

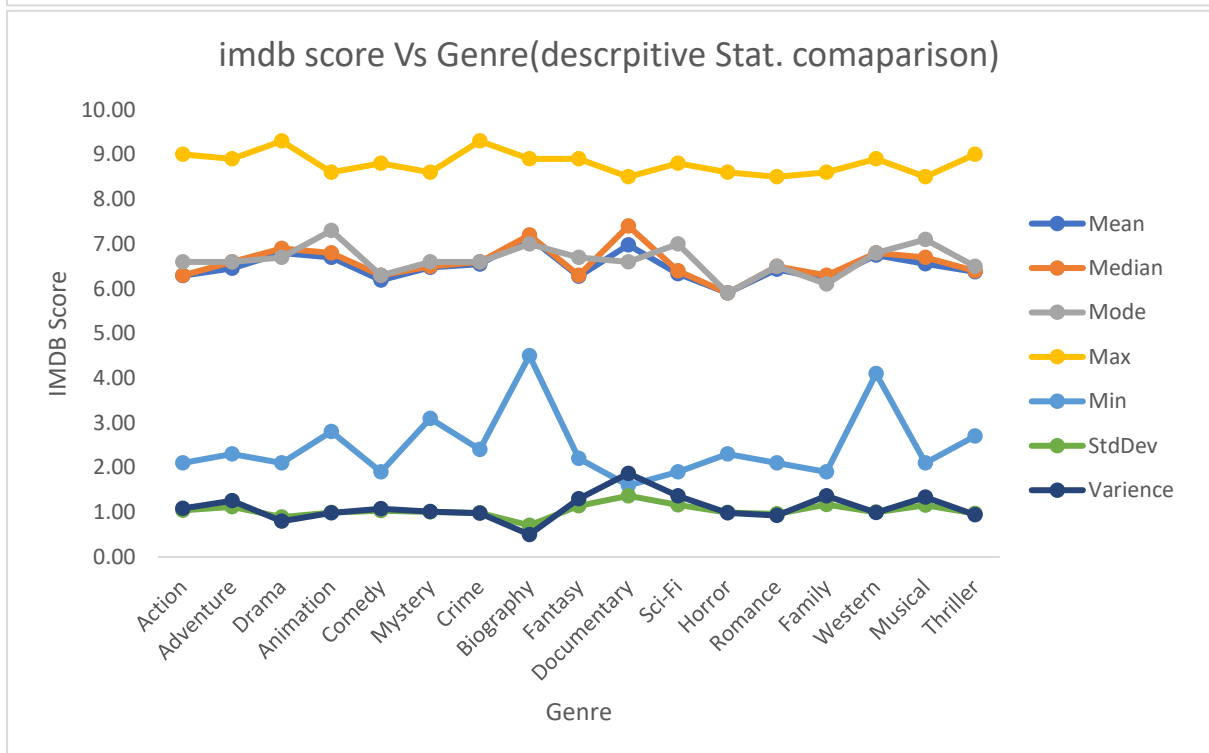
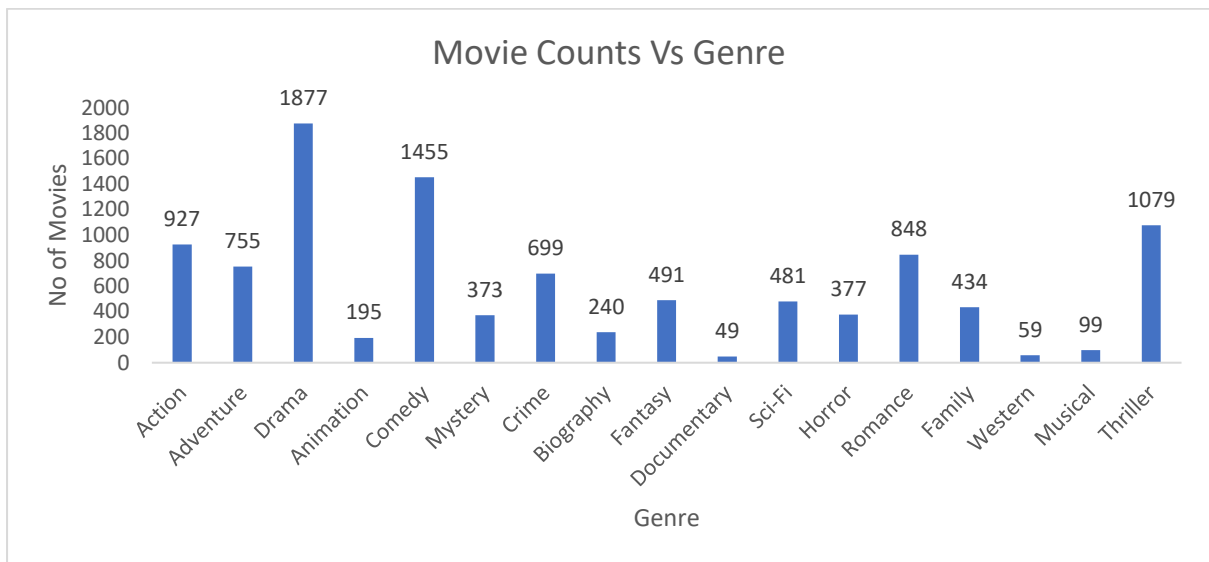
Way of doing analysis:

I trimmed irrelevant columns, such that the dataset now comprises only 8 columns: director's name, duration, movie title, genre, budget, gross, IMDb rating, language, and country. Subsequently, I systematically removed blank rows and duplicates, resulting in a streamlined dataset of 3702 rows for a more focused analysis.

A. Movie Genre Analysis:

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.**

Unique Genre	Movie Counts	Mean	Median	Mode	Max	Min	StdDev	Variance
Action	927	6.29	6.30	6.60	9.00	2.10	1.04	1.08
Adventure	755	6.45	6.60	6.60	8.90	2.30	1.12	1.26
Drama	1877	6.79	6.90	6.70	9.30	2.10	0.89	0.80
Animation	195	6.70	6.80	7.30	8.60	2.80	0.99	0.98
Comedy	1455	6.19	6.30	6.30	8.80	1.90	1.04	1.08
Mystery	373	6.48	6.50	6.60	8.60	3.10	1.01	1.01
Crime	699	6.55	6.60	6.60	9.30	2.40	0.99	0.98
Biography	240	7.15	7.20	7.00	8.90	4.50	0.71	0.50
Fantasy	491	6.28	6.30	6.70	8.90	2.20	1.14	1.30
Documentary	49	6.98	7.40	6.60	8.50	1.60	1.37	1.87
Sci-Fi	481	6.33	6.40	7.00	8.80	1.90	1.17	1.36
Horror	377	5.90	5.90	5.90	8.60	2.30	0.99	0.99
Romance	848	6.43	6.50	6.50	8.50	2.10	0.96	0.93
Family	434	6.20	6.30	6.10	8.60	1.90	1.17	1.37
Western	59	6.75	6.80	6.80	8.90	4.10	1.00	1.00
Musical	99	6.55	6.70	7.10	8.50	2.10	1.16	1.34
Thriller	1079	6.37	6.40	6.50	9.00	2.70	0.97	0.94



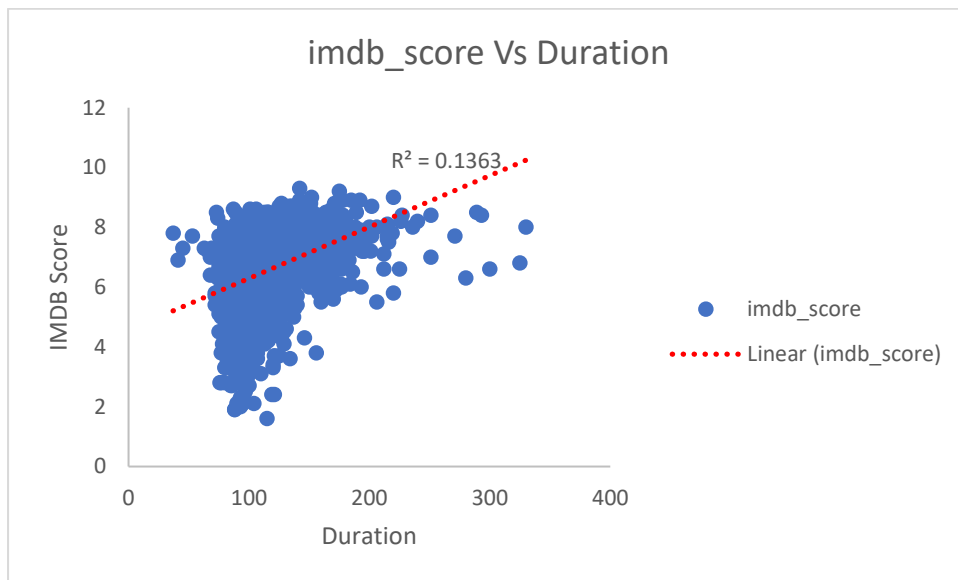
Insights-

Drama, Comedy, Thriller, and Action emerge as the most prevalent movie genres in the dataset.

B. Movie Duration Analysis:

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

Mean	Median	Mode	Max	Min	StdDev	Variance
110.12	106	101	330	37	22.74	516.91



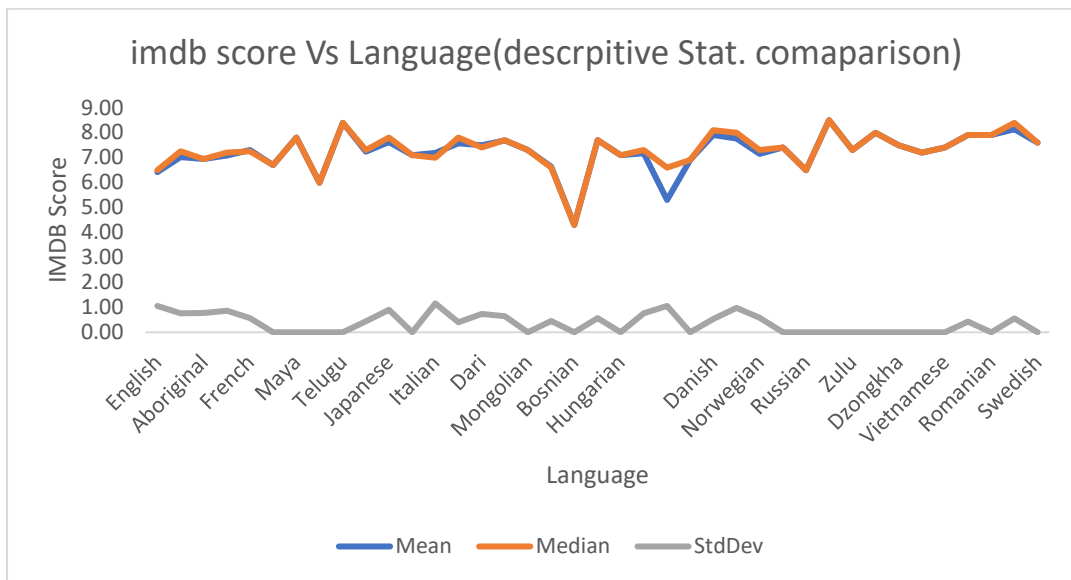
Insights-

The typical movie duration spans 110 minutes, and the correlation between duration and IMDb score exhibits an upward trend, with an R^2 value of 0.1363.

C. Movie Language Analysis:

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

Languages	movies
	Count
English	3527
French	36
Spanish	23
Mandarin	14
German	13
Japanese	12
Hindi	8
Cantonese	8
Italian	7
Korean	5
Portuguese	5



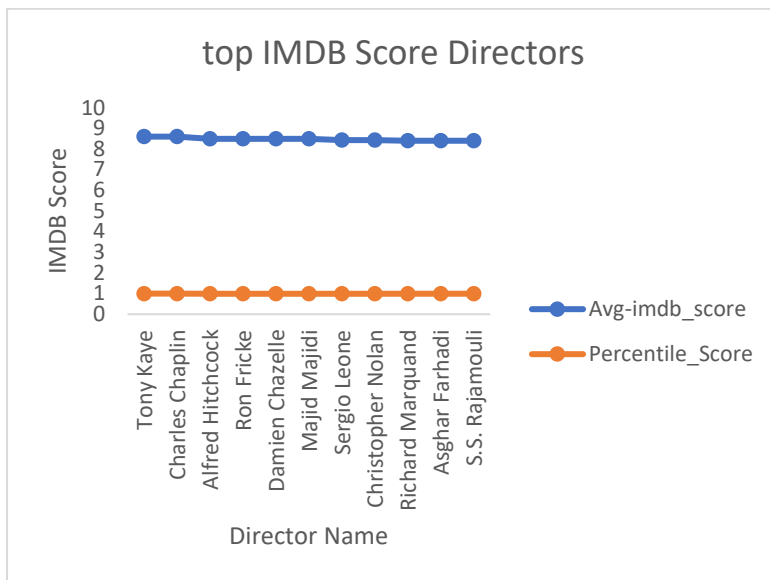
Insights-

English, French, Spanish, Mandarin, and German stand out as the most frequently utilized languages in films.

D. Movie Director Analysis:

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

director	Avg-imdb_score	Percentile_Score
Tony Kaye	8.6	0.998
Charles Chaplin	8.6	0.998
Alfred Hitchcock	8.5	0.996
Ron Fricke	8.5	0.996
Damien Chazelle	8.5	0.996
Majid Majidi	8.5	0.996
Sergio Leone	8.43	1.00
Christopher Nolan	8.425	0.995
Richard Marquand	8.4	0.993
Asghar Farhadi	8.4	0.993
S.S. Rajamouli	8.4	0.993



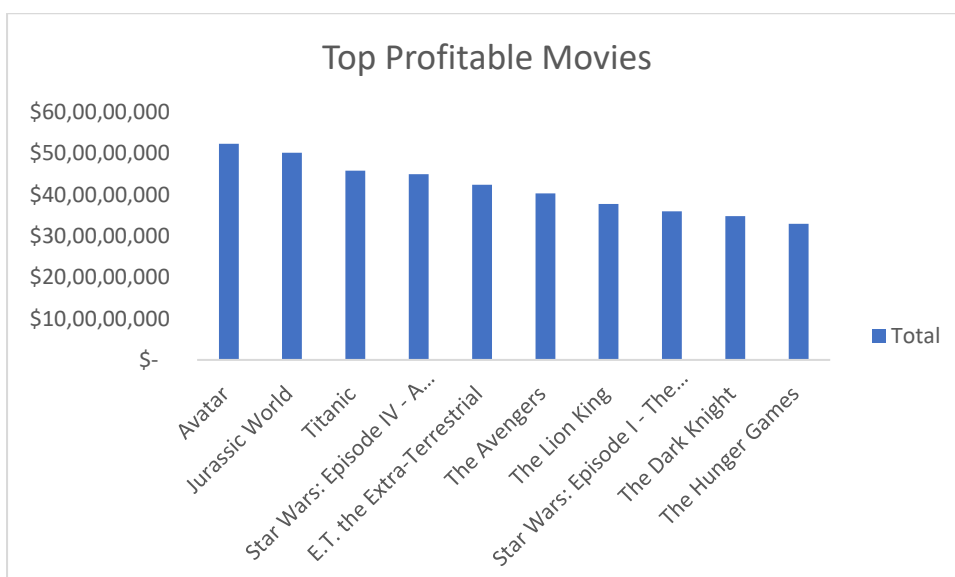
Insights-

Tony Kaye, Charles Chaplin, Alfred Hitchcock, Ron Fricke, Damien Chazelle, Majid Majidi, Sergio Leone, Christopher Nolan, SS Rajamouli, and Richard Marquand have been identified as the top 10 directors, each maintaining an average IMDb score of 8.4 or higher.

E. Movie Budget Analysis:

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

movie	Sum of Profit
Avatar	\$ 52,35,05,847
Jurassic World	\$ 50,21,77,271
Titanic	\$ 45,86,72,302
Star Wars: Episode IV - A New Hope	\$ 44,99,35,665
E.T. the Extra-Terrestrial	\$ 42,44,49,459
The Avengers	\$ 40,32,79,547
The Lion King	\$ 37,77,83,777
Star Wars: Episode I - The Phantom Menace	\$ 35,95,44,677
The Dark Knight	\$ 34,83,16,061
The Hunger Games	\$ 32,99,99,255



Insights-

Avatar, Jurassic World, Titanic, Star Wars: Episode IV - A New Hope, and E.T. The Extra-Terrestrial lead the charts as the top five most profitable movies. Additionally, there is a positive correlation between the budget and gross earnings.

Link for excel file-

<https://docs.google.com/spreadsheets/d/1a5J1euyGkxcZju4Uyi3zVaQyAqM2FYPw/edit?usp=sharing&ouid=101202744784682762684&rtpof=true&sd=true>