

MicroSoft goes into movie making

HOW A COMPANY SQUEEZES INTO "MAKING MOVIE BUSINESS" AND GETS SUCCESSFUL

FLATIRON SCHOOL

Phase 1 Final Project

Authors: Elena Kazakova

Cohort: DS02222021

Instructor: James Irving



Summary

2

This project analyzes the three publically available online databases, IMDb, TN, and BOM.

- IMDb (an acronym for Internet Movie Database) is an online database of information related to films, television programs, home videos, video games, and streaming content online.
- The Numbers (TN) is a film industry data website that tracks box office revenue in a systematic, algorithmic way.
- Box Office Mojo (BOM) is an American website that tracks box-office revenue in a systematic, algorithmic way.

Descriptive analysis suggests that a company looking into joining a lucrative movie business needs to partner with the studios producing the highest ROIs, and invest in making movies in Horror, Mystery, and Thriller genres, and carefully plan the release date of their products.



Outline of the presentation

3

- Business Problem
- Data
- Methods
- Results
- Conclusions



Business Problem

4

Microsoft sees all the big companies creating original video content, and they want to get in on the fun. They have decided to create a new movie studio, but they don't know anything about producing movies. They need recommendations on how to minimize the risk to their investment and maximize the profitability of their future products. Though this project's scope is limited and further in-depth analysis should be considered, the results of the analysis reveal several valuable recommendations.



Data

5

- Data used in the project was previously imported from IMDb, TN, and BOM databases.
- Additional data from Rotten Tomatoes (RT) and The Movie Database (TMdb) were explored and initially considered, and it has been decided not to use it in the analysis. This data might be regarded for future exploration.
- To help make this decision, analysis of the content in all the sources available was performed, a data diagram of all the tables was created, and as a result, only tables with movie financial data and easily joined tables have been chosen for the analysis.
- The decision has been made to limit the analysis considering only data for movies released within the last decade to minimize the influence of audience changes over time



Data (continued)

6

- The CSV zipped files were unzipped, and the data was imported into an SQLite database.
- To investigate in more detail data available in all of the available tables, an Entity Relation Diagram was created with LucidChart online application.
- Data in the tables were reviewed for duplicates and NULL values. NULL values in categorical columns were replaced with a "Missing" category where appropriate.
- No duplicates were found
- Date values saved as strings were converted by DateTime pandas method to separate columns as year, month, quarter
- Records for movies released before 2010 and after 2019 were deleted from the tables
- Three tables with financial data, studio and genres data to be used in the analysis process were joined using movie titles and year of release



Methods

7

A descriptive analysis of the data was performed. The analysis was intended to answer four main questions:

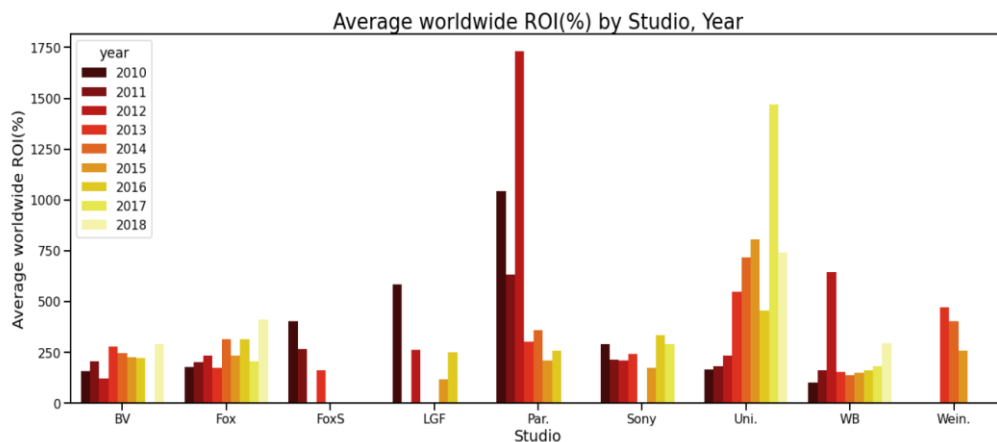
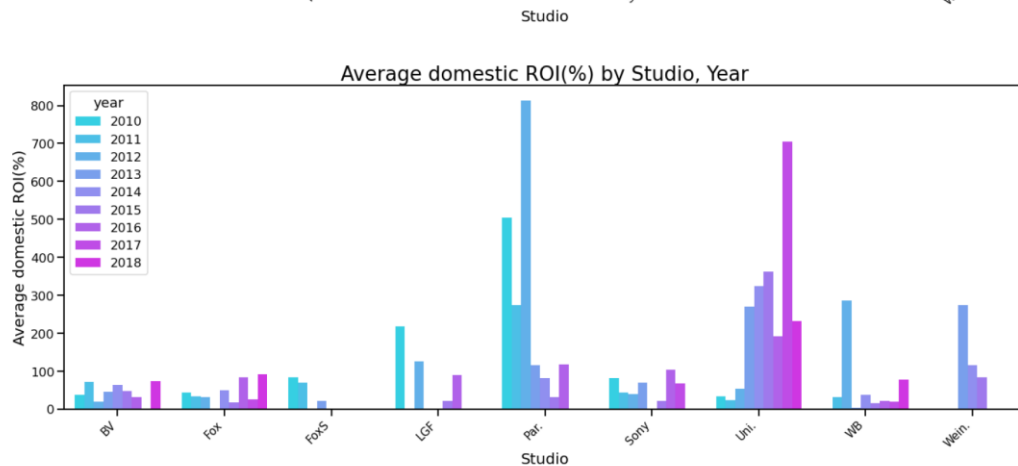
- What studios are most successful in the movie production business?
- Does the runtime of a movie affect the movie's profitability?
- How does the timing of a movie release affect its' profitability?
- What movie genres are most profitable considering Return of Investment measurement?



Methods (continued)

8

Descriptive analysis of studio profitability data



The analysis indicates that based on average ROI values

- Universal Studios
- Paramount Pictures
- The Weinstein Company
- Lions Gate Films Corporation studios

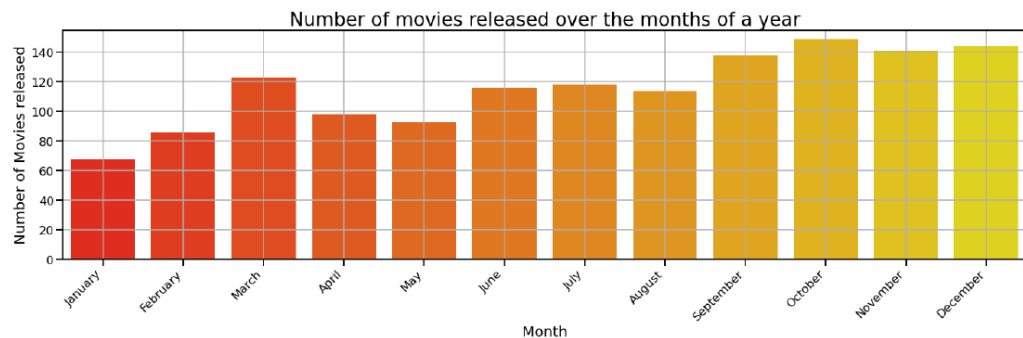
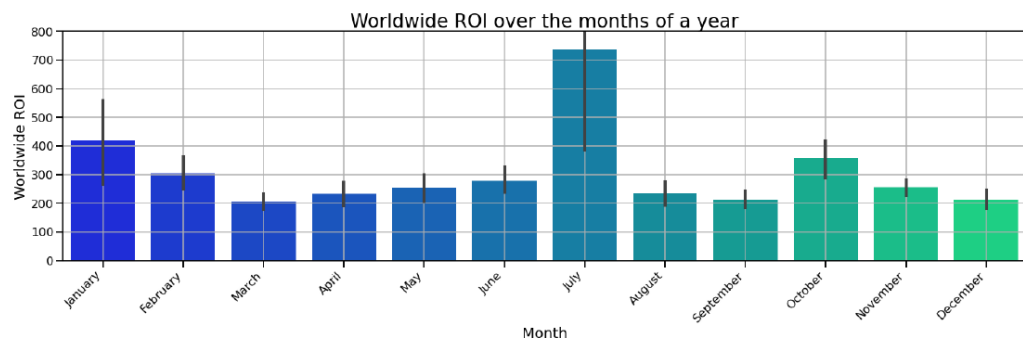
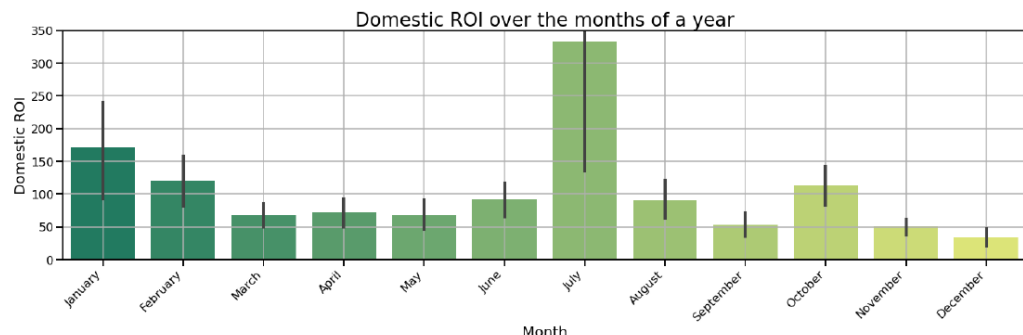
have been the most successful studios over the course of the last nine years



Methods (continued)

9

Descriptive analysis of month of release/ROI correlation



The analysis indicates that based on average ROI values over the course of a year

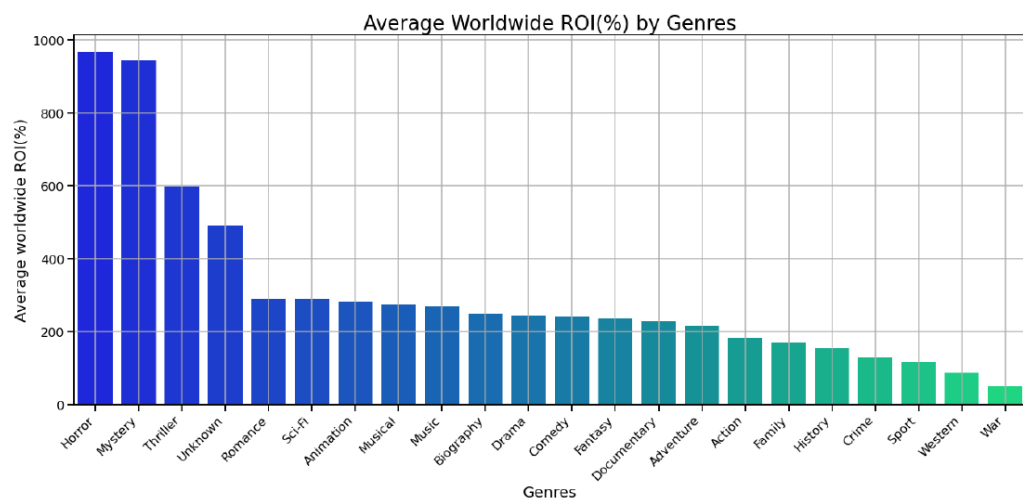
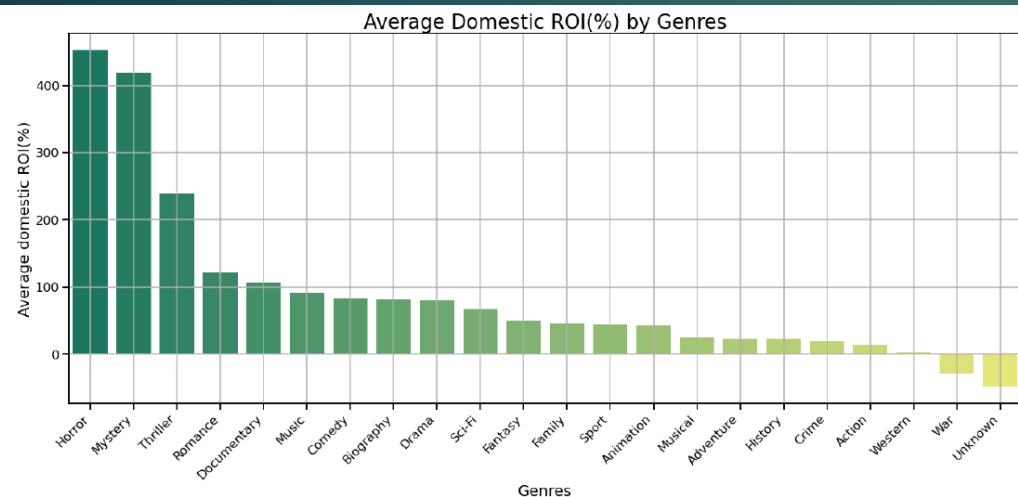
- There is a negative correlation between the number of movies released and the ROI (domestic and worldwide).
- The negative correlation between the number of movies released over a time period suggests that the customer should consider this factor when planning a release of their movie.
- The only exception is a month of July, an outlier among other months of a year. It seems that no matter how many movies are in the theaters, a movie is going to be more profitable than in any other months of a year.



Methods (continued)

10

Descriptive analysis of genre effect on movies profitability



The analysis indicates that based on average ROI values

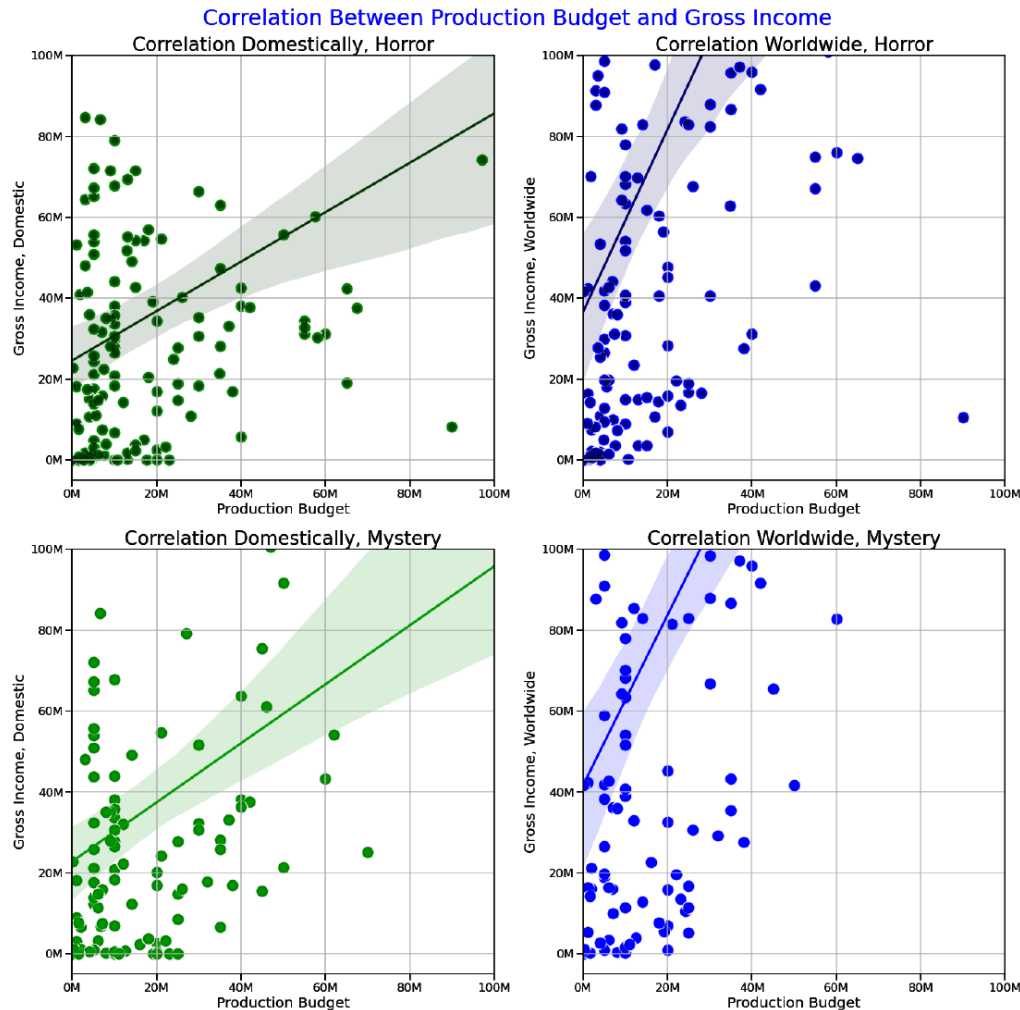
- Three most profitable genres are Horror, Mystery, and Thriller (in that order) in financial terms of an average ROI, both domestically and abroad.
- There is a strong presence of the "Unknown" category in the lower plot of average worldwide ROI. That might be in part due to the international practice of categorizing movies differently than in the United States.



Methods (continued)

11

Descriptive analysis of a size of production budget on gross income of a movie



The analysis indicates that based on the correlation between Gross Income and Production costs:

- One can expect a ratio between gross income and production for movies in Horror and Mystery categories with budgets of 20 million and lower close to 2.0 domestically and 3.5 internationally. However, as a budget grows, the return drops.
- Most movies in these two categories have lower budgets

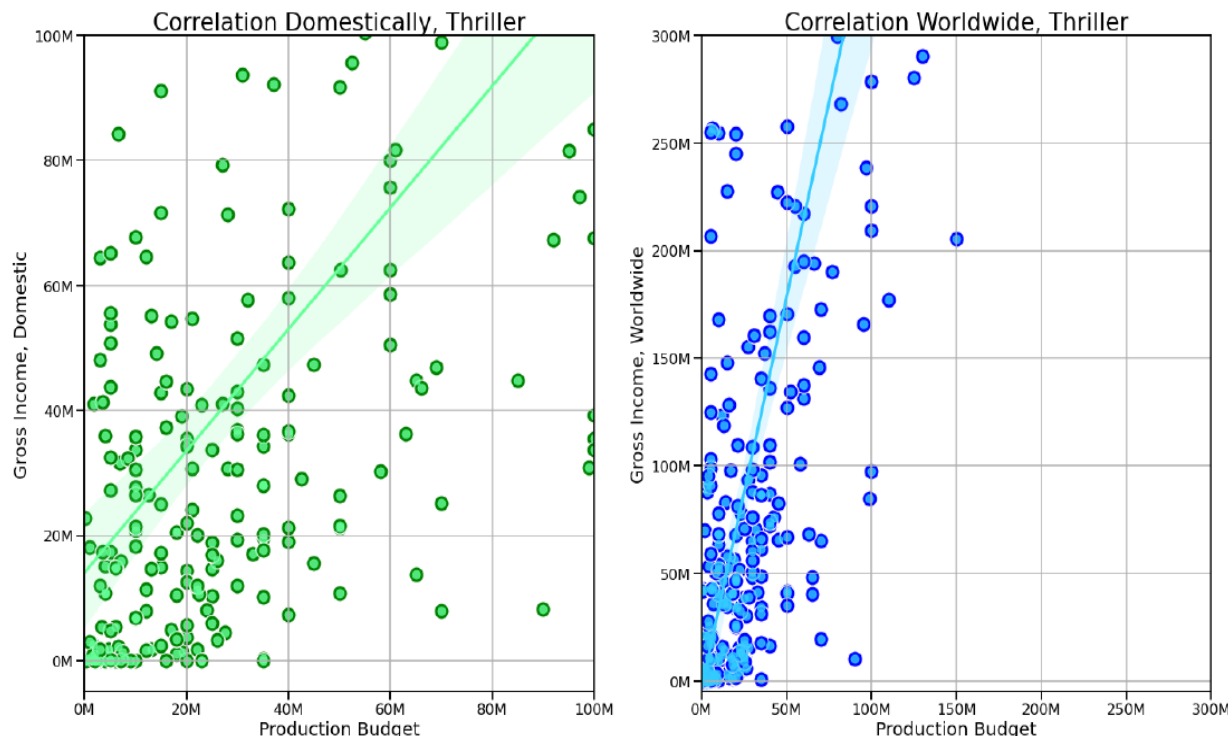


Methods (continued)

12

Descriptive analysis of a size of production budget on gross income of a movie

Correlation Between Production Budget and Gross Income



The analysis indicates that based on the correlation between Gross Income and Production costs:

- Thriller movies do not perform quite as well as horror and mystery movies domestically, i.e., the regression model suggests that 20 million investment would generate slightly above 30 million in gross income.
- Internationally, thriller movies tend to do much better, and an estimate of an average ratio of gross income to production budget ratio is about 3.5.
- The customer's recommendation is to release a movie both domestically and internationally to maximize the return on investment.



Conclusions

13

Recommendations to the customer:

- Either partner with Universal Studios, Paramount Pictures, The Weinstein Company, and Lions Gate Films Corporation studios (in that order) or invest in investigating their business practices and following their example.
- Carefully plan the timing of releasing their movies because ROI tends to be higher in the time periods when fewer movies are available to the viewers. The only exception is a month of July, an outlier among other months of a year.
- Invest in the three most profitable genres in terms of ROI, Horror, Mystery, and Thriller (in that order), both domestically and abroad. Though their production budgets tend to be lower than movies in other genres, however, the ratio between their gross incomes to the production costs is higher.



Conclusions (continued)

14

Additional analysis suggested:

- Update data available for analysis by either creating APIs with the sources or web scraping their sites
- Investigate the effect of a choice of directors/writers on profitability of a movie using additional tables in the database
- Use Rotten Tomatoes tables to analyze the correlations between the profitability of movies and their critics and viewers rating
- Use Rotten Tomatoes database tables to confirm the results and the conclusions of this project



Thank you!

15

Email: e.v.kazakova@gmail.com

GitHub: @sealaurel

LinkedIn: <https://www.linkedin.com/in/elena-v-kazakova/>

