Data Analysis of Movie Industry from 1980-2020

Prepared By :

GURUR AŞER

gururaser@hotmail.com

7th April 2023

Introduction:

This report presents a data analysis of movies from 1980-2020. This dataset contains various features of movies such as budget, scores, revenue, director, cast, genres, etc. This report provides insights into the budget, revenue, IMDb scores, genres, directors, cast, and production trends of movies on a yearly basis.

Methodology:

The dataset used in this report was obtained from data sources called Kagle. The data analysis tools used included Python programming language, Jupyter Notebook, and Python libraries such as pandas, NumPy, Seaborn and Matplotlib. Python code was written for cleaning, processing, visualizing, and analyzing the data.

Findings and Analysis:

1. The Most Successful Director-Star Pairs:

The most successful director-star pairs in the dataset are as follows (Based on IMDb ratings)

- Martin Scorsese - Robert De Niro (8.0)
- Martin Scorsese - Leonardo DiCaprio (8.0)
- Richard Linklater - Ethan Hawke (7.8)
- Tim Burton - Johnny Deep (7.2)
- Gore Verbinski - Johnny Deep (7.2)

The most successful director-star pairs in the dataset are as follows (Based on movie revenue)

- Gore Verbinski - Johnny Deep ($3.19 Billion)
- Ron Howard - Tom Hanks ($1.89 Billion)
- Dennis Dugan - Adam Sandler ($1.55 Billion)
- Martin Scorsese - Leonardo DiCaprio ($1.39 Billion)
- Tim Burton - Johnny Deep ($1.29 Billion)

2. The Most Popular Actors:

The most popular stars members in the dataset are as follows:

1. Nicolas Cage (43 movies)
2. Robert De Niro (41 movies)
3. Tom Hanks (41 movies)
4. Denzel Washington (37 movies)
5. Bruce Willis (34 movies)
6. Tom Cruise (33 movies)

7. Johnny Depp (33 movies)

8. Sylvester Stallone (32 movies)

9. John Travolta (30 movies)

10. Mel Gibson (29 movies)

3. Popular Movie Genres:

Looking at the data, we can say that Comedy, Drama, and Action genres are the most popular genres over the years. While the highest number of films in the Comedy genre was observed in 1994, it gradually began to decrease and fell to 2 in 2020. In 2020, the highest number of films was observed in the Comedy genre. Additionally, we can observe an increasing trend in Animation, Fantasy, and Sci-Fi genres in recent years. Even in 2019 and 2020, relatively high numbers of films were observed in these genres. On the other hand, Biography and History genres fall behind the other genres. Although there was an increase in the number of films in the Biography genre in the early 2000s, it has steadily decreased in subsequent years. The History genre, on the other hand, has had very low numbers since 2010.

4. Popular Trends by Year:

When examining the production trends of movies by year, it can be seen that there was a significant increase in movie production in the 1990s and 2000s. However, there was a significant drop in the number of films in 2020 compared to the previous year. The main reason for this significant drop is likely the COVID-19 pandemic. In 2020, most cinemas around the world were closed or had limited operating hours due to the pandemic. This led film producers and distributors to either postpone their film release dates or shift their focus to digital platforms. Additionally, the suspension or delay of film productions due to the pandemic may have resulted in fewer new film releases in 2020. Therefore, the observed decrease in the number of films during this period could be a consequence of the global situation.

5. Budget and Revenue of Movies:

When examining the budget and revenue data of movies, it can be seen that big-budget movies usually generate high revenue (Correlation is %71). For example, movies with a budget of over $100 million generally earn over $500 million in revenue. However, low-budget movies can also sometimes generate high revenue.

There are two movies that are examples of this situation:

- The Blair Witch Project (1999), which was made with a budget of $60,000 and earned approximately $250 million in revenue.

- Paranormal Activity (2009), which was made with a budget of $15,000 and earned approximately $195 million in revenue.

The interesting similarity between these two films is that both films are in the Horror category.

5.1. Budget and Revenue of Movies by Season:

When examining the dataset, we can see that the number of movies released by season varies. The highest number of movies were released in the fall season (1996 movies), while the lowest number of movies were released in the winter season (1434 movies). However, when looking at box office revenue by season, we can see that the summer season had the highest revenue ($181 billion), while the fall season had the lowest revenue ($128 billion), despite having the highest number of movie releases.

Conclusion:

This report presents a comprehensive data analysis of movies from 1980-2020. It provides insights into the budget, revenue, IMDb scores, genres, directors, cast, and production trends of movies on a yearly basis. Based on the analysis, it can be concluded that big-budget movies usually generate high revenue, and Comedy, Drama, and Action genres are the most popular genres over the years. Martin Scorsese-Robert De Niro and Gore Verbinski-Johnny Depp are the most successful director-star pairs based on IMDb ratings and movie revenue, respectively. Nicolas Cage, Robert De Niro, and Tom Hanks are the most popular cast members in the dataset.

The film industry has seen a significant increase in movie production in the 1990s and 2000s. However, the COVID-19 pandemic has led to a significant drop in the

number of films in 2020 compared to the previous year. This drop can be attributed to the closure of cinemas and the suspension or delay of film productions due to the pandemic.

Additionally, we observed that the number of movies released by season varies, with the highest number of movies released in the fall season. However, the summer season had the highest box office revenue, while the fall season had the lowest revenue, despite having the highest number of movie releases.

Overall, this report can be used as a basis for understanding industry trends in the movie industry and can assist decision-makers in the production of movies. It can be used to identify which genres, directors, and cast members are popular, and which ones are not, as well as to make informed decisions about budgeting, marketing, and distribution strategies.