IMDB Movie Analysis

Problem Statement:

The dataset provided is related to IMDB Movies. A potential problem to investigate could be: "What factors influence the success of a movie on IMDB?" Here, success can be defined by high IMDB ratings. The impact of this problem is significant for movie producers, directors, and investors who want to understand what makes a movie successful to make informed decisions in their future projects.

Approach:

Data Cleaning: Here are the steps for cleaning data in an IMDb movie analysis using Excel:

1. First I import the data and then analyze it.
2. Find the relevant columns for further use.
3. Find blank cell and delete it.
4. Remove the irrelevant columns from data.
5. Lastly find the duplicate data delete it.

Tech Task Used:

Microsoft Excel:

- **Version:** Microsoft 365 Excel is a part of this suite and offers the latest version of Excel, which is Excel 2019 or later.
- **Key Features:**
  - **Pivot Tables:** For summarizing data.
  - **Charts:** For data visualization.
  - **Formulas and Functions:** For calculations and data analysis (e.g., SUM, AVERAGE, CORREL, MEDIAN, STDEV, COUNTIF).

Data Analytics Tasks:

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.

ACTION	**Descriptive Statistics**	
	Column1	**Column2**
	AVERAGE	222.5
	MEDIAN	38
	MODE	3
	MAX	1008
	MIN	1
	VAR	116508.2647
	STDEV	341.33307

genres	**Count of genres**
Action	1008
Adventure	330
Animation	43
Biography	206
Comedy	997
Crime	257
Documentary	27
Drama	668
Family	4
Fantasy	38
Horror	165
Musical	2
Mystery	22
Romance	1
Sci-Fi	8
Thriller	3
Western	3
(blank)	
Grand Total	**3782**

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.

Descriptive Statistics	Values
Mean of Duration	110.1
Median	106
Standard Deviation	22.7

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 language	Average	Median	Standard Deviation
Aboriginal	2	7.0	6.6	1.1
Arabic	1	7.2	6.6	1.1
Aramaic	1	7.1	6.6	1.1
Bosnian	1	4.3	6.6	1.1
Cantonese	7	7.3	6.6	1.1
Czech	1	7.4	6.6	1.1
Danish	3	7.9	6.6	1.1
Dari	2	7.5	6.6	1.1
Dutch	3	7.6	6.6	1.1
English	3625	6.4	6.6	1.1
Filipino	1	6.7	6.6	1.1
French	34	7.4	6.6	1.1
German	10	7.8	6.6	1.1
Hebrew	1	8.0	6.6	1.1
Hindi	5	7.2	6.6	1.1
Hungarian	1	7.1	6.6	1.1
Indonesian	2	7.9	6.6	1.1
Italian	7	7.2	6.6	1.1
Japanese	10	7.7	6.6	1.1
Kazakh	1	6.0	6.6	1.1
Korean	5	7.7	6.6	1.1
Mandarin	14	7.0	6.6	1.1
Maya	1	7.8	6.6	1.1

Mongolian	1	7.3	6.6	1.1
None	1	8.5	6.6	1.1
Norwegian	4	7.2	6.6	1.1
Persian	3	8.1	6.6	1.1
Portuguese	5	7.8	6.6	1.1
Romanian	1	7.9	6.6	1.1
Russian	1	6.5	6.6	1.1
Spanish	23	7.1	6.6	1.1
Thai	3	6.6	6.6	1.1
Vietnamese	1	7.4	6.6	1.1
Zulu	1	7.3	6.6	1.1

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_name	Average of imdb_score	Percentile
Akira Kurosawa	8.70	8.70
Michael Curtiz	8.60	8.60
Charles Chaplin	8.60	8.60
Tony Kaye	8.60	8.60
Ron Fricke	8.50	8.50
Damien Chazelle	8.50	8.50
Majid Majidi	8.50	8.50
Sergio Leone	8.43	8.48
Christopher Nolan	8.43	8.48
Asghar Farhadi	8.40	8.48
Richard Marquand	8.40	8.48

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.

Correlation	Highest Profit Margin
0.10	523505847

Result:

In this way, I done IMDB Movie Analysis Project using Excel. All tasks and its result are clearly mentioned in the report above.