

IMDB Movie Analysis Report

Prepared by
Abdelrahman Hussein Elshimy

Project Description

- This dataset contains information about different films, detailing aspects related to their production, cast, and performance. The data is organized into several columns, each of which provides specific information about the movies as name of director, number of reviews by critics, gross revenue generated by the movie, title of the movie, num of users who voted the movie, number of reviews by users, country where the movie was produced, budget allocated for the production of the movie, year the movie was released and rating of the movie.

Project Approach

- First we got an understanding of the given data, then we cleaned this data by removing null values, duplicated data, deleting unnecessary columns. After cleaning the data, we answered some questions by using some pivot tables and charts for explain more informations about this this data to reach the main goal of its analysis and then provide recommendations if it found.

Data Cleaning

- In this process we deleted plot-keywords column as unnecessary column because if the missing values are deleted, it may affect the rest of the columns inside the data and removed missing data in languages column.
- We removed duplicated data but in this data we found two movies has title (the host) ,So we changed title for each movie manually to be un duplicated (the host (USA), the host (S.KOR)).

Calculated columns

- Net profit = gross - budget.
- Profit percentage = $\frac{\text{Net profit}}{\text{budget}} * 100$
- Range of years [(Between 1927 and 1999) , (Between 2000 and 2016)]

Analysis

❖ Pivot tables & charts

- What is name & count of movies and average rating for directors (top 10)?

Director	Count of movie_title	Average of imdb_score
⊖ Irvin Kershner		
Star Wars: Episode V - The Empire Strikes Back	1	8.8
⊕ Charles Chaplin	1	8.6
⊕ Tony Kaye	1	8.6
⊖ Alfred Hitchcock		
Psycho	1	8.5
⊖ Ron Fricke		
Samsara	1	8.5
⊖ Damien Chazelle		
Whiplash	1	8.5
⊖ Majid Majidi		
Children of Heaven	1	8.5
⊕ Sergio Leone	3	8.433333333
⊕ Christopher Nolan	8	8.425
⊕ Asghar Farhadi	1	8.4
⊕ Richard Marquand	1	8.4
⊕ Marius A. Markevicius	1	8.4
⊕ S.S. Rajamouli	1	8.4

In this pivot table we displayed top 10 of directors based on average rating and also displayed the count of movies for each director and through this table we found that Irvin Kershner has highest Avg. rating 8.8 although Sergio Leone has 3 movies with Avg. rating 8.43 and Christopher Nalon has 8 movies with Avg. rating 8.425 This indicates Irvin's success.

- What is name & count of movies and average rating for actors (top 10)?

Actor	Count of movie_title	Average of imdb_score
Takashi Shimura	1	8.7
Bunta Sugawara	1	8.6
Paulette Goddard	1	8.6
Bahare Seddiqi	1	8.5
Emilia Fox	1	8.5
Collin Alfredo St. Dic	1	8.5
Janet Leigh	1	8.5
Mathieu Kassovitz	1	8.4
Tamannaah Bhatia	1	8.4
Claude Rains	1	8.4
Mhairi Calvey	1	8.4
Tommy Sheppard	1	8.4
Jürgen Prochnow	1	8.4
Shahab Hosseini	1	8.4

In this pivot table we displayed top 10 of actors based on average rating and also displayed the count of movies for each actor and through this table we found that Takashi Shimura has highest Avg. rating 8.7.

- What is the net profit and profit percentage of movies (top 10)?

Movie	gross	budget	net profit	Profit %
Avatar	760505847	237000000	523505847	221%
Jurassic World	652177271	150000000	502177271	335%
Titanic	658672302	200000000	458672302	229%
Star Wars: Episode IV - A	460935665	11000000	449935665	4090%
E.T. the Extra-Terrestrial	434949459	10500000	424449459	4042%
The Avengers	623279547	220000000	403279547	183%
The Lion King	422783777	45000000	377783777	840%
Star Wars: Episode I - The	474544677	115000000	359544677	313%
The Dark Knight	533316061	185000000	348316061	188%
The Hunger Games	407999255	78000000	329999255	423%

In this pivot table we displayed top 10 of movies based on net profit and also displayed gross, budget, profit percentage for each movie and through this table we found that Avatar movie has highest net profit 523M but we found that some movies have a profit percentage greater than 100% and also some movies have a profit percentage less than 0%.

Recommendations: Movies that have profit percentage greater than 100% can save budget or make a little increase on it but Movies that have profit percentage less than 0% can make a decrease budget later in the next movies.

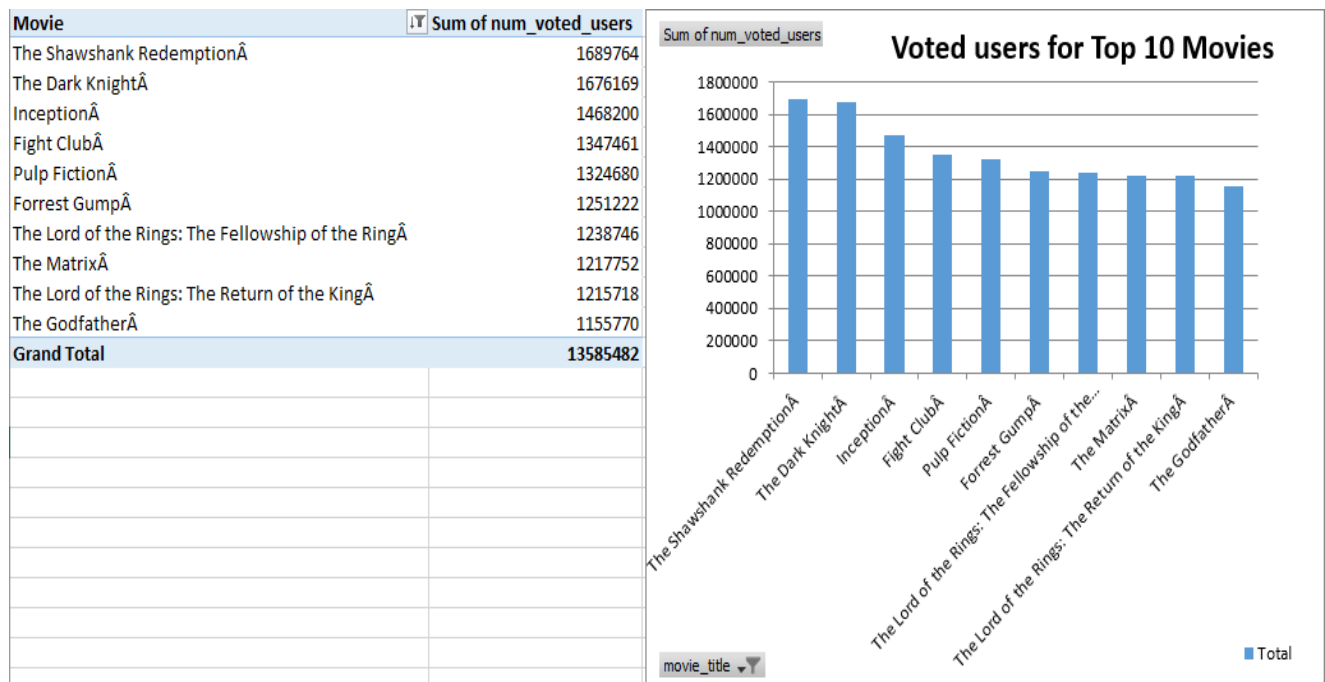
- What is the count of movies, net profit and avg. rating for movies before 2000 & after 2000?

Range of years	Count of movie_title	Sum of net profit	Average of imdb_score
Between 1927 and 1999	959	15936705818	6.642752868
Between 2000 and 2016	2561	6223086111	6.387934401

In this pivot table we displayed count of movies, net profit and avg. rating for movies before 2000 and after 2000 and through this table we found that sum of net profit and avg. rating in range between 1927 and 1999 greater than range between 2000 and 2016 although that count of movies after 2000 is 2561 and before 2000 is 959.

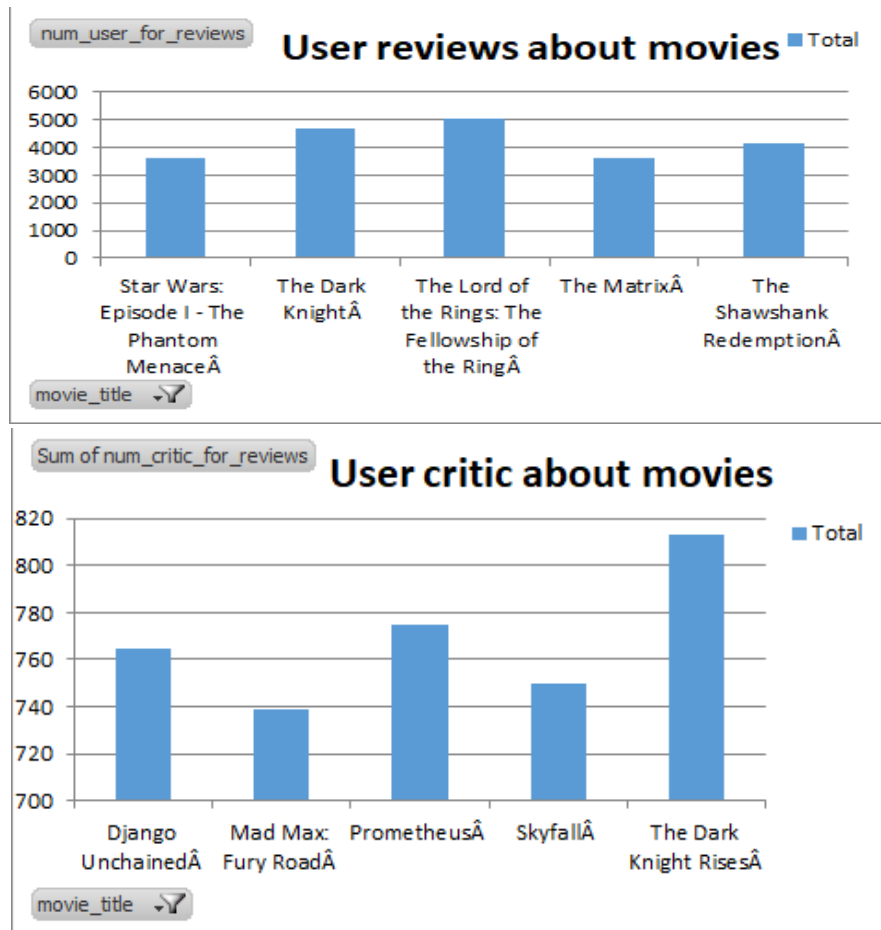
Recommendations: Finding the success factors that contributed to the success of films before 2000 and trying to apply them to current films in order to achieve the same success.

- What is the number of voted users for movies (top 10)?



In this chart we displayed number of voted users for movies and we found that The Shawshank Redemption movie has the highest number of voted users 1689764 and this indicates this movie success.

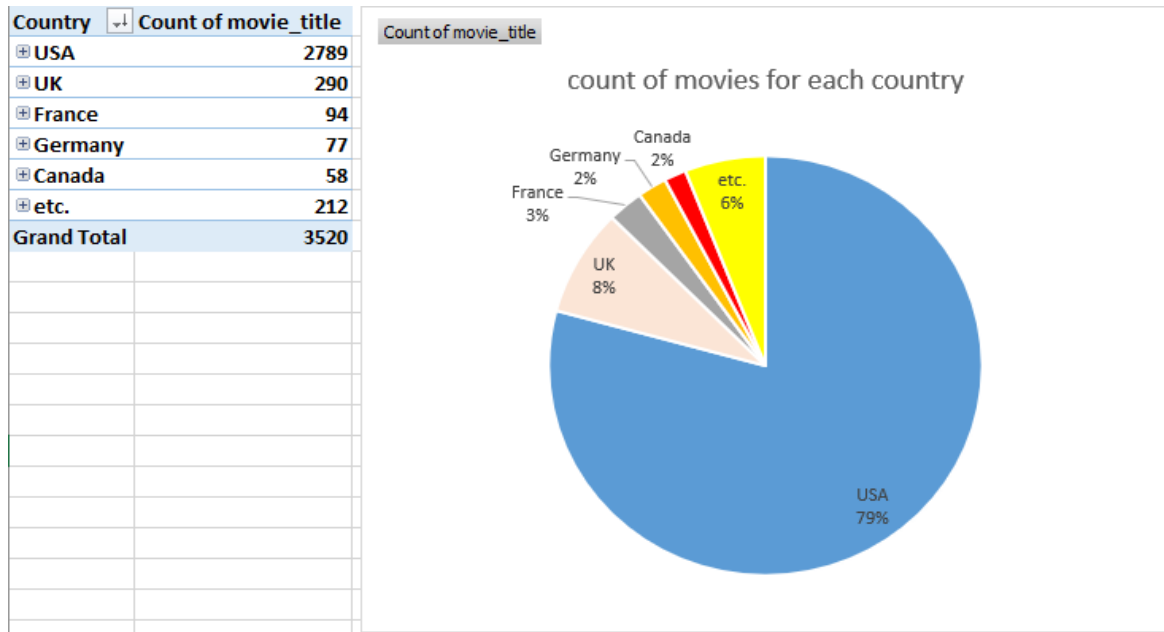
- What is number of user reviews and number of critic reviews for movies (top 5)?



In this chart we displayed number of user reviews and number of critic reviews for movies and we found that **The Lord of the Rings: The Fellowship of the Ring** movie has the highest number of user reviews 5060 and **The Dark Knight Rises** movie has the highest number of critic reviews 813 .

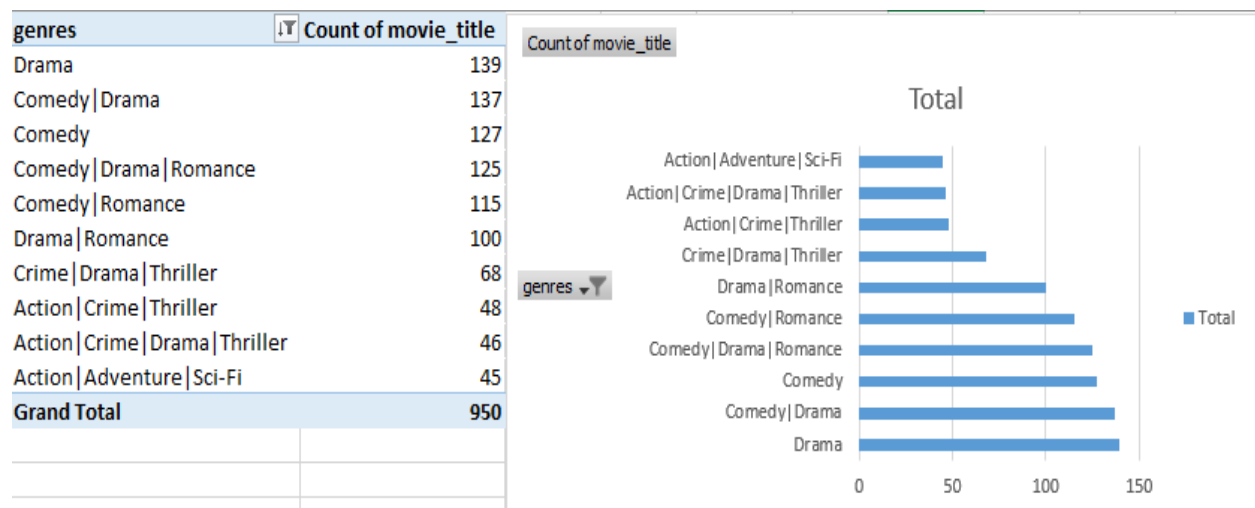
Recommendations: We must take user reviews, as well as critic reviews, whether they are positive or negative, and use them to maintain the success of films if they are positive. If they are negative, they should be used to try to improve upcoming works.

- What is count of movies for each country?



In this chart we displayed count of movies for each country and we found that **USA has the highest count of movies 2789 (79%)** and it comes after that **UK with 290 (8%)** and **France with 94 (3%)**.

- What is count of movies for genres?



In this chart we displayed count of movies for genres and we found that **Drama has the highest count of movies 139**.