

IMDB Movie Analysis

PROJECT DESCRIPTION:

In this project i.e. IMDB Movie Analysis, I provided support in the effort to produce business insights for the marketing, product, and by giving the data to the development teams working on this project using Microsoft Excel. I am utilising a variety of Excel queries during this process to obtain the relevant data. Based on the data obtained, I managed to figure out-

- Movies with highest profit.
- IMDB top 250.
- Best Directors.
- Popular Genres.
- Critic-favourite and audience-favourite actors.

APPROACH:

I initially examined the objective to find the relevant data that the team needed and then I imported the data into Excel and using different formulas in excel to find necessary conclusions about the company. By using the below Steps for EDA

- Understanding data columns and data
- Checking for missing data
- Clubbing columns with multiple categories
- Dropping columns.
- Removing null values.
- Drawing Data Summary

TECH-STACK USED:

I carried out the project using **MS Excel**.

INSIGHTS:

- I identify and clarify the main objective and by using Excel, I learn more about real-time analytics and obtain insights by applying various Excel formulas on this project. I identified the movie database info that was provided. I gained various insights into the Movies with highest profit, IMDB top 250, best directors, popular genres and critic-favourite and audience-favourite actors.

RESULTS:

By the completion of the project, I had improved my Excel skills, learned how to work with real-world data and experienced how to approach and resolve problems.

Based on the data provided, I was able to achieve a number of results which are listed below.

1. Cleaning the data: Dropping columns, removing null values, etc.

Task: Clean the data.

Result: In cleaning the data I have-

- Deleted all the duplicates record from the table.
- Removed the blank data rows.
- Removed unwanted special characters from the data.
- After cleaning the data the rows reduced from 5044 to 3724.

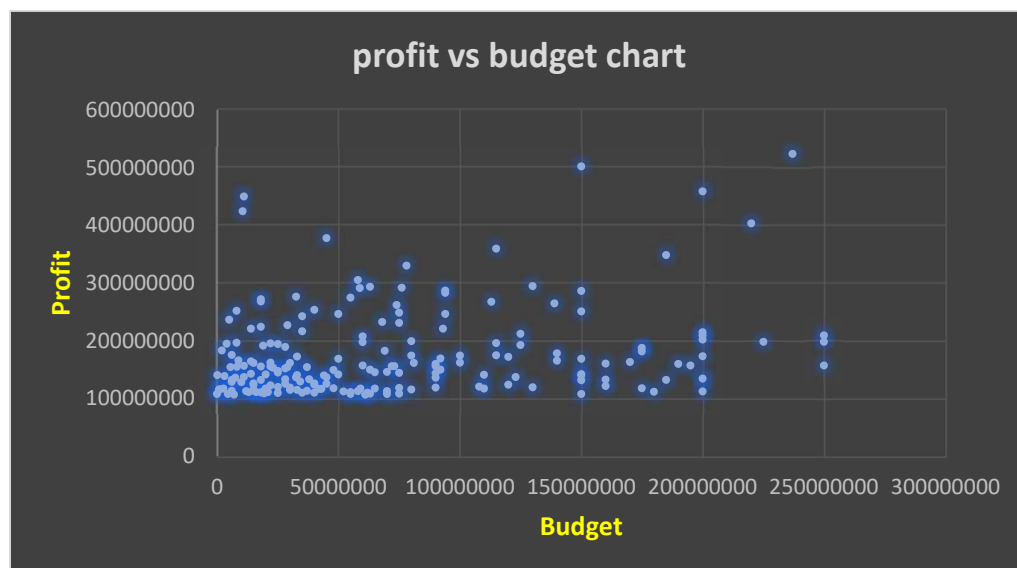
3704	921	21379315	Action Adventure Animation Comedy Family Sci-Fi	Elisabeth Harnois	Mars Needs Moms
3705	17000	47375327	Action Adventure Sci-Fi	Channing Tatum	Jupiter Ascending
3706	979	65171860	Adventure Fantasy	Eddie Marsan	Jack the Giant Slayer
3707	18000	38297305	Action Adventure Drama Fantasy	Keanu Reeves	47 Ronin
3708	14000	65173160	Action Adventure Sci-Fi Thriller	Liam Neeson	Battleship
3709	15000	665426	Adventure Animation Family Fantasy	Tom Hanks	The Polar Express
3710	778	11905519	Action Crime Drama Thriller	Jon Foo	The Protector
3711	640	73058679	Action Adventure Sci-Fi	Daryl Sabara	John Carter
3712	64	102055	Action	Nirut Sirichanya	Ong-bak 2
3713	51	14131298	Adventure Biography Drama History War	Paul Brooke	The Messenger: The Story of Joan of
3714	7	454255	Action Adventure Drama History War	Sarunyu Wongkrachang	The Legend of Suriyothai
3715	755	626809	Action Adventure Drama History War	Takeshi Kaneshiro	Red Cliff
3716	8000	3275443	Drama	Shah Rukh Khan	Kabhi Alvida Naa Kehna
3717	341	1687311	Drama Musical	MÃ-a Maestro	Tango
3718	43	10037390	Action Adventure Drama Sci-Fi Thriller	Hiroshi Abe	Godzilla 2000
3719	6	439162	Action Animation Sci-Fi	Mitsuo Iwata	Akira
3720	488	410388	Action Adventure Animation Family Sci-Fi Thriller	William Hootkins	Steamboy
3721	893	2298191	Adventure Animation Fantasy	Minnie Driver	Princess Mononoke
3722	9	195888	Drama Romance War	Marcell Nagy	Fateless
3723	717	211667	Crime Drama	Min-sik Choi	Lady Vengeance
3724	629	2201412	Comedy Drama Horror Sci-Fi	Doona Bae	The Host
3725					
3726					

2. Movies with highest profit: Create a new column called profit which contains the difference of the two columns: gross and budget.

Task: Find the movies with the highest profit.

Result:

movie_title	budget	gross	Profit
Avatar	237000000	760505847	523505847
Jurassic World	150000000	652177271	502177271
Titanic	200000000	658672302	458672302
Star Wars: Episode IV - A New Hope	11000000	460935665	449935665
E.T. the Extra-Terrestrial	10500000	434949459	424449459
The Avengers	220000000	623279547	403279547
The Lion King	45000000	422783777	377783777
Star Wars: Episode I - The Phantom Menace	115000000	474544677	359544677
The Dark Knight	185000000	533316061	348316061
The Hunger Games	78000000	407999255	329999255
Deadpool	58000000	363024263	305024263
The Hunger Games: Catching Fire	130000000	424645577	294645577
Jurassic Park	63000000	356784000	293784000
Despicable Me 2	76000000	368049635	292049635
American Sniper	58800000	350123553	291323553
Finding Nemo	94000000	380838870	286838870
Shrek 2	150000000	436471036	286471036
The Lord of the Rings: The Return of the King	94000000	377019252	283019252
Star Wars: Episode VI - Return of the Jedi	32500000	309125409	276625409
Forrest Gump	55000000	329691196	274691196
Star Wars: Episode V - The Empire Strikes Back	18000000	290158751	272158751
Home Alone	18000000	285761243	267761243
Star Wars: Episode III - Revenge of the Sith	113000000	380262555	267262555
Spider-Man	139000000	403706375	264706375
Minions	74000000	336029560	262029560
The Sixth Sense	40000000	293501675	253501675
Jaws	8000000	260000000	252000000



3. Top 250: Create a new column IMDb_Top_250 and store the top 250 movies with the highest IMDb Rating.

Result:

num_voted_users>25000	imdb_score	imdb_top250	Rank
1689764	9.3	The Shawshank Redemption	1
1155770	9.2	The Godfather	2
1676169	9	The Dark Knight	3
790926	9	The Godfather: Part II	4
1324680	8.9	Pulp Fiction	5
1215718	8.9	The Lord of the Rings: The Return of the King	6
865020	8.9	Schindler's List	7
503509	8.9	The Good, the Bad and the Ugly	8
1468200	8.8	Inception	9
1347461	8.8	Fight Club	10
1251222	8.8	Forrest Gump	11
1238746	8.8	The Lord of the Rings: The Fellowship of the Ring	12
837759	8.8	Star Wars: Episode V - The Empire Strikes Back	13
1217752	8.7	The Matrix	14
1100446	8.7	The Lord of the Rings: The Two Towers	15
911097	8.7	Star Wars: Episode IV - A New Hope	16
728685	8.7	Goodfellas	17
680041	8.7	One Flew Over the Cuckoo's Nest	18
533200	8.7	City of God	19
229012	8.7	Seven Samurai	20
1023511	8.6	Se7en	21
928227	8.6	Interstellar	22
887467	8.6	The Silence of the Lambs	23
881236	8.6	Saving Private Ryan	24
782437	8.6	American History X	25
num_voted_users>25000	imdb_score	imdb_top250	Rank
740918	8.6	The Usual Suspects	26
417971	8.6	Spirited Away	27
143086	8.6	Modern Times	28
1144337	8.5	Back to the Future	29
982637	8.5	The Green Mile	30
955174	8.5	Apocalypse Now	31
873649	8.5	Psycho	32
845580	8.5	The Prestige	33
844052	8.5	Whiplash	34
782610	8.5	Alien	35
744891	8.5	Gladiator	36
732212	8.5	Terminator 2: Judgment Day	37
661017	8.5	The Dark Knight Rises	38
644348	8.5	Raiders of the Lost Ark	39
563827	8.5	Django Unchained	40
497946	8.5	Star Wars: Episode VI - Return of the Jedi	41
450676	8.5	The Departed	42
422432	8.5	Memento	43
399138	8.5	The Lives of Others	44
259379	8.5	Children of Heaven	45
27882	8.5	The Pianist	46
822500	8.4	Aliens	47
736638	8.4	Reservoir Dogs	48
718837	8.4	A Separation	49
681857	8.4	American Beauty	50

num_voted_users>25000	imdb_score	imdb_top250	Rank
664719	8.4	Oldboy	51
573541	8.4	Das Boot	52
534262	8.4	Princess Mononoke	53
488537	8.4	WALL·E	54
356181	8.4	Requiem for a Dream	55
221552	8.4	Toy Story 3	56
221000	8.4	Amélie	57
192775	8.4	Once Upon a Time in America	58
168203	8.4	Lawrence of Arabia	59
151812	8.4	Braveheart	60
980946	8.3	Inglourious Basterds	61
885175	8.3	2001: A Space Odyssey	62
666937	8.3	Room	63
665575	8.3	Unforgiven	64
623757	8.3	The Sting	65
604904	8.3	Up	66
600996	8.3	Some Like It Hot	67
544884	8.3	Inside Out	68
537442	8.3	Eternal Sunshine of the Spotless Mind	69
515306	8.3	Good Will Hunting	70
427357	8.3	Amadeus	71
414219	8.3	Snatch	72
382240	8.3	The Hunt	73
345198	8.3	Toy Story	74
277505	8.3	Batman Begins	75
270790	8.3	L.A. Confidential	76
248354	8.3	Raging Bull	77
235133	8.3	Finding Nemo	78

num_voted_users>25000	imdb_score	imdb_top250	Rank
235133	8.3	Finding Nemo	78
175607	8.3	Indiana Jones and the Last Crusade	79
175196	8.3	Scarface	80
170155	8.3	Metropolis	81
161288	8.3	Monty Python and the Holy Grail	82
111841	8.3	Downfall	83
791783	8.2	Trainspotting	84
780588	8.2	V for Vendetta	85
692482	8.2	Gone with the Wind	86
610568	8.2	Die Hard	87
592582	8.2	How to Train Your Dragon	88
561773	8.2	A Beautiful Mind	89
537419	8.2	Incendies	90
485430	8.2	The Bridge on the River Kwai	91
469561	8.2	On the Waterfront	92
467234	8.2	The Secret in Their Eyes	93
461609	8.2	The Thing	94
426359	8.2	Lock, Stock and Two Smoking Barrels	95
414976	8.2	The Big Lebowski	96
333542	8.2	Warrior	97
332276	8.2	Howl's Moving Castle	98
272670	8.2	Gran Torino	99
258078	8.2	Casino	100
215340	8.2	Captain America: Civil War	101
214091	8.2	The Avengers	102
149444	8.2	Pan's Labyrinth	103
131831	8.2	The Wolf of Wall Street	104
100890	8.2	Into the Wild	105

num_voted_users>25000	imdb_score	imdb_top250	Rank
131831	8.2	The Wolf of Wall Street	104
100890	8.2	Into the Wild	105
80429	8.2	Blade Runner	106
995415	8.1	Deadpool	107
809474	8.1	Guardians of the Galaxy	108
786092	8.1	Stand by Me	109
735784	8.1	12 Years a Slave	110
704766	8.1	Monsters, Inc.	111
682155	8.1	The Help	112
667983	8.1	Groundhog Day	113
656640	8.1	The Grand Budapest Hotel	114
613473	8.1	The Sixth Sense	115
612060	8.1	The Revenant	116
600266	8.1	Spotlight	117
585659	8.1	Pirates of the Caribbean: The Curse of the Black Pearl	118
580999	8.1	Elite Squad	119
569841	8.1	Butch Cassidy and the Sundance Kid	120
552503	8.1	Amores Perros	121
491077	8.1	Rocky	122
482064	8.1	The Truman Show	123
479047	8.1	Jurassic Park	124
475518	8.1	The Terminator	125
472488	8.1	The Bourne Ultimatum	126
467613	8.1	Million Dollar Baby	127
439176	8.1	Annie Hall	128
437418	8.1	No Country for Old Men	129
406020	8.1	Shutter Island	130
383591	8.1	The Princess Bride	131
375240	8.1	Gone Girl	132
372990	8.1	Prisoners	133
318955	8.1	Platoon	134
312629	8.1	Tae Guk Gi: The Brotherhood of War	135
num_voted_users>25000	imdb_score	imdb_top250	Rank
338383	8	Doctor Zhivago	173
326494	8	Bohemia	174
307639	8	Mulholland Drive	175
284252	8	Aladdin	176
277451	8	Mystic River	177
266020	8	Bowling for Columbine	178
260939	8	The Incredibles	179
242599	8	Cinderella Man	180
241030	8	Serenity	181
235992	8	My Name Is Khan	182
190030	8	Sling Blade	183
186485	8	The Sound of Music	184
168398	8	Central Station	185
163492	8	Persepolis	186
152306	8	In Bruges	187
148238	8	Blood Diamond	188
148172	8	Rain Man	189
147566	8	Before Sunset	190
128455	8	Avatar	191
123090	8	Casino Royale	192
113472	8	The Artist	193
112671	8	Dead Poets Society	194
79330	8	Magnolia	195
72443	8	Dallas Buyers Club	196
70194	8	Big Fish	197
69759	8	Dancer in the Dark	198
66610	8	Shaun of the Dead	199
63733	8	Brazil	200
55816	8	Black Swan	201
46107	8	True Romance	202
29839	8	Kill Bill: Vol. 2	203
28951	8	Her	204
886204	7.9	E.T. the Extra-Terrestrial	205
696338	7.9	Toy Story 2	206
483756	7.9	Crouching Tiger, Hidden Dragon	207
483540	7.9	How to Train Your Dragon 2	208
467113	7.9	Iron Man	209
431620	7.9	The Hangover	210
.....

num_voted_users>25000	imdb_score	imdb_top250	Rank
483540	7.9	How to Train Your Dragon 2	208
467113	7.9	Iron Man	209
431620	7.9	The Hangover	210
407601	7.9	Big Hero 6	211
396396	7.9	Her	212
385871	7.9	Straight Outta Compton	213
361767	7.9	Edge of Tomorrow	214
361169	7.9	Halloween	215
357581	7.9	The Hobbit: The Desolation of Smaug	216
355810	7.9	The Untouchables	217
352695	7.9	The Remains of the Day	218
323353	7.9	Little Miss Sunshine	219
293304	7.9	Do the Right Thing	220
281842	7.9	Shrek	221
279093	7.9	My Fair Lady	222
275869	7.9	The Bourne Identity	223
272839	7.9	Once	224
260607	7.9	Nine Queens	225
251349	7.9	Hot Fuzz	226
221128	7.9	The Blues Brothers	227
219008	7.9	Crash	228
217740	7.9	Walk the Line	229
207287	7.9	The Notebook	230
189032	7.9	The Hateful Eight	231
188637	7.9	The Fighter	232
157863	7.9	Eduard Scissorhands	233
157857	7.9	Halloween	234
149414	7.9	The Insider	235
142448	7.9	Nightcrawler	236
142416	7.9	The World's Fastest Indian	237
133526	7.9	Children of Men	238
132149	7.9	The Right Stuff	239
119928	7.9	Taken	240
101889	7.9	Before Midnight	241
101888	7.9	Glory	242
95362	7.9	4 Months, 3 Weeks and 2 Days	243
90827	7.9	Glory	244
70382	7.9	Ed Wood	245
66959	7.9	Captain Phillips	246
59524	7.9	The Wrestler	247
45703	7.9	Boogie Nights	248
45271	7.9	Amour	249
44763	7.9	Moan	250

Top foreign language film

Top foreign lang film	num_voted_users	language	imdb_score
The Good, the Bad and the Ugly	503509	Italian	8.9
City of God	533200	Portuguese	8.7
Seven Samurai	229012	Japanese	8.7
Spirited Away	417971	Japanese	8.6
The Lives of Others	259379	German	8.5
Children of Heaven	27882	Persian	8.5
Amélie	534262	French	8.4
Oldboy	356181	Korean	8.4
Princess Mononoke	221552	Japanese	8.4
Das Boot	168203	German	8.4
A Separation	151812	Persian	8.4
Downfall	248354	German	8.3
The Hunt	170155	Danish	8.3
Metropolis	111841	German	8.3
Pan's Labyrinth	467234	Spanish	8.2
Howl's Moving Castle	214091	Japanese	8.2
The Secret in Their Eyes	131831	Spanish	8.2
Incendies	80429	French	8.2
Amores Perros	173551	Spanish	8.1
Akira	106160	Japanese	8.1
Elite Squad	81644	Portuguese	8.1
The Celebration	65951	Danish	8.1
The Sea Inside	64556	Spanish	8.1
Tae Guk Gi: The Brotherhood of War	31943	Korean	8.1
A Fistful of Dollars	147566	Italian	8
Persepolis	70194	French	8
My Name Is Khan	69759	Hindi	8
Waltz with Bashir	46107	Hebrew	8
Central Station	28951	Portuguese	8
Crouching Tiger, Hidden Dragon	217740	Mandarin	7.9
Hero	149414	Mandarin	7.9
Letters from Iwo Jima	132149	Japanese	7.9
Amour	70382	French	7.9
4 Months, 3 Weeks and 2 Days	44763	Romanian	7.9

4. Best Directors: Group the column using the director_name column.

Task: Find the best directors.

Result:

top_10_directors	mean_imdb_score
Akira Kurosawa	8.70
Tony Kaye	8.60
Charles Chaplin	8.60
Alfred Hitchcock	8.50
Damien Chazelle	8.50
Majid Majidi	8.50
Ron Fricke	8.50
Sergio Leone	8.43
Christopher Nolan	8.43
Richard Marquand	8.40

5. Popular Genres:

Task: Find popular genres.

Result:

popular_genres	Count of movie_title
Comedy Drama Romance	147
Drama	140
Comedy	137
Comedy Drama	137
Comedy Romance	131
Drama Romance	114
Crime Drama Thriller	81
Action Crime Thriller	55
Action Crime Drama Thriller	50
Action Adventure Sci-Fi	46

5. Charts:

Task: Find the critic-favorite and audience-favorite actors.

Result:

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

actor_1_name	movies	Meryl Streep	A Prairie Home Companion
Leonardo DiCaprio	Blood Diamond	Meryl Streep	Hope Springs
Leonardo DiCaprio	Body of Lies	Meryl Streep	It's Complicated
Leonardo DiCaprio	Catch Me If You Can	Meryl Streep	Julie & Julia
Leonardo DiCaprio	Django Unchained	Meryl Streep	Lions for Lambs
Leonardo DiCaprio	Gangs of New York	Meryl Streep	One True Thing
Leonardo DiCaprio	Inception	Meryl Streep	Out of Africa
Leonardo DiCaprio	J. Edgar	Meryl Streep	The Devil Wears Prada
Leonardo DiCaprio	Marvin's Room	Meryl Streep	The Hours
Leonardo DiCaprio	Revolutionary Road	Meryl Streep	The Iron Lady
Leonardo DiCaprio	Romeo + Juliet	Meryl Streep	The River Wild
Leonardo DiCaprio	Shutter Island	Brad Pitt	Babel
Leonardo DiCaprio	The Aviator	Brad Pitt	By the Sea
Leonardo DiCaprio	The Beach	Brad Pitt	Fight Club
Leonardo DiCaprio	The Departed	Brad Pitt	Fury
Leonardo DiCaprio	The Great Gatsby	Brad Pitt	Interview with the Vampire: The Vampire Chronicles
Leonardo DiCaprio	The Man in the Iron Mask	Brad Pitt	Killing Them Softly
Leonardo DiCaprio	The Quick and the Dead	Brad Pitt	Mr. & Mrs. Smith
Leonardo DiCaprio	The Revenant	Brad Pitt	Ocean's Eleven
Leonardo DiCaprio	The Wolf of Wall Street	Brad Pitt	Ocean's Twelve
		Brad Pitt	Seven Years in Tibet
		Brad Pitt	Sinbad: Legend of the Seven Seas
		Brad Pitt	Spy Game
		Brad Pitt	The Assassination of Jesse James by the Coward Robert Ford
		Brad Pitt	The Curious Case of Benjamin Button
		Brad Pitt	The Tree of Life
		Brad Pitt	Troy
		Brad Pitt	True Romance

mean of the num_critic_for_reviews and num_users_for_review

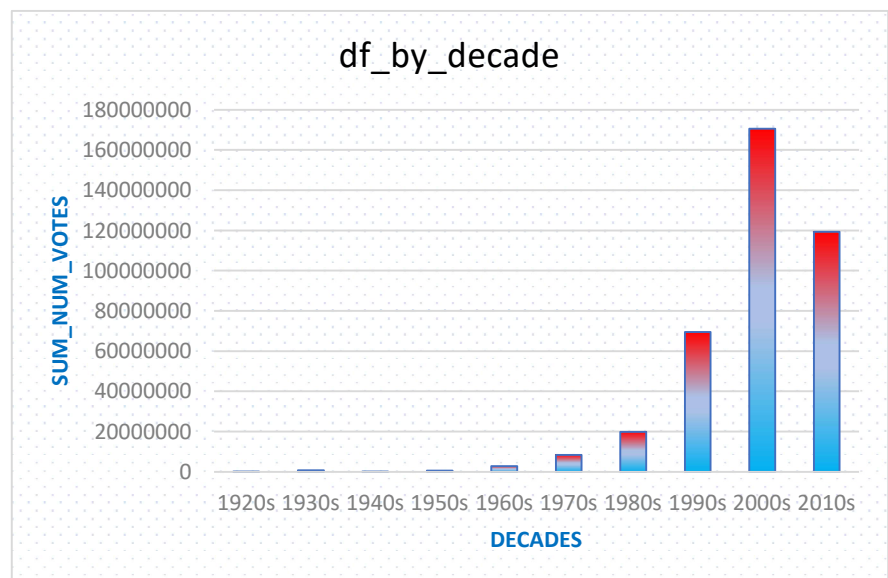
actor name	_num_user_for_reviews	_num_critic_for_reviews
Heather Donahue	3400	360
Christo Jivkov	2814	406
Steve Bastoni	2789	275
Phaldut Sharma	1885	738
Orlando Bloom	1842	259
Keir Dullea	1736	285
Eva Green	1708	389
Chen Chang	1641	287
Nick Stahl	1562	218
Albert Finney	1498	750
Kevin Rankin	1445	267
Noah Huntley	1441	224
Osama bin Laden	1416	288
Seychelle Gabriel	1382	280
Mathieu Kassovitz	1314	242
Essie Davis	1286	245
Sharlto Copley	1262	472
Giancarlo Giannini	1243	403
Christopher Lee	1237	155
Matt Frewer	1229	451
Luenell	1198	343
Micah Sleet	1189	409

actors which have the highest mean

Audience favourite actor	Heather Donahue
Critic favourite actor	Albert Finney

Decade column which represents the decade to which every movie belongs to

Decades	Users_voted
1920s	116387
1930s	804839
1940s	159517
1950s	678336
1960s	2982551
1970s	8523299
1980s	19987476
1990s	69581866
2000s	170711435
2010s	119432961



Link↓

<https://docs.google.com/spreadsheets/d/1S-keX9nexJ5UZyB3rH6DjPstfNQGipQe/edit?usp=sharing&ouid=115123843229850064657&rt=pof=true&sd=true>