# Are Action Movies more Successful for Disney?

## OBJECTIVE

The aim of our project was to determine if Action movies are more successful for Disney compared to Non-Action genres. For the purpose of our analysis, we use inflation adjusted gross as the metric for success - this is also our dependent variable. Our independent variables are genre, inflation adjusted budget, IMDB score, and star rating. We perform our analysis on a dataset of 350 rows.

## WHY A DATA DRIVEN APPROACH

Since the 1990's the silver screen has seen a rise in the popularity of Action movies. As technology evolved, the 2000's marked the start of many action packed superhero movies. Action movies grossed around $26 billion in the 90's in the 2000-2010 decade, this number climbed to $149 billion. It may follow that action movies generally gross higher than other genres, but in this project, we examine the data for the Walt Disney production house exclusively. In 2019, Disney posted an industry record of $13.2 billion at the global box office.The studio has also released five of the top ten highest-grossing films of all time worldwide, and the two highest-grossing film franchises of all time. However action is not Disney's flagship genre. Data driven approaches will shed light on whether Action grosses higher for Disney specifically.

## THE HYPOTHESIS

We test our null hypothesis which is that Action movies perform as well as the other genres - there is no difference between action and non action gross. The alternate hypothesis is that Action movies gross higher than other genres.

$H_0$ : Disney action movies perform as well as other genres

$H_a$ : Disney Action movies gross higher than other genres.

## THE IDEAL EXPERIMENT

The ideal experiment to test this is as follows. We consider two groups; Action movies and Non-Action movies. The other variables like the release date, budget, movie stars, IMDB score, movie runtime remain the same. This way, we can attribute the success of the film to its genre.

## DATA GENERATION PROCESS

To obtain data for the analysis, we merge three different datasets to get a list of Disney movies released from 1990-2016, that has data about each film's genre, stars, release date, budget, imdb score and gross.

## METHODOLOGY

To understand the data better, we initially performed exploratory data analysis.
The descriptive statistics tells us that Animation films gross the highest on average followed by Action and Adventure. We see that their average gross amounts are 100 million, 90 million and 52 million respectively.
To test the significance of these results and to see which other factors affect the gross, we perform regression analysis.

## REGRESSION ANALYSIS

We can see from our initial regression models that if the movie is of Action genre, then it earns 36 million higher than non Action movies. However, due to the low $R^2$ value of 0.2, it follows that genre alone is not an adequate predictor of inflation adjusted gross , and other variables seem to have an effect. We consider other factors such as Action (is the movie an action movie or not), star rating, budget and IMDB score. We interact with the variables star rating and action and also imdb score and action to determine whether the star rating or imdb scores have additional , non continuous effects on the movie's total gross. The regression analysis results are as follows -

1. We use the inflation adjusted gross as the dependent variable, and the independent variables are all the genres apart from action.
   The intercept value is the average that Action movies gross, and all the coefficient values indicate the difference between that genre and action. The values are negative (apart from animation), we can conclude that Action movies gross second the highest on average,

following Animation movies. The R^2 value of this model is low, we can see that genre alone is not an adequate predictor of inflation adjusted gross, and other variables may have an effect. We explore these as we go ahead.

2. We use the inflation adjusted gross as the dependent variable, and the independent variable is Action genre. If the movie is of Action genre, then it earns 36 million higher than non Action movies. This result is significant because of the low p-value (0.025). We see that the R^2 value of this model is low, we can see that Action genre alone is not an adequate predictor of inflation adjusted gross, and other variables have an effect. We explore these as we go ahead.
3. We use the inflation adjusted gross as the dependent variable, and the independent variables are Action genre, inflation adjusted budget, star rating, IMDB score.
   Controlling for inflation adjusted budget, star rating and IMDB score, an Action movie earns 23 million more than non-Action movies. However, the p value is high which makes this conclusion insignificant.
4. Here we use the inflation adjusted gross as the dependent variable, and the independent variables are Action genre, inflation adjusted budget, star rating, IMDB score, interacted variable (star rating and action), interacted term (IMDB score and action).
   The addition of interaction terms improves the significance of Action genre and increases the explainability of variation of inflation adjusted gross revenue.

CONCLUSION

Based on all the regressions that we ran we have the following conclusions:

1. If the genre of the movie is action and we don't take into account any other factor, the action genre highly impacts the inflation adjusted gross. An Action movie is expected to earn $36 million more than non Action movies.
2. If we take into account other factors that impact the gross earnings of a movie, the action genre alone is not that significant, but the IMDB rating and the budget of the movie become important factors.
3. Including all genres in the model we could see that when compared with action movies all other genres except animation garnered lower average gross income

LIMITATIONS

1) Animation is a broad genre and can also encompass other genres like comedy, action, etc.
2) Our dataset can be expanded to include movies after 2016 as well.
3) More factors, such as runtime and whether the movie is a sequel or a stand-alone could improve $R^2$ and make it more accurate.
4) We may better assess the significance level if we further categorized the data into animation-action .