Dabkowski, Michael
John, Sarah
Villarroel, Valentina
Daaleman, Rebecca

Movie Magic: Makings of a Successful Films

08/16/2023

Our project focuses on the question: what factors help make a successful film? For the scope of this project, we define "success" as awards won and nominations accrued. Utilizing a dataset called "Movie Database Alternative", our group has randomly selected 500 movies that premiered after the year 2000. The database includes information such as movie rating, release dates, genre, director, actors, award nominations & wins, as well as public data such as box office earnings, and reviews from sites such as Rotten Tomatoes, Metacritic, and IMDb. The awards include the well-known awards like the Oscars, but also include lesser known awards like the MTV Movie Awards, Teen Choice Awards, and even smaller awards like the Chicago Film Critics Association. After cleaning up the dataset to remove duplicate movies, the dataset includes a total of 266 unique movies. There were a total of 3996 awards handed out and 31.4% of the movies won at least one award.

Our group broke down the data and focused on three main data points to determine correlation with award wins or nominations accrued. We first focus on public reviews. Rotten Tomatoes, Metacritic, and IMDb all provide a public service where anybody can rate and review a movie and provide their viewer score. We look to see if the public's perception of a movie gives any indication on actual award wins or nominations accrued. We then focus on if a movie's length has any correlation with award wins. Is there a set time frame that wins the most awards? Do certain movie lengths basically mean there is no chance at winning awards? We then look at a movie's box office success. For movie studios, one of the driving factors in the decision to make a movie or not would be box office success. If a studio thinks a movie would not sell well, there is little chance of it being created. However, does big box office success turn into big award wins? We'll let individuals and studio heads debate what is more important between box office success and award success. We simply look to determine if there is any correlation between the two.

We started off our investigation by looking into how well different review sites (IMDb, Rotten Tomatoes and Metacritic) predict a film's success. A film was considered to have been deemed a "good" movie by these sites if they gave it a score of 70% or more. A bar graph was produced to compare how many total award nominations and wins corresponded to each sites' batch of "good" movies. Examining the resulting graph, Rotten Tomatoes and IMDb were the most generous with ratings above 70%, as their batches received more nominations than Metacritic's. This generosity did not reflect in being able to predict wins, however. IMDb, Metacritic and Rotten Tomatoes are not too far off from each other in regards to total wins. Metacritic's win to nomination ratio for reviews above 70% is much better than IMDb and Rotten Tomatoes. For this reason, one could argue that a high Metacritic rating is more likely to indicate an award-winning film.

By taking the data on runtime (in minutes) and placing them into different bins, we're able to categorize and display the information in a clear way. Breaking the bins into recognizable groups, we're able to quickly get across that most award wins happen in two categories. 40.2% of all total wins occur with movies that have a runtime between 90 – 119 minutes, and then an additional 38.4% of total wins occur for movies that have a runtime of 120 – 149 minutes. These two categories combine for 78.6% of all total wins. Perhaps more importantly it shows the difficulty in winning awards if you are below 90 minutes, or above 179. The percentage wins for movies with runtimes of 180 minutes or greater was negligible (1.2%). Movies that were below 90 minutes only accounted for 2.3% of wins. We visualized this data using a pie chart. To complement the pie chart, we also examined this data using boxplots for wins and nominations. These boxplots showed that movies with runtimes over 150 min and ranging 120 min to 150 min tend to do better than other movies as their median win amounts are above the others. Surprisingly, films with runtimes over 150 min tend to perform much better than all others, with this median being above the third quartile (Q3) for films with runtimes between 120 min - 150 min. This is true for both wins and nominations. Both forms of data analysis support the conclusion that films with runtimes below 90 minutes are the least successful and films with runtimes ranging from 120 min to 150 min are generally successful. Where they seem to disagree is the success of movie runtimes ranging from 90 min to 120 min and above 150 min. This may be due to the fact that the average runtime is about 116 min and there are many outliers for movies between 90 min to 120 min.

Similar to runtime (in minutes), we broke up Box Office into bins so that it was more clearly able to see trends between box office sales and wins. First we broke up the Box Office groupings by 50 million each to see if there were any noticeable trends in the data. When there didn't seem to be a correlation between awards and Box Office, we broke the bins further by 25 million to see if there were any correlations or trends. Unfortunately it did not look like there was any relationship between Awards and Box Office. This is not very surprising seeing that Awards are chosen by a group of movie critics and not based on ticket sales or movie popularity so it's unlikely that ticket sales would influence movie critics.

General perception and professional opinion often differ. While Rotten Tomatoes, Metacritic, and IMDb scores can often align with award success, it doesn't guarantee it. Reviewing the scatterplots, you can see a correlation between the higher the score, the more nominations and wins a movie is likely to receive, however there are a sizable amount of nominations and wins for movies that receive public scores of below 50%. When comparing the sites against each other though, Metacritic is the most likely site to predict success. There will always be outliers, but the slope for the Metacritic data is steepest, indicating the strongest correlation.

After reviewing the data on runtime, we can conclude that your best chances of winning awards is to have a movie with a runtime between 90 - 150 minutes, with the 120- 150 minute bucket being slightly more favorable. After determining the mean runtime length of 116 minutes, this isn't all that surprising. The most telling data from runtime is what to avoid. If a movie has a length under 90 minutes, or over 180 minutes, there's very little chance of winning an award.