

Predicting Disney's Rotten Tomatoes

Caitlin Alano

Abstract

Disney films have been very popular for decades, with people of all generations embracing *nostalgia* in their own favorites. Every film has received a Rotten Tomatoes critic rating. With this as our response variable, we sought to determine the significance of movie length and budget as predictors. Our study utilized a dataset from Kaggle called “Disney Movies and Films,” which contains a list of Disney films, with each film as one observation, and their respective Rotten Tomatoes critic ratings, as well as other film-related statistics. Through a test for significant overall regression, we found that the model containing movie length and budget was not statistically significant in predicting Rotten Tomatoes critic rating for Disney movies, with both predictors exhibiting collinearity with one another. However, budget appeared to be a lot more significant than movie length in predicting critic rating. Knowing that budget does not significantly explain Rotten Tomatoes critic rating, this could encourage small-budget film producers to make more films without the fear of being overshadowed by big-budget films. Since movie length is also not a significant predictor of Rotten Tomatoes critic rating, this allows for the integration of unconventionally long genres like indie films. Thus, we see an abundance of creative expression that maximizes versatility in the ever-evolving film industry.

Introduction

Films are a major part of the entertainment industry, especially with the growing use of technology globally. Through the use of linear regression models, Fan et al. (2014) found that explanatory variables collectively proved to be the best predictors for box office revenue. Huang et al. (2015) revealed small budget movie strategies, which big budget movies lack, that make them equally worthy of investment as big budget films, including minimizing sunk costs and relying on post-release advertising with the help of positive critic reviews. Disney films in particular have been around since the late 1930s, yet they still soar in popularity, attracting viewers of all ages. How are these films perceived in the eyes of a Rotten Tomatoes critic? What factors can we further investigate in determining their effect on critic ratings?

This study will investigate how significant movie length and budget are in the prediction of a Disney film's Rotten Tomatoes critic rating. Disney films have been very popular for decades, and every film has received a Rotten Tomatoes critic rating. How effective are certain variables in predicting this rating? We will mainly be using Rotten Tomatoes critic rating as a response variable to analyze the contribution of two predictors, movie length and budget, in predicting this response.

With the popularity of the motion picture industry combined with the world's eternal love for Disney movies, our study is one that is timeless and relatable for people of all ages. Those involved in Disney film production would particularly be interested in our study, as finding the relationship between movie length and budget in predicting critic rating could reveal a perfect formula for getting a high Rotten Tomatoes score. From past studies mentioned earlier, we can conclude that through finding this formula, we can determine overall movie success in terms of box office revenue and demand. In the past, studies have included critic rating as a predictor

variable for box office revenue. However, no one has researched the factors that affect critic rating as a response variable. Thus, our study seeks to investigate two such factors and determine their significance to contribute to existing research in the motion picture industry.

Literature Review

Fan et al. (2014)

In the research study conducted by Fan et al. (2014), they investigated the use of FRAME, a nonlinear regression approach, to predict box office revenue patterns based on factors including VSMs, or virtual stock markets, and consumers' word of mouth, which can be deduced from the shapes of the distributions through FRAME. Through the use of seven linear regression models, they utilized explanatory variables that are movie characteristics, including genre, sequel, production budget, MPAA rating, runtime, studios, and all of them collectively in the last model. With FRAME, they utilized explanatory variables including consumers' word of mouth and VSMs. For both approaches the response variable was the cumulative box office revenues since the release date of the films. They found that the explanatory variables collectively proved to be the best predictors for box office revenue when looking at linear regression models. VSMs proved to be an effective predictor of success in the motion picture industry but required better applications (Fan et al., 2014).

Huang et al. (2015)

Huang et al. (2015) discussed the paradox that exists in the motion picture industry, where only a few movies, typically big-budget films, turn out to be successful despite the large number of low budget movies costing under \$15 million to make that are still being produced in Hollywood. By analyzing business models that reflect profit sources of both low-budget and big-budget films, this research investigated which effective strategies small movies utilize in contrast to big-budget

productions that make them worth investing in, with various factors that contribute to revenue as the predictor variables and ROI, or return on investment, as the response variable, which will be discussed in detail in terms of significance in the conclusions (Huang et al., 2015). After analyzing the significance of certain factors in both cases, this research revealed that unlike previous sources' claims, low-budget movies are worth investing in when considering other strategies that big-budget films typically would not utilize. According to their research, small films' profitability proved to be worthy of investment when analyzing the response variable of ROI, or return on investment, whose measures of central tendency were slightly higher than those of big-budget productions. Through the use of hypothesis testing techniques, they revealed that strategies such as minimizing sunk costs, reliance on post-release advertising with the help of positive critic reviews, and revenue from home video rather than box office contribute enough to small-movie profits that would make them worthy of investment (Huang et al., 2015).

Yahav (2016)

In 2016, Yahav sought to contribute research in the area of movie revenue prediction that rises above the focus of “predicting a total revenue of a movie or its opening weekend” after citing certain films like *The Incredible Hulk*, which “grossed \$55.4 [million] in the USA and in Canada during its release weekend but dropped 60% in the following weekend” (p. 409). With the goal of being able to predict the dynamics in box-office movie demand, he investigated the importance of consumer behavior via the Movie-Similarity Network in contrast to what he calls “under-control predictors,” such as “genre, budget, and actors” (Yahav, 2016, p. 410). Thus, he utilized a variable selection model to avoid “the issue of overfitting that incurs in the case in which the number of variables is extremely high in comparison with the number of observations” (Yahav, 2016, p. 419). Some explanatory variables that were included through this technique

include budget and runtime. With the similarity network as the main predictor, they found that using the similarity network significantly increased the prediction accuracy of the model when used with under-control predictors, noting that it was more effective than the prediction power of under-control predictors alone (Yahav, 2016).

Comparisons

Fan et al. (2014), Huang et al. (2015), and Yahav (2016) both investigated the profitability of movies with a response variable that indicated this. Their response variables are different, with Huang et al. using ROI, Fan et al. using box office revenue, and Yahav measuring box office demand through four measures, including peak time, second peak, horizon, and peak gross. However, their models shared similar explanatory variables, including budget, genre, sequel, and critic reviews. What Yahav referred to as “under-control predictors” were used differently in each of the research articles (2016, p. 410). Fan et al. had combined all seven variables into one of the linear regression models. In Huang et al.’s research, they analyzed the effect of each of these predictor variables individually for both small-budget productions and big-budget productions. Yahav’s research was an extension of both previously mentioned studies, as he introduced the similarity network as a main predictor in addition to some of these predictors that his variable selection model selected .

Takeaways

Nevertheless, all three research articles sought to find a model that effectively measured the predicted success of films, helping investors and distributors in the motion picture industry. Fan et al. (2014) concluded that the linear regression model with all six predictor variables combined had the highest prediction power of box office revenue. However, Huang et al. (2015) mainly concluded that when all explanatory variables were taken into account separately, certain factors

favor small-budget and big-budget productions and provide an accurate snapshot of profitability. In contrast, Yahav (2016) emphasized the risk of overfitting if using all the predictor variables that Fan et al. (2014) had used in a single linear regression model. Thus, his variable selection model was effective in preventing this phenomenon, and when combined with the introduction of the similarity network, his model's prediction power proved to be much higher than models that simply used predictor variables like genre and budget (Yahav, 2016). Ultimately, Yahav concluded that his method that employed the use of the similarity network and a variable selection model in predicting box office demand dynamics would have an even higher prediction power if applied with virtual stock markets, as Fan et al. suggested, or critic reviews, which Huang et. al tested.

Methodology

The goal of this study is to determine factors that have an effect on the Rotten Tomatoes critic rating of Disney movies. We will be investigating how significant movie length and budget are in the prediction of a Disney film's Rotten Tomatoes critic rating. Our null hypothesis, H_0 , is that the model including movie length and budget is not significant in describing patient satisfaction levels. Our alternative hypothesis, H_A , is that the model containing movie length and budget as predictors is significant in describing patient satisfaction levels.

Participants

Each observation refers to a particular Disney film. We will only be analyzing Disney films in order to avoid conflict that arises when comparing films from lesser known studios. The data contains roughly 254 observations, which should be sufficient enough for performing multiple linear regression hypothesis testing.

The population of interest being studied is films from the Big Five, including Universal Pictures, Paramount Pictures, Warner Bros. Pictures, Columbia Pictures, and Walt Disney Pictures, which we will be utilizing in this study. Since small budget films and big budget films have different factors that contribute to their profitability according to Huang et al. (2015), we chose to focus on big-budget films from well-known production studios.

Data Collection

Through Kaggle, we obtained a dataset called “Disney Movies and Films,” which contains a list of Disney movies and their Rotten Tomatoes critic ratings, as well as other miscellaneous film-related information. The original poster of this dataset noted that the data was scraped from Wikipedia and cleaned afterwards, so it is reliable. We will be using all of the movies in this dataset, as they are all Disney films that we seek to analyze.

In order to clean the data, we needed to take into account that some observations lacked data that pertained to the predictor variables columns, movie length and budget, and the response variable column, Rotten Tomatoes critic rating. For rows with missing values in one of these columns, we will simply eliminate them. Similarly, we will eliminate the columns that contain data characteristics other than our three variables.

Variables

Our response variable is Rotten Tomatoes rating, which is a quantitative variable on a scale from 0 to 1.00, which critics use to rate the Disney films. Movie length is our first predictor, which is a quantitative variable describing the length of each Disney film measured in hours, rounded to the nearest hundredths. Budget is our second predictor, which is a quantitative variable describing the amount that Disney invested in making each film which is measured in millions of US dollars.

Description of the Analysis

Using already existing columns in the Disney dataset for every respective variable, we will be assessing the effects of both movie length and budget on Rotten Tomatoes critic rating for each Disney movie. The response variable (Y) represents the Rotten Tomatoes Rating of the Disney movie. We have two predictors, including movie length and budget. We generated the following multiple linear regression model:

where the variables are defined as follows:

- X_1 is the length of the movie in hours (Movie Length)
- X_2 is the budget of the movie in millions, rounded to the nearest thousandth (Budget)

Plan to Analyze Data

Primarily, we will be utilizing data wrangling skills in R, specifically through the ‘dplyr’ package, which allows us to transform tabular data. These skills will be used to select the columns corresponding to each of our variables and eliminate observations with missing values. We will also utilize R to rename columns to fit our desired variable names and ensure that variable types match our need for analysis, while also adjusting units to better fit our study. Lastly, we will use R to create a “Decade” variable in order to analyze the general trend in Rotten Tomatoes critic rating for Disney movies over the years. Using Minitab, we will then obtain outputs, including summary statistics for the response variable, which is Rotten Tomatoes rating, and explanatory variables, which are movie length and budget.

Afterwards, we will run regression diagnostics by obtaining diagnostic plots. After setting up the appropriate regression, we will obtain and analyze a histogram and normal probability plot of residuals for normality, a plot of residuals versus fits to check for homoscedasticity, and lastly, a

plot of residuals versus order to assess independence of order. To answer our main research question, we will be performing a test for significant overall regression in order to analyze the regression coefficients and in testing for the significance of movie length and budget in predicting Rotten Tomatoes rating. Lastly, we will obtain a time series of Rotten Tomatoes critic rating by decade.

Results

The main problem of this study was to analyze two specific explanatory variables, movie length and budget, and analyze their effect on the response variable, which is Rotten Tomatoes critic rating for each Disney movie. In terms of purpose, the outcome of this study could potentially be beneficial for those who are involved in Disney film production, as determining the prediction power of movie length and budget with relation to Rotten Tomatoes critic rating could ultimately reveal the best formula for what critics would deem a successful movie. We used a dataset from Kaggle called “Disney Movies and Films,” which consisted of a list of Disney films and their Rotten Tomatoes critic ratings, as other film details, including our two predictors under analysis, which were movie length and budget. After removing observations with missing values in R, we were left with 254 observations.

Descriptive Statistics

Based on the Minitab outputs, we found that the response variable was severely left skewed by looking at the histogram of Rotten Tomatoes Rating, which we can refer back to in Figure 1 below.

Figure 1

Histogram of Rotten Tomatoes Rating of Disney Movies

Kurtosis also confirmed the distribution of the response to be severely left skewed with our Minitab output returning a value of -1.03048. Generally, Disney movies received a moderately high Rotten Tomatoes rating with a mean of over 50 percent (Mean: 0.62736, SD: 0.26754). Few observations received Rotten Tomatoes ratings on the lower end of the spectrum, with the first quartile value of 0.40000 revealing that only 25 percent of Disney films received a rating of less than 40 percent. Further breakdown of the Rotten Tomatoes critic Rating response variable is presented in Table 1 below.

Table 1

Mean, Standard Deviation, Q1, Measures of Normality for Rotten Tomatoes Rating

Variable	Mean	St. Dev.	Q1	Kurtosis
Rotten Tomatoes Rating	0.62609	0.26730	0.40000	-1.03156

Minitab revealed that our first predictor, movie length, had a slightly right skewed distribution in the histogram, which we can refer to below in Figure 2.

Figure 2

Histogram of First Predictor: Movie Length

A value of 0.731 for kurtosis also revealed that the distribution for movie length is moderately right skewed. We also see that a majority of movies run for between one and two hours, with a mean of approximately 1.654 hours. We also confirm this through inspecting the first and third quartiles, finding that 50% of data lies between 1.417 and 1.808 hours. Further breakdown of movie length can be referred to below in Table 2.

Table 2

Mean, Standard Deviation, Quartiles, Measures of Normality for Movie Length (Hours)

Variable	Mean	St. Dev.	Q1	Q3	Kurtosis
Movie Length	1.65369	0.31976	1.41667	1.80833	0.731

Our second predictor, budget, displayed a severely right skewed distribution. We can deduce this from the histogram, which we can refer to below in Figure 3.

Figure 3

Histogram of Second Predictor: Budget

A value of 1.719 for kurtosis also revealed the severely right skewed nature of the distribution of budgets for Disney films. On average, the 254 films in this sample, had a budget of about \$66 million. However, there seems to be a large amount of variation from the mean, with a standard deviation of approximately \$70 million. Further breakdown of summary statistics for budget can be referred to below in Table 3.

Table 3

Mean, Standard Deviation., Quartiles, Measures of Normality for Budget (Millions US Dollars)

Variable	Mean	St. Dev.	Q1	Q3	Kurtosis
Budget (millions)	66.375	72.923	12.000	110.000	1.719

With regards to critic rating over the decades, we found that the average Rotten Tomatoes critic rating was on quite a steep decline from a peak average of 0.781 in the 1940s to a low of 0.5325 in the 1990s. However, we then see a recovery to 0.709851 in the 2010s that is soon followed by a decrease to 0.61 in the current decade. Refer to Figure 4 below for further analysis.

Figure 4

Time Series Plot of Average Rotten Tomatoes Rating Versus Decade

Inferential Statistics

The regression model defined in the methodology was not statistically significant based on the test statistic ($F_{2,250,0.05} = 3.03$, $p = 0.2276$, $r^2 = 0.018$), indicating that the model including both movie length and budget was not significant in predicting Rotten Tomatoes critic rating. There seemed to be issues of collinearity between both predictor variables, with a correlation of 0.412. In order to further analyze each variable's effect, all coefficients in Table 4 below were observed.

Table 4

Coefficients and Statistical Tests for Constant and Predictor Variables

Variable	Coefficient	SE	T-Value	p
Constant	0.6581	0.09134	7.20	0.000
Budget	0.0004	0.00025	1.72	0.086
Movie Length	-0.0368	0.05768	-0.64	0.5236

After obtaining all diagnostic plots, some regression conditions were not satisfied. In analyzing the residuals versus fits scatter plot in Figure 5 below, homoscedasticity seems to be violated as typically residuals seem to be larger towards the middle of the graph and smaller at the ends.

Figure 5

Assessing Homoscedasticity: Scatterplot of Residuals Versus Fitted Value

Because the normal probability plot appears to have a curve, which can be seen below in Figure 6, normality also appears to be violated. Furthermore, the histogram of residuals of the response in Figure 7 is severely left-skewed, confirming that normality is violated.

Figure 6

Assessing Normality I: Normal Probability Plot of Residuals

Figure 7

Assessing Normality II: Histogram of Residuals

However, independence was satisfied, as after observing residuals versus order in Figure 8 below, the residual plot does not seem to follow any kind of pattern.

Figure 8

Assessing Independence: Residuals Versus Order Plot

Conclusion

Through a test for significant overall regression, we found that the model containing movie length and budget was not statistically significant in predicting Rotten Tomatoes critic

rating for Disney movies, with both predictors exhibiting collinearity with one another. However, budget appeared to be a lot more significant than movie length in predicting critic rating. A time series plot revealed the trend in critic rating for Disney films over the past eight decades.

Interpretation of Results

With regards to the collinearity between our two predictors, movie length and budget, this can likely be attributed to a higher cost of production for longer movies and thus a lower cost of production for shorter movies. This can possibly be due to higher pay for actors who have longer scripts and thus spend more time on set for longer movies. Similarly, actors who do not have to spend as much time on set for shorter movies are typically paid less for their time, resulting in a lower budget allocated for the production of the film.

With a p-value of 0.086 for budget's coefficient in the model in comparison to 0.5236 for movie length, budget appears to be more significant than movie length in predicting Rotten Tomatoes critic rating for Disney movies. This could be attributed to the fact that having more money allocated to the production of a film allows for more advanced effects and editing, including computer generated imagery, in contrast to small budget films.

With regards to the insignificance of budget as a predictor of critic rating, we can refer to the histogram of budget. Since the distribution of budget was severely right skewed, we see many of the films being less than \$100 million. This could explain as to why our model was not able to capture a wide enough range of values for budget to describe critic rating. Movie length was also deemed insignificant from our overall test for regression, even more so than budget, which we can attribute to the fact that critics are accustomed to watching movies of all lengths. Thus, it would not be fair for them to deduct percentage points solely based on how long a movie is.

With relation to the time series plot of decade and average Rotten Tomatoes critic rating, it was surprising to see the sharp decline in average Rotten Tomatoes critic rating from the 1940s to the 1990s.

Implications

Knowing that budget does not significantly explain Rotten Tomatoes critic rating, this could encourage small-budget film producers to make more films without the fear of being overshadowed by big-budget films. Additionally, since movie length is not a significant predictor of Rotten Tomatoes critic rating, this allows for the integration of different genres like indie films, which seem to be unconventionally longer than popular films from big-budget production companies. Thus, the main takeaway of this study for the motion-picture industry is the encouragement of creative expression that fosters exploration of a realm of films outside of the conventional.

Limitations

Certain factors would have helped our analysis if made available to us. Primarily, it would have been helpful to know the number of reviews that went into each average Rotten Tomatoes critic rating that was provided in the dataset, as a smaller number of reviews leads to greater variability and thus, an inaccurate representation of the big picture that could throw off our regression model. Additionally, our initial plan of action included a comparison between average ratings of Pixar and Disney movies to see which films critics preferred of the two categories. However, there was no existing dataset that allowed for us to perform this analysis.

Lastly, after obtaining all necessary outputs in Minitab and performing the analyses, some values in the dataset from Kaggle appeared to be inaccurately entered but were inherently difficult to find because of the large number of observations. For example, one movie had a

budget of \$15 listed instead of \$15 million, which could negatively affect our results if this same mistake was made a lot.

Suggestions for Future Research

The motion picture industry is extremely popular in current day, as technology is always improving and evolving. For future research purposes, it would be interesting to explore the relationship between not only Pixar and Disney movies' Rotten Tomatoes critic ratings, but also the general interest from viewers. Furthermore, an adjustment that we could make to the study include taking a better look at the cleaned data. By doing this we would be able to avoid the possibility of inaccurate observations that would throw off our regression equation and deem our test for significance inaccurate.

Closing Remarks

Our study utilized a dataset of Disney films to explore the relationship between two predictors, including movie length and budget, and the response variable, which was Rotten Tomatoes critic rating. Ultimately, we wanted to see whether both predictors were capable of predicting critic ratings for Disney films. In essence, the model containing movie length and budget as predictors was not significant in predicting the Rotten Tomatoes critic rating. This knowledge benefits the motion picture industry in terms of encouraging creativity. If movie length and budget are not deemed significant factors in explaining critic rating, this allows for smaller-budget and unconventional genre movies that exceed the average movie length to have a fair chance and ultimately leads to a more versatile and accepting industry.

References

- Fan, Y., Foutz, N., James, G., & Jank, W. (2014). Functional response additive model estimation with online virtual stock markets. *The Annals of Applied Statistics*, 8(4), 2435-2460. Retrieved February 18, 2021, from <http://www.jstor.org/stable/24522390>
- Huang, D., Markovitch, D., & Strijnev, A. (2015). Exploring the small movie profitability puzzle. *Marketing Letters*, 26(1), 43-55. Retrieved February 17, 2021, from <http://www.jstor.org/stable/24571098>
- Patel, Sameer. (2021). *Disney Movies and Films* [Data set]. Kaggle. https://kaggle.com/therealsampat/disney-movies-dataset?utm_medium=social&utm_campaign=kaggle-dataset-share&utm_source=twitter
- Yahav, I. (2016) Network analysis: Understanding consumers' choice in the film industry and predicting pre-released weekly box-office revenue. *Appl. Stochastic Models Bus. Ind.*, 32: 409– 422. doi: 10.1002/asmb.2156

