

Project Investigate A TMDB Dataset

- **Introduction**

- **Summary**

Data Wrangling Process

I renamed the columns to make the data easier to investigate.

I deleted irrelevant columns such as “Tagline”, “Homepage”, “Budget Adj”, “Revenue Adj”, “ID”, “Release Date” and “IMDB ID”.

I also separated the date column to ‘year’, ‘month’, and ‘day’ to make it easier to to get.

Overview

I have decided to investigate the TMDB dataset, and I decided to find out what makes a movie earn more revenue, the following actions and questions were carried out and investigated, the following are the processes carried out and the questions asked.

1. Drop Irrelevant Columns
2. Separate data
3. What Are The Top 10 Most Popular Movies And Their Genres?
4. What Are The Top 10 Least Popular Movies ?
5. Who Are The Top 5 Directors With The Highest Revenue ?
6. Do movies with a longer runtime have more views ??
7. What Are The Movies With The Highest Budgets ?

Numpy, Pandas, Matplotlib and Plotly were used for this project.

1. What Were The Top Ten Most Popular Movies ?

The top 10 most popular movies were

1. Jurassic World
2. Mad Max: Fury Road
3. Interstellar
4. Guardians of the Galaxy
5. Insurgent
6. Captain America: The Winter Soldier
7. Star Wars
8. John Wick
9. Star Wars: The Force Awakens
10. The Hunger Games: Mockingjay - Part 1

2. What Were The Top Ten Least Popular Movies ?

The top 10 most popular movies were

1. North and South, Book I
2. The Hospital
3. G.B.F
4. Mon petit doigt m'a dit..
5. Khosla Ka Ghosla!
6. Born into Brothels
7. Soupçons
8. Fuera de carta
9. Slashers
10. The Invisible War

3. Do Movies With Certain Directors Have More Revenue ?

From my analysis I discovered that certain directors like James Cameron and J.J. Abrams have multiple movies among the top ten when considering revenue, however, the director of a movie has no serious impact on the view count as the top ten directors according to revenue are not the top 10 directors according to vote count.

4. Do movies with a longer runtime have more views ?

Movies with longer runtime such as “Pirates of the Caribbean: On Stranger Tides” and “The Warriors Way” tend to have more views than movies with a shorter runtime.

5. What Are The Movies With The Highest Budgets ?

1. The Warrior's Way
2. Pirates of the Caribbean: On Stranger Tides
3. Pirates of the Caribbean: At World's EndAvengers: Age of Ultron
4. Avengers: Age of Ultron
5. Superman Returns

Conclusion

The data used for this project is outdated and cannot be used to represent the current general population , due to the nature of the data, it cannot be used to make accurate predictive analysis.