Does The Budget of a Movie Have an Impact on Income?

Fiorella Mendez & Luciana Mendoza, Fort Hays State University

Abstract

Numerous sources in the field have discussed the positive correlation between movie budgets and the box office income of movies. This research paper aims to examine further the effects budget has on the income success of movies. To achieve this objective, we have extracted data on the most popular movies globally from the International Movie Database (IMDb). The data utilizes information between the years 2019 to 2022. The dependent variable, INCOME, is controlling for various independent variables, which are BUDGET (in millions of dollars), RATING (on a scale of 1-10), and RUNTIME (in minutes). Furthermore, our categorical variables are MONTH and YEAR.

The study obtains the results using the Ordinary Least Squares Methodology (OLS.) We found a few potential challenges that violate the OLS assumptions. However, they are solved by transforming our model. With this, we can say that our model could have an unbiased, efficient, and consistent estimator.

The findings indicate a potential positive relationship between budget and income. Even when adding the independent variables, the relationship becomes stronger; suggesting that budget is not necessarily the most critical factor in predicting the revenue success of movies.

Introduction

According to the journal "The Impact of Movie Titles on Box Office Success" (Kim and Bae, 2019), movies that invest more money in their making generate higher revenues. However, other sources contradict this assumption, stating that a high production budget is unnecessary for a movie to succeed. We find it very interesting to study the correlation between these movie budgets and income because that would be a significant finding for the movie industry to help them define their efforts better. Therefore, the hypothesis we are testing is whether a positive relationship exists between budget and income by increasing the budget amount of movie production. Having income as our dependent variable and budget as our independent variable, we also control four explanatory variables: rating, runtime, release month, and release year.

Another essential factor of our research is measuring income performance before, during, and after-effects of the COVID-19 Pandemic. According to the Forbes article "The Impact COVID-19 Had On The Entertainment Industry In 2020," the movie industry suffered greatly. As strict quarantine laws were implemented, many movie theaters were forced to limit entrance or completely close. Leading many to turn to illegal online streaming platforms. Because of this, we predict that there should be some outstanding results during 2020 and 2021, which need to be controlled for (House, 2023).

Likewise, the release month of the movie could play an essential role in its profitability. It is expected that people enjoy movie marathons during the holiday season, summer breaks, winter breaks, etc.

Summary Statistics of Numerical Variables

Variable	n	Mean	S.D.	Min	Mdn	Max
INCOME (in millions of dollars)	205	218.6	392.0	0.0~	70.0	2797.5
BUDGET (in millions of dollars)	205	71.1	73.9	0.3	40.0	356.0
RATING (in scale of 1-10)	205	6.8	0.8	2.8	6.8	8.5
RUNTIME (in minutes)	205	122.5	21.6	84.0	120.0	209.0

Bar Chart – Income Per Movies

Income (in millions of dollars)
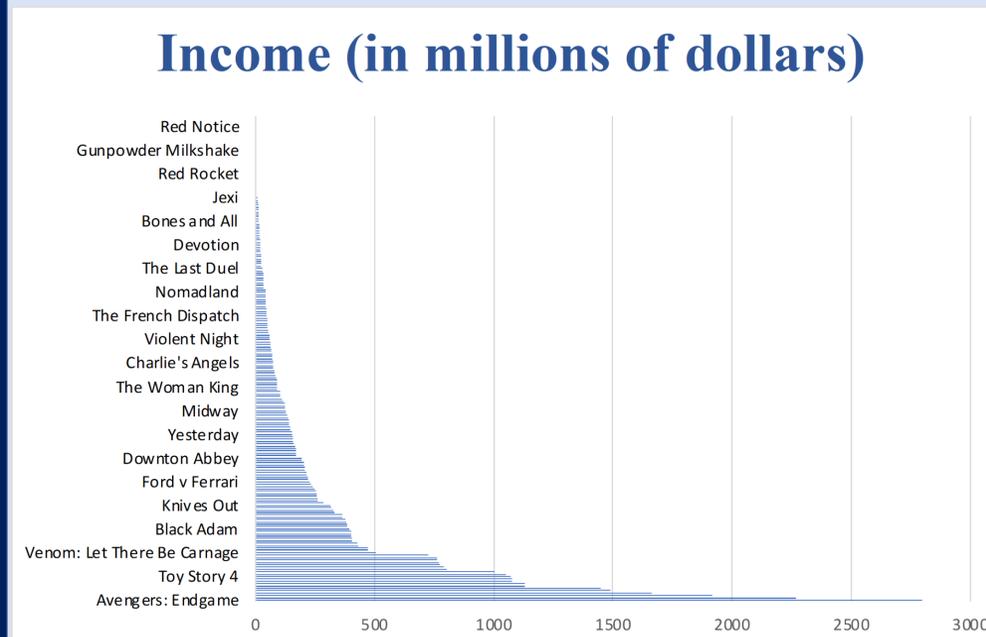
Methodology

This study's data was retrieved from the database platform "Kaggle," which was put together by George Scutelnicu. The author collected data from the IMDb for 2000 movies from 2003 to 2022, from which we decided to study only the data from 2019 to 2022. Following that parameter, we analyze 205 observations.

The mean income for the movies observed is $218,600,000. As we can observe, there is a considerable distance between the median and the mean. Some movies generate income around that median value, while others generate higher revenue.

Regarding the bar chart above, we can observe that the movie *Avengers: Endgame* generated the most income (bottom of the graph), while *Red Notice* generated the least (top of the graph.)

We test for multicollinearity, linearity, homoskedasticity, normality, and influential variables. Among these, the only assumption violated is homoskedasticity; therefore, we modify our model with a log-log transformation.

Selected Results

As we predicted at the beginning of our study, the regression results show a positive relationship between budget and income. A 1% increase in the variable budget increases the income by 0.797%. We can also observe that this effect increases as more controlled variables are added. The full model can be seen in Regression 5. A 1% increase in the budget raises income by 0.844%. Regression 6 excludes the top three influential variables, and the slope results appear even higher. All the relationships between budget and income are statistically significant under the 1% level. However, many results regarding RATING and RUNTIME are not statistically significant. Lastly, it is interesting to observe that among all the years included, 2020 has the slightest increase. In 2020, the income increase was 1.154% less compared to 2019. This shows a possible direct effect of COVID-19 on the movie industry's income.

Regression Outputs

VARIABLES	(1) lincome	(2) lincome	(3) lincome	(4) lincome	(5) lincome	(6) lincome
BUDGET	0.797***	0.801***	0.846***	0.834***	0.844***	0.922***
	(-0.101)	(-0.101)	(-0.114)	(-0.115)	(-0.114)	(-0.104)
RATING		1.156	1.538	1.235	1.117	0.556
		(-0.991)	(-1.086)	(-1.115)	(-1.099)	(-0.998)
RUNTIME			-0.829	-0.131	-0.213	-0.146
			(-0.96)	(-1.013)	(-1.026)	(-0.955)
MONTH	No	No	No	Yes	Yes	Yes
YEAR	No	No	No	No	Yes	Yes
Constant	1.047***	-1.172	1.913	0.0695	1.162	1.678
	(-0.389)	(-1.941)	(-4.068)	(-4.226)	(-4.34)	(-4.07)
Observations	205	205	205	205	205	202
R-squared	0.233	0.238	0.241	0.289	0.324	0.404
Adj R-squared	0.229	0.231	0.23	0.237	0.262	0.349

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Regression (6)-excludes observations: 27,75,316

Conclusion

Overall, the results are promising, and they match our initial prediction. Identifying a relationship between budget and income could be groundbreaking for the movie industry. In this study, we are not stating that budget is the most determinant factor in understanding income, nor are we claiming that this proves the existence of this relationship. However, we have found strong results pointing to this relationship being trustworthy. We also saw in our results that income went down in 2020, and then increased, but it has not reached the before-pandemic levels.

One of our shortcomings is that our data did not report income or budget for many movies. Having these details for all of them would give as more precise findings. Another thing is that not a high percentage of our data variation is explained by our model; We can know this by looking at our Adjusted R-squared 0.262. One of the things we can do to improve our model is to include more numerical variables, such as the number of A-list actors. It would also be worth analyzing the impact on income if the movie were directed by a famous producer.

Selected References

Bae, G., & Kim, H.-jin. (2019). The impact of movie titles on box office success. *Journal of Business Research*, *103*, 100–109. https://doi.org/10.1016/j.jbusres.2019.06.023

House, F. (2023, March 1). *Forbes EQ Brandvoice: The key to tackling our mental health crisis: Community-based care*. Forbes. Retrieved May 4, 2023, from https://www.forbes.com/sites/forbeseq/2023/02/27/the-key-to-tackling-our-mental-health-crisis-community-based-care/?sh=7bcf9ba67f1c

Kim, I. K. (2020). The impact of social distancing on box-office revenue: Evidence from the COVID-19 pandemic. *Quantitative Marketing and Economics*, *19*(1), 93–125. https://doi.org/10.1007/s11129-020-09230-x

Scutelnicu, G. (2023, March 15). *Top 100 popular movies from 2003 to 2022 (IMDb)*. Kaggle. Retrieved May 3, 2023, from https://www.kaggle.com/datasets/georgescutelnicu/top-100-popular-movies-from-2003-to-2022-imdb