



Analysis of the Movie Industry

Module 1 - Final Project
Online-DS-PT-032020
Kevin Machine



A brief background on this presentation:

Microsoft wants to create a film studio to produce original movies, so as to compete with the big film companies. To this end, they've hired us to help them better understand the movie industry.



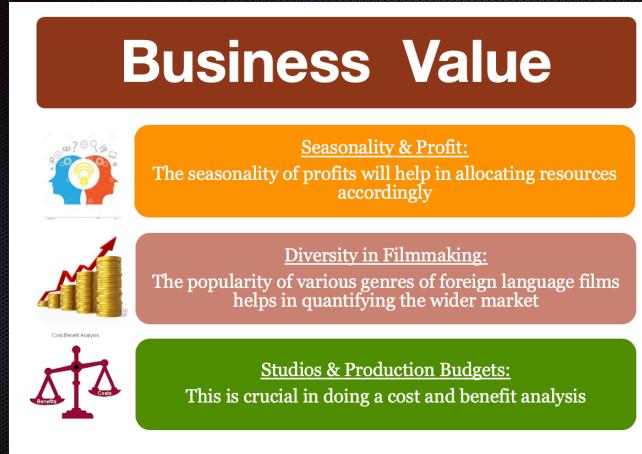
Problem Statement:

What types of films are currently doing the best at the box office?

2

Our team is charged with doing a data analysis, to explore what type of films are do the best at the box office.

Business Value



Seasonality & Profit:
The seasonality of profits will help in allocating resources accordingly

Diversity in Filmmaking:
The popularity of various genres of foreign language films helps in quantifying the wider market

Studios & Production Budgets:
This is crucial in doing a cost and benefit analysis

3

We will need to translate our findings into actionable insights that the CEO can use when deciding what type of films they should be creating.

We believe by exploring these aspects of the movie industry will add most value to the decision making process.

1. **Seasonality & Profit:** This will help in a timely allocation of resources.
2. **Diversity in Filmmaking:** This helps in quantifying the diverse and growth market of the industry.
3. **Studios & Production Budgets:** This is crucial in doing a cost and benefit analysis of creating a certain type of studio.



Methodology



Used dataset imported from IMDB, Box Office Mojo, Rotten Tomatoes and TheMovieDB.org



Created new columns and merged columns in dataset to get insight on profits, genres, movie language and release months



Review dataset to see historical studio information to help in determining number of movies produced and at what cost

In this project, we had access to several databases that needed to be analyzed. The datasets included a wide range of related data imported from The Box Office Mojo, IMDB, TMDB, The Numbers and Rotten Tomatoes.

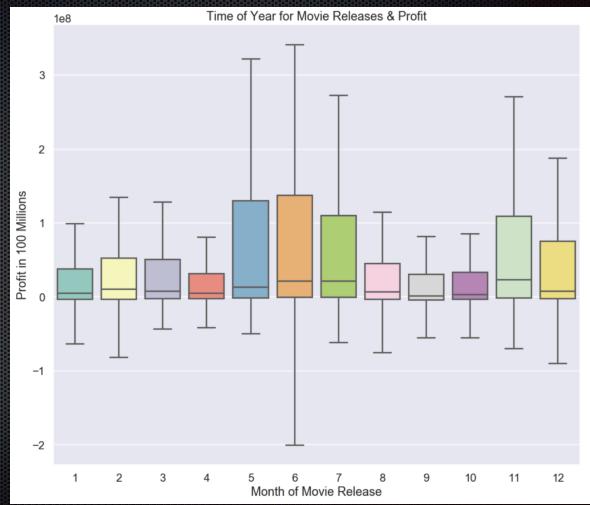
Given the breadth of the data already provided, I figured we had enough data and there was no need for scraping additional data from the websites.

- I created new columns and merged columns in dataset to get insight on profits, genres, movie language and release months.
- I reviewed the dataset to see historical studio information to help in determining the number of movies produced by each studio and at what cost.

Findings

- Seasonality of movie profits

Source for Budget, Gross Revenue and Release Date: www.the-numbers.com
Source for Original Language: www.themoviedb.org

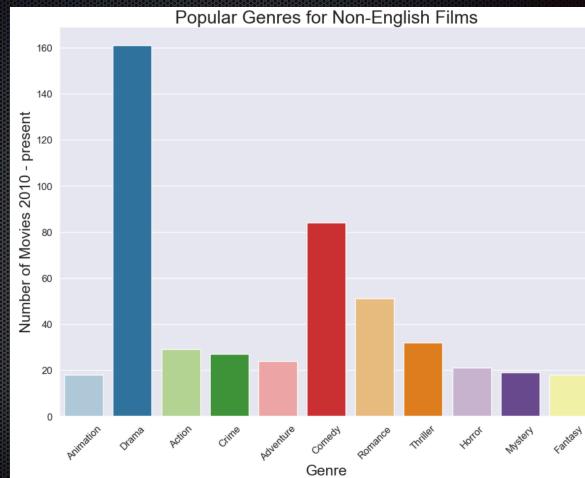


From our box plot, we see there is dual seasonality of higher profits during early summer and at the end of the year. From our data, the highest profit in summer was realized during the month of June and in during the end of the year in November. It is important to point out that the month of June also had instances of the largest losses.

Findings

- Popular genres for foreign language films

*Present - As of end 2019
Source for Genres and Movies: www.imdb.com

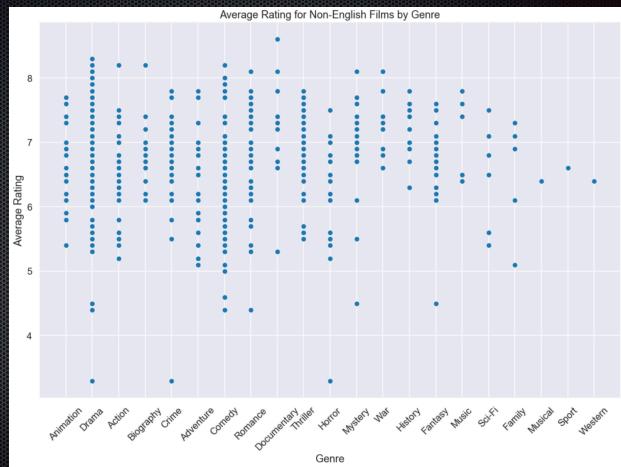


On the foreign language films, Non-English Drama, Comedy and Romance are likely to attract more viewers compared with other movie genres.

Findings

- Average rating for various genre of foreign language films

Source for Genres and Movies: www.imdb.com



This scatter plot shows the average voter rating in each genre.

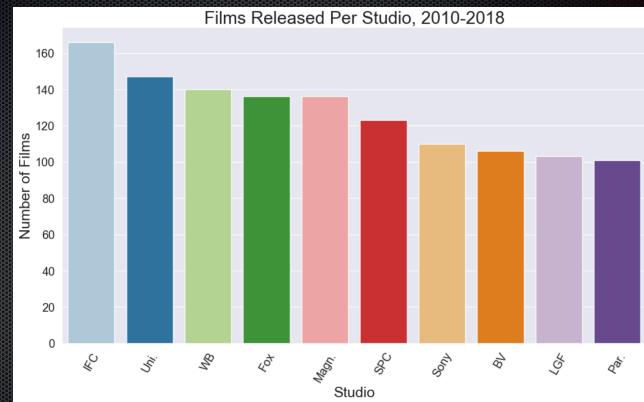
Our top 3 genres (Drama, Comedy, Romance) didn't make the top 5 average voter ratings by genre, with War films having the highest vote average at 7.3 and Horror films having the lowest at 6.2. The rankings of our top three genres of foreign language films, out of a complete list of 21 genres.

- 6 Drama 6.9
- 9 Romance 6.7
- 19 Comedy 6.4

Findings

- Number of films released by top 10 film studios

Source for movie studios: www.the-numbers.com

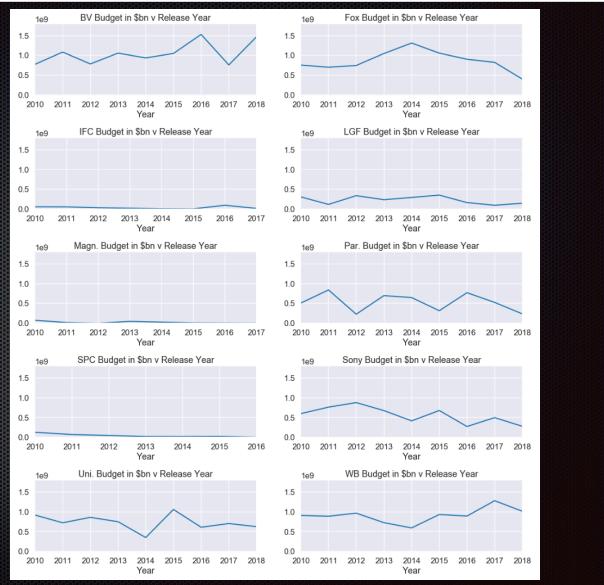


We found that some independent movie studios such as IFC, Magnolia Studios and Sony Pictures Classics, release large volume of movies, almost equal to the big film companies such as Universal, Fox and Warner Brothers Studios.

Findings

- Movie production budgets of major and acquire/distributor film studios

Source for movie budgets www.the-numbers.com



The major film studios (such as Warner Bros, Universal, Paramount) generally have bigger annual production budgets (of approximately \$1bn) compared with the independent distributors that acquire and distribute films. This is stark revelation as some of these independents release as many movies as the major film studios.



Business Recommendations

1. We recommend entering the movie industry through an [acquisition/distribution](#) model

This model shows that more top movies can also be produced at a lower budget.



Business Recommendations

2. We recommend going for **Drama** or **Comedy** genre, if the business opts for diversity in films



Business Recommendations

3. We recommend planning to release movies towards the
summer or at the **end of the year**.



Future Work

1. Which advertising medium is more effective?
2. What is the ideal length of a popular movie
3. What are the profit margins of the top 10 film studios?



Future Work

Which advertising medium is the most effective ?

With the typical marketing budget being 50% of production, by reducing the advertising cost, a movie can increase its profitability:*

- Machine learning and Algorithmic recommendations based on past viewership (for streaming media)
- In-Theater movie previews
- Social Media

*Source: Howstuffworks - Entertainment: <https://entertainment.howstuffworks.com/movie-cost1.htm>

1. With the typical marketing budget being 50% of production, by reducing the advertising cost, a movie can increase its profitability:

- Machine learning and Algorithmic recommendations based on past viewership (for streaming media)
- In-Theater movie previews
- Social Media



Future Work

What is the ideal length of a popular movie ?

Referred to as “runtime”, the popularity of movies with various runtimes should be explored.

From the data already collected, this insight will be crucial in determining the production budget, as a longer movie would potentially cost more to produce.

2. Referred to as “runtime”, the popularity of movies with various runtimes should be explored. From the data already collected, this insight will be crucial in determining the production budget, as a longer movie would potentially cost more to produce.



Future Work

What are the profit margins of the top 10 film studios?

Although having big budgets might mean larger gross sales, and more movies, a further analysis from our data can be done to see how much **profit** did each **top studio make** and what is the **ideal number of movies** to attain a target profit.

3. Although having big budgets might mean large gross sales and more movies, an analysis from the gather data can be easily done to see how much **profit** did each top studio make and what is the ideal number of movies to attain a target profit.

Thank you for your attention

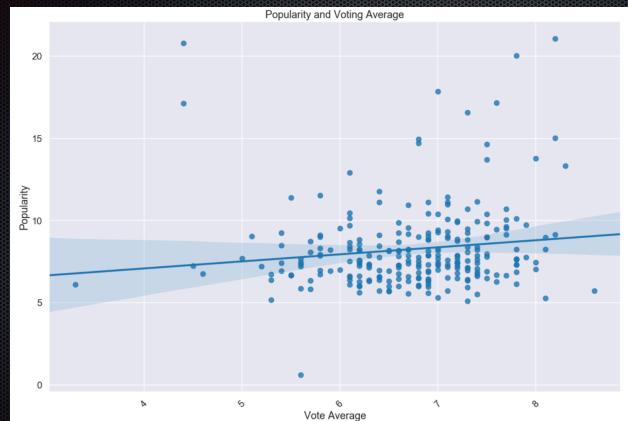
Questions ?

Comments ?

Feedback ?



Appendix 1 — Popularity & Voting Average



The scatter plot shows the correlation between vote average and popularity of movies.



Appendix 1 — Correlation between Popularity & Voting Average

	start_year	id	popularity	vote_average	vote_count	Animation
start_year	1.000000	0.782352	0.059249	0.013933	-0.070101	
id	0.782352	1.000000	0.192613	0.048509	-0.047919	
popularity	0.059249	0.192613	1.000000	0.132982	0.384245	
vote_average	0.013933	0.048509	0.132982	1.000000	0.246060	
vote_count	-0.070101	-0.047919	0.384245	0.246060	1.000000	

Out[69]:

We confirmed that there is no correlation between vote average and popularity by running the data through the correlation function.

The data on Popularity (from TMDB), is based on number of votes, views, users adding a title to their watchlist or favorites, as well as previous day's popularity.

As you can see, the correlation is 0.132982, so there is no correlation.