Movie Data Analysis for Producers

Group 10

Group Members: Elliana Raines, Mukta Marathe, Evan Sharpe, Nicolas Peterson, Chancellor Harper

The goal of this project is to gain a real-world understanding of a data set, in our case, movie data. Our team at Arizona State University is going to take our selected data and gain an insightful understanding of it using in-class tools such as Excel and JMP to develop an in-depth analysis. The goal of this proposal is to create and present recommendations for movie producer companies based on the movie data set. Our team wants to address various questions regarding the data, including the following four important questions:

1. Which factors are the best predictors of a movie's success? Thereby also what are the priorities of movie producer companies with respect to factors it can choose in production. This includes but is not limited to the vision of the movie, quality of the story, budget and casting. These factors directly influence each other and the ability to make a great movie, so it's important to prioritize which ones deserve the most attention while making a movie.
2. What is the extent of external reception versus internal production on the movies' success? For example, what is the weight of critics reviews versus selected actors on a movie's success? How much do people value or like certain actors in the movie? Will that influence them to come to the theater more? The extent of the external reception versus internal production impacts are both very critical and equally important when it comes to the success of a film. Furthermore, managing the internal production such as picking reliable actors or maintaining an encouraging workspace will directly correlate to the external reception you receive.
3. What would we tell our marketing team given the task of advertising a certain movie, how would we do it, and what factors would be most important to touch on/what matters to audiences the most?
4. What are the most common or most popular plots and genres for movies?

Our data set contains information about various movies. For each movie, there is a director name, names of two actors, gross total made from the movie, a genre and key plot phrasing, link to an IMDB page as well as a score, and some other relevant data.

In order to answer our questions regarding the largest contributing factors and marketing, we will analyze the financial aspects as well as IMDB ratings to better understand what appeals to customers. For the question about popular genres and topics, we will analyze gross income as well as actors in the movie and IMDB pages. For internal versus external impacts, we will analyze and understand the IMDB pages and scores for what critics have commented on, as well as the number of movies Facebook likes to understand what customers and viewers think. Overall, the movie budget and financial data, as well as IMDB scores and pages seem to be very crucial aspects that we'll use to better analyze the data and answer our questions.

In order to answer our questions, we are going to organize data and graph it, then apply a correlation coefficient and correlation rating scale to understand what factors correlate with viewing rates, marketing, and movie popularity. For instance, we will graph movie IMDB scores versus Facebook likes to understand the difference or impact of critic ratings and audience ratings. Additionally, we will graph the amount spent on the money with the ratings and compare that to the amount spent and gross amount made from the movie, to understand what audiences enjoy. To further shed light on audience trends, we will also graph which actors and directors created the movie to IMDB and Facebook likes. This will help us understand which factors are

the most important for marketing techniques as well. To understand what the most common plots or genres are, we can graph the genre and plot description to amount of likes, gross income, and IMDB rankings. Overall, comparing correlation coefficients will help us understand which factors are most influential.

When addressing who our data benefits, we would be discussing shareholders and stakeholders of the specified movie producer company. Moreover in addition to our shareholders of the company, our stakeholders to this analysis would be the staff on the company's team, movie theater companies, actors and actresses, and ultimately the consumers. By being able to expertly leverage our data to provide a clear analysis to create a logistically sound movie, the producer company stands to make substantial revenue from the proceeds. Therefore besides the obvious monetary gain those hired by the company would receive, the people within the company would benefit with regard to an increase in reputational benefits. This could look like nominations at award ceremonies for outstanding performances both on and off screen following the execution of our suggestions.

After analyzing studies others have done regarding movie data for producer recommendations, we found that having an in-depth understanding of past analytics creates an environment possible for the success of a movie. This information would be most useful to movie marketers as they realize which factors in a movie are best to highlight.

Here, as an example of what our analysis could look like, we have included some preliminary analysis of movie Facebook likes versus IMDB score. As you can see, interestingly enough, the lower IMDB scores have a higher amount of Facebook likes. This shows a discrepancy between what critics and audiences think.