西安交通大學

If the availability of the OTT platform affects the movie's achievement?

3121999272 Chanisa Sasivongpakdi

3121999301 Kim JinGu

Master Degree Program for International Students, School of Management,

Xi'an Jiao Tong University

28 Xianning W Rd, Jiaoda Commerce Block, Beilin, Xi'an, Shaanxi, China

November 2021

If the availability of OTT platforms affects the movie's achievement?

ABSTRACT

Nowadays, Over The Top (OTT) platform show rapid growth and strong escalation within these few years due to many factors and one of them is COVID-19. The pandemic has shut down and postpone the film production. Many consumers move from the traditional to OTT platform. This motivates us to examine whether the availability of OTT platforms affects the achievement of the movies which we define as movie's IMDb rating and popularity. Analyze by developing multiple regression model via R coding. The total of 108,492 movies data were harvested for secondary data resources which are IMDb, TMDB, and Kaggle

The result indicates that availability of OTT platforms and if it is the original series show significantly affect the result of the regression model. The R square values of each model range between 30-40%. The parameters are jointly statistically significant at significance level 0.05 and 0.01. Moreover, we found various significance p-value among each factor. Accordingly, this research aims to benefit in OTT and movie production industry and contribute in managerial level.

Keyword: OTT Platforms, Availability of OTT Platforms, Multiple Regression Model

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATION

	ABBREVIATION	Meaning
1	OTT	Over the top
2	MLR	Multiple linear regression

1. Introduction

1.1 Problem Background

OTT (over-the-top) platforms providing television and film content over the internet specifically suit the requirements of the individual consumer (*What Is OTT? | Telestream*, n.d.). Traditional broadcast, cable, and satellite pay-TV providers are no longer required (*What Is OTT(Over-the-Top)? Everything You Need to Know | Endavo*, n.d.). Netflix, Amazon Prime, and Disney Plus are just a few of the many OTT platforms available.

Nowadays OTT platforms show rapid growth within these few years. According to Statista, revenue in the OTT Video segment is estimated to reach US$178,012 million in 2021, and user penetration will be 40.9 percent in 2021, rising to 48.3 percent by 2025 (*OTT Video - Worldwide | Statista Market Forecast*, n.d.). These platforms provide a plethora of possibilities for viewers to enjoy a wide range of TV episodes and films. Audiences are not required to adhere to the television station's show time slot. Viewers can visit the site at any time and from any location. People are more likely to switch from traditional platforms to OTT platforms as a result of this (Hamm, n.d.).

One of the factors contributing to the rise in popularity of OTT platforms is the shortage of entertainment media options caused by the COVID-19 pandemic. All media production and film screenings have been suspended. According to an article published in the Journal of Content, Community, and Communication, at COVID-19, consumers in India were more accepting of OTT video streaming platforms, emphasizing the impact of OTT platforms on consumer behavior as the new normal (Malewar & Bajaj, 2020).

And how does this new normal behavior coexist with pre-existing behavior? As a consequence of the epidemic, many production companies had decided to postpone their film's launch date or switch to an OTT platform to reduce commercial losses. One example is the film "Cruella," which was scheduled to premiere in Thailand on June 24, 2021, but then was rescheduled to premiere on Disney+ Hotstar, Disney's OTT platform, on August 27, 2021, instead (ชูวาธิวัฒน์, 2021).

The content available on various OTT platforms varies. As researchers, this raises our awareness and highlights the topic of the impact of OTT platforms on the modern media industry. Many studies have focused on the revenue generated by OTT platforms, as well as the changes in customer behavior that have accompanied their expansion. However, in this study, we try to focus on how the film's availability on OTT platforms affects its success.

All of this has motivated us to investigate if the existence of an OTT platform has an impact on the success of films. After passing the data normality test, factors which are release year,

nation, language, genre, runtime, is adult, budget, director, authors, cast, and OTT availability will be examined using a multiple regression model. The popularity of the original series produced by the OTT platform and the other films will also be compared. In addition, because of the international COVID-19 pandemic, the OTT platform is growing.

This study's findings are intended to benefit the OTT and movie production industries by raising awareness of the collaboration between traditional and new normal behavior. Contribute at the managerial level to guiding on the effective of the OTT platform in film industry.

1.2 Research Framework

1.2.1 Research Purpose

Our research aims to declare the impact of the OTT platforms on movie achievement. Whether it should be considered one of the factors or not. As well as, how OTT production effect to overall film industry. By using movie's ratings and popularity as our dependent variables. Employing multiple regression model method on collected secondary data which will be explain more in Data collection and preprocessing section.

1.2.2 Research Objectives

1. Examine the popularity disparities between movies available on OTT platforms and not.
2. Examine the popularity disparities between original series (produced by platforms) and other films.
3. Examine the impact of variables, such as age, release year, place of origin, genre, availability of OTT platforms, and etc. on movie's ratings and popularity.
4. Examine the impact of variables over time.
5. Create models that can address all mentioned objectives

1.2.3 Research Question

If the availability of OTT platforms affects the achievement of the movies?

1.3 Research Significances

As the technology growth rapidly, this impacts every industry for instance in medical there is the Telemedicine, in publishing there is e-book, and film there is OTT platforms. These impact many companies to change their strategies and establish of many start-up companies. This research hope to contribute in OTT and movie production industry by help rising the awareness of the cooperate between the traditional and new normal behavior. As well as making decision in managerial level to guide how the industry post COVID-19 is going to be.

1.4 Research Methodology

1.4.1 Multiple Linear Regression (MLR)

Multiple linear regression or multiple regression is one of the statistical techniques that is a type of linear regression which used to study a relationship between the dependent variable and independent variables to predict the outcome in model from that indicate the linear relationship between dependent variable and independent variables which show in generalizing functional. This will benefit in investigating the cause of the variation in the dependent variable suitable for predicting the dependent variable and incorporate general functional form relationships. In essence, the multiple linear regression is the extension of ordinary least-square (OLS), which use only one explanatory variable, but the multiple linear regression involves in more than one independent variables. This method is benefits and wildly use in econometrics and financial inference. However, the multiple linear regression result is based on the previous information. It does not give the exact result of what will transpire in the future.

The differences between the multiple liner regression model and the simple linear regression model are the complexity and variety of variables, the multiple linear regression composed of more than one independent variable. The general form of the multiple linear regression is the following:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + U$$

Where:

Y = Dependent Variable

β_0 = Slope parameter or the intercept value of Y when all independent variables are 0

β_n = The intercept value of Y with respect to x_n

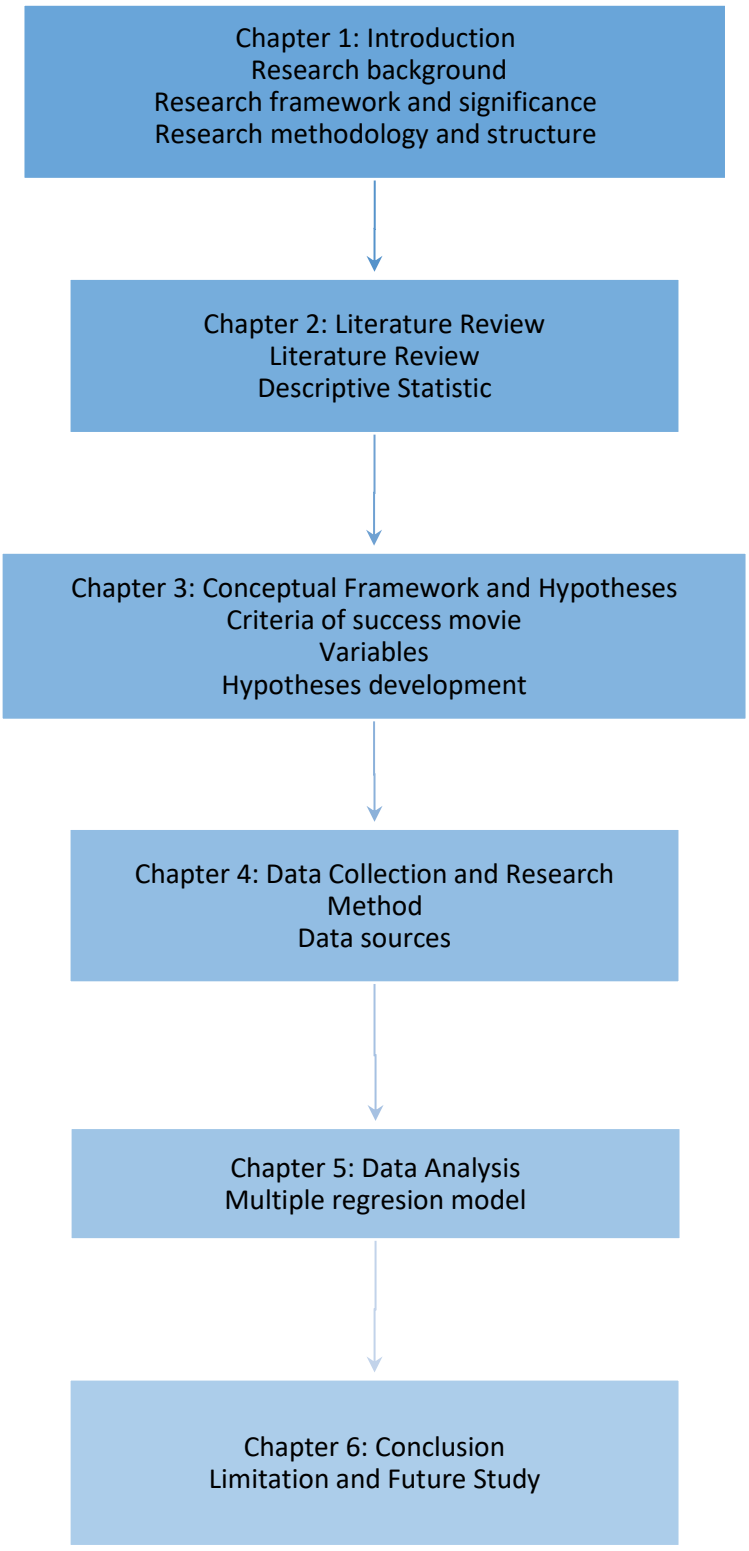
x_n = Independent Variable

U = Error term

1.4.2 RStudio

RStudio is a software provide an integrated development environment (IDE) for R coding in both statistical computing and graphics. It includes a console, syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging and workspace management. The software available in open source and commercial editions and runs on the desktop or in a browser connected to RStudio Server or RStudio Workbench (Debian/Ubuntu, Red Hat/CentOS, and SUSE Linux).

1.5 Research Structure



2. Literature Review

In the prior researches, most researchers defined the success of the movie with its financial success, especially earning from theater. In 1983 creative sphere, the scheduling and release pattern, and the marketing effort were use as three crucial decisions in determining the success of a theatrical motion picture (Litman, 1983). In post-studio era: 1960–1995, the star power was examined with box-office appeal and marking power as dependent variables to present the criteria of movie success (Albert, 1998). Later, during 1980-1990 the video rental industry was boomed and enjoyed its peak in the early 2000 (Contributor, 2021). Although the public can access almost all movies with video rental, in their own house, the financial result of the movies was still an important factor to measure the movies' achievement in the early of 2000's. The film sequential distribution channels were examined to study the differences between short-term box office, long-term box office, and video rental revenues. This indicated the differences impact of the success factor across the channel (Hennig-Thurau et al., 2006). All these years with many revolutions in film industry but the movie achievement still mostly rely on their revenue.

As the movie industry gets bigger and with the rapid globalization through many industries, the border of media is getting blurred. People can get various media contents through the internet. American media culture shows global influence effect the need for global box office receipts in other countries (Crane, 2014). More and more people can get access to the movies from all over the world, we can get some more results regarding the success of the movies. We can get the evaluation of the movies from some websites, such as IMDb, RottenTomatoes, and Metacritic, especially IMDb has more than 54 million users from all over the world including movie critics and normal viewers. These kinds of evaluations become important considerations for movie viewers whether they will see the movie. There may type of reviewers from novices to experts in terms of rating, readability and usefulness (Banerjee & Chua, 2018). Accordingly, we define movie achievement with IMDb rating and movie's popularity.

In addition to this, the characteristics of successful films have changed over the past forty years. The criteria attributes usually according to likelihood of consumer demand (Oat, 2013). With the world modernization and technology take part in every field. These given us an interest to focus in introducing the availability of the OTT as one of the attributes. The factors that empirically in the model are budget (Addis & Holbrook, 2018; Simonton, 2009), genre (Addis & Holbrook, 2018; Deuchert et al., 2005; Lash & Zhao, 2016; Simonton, 2009), runtime (Simonton, 2009; Zhang & Dai, 2021), language (Addis & Holbrook, 2018; Zhang & Dai, 2021), year of release (Addis & Holbrook, 2018; Lash & Zhao, 2016; Simonton, 2009), and power of director and star(Addis & Holbrook, 2018; Albert, 1998; Deuchert et al., 2005; Elberse, 2018; Hofmann et al., 2017; Inc, 2015; Lash & Zhao, 2016; Simonton, 2009). These factors have been discussed in many prior research the factors that we focus on to examining in this study are the availability of OTT platform and the is original which refer to the content the OTT platform had produced.

From (Simonton, 2009), First the budget factor will determine how much the entire production will cost, and it will place limits on all other factors. For example, in order to attract audience attention, more money will be spent on lavish special effects. The budget expansion is also due to the cast (Addis & Holbrook, 2018). However, not many films provide this information; instead, they mostly provide the gross box office. It is a rare and unclear piece of information. Second, the genre factor, there are numerous movie genres that can be considered as one unique value, but they are not mutually exclusive. For example, romantic comedy is a combination of comedy and romance. Different genres have varying degrees of effect on their audiences, which some prior research consider this as a control variable (Addis & Holbrook, 2018). Third runtime factor, it shows their affect when combining with other factors. The runtime positively correlated with critical evaluations, financial performance, and awards nominations. Fourth language factor, usually prior research focus on movie which in English (Lash & Zhao, 2016) but with the globalization the language is not the barrier anymore. There are more mobility and at the same time show characteristic through language (Zhang & Dai, 2021). Fifth year of release factor, also work as combination effect as well (Lash & Zhao, 2016). For example, different movie production technique throughout the time which usually use as a control variable to illustrate event in different years (Addis & Holbrook, 2018). Sixth the power of star and director factor, the director and star success, box office and awards winning, help increase the audience trust in quality of the film. Many prior research has examined this factor as a major influence that drive to movie success (Elberse, 2018). For example, consumer dress according to cast's outfits and appearance in their daily basis and festival like Halloween (Inc, 2015).

Accordingly, using the availability of OTT platform still relatively new and still room to be explore. This motivates us to conduct this research to develop the multiple linear regression model with R coding via RStudio software to answer the question whether the availability of OTT platforms affects the achievement of the movies. Following is Table 1 show descriptive statistics.

Table 1: Descriptive statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
release_year	25825	115.000	1906.000	2021.000	1999.676	20.863	435.265	-1.713	0.015	2.368	0.030
region_FR	25825	1.000	0.000	1.000	0.038	0.192	0.037	4.812	0.015	21.156	0.030
region_GB	25825	1.000	0.000	1.000	0.062	0.241	0.058	3.643	0.015	11.270	0.030
region_IN	25825	1.000	0.000	1.000	0.053	0.223	0.050	4.013	0.015	14.103	0.030
region_US	25825	1.000	0.000	1.000	0.431	0.495	0.245	0.278	0.015	-1.923	0.030
region_others	25825	1.000	0.000	1.000	0.369	0.483	0.233	0.541	0.015	-1.707	0.030
language_en	25825	1.000	0.000	1.000	0.697	0.459	0.211	-0.858	0.015	-1.264	0.030
language_es	25825	1.000	0.000	1.000	0.027	0.164	0.027	5.780	0.015	31.408	0.030
language_fr	25825	1.000	0.000	1.000	0.037	0.189	0.036	4.896	0.015	21.972	0.030
language_hi	25825	1.000	0.000	1.000	0.022	0.146	0.021	6.550	0.015	40.901	0.030
language_others	25825	1.000	0.000	1.000	0.200	0.400	0.160	1.497	0.015	0.241	0.030
genre_Comedy	25825	1.000	0.000	1.000	0.058	0.233	0.054	3.802	0.015	12.454	0.030
genre_Crime	25825	1.000	0.000	1.000	0.002	0.040	0.002	25.039	0.015	625.001	0.030
genre_Drama	25825	1.000	0.000	1.000	0.095	0.294	0.086	2.756	0.015	5.596	0.030
genre_Horror	25825	1.000	0.000	1.000	0.031	0.172	0.030	5.460	0.015	27.810	0.030
genre_other	25825	1.000	0.000	1.000	0.809	0.393	0.155	-1.572	0.015	0.471	0.030
budget_dol_log	25825	32.053	-4.605	27.448	14.685	2.527	6.386	-0.859	0.015	3.062	0.030
runtime_min_log	25825	2.371	3.714	6.084	4.613	0.194	0.038	0.607	0.015	2.036	0.030
is_adult	25825	1.000	0.000	1.000	0.001	0.025	0.001	40.140	0.015	1609.375	0.030
dircetor_total_num_movies	25825	79.000	0.000	79.000	2.038	6.001	36.007	3.779	0.015	17.116	0.030
director_num_academy_winner	25825	2.000	0.000	2.000	0.035	0.189	0.036	5.378	0.015	29.156	0.030
director_total_popluarity	25825	21.678	0.000	21.678	1.444	3.267	10.674	1.978	0.015	2.529	0.030
stars_num_academy_winner	25825	4.000	0.000	4.000	0.133	0.408	0.166	3.411	0.015	12.966	0.030
stars_popularity_point_log	25825	45.646	0.000	45.646	7.166	10.710	114.699	1.480	0.015	1.297	0.030
is_original_movie	25825	1.000	0.000	1.000	0.003	0.055	0.003	17.998	0.015	321.964	0.030
Avail_OTT	25825	1.000	0.000	1.000	0.118	0.322	0.104	2.373	0.015	3.629	0.030
movie_popularity_log	25825	9.271	-0.511	8.760	1.443	1.183	1.399	0.615	0.015	0.700	0.030
num_votes_2019_log	25825	14.639	0.000	14.639	7.425	2.751	7.566	-0.411	0.015	0.774	0.030
num_votes_2021_log	25825	10.116	4.605	14.721	7.880	2.200	4.841	0.561	0.015	-0.625	0.030
num_votes_2017_log	25825	9.458	0.000	9.458	2.245	2.245	5.040	0.799	0.015	-0.372	0.030
Valid N (listwise)	25825										

3. Hypothesis

3.1 Criteria for success of Movie (Independent Variables)

In this research, we define IMDb ratings and popularity of movies as the main achievement of movies. Since the distribution channel of the movies changed in the period of globalization, expanded from online world to offline world, financial success from box-office is not the only measurement for the achievement of the movies anymore.

The recent growth of OTT platform services, such as Netflix, Amazon Prime, is very remarkable, and the impact of this service is quite huge to the movie industry as well. OTT platforms help domestic movie production companies financially, and to enter overseas markets

(Shon et al., 2021). Since the earning model of OTT platforms is to make more subscribers, those companies keep trying to grasp the interests of customers with exclusive or famous contents (Oat, 2013), and it is becoming important for movie producers to meet the conditions provided by OTT platforms. And also, the earning sources, such as merchandise of movies, license of the movies, take a portion for the total profit for movie production companies.

However, getting information about a movie's profit is almost impossible these days. The amount of earnings from OTT platforms are not available following the contract regulation of production companies and service providers, so it is really hard to estimate the total financial success of the Movie. Moreover, with the situation of COVID-19 pandemic, the global box office came in at $2.2billion, down 72% compared to the box office of 2019, in 2020 (ZIPIN, 2021). Some of the big movies like “Cruella” are only available on the OTT platforms since movie theaters are closed, and this means it doesn't have box-office scores (ชูวาธิวัฒน์, 2021).

Therefore, when considering dramatic changes of the film industry, IMDb ratings, which infer if those movies are recognized by many movie viewers, and the popularity, which infer if the movies can attract interests of potential customers, seems to be the proper variables to measure the success of the Movies.

3.2 Variables affects to the success of the Movies (Dependent Variables)

We used totally 8 independent variables to build our regression model and most of the variables are used in prior research. For example, Litman (Litman, 1983) used release year, Power of directors and stars (Academy award information), and MPAA ratings, and Addis (Addis & Holbrook, 2018) used *MPAA ratings, release year, language, genres, runtime, production budget, opening screens* as variables in his papers. In addition to this, since we have different conditions for measuring the success of the movies and film industry, we added *Availability of OTT platform* as the independent variables of our regression model.

(a) Release year. Release year is the year when movies are released to the public. This variable was the variable from the early paper and still considered as one of the important variables.

(b) Production Region. This is the nationality of the movie production companies. Many prior researches concentrate on the language of the movies, but production companies also take a big part of the movie, such as the cultural reflection of the movies, even sometimes take a part on the distribution of the movies. We used 5 categories for these variables. They are the top 4 dominant regions, “France”, “Great Britain”, “India”, “United states”, and took “Other'' for the remaining countries.

(c) Language. This is the main language of the movies. Many researchers, such as Addis (18), used this variable if it is English or not as a categorical variable. With our dataset, we have a lot of primary languages. We used the 5 categories for this variable. It is the top 4

dominant languages, “English”, “Spanish”, “French”, “Hindi” and took “Other” for the remaining languages, for the categorical variable.

(d) Budget(log). Most of the movies don’t have the budget information, since it is kind of restricted information that is not open to the public. However, still some of the movies make public their budget information, and we change all the currency to dollars and take a natural log since it is seriously skewed among the data.

(e) Is adult. From the early research, Litman (Litman, 1983) started to use MPAA (Motion Picture Association of America) rating as the variable to predict the success of the movies. This is the ratings system of the movies which classify them as G (General Audiences), PG (Parental Guidance Suggested), PG-13 (Parents Strongly Cautioned), R (Restricted), NC-17 (No One 17 And Under Admitted). However, this system was built in 1945, and has quite complex rating categories, we simply use if the movies are for adults or not in our research.

(f) Genres. Genres from our metadata, which is from IMDb datasets, have 23 unique values. Also, each movie has more than one genre such as “Action, Romance”. Therefore, it gives us thousands of genre categories when we make it as dummy values. Therefore, we used only the primary genre and chose 4 dominant primary genres of the movies. The list of genres is 'Drama’, ‘Comedy', 'Crime', 'Horror' and defines ‘other’ for genres which are not dominant.

(g) Runtime(log). This is the runtime of movies. Since the length of the movies varies, and this variable has high skewness. For better results, we take a natural log with this variable.

(h) Power of directors and stars. There is a lot of prior research using this variable and also some of the researchers try to figure out the power of cast, stars and directors. Elberse A. (Elberse, 2018) tried to measure the power of stars with the average income of stars, prize history of them and the number of cast members to predict the box-office revenue. In our research, we used the popularity and academy prize history of stars, directors to measure the power of them.

(i) Is original. As the size of corporations gets bigger, OTT platforms like Netflix try to make their own contents by themselves. These original contents naturally exclusive for its OTT platforms and were not considered for the factor of research since most prior researches focus on the box-office revenue. However, it seems that the influence of these original contents getting bigger, and the quality of them are also getting greater,

(j) Availability of OTT platform. This is our main variable to evaluate our hypothesis if the availability of OTT platforms affects the success of movies. We get the content list of Amazon Prime, Disney Plus, Netflix to make this independent variable. This variable has 8,000 available contents out of a total 24,000 movies in our datasets.

3.3 Hypotheses Development and conceptual framework

The factors including release_year, region_FR, region_GB, region_IN, region_US, region_others, language_en, language_es, language_fr, language_hi, language_others, genre_Comedy, genre_Drame, genre_Horror, genre_other, budget_dol_log, runtime_min_log, is_adult, director_total_num_movies, director_num_academy_winner, director_total_popularity, stars_num_academy_winner, star_popularity_point_log, is_original_movie, and Avail_OTT are designed to be factors of the models that caused the different level of movie's achievement base on prior research as mentioned on literature review section.

$\boldsymbol{H_0}$: The availability of OTT platforms factor show insignificantly affects to the IMDb ratings of movie.

$\boldsymbol{H_1}$: The availability of OTT platforms factor show significantly affects to the IMDb ratings of movie.

$\boldsymbol{H_2}$: The impact of OTT platforms on IMDb ratings of movie has been growth over time.

4. Data collection and preprocessing

4.1 Data sources

In this research, there are 3 main data sources. First is the IMDb dataset. IMDb is an online database of information related to films, television programs, home videos, video games, and streaming content online. As of June 2021, IMDb has approximately 8 million titles and 10.4 million personalities in its database, as well as 83 million registered users. We used IMDb dataset to build a metadata of the movies. We got the basic information of movies from this resource, and it consists of*, title, genres, region of production company, main language, released year, runtimes, average rating, number of votes.* In order to get a more accurate result, we chose the data typed "Movie, tvMovie", genre without short film, runtime more than 40 minutes and more than 100 number of votes.

With the basic information of movies, we used additional data from IMDb to get the quantitative figures of power of director and stars as well. First is the popularity of directors and stars. We used the poll data of IMDb for choosing favorite directors and stars, and got the list of 677 directors and 1,479 stars that have over 500 points from IMDb's poll. Second data is the history of academy prizes, which is the most influential prize among movie prizes. Last is the number of movies taken by top 500 directors and average income of their movies.

Second resource is TMDB. TMDB is an online database of information about movies which has millions of users from all over the world. We can get additional data from TMDB which we could not find in IMDb, that are *movie budget, movie popularity, movie revenue*. Specifically for

the movie popularity, TMDB has its own algorithm for calculate this and it is made with Number of votes for the day, Number of views for the day, Number of users who marked it as a "favourite" for the day, Number of users who added it to their "watchlist" for the day, Release date, Number of total votes, Previous days score.

Final source of the data is Kaggle. Kaggle is one of the biggest platforms of data analysis competition, and also can find a lot of users oriented and companies' primary datasets from this website. Since we don't have the past data of IMDb ratings, we got the IMDb data scrapped on 2017 and 2019 from Kaggle, and the list of content available on OTT platform as well. Our datasets consist of 108,492 movies and all the data from IMDb, TMDB, Kaggle were scrapped in October 2021.

4.2 Data structure

Since we have some categorical data, skewed data we did the manipulation for better results. Table is the description of our data. Also, we posted all the code and process to make this on Github (16). Table 2 illustrate variable name, data type, their description and how we manipulate it for this research as following;

Table 2: Data description

Variable Name	Data Type	Description	Manipulation
imdb_rate_2021	numerical	IMDb ratings of 2021	
movie_popularity_log	numerical	populariry of movies	
imdb_rate_2019	numerical	IMDb ratings of 2019	
genre_	categorical	Genre of the movie	dummy, top 4 + 'other'
is_adult	categorical	if the movie is for adult	
region_	categorical	The region of production company	dummy, top 4 + 'other'
language_	categorical	The primary language of movies	dummy, top 4 + 'other'
budget_dol_log	numerical	budget of the movies which made public by production company	natural log
runtime_min_log	numerical	runtime of the movie	natural log
director_num_academy_winner	numerical	The number of academy awards for whole participated directors	
director_total_popluarity_log	numerical	The total popularity point for whole participated directors	natural log
stars_num_academy_winner	numerical	The number of academy awards for participated primary stars	
stars_popularity_point_log	numerical	The total popularity point for participated primary stars	natural log
is_original_movie	categorical	if the movie is produced by OTT platform itself	
Avail_OTT	categorical	If the movie is available on OTT platform	

5. Multiple Linear Regression model for hypothesis testing

5.1 Regression model for measuring the success of movies

We made a regression model to identify if the availability of OTT platforms significantly affects the success movies. Model 1 and 3 imply the regression models without the variables about the OTT platforms which is if it is the original series produced by OTT platforms and if it is

available on OTT platforms. In this model, we found that most of the variables from prior research significantly affect the IMDb ratings and popularity of Movies.

Surprisingly, regarding the correlation coefficient of the directors and stars power, it revealed that the number of movies that directors produced affected negatively to both IMDb ratings and movies' popularity (p_value < 0.01, estimate = -0.008, p_value < 0.01, estimate = -0.009). Furthermore, the history of getting academy prizes for directors and stars affect the popularity negatively (p_value < 0.01, estimate = -0.111, p_value < 0.01, estimate = -0.141). Finally, they do not even significantly affect the IMDb ratings.

Model 2 and 4 imply the regression models with the variables related to the OTT platforms. As we can see on table2, we found that the availability of OTT platforms significantly has an impact (p_value < 0.01, estimate = 0.25) on the IMDb ratings. Also, it has an impact (p_value < 0.01, estimate = 0.401) on the popularity of the movies. In addition to this, the the fact that the movies are the original series from OTT platform has significantly affects the IMDb ratings and the popularity of movies (p_value < 0.05, estimate = 0.311, p_value < 0.05, estimate = 0.205)

5.2 Comparison the models without variables about OTT platform and with OTT platform variables

Table 3 and 4: ANOVA for the two models with different dependent variables.

Model comparison for IMDb ratings

	Res.Df	RSS	Df	Sum of Sq	F	Pr(> F)
1	25,800	29,458.630				
2	25,798	29,284.570	2	174.063	76.670	0

Model comparison for popularity

	Res.Df	RSS	Df	Sum of Sq	F	Pr(> F)
1	25,800	18,352.900				
2	25,798	17,946.120	2	406.776	292.375	0

In order to get more evidence to prove that the variables about OTT platforms affect the success of movies, we conduct ANOVA for two regression models. The result of ANOVA, table4, shows the significantly low p-value for both independent variables. This means each of two regression models of two independent variables are significantly different. The only difference for two regression models was the variables about OTT platforms. Therefore, we can find that "Availability of OTT platforms" and "If it is the original series" significantly affect the result of the regression model.

5.3 Regression model to identify if the impact of OTT platform availability getting bigger

With the world-wide spread of the Covid-19 and lock down policy in many countries, it was clear that more people spent their time in their own house, and the number of Netflix users gets 16 million new subscribers in the early of 2020. Even some of the big movies were released

only on OTT platforms, because of the lock down, they can't be distributed through movie theaters.

Reflecting this trend on the film industry, we expected that the impact of the OTT platforms, which are the availability of the OTT platform and the fact that movies are produced by the OTT platforms in this research, is growth significantly for about 2 years. However, when we analyze the regression model, model one has IMDb ratings of 2021 and model two has IMDb ratings of 2019, we found that the impact of OTT platform didn't increase following 2 years. The estimate of 2 variables (Availl_OTT = 0.254, 0.251, is_original = 0.311, 0.299) don't show a big difference.

We thought that it is because the two years gap of the independent variables (2019 and 2021) is not enough to show the significant difference on IMDb ratings. Secondly, there are only 3,040 contents on OTT platform which is 12% of our whole dataset, and this portion can't change the whole trend of the regression model. Lastly, when some of the movies' license contracts expire, those movies are no longer available on OTT platforms. It means that the list of contents on OTT platforms would be changed every year. However, we used the list of 2021, so some of the movies may not be available in 2019, therefore, the availability of OTT platform may not actually have an impact on the IMDb rating of 2017 even though some movies seem to be available on OTT platform on our dataset. This is one of our limitations that we can't get the prior data by ourselves directly from the source of it.

Table 5: Regression model for the success of Movies

	Dependent variable:			
	y_2021		y_pop	
	(1)	(2)	(3)	(4)
release_year	-0.017***	-0.017***	0.009***	0.008***
	(0.0004)	(0.0004)	(0.0003)	(0.0003)
region_FR	0.026	0.030	-0.028	-0.021
	(0.049)	(0.049)	(0.039)	(0.038)
region_GB	0.112***	0.111***	0.118***	0.118***
	(0.041)	(0.041)	(0.033)	(0.032)
region_IN	0.227***	0.221***	-0.755***	-0.765***
	(0.053)	(0.053)	(0.042)	(0.041)
region_others	-0.043	-0.036	-0.474***	-0.463***
	(0.035)	(0.035)	(0.028)	(0.027)
region_US	-0.082**	-0.090***	0.208***	0.197***
	(0.033)	(0.033)	(0.026)	(0.026)
language_en	-0.478***	-0.497***	0.202***	0.173***
	(0.055)	(0.055)	(0.043)	(0.043)
language_es	0.144**	0.126*	0.468***	0.439***
	(0.066)	(0.066)	(0.052)	(0.052)
language_fr	0.103	0.107	0.054	0.060
	(0.068)	(0.068)	(0.054)	(0.053)
language_hi	-0.397***	-0.481***	0.312***	0.180***
	(0.077)	(0.077)	(0.061)	(0.061)
language_others	0.242***	0.235***	0.045	0.034
	(0.055)	(0.055)	(0.043)	(0.043)
genre_Comedy	0.745***	0.745***	0.208***	0.210***
	(0.090)	(0.090)	(0.071)	(0.070)
genre_Crime	0.437**	0.451**	0.036	0.059
	(0.188)	(0.187)	(0.148)	(0.146)
genre_Drama	1.457***	1.456***	0.019	0.018
	(0.088)	(0.088)	(0.070)	(0.069)
genre_Horror	-0.105	-0.102	0.432***	0.437***
	(0.094)	(0.094)	(0.074)	(0.073)
genre_other	0.952***	0.945***	0.462***	0.451***
	(0.086)	(0.086)	(0.068)	(0.067)
budget_dol_log	0.031***	0.028***	0.151***	0.145***
	(0.003)	(0.003)	(0.002)	(0.002)
runtime_min_log	1.257***	1.251***	0.336***	0.328***
	(0.043)	(0.042)	(0.034)	(0.033)
is_adult	-0.075	-0.058	-0.011	0.015
	(0.267)	(0.267)	(0.211)	(0.209)
dircetor_total_num_movies	-0.008***	-0.008***	-0.009***	-0.009***
	(0.002)	(0.002)	(0.002)	(0.002)
director_num_academy_winner	0.045	0.047	-0.115***	-0.111***
	(0.043)	(0.043)	(0.034)	(0.033)
director_total_popluarity	0.079***	0.078***	0.071***	0.068***
	(0.004)	(0.004)	(0.003)	(0.003)
stars_num_academy_winner	-0.013	-0.012	-0.143***	-0.141***
	(0.020)	(0.019)	(0.015)	(0.015)
stars_popularity_point_log	0.021***	0.020***	0.021***	0.020***
	(0.001)	(0.001)	(0.001)	(0.001)
is_original_movie		0.311**		0.205**
		(0.122)		(0.095)
Avail_OTT		0.254***		0.401***
		(0.022)		(0.017)
Constant	32.348***	33.538***	-20.367***	-18.591***
	(0.755)	(0.759)	(0.596)	(0.594)
Observations	25,825	25,825	25,825	25,825
R^2	0.330	0.334	0.492	0.503
Adjusted R^2	0.330	0.333	0.492	0.503
Residual Std. Error	1.069 (df = 25800)	1.065 (df = 25798)	0.843 (df = 25800)	0.834 (df = 25798)
F Statistic	529.786*** (df = 24; 25800)	497.799*** (df = 26; 25798)	1,041.713*** (df = 24; 25800)	1,005.791*** (df = 26; 25798)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 6: Regression model with IMDb ratings of 2021 and 2019

	Dependent variable:	
	y_2021	y_2019
	(1)	(2)
release_year	-0.017***	-0.018***
	(0.0004)	(0.0004)
region_FR	0.030	-0.001
	(0.049)	(0.050)
region_GB	0.111***	0.087**
	(0.041)	(0.043)
region_IN	0.221***	0.164***
	(0.053)	(0.054)
region_others	-0.036	-0.074**
	(0.035)	(0.036)
region_US	-0.090***	-0.102***
	(0.033)	(0.035)
language_en	-0.497***	-0.516***
	(0.055)	(0.057)
language_es	0.126*	0.130*
	(0.066)	(0.068)
language_fr	0.107	0.097
	(0.068)	(0.070)
language_hi	-0.481***	-0.469***
	(0.077)	(0.079)
language_others	0.235***	0.231***
	(0.055)	(0.057)
genre_Comedy	0.745***	0.676***
	(0.090)	(0.092)
genre_Crime	0.451**	0.352*
	(0.187)	(0.190)
genre_Drama	1.456***	1.402***
	(0.088)	(0.090)
genre_Horror	-0.102	-0.152
	(0.094)	(0.096)
genre_other	0.945***	0.874***
	(0.086)	(0.088)
budget_dol_log	0.028***	0.025***
	(0.003)	(0.003)
runtime_min_log	1.251***	1.357***
	(0.042)	(0.044)
is_adult	-0.058	1.201
	(0.267)	(1.064)
dircetor_total_num_movies	-0.008***	-0.008***
	(0.002)	(0.002)
director_num_academy_winner	0.047	0.036
	(0.043)	(0.043)
director_total_popluarity	0.078***	0.078***
	(0.004)	(0.004)
stars_num_academy_winner	-0.012	-0.025
	(0.019)	(0.020)
stars_popularity_point_log	0.020***	0.021***
	(0.001)	(0.001)
is_original_movie	0.311**	0.299**
	(0.122)	(0.135)
Avail_OTT	0.254***	0.261***
	(0.022)	(0.022)
Constant	33.538***	34.001***
	(0.759)	(0.776)
Observations	25,825	24,570
R^2	0.334	0.339
Adjusted R^2	0.333	0.338
Residual Std. Error	1.065 (df = 25798)	1.064 (df = 24543)
F Statistic	497.799*** (df = 26; 25798)	483.831*** (df = 26; 24543)
Note:		*p<0.1; **p<0.05; ***p<0.01

6. Conclusion

6.1 Conclusion

According to our research, factors which are release year, production region, language, budget, is adult, genres, runtime, power of directors and stars, is original, and availability of OTT platforms are potential factors that can influence the movie's achievement.

After developing Multiple regression model to study the impact of the OTT platforms availability on movie's achievement, considering the comparison between the models without variables about OTT platform and with OTT platform variables. We employ 2021 IMDb rating and movie popularity and our dependent variables. We found similarity value between R square and adjusted R square but R square display slightly more. The R square values of 2021 IMDb rating the is 33% (0.33) and 33.4% (0.334) respectively, first value is model without OTT platforms variables and the second value is the model with the OTT platforms variables. Movie popularity, R square show 49.2% (0.492) and 50.3% (0.503). Then, consider the impact of OTT platforms availability growth over time respective to COVID-19. The R square value of IMDb rating year 2021 and 2019 is 33.4% (0.334) and 33.9% (0.339) respectively. These indicate how much movie's achievement can be explained by the mentioned factors.

From ANOVA table it indicates that availability of OTT platforms and if it is the original series show significantly affect the result of the regression model. As for the OTT platform growth it shows insignificance relationship which may due to the fact that 2 years gap is not enough. We can conclude that the parameters are jointly statistically significant at significance level 0.05 and 0.01. Moreover, we found many significances relationship between each factor.

Therefore, these models may not show the high significance value but the models could be use as the guideline to make the model that can predict the movie's achievement as relate to the impact of the OTT platform. Then find all the causes and effect of the OTT platform and to improve in managerial level in movie industry.

6.2 Limitation and Future Study

This study has several implications. First, in this research the OTT platforms that we consider are Netflix, Amazon Prime, and Disney+ which may not generalize to all OTT platforms. Second, there are still some factors left that should take to consideration such as season of release, what are trended in that time, and etc. Third, we conduct this research by employ information internationally which may not able to represent all the areas and cultures. Fourth, as our data come from secondary sources. There is a lot of missing information which cause the accuracy of the data to be low but if the data were collect by ourselves the problem would be minimize. Moreover, the method to improve the model still left to be study furthermore.

References

Addis, M., & Holbrook, M. B. (2018). Is movie success a judgment device? When more is not better. *Psychology & Marketing*, *35*(12), 881–890. https://doi.org/10.1002/MAR.21141

Albert, S. (1998). Movie Stars and the Distribution of Financially Successful Films in the Motion Picture Industry. *Journal of Cultural Economics 1998 22:4*, *22*(4), 249–270. https://doi.org/10.1023/A:1007511817441

Banerjee, S., & Chua, A. Y. K. (2018). Tracing the growth of IMDb reviewers in terms of rating, readability and usefulness. *2018 4th International Conference on Information Management, ICIM 2018*, 57–61. https://doi.org/10.1109/INFOMAN.2018.8392809

Contributor, C. (2021). *Movie Rental Industry Life Cycles*. https://smallbusiness.chron.com/movie-rental-industry-life-cycles-63860.html

Crane, D. (2014). Cultural globalization and the dominance of the American film industry: cultural policies, national film industries, and transnational film. *Https://Doi.Org/10.1080/10286632.2013.832233*, *20*(4), 365–382. https://doi.org/10.1080/10286632.2013.832233

Deuchert, E., Adjamah, K., & Pauly, F. (2005). For Oscar Glory Or Oscar Money? *Journal of Cultural Economics 2005 29:3*, *29*(3), 159–176. https://doi.org/10.1007/S10824-005-3338-6

Elberse, A. (2018). The Power of Stars: Do Star Actors Drive the Success of Movies?: *Https://Doi.Org/10.1509/Jmkg.71.4.102*, *71*(4), 102–120. https://doi.org/10.1509/JMKG.71.4.102

Hamm, P. (n.d.). *Top 3 Reasons Why OTT Platforms are Better Than Cable | Endavo Media*. ENDAVO. Retrieved November 26, 2021, from https://www.endavomedia.com/netflix-versus-cable/

HAYES, A. (2021). *Multiple Linear Regression (MLR) Definition*. https://www.investopedia.com/terms/m/mlr.asp

Hennig-Thurau, T., Houston, M. B., & Walsh, G. (2006). The differing roles of success drivers across sequential channels: An application to the motion picture industry. *Journal of the Academy of Marketing Science 2006 34:4*, *34*(4), 559–575. https://doi.org/10.1177/0092070306286935

Hofmann, J., Clement, M., Völckner, F., & Hennig-Thurau, T. (2017). Empirical generalizations on the impact of stars on the economic success of movies. *International Journal of Research in Marketing*, *34*(2), 442–461. https://doi.org/10.1016/J.IJRESMAR.2016.08.006

Inc, W. B. E. (2015). *WarnerBros.com | Oscar-Winning Costumes | Articles*. Warnerbros. https://www.warnerbros.com/news/articles/2015/02/20/oscar-winning-costumes

Lash, M. T., & Zhao, K. (2016). Early Predictions of Movie Success: The Who, What, and When of Profitability. *Journal of Management Information Systems*, *33*(3), 874–903.

https://doi.org/10.1080/07421222.2016.1243969

Litman, B. R. (1983). Predicting Success of Theatrical Movies: An Empirical Study. *The Journal of Popular Culture*, *16*(4), 159–175. https://doi.org/10.1111/J.0022-3840.1983.1604_159.X

Malewar, D. S., & Bajaj, D. S. (2020). ACCEPTANCE OF OTT VIDEO STREAMING PLATFORMS IN INDIA DURING COVID-19: EXTENDING UTAUT2 WITH CONTENT AVAILABILITY. *Journal of Content, Community & Communication*, *12*. https://doi.org/10.31620/JCCC.12.20/09

Oat, E. (2013). Analysis of Netflix architecture and business model. *Aalto University T-110.5191 Seminar on Internetworking*.

OTT Video - Worldwide | Statista Market Forecast. (n.d.). Statista. Retrieved November 26, 2021, from https://www.statista.com/outlook/amo/media/tv-video/ott-video/worldwide

RStudio. (n.d.). *RStudio - RStudio*. Retrieved November 26, 2021, from https://www.rstudio.com/products/rstudio/

RStudio - Wikipedia. (2011). https://en.wikipedia.org/wiki/RStudio

Shon, M., Lee, D., & Kim, J. (2021). Are global over-the-top platforms the destroyers of ecosystems or the catalysts of innovation? *Telematics and Informatics*, *60*, 101581. https://doi.org/10.1016/J.TELE.2021.101581

Simonton, D. K. (2009). Cinematic success criteria and their predictors: The art and business of the film industry. *Psychology & Marketing*, *26*(5), 400–420. https://doi.org/10.1002/MAR.20280

What is OTT? | Telestream. (n.d.). Telestream. Retrieved November 26, 2021, from https://www.telestream.net/video/solutions/what-is-ott.htm

What is OTT(Over-the-Top)? Everything You Need to Know | Endavo. (n.d.). ENDAVO. Retrieved November 26, 2021, from https://www.endavomedia.com/what-is-ott/

Zhang, X., & Dai, J. (2021). Cultural and creative production in the era of globalization: Exploring the trans-border mobility of Chinese media and entertainment celebrities. *Geoforum*, *120*, 198–207. https://doi.org/10.1016/J.GEOFORUM.2021.02.001

ZIPIN, D. (2021). *How Exactly Do Movies Make Money?* Investopedia. https://www.investopedia.com/articles/investing/093015/how-exactly-do-movies-make-money.asp

ชูวาธิวัฒน์, ธ. (2021). *ปวดใจ เมื่อหนังฟอร์มใหญ่อาทิ Cruella และ Black Widow อาจไม่ได้ฉายโรง*. https://www.sanook.com/movie/113333/