

IMDB MOVIE ANALYSIS

PROJECT DESCRIPTION

- Given dataset of IMDB movies having information about movie actors, directors, critics, audience, genre, rating, facebook likes etc.
- Our task is to analyse the data and find major trends, popular movies, genres, directors and give insights on the basis of given dataset.

APPROACH

- Understanding the data
- Getting familiar with the data values and attributes
- Analysing the data by applying filters and formulas
- Visualise the data with graph/ charts to get better understanding

TECH STACK USED

Microsoft Excel

A. CLEANING THE DATA

Dropped columns which are not needed for the analysis

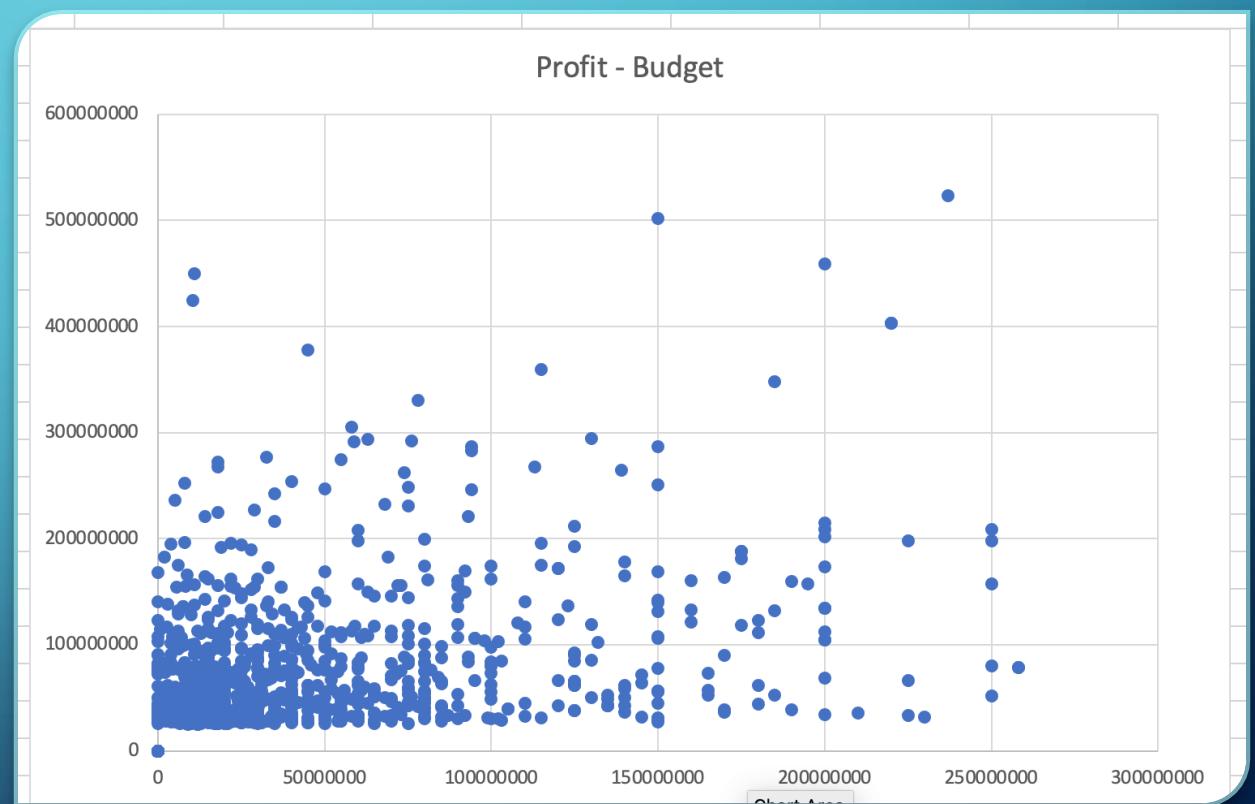
Replaced blanks with values like 0 or 'zzz' so that they don't interfere in analysis

After cleaning data set has these columns - director_name,
num_critic_for_reviews, gross, genres, actor_1_name, movie_title,
num_voted_users, num_user_for_reviews, language, budget, title_year,
imdb_score

B. MOVIES WITH HIGHEST PROFIT

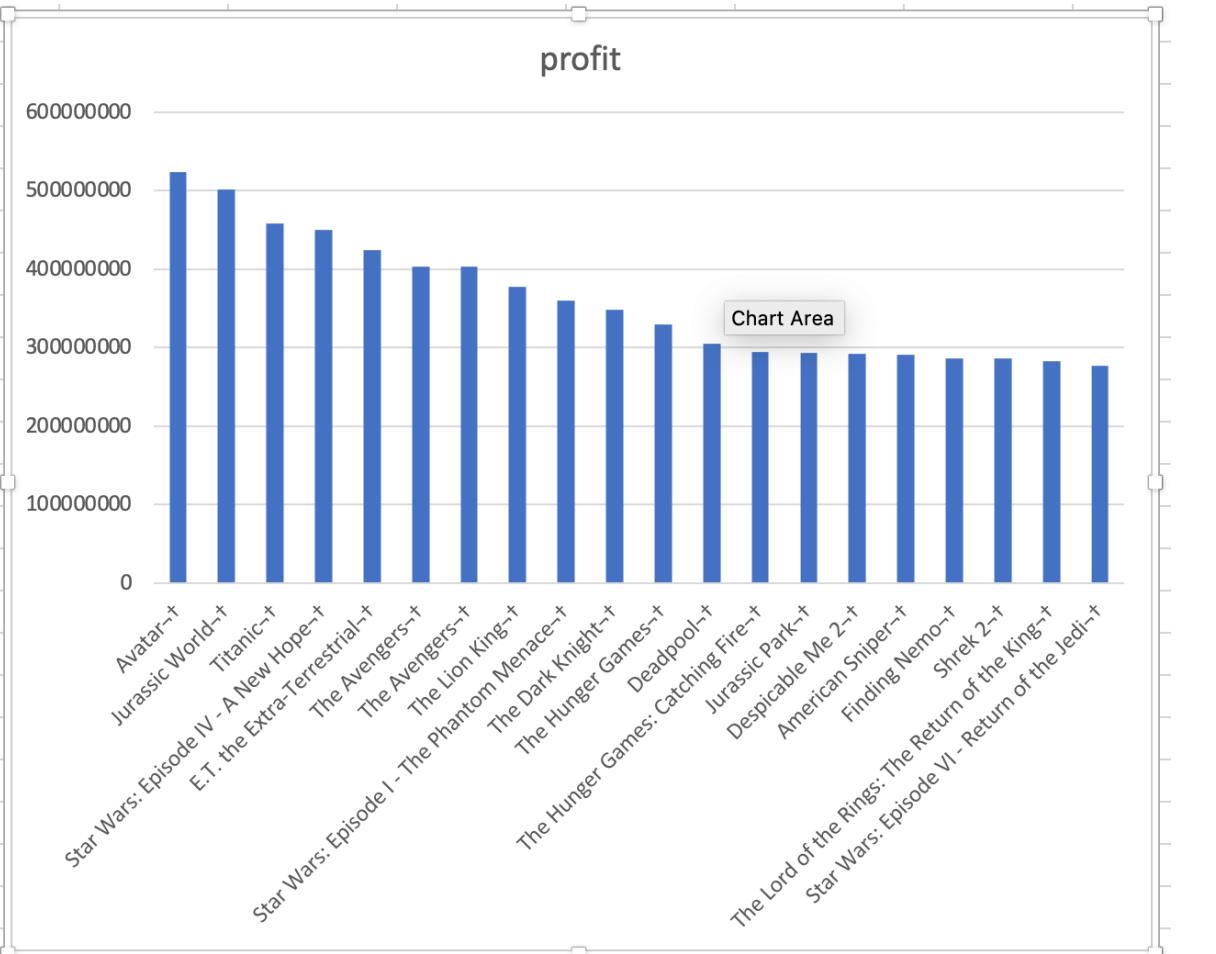
- Profit – budget chart of top 1000 movies according to profit
- Y axis – profit
- X axis – budget

From this graph we can infer that most of the movies which are low budget are able to earn low to medium range profits. On the other hand, there are few movies which are high budget and are able to earn very high profits.



TOP 20 MOVIES WITH HIGHEST PROFIT

'Avatar' is the most profitable movie according to given dataset.

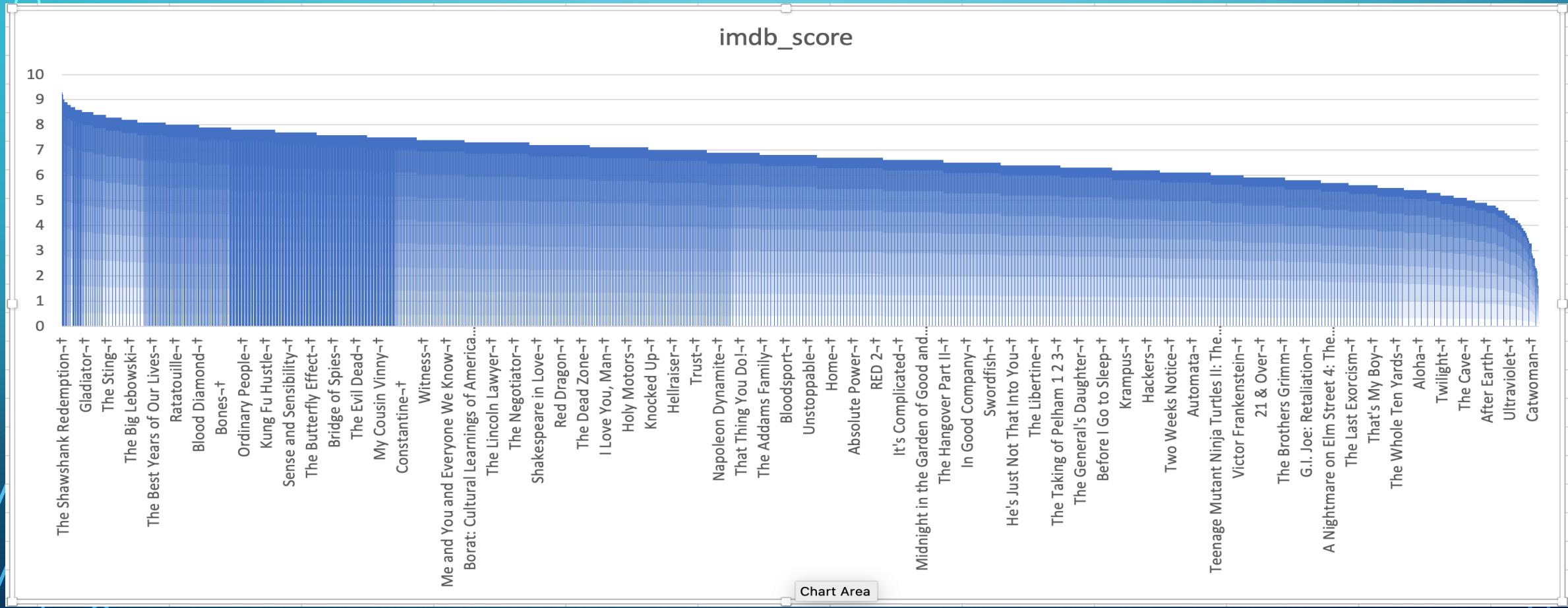


C. TOP 250 MOVIES WITH HIGHEST IMDB RATING

- NUMBER OF VOTED USERS > 25000
- Highest IMDB rating movie is 'The Shawshank Redemption' having 9.3 IMDB rating, in English language.

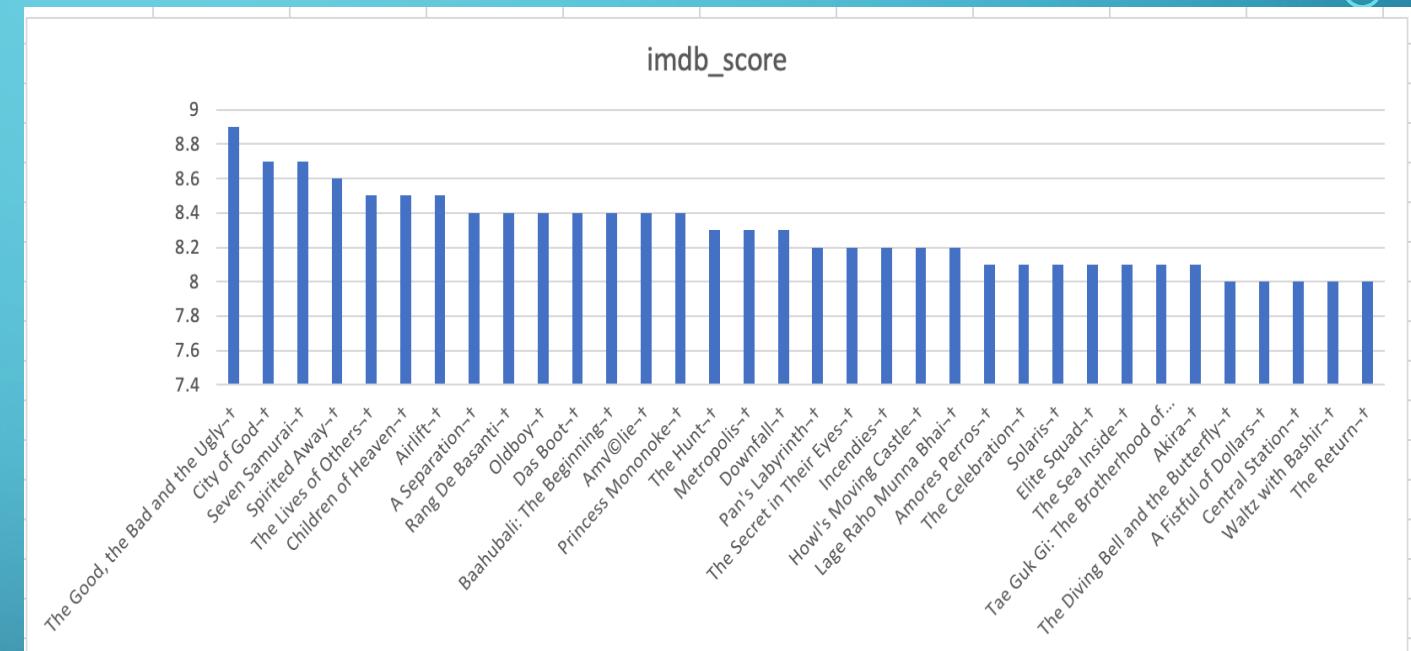
IMDb_Top_250	num_voted_users	imdb_score	language	Rank
The Shawshank Redemption~†	1689764	9.3	English	1
The Godfather~†	1155770	9.2	English	2
The Dark Knight~†	1676169	9	English	3
The Godfather: Part II~†	790926	9	English	4
Fargo~†	170055	9	English	5
The Lord of the Rings: The Return of the King~†	1215718	8.9	English	6
Pulp Fiction~†	1324680	8.9	English	7
Schindler's List~†	865020	8.9	English	8
The Good, the Bad and the Ugly~†	503509	8.9	Italian	9
12 Angry Men~†	447785	8.9	English	10
Forrest Gump~†	1251222	8.8	English	11
Star Wars: Episode V - The Empire Strikes Back~†	837759	8.8	English	12
The Lord of the Rings: The Fellowship of the Ring~†	1238746	8.8	English	13
Inception~†	1468200	8.8	English	14
Daredevil~†	213483	8.8	English	15
It's Always Sunny in Philadelphia~†	133415	8.8	English	16
Fight Club~†	1347461	8.8	English	17
Star Wars: Episode IV - A New Hope~†	911097	8.7	English	18
The Lord of the Rings: The Two Towers~†	1100446	8.7	English	19
The Matrix~†	1217752	8.7	English	20
One Flew Over the Cuckoo's Nest~†	680041	8.7	English	21
Goodfellas~†	728685	8.7	English	22
City of God~†	533200	8.7	Portuguese	23
Friday Night Lights~†	42746	8.7	English	24
Seven Samurai~†	229012	8.7	Japanese	25
Saving Private Ryan~†	881236	8.6	English	26
The Silence of the Lambs~†	887467	8.6	English	27
Se7en~†	1023511	8.6	English	28
Interstellar~†	928227	8.6	English	29
The Usual Suspects~†	740918	8.6	English	30
Hannibal~†	159910	8.6	English	31
Luther~†	70568	8.6	English	32
Spartacus: War of the Damned~†	173172	8.6	English	33
American History X~†	782437	8.6	English	34
Casablanca~†	387508	8.6	English	35
Modern Times~†	143086	8.6	English	36
It's a Wonderful Life~†	275720	8.6	English	37
Once Upon a Time in the West~†	216005	8.6	English	38
Spirited Away~†	417971	8.6	Japanese	39
The Lion King~†	644348	8.5	English	40
Raiders of the Lost Ark~†	661017	8.5	English	41
The Dark Knight Rises~†	1144337	8.5	English	42
Back to the Future~†	732212	8.5	English	43
Terminator 2: Judgment Day~†	744891	8.5	English	44
Gladiator~†	982637	8.5	English	45
The Green Mile~†	782610	8.5	English	46
Alien~†	563827	8.5	English	47

TOP 250 MOVIES WITH HIGHEST IMDB RATING



TOP FOREIGN LANGUAGE FILMS

Top_Foreign_Lang_Film	num_voted_use	imdb_sco	language	Rank
The Good, the Bad and the Ugly-†	503509	8.9	Italian	9
City of God-†	533200	8.7	Portuguese	23
Seven Samurai-†	229012	8.7	Japanese	25
Spirited Away-†	417971	8.6	Japanese	39
The Lives of Others-†	259379	8.5	German	55
Children of Heaven-†	27882	8.5	Persian	56
Airlift-†	30977	8.5	Hindi	61
A Separation-†	151812	8.4	Persian	66
Rang De Basanti-†	70233	8.4	Hindi	68
Oldboy-†	356181	8.4	Korean	74
Das Boot-†	168203	8.4	German	78
Baahubali: The Beginning-†	62756	8.4	Telugu	81
Amv@lie-†	534262	8.4	French	84
Princess Mononoke-†	221552	8.4	Japanese	85
The Hunt-†	170155	8.3	Danish	112
Metropolis-†	111841	8.3	German	114
Downfall-†	248354	8.3	German	115
Pan's Labyrinth-†	467234	8.2	Spanish	125
The Secret in Their Eyes-†	131831	8.2	Spanish	126
Incidencies-†	80429	8.2	French	134
Howl's Moving Castle-†	214091	8.2	Japanese	145
Lage Raho Munna Bhai-†	27569	8.2	Hindi	146
Amores Perros-†	173551	8.1	Spanish	185
The Celebration-†	65951	8.1	Danish	187
Solaris-†	54057	8.1	Russian	189
Elite Squad-†	81644	8.1	Portuguese	194
The Sea Inside-†	64556	8.1	Spanish	196
Tae Guk Gi: The Brotherhood of War-†	31943	8.1	Korean	199
Akira-†	106160	8.1	Japanese	202
The Diving Bell and the Butterfly-†	89906	8	French	235
A Fistful of Dollars-†	147566	8	Italian	238
Central Station-†	28951	8	Portuguese	240
Waltz with Bashir-†	46107	8	Hebrew	242
The Return-†	31589	8	Russian	243



Highest IMDB rated foreign language movie is 'The Good, the Bad and the Ugly' having 8.9 rating in Italian language. Overall rank of this movie is 9th.

D. TOP 10 DIRECTORS BY IMDB RATING

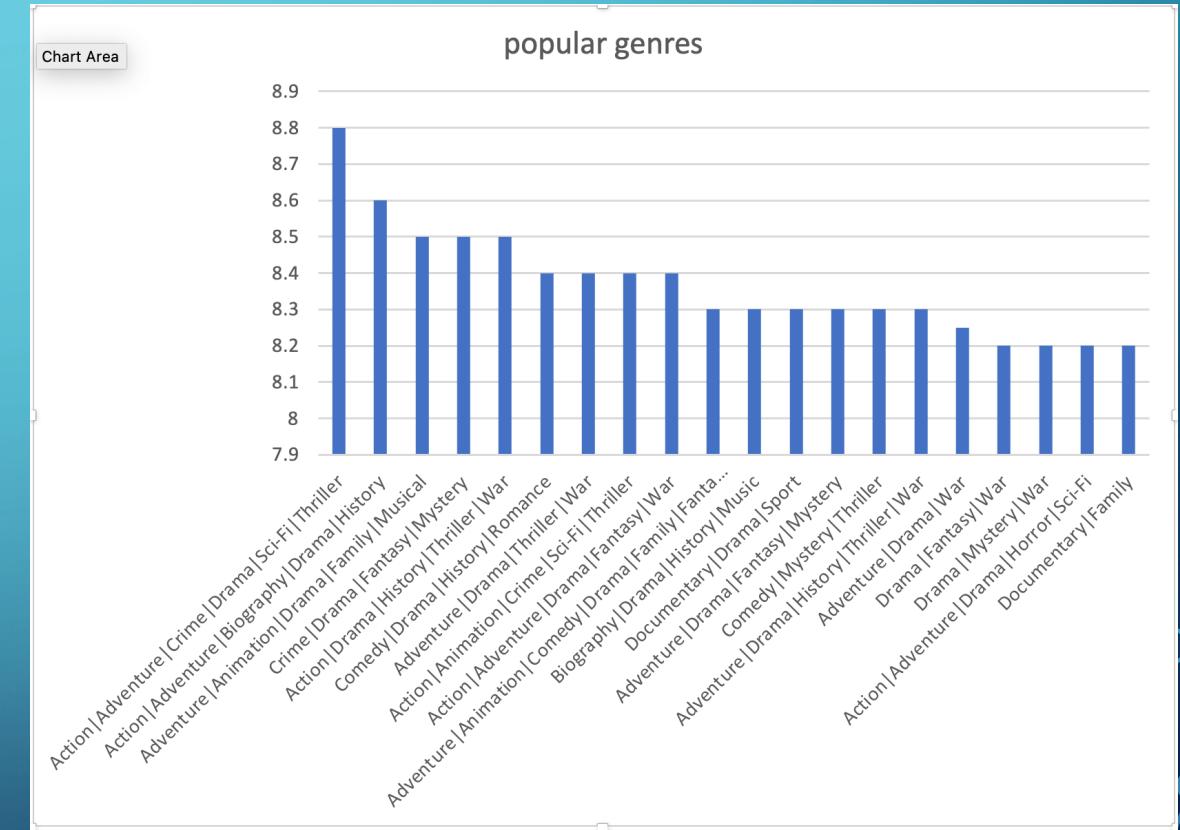


director_name	avg imdb rating
John Blanchard	9.5
Cary Bell	8.7
Mitchell Altieri	8.7
Sadyk Sher-Niyaz	8.7
Charles Chaplin	8.6
Mike Mayhall	8.6
Damien Chazelle	8.5
Majid Majidi	8.5
Raja Menon	8.5
Ron Fricke	8.5

Best director is John Blanchard having 9.5 avg. IMDB rating.

E. POPULAR GENRES BY IMDB RATING

genre	rating
Action Adventure Crime Drama Sci-Fi Thriller	8.8
Action Adventure Biography Drama History	8.6
Adventure Animation Drama Family Musical	8.5
Crime Drama Fantasy Mystery	8.5
Action Drama History Thriller War	8.5
Comedy Drama History Romance	8.4
Adventure Drama Thriller War	8.4
Action Animation Crime Sci-Fi Thriller	8.4
Action Adventure Drama Fantasy War	8.4
Adventure Animation Comedy Drama Family Fantasy	8.3
Biography Drama History Music	8.3
Documentary Drama Sport	8.3
Adventure Drama Fantasy Mystery	8.3
Comedy Mystery Thriller	8.3
Adventure Drama History Thriller War	8.3
Adventure Drama War	8.25
Drama Fantasy War	8.2
Drama Mystery War	8.2
Action Adventure Drama Horror Sci-Fi	8.2
Documentary Family	8.2

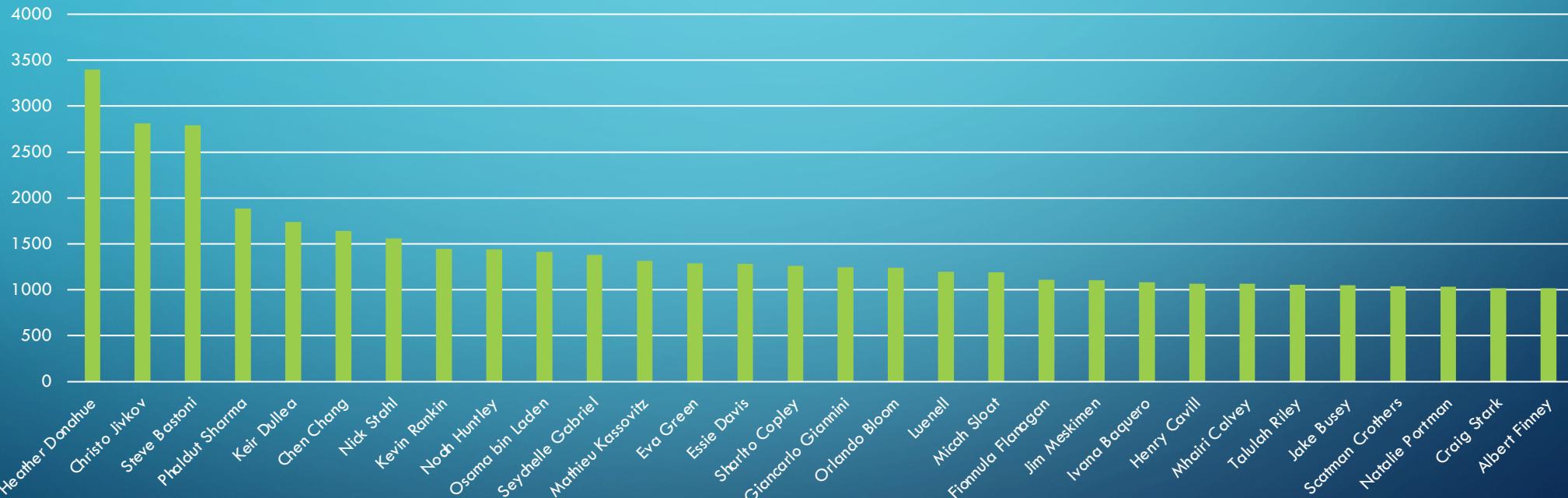


F. POPULAR ACTORS

Brad_Pitt	Leo_Caprio	Meryl_Streep
Ocean's Eleven-†	Titanic-†	The Devil Wears Prada-†
Mr. & Mrs. Smith-†	Inception-†	Out of Africa-†
Interview with the Vampire: The Vampire Chronicles-†	Catch Me If You Can-†	Julie & Julia-†
Fury-†	Django Unchained-†	Hope Springs-†
Ocean's Twelve-†	The Revenant-†	It's Complicated-†
Babel-†	Shutter Island-†	The Iron Lady-†
Killing Them Softly-†	The Departed-†	The Hours-†
Johnny Suede-†	The Great Gatsby-†	A Prairie Home Companion-†
True Romance-†	The Great Gatsby-†	The River Wild-†
By the Sea-†	Romeo + Juliet-†	One True Thing-†
The Tree of Life-†	The Man in the Iron Mask-†	Julia-†
The Curious Case of Benjamin Button-†	The Wolf of Wall Street-†	Lions for Lambs-†
Fight Club-†	J. Edgar-†	Florence Foster Jenkins-†
The Assassination of Jesse James by the Coward Robert Ford-†	The Aviator-†	
Seven Years in Tibet-†	Marvin's Room-†	
Sinbad: Legend of the Seven Seas-†	The Beach-†	
Troy-†	Revolutionary Road-†	
Spy Game-†	The Quick and the Dead-†	
	Gangs of New York-†	
	Body of Lies-†	
	Blood Diamond-†	

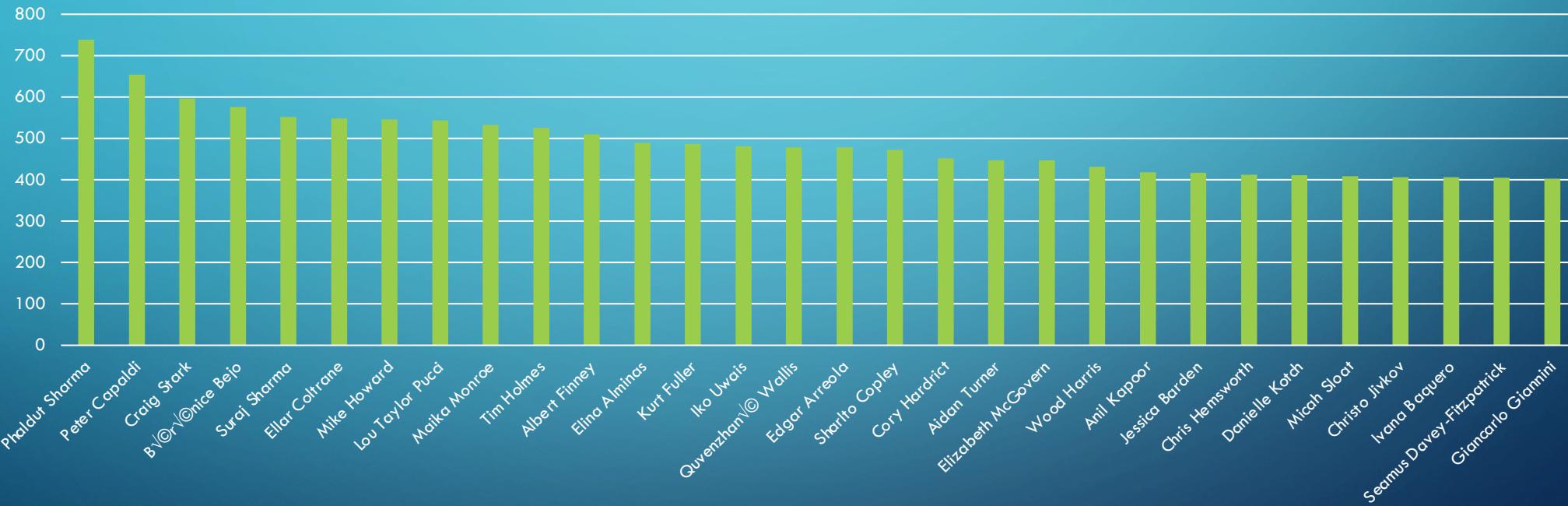
AUDIENCE FAVORITE ACTORS

avg num_users_for_review



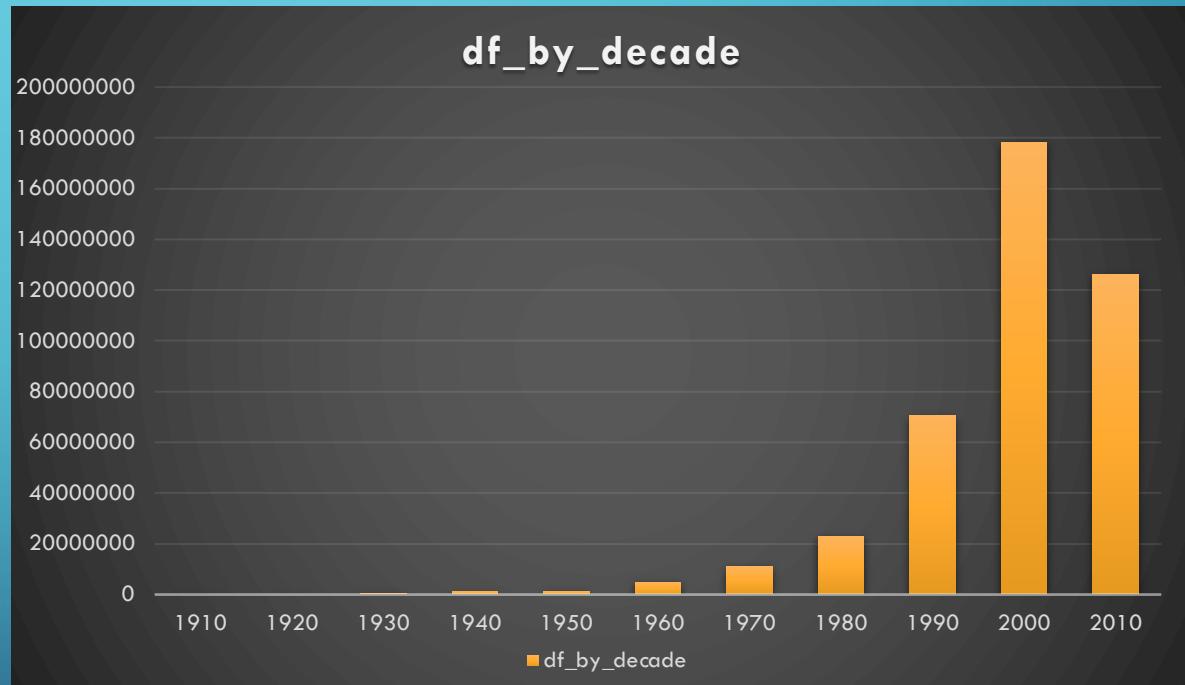
CRITIC FAVORITE ACTORS

avg num_critic_for_review



SUM OF USERS VOTED IN EACH DECADE

decade	df_by_decade
1910	10718
1920	128672
1930	984397
1940	1211888
1950	1638504
1960	5153052
1970	11312705
1980	23176169
1990	70633270
2000	178354686
2010	126202706



From 1910s to 2000s, no. of users voted have increased but their number decreased in 2010s.

RESULT

- Using Microsoft excel, I analyzed IMDB movie dataset and I find out that on average, low budget movies have more possibility to earn good profits whereas very few high budget movies can earn high profits.
- I listed out top directors, movies, popular actors and genres according to IMDB rating.
- As the years pass by, a greater number of users voted for movies, and we get more correct ratings.