IMDB Movie Analysis

DONE BY: G.REVAN

Description:

The IMDB Movie Analysis is a data-driven endeavor aimed at exploring and extracting insights from the vast collection of movies and related information available in teh given dataset. Through this analysis, we can uncover various trends, patterns, and correlations within the movie industry. They might delve into areas like the relationship between a movie's budget and its box office success, the impact of genres on audience preferences, the influence of actors and directors on a film's reception, and the evolution of cinematic trends over time.

DATA CLEANING:

This step involves preprocessing the data to make it suitable for analysis.

--->In the given dataset,there are missing values and null values,So remove all the rows containing null values.

--->There are no duplicates in the given dataset.

DATA ANALYSIS:

- In this step,check the datatypes of objects and the relationships between the data.
- The correlation among the genres ,director,budget,ratings.

Techstack used: MS EXCEL

INSIGHTS:

- We have analyzed the different genres and most common genres people like to watch.
- The highest number of movies in each language and their imdb score.
- The duration of movies affecting Imdb score and ratings of movies.
- The budget and profit margins analysis helps to find the factors for getting higher profits.

TASK 1:

Movie Genre Analysis: Analyze the distribution of movie genres and their impact on the IMDB score.

Task:

Determine the most common genres of movies in the dataset. Then, for each genre, calculate descriptive statistics (mean, median, mode, range, variance, standard deviation) of the IMDB scores.

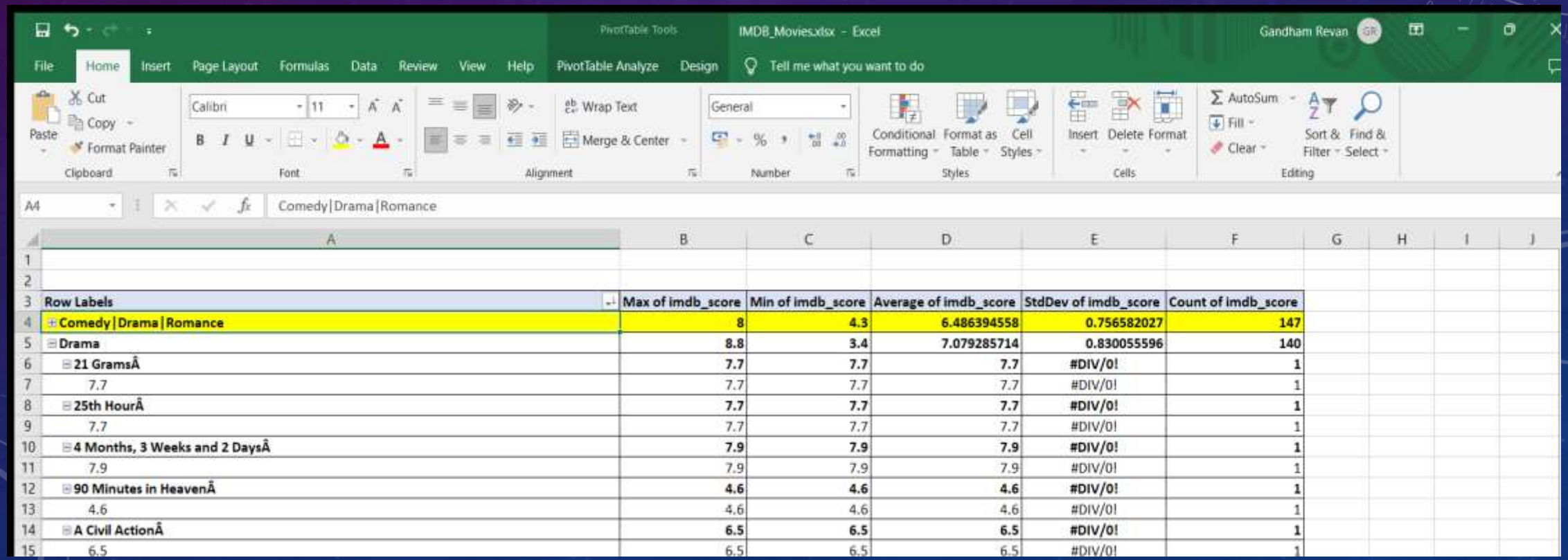
Result: **The most common genres are**

->Comedy|Drama|Romance(Movie count:147)

->Drama(Movie count: 140)

TASK 1:

Movie Genre Analysis: Analyze the distribution of movie genres and their impact on the IMDB score.



Row Labels	Max of imdb_score	Min of imdb_score	Average of imdb_score	StdDev of imdb_score	Count of imdb_score
Comedy|Drama|Romance	8	4.3	6.486394558	0.756582027	147
Drama	8.8	3.4	7.079285714	0.830055596	140
21 GramsÂ	7.7	7.7	7.7	#DIV/0!	1
7.7	7.7	7.7	7.7	#DIV/0!	1
25th HourÂ	7.7	7.7	7.7	#DIV/0!	1
7.7	7.7	7.7	7.7	#DIV/0!	1
4 Months, 3 Weeks and 2 DaysÂ	7.9	7.9	7.9	#DIV/0!	1
7.9	7.9	7.9	7.9	#DIV/0!	1
90 Minutes in HeavenÂ	4.6	4.6	4.6	#DIV/0!	1
4.6	4.6	4.6	4.6	#DIV/0!	1
A Civil ActionÂ	6.5	6.5	6.5	#DIV/0!	1
6.5	6.5	6.5	6.5	#DIV/0!	1

TASK 2:

Movie Duration Analysis: Analyze the distribution of movie durations and its impact on the IMDB score.

Task:

Analyze the distribution of movie durations and identify the relationship between movie duration and IMDB score.

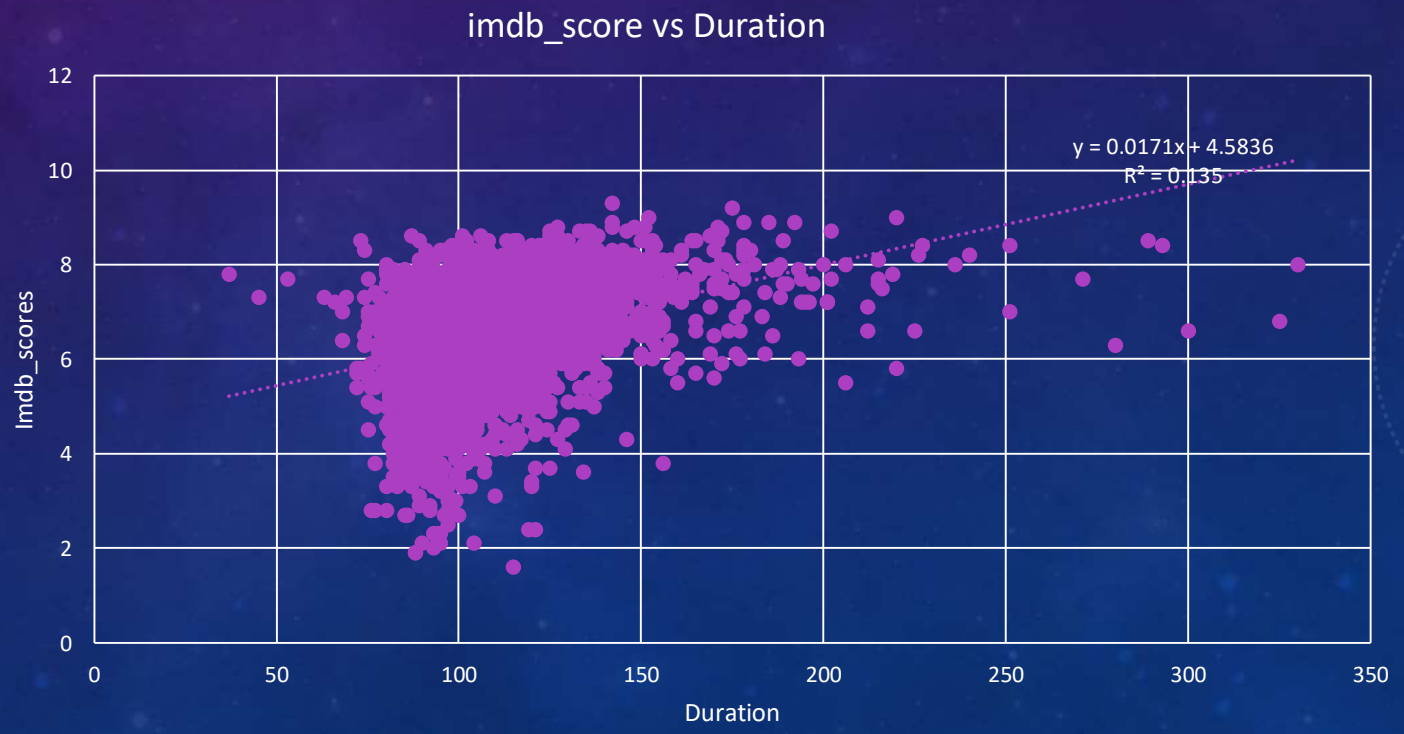
Result: The correlation between duration and score is approx. 0.3 indicating that if duration increases score also increases but as the correlaton is weak other also affecting score rather than duration alone like actors,director and budget.

Average duration most of users liked is 110minutes.

TASK 2:

Movie Duration Analysis: Analyze the distribution of movie durations and its impact on the IMDB score.

110.2635	Mean of duration
106	Median of duration
22.67832	Stddev of duration



TASK 3:

Language Analysis:

Situation: Examine the distribution of movies based on their language.

Task:

Determine the most common languages used in movies and analyze their impact on the IMDB score using descriptive statistics.

Result: More number of movies are in English and other than English the next top language movies are French , Spanish

TASK 3:

Language Analysis:

Situation: Examine the distribution of movies based on their language.

language	Number of movies	Mean of IMDB	Std dev of IMDB
English	3566	6.427509815	1.053502018
French	34	7.355882353	0
Spanish	23	7.082608696	0
Mandarin	14	7.021428571	0
Japanese	10	7.66	0
German	10	7.77	0
Italian	7	7.185714286	0
Cantonese	7	7.342857143	0
Portuguese	5	7.76	0
Korean	5	7.7	0
Hindi	5	7.22	0
Norwegian	4	7.15	0
Persian	3	8.133333333	0
Danish	3	7.9	0

TASK 4:

Director Analysis: Influence of directors on movie ratings.

Task:

Identify the top directors based on their average IMDB score and analyze their contribution to the success of movies using percentile calculations.

Result: The Top director Akira Kurosawa whose average IMDB score is 8.7.

TASK 4:

Director Analysis: Influence of directors on movie ratings.

director_name	imdb_score	Average imdb score for each director	imdbscore > percentile value
Akira Kurosawa	8.7	8.7	Yes
Tony Kaye	8.6	8.6	Yes
Charles Chaplin	8.6	8.6	Yes
Damien Chazelle	8.5	8.5	Yes
Majid Majidi	8.5	8.5	Yes
Ron Fricke	8.5	8.5	Yes
Alfred Hitchcock	8.5	8.5	Yes
Sergio Leone	8	8.433333333	Yes
Sergio Leone	8.9	8.433333333	Yes
Sergio Leone	8.4	8.433333333	Yes

TASK 5:

Budget Analysis: Explore the relationship between movie budgets and their financial success.

Task:

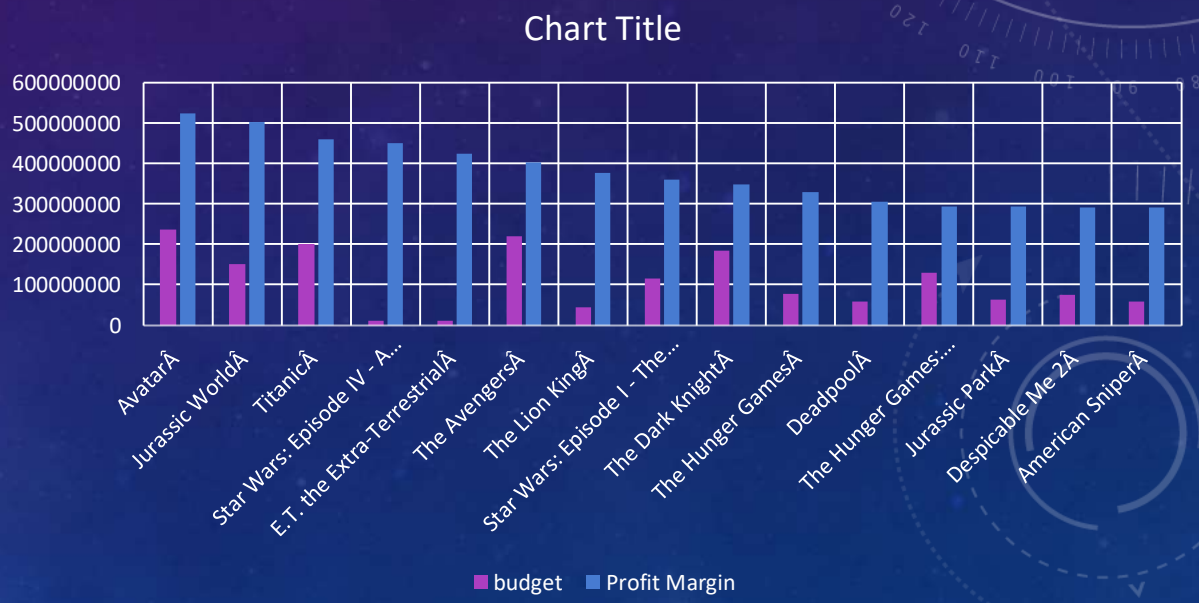
Analyze the correlation between movie budgets and gross earnings and identify the movies with the highest profit margin.

Result: One factor influencing profit margin is the higher budget and the other margin is the Genre of movie and actors.

TASK 5:

Director Analaysis: Explore the relationship between movie budgets and their financial success.

movie_title	budget	Profit Margin
AvatarÂ	237000000	523505847
Jurassic WorldÂ	150000000	502177271
TitanicÂ	200000000	458672302
Star Wars: Episode IV - A New HopeÂ	11000000	449935665
E.T. the Extra-TerrestrialÂ	10500000	424449459
The AvengersÂ	220000000	403279547
The Lion KingÂ	45000000	377783777
Star Wars: Episode I - The Phantom MenaceÂ	115000000	359544677
The Dark KnightÂ	185000000	348316061
The Hunger GamesÂ	78000000	329999255
DeadpoolÂ	58000000	305024263
The Hunger Games: Catching FireÂ	130000000	294645577
Jurassic ParkÂ	63000000	293784000
Despicable Me 2Â	76000000	292049635
American SniperÂ	58800000	291323553



RESULT:

- We have analyzed the dataset and answered all the tasks provided with charts and keyfindings.
- The excel sheet link for the analysis and findings: https://docs.google.com/spreadsheets/d/1Sk3N6cYH7j7zAswwd5ySex5a-Xl-CyxA/edit?usp=drive_link&ouid=107111635379336937367&rtpof=true&sd=true
- From this project we can get meaningful insights like what factors influence movie ratings and its profit.