

# IMDB Movie Analysis

## PROJECT DESCRIPTION :

The IMDB Movie Analysis project is designed to explore and examine a rich dataset of movies sourced from the IMDb platform. This dataset includes key details such as directors, movie titles, duration, genres, budgets, gross earnings, IMDb ratings, and more. By applying in-depth data analysis, data visualization, and statistical techniques in Excel, the project aims to uncover meaningful insights and trends that help explain what drives a movie's success.

As part of this project, I was tasked with preparing a detailed report for the provided data record, addressing the questions that follow.

A. 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.

B. 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.

C. 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.

. D. 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

E. 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. Remember, these tasks are designed to progressively explore different aspects of the dataset and uncover meaningful insights. Each task builds upon the previous one to provide a comprehensive analysis of the IMDB movie data.

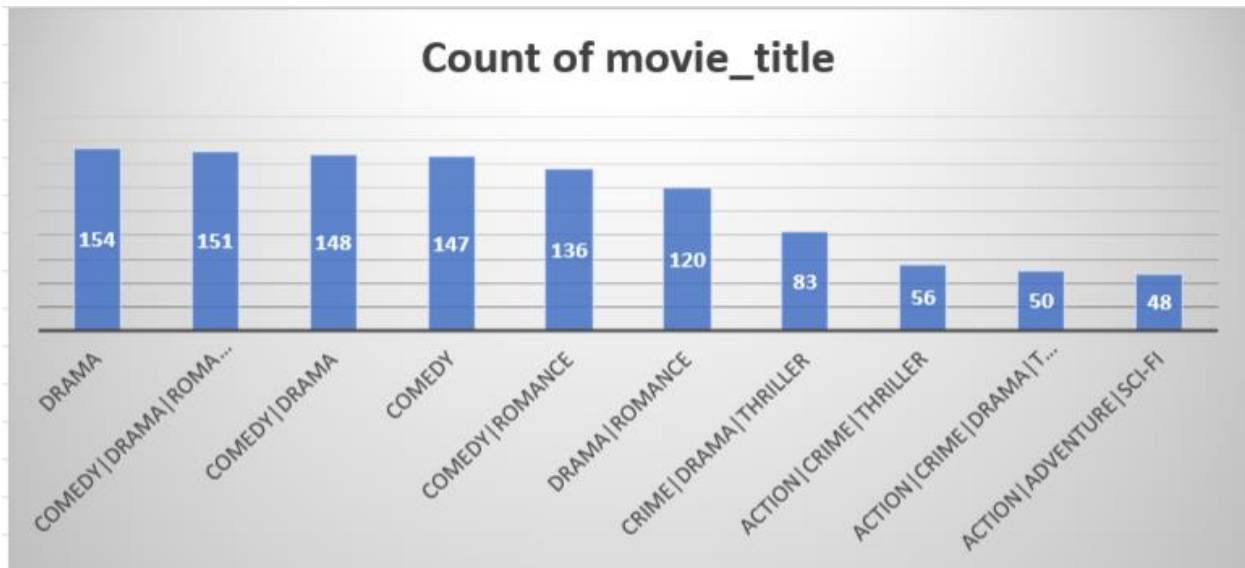
## **Cleaning the Data :**

- **In this data analysis project, the first and most vital step is data cleaning.**
- **Begin by dropping columns that do not contribute to the analysis, followed by removing any null values. These columns are removed because they do not affect the insights we want to extract.**
- **After dropping irrelevant columns, check for blank rows. Blank rows can distort results, so they must be removed. Apply filters to find blanks in each column and delete the entire row where they occur.**
- **Lastly, check for duplicate rows and remove them to ensure the dataset remains accurate and reliable.**

## A. 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.

Genres	Count of movie_title	MEAN	MAX	MIN	MEDIAN	MODE	VARIANCE	STD DEV
Drama	154	7.0	8.8	3.4	7.2	7.3	0.68630	0.828430
Comedy Drama Romance	151	6.5	8	4.3	6.5	6.5	0.56277	0.750181
Comedy Drama	148	6.6	8.8	3.3	6.7	6.7	0.80266	0.895911
Comedy	147	5.9	8	1.9	6	6.5	1.48301	1.217790
Comedy Romance	136	5.9	8.4	2.7	6	6.1	0.76288	0.873432
Drama Romance	120	7.0	8.1	4.1	7.1	7.2	0.55807	0.747039
Crime Drama Thriller	83	6.9	8.5	5.1	7	6.1	0.61335	0.783167
Action Crime Thriller	56	6.4	7.6	4.4	6.5	6.5	0.39361	0.627384
Action Crime Drama Thriller	50	6.5	9	5.1	6.5	6.1	0.51408	0.716992
Action Adventure Sci-Fi	48	6.7	8.4	2.4	6.8	6.6	1.54170	1.241651

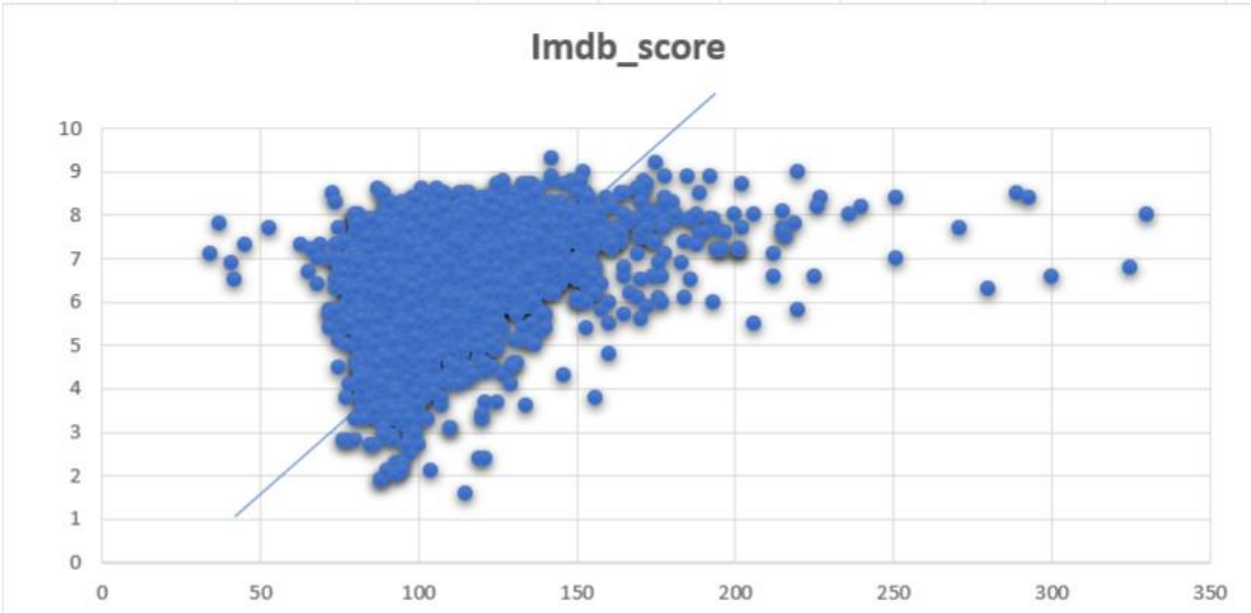


The IMDb movie ratings show that drama and drama-romance movies usually get the highest scores, meaning people like them the most. Crime and thriller movies also do well. Pure comedies often get lower ratings and have mixed reviews — some are great, but many don't do well. Action and sci-fi movies can be big hits or big misses, so they're less predictable. Overall, adding drama or romance seems to make movies more popular. So, if someone wants to make a well-liked movie, choosing strong stories with drama or thrillers is a safer bet than making only comedies.

**B. 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.

	Duration	Imdb_score
Mean	109.915036	6.463542739
Median	106	6.6
Mode	101	6.7
variancve	515.9696013	1.11517323
STD DEV	22.71496426	1.056017628

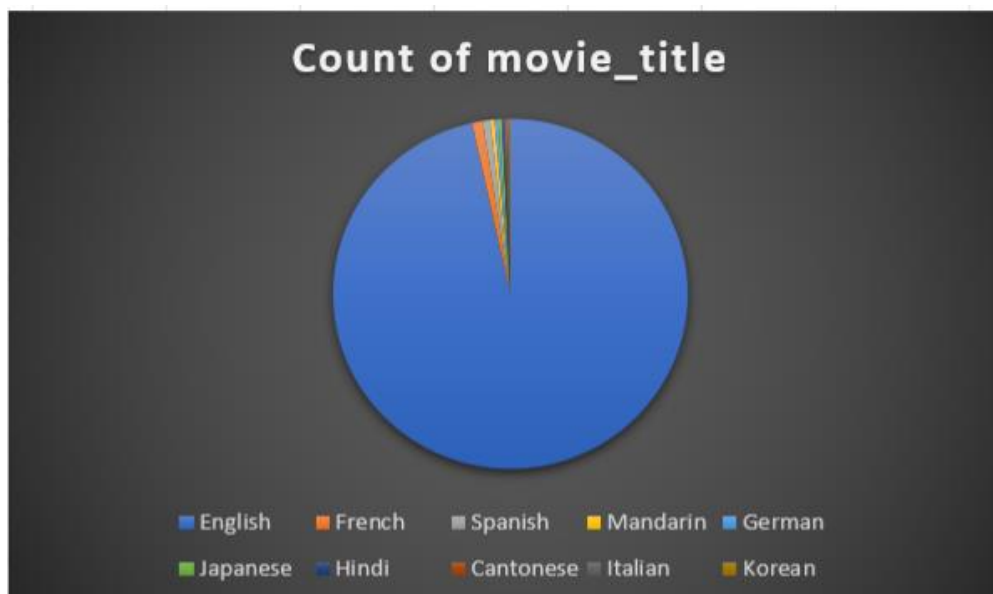


The average movie in this dataset is about 110 minutes long, with most movies around 106 minutes. Many movies have a runtime close to 101 minutes, showing that movies tend to stay near the 1.5–2 hour mark. The average IMDb score is about 6.5, with most movies rated between 6.6 and 6.7. This means most movies are rated as average to good. The variation in scores is low, so ratings don’t change wildly. Overall, movies are mostly of medium length and get decent but not outstanding audience scores.

### C. 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.

Language	Count of movie_title	Mean	Median	STD DEV
English	3703	6.423251418	6.5	1.050986526
French	37	7.286486486	7.2	0.561328861
Spanish	26	7.05	7.15	0.826196103
Mandarin	15	7.08	7.4	0.772010363
German	13	7.692307692	7.7	0.640912811
Japanese	12	7.625	7.8	0.899621132
Hindi	10	6.76	7.05	1.111755369
Cantonese	8	7.2375	7.3	0.440575922
Italian	7	7.185714286	7	1.155318962
Korean	5	7.7	7.7	0.570087713
Portuguese	5	7.76	8	0.978774744



English movies make up most of the dataset but have a lower average score (6.4) compared to movies in other languages. Movies in French, German, Japanese, Korean, and Portuguese have higher average IMDb scores, mostly above 7.5, showing they are often well-received. German and Portuguese movies top the list with averages near 7.7–7.8 and high medians too. Hindi movies have a slightly above-average score of 6.7. Overall, while Hollywood produces the most films, smaller sets of foreign-language films tend to score higher, suggesting they're often more appreciated by viewers.

English movies have the widest variation in ratings, showing both hits and flops. Smaller language groups like German, French, and Cantonese have lower standard deviations, meaning their ratings are more consistent. Italian and Hindi films show mixed reviews, but overall, foreign-language movies tend to get steadier audience scores.

#### D. 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.

Director	MEAN	Percentile
Charles Chaplin	8.6	0.999
Tony Kaye	8.6	0.999
Alfred Hitchcock	8.5	0.997
Damien Chazelle	8.5	0.997
Majid Majidi	8.5	0.997
Ron Fricke	8.5	0.997
Sergio Leone	8.4	0.996
Christopher Nolan	8.4	0.995
Asghar Farhadi	8.4	0.993
Marius A. Markevicius	8.4	0.993

Directors like Charles Chaplin and Tony Kaye score an average IMDb rating of 8.6, putting them in the top 90th percentile of all directors — meaning their movies are among the best rated. Big names like Hitchcock and Nolan also rank high with scores above 8.4. These directors have high scores because they tell unique stories, create memorable characters, and often make movies that stand the test of time. They rarely make bad films, so people trust their work. Their movies often have strong plots, good acting, and emotional impact, which keeps their ratings consistently high.

Charles Chaplin is loved for timeless classics like *Modern Times* and *The Great Dictator*. Tony Kaye is known for *American History X*, a powerful drama. Alfred Hitchcock made legendary thrillers like *Psycho* and *Rear Window*. Damien Chazelle directed the award-winning *La La Land* and *Whiplash*. Majid Majidi's *Children of Heaven* is an acclaimed Iranian film. Ron Fricke created stunning documentaries like *Baraka*. Sergio Leone is famous for westerns like *The Good, the Bad and the Ugly*. Christopher Nolan made hits like *Inception* and *The Dark Knight*. Asghar Farhadi directed *A Separation*. Marius A. Markevicius made *The Other Dream Team*.



## E. 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. Remember, these tasks are designed to progressively explore different aspects of the dataset and uncover meaningful insights. Each task builds upon the previous one to provide a comprehensive analysis of the IMDB movie data.

NO.	movie_title	Budget	Gross	Profit
1	Avatar	237000000	760505847	523505847
2	Jurassic World	150000000	652177271	502177271
3	Titanic	200000000	658672302	458672302
4	Star Wars: Episode IV - A New Hope	11000000	460935665	449935665
5	E.T. the Extra-Terrestrial	10500000	434949459	424449459
6	The Avengers	220000000	623279547	403279547
7	The Avengers	220000000	623279547	403279547
8	The Lion King	45000000	422783777	377783777
9	Star Wars: Episode I - The Phantom Menace	115000000	474544677	359544677
10	The Dark Knight	185000000	533316061	348316061



This list shows blockbuster movies that earned huge profits. *Avatar* leads with over \$520 million profit, showing how big-budget films can pay off. *Jurassic World* and *Titanic* follow close behind with massive earnings. Interestingly, older classics like *Star Wars: A New Hope* and *E.T.* made huge profits with much smaller budgets, proving that great stories can succeed without massive spending. Animated hit *The Lion King* and superhero movies like *The Avengers* and *The Dark Knight* also made strong profits. Overall, successful movies often mix big budgets, famous franchises, and universal appeal to reach large audiences and earn huge returns.

Movies with bigger budgets often earn bigger profits, as seen with *Avatar*, *Jurassic World*, and *Titanic*. However, some films like *Star Wars: A New Hope* and *E.T.* made huge profits despite smaller budgets, showing that a good story and strong audience connection can deliver big returns without needing massive spending.

Big-budget films like *Avatar*, *Jurassic World*, and *Titanic* show that spending more can bring huge profits — but it's not just the money. These movies also have visionary directors: James Cameron (*Avatar*, *Titanic*) is known for creating groundbreaking hits. *Star Wars* (*A New Hope*, *The Phantom Menace*) under George Lucas proves that an iconic story and creative direction can make huge profits even with a modest budget. Spielberg's *E.T.* and Nolan's *The Dark Knight* show how skilled directors turn strong ideas into blockbuster successes. So, budget helps, but the director's talent and storytelling make the real difference.