

IMDB Movie Analysis

Name- Mukesh Chandra Kamila

Description:

This project is all about carry out the in-depth analysis of important underlying insights of IMDB dataset. Being a Data Analyst, my job was to provide actionable insights that can help stakeholders make informed decisions.

Approach:

First, I have done data cleaning, removed blank cell / null values, removed duplicates and dropped unnecessary columns. Then I have done the EDA part using excel with the help of data analysis tool pack. And this analysis was done with the help of various functions, formula and tools.

Tech-Stack: Microsoft Excel 2019

Used:

Microsoft Excel 2019 was used in this project execution. The ease of access and set up with convenient user interface made it a good tool for the project.

Insights:

In this project, I learned about advanced EXCEL and statistics. And how to analyse the problem statement and the functions that I can use in EXCEL to solve the problem statement. All the questions asked has been answered or solved through Excel and statistics.

Result:

In this project, I have achieved and gained knowledge how to deal and analyse the data with help of EXCEL and apply statistics logic and function. And how to interact and run different functions to get desired output from data.

Analysis

Data Cleaning:

1. Dropping unnecessary columns:

(Color, director_facebook_likes, actor_3_facebook_likes, actor_2_name, actor_1_facebook_likes, cast_total_facebook_likes, actor_3_name, facenumber_in_posts, plot_keywords, movie_imdb_link, content_rating, actor_2_facebook_likes, aspect_ratio, movie_facebook_likes)

2. Removed blank cell / null values.

3. Removed duplicates.

Excel Workbook-

https://docs.google.com/spreadsheets/d/17U2RO_ID1NTXT44bxe1Vs6V9BAHBNGI1/edit?usp=sharing&oid=110645243858390813184&rtpof=true&sd=true

A. Movie Genre Analysis

By using CountIF and AverageIF formula, got this output.

Genre	Movies count	Avg. IMDB
Horror	391	5.93
Western	58	6.77
Thriller	1117	6.38
Drama	1939	6.79
Sci-Fi	497	6.32
Romance	877	6.43
Documentary	64	6.99
Action	960	6.29
Adventure	786	6.46
Comedy	1503	6.18

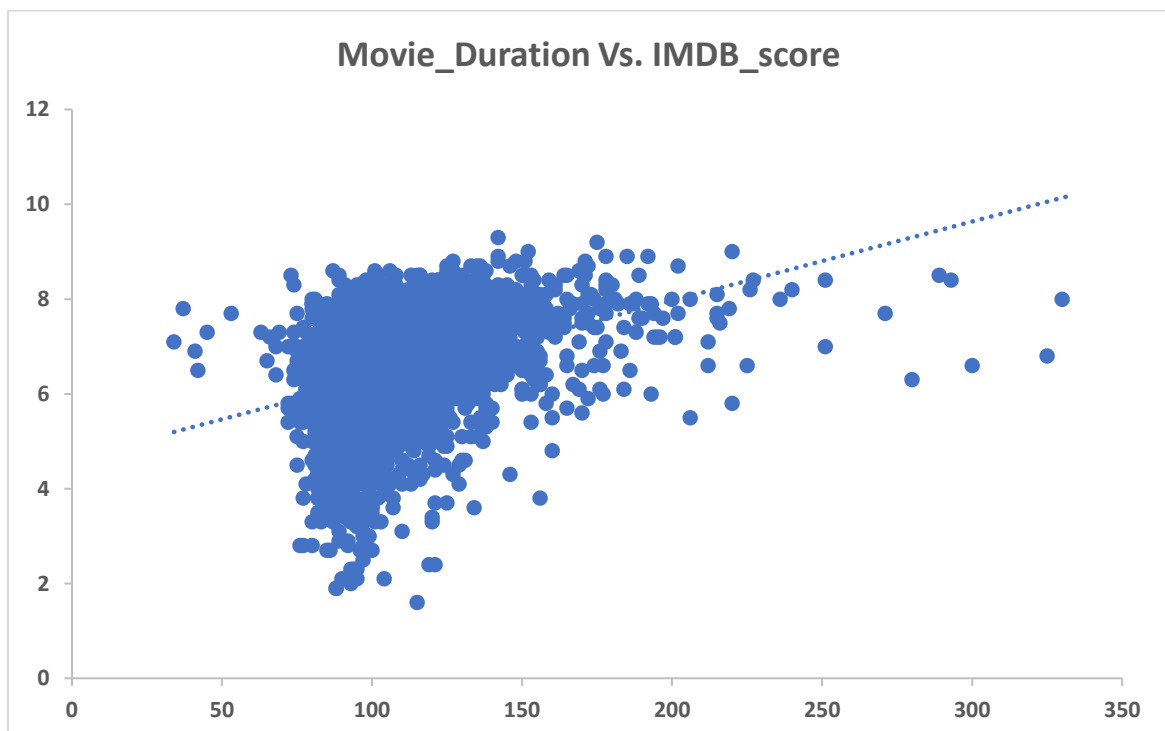
B. Movie Duration Analysis

Output with help of formula:

Avg. Duration	Median Duration	Stdev_Duration
111.8719915	106	28.72505149

Output with the help of Data Analysis tool:

<i>Descriptive of duration</i>	
Mean	111.8719915
Standard Error	0.467151209
Median	106
Mode	100
Standard Deviation	28.72505149
Sample Variance	825.128583
Kurtosis	32.06245676
Skewness	3.711479383
Range	569
Minimum	34
Maximum	603
Sum	422988
Count	3781



C. Language Analysis

Language Analysis Using Pivot, CountIF, SumIF and Descriptive Statistics.

Languages	Movie's Counts	Sum_IMDB_score	Mean_IMDB score	Median_IMDB score	Standard Dev._IMDB score
English	3668	23562.9	6.42	6.6	1.05
French	37	269.6	7.29	6.8	0.67
Spanish	26	183.3	7.05	6.7	0.79
Mandarin	14	98.3	7.02	6.8	0.70
German	13	100	7.69	6.9	0.61
Japanese	12	91.5	7.63	6.9	0.63
Hindi	10	67.6	6.76	6.6	0.85

Overall IMDB score descriptive output:

<i>Descriptive of imdb_score</i>	
Mean	7.259886312
Standard Error	0.117135594
Median	7.3
Mode	7.5
Standard Deviation	0.722072298
Sample Variance	0.521388404
Kurtosis	6.780018592
Skewness	-1.830371558
Range	4.2
Minimum	4.3
Maximum	8.5
Sum	275.8756799
Count	38

D. Director Analysis

Top directors based on their average IMDB score:

Directors Name	Sum of Profit	Average of imdb_score
Tony Kaye	-787759.00	8.60
Charles Chaplin	-1336755.00	8.60
Alfred Hitchcock	31193053.00	8.50
Ron Fricke	-1398153.00	8.50
Damien Chazelle	9792000.00	8.50
Majid Majidi	745402.00	8.50
Sergio Leone	-16500000.00	8.43
Christopher Nolan	808227576.00	8.43
S.S. Rajamouli	-11528148.00	8.40
Richard Marquand	276625409.00	8.40
Marius A. Markevicius	-366222.00	8.40
Asghar Farhadi	6598492.00	8.40

E. Budget Analysis

Top 10 movies with the highest profit margin:

director_name	genres	actor_1_name	movie_title	language	country	title_year	imdb_score	Profit
James Cameron	Action Adventure Fantasy	CCH Pounder	Avatar	English	USA	2009	7.9	523505847
Joss Whedon	Action Adventure Sci-Fi	Chris Hemsworth	The Avengers	English	USA	2012	8.1	403279547
James Cameron	Drama Romance	Leonardo DiCaprio	Titanic	English	USA	1997	7.7	458672302
Colin Trevorrow	Action Adventure Sci-Fi Thriller	Bryce Dallas Howard	Jurassic World	English	USA	2015	7	502177271
Christopher Nolan	Action Crime Drama Thriller	Christian Bale	The Dark Knight	English	USA	2008	9	348316061
George Lucas	Action Adventure Fantasy	Natalie Portman	Star Wars: Episode I - The Phantom Menace	English	USA	1999	6.5	359544677
Gary Ross	Adventure Drama Sci-Fi Thriller	Jennifer Lawrence	The Hunger Games	English	USA	2012	7.3	329999255
Roger Allers	Adventure Animation Drama	Matthew Broderick	The Lion King	English	USA	1994	8.5	377783777
George Lucas	Action Adventure Fantasy	Harrison Ford	Star Wars: Episode IV - A New Hope	English	USA	1977	8.7	449935665
Steven Spielberg	Family Sci-Fi	Henry Thomas	E.T. the Extra-Terrestrial	English	USA	1982	7.9	424449459

The movies with the highest profit margin:

Movie With Highest Profit		
523505847		
Director name	Movies name	Profit
James Cameron	Avatar	523505847

The correlation between movie budgets and gross earnings:

With Formula-

Correlation Between Movie Budgets and Gross Earnings
0.122979125

With Data Analysis tool-

	Sum of budget	Sum of gross
Sum of budget	1	
Sum of gross	0.122979125	1