IMDB Movie Analysis

 <u>Project Overview</u> :

This project focuses on analyzing an IMDB movie dataset to uncover the factors that influence a movie's success, defined by high IMDB ratings.

<u>Objective</u> : To identify key elements such as budget, genre, actors, directors, and other factors that contribute to higher movie ratings on IMDB.

<u>GOAL</u> : Provide actionable insights for movie producers, directors, and investors to make informed decisions about future film projects

<u>EXCEL SHEET⬇</u>

[Excel Sheet (Click Here)](#)

Approach and Execution

Data Collection & Cleaning:

The dataset was cleaned by handling missing values, converting data types, and removing duplicates.

Exploratory Data Analysis (EDA):

Analyzed relationships between different factors like budget, genre, director, and IMDB ratings. Various visualizations were created using tools like histograms, bar charts, and scatter plots.

Techniques Used:

- Correlation matrix to find relationships between variables.

- Regression analysis to determine which factors have the most influence on IMDB ratings.

Tech- Stack Used

Microsoft Excel 2022:

- Data Cleaning: Used to handle missing data, remove duplicates, and convert data types for consistency.

Data Analysis:

- Applied filters, pivot tables, and conditional formatting to explore relationships between variables like budget, ratings, and genres.

- Utilized Excel formulas (such as SUMIF, AVERAGEIF, VLOOKUP) for calculations and summaries.

Visualizations: Created charts such as bar graphs, scatter plots, and pie charts to visualize trends and insights in the dataset.

Trend Analysis: Used Excel's built-in functions for trend lines and regression to analyze how factors like budget, cast, and director influence IMDB ratings.

Insights

Top Factors for Success:
- **High Budget** and **Popular Cast/Directors** lead to higher ratings.

Trends Identified:
- **Action/Adventure genres** perform better in ratings.
- **Newer films** slightly outperform older ones in average ratings.

- ## Conclusion:
  The analysis provides valuable insights into the factors influencing movie success, helping stakeholders make informed decisions.
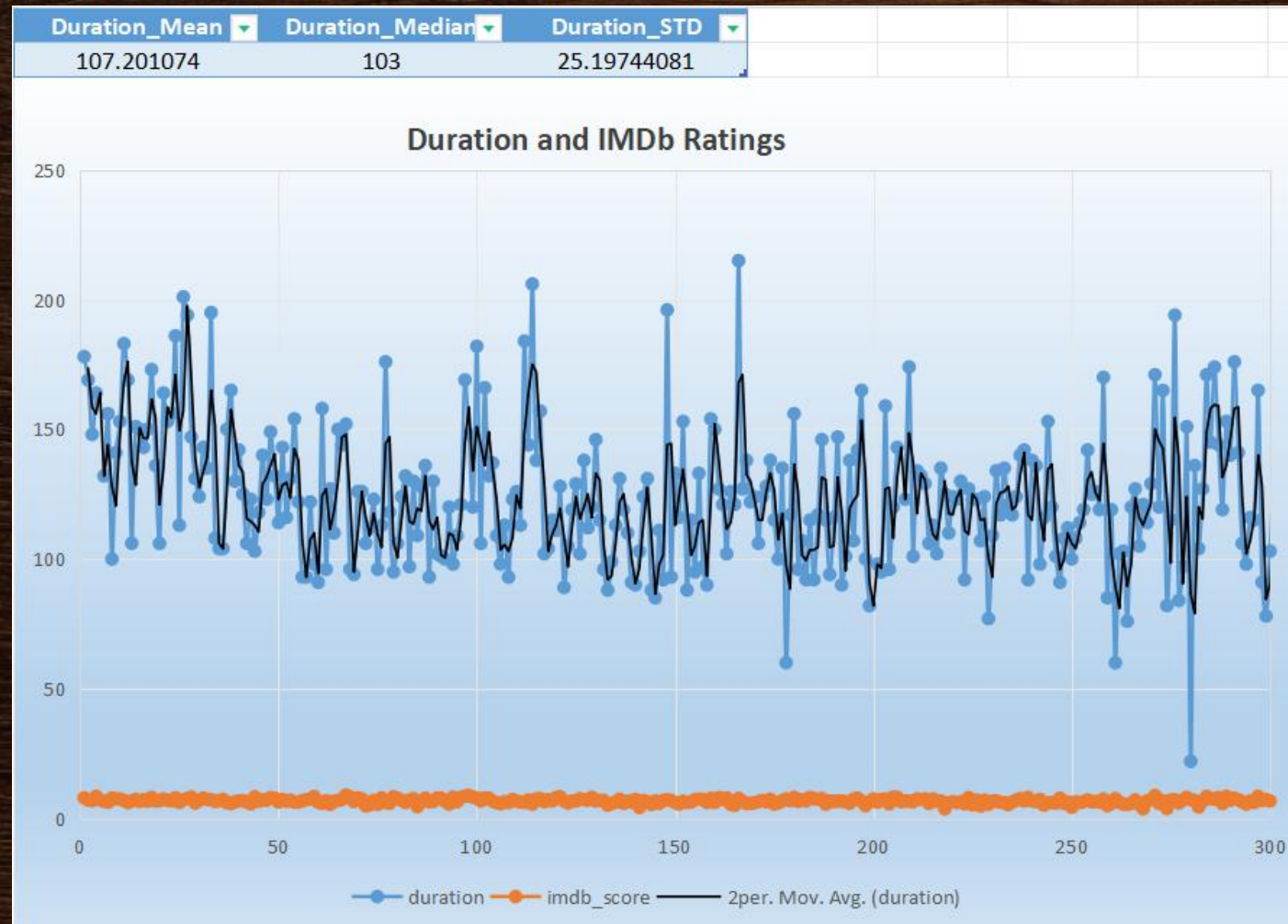
RESULTS

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

genres	Count of genres	MEAN	MEDIAN	MODE	RANGE	VARIANC	STD
Action	1153.00	6.24	6.30	6.70	5.40	1.25	1.12
Adventure	453.00	6.44	6.60	6.10	4.50	1.28	1.13
Animation	61.00	6.58	6.70	6.70	1.10	1.30	1.14
Biography	252.00	7.15	7.20	7.00	0.00	0.52	0.72
Comedy	1329.00	6.20	6.30	6.70	7.60	1.19	1.09
Crime	349.00	6.56	6.60	6.60	0.00	1.05	1.03
Documenta	84.00	7.18	7.40	7.50	6.00	1.12	1.06
Drama	972.00	6.76	6.90	7.20	5.80	0.92	0.96
Family	11.00	6.25	6.40	6.70	5.40	1.44	1.20
Fantasy	53.00	6.31	6.40	6.70	0.40	1.35	1.16
Film-Noir	1.00	7.63	7.65	6.60	0.00	0.19	0.43
Game-Show	1.00	2.90	2.90	7.40	0.00	1.27	1.15
History	1.00	7.08	7.20	7.50	0.00	0.79	0.89
Horror	232.00	5.84	5.90	6.20	5.80	1.28	1.13
Music	1.00	6.41	6.60	6.50	0.00	1.39	1.18
Musical	4.00	6.51	6.70	7.00	0.00	1.50	1.23
Mystery	33.00	6.49	6.60	6.60	0.00	1.19	1.09
Romance	6.00	6.45	6.50	6.50	1.50	0.99	1.00
Sci-Fi	13.00	6.28	6.40	6.70	3.50	1.47	1.21
Thriller	22.00	6.31	6.40	6.10	4.10	1.11	1.05
Western	12.00	6.69	6.80	6.50	5.10	1.09	1.04
(blank)							
Grand Total	5043.00						

RESULTS :

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

Results

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

language	Count_Of_Lang	Mean	Median	STD
English	4704	6.40	6.50	1.12
Japanese	18	7.39	7.60	0.96
French	73	7.04	7.20	0.72
Mandarin	26	6.79	7.05	1.02
Aboriginal	2	6.95	6.95	0.55
Spanish	40	6.94	7.15	0.84
Filipino	1	6.70	6.70	0.00
Hindi	28	6.63	6.95	1.37
Russian	11	6.36	6.50	1.32
Maya	1	7.80	7.80	0.00
Kazakh	1	6.00	6.00	0.00
Telugu	1	8.40	8.40	0.00
Cantonese	11	6.95	7.20	0.67
Icelandic	2	7.55	7.55	0.65
German	19	7.34	7.60	0.93

More In Excel Sheet...

Results

D. Director Analysis: Influence of directors on movie ratings.

director_name	Average of imdb_score	Percentile
Akira Kurosawa	8.10	8.38
Alejandro G. Iñárritu	7.78	8.39
Alex Garland	7.70	8.4
Alex Gibney	7.70	8.4
Alfonso Cuarón	7.80	8.4
Alice Wu	7.60	8.4
Amal Al-Agroobi	8.20	8.4
Andrei Tarkovsky	8.10	8.4
Andrew Haigh	7.70	8.4

Results

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

Correlation	Profit Margin
0.102179454	523505847
	9404152
Highest Profit	-44925825
523505847	198130642
	0
Hight Profit Movie	-190641321
1	78530303
	-59192738
Movie name	208991599
Avatar	51956980
	80249062
Correlation Coefficient	-8930592
0.102179454	-31631573
	198032628