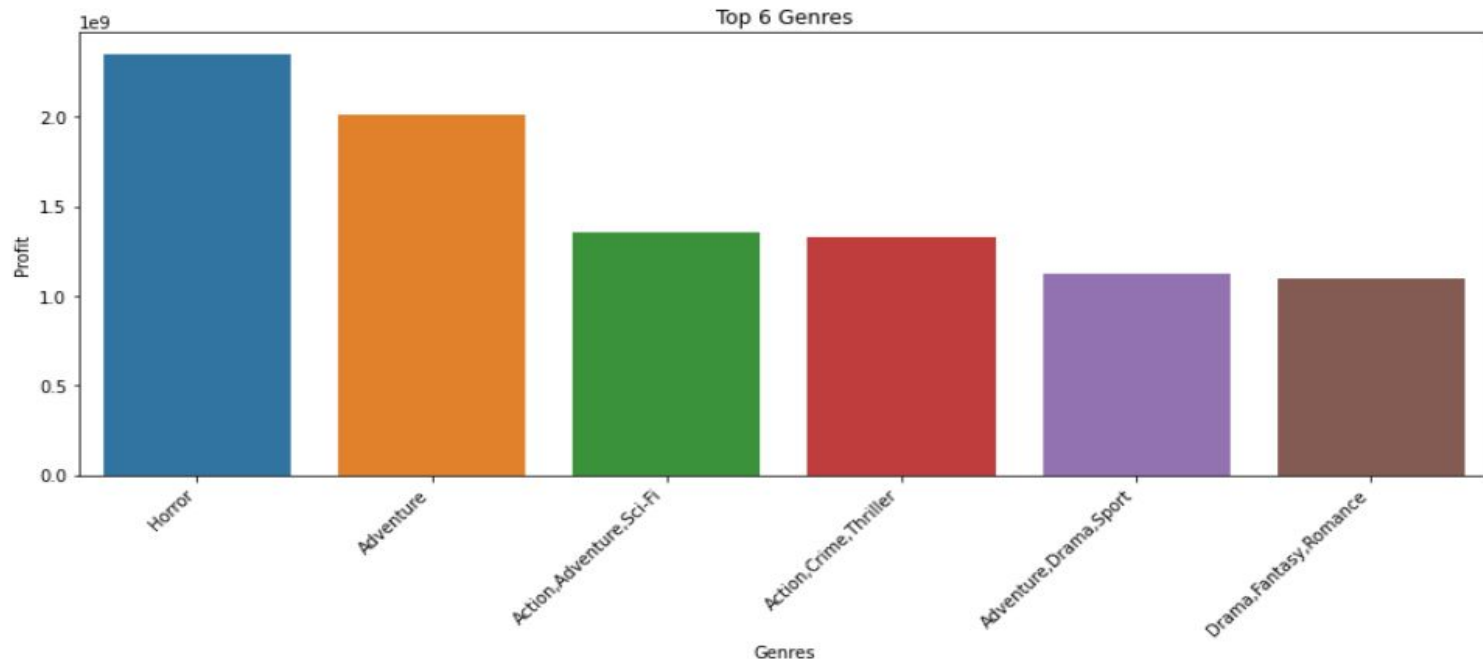
## Correlation between production\_budget, worldwide\_gross, worldwide\_gross and profit



From the heatmap above we can see that there is a good correlation between production\_budget, worldwide\_gross, worldwide\_gross and profit, We can predict that the higher the budget the higher the income for the different movies.

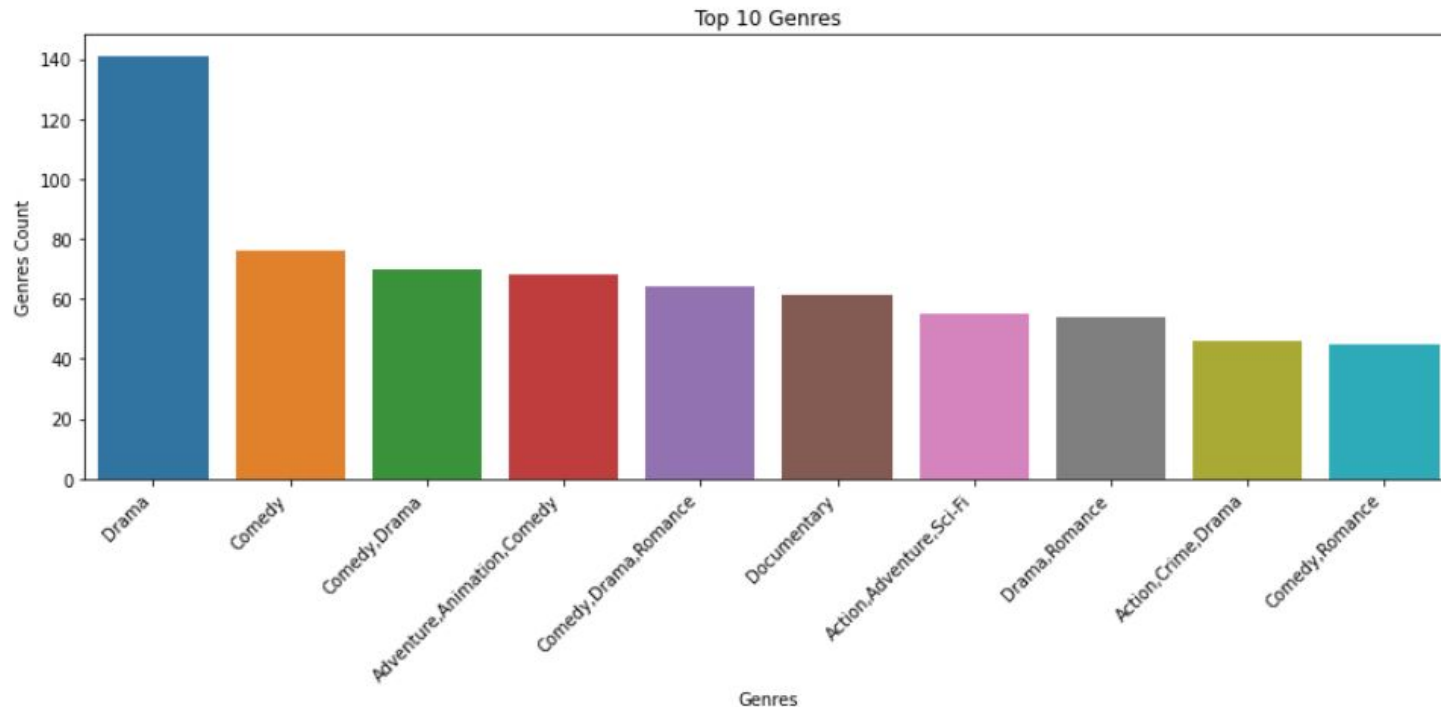


## A bar plot showing the top 6 genres making the most profits



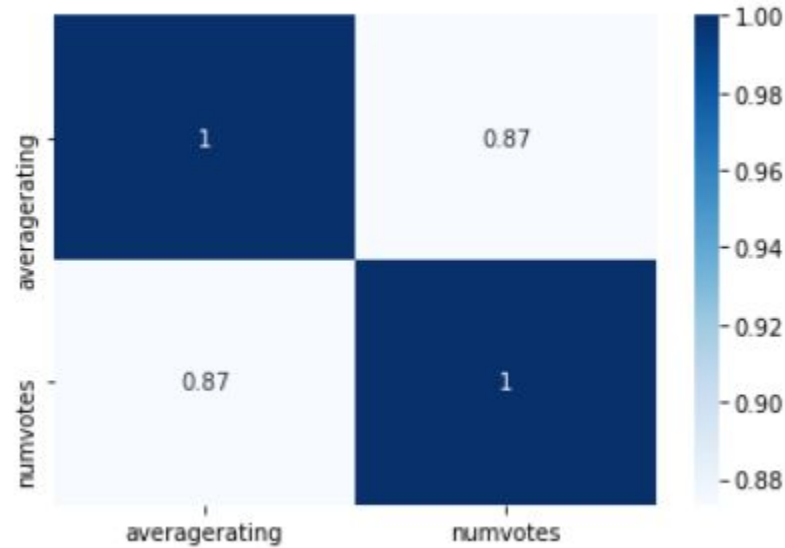
From the bar plot above we can see that Horror movies make the most profits followed by adventure movies. We can also see that there are certain movies that fall under more than one category and are making profits, especially if it also falls under action and drama movies.

A bar plot showing the top 10 genres with the most value counts



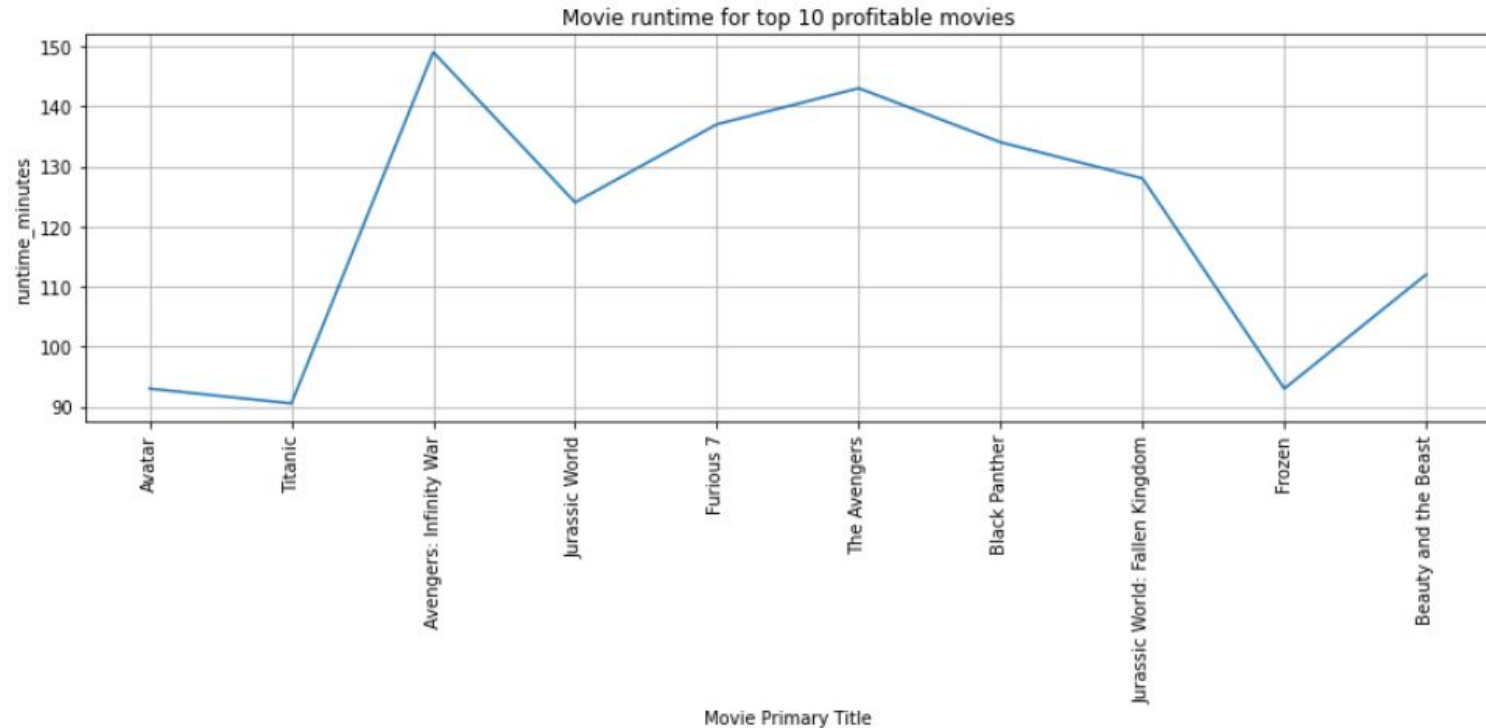
From the above bar plot we can see that the most popular genre in terms of the number of times movies have been created in that genre, is Drama followed by comedy. From the previous analysis we saw that horrors and adventures make the most profits but drama and comedy are created the most. This might be attributed to a number of things, eg. online streaming shows might be airing these kind of movies because of the demand, The cost of making drama and comedy shows might be lower hence easier to produce more.

A heatmap to show the correlation of the average rating and the number of vote



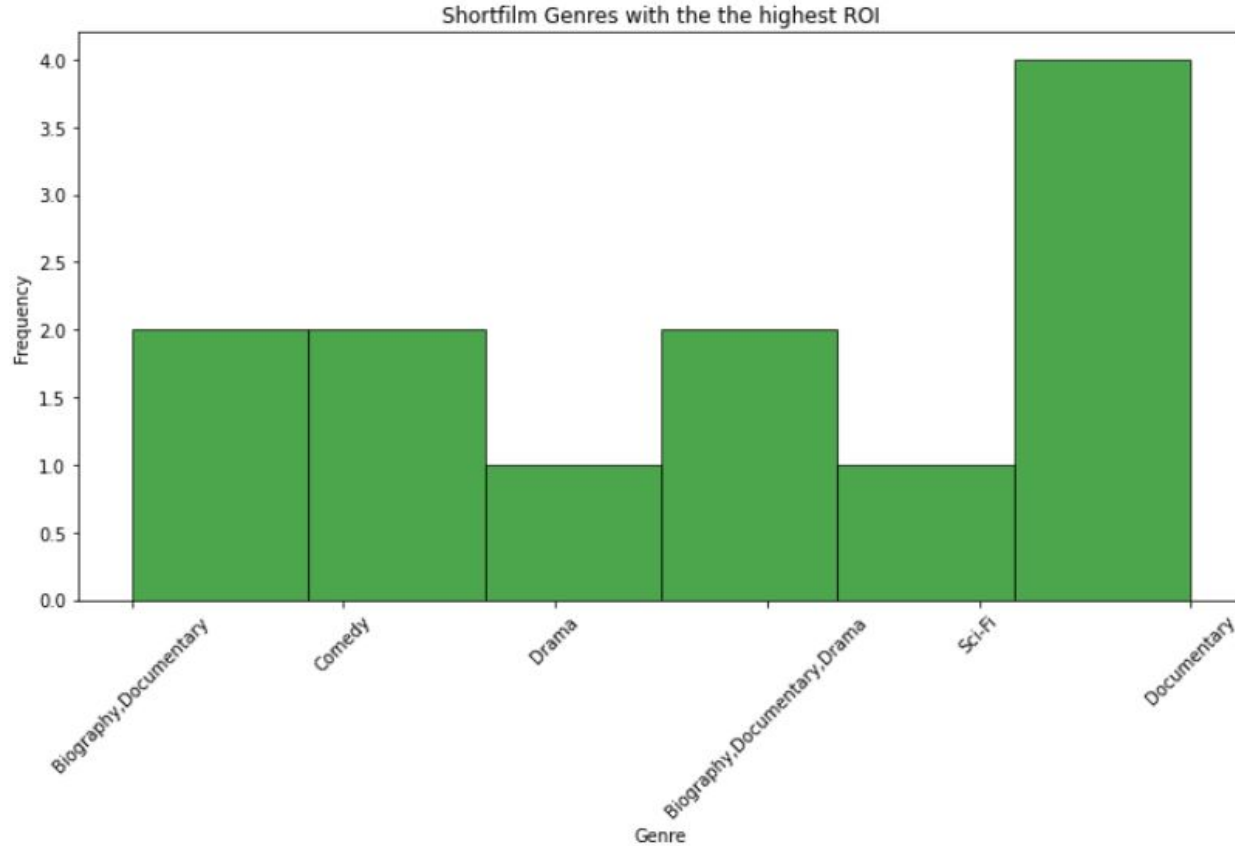
From the heatmap above we can see that there is a good correlation between the average rating and the number of votes. We can predict that the higher the rating the higher the number of votes for the different movies.

## A lineplot showing Movie runtime for the top 10 profitable movies



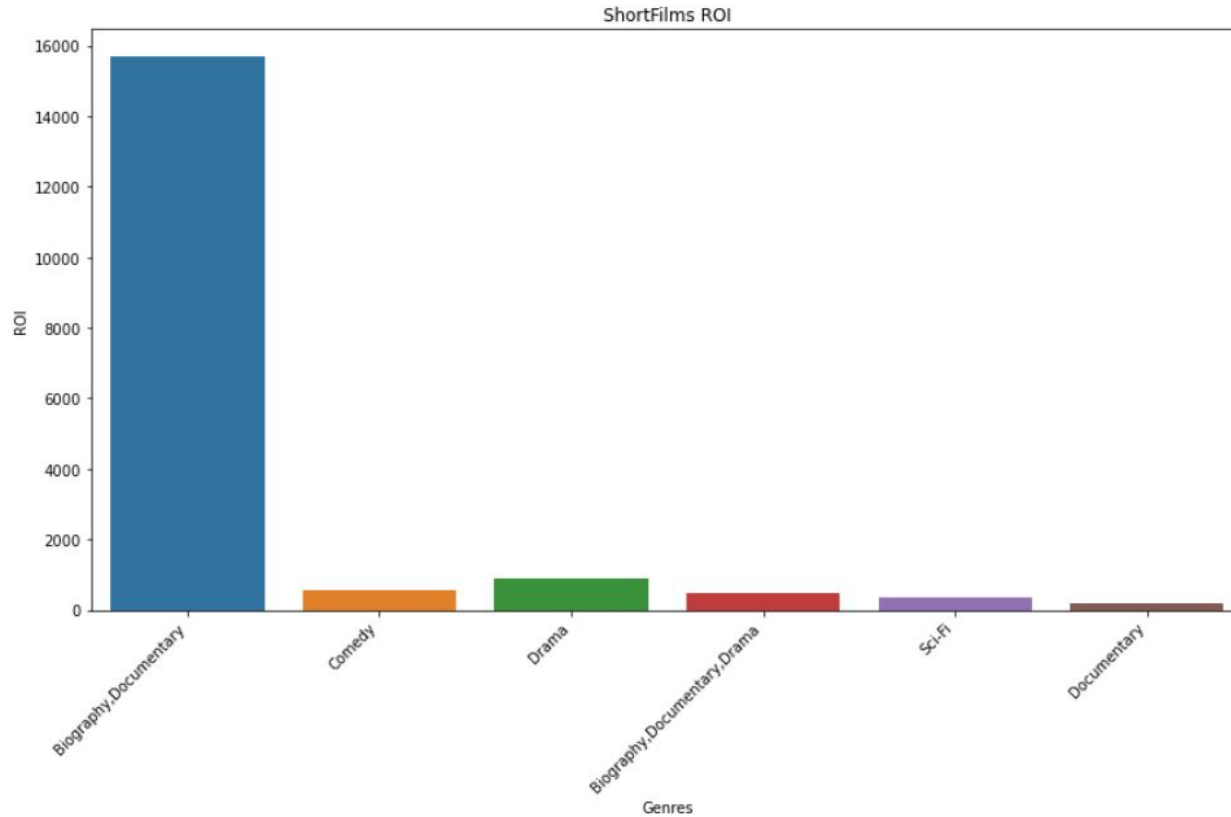
From the line plot above we can see that the most profitable movie ran for less than 95 minutes, the longest movie was Avengers which ran for almost 150 minutes, while the shortest, Titanic, ran for around 90 minutes. So we can conclude that averagely, a movie should last between 90 to 150 minutes

## A histogram showing the frequency of the Short Film Genres with the highest ROI



From the histogram above, we can conclude that in short films, Documentaries are the most preferred short films.

## A bar graph showing the Short Film genres with ROI greater than 20%



From the barplot above, we can clearly see that Short films in Documentary and Biography have the highest rate of ROI.

---

# Conclusion

Below are the findings from the analysis conducted above:

- The budget allocated for the production of the film will play a huge role in the profitability of the film alongside other factors like the cast, producers, marketing, language of the film etc. Also note that it's not guaranteed that the budget allocated will automatically result in the success of the film.
- Popularity of a film doesn't guarantee the profitability of the film. From our analysis, there are some genres that were very popular in terms of vote counts but the profitability was slightly lower.
- Films that fall under more than one category are popular and profitable too.
- There's a good ROI in short films specifically Documentaries and Biographies.



# Recommendations:

- Consider having a good budget for the film production, since we have observed there's a positive correlation between the budget and the income. But it should be noted that there needs to have a very good marketing strategy for the film, while factoring in the producers and cast to work with.
- Short Documentary films have a high ROI, therefore it is a good genre to focus on too. Diversifying the genres and the kind of films they create is very important as they'll appeal to a wider scope of viewers.
- There are certain external factors that affect the success of a film, therefore understanding the market and the current social-demographic issues such as - 'Black lives matter', 'Gender Inclusion', 'race' etc.- is very important. It would be good to help push the agendas the company resonates with so as to appeal to its target market.
- The runtime minutes for films is averagely between 90 - 150 minutes, therefore if there's a probability of having a longer film, it's best to create a limited series.



# Questions?

---

# Thank You!!!

Name: Lucy Muthoni Munge

Email: [mungelucy07@gmail.com](mailto:mungelucy07@gmail.com)

Contact: 0725277106