IMDB Movie Analysis

Project Description

The IMDB Movie Analysis project aimed to uncover the factors that influence a movie's success as defined by high IMDB ratings. By analyzing key aspects such as genres, durations, languages, directors, and budgets, this project provides valuable insights to producers, directors, and investors. The ultimate objective was to identify trends and patterns that contribute to better decision-making in the movie industry.

Approach

To achieve the objectives, the following steps were undertaken:

1. **Data Cleaning:**

   o   Handled missing values by imputing or removing them as appropriate.

   o   Removed duplicate entries to ensure data integrity.

2. **Data Analysis:**

   o   Conducted descriptive statistics for numerical columns.

   o   Used filters and Excel functions (e.g., COUNTIF, AVERAGE, MEDIAN, etc.) to analyze categorical data.

   o   Created visualizations (bar graphs, scatter plots) to better understand relationships.

3. **Feature Engineering:**

   o   Derived new columns such as profit margin (gross earnings minus budget).

   o   Categorized durations and budgets into intervals for further analysis.
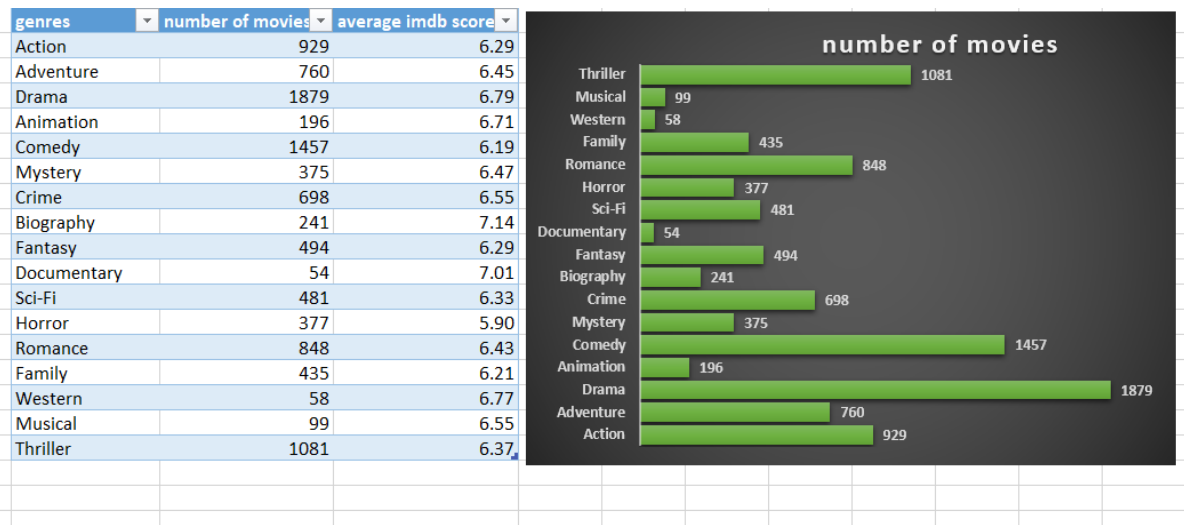
Tech-Stack Used

- **Microsoft Excel 2022:**

   o   Used for data cleaning, calculations, and creating visualizations.

   o   Functions like COUNTIF, AVERAGE, MEDIAN, STDEV, and CORREL were extensively utilized.

Insights
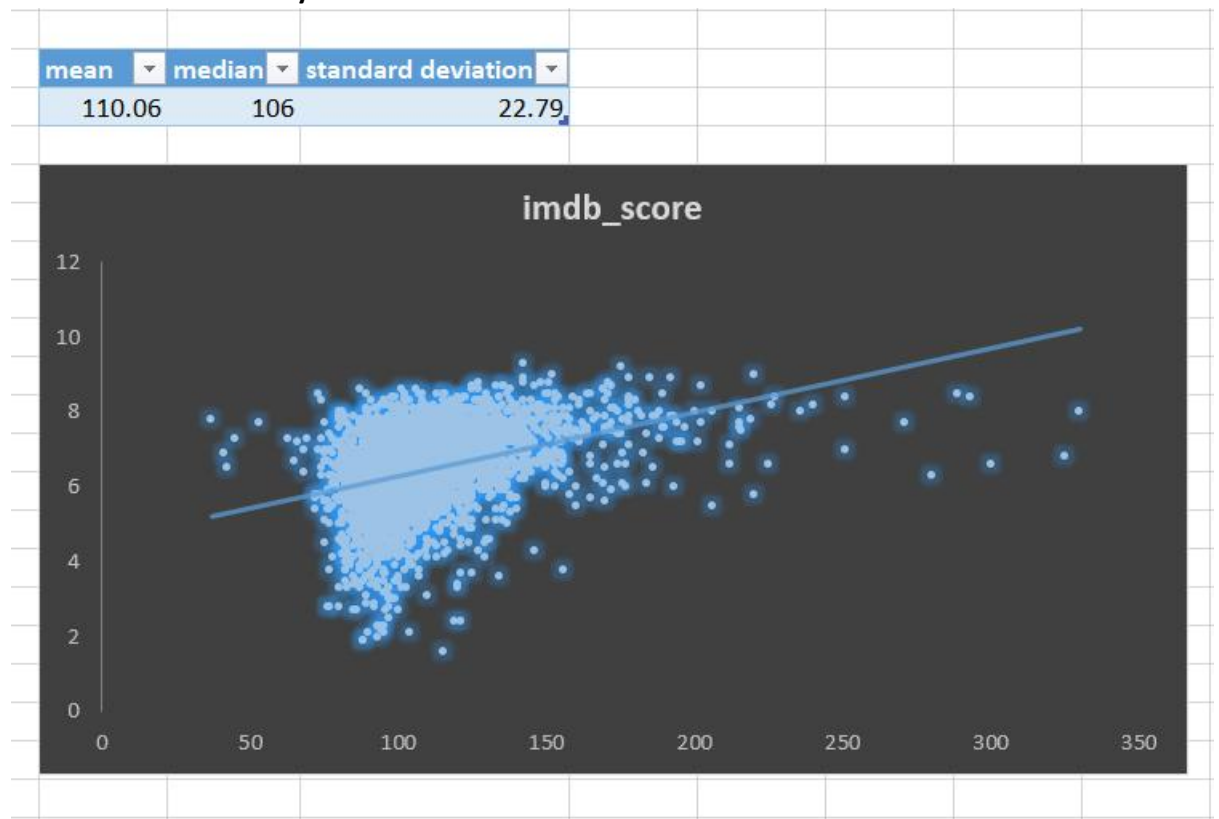
   A. Movie Genre Analysis:

Analyze the distribution of movie genres and their impact on the IMDB score.

genres	number of movies	average imdb score
Action	929	6.29
Adventure	760	6.45
Drama	1879	6.79
Animation	196	6.71
Comedy	1457	6.19
Mystery	375	6.47
Crime	698	6.55
Biography	241	7.14
Fantasy	494	6.29
Documentary	54	7.01
Sci-Fi	481	6.33
Horror	377	5.90
Romance	848	6.43
Family	435	6.21
Western	58	6.77
Musical	99	6.55
Thriller	1081	6.37

number of movies

Most common genres: Drama, and Comedy.

Genres like Biography and Documentary had higher mean IMDB scores, indicating their popularity.

B. Movie Duration Analysis:

mean	median	standard deviation
110.06	106	22.79

imdb_score

Mean duration of movies was approximately 110 minutes.

Scatter plot analysis showed a slight positive correlation between longer durations and higher IMDB scores.

C. Language Analysis:

language	no of movies	mean	median	mode
English	3543	6.4	6.6	6.7
French	36	7.3	6.6	6.7
Spanish	23	7.1	6.6	6.7
Mandarin	14	7.0	6.6	6.7
German	13	7.7	6.6	6.7
Japanese	12	7.6	6.6	6.7
Cantonese	8	7.2	6.6	6.7
Hindi	8	7.2	6.6	6.7
Italian	7	7.2	6.6	6.7

English dominated the dataset, but non-English movies often had comparable or higher average IMDB scores.

D. Director Analysis:

90th percentile	7.5
7.5 refers to	90% of the imdb scores are equal to or below 7.5
top direcotors	183

director_name	average imdb score
Tony Kaye	8.6
Charles Chaplin	8.6
Alfred Hitchcock	8.5
Ron Fricke	8.5
Damien Chazelle	8.5
Majid Majidi	8.5
Sergio Leone	8.4
Christopher Nolan	8.4
S.S. Rajamouli	8.4
Richard Marquand	8.4
Asghar Farhadi	8.4
Lee Unkrich	8.3
Fritz Lang	8.3
Lenny Abrahamson	8.3
Billy Wilder	8.3
Charles Ferguson	8.3
Pete Docter	8.2
Hayao Miyazaki	8.2
Quentin Tarantino	8.2
George Roy Hill	8.2
Juan José Campanella	8.2
Joshua Oppenheimer	8.2

Elia Kazan	8.2
Victor Fleming	8.2
Milos Forman	8.1
Akira Kurosawa	8.1
Tim Miller	8.1
Terry George	8.1
Je-kyu Kang	8.1
William Wyler	8.1
David Sington	8.1
Michael Wadleigh	8.1
David Lean	8.0
Michel Hazanavicius	8.0
Stephen Chbosky	8.0
Vincent Paronnaud	8.0
Ari Folman	8.0
Frank Darabont	8.0
Denis Villeneuve	8.0
James Cameron	7.9
Jacques Perrin	7.9
Tom McCarthy	7.9
Don Hall	7.9
George Cukor	7.9
StÃ©phane Aubier	7.9
Dan Gilroy	7.9
Jonathan Dayton	7.9
Christophe Barratier	7.9
FabiÃ¡n Bielinsky	7.9
Anna Muylaert	7.9
Cristian Mungiu	7.9
Peter Jackson	7.9
Joss Whedon	7.9
Morten Tyldum	7.9
Alejandro G. IÃ±Ã¡rritu	7.8
Stanley Kubrick	7.8
Alfonso CuarÃ³n	7.8
Nathan Greno	7.8
Rich Moore	7.8
Bernardo Bertolucci	7.8
Giuseppe Tornatore	7.8
Christian Carion	7.8
Mark Herman	7.8
Josh Boone	7.8
Robert Stevenson	7.8
Mike van Diem	7.8
Jim Abrahams	7.8
Sylvain Chomet	7.8
Ralph Ziman	7.8
Ritesh Batra	7.8
Mark Sandrich	7.8
Henry Alex Rubin	7.8

Dean DeBlois	7.8
Mel Gibson	7.8
Michael Moore	7.8
Richard Curtis	7.8
James Marsh	7.8
David Fincher	7.8
Andrew Stanton	7.7
Tom Hooper	7.7
Peter Weir	7.7
Shane Meadows	7.7
Eric Bress	7.7
Alex Garland	7.7
Brian Henson	7.7
Paolo Sorrentino	7.7
Chuan Lu	7.7
Philip Saville	7.7
FranÃ§ois Girard	7.7
Tomm Moore	7.7
Caroline Link	7.7
Wolfgang Becker	7.7
Kevin Macdonald	7.7
Denys Arcand	7.7
Jorge RamÃrez SuÃ¡rez	7.7
Fernando LeÃ³n de Aranoa	7.7
Orson Welles	7.7
William Cottrell	7.7
Chris Paine	7.7
Alex Gibney	7.7
Lloyd Bacon	7.7
Ricki Stern	7.7
Andrew Haigh	7.7
Mark Osborne	7.7
Steve McQueen	7.7
Martin Scorsese	7.7
Thomas Vinterberg	7.7
Francis Ford Coppola	7.7
Matthew Vaughn	7.7
Ben Affleck	7.7
John G. Avildsen	7.7
Paul Haggis	7.7
Philip Kaufman	7.7
John Carney	7.7
Todd Solondz	7.7
Alejandro AmenÃ¡bar	7.6
Mel Brooks	7.6
Chan-wook Park	7.6
Wes Anderson	7.6
Edgar Wright	7.6
Pierre Coffin	7.6
Ryan Coogler	7.6

Martin McDonagh	7.6
Chris Buck	7.6
Ted Demme	7.6
Tim McCanlies	7.6
Thea Sharrock	7.6
Brian Percival	7.6
Stanley Kramer	7.6
Don Siegel	7.6
Mary Harron	7.6
Tom Ford	7.6
Jerome Robbins	7.6
Michael McGowan	7.6
Eugenio Derbez	7.6
Luc Jacquet	7.6
Isabel Coixet	7.6
Joshua Tickell	7.6
Petter NÃ¦ss	7.6
Chia-Liang Liu	7.6
Gareth Evans	7.6
Anthony Powell	7.6
Barry W. Blaustein	7.6
Kay Pollak	7.6
Pan Nalin	7.6
Kevin Jordan	7.6
Fernando Meirelles	7.6
Paul Greengrass	7.6
Brad Bird	7.6
Spike Jonze	7.6
Duncan Jones	7.6
Steve James	7.6
James Algar	7.6
J.A. Bayona	7.6
Stephen Chow	7.6
Todd Field	7.6
Steven Spielberg	7.5
David Lynch	7.5
Yimou Zhang	7.5
Paul Thomas Anderson	7.5
Lars von Trier	7.5
Bennett Miller	7.5
Sean Penn	7.5
Irvin Kershner	7.5
Tate Taylor	7.5
Katsuhiro Ã"tomo	7.5
Tony Bancroft	7.5
Norman Ferguson	7.5
Steve Box	7.5
Kirsten Sheridan	7.5
Dexter Fletcher	7.5
John Patrick Shanley	7.5

Charlie Kaufman	7.5
Gilles Paquet-Brenner	7.5
Timothy Hines	7.5
John Crowley	7.5
Carol Reed	7.5
Lynne Ramsay	7.5
Joshua Marston	7.5
Khyentse Norbu	7.5
Whit Stillman	7.5
Tom Putnam	7.5
Jafar Panahi	7.5
Darren Aronofsky	7.5
Jean-Marc VallÃ©e	7.5
Sam Mendes	7.5
Stephen Daldry	7.5
Nicholas Meyer	7.5
Marco Kreuzpaintner	7.5
J.J. Abrams	7.5
Denzel Washington	7.5

Top-performing directors belonged to the 90th percentile of average IMDB scores.

Successful directors showed a consistent trend of high-budget movies with unique storytelling.

E. Budget Analysis:

Row Labels	Sum of profit margin
Avatar	523505847
E.T. the Extra-Terrestrial	424449459
Jurassic World	502177271
Star Wars: Episode I - The Phantom Menace	359544677
Star Wars: Episode IV - A New Hope	449935665
The Avengers	403279547
The Dark Knight	348316061
The Hunger Games	329999255
The Lion King	377783777
Titanic	458672302

Movies with higher budgets generally yielded higher profit margins.

Result

This project provided a comprehensive understanding of the factors contributing to a movie's success. Key achievements included:

- Identifying actionable trends in genres, durations, and budgets.

- Highlighting the role of directors and languages in determining IMDB ratings.

- Demonstrating a systematic approach to leveraging descriptive statistics and visualizations for impactful insights.

<u>Drive Link</u>

[Click here to access the full report](#)