

Investigate the Secrets of Successful Movies

Data Selection

As a James Bond fan, I went to the cinema two weeks ago to see the film *No Time to Die*, which was brilliant and sold well, and it also had a 7.5/10 rating on IMDB. Looking at the success of the 007 series, I want to know what kind of movies are more successful. In other words, what kind of movie gets a higher IMDB score. So, for my project, I selected IMDB movies dataset with link [movie.csv](#).

Formulate a Question

Since my goal is to look for the secrets of successful movies, I decided to start my data exploration by looking at the top 10 movies with the highest IMDB rating.

Question 1: What are the top 10 movies with the highest IMDB rating? And how many IMDB users vote for each of these?

Question 2: Which movies have won both commercial success and critical acclaim?

Question 3: Which director's work gets the higher IMDB rating?

Question 4: What creative type and major genre of movies get a higher IMDB rating?

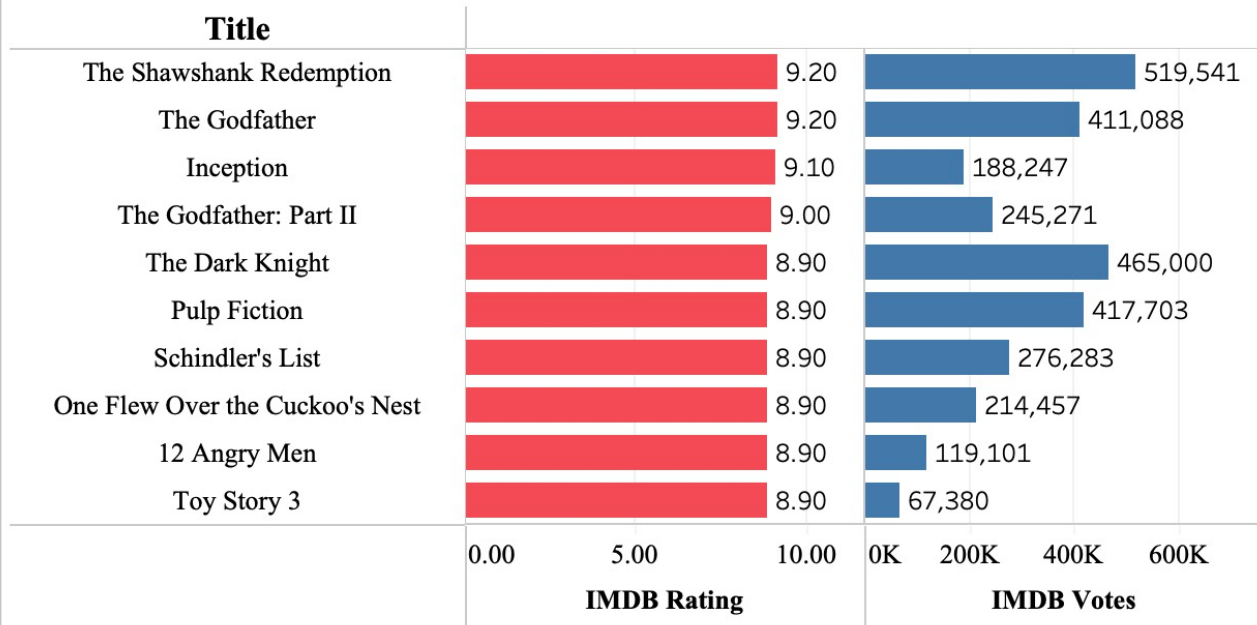
Inspect the Dataset

This dataset contains important statistics from movies, with 3,177 observations and 16 variables: Worldwide Gross, Production Budget, Major Genre, Creative Type, Director, IMDB Rating, etc. Most columns contain “null”, and if I eliminated all null-containing rows, there wouldn’t be many movies left to analyze. So, I decided to keep “null” but exclude them from the visualization.

Exploratory Visual Analysis

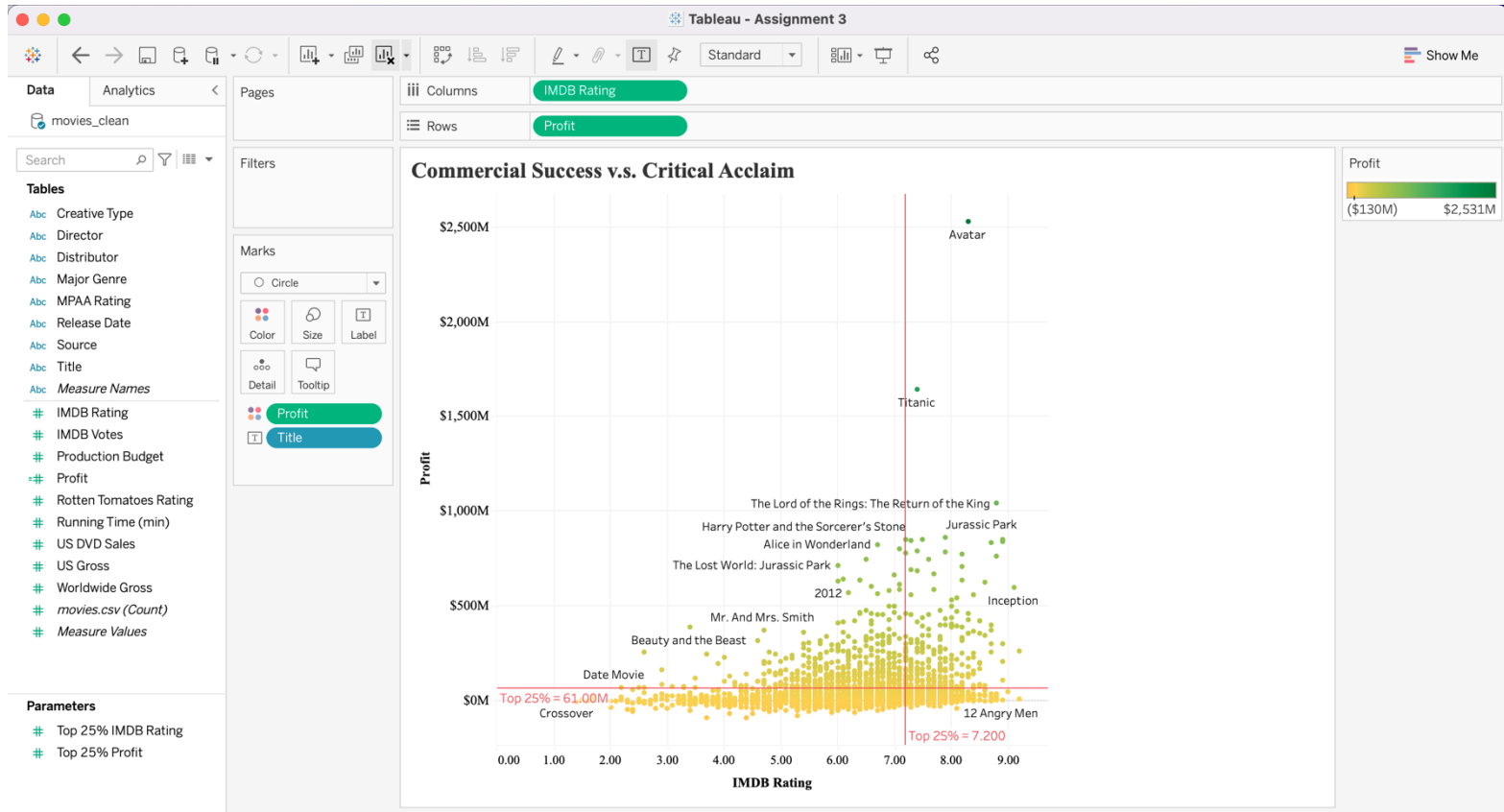
Question 1: What are the top 10 movies with the highest IMDB rating? And how many IMDB users vote for each of these?

Top 10 Movies with Highest IMDB Rating with the number of people vote



I decided to start my data exploratory by looking at the top 10 movies with the highest IMDB rating and voting number. For the same rating score, the more people vote, the better their reputation. I found that among the top 10 rated movies, there are some movies with high worldwide gross (*The Dark Knight*, *Toy Story 3*), some directed by famous directors (*Inception* and *The Dark Knight* by Christopher Nolan, *Pulp Fiction* by Quentin Tarantino), and some dramatic movies (*Schindler's List*, *12 Angry Men*). So, I will explore the secret of high IMDB ratings in terms of revenue, director, and genre of movies.

Question 2: Which movies have won both commercial success and critical acclaim?



How these new variable and parameters come from?

Profit = Worldwide Gross – Production Budget

Top 25% IMDB Rating = 7.2 (the value gets from the box plot of IMDB Rating)

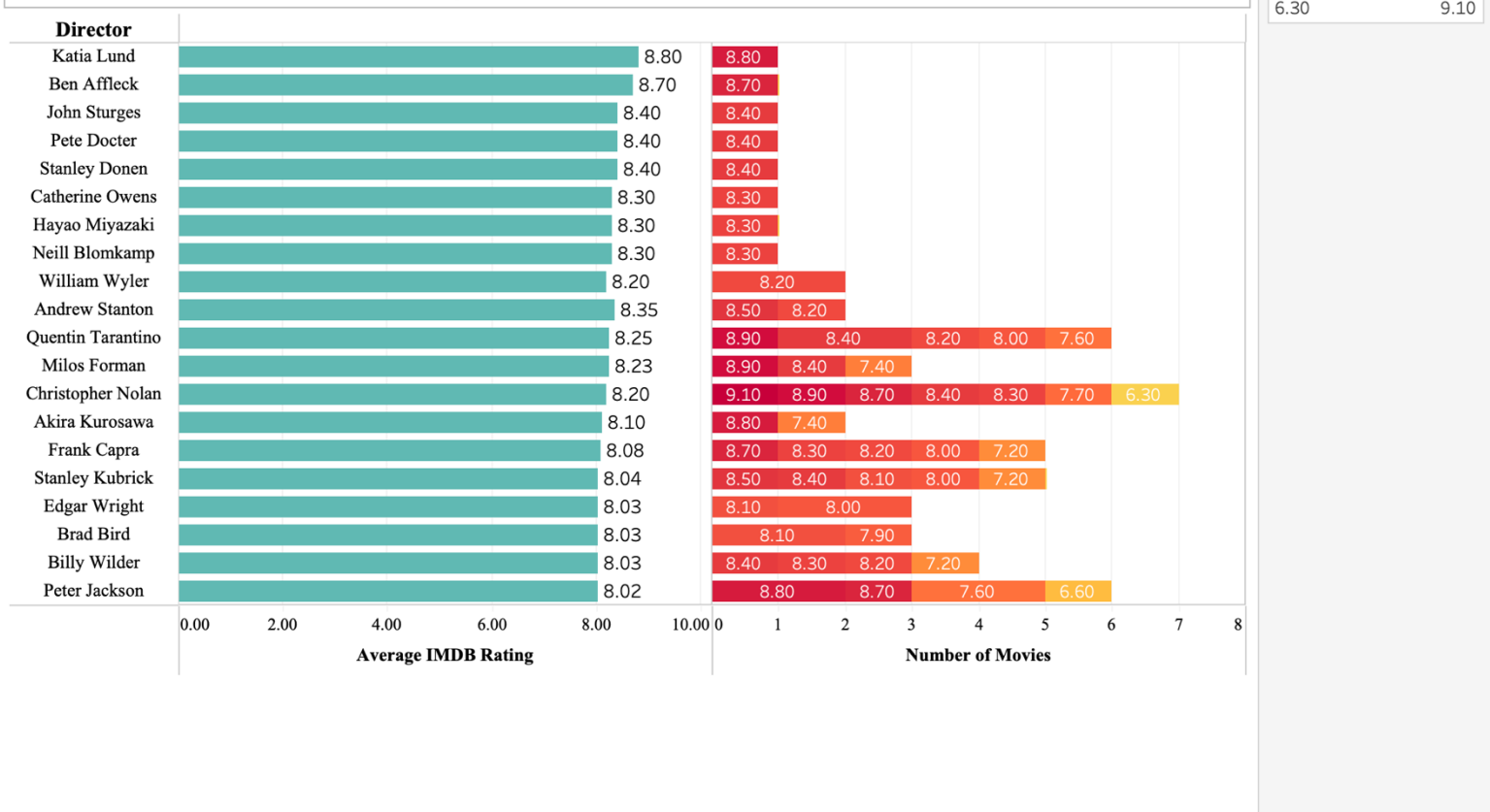
Top 25% Profit = 61M (the value gets from the box plot of Profit)

Let's take a look at some of the movies that have critical acclaim and commercial success at the same time. To highlight the success of these movies, I added two parameters: Top 25% IMDB Rating and Top 25% Profit. So, in the figure above, when a movie's IMDB rating exceeds 7.2, and its profit exceeds 61million, it exceeds 75% movies in terms of rating and profit. Such as *Avatar*, *Titanic*, *Jurassic Park*, etc. To my surprise, quite a few movies have been a huge commercial success, far higher than 75-percentile of profit.

Question 3: Which director's work get the higher IMDB rating?

Top 20 Directors with Highest Average IMDB Rating

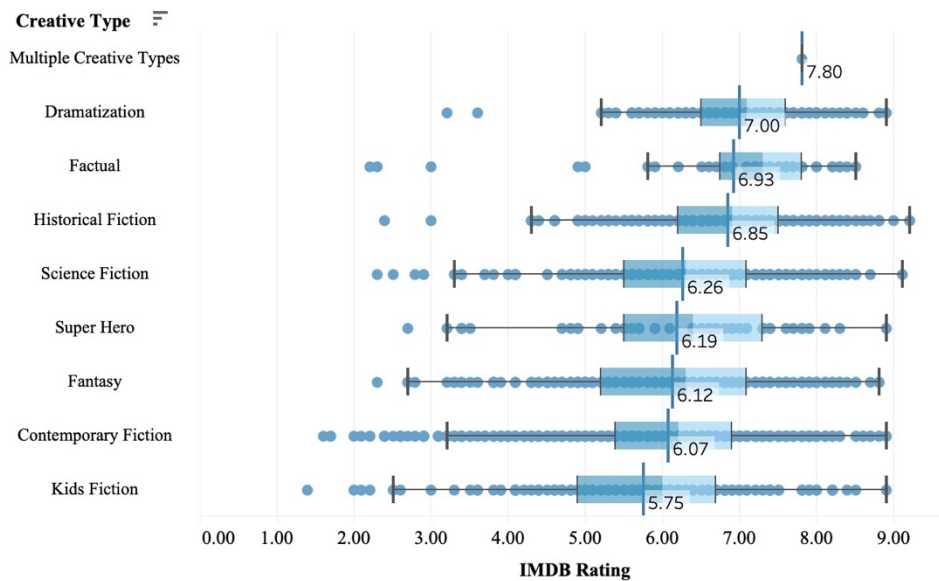
with the Number of Movies & the IMDB Rating of each Movie



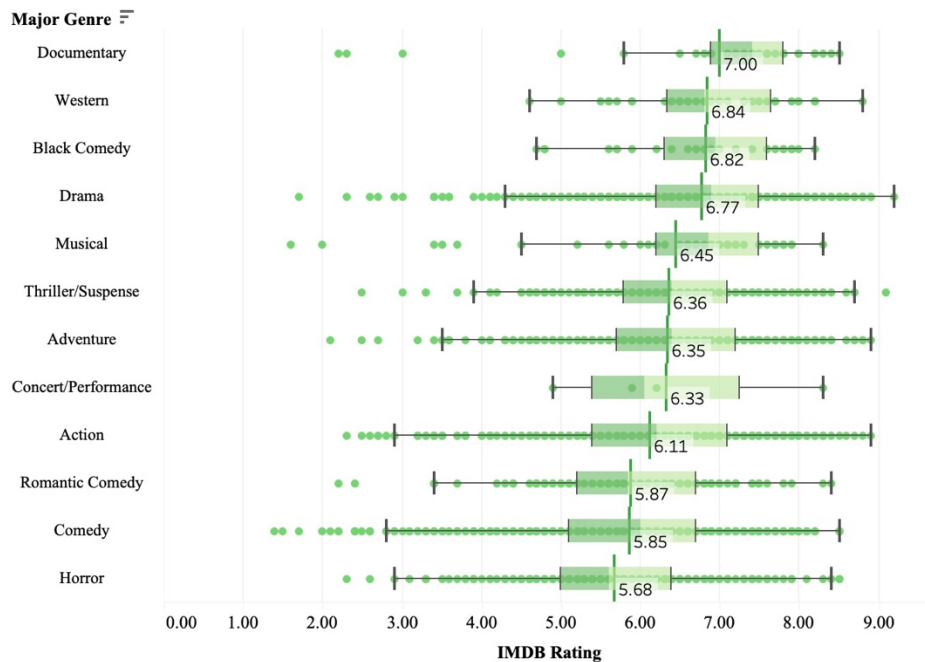
Next, I'm going to explore which director makes the highly rated movie. Because some directors are famous for one movie and some directors direct couples of the movie but not all of their works have an excellent rating, I create bar charts to present the top 20 directors with the highest average IMDB rating, their works, and the IMDB rating of each movie. It offers a more detailed understanding of the composition of the high reputation of directors. For example, we are familiar with the famous director Christopher Nolan, who directed *Inception*, *The Dark Knight*, *Memento*, and other high-scoring movies. Even though his work *Insomnia* does not have a good reputation as his other works, we cannot deny his commercial value.

Question 4: What creative type and major genre of movies get the higher IMDB rating?

The Average IMDB Rating of Creative Type



The Average IMDB Rating of Major Genre



Finally, I explored what kinds of creative types and genres high IMDB rating movies focus on and created box plots and labeled the average IMDB rating for each creative type and major genre. Multiple Creative Types have the highest average IMDB rating, but there is only one movie of this type, which is not sufficient to explain. So, the average IMDB rating of movies adapted from dramas is the highest. For major genre, documentaries have the highest average IMDB rating. On the other hand, the box plot can also vertically observe the distribution of

different creative types and genres of movies within each IMDB rating segment. For example, action movies don't have a high average IMDB rating, but the number of action movies with above 8 scores are greater than documentaries that have a rating above 8.

Conclusion

As discussed, commercially successful, directed by famous directors, and dramatic adaptations may be the secret to a movie getting a high IMDB rating.