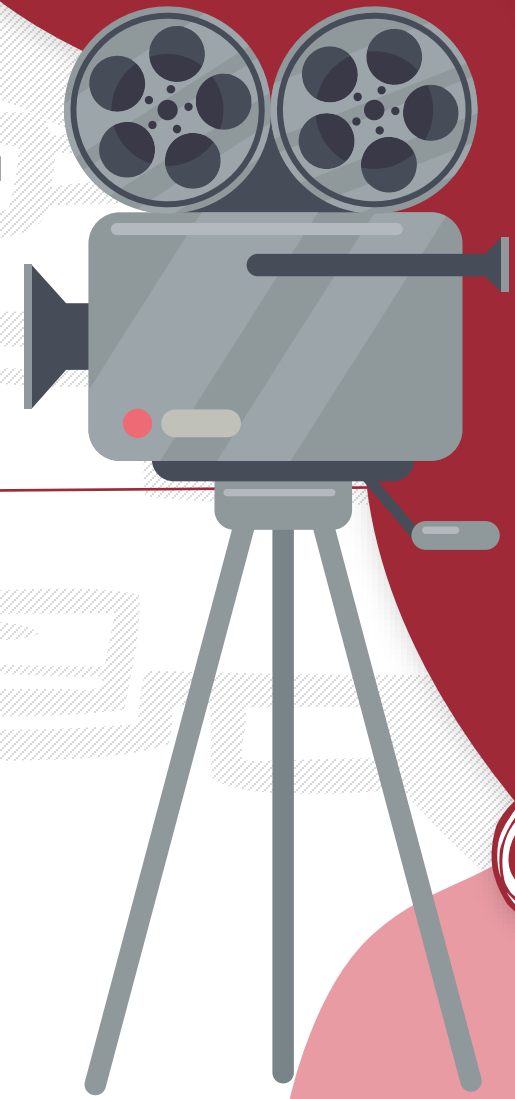


# IMDB MOVIE ANALYSIS



BY ANJALI GUPTA



# CONTENT



Description

---



Approach

---



Tech Stack Used

---



Insights



# PROJECT DESCRIPTION

A dataset is provided having various columns of different IMDB Movies. We are required to answer the questions given below. And provide our insights on it providing solutions to the problem.



# APPROACH

First, take the dataset and clean the dataset using Microsoft Excel.

Read the problems carefully and try to propose a solution to the problem by manipulating the data in MS Excel.

Finally provide our insight on the proposed result.

# TECH STACK USED



Microsoft Excel



Microsoft PowerPoint Presentation





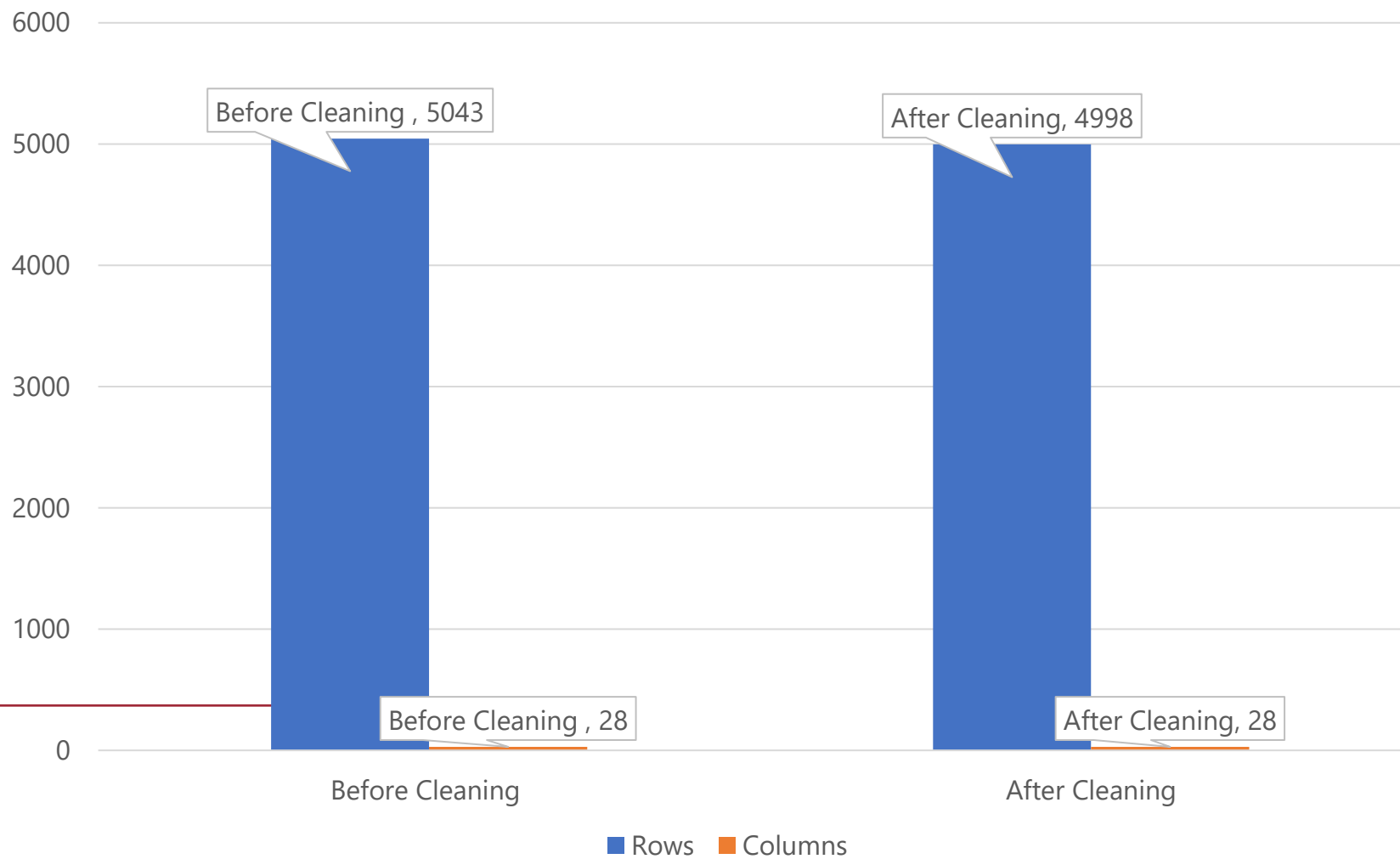
# INSIGHTS

# Cleaning the data

*Cleaning the data is one of the most significant step in data analysis.  
It gives us a clear picture of the dataset given to us  
We can clean the data by removing the duplicate values present in the data.  
We can first form the table of the dataset using the format table command and then remove duplicates present using the remove duplicate command hence cleaning the data, which reduced the number of rows present in the data*



## Number of rows decreased after cleaning



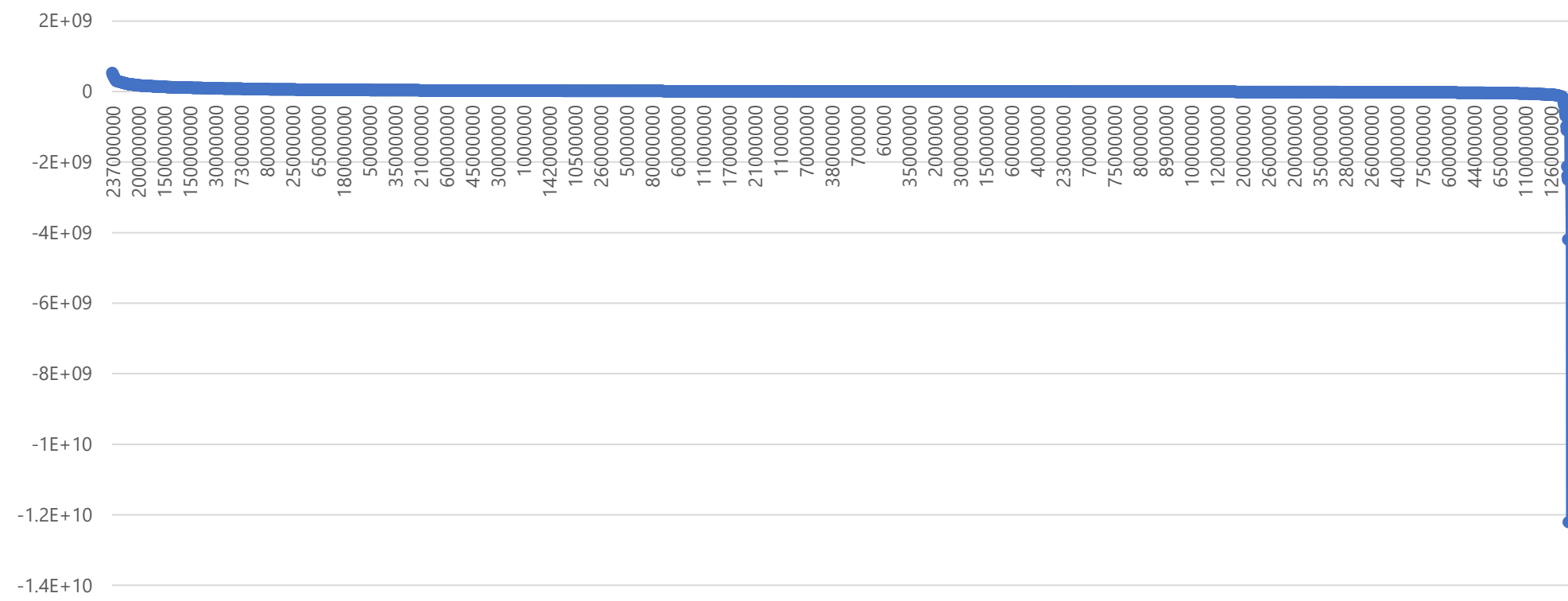


# Movies with the highest profit

*Movies were sorted in order of their profits which is nothing but the difference between the gross value and the budget.  
A graph was plotted with profit on y-axis and budget on x-axis  
From the data, we can infer that Avatar<sup>Â</sup> was the movie with the highest profit, with a profit of 523505847 from the scatter plot graph it is visible that there are few outliers in the data*



# Budget vs Profit



## Budget vs Profit



# Top 250

*The top 250 IMDB movies are shown in the following slides.*

*We can see from the result that The Shawshank Redemption is the top IMDb-rated movie with more than 25000 votes. It is an English movie.*

*Amongst foreign languages, i.e. other than the English Language. The Good, the Bad and the Ugly, an Italian movie, is at the top.*



Rank	IMDb rating	Movie Title	num_voted_users
1	9.3	The Shawshank Redemption	1689764
2	9.2	The Godfather	1155770
3	9	The Dark Knight	1676169
4	9	The Godfather: Part II	790926
5	9	Fargo	170055
6	8.9	The Lord of the Rings: The Return of the King	1215718
7	8.9	Pulp Fiction	1324680
8	8.9	Schindler's List	865020
9	8.9	The Good, the Bad and the Ugly	503509
10	8.9	12 Angry Men	447785
11	8.8	Forrest Gump	1251222
12	8.8	Star Wars: Episode V - The Empire Strikes Back	837759
13	8.8	The Lord of the Rings: The Fellowship of the Ring	1238746
14	8.8	Inception	1468200
15	8.8	Daredevil	213483
16	8.8	It's Always Sunny in Philadelphia	133415
17	8.8	Fight Club	1347461
18	8.7	Star Wars: Episode IV - A New Hope	911097
19	8.7	The Lord of the Rings: The Two Towers	1100446
20	8.7	The Matrix	1217752
21	8.7	One Flew Over the Cuckoo's Nest	680041
22	8.7	Goodfellas	728685
23	8.7	City of God	533200
24	8.7	Friday Night Lights	42746
25	8.7	Seven Samurai	229012
26	8.6	Saving Private Ryan	881236
27	8.6	The Silence of the Lambs	887467
28	8.6	Se7en	1023511
30	8.6	The Usual Suspects	740918





31	8.6	HannibalÃ	159910
32	8.6	LutherÃ	70568
33	8.6	Spartacus: War of the DamnedÃ	173172
34	8.6	American History XÃ	782437
35	8.6	CasablancaÃ	387508
36	8.6	Modern TimesÃ	143086
37	8.6	It's a Wonderful LifeÃ	275720
38	8.6	Once Upon a Time in the WestÃ	216005
39	8.6	Spirited AwayÃ	417971
40	8.5	The Lion KingÃ	644348
41	8.5	Raiders of the Lost ArkÃ	661017
42	8.5	The Dark Knight RisesÃ	1144337
43	8.5	Back to the FutureÃ	732212
44	8.5	Terminator 2: Judgment DayÃ	744891
45	8.5	GladiatorÃ	982637
46	8.5	The Green MileÃ	782610
47	8.5	AlienÃ	563827
48	8.5	Django UnchainedÃ	955174
49	8.5	Apocalypse NowÃ	450676
50	8.5	The DepartedÃ	873649
51	8.5	PsychoÃ	422432
52	8.5	MementoÃ	845580
53	8.5	The PrestigeÃ	844052
54	8.5	WhiplashÃ	399138
55	8.5	The Lives of OthersÃ	259379
56	8.5	Children of HeavenÃ	27882
57	8.5	OutlanderÃ	50391
58	8.5	EntourageÃ	135643
59	8.5	Dr. Strangelove or: How I Learned to Stop Worrying and Love the BombÃ	342585
60	8.5	The PianistÃ	497946





61	8.5	AirliftÂ	30977
62	8.4	Star Wars: Episode VI - Return of the JediÂ	681857
63	8.4	American BeautyÂ	822500
64	8.4	AliensÂ	488537
65	8.4	WALLÂ-EÂ	718837
66	8.4	A SeparationÂ	151812
67	8.4	BraveheartÂ	736638
68	8.4	Rang De BasantiÂ	70233
69	8.4	Reservoir DogsÂ	664719
70	8.4	Veronica MarsÂ	55524
71	8.4	The InbetweenersÂ	55987
72	8.4	PsychÂ	67315
73	8.4	M*A*S*HÂ	36265
74	8.4	OldboyÂ	356181
75	8.4	Requiem for a DreamÂ	573541
76	8.4	Stargate SG-1Â	63982
77	8.4	To Kill a MockingbirdÂ	215086
78	8.4	Das BootÂ	168203
79	8.4	Batman: The Dark Knight Returns, Part 2Â	30434
80	8.4	Lawrence of ArabiaÂ	192775
81	8.4	Baahubali: The BeginningÂ	62756
82	8.4	The ShiningÂ	610333
83	8.4	Once Upon a Time in AmericaÂ	221000
84	8.4	AmÂ©lieÂ	534262
85	8.4	Princess MononokeÂ	221552
86	8.3	Toy Story 3Â	544884
87	8.3	Inside OutÂ	345198
88	8.3	Toy StoryÂ	623757
89	8.3	The StingÂ	175607
90	8.3	Indiana Jones and the Last CrusadeÂ	515306



91	8.3	Good Will Hunting	604904
92	8.3	Up	665575
93	8.3	Unforgiven	277505
94	8.3	Batman Begins	980946
95	8.3	Inglourious Basterds	885175
96	8.3	2001: A Space Odyssey	427357
97	8.3	Amadeus	270790
98	8.3	L.A. Confidential	414219
99	8.3	Snatch	600996
100	8.3	Some Like It Hot	175196
101	8.3	Scarface	537442
102	8.3	Eternal Sunshine of the Spotless Mind	666937
103	8.3	Inside Job	55382
104	8.3	Room	161288
105	8.3	Monty Python and the Holy Grail	382240
106	8.3	Life	29450
107	8.3	Taxi Driver	507063
108	8.3	Singin' in the Rain	150020
109	8.3	The Apartment	109335
110	8.3	Judgment at Nuremberg	44457
111	8.3	The Hunt	170155
112	8.3	The Great Escape	165638
113	8.3	Metropolis	111841
114	8.3	Downfall	248354
115	8.3	Raging Bull	235133
116	8.2	Finding Nemo	692482
117	8.2	Gone with the Wind	215340
118	8.2	Captain America: Civil War	272670
119	8.2	Gran Torino	561773
120	8.2	A Beautiful Mind	610568







121	8.2	Die Hard	592582
122	8.2	How to Train Your Dragon	485430
123	8.2	The Bridge on the River Kwai	149444
124	8.2	Pan's Labyrinth	467234
125	8.2	The Secret in Their Eyes	131831
126	8.2	The Wolf of Wall Street	780588
127	8.2	V for Vendetta	791783
128	8.2	Trainspotting	469561
129	8.2	On the Waterfront	100890
130	8.2	Into the Wild	426359
131	8.2	Lock, Stock and Two Smoking Barrels	414976
132	8.2	The Big Lebowski	537419
133	8.2	Incendies	80429
134	8.2	It Happened One Night	64888
135	8.2	Blade Runner	461609
136	8.2	The Thing	258078
137	8.2	Rebecca	87424
138	8.2	Mr. Smith Goes to Washington	77392
139	8.2	Buffy the Vampire Slayer	101902
140	8.2	The Elephant Man	161972
141	8.2	Casino	333542
142	8.2	Warrior	332276
143	8.2	The Deer Hunter	232577
144	8.2	Howl's Moving Castle	214091
145	8.2	Lage Raho Munna Bhai	27569
146	8.1	The Avengers	995415
147	8.1	Deadpool	479047
148	8.1	Jurassic Park	613473
149	8.1	The Sixth Sense	704766
150	8.1	Monsters, Inc.	585659
151	8.1	Pirates of the Caribbean: The Curse of the Black Pearl	809474
152	8.1	Guardians of the Galaxy	682155
153	8.1	The Help	318955
154	8.1	Platoon	291603
155	8.1	The Martian	472488





156	8.1	The Bourne UltimatumÂ	491077
157	8.1	RockyÂ	375240
158	8.1	Gone GirlÂ	569841
159	8.1	Butch Cassidy and the Sundance KidÂ	152089
160	8.1	The Imitation GameÂ	467613
161	8.1	Million Dollar BabyÂ	482064
162	8.1	The Truman ShowÂ	667983
163	8.1	Groundhog DayÂ	437418
164	8.1	No Country for Old MenÂ	612060
165	8.1	The RevenantÂ	406020
166	8.1	Shutter IslandÂ	786092
167	8.1	Stand by MeÂ	271794
168	8.1	Kill Bill: Vol. 1Â	735784
169	8.1	12 Years a SlaveÂ	439176
170	8.1	Annie HallÂ	192940
171	8.1	Sin CityÂ	656640
172	8.1	The Grand Budapest HotelÂ	475518
173	8.1	The TerminatorÂ	600266
174	8.1	SpotlightÂ	195333
175	8.1	The Best Years of Our LivesÂ	40359
176	8.1	The Wizard of OzÂ	291875
177	8.1	There Will Be BloodÂ	372990
178	8.1	PrisonersÂ	383591
179	8.1	The Princess BrideÂ	294163
180	8.1	Hotel RwandaÂ	264533
181	8.1	Touching the VoidÂ	26926
182	8.1	Mad Max: Fury RoadÂ	552503
183	8.1	Amores PerrosÂ	173551
184	8.1	Before SunriseÂ	183288
185	8.1	The CelebrationÂ	65951
186	8.1	High NoonÂ	80193
187	8.1	SolarisÂ	54057
188	8.1	Cat on a Hot Tin RoofÂ	33741
189	8.1	The Man Who Shot Liberty ValanceÂ	53741
190	8.1	Donnie DarkoÂ	580999





191	8.1	NetworkÂ	103493
192	8.1	Elite SquadÂ	81644
193	8.1	A Christmas StoryÂ	104908
194	8.1	The Sea InsideÂ	64556
195	8.1	Barry LyndonÂ	101627
196	8.1	RushÂ	312629
197	8.1	Tae Guk Gi: The Brotherhood of WarÂ	31943
198	8.1	Hachi: A Dog's TaleÂ	155249
199	8.1	GandhiÂ	171726
200	8.1	AkiraÂ	106160
201	8	JawsÂ	412454
202	8	The ExorcistÂ	284252
203	8	AladdinÂ	260939
204	8	The IncrediblesÂ	479166
205	8	Dances with WolvesÂ	186485
206	8	The Sound of MusicÂ	148172
207	8	Rain ManÂ	383784
208	8	Slumdog MillionaireÂ	641997
209	8	The King's SpeechÂ	503631
210	8	Catch Me If You CanÂ	525801
211	8	Star TrekÂ	504419
212	8	The Pursuit of HappynessÂ	338383
213	8	Doctor ZhivagoÂ	55816
214	8	Black SwanÂ	551363
215	8	District 9Â	531737
216	8	Young FrankensteinÂ	112671
217	8	Dead Poets SocietyÂ	277451
218	8	Mystic RiverÂ	338415
219	8	RatatouilleÂ	473887
220	8	Fiddler on the RoofÂ	29839
221	8	Kill Bill: Vol. 2Â	512749
222	8	X-Men: Days of Future PastÂ	514125
223	8	JFKÂ	113472
224	8	The ArtistÂ	190030
225	8	Sling BladeÂ	72443





226	8	Dallas Buyers Club	326494
227	8	Boyhood	266020
228	8	Bowling for Columbine	123090
229	8	Casino Royale	470483
230	8	Casino Royale	470501
231	8	Sicko	66610
232	8	Shaun of the Dead	395921
233	8	The Diving Bell and the Butterfly	89906
234	8	Life of Pi	440084
235	8	The Perks of Being a Wallflower	351274
236	8	A Fistful of Dollars	147566
237	8	Before Sunset	168398
238	8	Central Station	28951
239	8	Her	355126
240	8	Waltz with Bashir	46107
241	8	The Return	31589
242	8	Night of the Living Dead	87978
243	8	The Man from Earth	129799
244	8	True Romance	163492
245	8	A Streetcar Named Desire	78454
246	8	The Hustler	62860
247	8	Rosemary's Baby	140527
248	8	Persepolis	70194
249	8	Days of Heaven	37594
250	8	Big Fish	350698



# TOP FOREIGN FILMS

IMDb	IMDb rating	Movie Title	num_voted_users	language
9	8.9	The Good, the Bad and the Ugly	503509	Italian
23	8.7	City of God	533200	Portuguese
25	8.7	Seven Samurai	229012	Japanese
39	8.6	Spirited Away	417971	Japanese
55	8.5	The Lives of Others	259379	German
56	8.5	Children of Heaven	27882	Persian
61	8.5	Airlift	30977	Hindi
66	8.4	A Separation	151812	Persian
68	8.4	Rang De Basanti	70233	Hindi
74	8.4	Oldboy	356181	Korean
78	8.4	Das Boot	168203	German
81	8.4	Baahubali: The Beginning	62756	Telugu
84	8.4	Amélie	534262	French
85	8.4	Princess Mononoke	221552	Japanese
111	8.3	The Hunt	170155	Danish
113	8.3	Metropolis	111841	German
114	8.3	Downfall	248354	German
124	8.2	Pan's Labyrinth	467234	Spanish
125	8.2	The Secret in Their Eyes	131831	Spanish
133	8.2	Incendies	80429	French
144	8.2	Howl's Moving Castle	214091	Japanese
145	8.2	Lage Raho Munna Bhai	27569	Hindi
183	8.1	Amores Perros	173551	Spanish
185	8.1	The Celebration	65951	Danish
187	8.1	Solaris	54057	Russian
192	8.1	Elite Squad	81644	Portuguese
194	8.1	The Sea Inside	64556	Spanish
197	8.1	Tae Guk Gi: The Brotherhood of War	31943	Korean
200	8.1	Akira	106160	Japanese
233	8	The Diving Bell and the Butterfly	89906	French
236	8	A Fistful of Dollars	147566	Italian
238	8	Central Station	28951	Portuguese
240	8	Waltz with Bashir	46107	Hebrew
241	8	The Return	31589	Russian
248	8	Persepolis	70194	French



# Best Directors

*The top 10 director list is given in the next slide which is found by doing the average IMDb rating for that director using a pivot table. Amongst the top 10 directors, John Blanchard had the highest IMDb average.*





Top 10 directors	Average IMDb
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
Majid Majidi	8.5
Raja Menon	8.5
Ron Fricke	8.5



# Popular Genres

*Amongst all the genres, the genre with the highest popularity is Drama. Therefore, it is the most popular genre among all other genres.  
Therefore, a movie made in Drama genre is sought to be well-received by the audience.*







Row Labels	Sum of Count
Drama	2571
Comedy	1862
Thriller	1390
Action	1143
Romance	1092
Adventure	914
Crime	883
Sci-Fi	604
Fantasy	602
Horror	556
Family	544
Mystery	492
Biography	292
Animation	242
Music	212
War	210
History	205
Sport	180
Musical	130
Documentary	121
Western	94
Film-Noir	6
Short	5
News	3
Reality-TV	2
Game-Show	1
(blank)	
<b>Grand Total</b>	<b>14356</b>



# Charts

*Here the best actors, popular and critics are found with the help of a pivot table. And the movies of some actors were listed separately. The decade of the movies was found using rounddown function in excel, which is further examined with the help of the pivot table and the number of users who voted in a given decade is observed.*

# MOVIES OF THE ACTORS MENTIONED



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



# CRITICS CHOICE BEST ACTORS

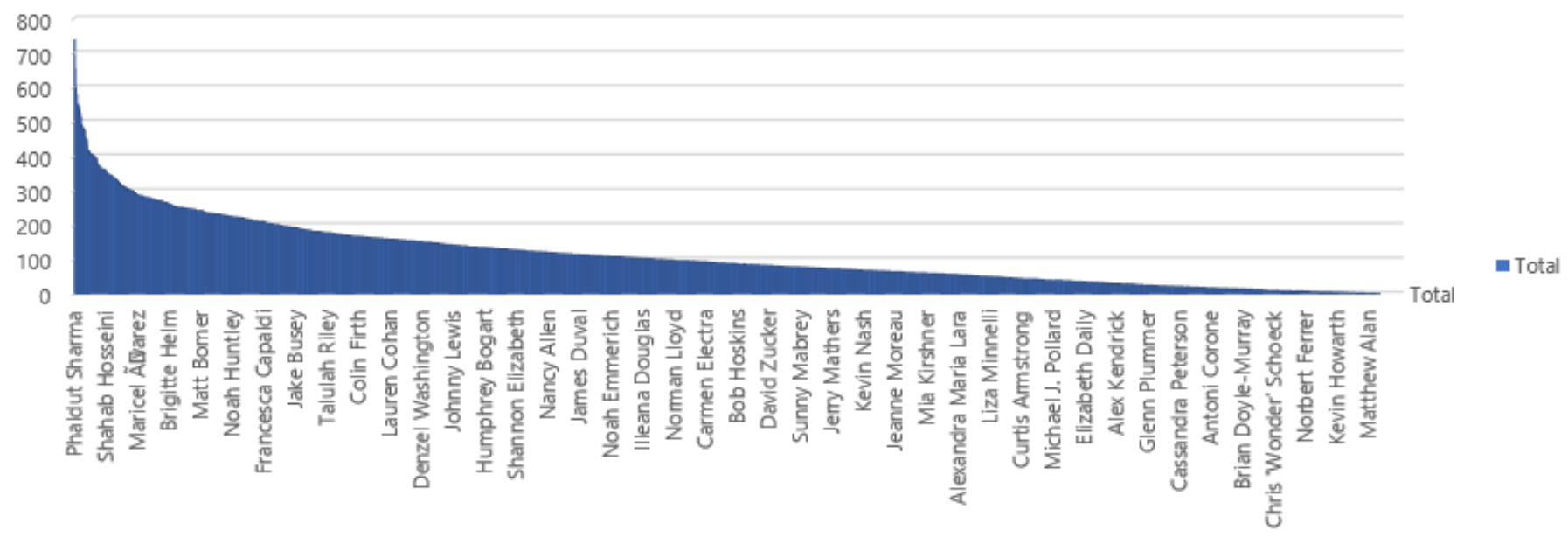


Row Labels	Average of num_critic_for_reviews
Phaldut Sharma	738
Gravity	738
Peter Capaldi	654
World War Z	654
Craig Stark	596
The Hateful Eight	596
BÃ©nice Bejo	576
The Artist	576
Suraj Sharma	552
Life of Pi	552
Ellar Coltrane	548
Boyhood	548
Mike Howard	546
The Master	546
Lou Taylor Pucci	543
Evil Dead	543
Maika Monroe	533
It Follows	533
Tim Holmes	525
Oz the Great and Powerful	525
Albert Finney	510
Skyfall	750
Tom Jones	30
Elina Alminas	489
Ex Machina	489
Kurt Fuller	487
Midnight in Paris	487





Mean num\_critic\_for\_review



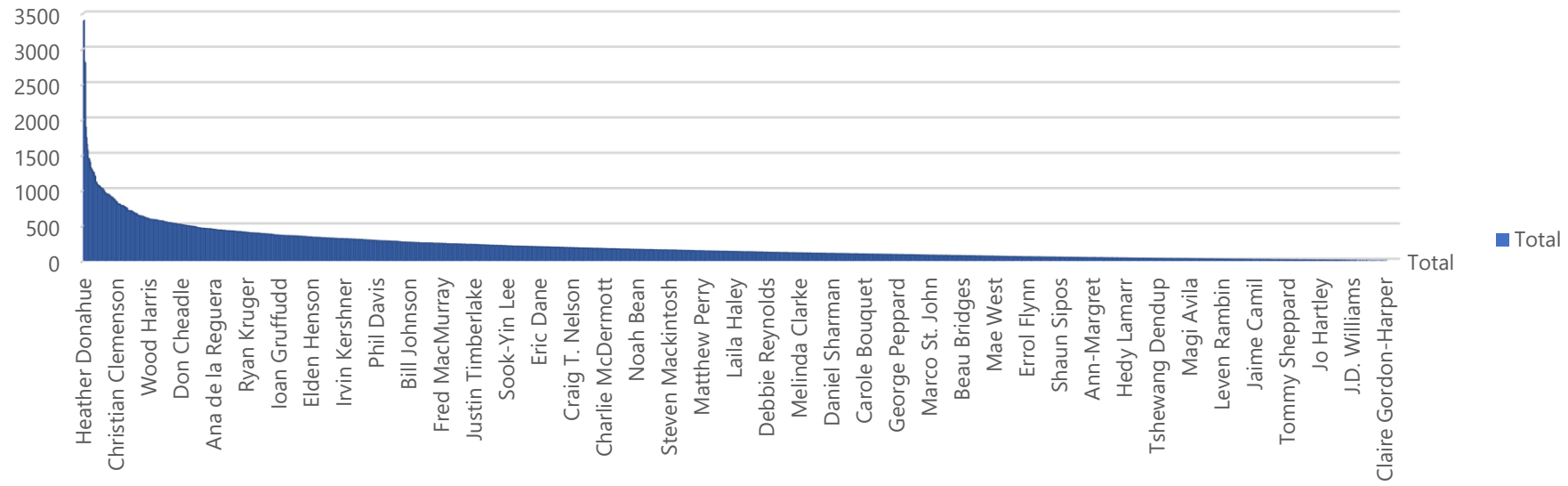
# AUDIENCE CHOICE BEST ACTOR



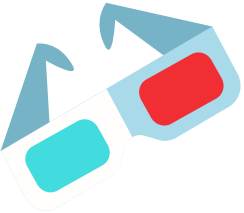
Row Labels	Average of num_user_for_reviews
[-] <b>Heather Donahue</b>	<b>3400</b>
The Blair Witch Project	3400
[-] <b>Christo Jivkov</b>	<b>2814</b>
The Passion of the Christ	2814
[-] <b>Steve Bastoni</b>	<b>2789</b>
The Matrix Reloaded	2789
[-] <b>Phaldut Sharma</b>	<b>1885</b>
Gravity	1885
[-] <b>Keir Dullea</b>	<b>1736</b>
2001: A Space Odyssey	1736
[-] <b>Chen Chang</b>	<b>1641</b>
Crouching Tiger, Hidden Dragon	1641
[-] <b>Nick Stahl</b>	<b>1562</b>
Terminator 3: Rise of the Machines	1676
The Thin Red Line	1448
[-] <b>Kevin Rankin</b>	<b>1445</b>
Hulk	1445
[-] <b>Noah Huntley</b>	<b>1441</b>
28 Days Later...	1441
[-] <b>Osama bin Laden</b>	<b>1416</b>
Fahrenheit 9/11	1416
[-] <b>Seychelle Gabriel</b>	<b>1382</b>
The Last Airbender	1382
[-] <b>Mathieu Kassovitz</b>	<b>1314</b>
Amélie	1314



## Mean num\_user\_for\_reviews



# CRITICS AND AUDIENCE FAVOURITE ACTORS



The award for the critic and audience's favourite actors is found by forming the pivot table of the given dataset and then finding the average of num\_users\_for\_review and average of num\_critic\_for\_review hence finding the actors name for the same



**Heather Donahue**

Heather Donahue gets the best actor award audience favourite for the movie title The Blair Witch Project.



**Phaldut Sharma**

Phaldut Sharma gets the best actor critics for Gravity





# BAR CHART ANALYSIS



The bar chart given in the next slide shows the number of users who voted in the given decade. It can be observed from the graph that the maximum number of user voted in the decade 2000s followed 2010s



**0.17B**

**2000s**

In the decade  
2000s nearly  
176260317  
people voted

**0.12B**

**2010s**

In the decade  
2010s nearly  
124384369  
People voted

**0.07B**

**1990s**

In the decade  
1990s nearly  
70278714  
People voted

**0.02B**

**1980s**

In the decade  
the 1980s  
nearly  
23021112  
people voted



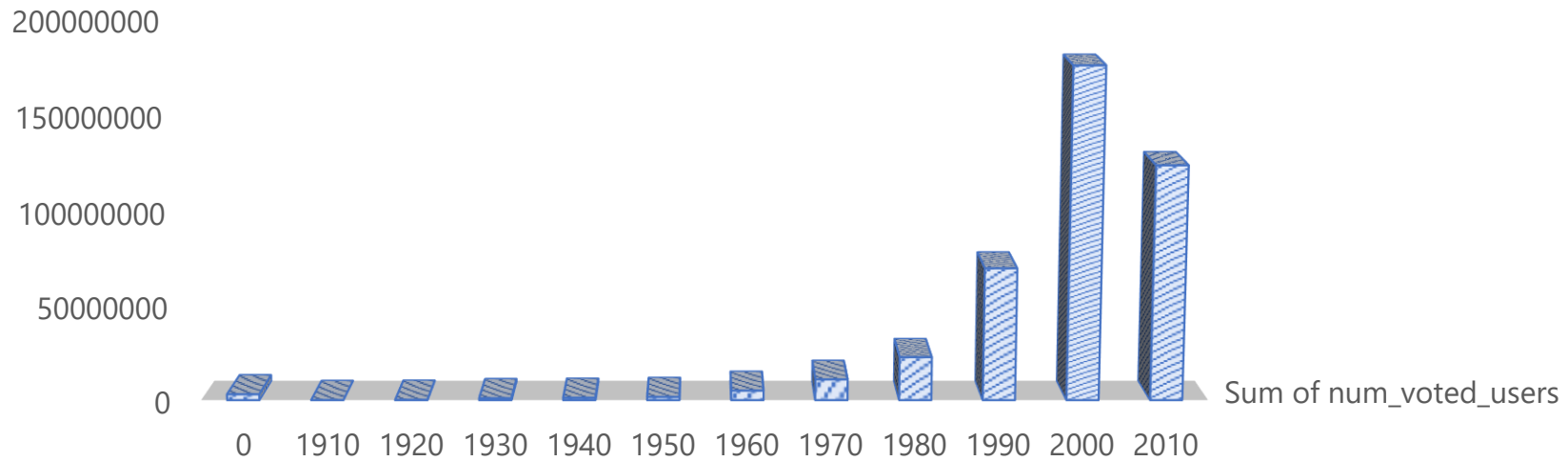


Row Labels ▾	Sum of num_voted_users
0	3127916
1910	10718
1920	128672
1930	984397
1940	1211888
1950	1638504
1960	5065074
1970	11072372
1980	23021112
1990	70278714
2000	176260317
2010	124384369
<b>Grand Total</b>	<b>417184053</b>



# SUM OF NUM\_VOTED\_USERS

Sum of num\_voted\_users



# RESULT

Through this project, we understood the importance of data analytics and its use in our daily life. We also understood how to use excel more efficiently. How to solve problems using data analytics



# THANK YOU

