



IMDB Movie Analysis

Sanjana Kumari Yadav



Agenda

- Movie Genre Analysis
- Movies Duration Analysis
- Language Analysis
- Director Analysis
- Budget Analysis

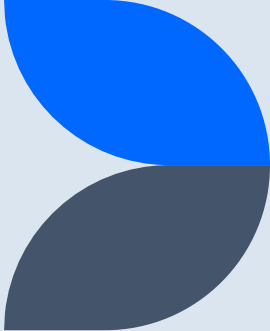
Introduction

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.

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.



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

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.

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.

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.

Hence the most
Common genres of
Movies is **Drama**.

Average	Median	Max	Min	Range	Mode	Var	Std
520.9048	392	1893	1	1892	59	258307.9	508.24

Name	count
Action	959
Adventure	780
Fantasy	505
Sci-Fi	495
Thriller	1115
Romance	859
Musical	96
Western	59
Family	441
Mystery	383
History	149
Comedy	1460
Western	59
Drama	1893
Biography	239
War	152
Document	45
Sport	148
Film-Noir	1
Crime	709
Horror	392

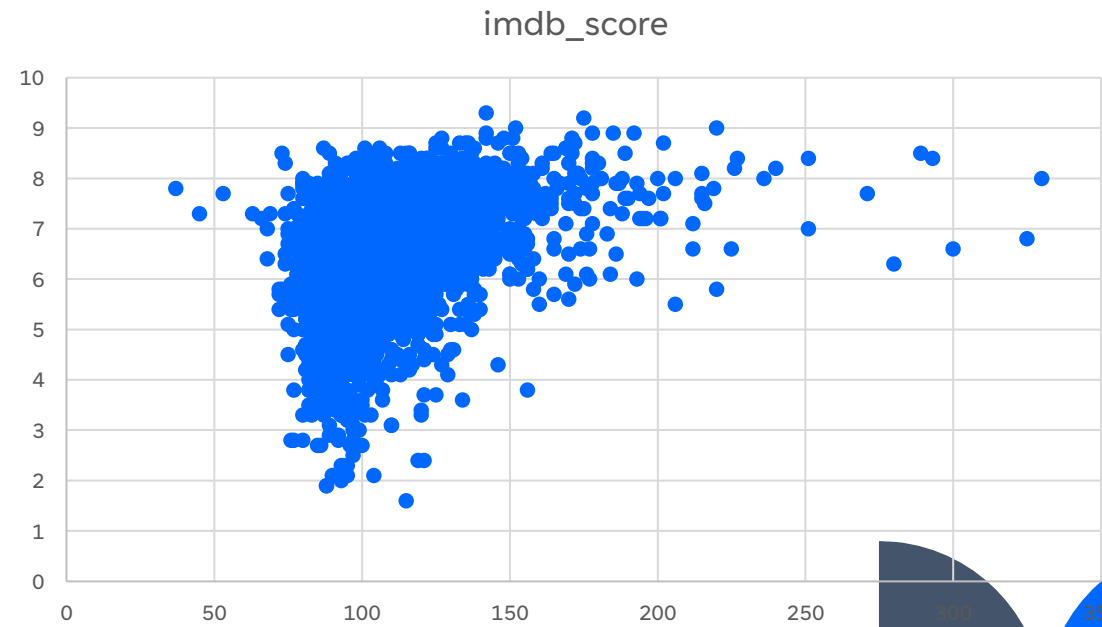
Movie Duration Analysis

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

Average :110.258

Median : 106

Standard Deviation:
26.6467



Language Analysis

- Task:** Determine the most common languages used in movies and analyze their impact on the IMDB score using descriptive statistics

The most common language is English.

Median : 15.4

Row Labels	Count of language	Average of imdb_score	StdDevp of imdb_score	Sum of imdb_score
Aboriginal	2	6.95	0.55	13.9
Arabic	1	7.2	0	7.2
Aramaic	1	7.1	0	7.1
Bosnian	1	4.3	0	4.3
Cantonese	7	7.342857143	0.324509048	51.4
Czech	1	7.4	0	7.4
Danish	3	7.9	0.43204938	23.7
Dari	2	7.5	0.1	15
Dutch	3	7.566666667	0.329983165	22.7
English	3598	6.427042802	1.050392173	23124.5
Filipino	1	6.7	0	6.7
French	34	7.355882353	0.51173935	250.1
German	10	7.77	0.67535176	77.7
Hebrew	1	8	0	8
Hindi	5	7.22	0.716658915	36.1
Hungarian	1	7.1	0	7.1
Indonesian	2	7.9	0.3	15.8
Italian	7	7.185714286	1.069617517	50.3
Japanese	10	7.66	0.939361485	76.6
Kazakh	1	6	0	6
Korean	5	7.7	0.509901951	38.5
Mandarin	15	7.08	0.745832868	106.2
Maya	1	7.8	0	7.8

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.

Percentile : 8.61

Name	score
Akira Kurosawa	8.7
Cedric Nicolas-Troyat	8.6
Alison Maclean	8.5
Damian Nieman	8.5
Maggie Greenwald	7.3
Audrey Wells	8.4
Ron Clements	7.4
Rick Bieber	5.7
Tomas Alfredson	6.2
Seth MacFarlane	6.5

Budget Analysis

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

The Avengers

Highest profit margin:

806559094

Correlation : 0.124941666

For correlation we use:

=CORREL(B4:B3659,C4:C365)

	A	B	C	D
1	Title	Sum of budget	Sum of gross	profit margin
2	The Avengers	440000000	1246559094	806559094
3	Avatar	237000000	760505847	523505847
4	Jurassic World	150000000	652177271	502177271
5	Titanic	200000000	658672302	458672302
6	Star Wars: Episode IV - A New Hope	11000000	460935665	449935665
7	E.T. the Extra-Terrestrial	10500000	434949459	424449459
8	The Lion King	45000000	422783777	377783777
9	The Jungle Book	35000000	725290282	375290282
10	Star Wars: Episode I - The Phantom Menace	115000000	474544677	359544677
11	The Dark Knight	185000000	533316061	348316061
12	The Twilight Saga: Breaking Dawn - Part 2	240000000	584597846	344597846
13	The Hunger Games	78000000	407999255	329999255
14	The Fast and the Furious	114000000	433536930	319536930
15	Twilight	74000000	382898950	308898950
16	Deadpool	58000000	363024263	305024263
17	The Hunger Games: Catching Fire	130000000	424645577	294645577
18	Jurassic Park	63000000	356784000	293784000
19	Despicable Me 2	76000000	368049635	292049635
20	American Sniper	58800000	350123553	291323553
21	Finding Nemo	94000000	380838870	286838870
22	Shrek 2	150000000	436471036	286471036
23	The Lord of the Rings: The Return of the King	94000000	377019252	283019252
24	Star Wars: Episode VI - Return of the Jedi	32500000	309125409	276625409
25	Forrest Gump	55000000	329691196	274691196
26	Star Wars: Episode V - The Empire Strikes Back	18000000	290158751	272158751
27	Jurassic	15000000	386085688	271085688



Thank you

Here is the link of the sheet:

https://drive.google.com/file/d/1bxS2qn2StMzCiwsRNO4XsIIKyP1K-o9g/view?usp=drive_link