Optimizing Factors to Maximize the Success of Movies

By Lucas Fishbein
(Flatiron DS Phase 1 Project)

Project Overview

- Microsoft is opening up a new movie studio
- How do we maximize the likelihood of success from the get-go?
- The culture and trends of movies globally is in constant flux
- To provide recommendations to maximize the likelihood of success, **a descriptive analysis has been completed on a set of movies released after 2010** in order to elucidate factors from these movies that may have contributed to their success

Analysis Criteria And Source Databases

Analysis Criteria:

In order to maximize Microsoft's investments our success metric will be Return on Investment (ROI)

Movie Factors explored to obtain a high ROI:

- Production Budget
- Genre
- Release Month

Data Sources:

- IMDB.com - A repository of descriptive features of movies
- The-Numbers.com - a repository of data on the financials of Movies

From these sources, **a combined database of 1,859 movies released after 2010** was utilized for the preceding analyses

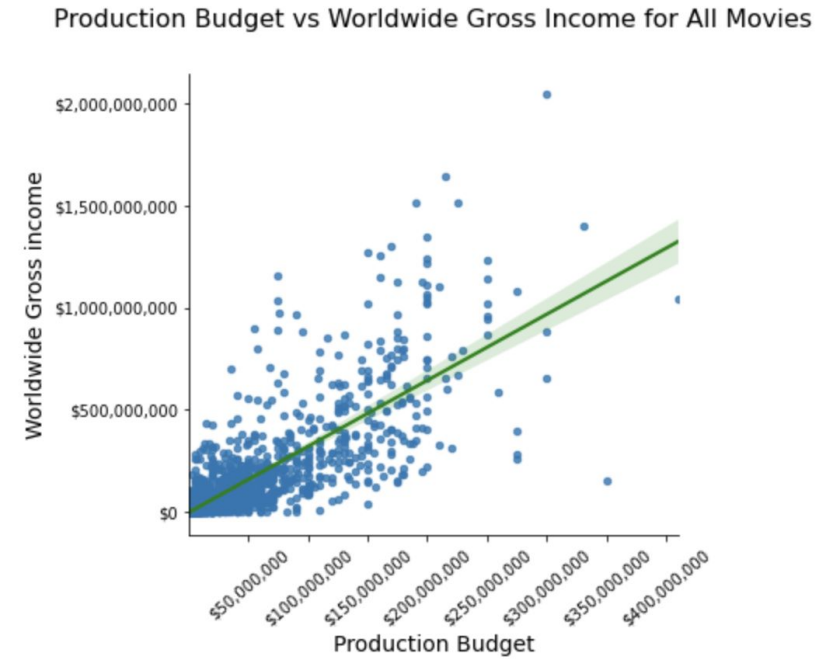
Data Analysis Basis

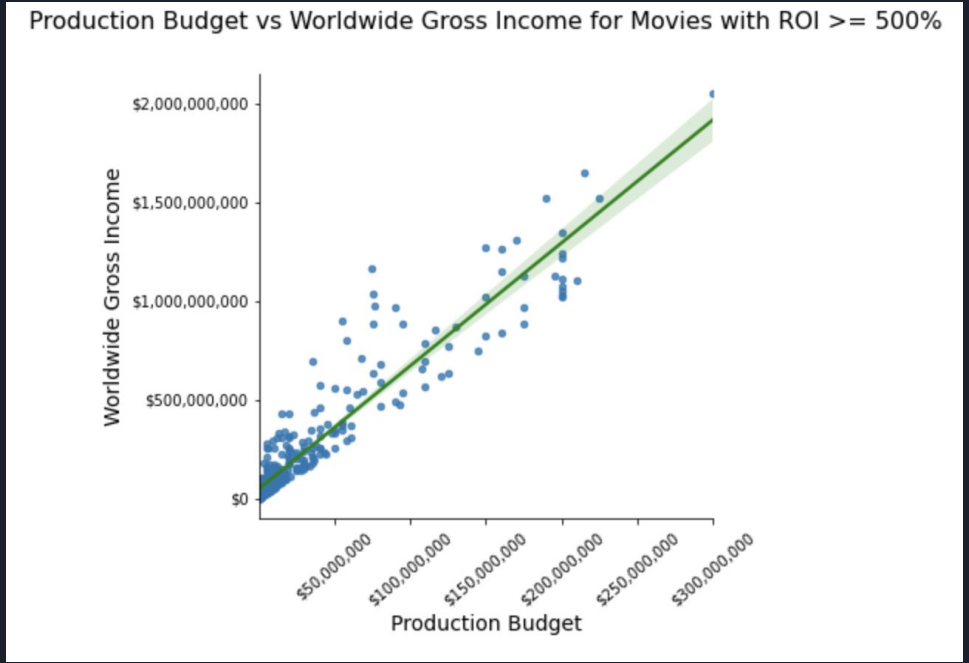
- To determine the rate of success for each of the variables explored the percentage of movies the became “Highly Successful” versus the overall data set was utilized

- Highly Successful movies have an ROI >=500%
  - 320 movies in the present dataset were highly successful
  - These will be used as the “Top Movies” dataset

Production Budget vs. Worldwide Gross Income

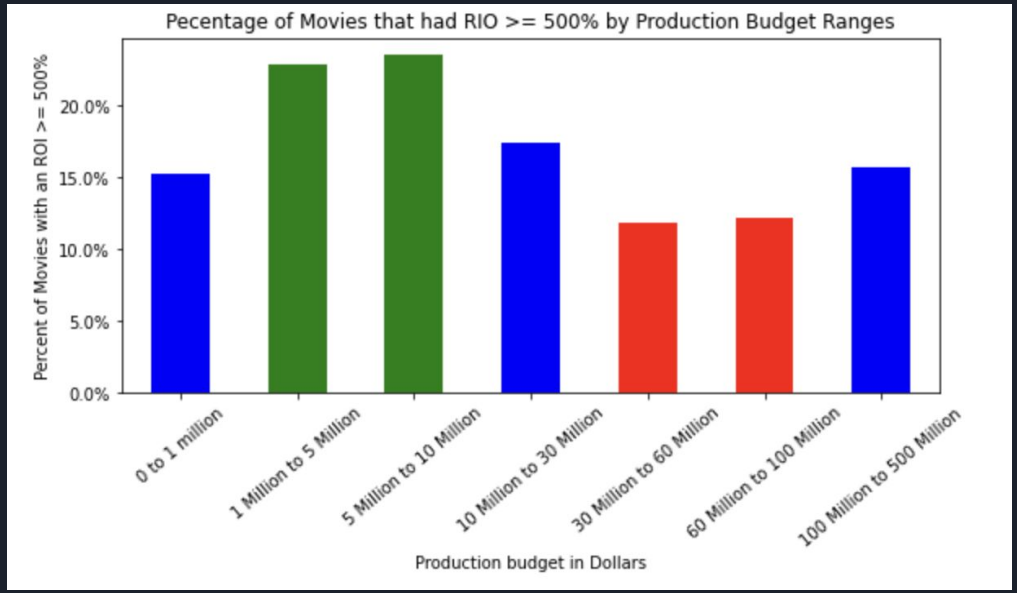
The general trend across both graphs is **the more you invest the higher the worldwide gross**





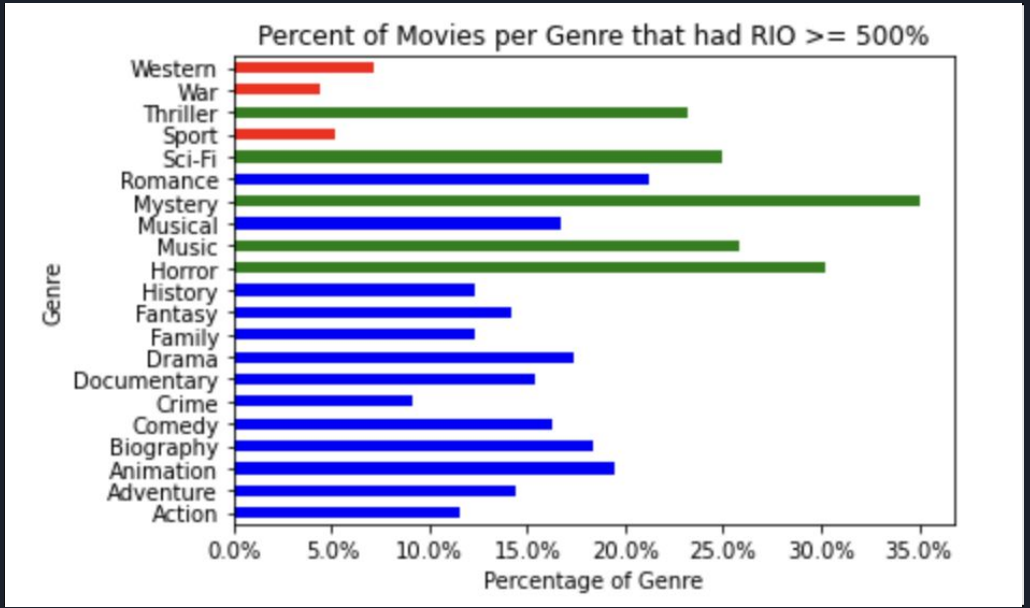
Production Budget Analysis

- One to Ten Million range has highest rate of success
- Thirty to One Hundred Million is the range to avoid



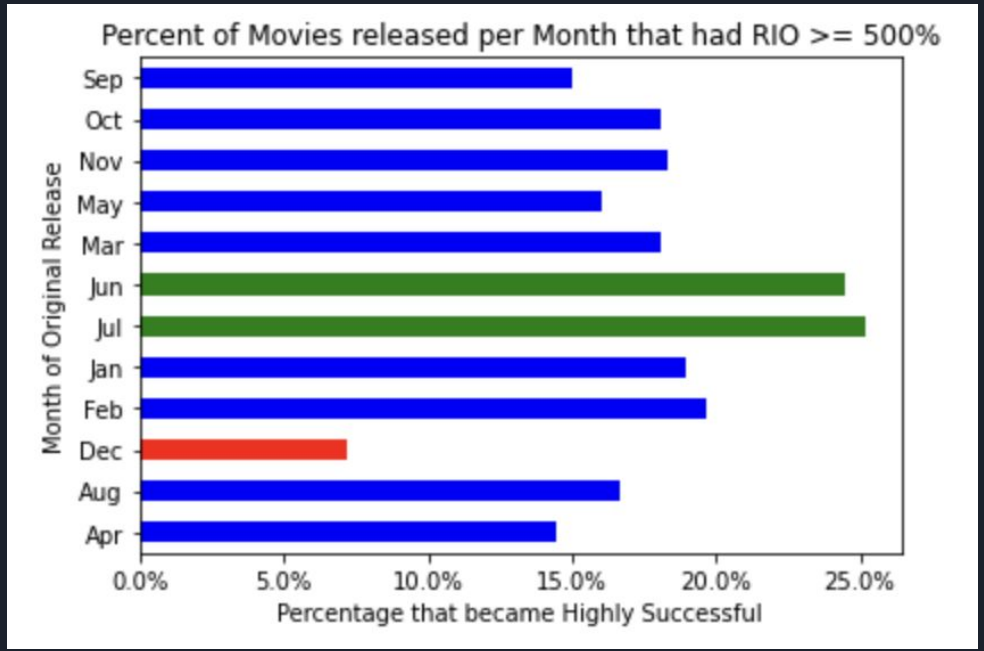
Genre Analysis

- Top Genres were Mystery and Horror
- Other High performing genres included Music, Sci-Fi, Thrill
- Poorest Performers were War, Sport and Western



Release Month Analysis

- June and July had the highest rate of success
- December should be avoided



Summary of Results and Recommendations

- The factors to include based on this analysis to maximize the chance of an ROI >= 500%:
  - Production Budget:
    - Between One and Ten Million Dollars were top performer
    - Avoid Thirty to One Hundred Million dollar range
  - Genre
    - Mystery and Horror were top performers
    - Avoid War, Sport & Western
  - Release Month
    - June and July were top performers
    - Avoid December

Limitations and future improvements

- Do not fully understand the origins of the databases
  - How movies were selected for inclusion
- A much larger dataset could be built beyond the limitations of the provided datasets.
- This analysis does not actual measure the how any of these factors impact the results as it is only a descriptive analysis
- It is possible that two movies with the same name came out in the same year
  - The-Numbers.com database sometimes used the “Video Release” date in place of the original release date, to compensate a tolerance of 1 year was used during the merge.
- The same movie may have been released at different times in different countries
- Investment portion of ROI does not include all factors of interest such as advertising budget
- Worldwide Gross does not contain factors such as merchandise and other peripheral sales

Contact Info

Lucas Fishbein of the Flatiron School Data Science Program

Locate at 248 Address Street, New York NY 10007

Github Repository: https://github.com/LayFish21/Microsoft_Movie_Analysis

FishbeinLucas@gmail.com

(129) 921-1234