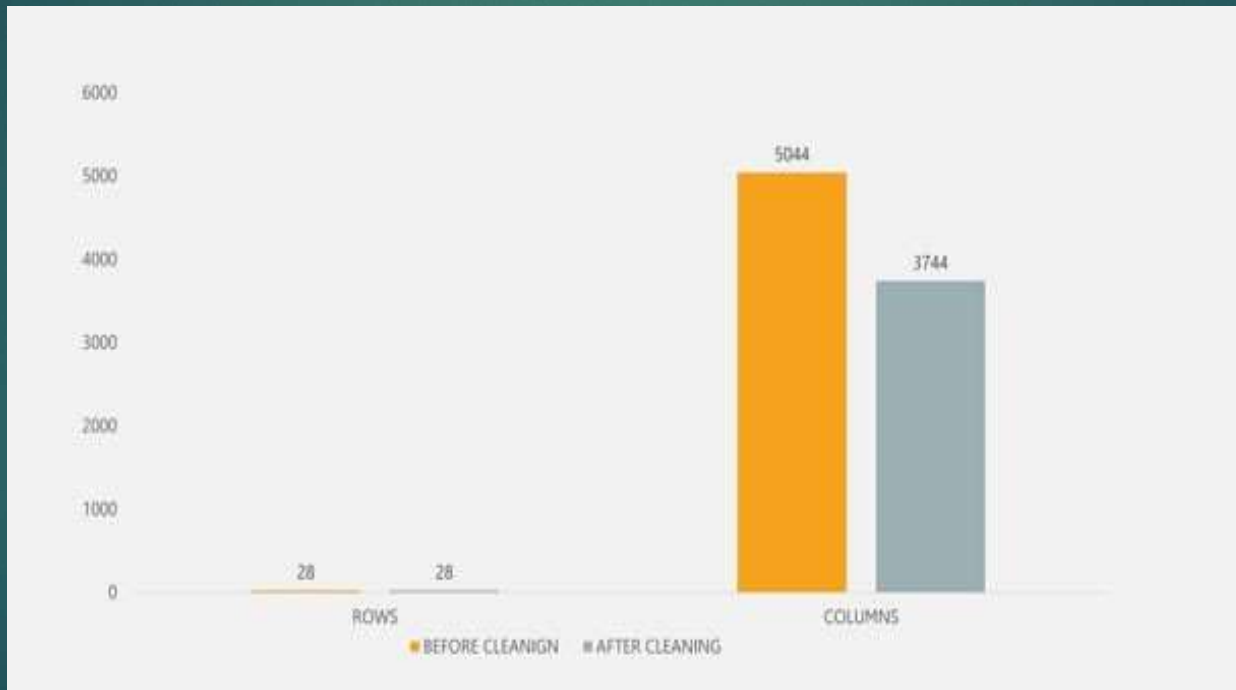




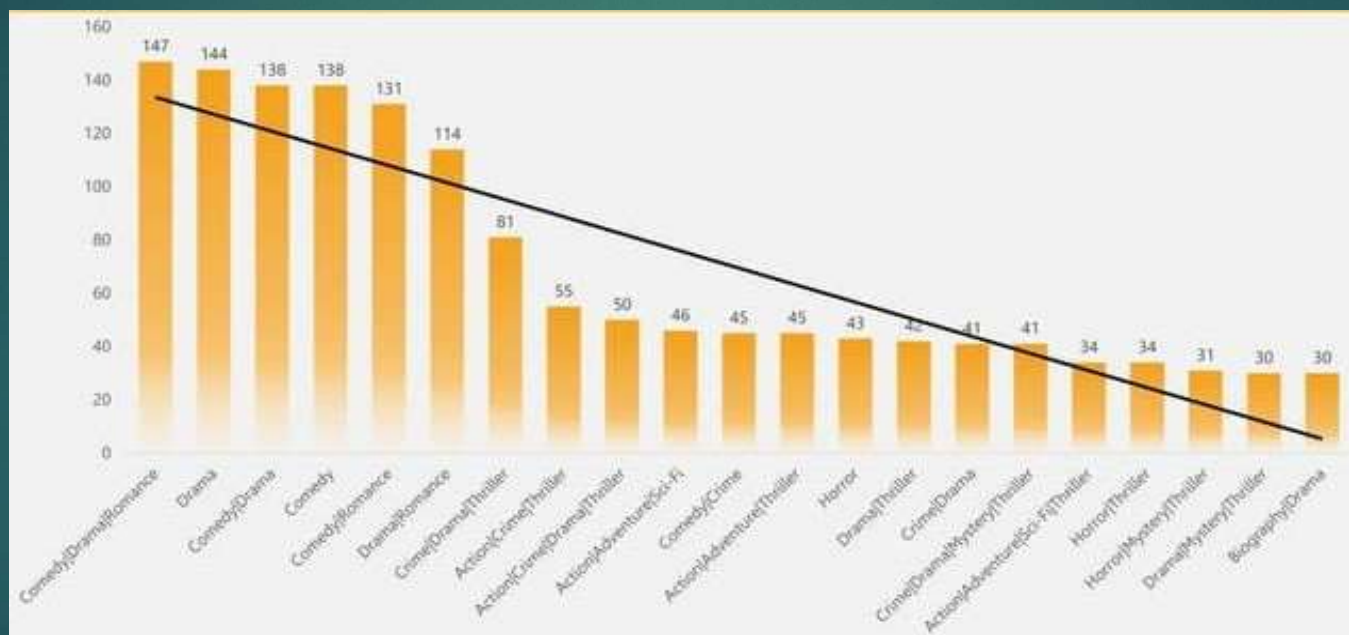
# IMDB MOVIE ANALYSIS

BY MD ANJAR

- ▶ **Data Cleaning:-** This step involves preprocessing the data to make it suitable for analysis. It includes handling missing values, removing duplicates, converting data types if necessary, and possibly feature engineering.



- **Movie Genre Analysis:-**Analyze the distribution of movie genres and their impact on the IMDB score.



- **Movie Duration Analysis:-**Analyze the distribution of movie durations and its impact on the IMDB score.

Movie Title	IMDb Top 250	Num_Voted_Users > 25,000
The Shawshank Redemption	9.3	1689764
The Godfather	9.2	1155770
The Dark Knight	9	1676169
The Godfather: Part II	9	790926
The Lord of the Rings: The Return of the King	8.9	1215718
Pulp Fiction	8.9	1324680
Schindler's List	8.9	865020
The Good, the Bad and the Ugly	8.9	503509
Forrest Gump	8.8	1251222
Star Wars: Episode V - The Empire Strikes Back	8.8	837759
The Lord of the Rings: The Fellowship of the Ring	8.8	1238746
Inception	8.8	1468200
Fight Club	8.8	1347461
Star Wars: Episode IV - A New Hope	8.7	911097
The Lord of the Rings: The Two Towers	8.7	1100446
The Matrix	8.7	1217752
One Flew Over the Cuckoo's Nest	8.7	680041
Goodfellas	8.7	728685
City of God	8.7	533200
Seven Samurai	8.7	229012

- **Language Analysis:-** Situation: Examine the distribution of movies based on their language.

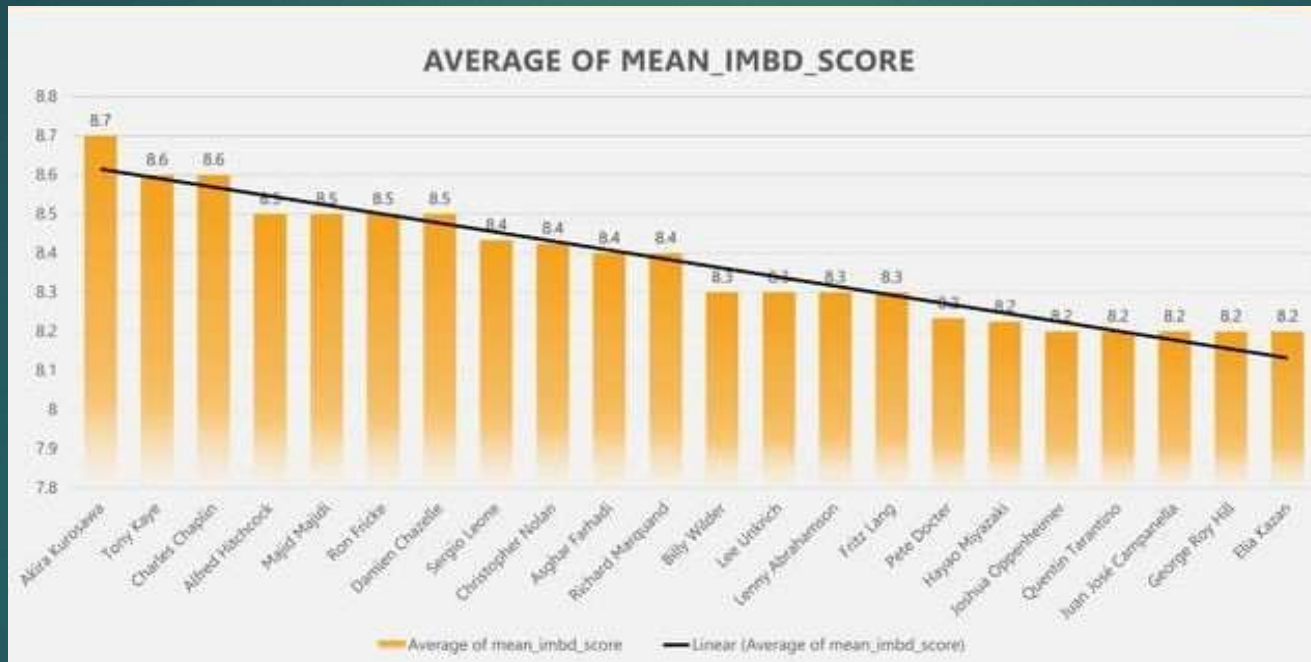
### Top 250 Movies:

director	num_c	durato	gross	genres	actor_1	movie_title	num_v	num_u	language	country	budget	title_year	imdb_s	Profit	Rank
Frank Darab	199	142	28341469	Crime	Dra Morgan Fr	The Shawshank Redemptbn	1689764	4144	English	USA	25000000	1994	9.3	3341469	1
Francis Fo	208	175	134821952	Crime	Dra Al Pacino	The Godfather	1155770	2238	English	USA	6000000	1972	9.2	128821952	2
Christophe	645	152	533316061	Action	Cri Christian B	The Dark Knight	1676169	4667	English	USA	185000000	2008	9	348316061	3
Francis Fo	149	220	57300000	Crime	Dra Robert De	The Godfather: Part II	790926	650	English	USA	13000000	1974	9	44300000	4
Peter Jack	328	192	377019252	Action	Adv Orlando B	The Lord of the Rings: The Return of the King	1215718	3189	English	USA	94000000	2003	8.9	283019252	5
Quentn Tai	215	178	107930000	Crime	Dra Bruce Will	Pulp Fictbn	1324680	2195	English	USA	8000000	1994	8.9	99930000	6
Steven Spi	174	185	96067179	Biography	Liam Nees	Schindler's List	865020	1273	English	USA	22000000	1993	8.9	74067179	7
Sergio Leo	181	142	6100000	Western	Clint Eastw	The Good, the Bad and the Ugly	503509	780	Italian	Italy	1200000	1966	8.9	4900000	8
Robert Ztr	149	142	329691196	Comedy	C Tom Hank	Forrest Gump	1251222	1398	English	USA	55000000	1994	8.8	274691196	9

### Top 250 Foreign Language Movies:

director_n	num_crit	duratbn	gross	genres	actor_1_n	movie_ttk	num_vote	num_user	language	country	budget	title_year	imdb_sco	Profit	Rank
Sergio Leo	181	142	6100000	Western	Clint Eastw	The Good,	503509	780	Italian	Italy	1200000	1966	8.9	4900000	8
Fernando l	214	135	7563397	Crime	Dra Alice Brag	City of Go	533200	748	Portugues	Brazil	3300000	2002	8.7	4263397	19
Akira Kuro	153	202	269061	Action	Adv Takashi Sh	Seven Sam	229012	596	Japanese	Japan	2000000	1954	8.7	-1730939	20
Hayao Miy	246	125	10049886	Adventure	Bunta Sug	Spirited Av	417971	902	Japanese	Japan	19000000	2001	8.6	-8950114	28
Florian He	215	137	11284657	Drama	Th Sebastin k	The Lives c	259379	407	German	Germany	2000000	2006	8.5	9284657	44
Majid Maji	46	89	925402	Drama	Fai Bahare	Sex Children o	27882	130	Persian	Iran	180000	1997	8.5	745402	45
Asghar Far	354	123	7098492	Drama	Mj Shahab Ho	A Separatt	151812	264	Persian	Iran	500000	2011	8.4	6598492	51
Chan-woo	305	120	2181290	Drama	Mj Min-sik Ch	Oldboy	356181	809	Korean	South Kore	3000000	2003	8.4	-818710	54

► **Director Analysis:-**Influence of directors on movie ratings.





- **Budget Analysis:-** Explore the relationship between movie budgets and their financial success.



## PROJECT DESCRIPTION

- ▶ Dataset having 28 columns, 5044 Rows of different IMDB Movies.
- ▶ Define a problem, clean the data as necessary remove null and duplicates values.
- ▶ Explore the data set and derive insights.
- ▶ use root cause analysis five 'whys' approach.
- ▶ Create charts pivot table and graphs to answer given questions .
- ▶ Give answers to asked questions and create reports for data – driven decision.



## Approach

- ▶ Download data and open in excel.
- ▶ Find null and duplicate values and clean all data.
- ▶ Data processing and solved as per asked problem.
- ▶ Use filter, pivot table, sum if, average if, count if and other functions to give answer of asked questions
- ▶ Create charts to for easy and meaningful data representations
- ▶ Create reports in the ppt format and submit the project.

## Tech-Stack Used



# Result

- ▶ Through the IMBD movie Analysis projects, I have done my logical , statistical , and technical skills to extract meaningful insight from the dataset , concepts such as averages , frequency tables, and identifying outliers have enabled me to establish a deeper connection with the data and conduct through analysis.
- ▶ By applying statistical techniques , and leveraging the technical capabilities of Microsoft excels , I have been able to streamline data analytics tasks and simplify complex calculations , additionally , the data visualizations functionalities of excel have allow me to effectively communicate insights by presenting data in a visually appealing manners, I have also gained understanding of how to choose the appropriate visualizations technique based on the nature of the data and the desired outcome.