Bret Peterson

MIS-505

Dr. Sibona

Final Project: Profits in the Movie Industry 2007 – 2011

In starting this project, the first challenge I crossed was selecting a data set with a topic that I was not only interested in, but also thought that I could develop an entertaining story with data about. After scrolling through multiple data websites, the data set “Hollywood’s Most Profitable Stories” continued to stick out in my head. Being that many people can relate to the films being released into theaters, I thought this would be an interesting topic to do some analysis on. The dataset, “Hollywood’s Most Profitable Films” was easy to download and connected cleanly with Tableau. It also contained several variables I thought would be good to conduct analysis on like; Genre, Lead Studio, Worldwide Gross, Profitability, and both Audience and Critics review ratings on the film. Once I began to work with the data and try to tell the story about profits within the film industry, I quickly realized that Profitability and Profits are not the same measure and as such, was required to hunt down a second data set to compliment the first one.

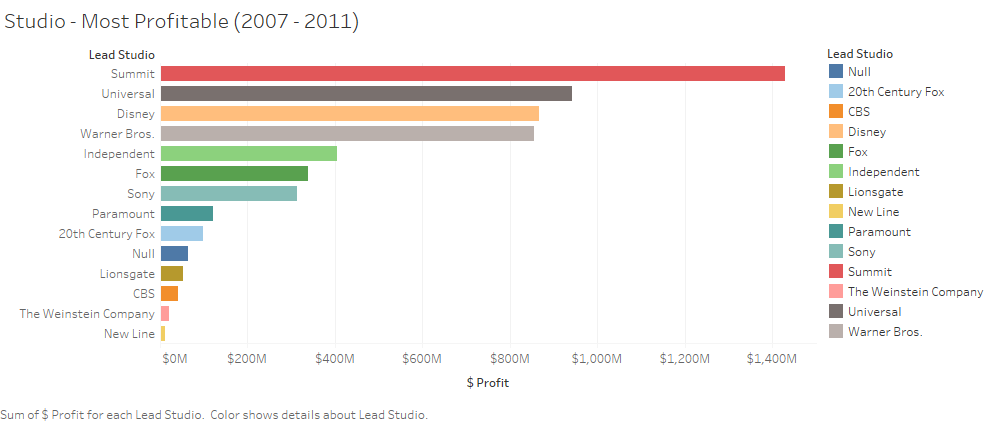
After searching several of the data websites that were provided for the project, I finally came across a dataset that both matched the films and years I was analyzing and provided much needed extra data about the films like Budget. Although the dataset provided the missing link for the data I was searching for, it was not nearly as easy to download and bring into Tableau to utilize. The biggest problem that I encountered with the data was that it was condensed into a ZIP 7 file format, something I had rarely worked with. After doing some research, I was able to decompress the .CSV file, “movie\_metadata” and then connect it with Tableau. Connecting the data to Tableau presented some problems as well in that, the column headers for the two different files were different. In order to cleanly connect the new data to the original dataset, I went with an inner-join connecting on “Film” from the original dataset and “Movie Title” from the second dataset.

Once the datasets were connected within Tableau, I began to create my visualization story of profits within the movie industry. I wanted to create a story to an audience who may be interested in investing in the film industry and show them:  
**A) What lead studios have had the most profits in the past? Does the size of the studio have any effect on the amount of profits?**

**B) What genres of film, if any, have had the most success with bringing in profit?**

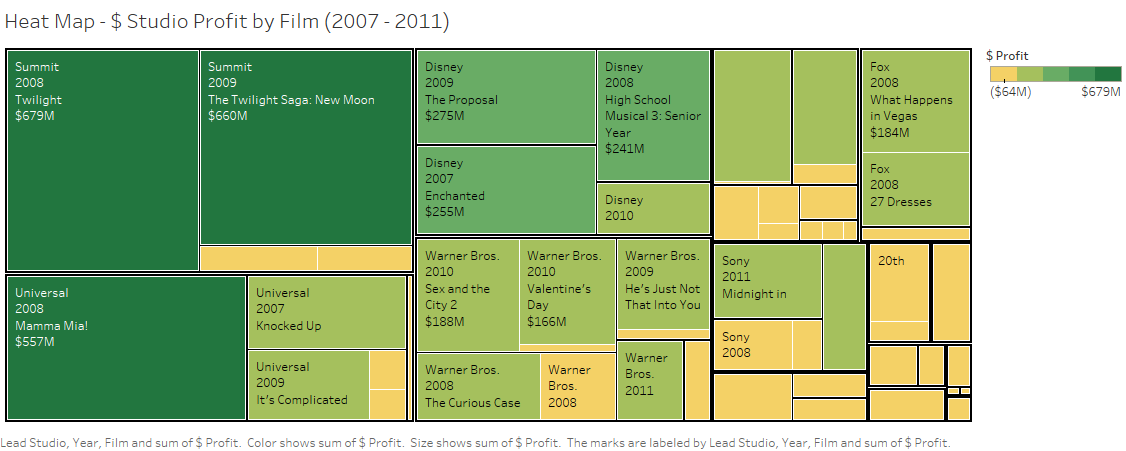
**C) Does the audience or critic rating of a film have any impact on the overall profits, and if so which has a larger impact?**

My first dashboard, “Most Profitable Studio” begins the story by creating a bar graph of profit dollars in millions by Lead Studio and sorts them in descending order.



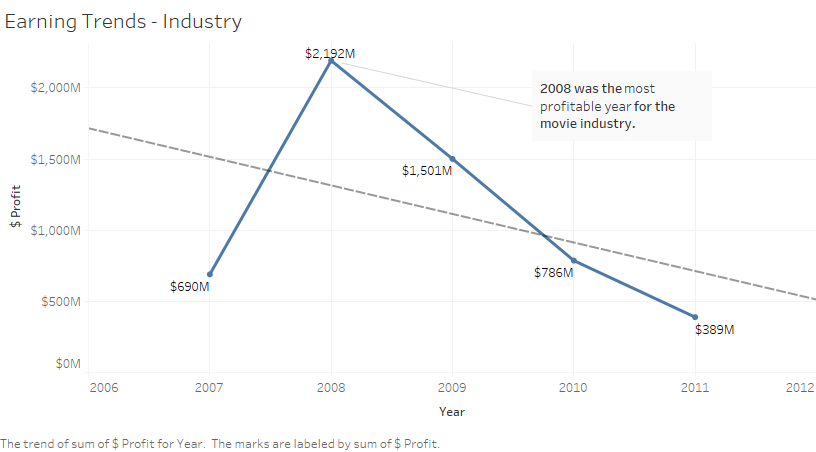
This graphic provided a simple, but clear visual as to what studios have earned the most profit dollars during the time period 2007 – 2011. (When analyzing the graph, it became evident to me that there were “tiers” to the amount of profits Studios were bringing in. As such, I decided to create different groups to represent the tiers; one for the Top 4 Studios, the Middle Studios, the Bottom Studios, and also one for Independent Studios as they change from film-to-film.)

Although the first chart in the Dashboard provided a clear visual of what studio profits the most, I wanted to provide a more detailed visual of what movies contributed to each Studio’s profit. In order to do so, I developed the following heat chart:



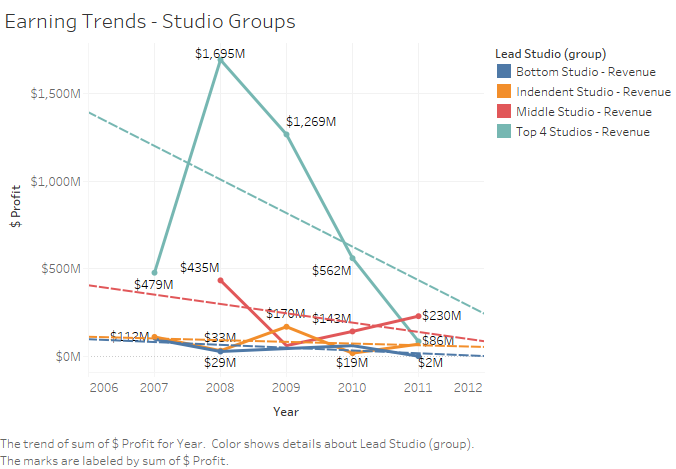
Each of the boxes within the heat chart represent a different Studio within the industry and are sorted in descending order from the overall most profitable studio to the least profitable. To emphasize the amount of dollars added towards profits, I heat mapped the chart using a red-gold-green stepped color scale with dark green representing the most profit dollars.

The next dashboard is used to continue the story of what Studio is the most profitable in the film industry, but is designed to provide insight into the year-over-year trends of the industry as a whole, by the size of the studios grouped, and by the individual studio. Being that time is being reported, I decided to use line graphs for all the charts:



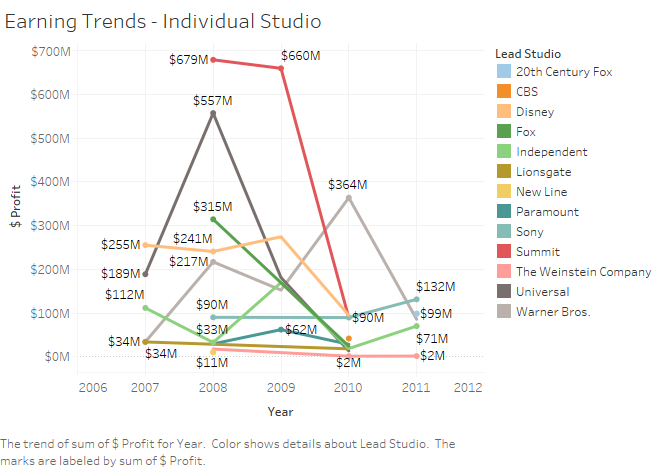
There is a clear downward trend in the amount of profits seen in the movie industry year-over-year, with 2008 marking the peak profits. (We can assume that the 2008 financial crisis would have something to do with this downward trend. Another thing to consider would be the increasing popularity of streaming services like Netflix and Hulu over this time period.)

Although, there was an obvious downward trend in the profits of the movie industry as a whole, I wanted to determine if this was consistent across different studio sizes. To help determine this, I used the groups I had created; Top 4 Studios, Middle Studio, Bottom Studio, and Independent, to track the trend in profits:



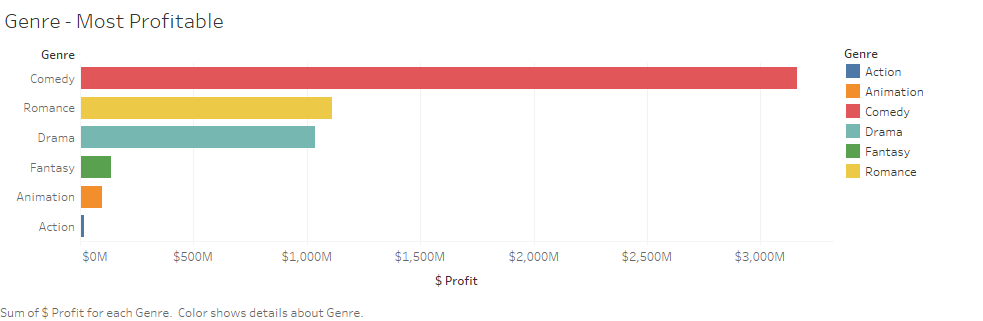
When looking at year-over-year profit trends by studio size groups, the downward trend is still evident; however, it is evident that the big and middle-sized studios are taking a far bigger cut in profits as compared to the smaller and independent studios.

Although looking at the year-over-year profit trends by studio size grouped gave more insight which players are taking the biggest hit; I still wanted to home in further on what individual studios are suffering the most:

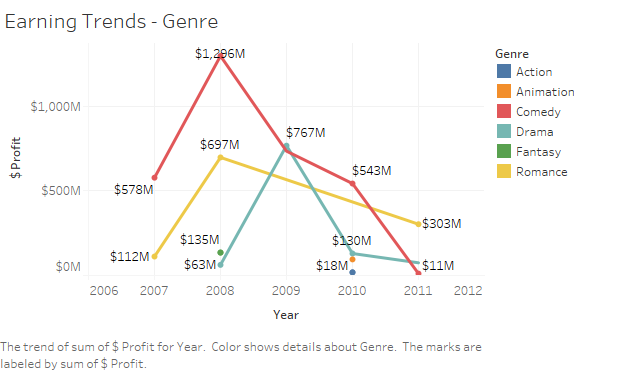


Despite the graph being somewhat crowded, the visual provides further insight into the fact that certain players are taking a bigger cut than others.

After seeing that there was an impact on profits year-over-year based on the size of the studio, I also wanted to provide my audience with some further insight as to what genres were the most profitable and rather or not there were any long term trends. My fourth Dashboard, “Most Profitable Genre” provides insight into these profits:

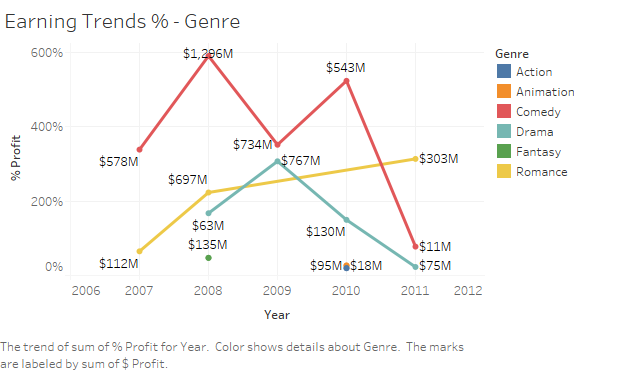


The first chart in the Dashboard was selected to show overall profit dollars by genre. It is clear to the audience that “Comedy” brings in substantially more money than other genres; however, are the profits impacted by the negative trend we saw in the film industry as a whole? In order to answer this question, I looked at year-over-year profit in both actual dollars, as well as, profit percent.



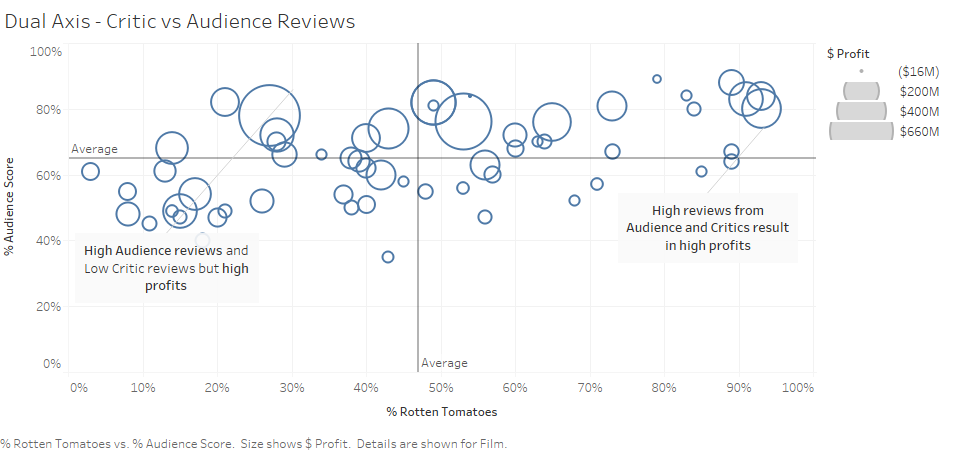
In analyzing the actual dollar profits by genre, no genre is exempt from the negative growth of the industry. There are some genres that have taken a bigger hit than others though. For example, “Comedy” has a much steeper drop in profits as compared to “Romance.”

In order to better show how big of a hit each genre has taken year-over-year, I also wanted to show the profit percent:



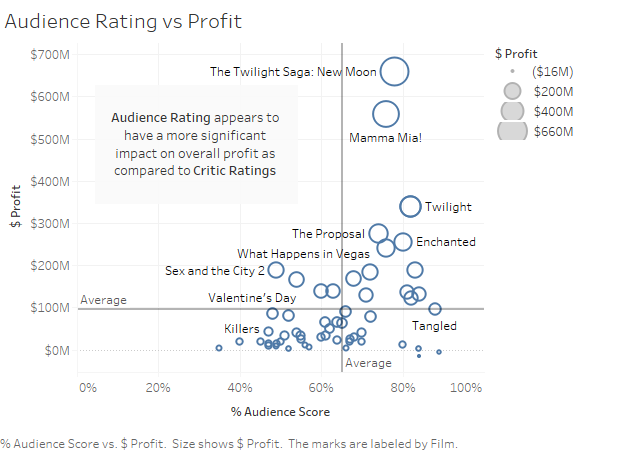
This story is a little less depressing and shows a growth in the “Romance” genre.

After determining that both the size of the studio and the genre of the film had some impact on the profit dollars, I also wanted to analyze the effects of reviews on profits for both audience and critic reviews. My fifth Dashboard, “Ratings vs Profits” looks at both audience and critic reviews together and then focus in on each review type individually:



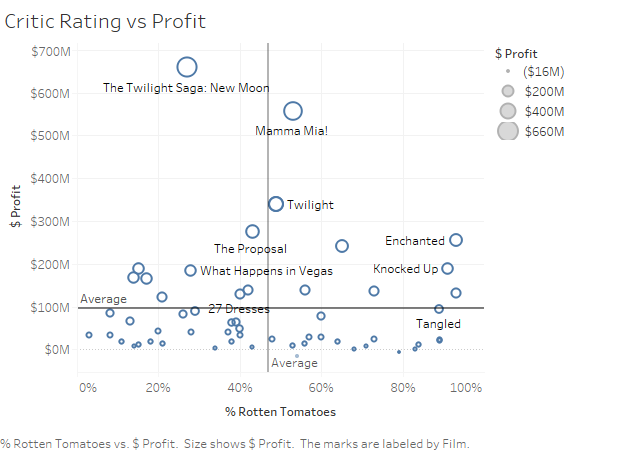
In the above chart, the dual-axis is labeled with “% Audience” on the Y-axis and “% Rotten Tomatoes” (or critic) on the X-axis. The chart also has average lines for both audience and critic reviews, showing that audiences tend to be less harsh in their reviews. The circles on the chart are used to represent individual films and the size of the circle represents the amount of profit earned (the bigger the circle, the bigger the profit.) In analyzing the results, when both Audiences and Critics give the film high reviews, the profits are going to be high (reference the top right quadrant of chart). It is also evident that Audience Reviews being high has a much more significant impact on profits as compared to Critics Reviews (reference top left of above chart.)

In order to provide more clear insight on Audience/Critic Reviews on the profit dollars, the last two charts look at each individually.



Looking at the impact on audience score and profit dollars, there is a clear correlation between a higher rated film (being further right on the x-axis) and more profit dollars (a bigger circle).

There seems to be a different impact when analyzing Critic Reviews:



When looking at Critic Reviews and Profit dollars, there doesn’t seem to be as strong as a correlation as there was with Audience Reviews. This is evident by the small size of the circles despite the high percentage review.

After analyzing my questions and creating my Dashboards, there were still a variety of charts that I excluded from my results. Most of these charts were left out due to their size and inability to digest quickly. I wanted the audience to be able to look at my graphs with very little analytical knowledge and be able to see a story. I do feel like the audience charts became a little confusing to your average user; but the insight the provided was too valuable to leave out.