

Disney's Gross Prediction before Covid

Thip Rattanaivilay

DSC680 - Fall 2021

<https://github.com/thiprattanaivilay/DSC680>

Which Domain?

The reason why I've selected The Disney Company for project 1, is because during covid streaming Disney+ was our go to entertainment during lock down and of course my kids love Disney. I wanted to use data science to predict and of course to see what Disney's gross profit was before covid-19 and compare regression. As we know The Walt Disney Company are one of the most influential corporations of the century. The company has been successful with business models for animated and live action films and television shows, immersive theme parks and cruise lines, and themed merchandise, while also successfully acquiring entertainment companies such as Marvel, Lucasfilm, ABC, ESPN, and soon, 21st Century Fox.

Here I have list 10 references that I will be gathering my information from to help with this project.

1. The Walt Disney Company. (n.d.). Retrieved from <https://www.thewaltdisneycompany.com/>
2. Forte, D. (2017, September 28). Walt Disney launches new online, store shopping experiences. Retrieved March 17, 2019, from <https://multichannelmerchant.com/ecommerce/walt-disneylaunches-new-online-store-shopping-experiences/>
3. Bonomolo, C. (2018, August 29). Disney's streaming service gets its name as new details emerge. Retrieved October 8, 2018, from <https://comicbook.com/movies/2018/08/26/disneystreaming-service-new-details-disney-play/>
4. Alexander, J. (2018, February 06). Disney's stand-alone streaming service won't compete with Netflix in scale. Retrieved October 8, 2018, from <https://www.polygon.com/2018/2/6/16981884/disneystreaming-netflix-marvel-star-wars-pixar-lucasfilm-svod>
5. Alvarez, E. (2018, April 12). ESPN offers a first look at Disney's big plans for streaming. Retrieved October 8, 2018, from <https://www.engadget.com/2018/04/12/espn-plus-disney-streamingservice-bamtech/>

6. Anderson, M. (2018, February 6). Will Disney's streaming service roar - or squeak? Retrieved October 8, 2018, from <https://www.usnews.com/news/business/articles/2018-02-06/will-disneysstreaming-service-roar-or-squeak>
7. Bartiromo, M. (2019, March 13). Walt Disney World increases ticket prices for most popular days. Retrieved March 14, 2019, from <https://www.foxnews.com/travel/walt-disney-world-increasesticket-prices-for-most-popular-days>
8. Thompson, D. (2018, May 17). Disneyflix is coming. And Netflix should be scared. Retrieved October 8, 2018, from <https://www.theatlantic.com/magazine/archive/2018/05/disneyflix-netflix/556895/>
9. Mitovich, M. W. (2018, February 09). Disney streaming service FAQ: Will Marvel heroes leave Netflix? Multiple Star Wars series? And more answers. Retrieved October 8, 2018, from <https://tvline.com/2018/02/08/star-wars-tv-series-disney-streaming-service-2019/>
10. Krämer, P. (2000). Entering the Magic Kingdom: The Walt Disney company, the Lion King and the limitations of criticism. *Film Studies*, (2), 44-50,2.

Which Data?

I will be selecting dataset from Kaggle which include what I am looking for from Disney's database. The data contains 579 Disney movies with six features: movie title, release date, genre, MPAA rating, total gross, and inflation-adjusted gross. While many of its films were big hits, some of them were not.

In this project:

* I will analyze data to see how Disney movies have changed in popularity since its first movie release

* Also, W will perform hypothesis testing to see what aspects of a movie contribute to its success.

The variables that I will be using for this project are:

- movie_title
- release_data
- genre
- mpaa_rating
- total_gross
- inflation_adjusted gross

Dataset link:

https://data.world/siyeh/walt-disney-movies/workspace/file?filename=disney_revenue_1991-2016.csv

<https://www.kaggle.com/rashikrahmanpritom/disney-movies-19372016-total-gross>

Research Questions? Benefits? Why analyzes these data?

I am proposing to analyze this Disney dataset, there will be 11 variables, one for each genre except the action genre which we will use as a baseline. For example, if a movie is an adventure movie, like The Jungle Book, the adventure variable will be 1 and other variables will be 0. Since the action genre is our baseline, if a movie is an action movie, such as The Marvel's Avengers, all those variables will be 0.

I assume my findings will benefit Disney greatly and with all the data gathered Disney will consider from all aspects of the company, but especially in relation to Disney+, is the increase in children's purchasing power following the rise in power and population dominance of Generation Z, the post-millennial age group. More children are being allowed to choose vacation spots, choose what toys they want to play with, and choose what they watch on television.

Research Question:

Do competitive platforms like Netflix, Amazon, or HBO be able to imitate Disney's existing library of content?

What is Disney internal weakness?

What is Disney external weakness?

How many genres are there and are they mixed?

Which production group made the most revenue for Disney?

What Method?

I am planning to use different approaches to predicting the gross profit for Disney. Boiling down this project, I see this being a classification problem in predicting genre among adventure and action movies. I am planning on using the following algorithms like random forests, linear regression, XGBoosting, and decision trees. Using different approaches will give an insightful image of which one works and the differences of the results. This will lead me to have a better understanding of the results and how to handle them.

Potential Issues?

One obvious issue that comes to mind is how to deal with wildly different results from using different approaches that might have many limitations.

I do fear that I won't complete the coding portion effectively with the given timeline. The data found on Kaggle are both numerous and high quality. There's no missing or fails data that needs to be removed, and some of the noise where a wait time will take longer and in the course of a couple minutes is smoothed by the prediction model.

Concluding Remarks

My goal was to provide general users with interesting, interactive visualizations that would allow them to explore and learn about different aspects of the Walt Disney Company and Disney studio. They attempted to predict gross revenue with the help of different sets of variables. For instance, budget, running time, star-power, MPAA ratings etc. are going to be used to predict the movie revenue from the standpoint of Movies. Whereas the fact that approximately 25% of the gross revenue gets accumulated in the first weekend of screening, I will have to take that in consideration to ensure my prediction is accurate.