

IMDB MOVIE ANALYSIS



PROJECT DESCRIPTION

- The project is about IMDB movies analysis dataset which contains the data regarding the movies information like actor's name , movie title and budget.. Etc.
- As an analyst extracting useful and meaningful insights from the data is important.
- So that we can get insights like profit of the movie .. etc.

APPROACH

- My approach for this project is to understand the dataset and thinking of methods on how to get particular answer for that particular question.
- Checking for the error's in the dataset like missing of the null values.

TECH-STACK USED:

- To Analyze and extract meaning full insights from the data I have used
Microsoft Excel 2022





INSIGHTS

- By analyzing the data of IMDB movies I have gained various insights from the data like top director's , top 250 movies.. Etc.
- Using the pivot tables for solving various problems.
- Individual rating's of the particular star's (or) actor and presenting them in the visualization form.

CLEANING THE DATA

director_name	num_critic_for_rev	gross	genres	actor_1_name	movie_title	num_voted_us	language
Frank Darabont	199	2.8E+07	Crime Drama	Morgan Freeman	The Shawshank Redemption	1689784	English
Francis Ford Coppola	208	1.3E+08	Crime Drama	Al Pacino	The Godfather	1155770	English
Christopher Nolan	645	5.3E+08	Action Crime Drama Thriller	Christian Bale	The Dark Knight	1676169	English
Francis Ford Coppola	149	5.7E+07	Crime Drama	Robert De Niro	The Godfather: Part II	790926	English
Peter Jackson	328	3.8E+08	Action Adventure Drama Fantasy	Orlando Bloom	The Lord of the Rings: The Return of the King	1215718	English
Quentin Tarantino	215	1.1E+08	Crime Drama	Bruce Willis	Pulp Fiction	1324680	English
Steven Spielberg	174	9.6E+07	Biography Drama History	Liam Neeson	Schindler's List	865020	English
Sergio Leone	181	6100000	Western	Clint Eastwood	The Good, the Bad and the Ugly	503509	Italian
Robert Zemeckis	149	3.3E+08	Comedy Drama	Tom Hanks	Forrest Gump	1251222	English
Irvin Kershner	223	2.9E+08	Action Adventure Fantasy Sci-Fi	Harrison Ford	Star Wars: Episode V - The Empire Strikes Back	837759	English
Peter Jackson	297	3.1E+08	Action Adventure Drama Fantasy	Christopher Lee	The Lord of the Rings: The Fellowship of the Ring	1238746	English
Christopher Nolan	642	2.9E+08	Action Adventure Sci-Fi Thriller	Leonardo DiCaprio	Inception	1468200	English
David Fincher	315	3.7E+07	Drama	Brad Pitt	Fight Club	1347461	English
George Lucas	282	4.6E+08	Action Adventure Fantasy Sci-Fi	Harrison Ford	Star Wars: Episode IV - A New Hope	911037	English
Peter Jackson	294	3.4E+08	Action Adventure Drama Fantasy	Christopher Lee	The Lord of the Rings: The Two Towers	1100446	English
Lana Wachowski	313	1.7E+08	Action Sci-Fi	Keanu Reeves	The Matrix	1217752	English
Milos Forman	149	1.1E+08	Drama	Scatman Crothers	One Flew Over the Cuckoo's Nest	680041	English
Martin Scorsese	192	4.7E+07	Biography Crime Drama	Robert De Niro	Goodfellas	728685	English
Fernando Meirelles	214	7563397	Crime Drama	Alice Braga	City of God	533200	Portuguese
Akira Kurosawa	153	269061	Action Adventure Drama	Takashi Shimura	Seven Samurai	229012	Japanese
Steven Spielberg	219	2.2E+08	Action Drama War	Tom Hanks	Saving Private Ryan	881236	English
Jonathan Demme	185	1.3E+08	Crime Drama Horror Thriller	Anthony Hopkins	The Silence of the Lambs	887467	English
David Fincher	216	1E+08	Crime Drama Mystery Thriller	Morgan Freeman	Se7en	1023511	English
Christopher Nolan	712	1.9E+08	Adventure Drama Sci-Fi	Matthew McConaughey	Interstellar	928227	English
Bryan Singer	162	2.3E+07	Crime Drama Mystery Thriller	Kevin Spacey	The Usual Suspects	740918	English
Tony Kaye	162	6712241	Crime Drama	Ethan Suplee	American History X	782437	English
Charles Chaplin	120	163245	Comedy Drama Family	Paulette Goddard	Modern Times	143086	English
Hayao Miyazaki	246	1E+07	Adventure Animation Family Fantasy	Bunta Sugawara	Spirited Away	417971	Japanese
Roger Allers	186	4.2E+08	Adventure Animation Drama Family Musical	Matthew Broderick	The Lion King	644348	English
Steven Spielberg	234	2.4E+08	Action Adventure	Harrison Ford	Raiders of the Lost Ark	661017	English
Christopher Nolan	813	4.5E+08	Action Thriller	Tom Hardy	The Dark Knight Rises	1144337	English
Robert Zemeckis	198	2.1E+08	Adventure Comedy Sci-Fi	Lea Thompson	Back to the Future	732212	English
James Cameron	210	2E+08	Action Sci-Fi	Joe Morton	Terminator 2: Judgment Day	744891	English
Paul Verhoeven	265	1.8E+08	Action Drama	Dirk Bogarde	Black Robe	889697	English

- For the cleaning the data we use filter feature in the excel as the raw data contains like null (or) black values and unwanted columns like facebook likes.. ect in them and those columns and rows are not useful for analysis.

MOVIES WITH HIGHEST PROFIT

movie_title	num_voted_users	language	budget	title_year	imdb_score	PROFIT
Avatar	886204	English	237000000	2009	7.9	523505847
Jurassic World	418214	English	150000000	2015	7	502177271
Titanic	793059	English	200000000	1997	7.7	458672302
Star Wars: Episode IV - A New Hope	911097	English	11000000	1977	8.7	449935665
E.T. the Extra-Terrestrial	281842	English	10500000	1982	7.9	424449459
The Avengers	995415	English	220000000	2012	8.1	403279547
The Avengers	995415	English	220000000	2012	8.1	403279547
The Lion King	644348	English	45000000	1994	8.5	377783777
Star Wars: Episode I - The Phantom Menace	534658	English	115000000	1999	6.5	359544677
The Dark Knight	1676169	English	185000000	2008	9	348316061
The Hunger Games	701607	English	78000000	2012	7.3	329999255
Deadpool	479047	English	58000000	2016	8.1	305024263
The Hunger Games: Catching Fire	498397	English	130000000	2013	7.6	294645577
Jurassic Park	613473	English	63000000	1993	8.1	293784000
Despicable Me 2	286877	English	76000000	2013	7.5	292049635
American Sniper	325264	English	58800000	2014	7.3	291323553
Finding Nemo	692482	English	94000000	2003	8.2	286838870
Shrek 2	314630	English	150000000	2004	7.2	286471036
The Lord of the Rings: The Return of the King	1215718	English	94000000	2003	8.9	283019252
Star Wars: Episode VI - Return of the Jedi	681857	English	32500000	1983	8.4	276625409
Forrest Gump	1251222	English	55000000	1994	8.8	274691196
Star Wars: Episode V - The Empire Strikes Back	837759	English	18000000	1980	8.8	272158751
Home Alone	311075	English	18000000	1990	7.5	267761243
Star Wars: Episode III - Revenge of the Sith	520104	English	113000000	2005	7.6	267262555
Spider-Man	544665	English	139000000	2002	7.3	264706375
Minions	142403	English	74000000	2015	6.4	262029560
The Sixth Sense	704766	English	40000000	1999	8.1	253501675
Jaws	412454	English	8000000	1975	8	252000000
Frozen	421658	English	150000000	2013	7.6	250736600

- For the finding the profit of the movies we consider the difference between the gross and budget.
- Then we got the profit of that particular movie and movies with highest profit are avatar , Jurassic world , Titanic and .. etc.

MOVIES WITH HIGHEST PROFIT

The outliers are:-

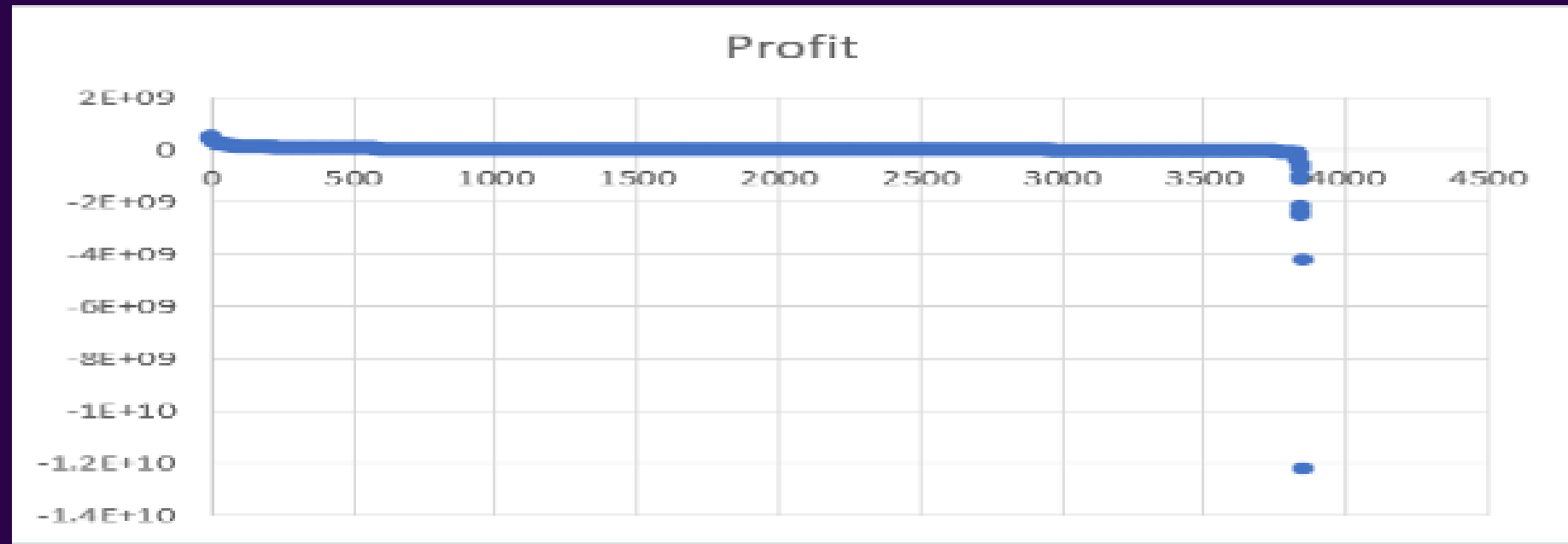
-12213298588

-4199788333

-2499804112

-2127109510

-1099560838



IMDB TOP 250 FILMS

IMDb_Top_250_MOVIES	IMDB_SCORE	RANK
The Shawshank Redemption	9.3	1
The Godfather	9.2	2
The Dark Knight	9	3
The Godfather: Part II	9	4
Star Wars: Episode VII - The Force Awakens	7.1	5
The Lord of the Rings: The Return of the King	8.9	6
Pulp Fiction	8.9	7
Schindler's List	8.9	8
The Good, the Bad and the Ugly	8.9	9
12 Angry Men	8.9	10
Forrest Gump	8.8	11
Star Wars: Episode V - The Empire Strikes Back	8.8	12
The Lord of the Rings: The Fellowship of the Ring	8.8	13
Inception	8.8	14
Fight Club	8.8	15
Star Wars: Episode IV - A New Hope	8.7	16
The Lord of the Rings: The Two Towers	8.7	17
The Matrix	8.7	18
One Flew Over the Cuckoo's Nest	8.7	19
Goodfellas	8.7	20
City of God	8.7	21
Seven Samurai	8.7	22
Saving Private Ryan	8.6	23
The Silence of the Lambs	8.6	24
Se7en	8.6	25
Interstellar	8.6	26
The Usual Suspects	8.6	27
American History X	8.6	28
Casablanca	8.6	29
Modern Times	8.6	30

- For the finding the imdb top 250 films I have used the filter feature in excel where I have sorted the films based on imdb_score and conditions like num_voted_users are above 25000 and ranking them by imdb_score.
- The movies with highest imdb_score are the Shawshank redemption , the godfather.. etc.

IMDB TOP FOREIGN FILMS

TOP_FOREIGN_LANG_FILM
The Good, the Bad and the Ugly
City of God
Seven Samurai
Spirited Away
The Lives of Others
Children of Heaven
Airlift
A Separation
Oldboy
Das Boot
Baahubali: The Beginning
Amélie
Princess Mononoke
The Hunt
Metropolis
Downfall
Pan's Labyrinth
The Secret in Their Eyes
Incendies
Howl's Moving Castle
Lage Raho Munna Bhai
Amores Perros
The Celebration
Solaris
Elite Squad
The Sea Inside
Tae Guk Gi: The Brotherhood of War
Akira
A Fistful of Dollars
Central Station

- For finding the top foreign language films we have used to filter feature in the excel.
- Where the condition for filtering is imdb_score and language of film which is language's which are other than english

TOP 10 DIRECTOR'S BY IMDB SCORE

Top 10 Director's	Mean of imdb_score
John Blanchard	9.5
Cary Bell	8.7
Mitchell Altieri	8.7
Sadyk Sher-Niyaz	8.7
Charles Chaplin	8.6
Mike Mayhall	8.6
Damien Chazelle	8.5
Majid Majidi	8.5
Raja Menon	8.5
Ron Fricke	8.5

- For finding the top 10 director by mean of imdb_score where I have used the pivot table feature in the excel.
- Where I have calculated the mean of imdb_score for their movies and sorted them in terms of the imdb_score if the two director's have score I have arranged them by alphabetical order.

POPULAR GENRE'S

GENRE	Count of Movies_Title
Drama	235
Comedy	208
Comedy Drama	190
Comedy Drama Romance	187
Comedy Romance	158
Drama Romance	152
Crime Drama Thriller	101
Horror	70
Action Crime Drama Thriller	67
Action Crime Thriller	65
Drama Thriller	64
Crime Drama	62
Horror Thriller	56
Crime Drama Mystery Thriller	55
Action Adventure Sci-Fi	51
Comedy Crime	51
Documentary	50
Action Adventure Thriller	46
Drama Mystery Thriller	37
Action Adventure Sci-Fi Thriller	35
Biography Drama	35
Horror Mystery Thriller	35
Action Adventure Fantasy	30
Action Comedy Crime	30

- For finding the popular genres I have used the pivot table feature in excel where I have counted the no . of movies which are under that particular genre.
- After that I have sorted them by higghest movies under that particular genre.

CHART

DATASET FOR MERYL STREEP

director_name	num_critic_for_rev	genres	actor_1_name	movie_title	num_voted_user	num_user_for_review	language	country
Nancy Meyers	187	Comedy Drama Romance	Meryl Streep	It's Complicated	69860	214	English	USA
Curtis Hanson	42	Action Adventure Crime Thriller	Meryl Streep	The River Wild	32544	69	English	USA
Nora Ephron	252	Biography Drama Romance	Meryl Streep	Julie & Julia	79264	277	English	USA
David Frankel	208	Comedy Drama Romance	Meryl Streep	The Devil Wears Prada	286178	631	English	USA
Robert Redford	227	Drama Thriller War	Meryl Streep	Lions for Lambs	41170	298	English	USA
Sydney Pollack	66	Biography Drama Romance	Meryl Streep	Out of Africa	52339	200	English	USA
David Frankel	234	Comedy Drama Romance	Meryl Streep	Hope Springs	34258	178	English	USA
Carl Franklin	64	Drama	Meryl Streep	One True Thing	9283	112	English	USA
Stephen Frears	87	Biography Comedy Drama Music Romance	Meryl Streep	Florence Foster Jenkins	2167	32	English	UK
Stephen Daldry	174	Drama Romance	Meryl Streep	The Hours	102123	660	English	USA
Phyllida Lloyd	331	Biography Drama History	Meryl Streep	The Iron Lady	82327	350	English	UK
Robert Altman	211	Comedy Drama Music	Meryl Streep	A Prairie Home Companion	19655	280	English	USA
Fred Zinnemann	38	Drama	Meryl Streep	Julia	6454	44	English	USA

CHART

DATASET FOR LEONARDO DICAPRIO

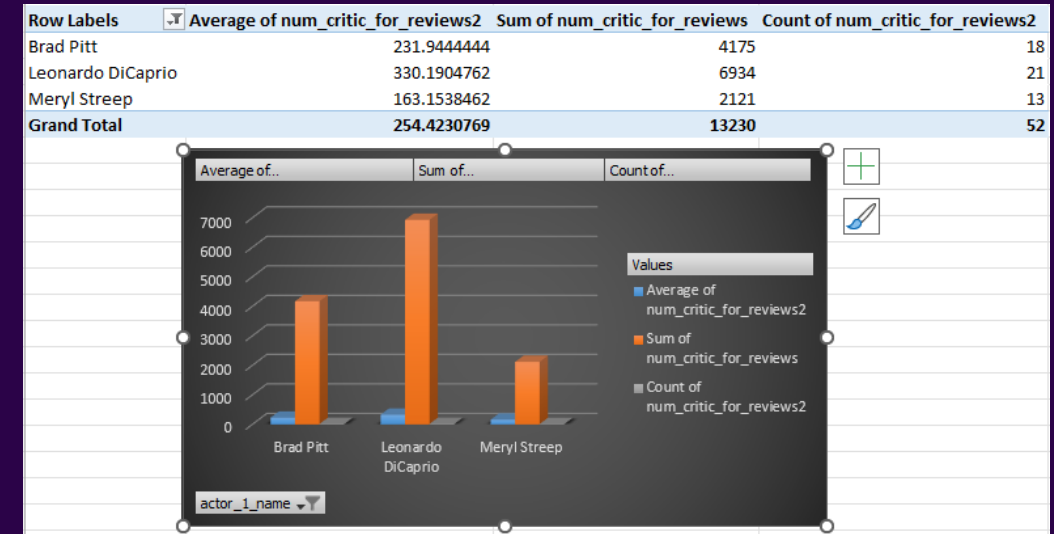
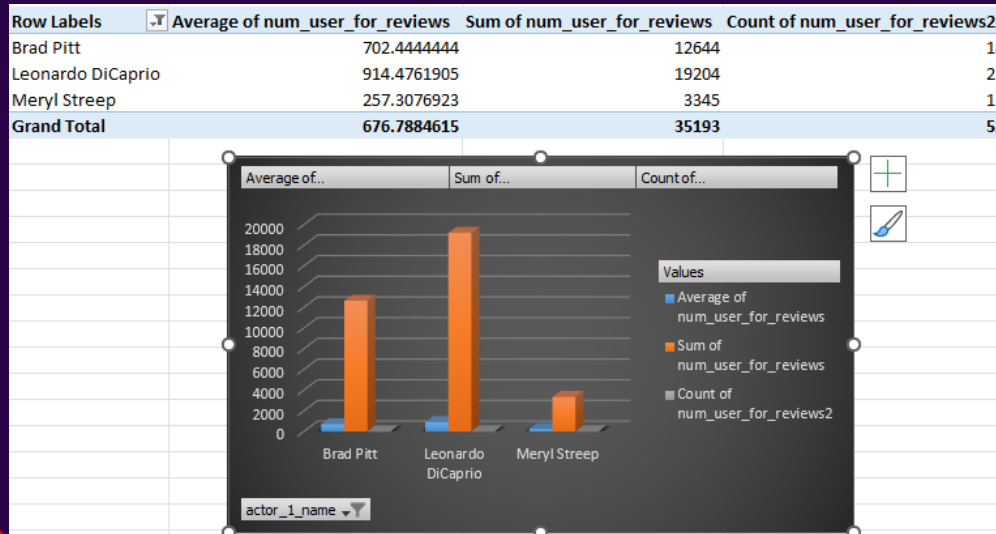
director_name	um_critc_for_rev	genres	actor_1_name	movie_title	num_voted_user	num_user_for_review	language	country
James Cameron	315	Drama Romance	Leonardo DiCaprio	Titanic	793059	2528	English	USA
Baz Luhrmann	490	Drama Romance	Leonardo DiCaprio	The Great Gatsby	362912	753	English	Australia
Christopher Nolan	642	Action Adventure Sci-Fi Thriller	Leonardo DiCaprio	Inception	1468200	2803	English	USA
Alejandro G. Iñárritu	556	Adventure Drama Thriller Western	Leonardo DiCaprio	The Revenant	406020	1188	English	USA
Martin Scorsese	267	Biography Drama	Leonardo DiCaprio	The Aviator	264318	799	English	USA
Quentin Tarantino	765	Drama Western	Leonardo DiCaprio	Django Unchained	955174	1193	English	USA
Edward Zwick	166	Adventure Drama Thriller	Leonardo DiCaprio	Blood Diamond	400292	657	English	Germany
Martin Scorsese	606	Biography Comedy Crime Drama	Leonardo DiCaprio	The Wolf of Wall Street	780588	1138	English	USA
Martin Scorsese	233	Crime Drama	Leonardo DiCaprio	Gangs of New York	314033	1166	English	USA
Martin Scorsese	352	Crime Drama Thriller	Leonardo DiCaprio	The Departed	873649	2054	English	USA
Martin Scorsese	490	Mystery Thriller	Leonardo DiCaprio	Shutter Island	786092	964	English	USA
Ridley Scott	238	Action Drama Thriller	Leonardo DiCaprio	Body of Lies	174248	263	English	USA
Steven Spielberg	194	Biography Crime Drama	Leonardo DiCaprio	Catch Me If You Can	525801	667	English	USA
Danny Boyle	118	Adventure Drama Thriller	Leonardo DiCaprio	The Beach	176169	548	English	USA
Sam Mendes	323	Drama Romance	Leonardo DiCaprio	Revolutionary Road	152591	414	English	USA
Randall Wallace	83	Action Adventure	Leonardo DiCaprio	The Man in the Iron Mask	125219	244	English	USA
Clint Eastwood	392	Biography Crime Drama	Leonardo DiCaprio	J. Edgar	102728	279	English	USA
Sam Raimi	63	Action Thriller Western	Leonardo DiCaprio	The Quick and the Dead	69197	216	English	Japan
Jerry Zaks	45	Drama	Leonardo DiCaprio	Marvin's Room	20163	71	English	USA
Baz Luhrmann	106	Drama Romance	Leonardo DiCaprio	Romeo + Juliet	167750	506	English	USA
Baz Luhrmann	490	Drama Romance	Leonardo DiCaprio	The Great Gatsby	362933	753	English	Australia

CHART

DATASET FOR BRAD PITT

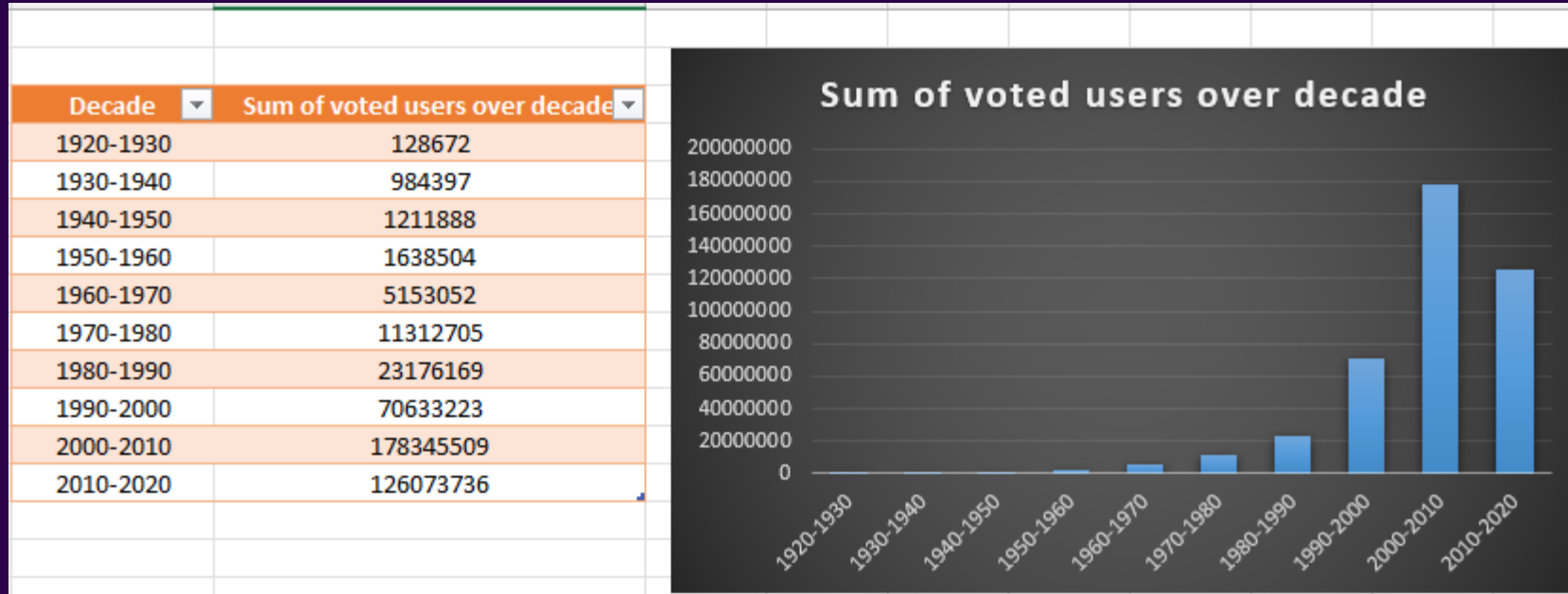
director_name	um_critic_for_rev	genres	actor_1_name	movie_title	num_voted_user	num_user_for_review	language	country
David Fincher	362	Drama Fantasy Romance	Brad Pitt	The Curious Case of Benjamin Button	459346	822	English	USA
Wolfgang Petersen	220	Adventure	Brad Pitt	Troy	381672	1694	English	USA
Steven Soderbergh	198	Crime Thriller	Brad Pitt	Ocean's Twelve	284852	627	English	USA
Doug Liman	233	Action Comedy Crime Romance Thriller	Brad Pitt	Mr. & Mrs. Smith	348861	798	English	USA
Tony Scott	142	Action Crime Thriller	Brad Pitt	Spy Game	121259	361	English	Germany
Steven Soderbergh	186	Crime Thriller	Brad Pitt	Ocean's Eleven	402645	845	English	USA
David Ayer	406	Action Drama War	Brad Pitt	Fury	303185	701	English	USA
Jean-Jacques Annaud	76	Adventure Biography Drama History War	Brad Pitt	Seven Years in Tibet	96385	119	English	USA
David Fincher	315	Drama	Brad Pitt	Fight Club	1347461	2968	English	USA
Patrick Gilmore	98	Adventure Animation Comedy Drama Family Fantasy	Brad Pitt	Sinbad: Legend of the Seven Seas	36144	91	English	USA
Neil Jordan	120	Drama Fantasy Horror	Brad Pitt	Interview with the Vampire: The Vamp	239752	406	English	USA
Terrence Malick	584	Drama Fantasy	Brad Pitt	The Tree of Life	136367	975	English	USA
Andrew Dominik	273	Biography Crime Drama History Western	Brad Pitt	The Assassination of Jesse James by t	136104	415	English	USA
Alejandro G. Iñárritu	285	Drama	Brad Pitt	Babel	243799	908	English	France
Angelina Jolie Pitt	131	Drama Romance	Brad Pitt	By the Sea	7976	61	English	USA
Andrew Dominik	414	Crime Thriller	Brad Pitt	Killing Them Softly	111625	369	English	USA
Tony Scott	122	Action Crime Drama Romance Thriller	Brad Pitt	True Romance	163492	460	English	USA
Tom DiCillo	10	Comedy Music Romance	Brad Pitt	Johnny Suede	3013	24	English	Switzerland

CHART



- For finding the actor who's has highest mean of num_user_for_reviews and num_critic_for_reviews we have used the pivot table in excel.
- Where we have used three actor's name and found the mean of num_user_for_reviews and num_critic_for_reviews for those particular actor's.
- By analyzing the data we got know The critic and user favorite actor is **LEONARDO DICAPRIO**.

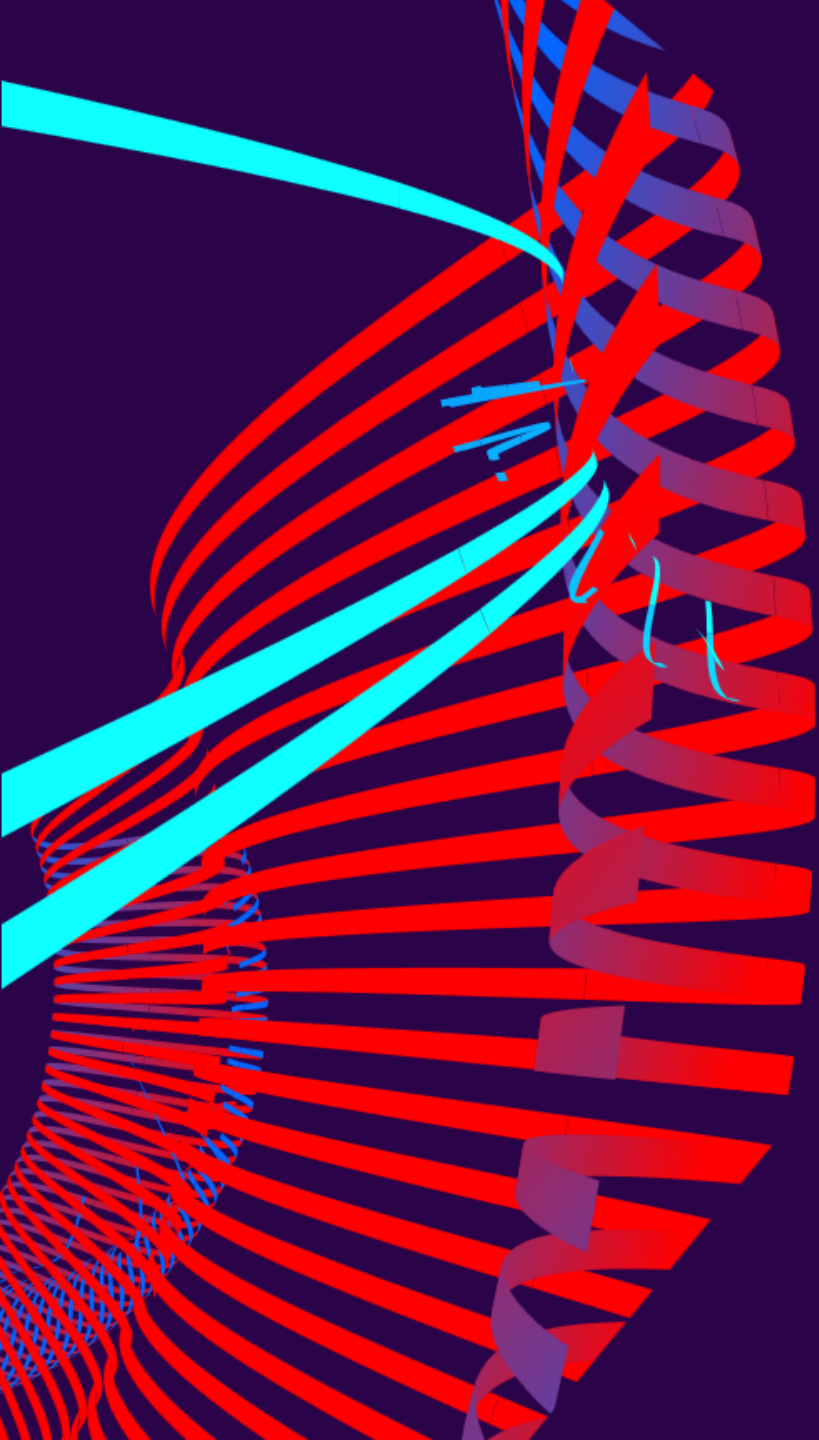
CHART



- For finding the sum of user voted over the decade we have used the pivot table feature in the excel where we have calculated the sum of voted users per year and then calculated in terms of decade.
- After that visualized the data in charts.

RESULT

- This project really helped me to understand the in-depth knowledge of the Microsoft excel.
- Where I have learned to use the pivot table feature efficiently and smartly to solve problems.
- Another excellent filter of the excel is filter where we can sort the data according to the requirements and needs.
- Data visualization in excel is also quite easy and this project has helped to improve my logical and data analytical skill.



THANK YOU