



Movie Data Analysis

Helping Microsoft Venture Into
The Film Industry



By Nagib Gonzalez





| What Should Our Budget Be And Expected Return?

- We should start with three movies and spend 150 million per movie totaling 450 million dollars.
- We should aim for a return of 250-275 percent which is a profit of 1.125-1.237 billion dollars.



| What Should We Focus On Next?

- Actors and directors with previous experience in Animation, Adventure and Sci-Fi.
- The relationship with profit and movie release date.
- The relationship with movie length and profit.



| Summary

After analyzing relevant movie data from IMDb and The Numbers I was able to derive insights that brought me to the following conclusions,

- The movie genres with highest return on investment percentage were Mystery, News and Animation.
- The movie genres that profited in terms of dollar amount were Animation, Adventure, and Sci-fi.
- The average movie budget was about 50 million dollars.



| Business Problems

- What type of movie genres should we focus on making?
- What should our budget be for the films?
- What should our expected return be given our budget.



| Data



The Internet Movie Database (IMDb):

IMDb is the world's most popular and authoritative source for movie, TV and celebrity content. Find ratings and reviews for the newest movie and TV shows. The table that I used in this database was the movie basics which contained the movie titles and genres.

The Numbers:



The Numbers provides detailed movie financial analysis, including box office, DVD and Blu-ray sales reports, and release schedules. This data set provides the financial data that we will be working with including production budget and worldwide gross. From this data set we were able to find the ROI percentage and profit.



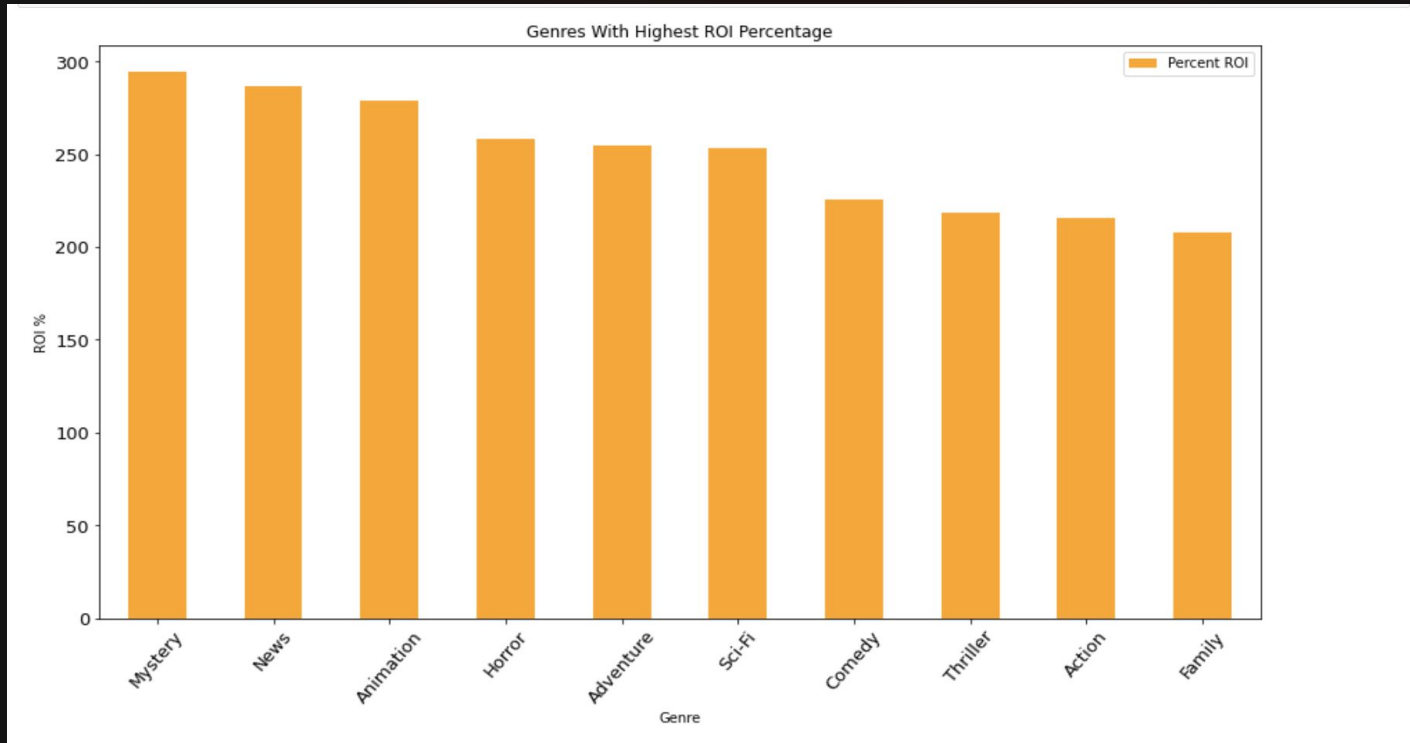
| Methods

In this analysis we focused on ROI which is calculated $(\text{worldwide gross} / \text{production budget}) \times 100$. We also focused on profit which is calculated $(\text{world wide gross} - \text{production budget})$.

As a measure of central tendency due to the skewness of the distributions and the high number of outlier movies we used median instead of mean for the most part. I did make a visualization which shows the average budget per genre but the median is also displayed.

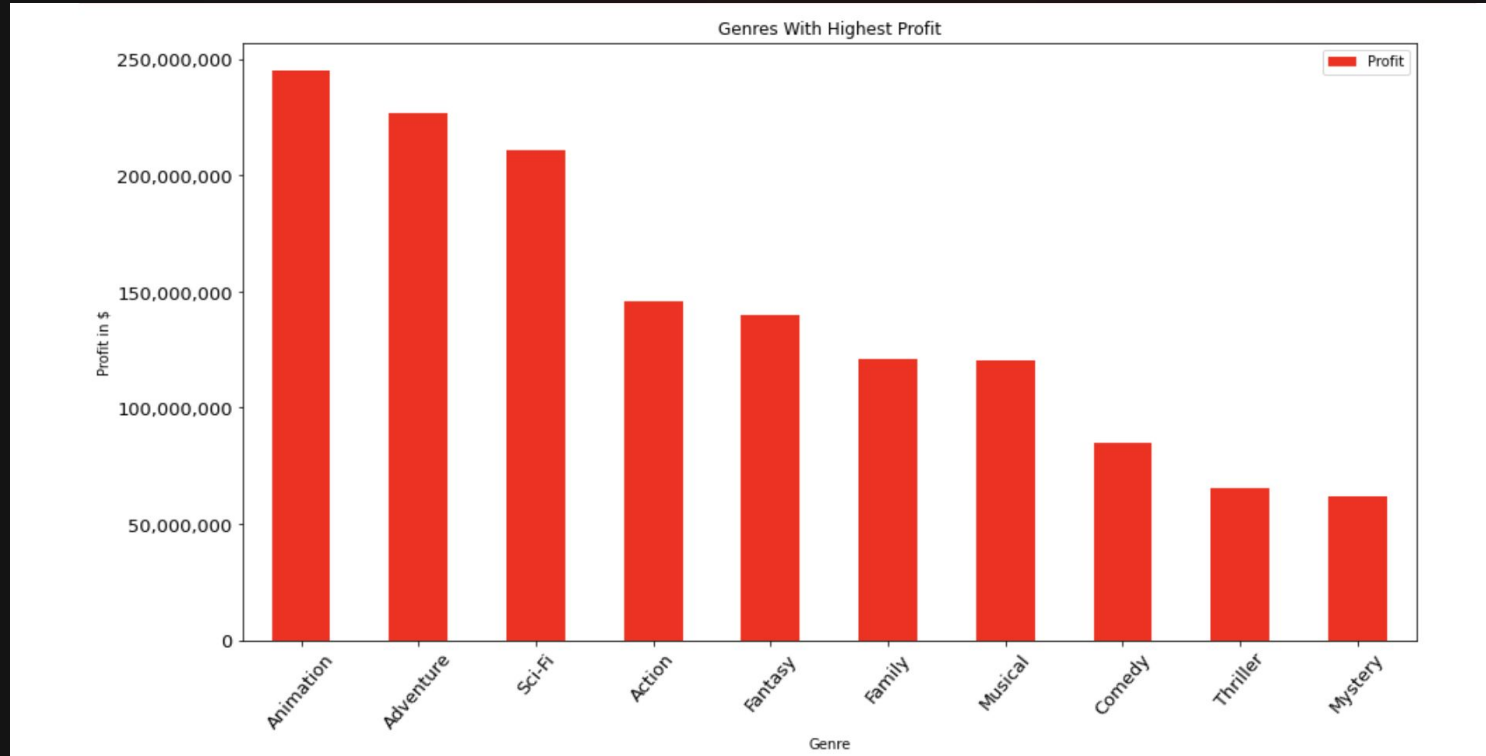


| Genres with Highest ROI





| Genres With Highest Profit





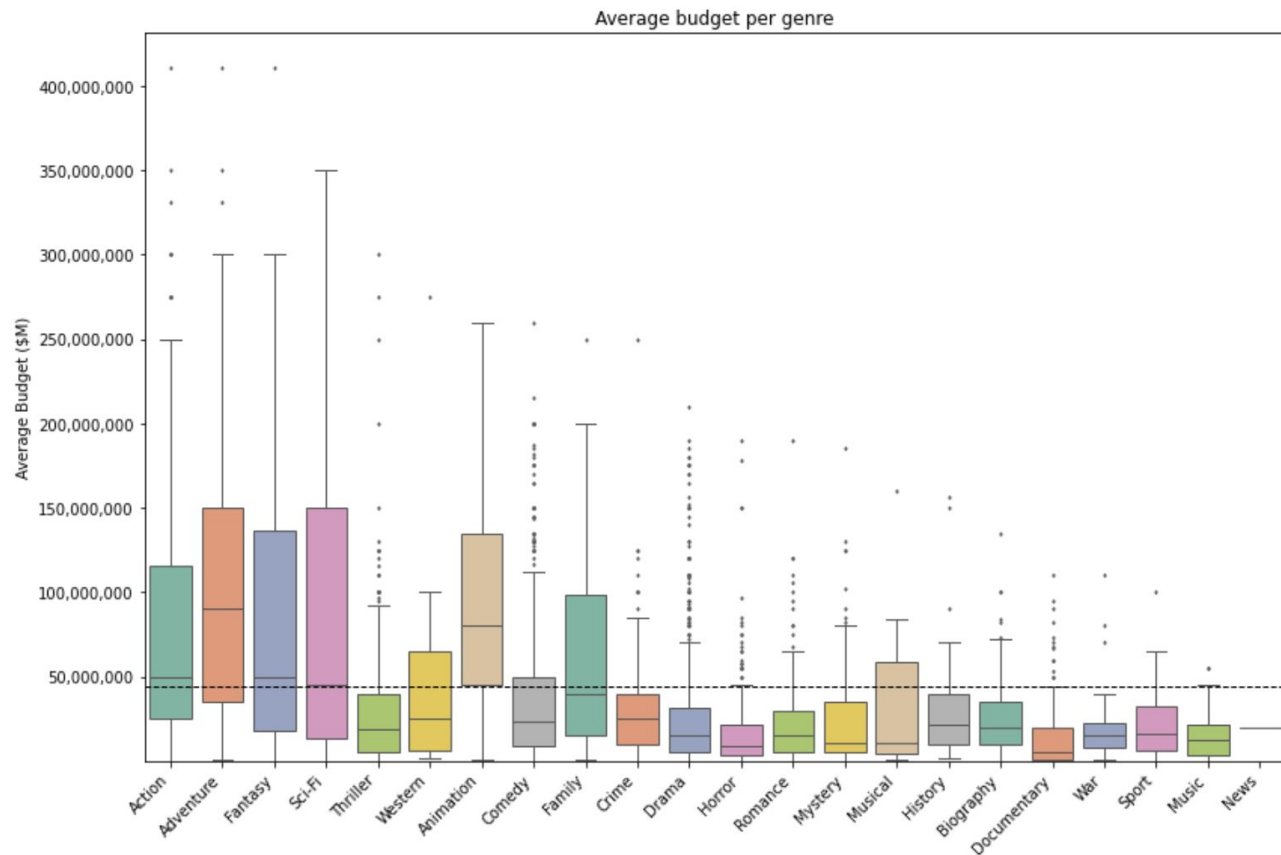
| What Genres Should We Make and Why?

The genres that we should make are:

- Animation
- Adventure
- Sci-Fi

Each of these genres has an roi of at least 250% and bring in over 200 million in profit. Although Mystery and News both have a slightly higher roi, they both bring in a significantly less profit.

Average budget per genre and median





| Thank you

- If you would like to see a more detailed analysis please visit the link below:
- <https://github.com/inagib21/Movie-Data-Analysis>