Hollywood's Most Profitable Stories Dataset

Amiya Renavikar

Data Source Information

Data Source: This dataset is widely available online however, the one I chose is a smaller subset from a Github repository source

Dataset Background: The major film studios of Hollywood are the primary source of the most commercially successful and most ticket selling movies in the world. Moreover, many of Hollywood's highest-grossing movies have generated more box-office revenue and ticket sales outside the United States than films made elsewhere (Wikipedia). Ultimately, we all love movies!

Each row in the dataset contains details of a single movie that was released from 2007-2012. The data includes information about the movie name, lead studio, rotten tomato rating, audience score, genre, worldwide gross, profitability and year of release.

Total Records: 640

Dataset Link:
https://github.com/gchan/hollywood-budgets/blob/master/public/data/Most%20Profitable%20Hollywood%20Stories.csv

Initial Questions for EDA

Since I chose a subset of the original dataset to work with, I was able to curate detailed questions to ask for exploratory data analysis.

1

Which movie genres are highly rated by the audience vs the critics?

2

What was the profitability trend of movie genres over the years?
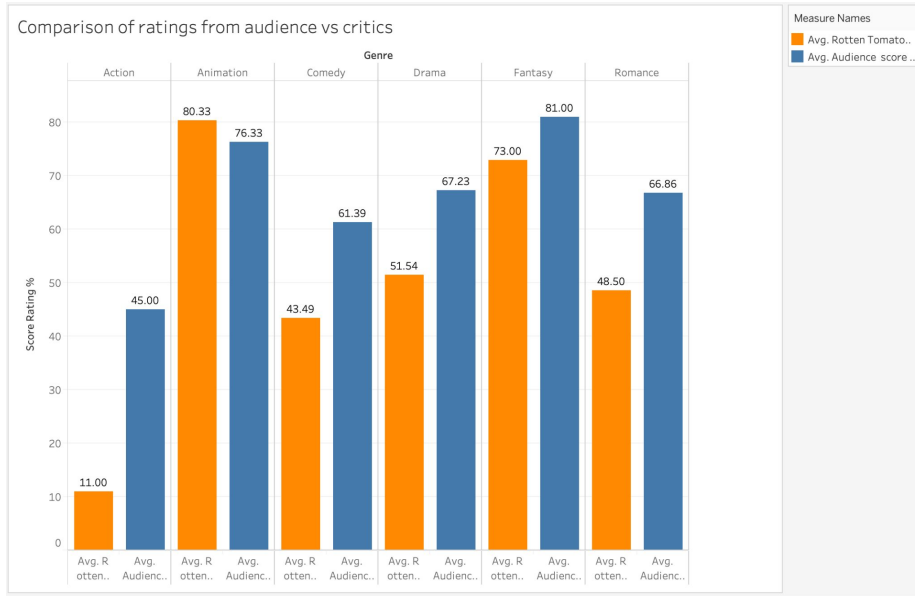
3

What were the top 20 most profitable movies in this time period?

Data Cleaning & Transformation

As mentioned in Slide 2, I used a subset of this data to perform my analysis. I performed the initial EDA by examining the data using Tableau Prep. Once I had identified null and/or missing values, I dropped them from my dataset. I used Tableau prep for the same. Then I loaded the data into Tableau and performed a visual analysis of the data.

Audience vs Critic Ratings of Movie Genres

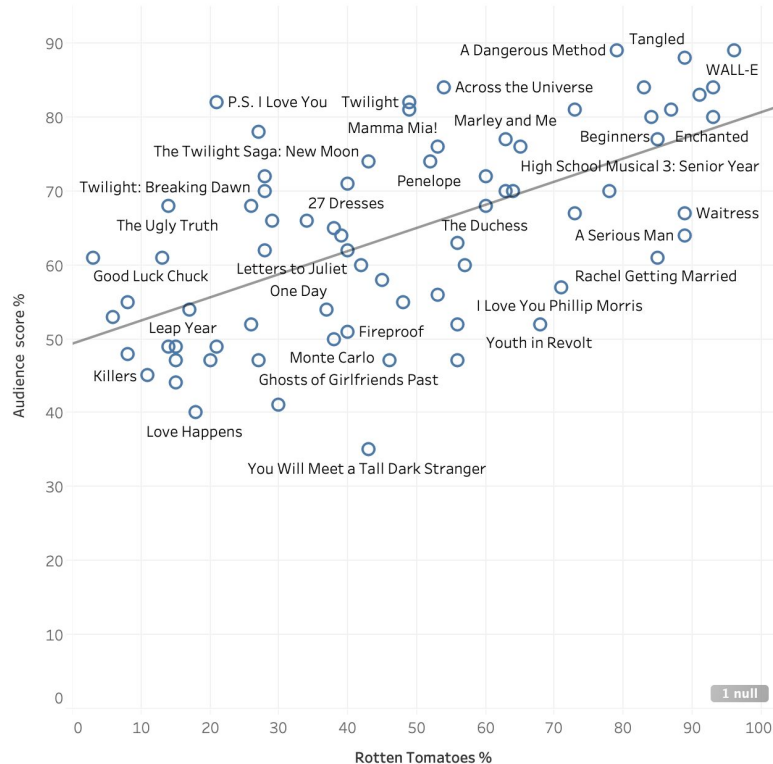Comparison of ratings from audience vs critics

With any budgeted film, it is important to understand the selection process of what kind of film it will be perceived as and whether it will do well in the box office. Moreover, it is important for filmmakers to know which genres would do well with the audience and revenue. Reviews and ratings from the audience as well as critics play a huge role in influencing viewers choice in entertainment. Hence, in this graph, I have plotted the % score of rating for Rotten Tomatoes and Audience score for each genre.

We can see that Fantasy and Animation genres have high ratings from both the audience and critics where as Action surprisingly has the lowest rating scores.
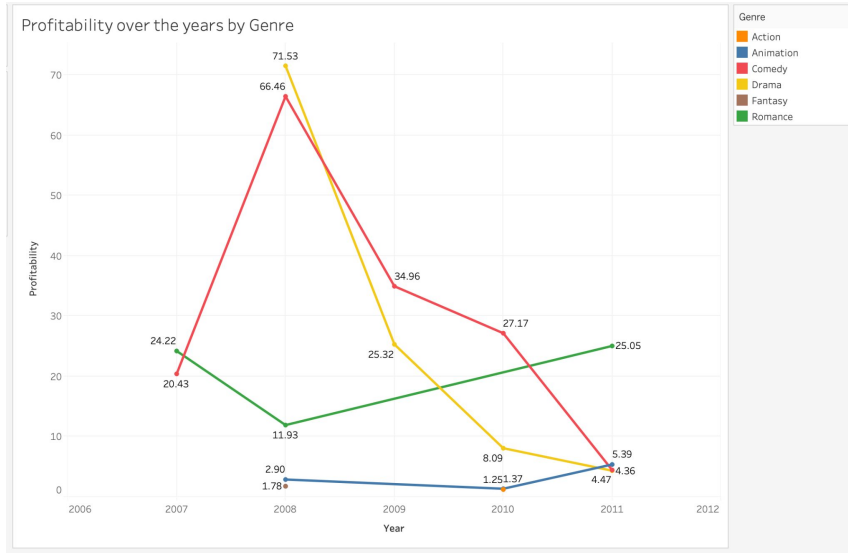
Audience vs Critic Ratings of Movies

Audience vs Critics Ratings for Films

This graph shows the correlation between audience score % and the Rotten tomatoes % ratings for all movies made during 2007-2012. There was a higher correlation between these two variables for movies such as Enchanted, Beginners, The Duchess, Letters to Juliet and Leap Year, which means that these movies were equally rated about the same by both the audience as well as critics.

Wall-E received the highest ratings from both the audience as well as the critics.
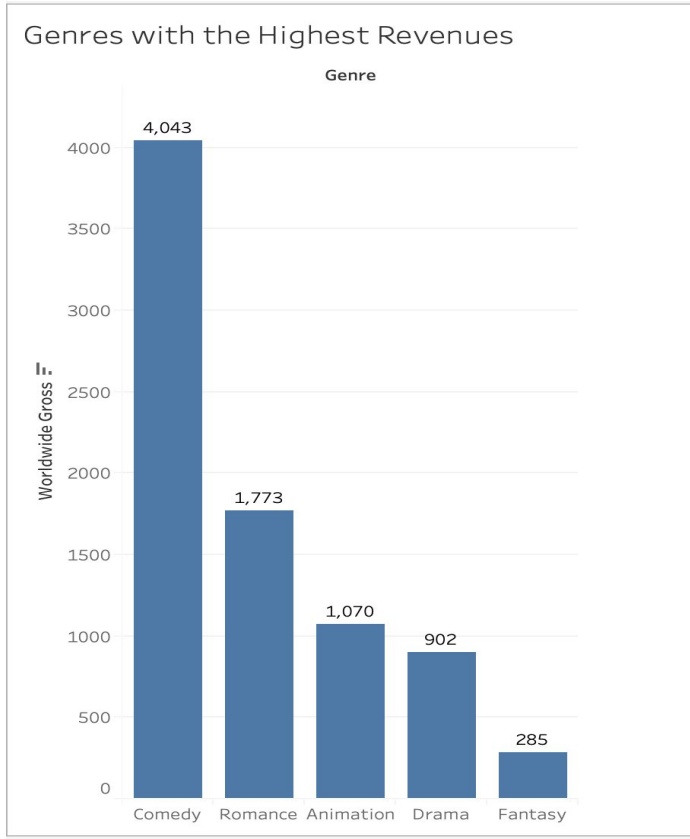
Movie Genre Profitability Trend

Profitability over the years by Genre

In order to further analyze which movie genres would do well at the box office, I plotted the profitability of each of the genres of the movies made between 2007-2012.

By looking at the visualization, Drama was the most profitable genre in 2008 however it's profitability decreased significantly in the following year. The trend for Comedy genre also declined after peaking in 2008, making romance the most profitable genre in 2012. Thus, we can see that the movie genre profitability changed significantly during this time period.

Genres with the Highest Overall Revenues

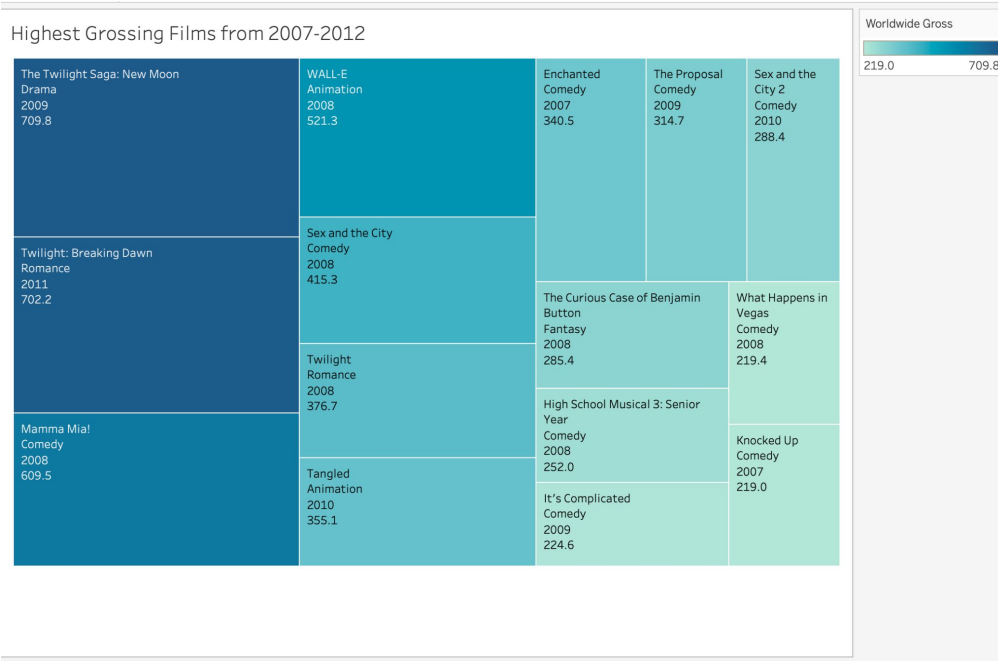## Genres with the Highest Revenues

Genre

In this data viz, I plotted the genres that generated the highest overall revenue from 2007-2012.

Based on the graph, Comedy seems to have generated the most amount of revenue at $4043M and Fantasy is the lowest at $285M.

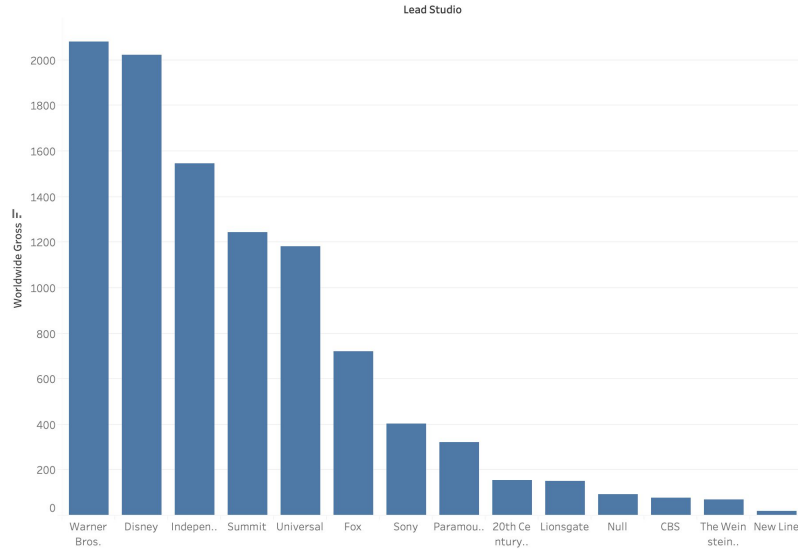This could be a deciding factor is choosing which genre to invest in to make high grossing films.

Highest Worldwide grossing Movies from 2007-2012

Highest Grossing Films from 2007-2012

Worldwide Gross
219.0   709.8

The Twilight Saga: New Moon Drama 2009 709.8	WALL-E Animation 2008 521.3	Enchanted Comedy 2007 340.5	The Proposal Comedy 2009 314.7	Sex and the City 2 Comedy 2010 288.4
Twilight: Breaking Dawn Romance 2011 702.2	Sex and the City Comedy 2008 415.3	The Curious Case of Benjamin Button Fantasy 2008 285.4		What Happens in Vegas Comedy 2008 219.4
	Twilight Romance 2008 376.7	High School Musical 3: Senior Year Comedy 2008 252.0		
Mamma Mia! Comedy 2008 609.5	Tangled Animation 2010 355.1	It's Complicated Comedy 2009 224.6		Knocked Up Comedy 2007 219.0

In order to better understand which films generated the most amount of revenue during this period, I created a heat map of the the top 15 highest worldwide grossing movies. Based on the data analysis, The Twilight Saga: New Moon was the highest grossing movies with $709.8M in revenue, followed by Twilight: Breaking Dawn. This definitely explains the slow increase in Romance genre profitability in the previous graph.
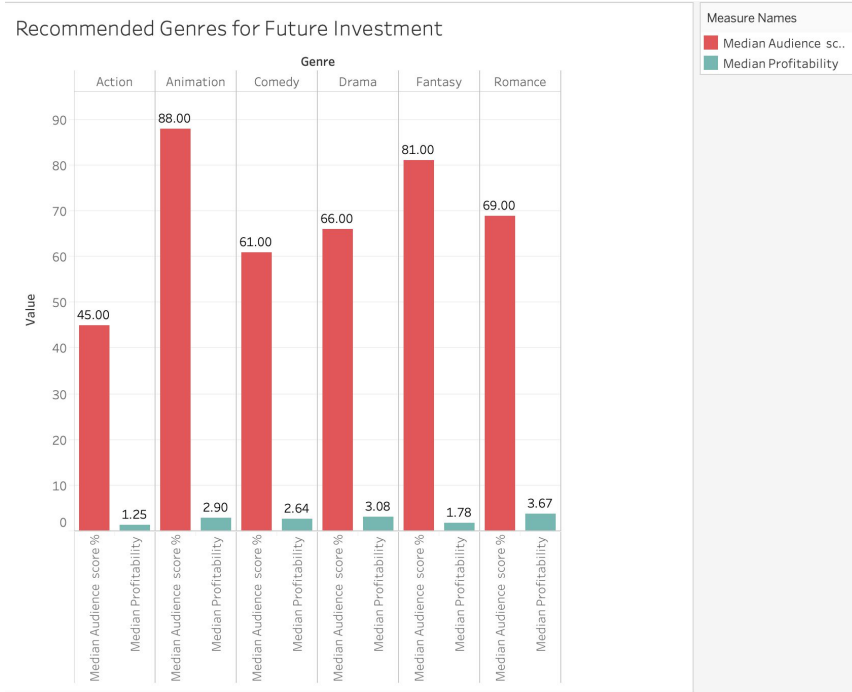
One of the top 3 films is Mamma Mia! which is included in the Comedy genre, which definitely explains Comedy being one of the highest grossing genres, followed by Romance.

WorldWide Gross Per Studio

Lead Studio

Worldwide Gross

In this viz, I have plotted the lead studios by their world-wide gross revenue. Based on the plot, we can see that Warner Bros and Disney earned $2000M or more during this 5 year time period. This means that they probably made good business decisions in terms of choosing the right films and genres to invest in.

Recommended Genres for Future Investments

Recommended Genres for Future Investment

Measure Names
■ Median Audience sc..
■ Median Profitability

Knowing which genres to invest in is an important factor to making a successful, high grossing film. Taking a look at the plot, I have plotted the median of Audience score and the median profitability for each genre. Based on the analysis, Romance seems to be the best genre for studios to invest in with a good balance of high Audience scores as well a decent profitability. On the other hand, Anime has average audience rating but has a lower profitability, so it is not a recommended genre to invest in.
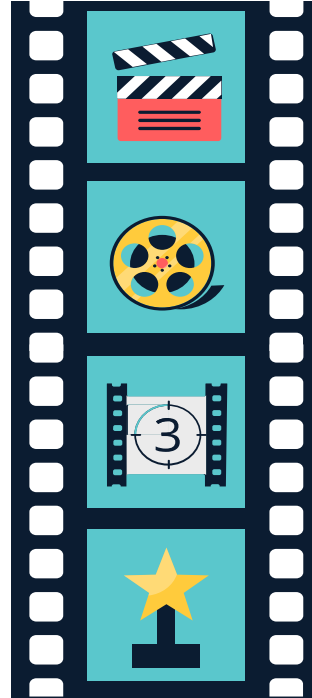
Key Insights & Future Scope

1 **WALL-E was the highest rated film by both audience & critics**

2 **Comedy was by far the highest revenue generating movie genre**

3 **The Twilight Saga series generated the highest revenues world-wide**

Warner Bros made the highest revenues during 2007-2012 **4**

Romance is the best genre to invest in for future investments **5**

As for future scope, it would be interesting to analyze budgets allotted to each film to get a better understanding of how much money is being invested in each film or genre.