IMDB MOVIE ANALYSIS

PROJECT 05 TRAINITY   EXCEL FILE

PROJECT DESCRIPTION: In this project, I have used Excel as my tool to analyse the imdb movie data given in the project. My goal in this project is to find details insight from the imdb data .I found 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: I have looked into the data given in the project. Then I have downloaded the IMDB data given in the project. Then cleaned the data properly ,filled the missing value after that I have performed operation.
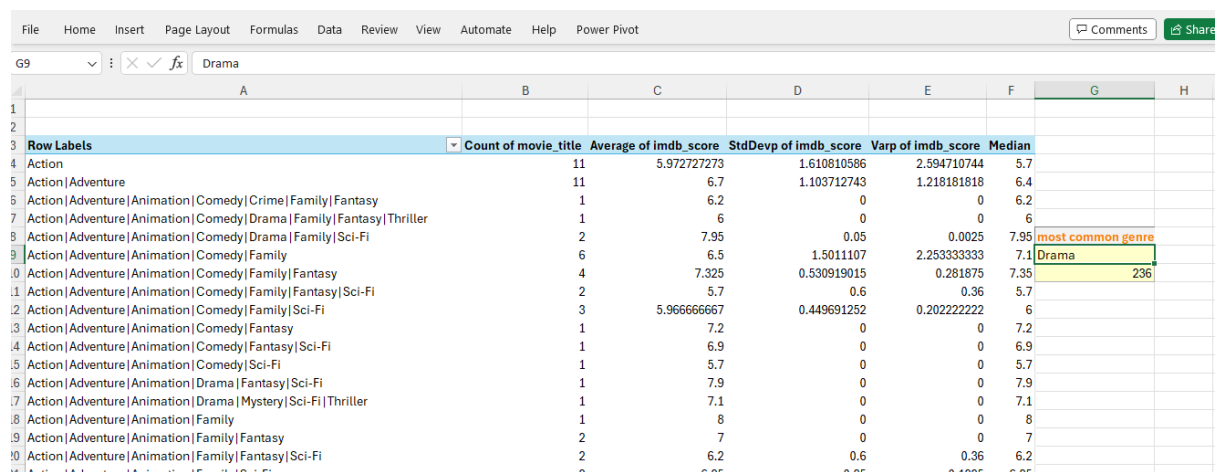
DATA CLEANING: At first I observe in the data. Then after looking the task I dropped the unnecessary columns. Then in the columns. If there is numbers I filled the blank cells with mean value. If there is string I filled the data with most used string used in the entire column .

TECH-STACK USED: Excel

PROJECT DETAILS AND DATA ANALYSIS

Movie Genre Analysis:

Here my task is to 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.. Here is the snapshot of the result on Excel.
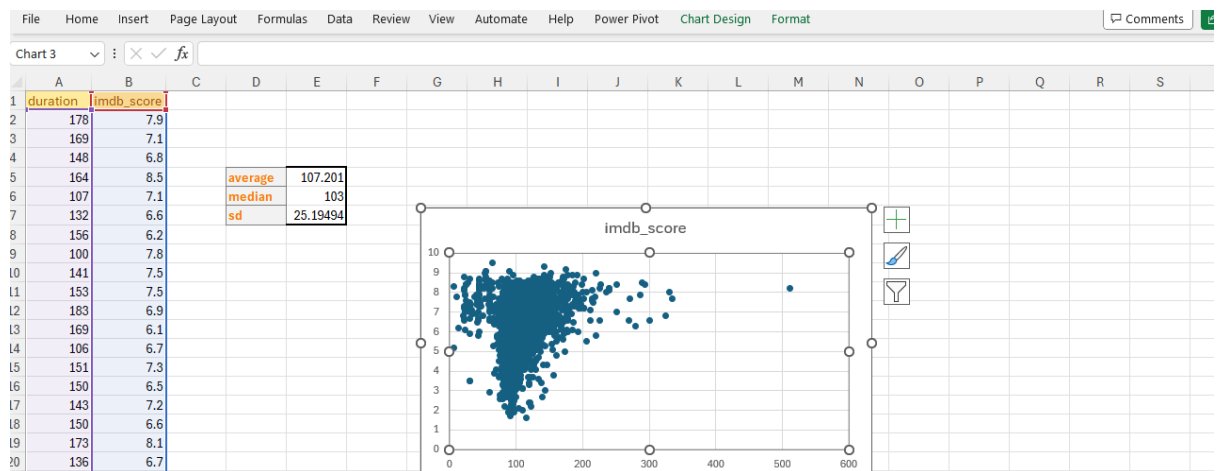

Row Labels	Count of movie_title	Average of imdb_score	StdDevp of imdb_score	Varp of imdb_score	Median		
Action	11	5.972727273	1.610810586	2.594710744	5.7		
Action|Adventure	11	6.7	1.103712743	1.218181818	6.4		
Action|Adventure|Animation|Comedy|Crime|Family|Fantasy	1	6.2	0	0	6.2		
Action|Adventure|Animation|Comedy|Drama|Family|Fantasy|Thriller	1	6	0	0	6		
Action|Adventure|Animation|Comedy|Drama|Family|Sci-Fi	2	7.95	0.05	0.0025	7.95	most common genre	
Action|Adventure|Animation|Comedy|Family	6	6.5	1.5011107	2.253333333	7.1	Drama	
Action|Adventure|Animation|Comedy|Family|Fantasy	4	7.325	0.530919015	0.281875	7.35		236
Action|Adventure|Animation|Comedy|Family|Fantasy|Sci-Fi	2	5.7	0.6	0.36	5.7		
Action|Adventure|Animation|Comedy|Family|Sci-Fi	3	5.966666667	0.449691252	0.202222222	6		
Action|Adventure|Animation|Comedy|Fantasy	1	7.2	0	0	7.2		
Action|Adventure|Animation|Comedy|Fantasy|Sci-Fi	1	6.9	0	0	6.9		
Action|Adventure|Animation|Comedy|Sci-Fi	1	5.7	0	0	5.7		
Action|Adventure|Animation|Drama|Fantasy|Sci-Fi	1	7.9	0	0	7.9		
Action|Adventure|Animation|Drama|Mystery|Sci-Fi|Thriller	1	7.1	0	0	7.1		
Action|Adventure|Animation|Family	1	8	0	0	8		
Action|Adventure|Animation|Family|Fantasy	2	7	0	0	7		
Action|Adventure|Animation|Family|Fantasy|Sci-Fi	2	6.2	0.6	0.36	6.2		
Action|Adventure|Animation|Family|Sci-Fi	2	6.25	0.35	0.1225	6.25		

Report: The most common genre is drama which is repeated 236 times in the data. Here I made a simple pivotable from original table the just calculated mean of imdb_score ,counted movies, stdDev of imdb_score,varp of imdb etc.With the help of power pivot I calculated the median also. With help of the data one can see which types of movie get more rating.

Movie Duration Analysis:

Here is my task to analyse the distribution of movie durations and identify the relationship between movie duration and IMDB score.. Here is the snapshot of the operation on Excel.

Report : Here we can clearly see the relationship between movie duration and imdb_score by a scatter plot. I have also calculated average duration with top 10 imdb .That is 128 minutes .Movies with These length get highest rating.

Language Analysis:

Here I have to determine the most common languages used in movies and analyse their impact on the IMDB score using descriptive statistics.

Report : Most common language is English which is used 4704 times. I made a pivot table with language and counting movie title and imdb_score. I found that It is all about quality the impact the rating not the language .Tamil has highest average rating it has only one movie in the data. So we can not judge based on that.

Director Analysis:

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

Here is the snapshot of to ten director with highest imdb ratings.

	A	B	C	D	E
1					
2	top 10 director				
3	director_name	Average of imdb_score			
4	Cary Bell	8.7			
5	Charles Chaplin	8.6			
6	Damien Chazelle	8.5			
7	John Blanchard	9.5			
8	Majid Majidi	8.5			
9	Mike Mayhall	8.6			
10	Mitchell Altieri	8.7			
11	Raja Menon	8.5			
12	Ron Fricke	8.5			
13	Sadyk Sher-Niyaz	8.7			
14	Grand Total	8.68			
15					
16					
17	Large	9.5			
18	percentrank	1	100%		
19	pecentile	9.5			
20					

Report : Carry bell has been most successful director based on data. As per the list impact of a director for success of a movie is 100%.

Budget Analysis:

Here I have to analyse the correlation between movie budgets and gross earnings, and identify the movies with the highest profit margin.

	A	B	C	D	E	F	G	H	I	J
1	movie_title	budget	gross	profit		CORRELATION				
2	AvatarÂ	237000000	760505847	523505847		0.101437273				
3	Pirates of the Caribbean: At World's EndÂ	300000000	309404152	9404152						
4	SpectreÂ	245000000	200074175	-44925825		HIGH PROFIT MARGIN				
5	The Dark Knight RisesÂ	250000000	448130642	198130642		523505847				
6	Star Wars: Episode VII - The Force AwakensÂ	39752620	48468407	8715787		AvatarÂ				
7	John CarterÂ	263700000	73058679	-1.91E+08						
8	Spider-Man 3Â	258000000	336530303	78530303						
9	TangledÂ	260000000	200807262	-59192738						
10	Avengers: Age of UltronÂ	250000000	458991599	208991599						
11	Harry Potter and the Half-Blood PrinceÂ	250000000	301956980	51956980						
12	Batman v Superman: Dawn of JusticeÂ	250000000	330249062	80249062						
13	Superman ReturnsÂ	209000000	200069408	-8930592						
14	Quantum of SolaceÂ	200000000	168368427	-31631573						
15	Pirates of the Caribbean: Dead Man's ChestÂ	225000000	423032628	198032628						
16	The Lone RangerÂ	215000000	89289910	-1.26E+08						
17	Man of SteelÂ	225000000	291021565	66021565						
18	The Chronicles of Narnia: Prince CaspianÂ	225000000	141614023	-83385977						
19	The AvengersÂ	220000000	623279547	403279547						
20	Pirates of the Caribbean: On Stranger TidesÂ	250000000	241063875	-8936125						

Report: Movie with highest profit is AvatarA .

Correlation between movie budget and profit is 0.101437273 which is significantly low. Profit depends on budget but not that much . Some time low budget movie earns a lot.

Result: After all the operation on excel I got that .Drama movies are highest in number, Movie with 2 hours and some minutes gets highest rating, a good director is must for a movie success, good budget is necessary but not always good budget movie do good on market.