|  |
| --- |
| Impact of Social Media on Movies |

|  |
| --- |
| Adarsh S  9-20-2017 |

Table of contents

[1. Introduction 2](#_Toc493714602)

[1.1 Source of Data 2](#_Toc493714603)

[1.2 Preparing the Data 2](#_Toc493714604)

[1.3 Agenda 2](#_Toc493714605)

[1.4 General observations 3](#_Toc493714606)

# Introduction

Impact of social media and other factors on movie’s performance at the box office and ratings.

## Source of Data

Public dataset available at Kaggle

<https://www.kaggle.com/deepmatrix/imdb-5000-movie-dataset/data/movie_metadata.csv>

## Preparing the Data

* Eliminating null values and zeros
* Removed irrelevant columns/data
* Transformed some data (content rating)
* Reduced the data to a representative set for analysis

## Agenda

Below are some of the queries that I will try to address as part of the forthcoming analysis

1. Do good IMDB ratings guarantee a movie's success at the box office?
2. Top 10 popular Lead Actor- how do their movies do?
3. Movies’ Facebook likes - how does it affect the movie's performance at the box office?
4. Any correlations observed from the data?

## 1.4 General observations

The below observations suggest that most of the movies have a rating of 5 or more and 8 or less.

The data set comprised of movie data, most of which were comedy, action or drama.

1. Do good IMDB ratings guarantee a movie's success at the box office?

From what was analysed and found in the data, the answer to the above question seems to be “Yes” but that comes with a certain set of conditions. From what we can observe below, it seems that a movie on average needs to be atleast rated “6” by IMDB for it to break even, and trend shows exponential growth with ratings going up from here onwards.

We did find outliers in both the extreme ends of this graph, where a movie with rating over 9 didn’t make profits and a movie with 1.5 rating making 50 million worth of profits. The former being “Shawshank Redemption” and the Latter being Justin Beiber’s “Never say never”.

There are various social reasons that contributed to these extreme situations that speak a different language than what the majority of data represents.

To summarise Mathematically, 91 % of the movies that have been rated over 5 have had positive profits which suggests that IMDB ratings and movie’s success (Net profit) are positively correlated.

1. Top 10 popular Lead actors- how do their movies do?

There is a positive correlation between the number of Facebook likes of a lead actor and the net profit their movie generates. This only makes sense as a lot of the fans go watch movies of their favorite actors which they tend to even like on Facebook to stay updated about them.

1. Movies’ Facebook likes - how does it affect the movie's performance at the box office?

From the below mentioned data there doesn’t seem to be much correlation between the number of likes a movie and the net profit it makes. As we can see, most of the data lies in the initial half of the X axis. We also see that the data on the Y axis(Profits) are distributed nearly equally on both positive and the negative side of the Y axis and this trend is uniform with growing number of Facebook likes too which only strengthens the argument we are presenting.

1. Any correlations observed from the data?

This data was observed to be positively correlated (coefficient of correlation was close to 1) which means the greater the number of critic reviews, the higher are the Facebook likes. This was a revelation of the data and there isn’t much that can be said about why this happens as a lot of the viewers that watch movies may or may not know or follow these critics so to be able to predict the root cause of this seems a little premature with the data at hand. To have such a definite relationship between these data points means that these data points needs further analysis on their own for us to drill down to the actual cause of this.