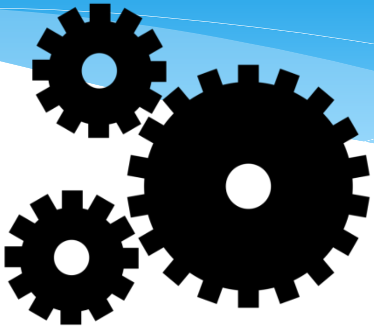
IMDB MOVIE ANALYSIS

Project By
Sumit Gope

Project Description

IMDb is one of the most popular online databases for movies, television shows, and video games. In this project, we will analyze a dataset of IMDb movies to gain insights and answer some interesting questions about the movie industry. The dataset contains various features of movies such as their title_movie , genres, rating, gross, and more.

Approach



Data Cleaning: The first step is to clean the data, which involves handling missing values, removing duplicate records, and converting data types.

Exploratory Data Analysis (EDA): In this step, we will explore the data to understand the relationships and patterns between different features. We can use various data visualization techniques such as Bar chart, pivot table, and many more.

Feature Engineering: We can create new features from the existing ones to get more meaningful insights. For example, we can calculate the profit made by a movie by subtracting the budget from the gross revenue.

Answering Questions: After completing the EDA and feature engineering, we can answer some interesting questions such as:

1.Which genres are the most popular?
2.What is the correlation between the budget and the revenue?
3.Which actors and directors have the highest average rating?
4.Which movies have the highest profit?
5.How has the movie industry evolved over the years?

Conclusion: Finally, we will summarize the findings from the analysis and draw conclusions. We can also provide recommendations for movie makers to improve their chances of making a successful movie.

Tech-Stack Used

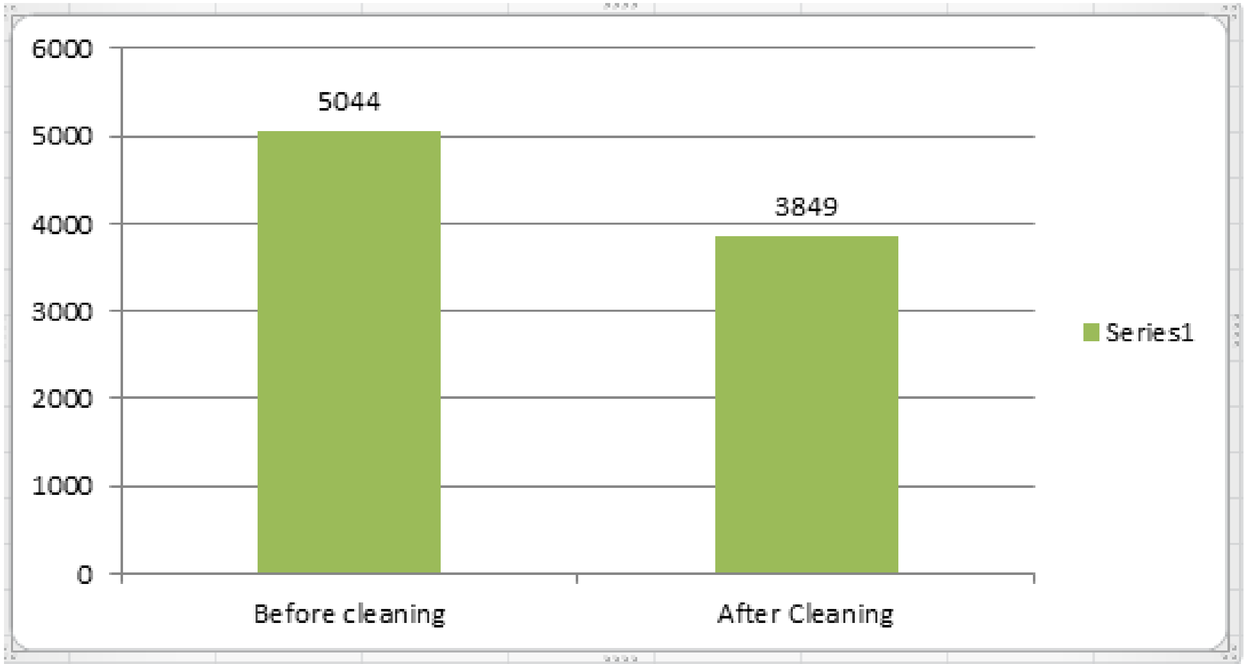
Insights

A. Your task: Clean the data

Solution: First I removed the duplicates value from the dataset from Data->Remove Duplicates using excel then removed the null values from the dataset using F5 then click on ‘special’ which direct you to ‘Go to special’ where we have to click blank and its shows all blank spots mark. To delete all the blanks spot’s row click ctrl+ - (minus) then click ‘entire row’.

After this I have done the remove unwanted columns which are not necessary for this project tasks.

After clean the data we’ve got this in output:

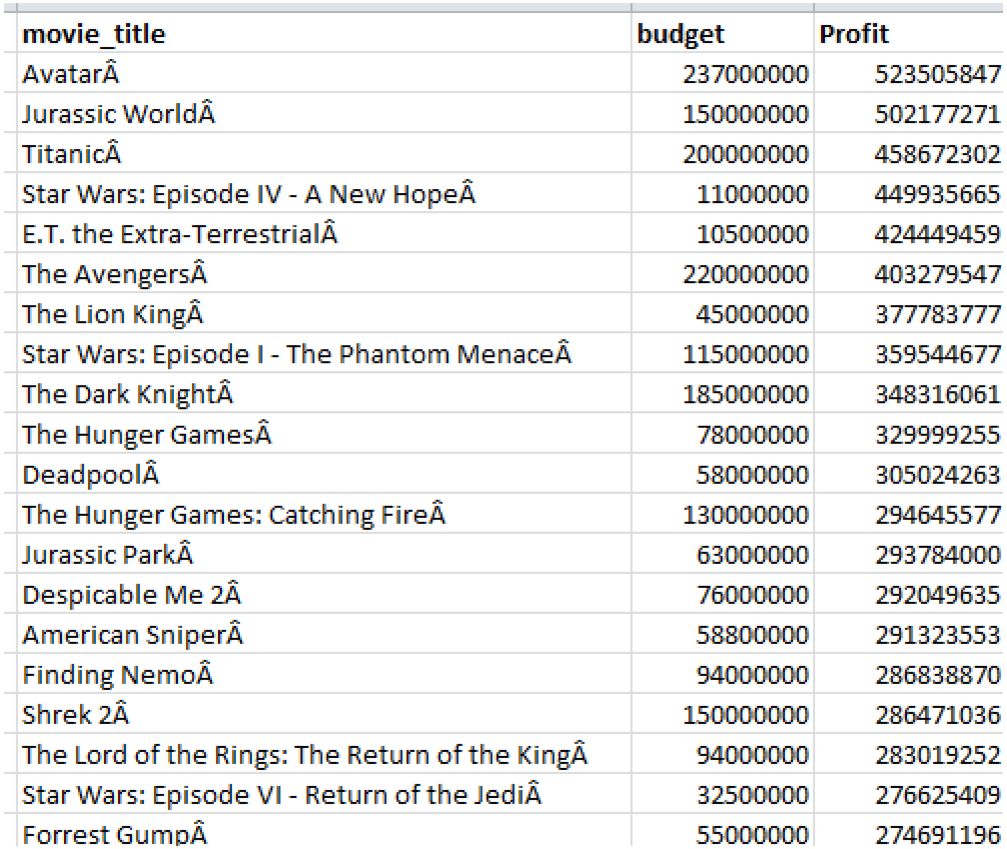


B. Your task: Find the movies with the highest profit?

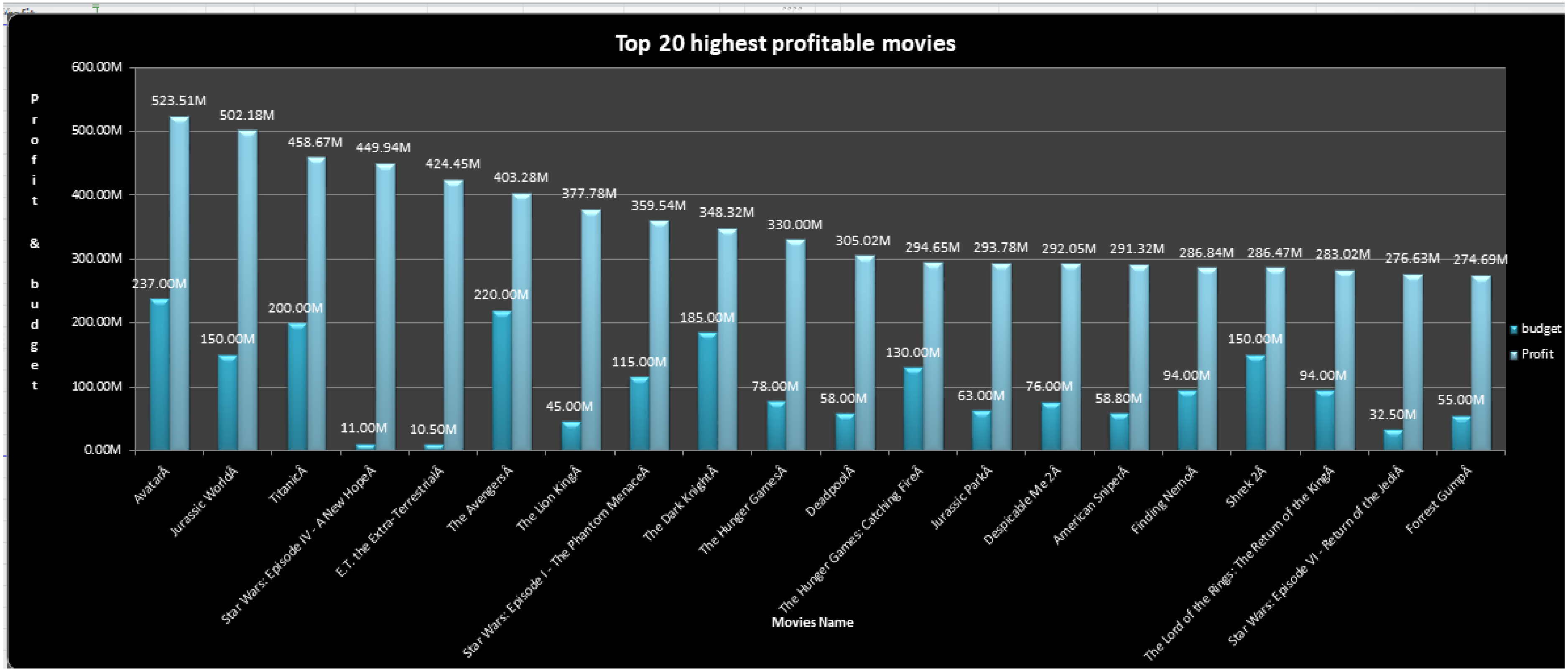
Solution:

I have created a new column called Profit and then used the formula " =gross cell-budget" . Then sort the column largest to smallest.

Output:



movie_title	budget	Profit
AvatarÂ	237000000	523505847
Jurassic WorldÂ	150000000	502177271
TitanicÂ	200000000	458672302
Star Wars: Episode IV - A New HopeÂ	11000000	449935665
E.T. the Extra-TerrestrialÂ	10500000	424449459
The AvengersÂ	220000000	403279547
The Lion KingÂ	45000000	377783777
Star Wars: Episode I - The Phantom MenaceÂ	115000000	359544677
The Dark KnightÂ	185000000	348316061
The Hunger GamesÂ	78000000	329999255
DeadpoolÂ	58000000	305024263
The Hunger Games: Catching FireÂ	130000000	294645577
Jurassic ParkÂ	63000000	293784000
Despicable Me 2Â	76000000	292049635
American SniperÂ	58800000	291323553
Finding NemoÂ	94000000	286838870
Shrek 2Â	150000000	286471036
The Lord of the Rings: The Return of the KingÂ	94000000	283019252
Star Wars: Episode VI - Return of the JediÂ	32500000	276625409
Forrest GumpÂ	55000000	274691196



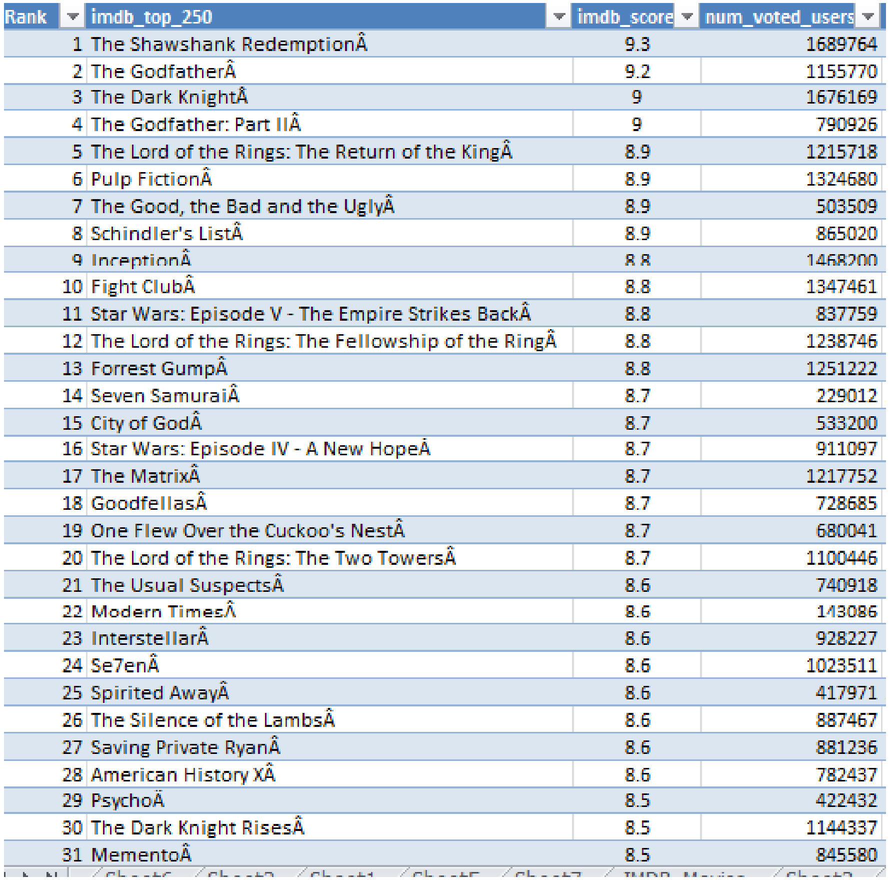
Top 20 highest profitable movies

C. Your task: Find IMDB Top 250

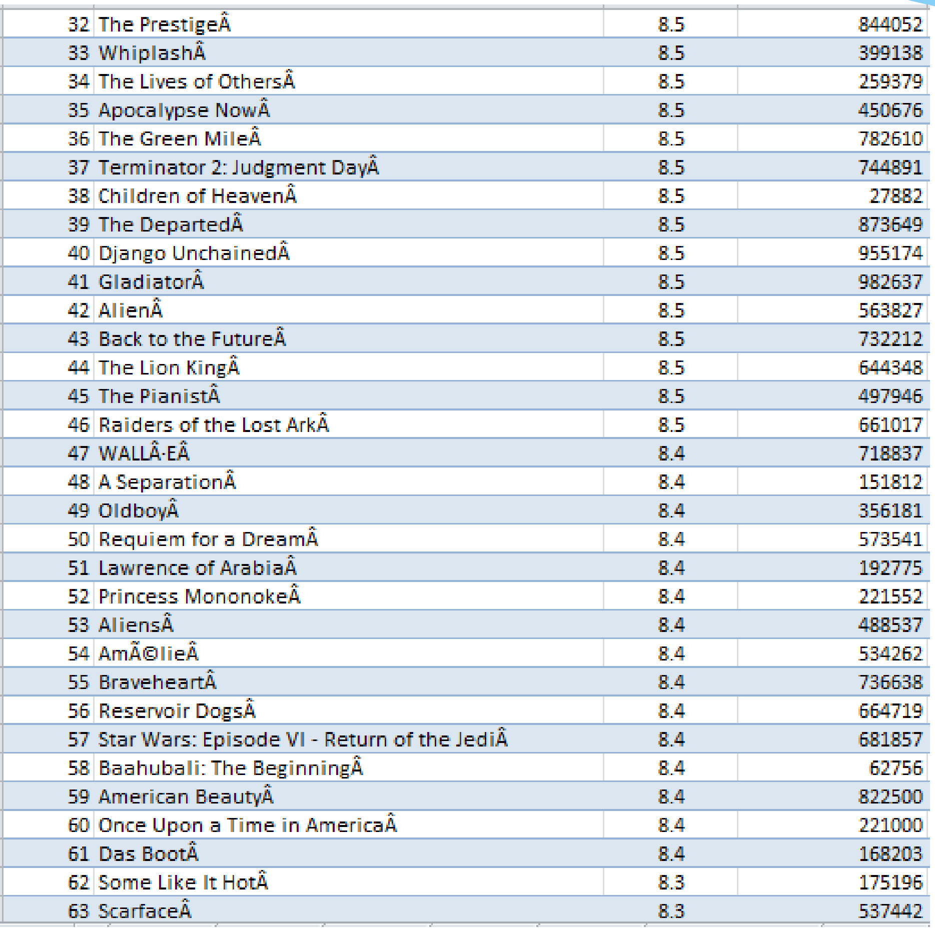
Solution: Created a new column called IMDB_Top_250 . Then sorted the Imdb_score largest to smallest and put the condition on num_voted_users value greater than and equal to 25000.

Then filter the language column by not deselecting english language. Which gave us non english movies which are actually top foreign language films. Extract on top_foreign_lang column.

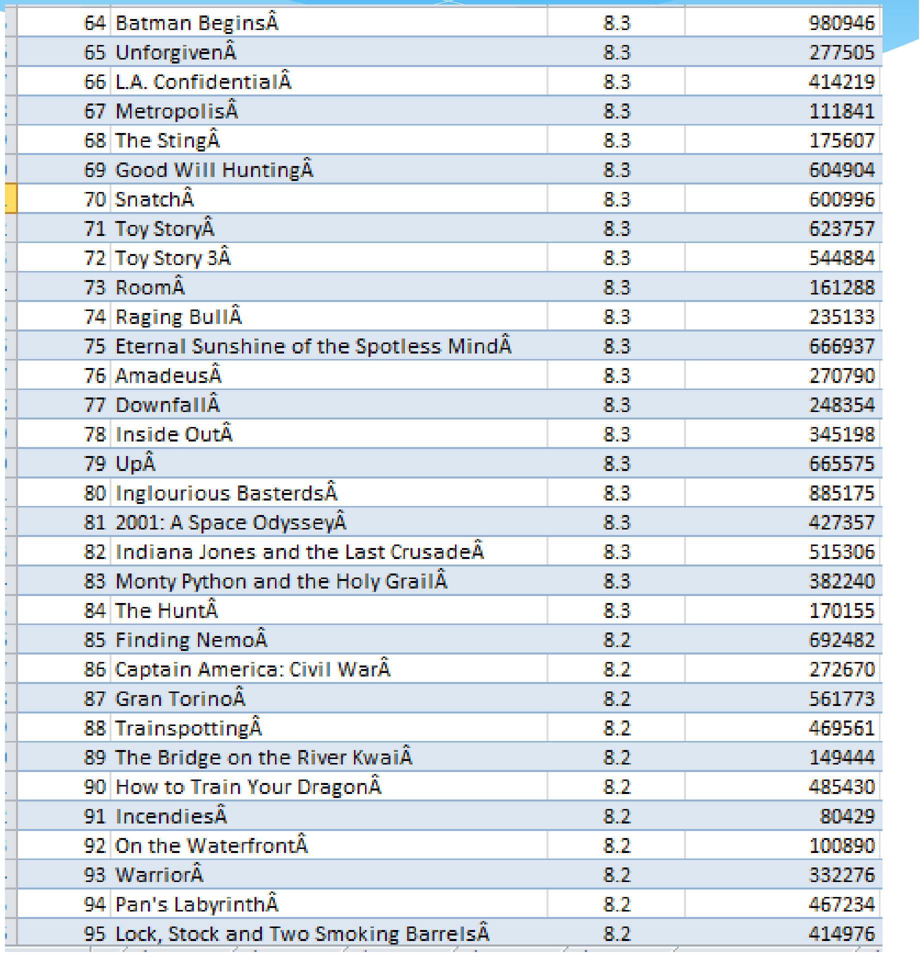
Output:



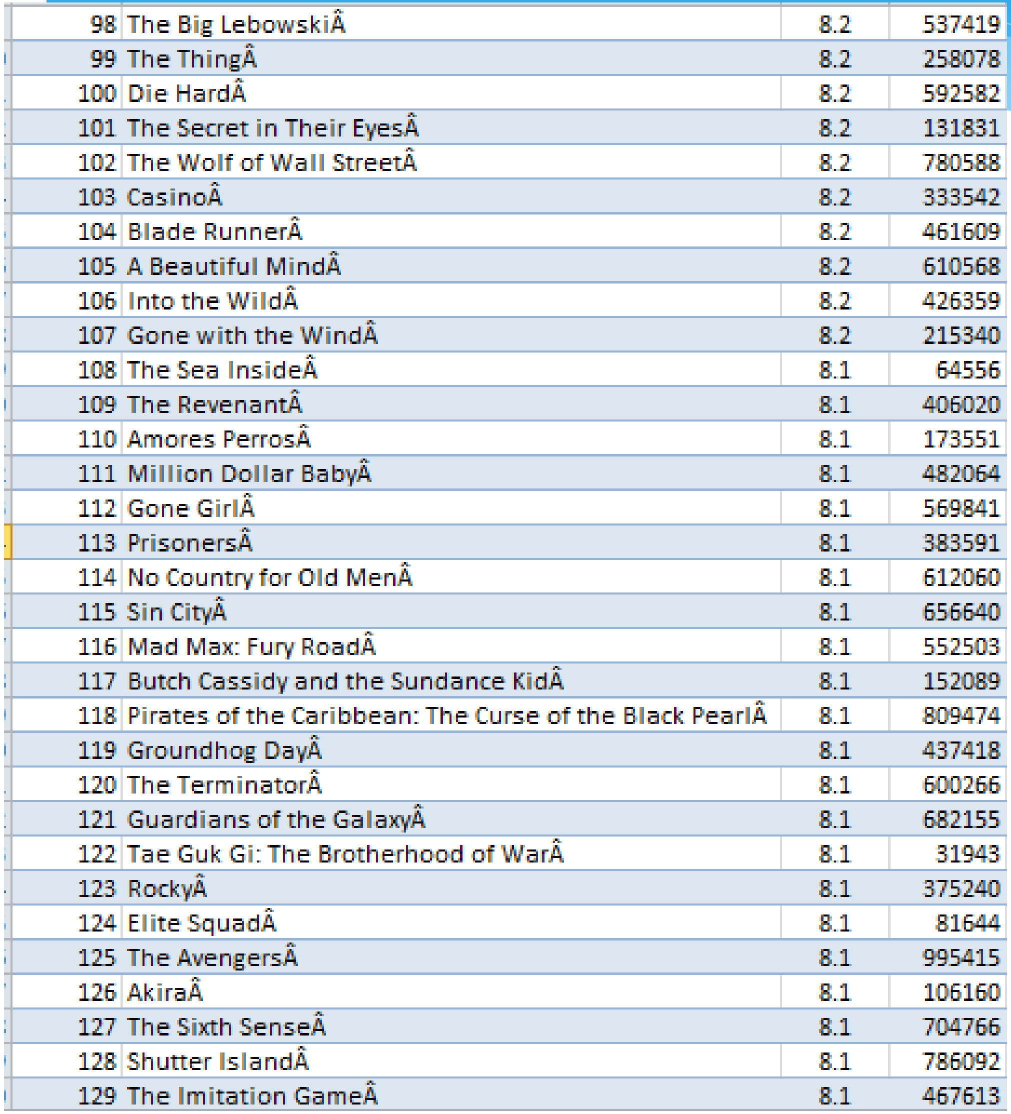
Rank	imdb_top_250	imdb_score	num_voted_users
1	The Shawshank RedemptionÂ	9.3	1689764
2	The GodfatherÂ	9.2	1155770
3	The Dark KnightÂ	9	1676169
4	The Godfather: Part IIÂ	9	790926
5	The Lord of the Rings: The Return of the KingÂ	8.9	1215718
6	Pulp FictionÂ	8.9	1324680
7	The Good, the Bad and the UglyÂ	8.9	503509
8	Schindler's ListÂ	8.9	865020
9	InceptionÂ	8.8	1468200
10	Fight ClubÂ	8.8	1347461
11	Star Wars: Episode V - The Empire Strikes BackÂ	8.8	837759
12	The Lord of the Rings: The Fellowship of the RingÂ	8.8	1238746
13	Forrest GumpÂ	8.8	1251222
14	Seven SamuraiÂ	8.7	229012
15	City of GodÂ	8.7	533200
16	Star Wars: Episode IV - A New HopeÂ	8.7	911097
17	The MatrixÂ	8.7	1217752
18	GoodfellasÂ	8.7	728685
19	One Flew Over the Cuckoo's NestÂ	8.7	680041
20	The Lord of the Rings: The Two TowersÂ	8.7	1100446
21	The Usual SuspectsÂ	8.6	740918
22	Modern TimesÂ	8.6	143086
23	InterstellarÂ	8.6	928227
24	Se7enÂ	8.6	1023511
25	Spirited AwayÂ	8.6	417971
26	The Silence of the LambsÂ	8.6	887467
27	Saving Private RyanÂ	8.6	881236
28	American History XÂ	8.6	782437
29	PsychoÂ	8.5	422432
30	The Dark Knight RisesÂ	8.5	1144337
31	MementoÂ	8.5	845580



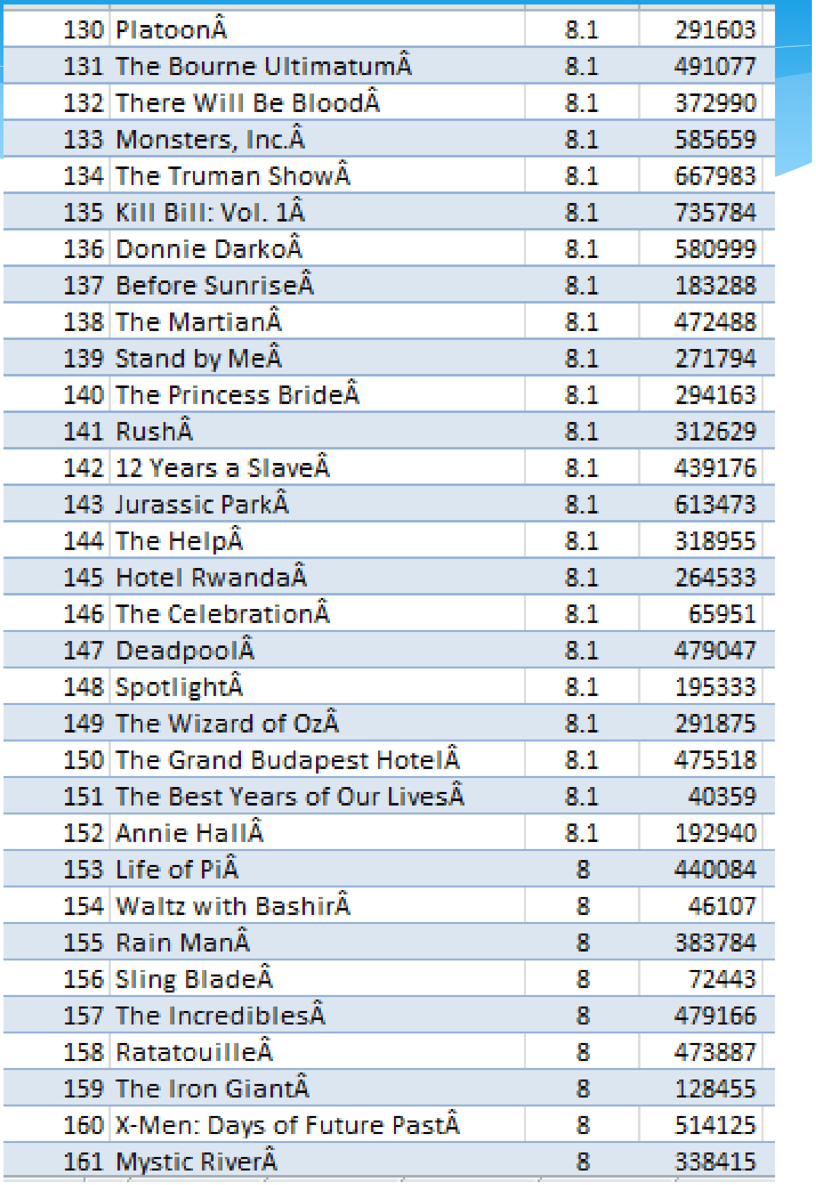
32	The PrestigeÂ	8.5	844052
33	WhiplashÂ	8.5	399138
34	The Lives of OthersÂ	8.5	259379
35	Apocalypse NowÂ	8.5	450676
36	The Green MileÂ	8.5	782610
37	Terminator 2: Judgment DayÂ	8.5	744891
38	Children of HeavenÂ	8.5	27882
39	The DepartedÂ	8.5	873649
40	Django UnchainedÂ	8.5	955174
41	GladiatorÂ	8.5	982637
42	AlienÂ	8.5	563827
43	Back to the FutureÂ	8.5	732212
44	The Lion KingÂ	8.5	644348
45	The PianistÂ	8.5	497946
46	Raiders of the Lost ArkÂ	8.5	661017
47	WALLÂ·EÂ	8.4	718837
48	A SeparationÂ	8.4	151812
49	OldboyÂ	8.4	356181
50	Requiem for a DreamÂ	8.4	573541
51	Lawrence of ArabiaÂ	8.4	192775
52	Princess MononokeÂ	8.4	221552
53	AliensÂ	8.4	488537
54	AmÃ©lieÂ	8.4	534262
55	BraveheartÂ	8.4	736638
56	Reservoir DogsÂ	8.4	664719
57	Star Wars: Episode VI - Return of the JediÂ	8.4	681857
58	Baahubali: The BeginningÂ	8.4	62756
59	American BeautyÂ	8.4	822500
60	Once Upon a Time in AmericaÂ	8.4	221000
61	Das BootÂ	8.4	168203
62	Some Like It HotÂ	8.3	175196
63	ScarfaceÂ	8.3	537442



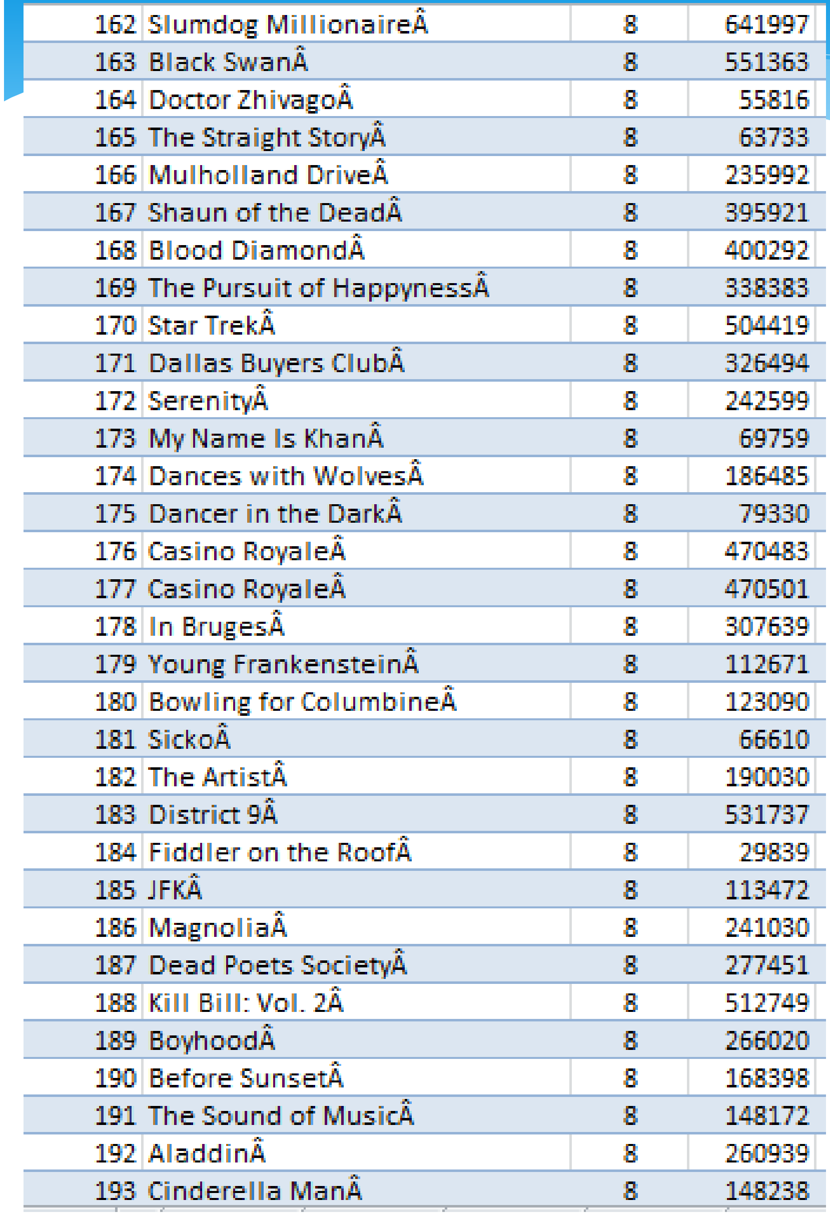
64	Batman BeginsÂ	8.3	980946
65	UnforgivenÂ	8.3	277505
66	L.A. ConfidentialÂ	8.3	414219
67	MetropolisÂ	8.3	111841
68	The StingÂ	8.3	175607
69	Good Will HuntingÂ	8.3	604904
70	SnatchÂ	8.3	600996
71	Toy StoryÂ	8.3	623757
72	Toy Story 3Â	8.3	544884
73	RoomÂ	8.3	161288
74	Raging BullÂ	8.3	235133
75	Eternal Sunshine of the Spotless MindÂ	8.3	666937
76	AmadeusÂ	8.3	270790
77	DownfallÂ	8.3	248354
78	Inside OutÂ	8.3	345198
79	UpÂ	8.3	665575
80	Inglourious BasterdsÂ	8.3	885175
81	2001: A Space OdysseyÂ	8.3	427357
82	Indiana Jones and the Last CrusadeÂ	8.3	515306
83	Monty Python and the Holy GrailÂ	8.3	382240
84	The HuntÂ	8.3	170155
85	Finding NemoÂ	8.2	692482
86	Captain America: Civil WarÂ	8.2	272670
87	Gran TorinoÂ	8.2	561773
88	TrainspottingÂ	8.2	469561
89	The Bridge on the River KwaiÂ	8.2	149444
90	How to Train Your DragonÂ	8.2	485430
91	IncendiesÂ	8.2	80429
92	On the WaterfrontÂ	8.2	100890
93	WarriorÂ	8.2	332276
94	Pan's LabyrinthÂ	8.2	467234
95	Lock, Stock and Two Smoking BarrelsÂ	8.2	414976



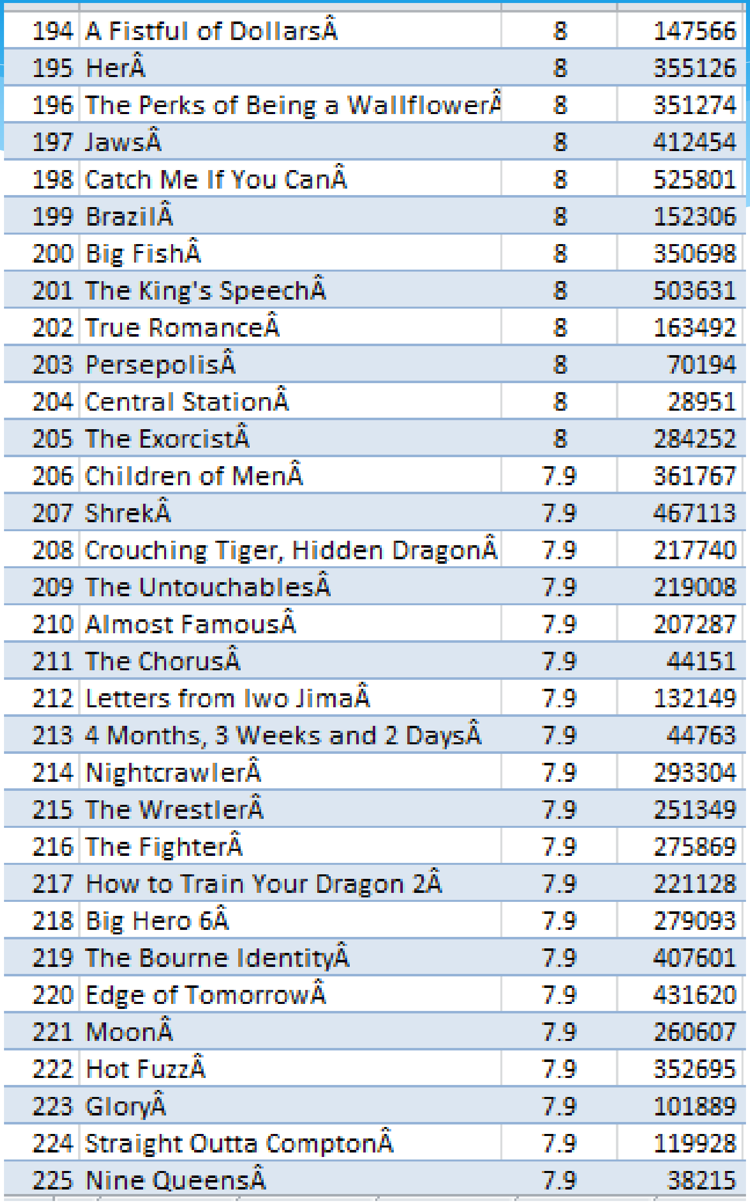
98	The Big LebowskiÂ	8.2	537419
99	The ThingÂ	8.2	258078
100	Die HardÂ	8.2	592582
101	The Secret in Their EyesÂ	8.2	131831
102	The Wolf of Wall StreetÂ	8.2	780588
103	CasinoÂ	8.2	333542
104	Blade RunnerÂ	8.2	461609
105	A Beautiful MindÂ	8.2	610568
106	Into the WildÂ	8.2	426359
107	Gone with the WindÂ	8.2	215340
108	The Sea InsideÂ	8.1	64556
109	The RevenantÂ	8.1	406020
110	Amores PerrosÂ	8.1	173551
111	Million Dollar BabyÂ	8.1	482064
112	Gone GirlÂ	8.1	569841
113	PrisonersÂ	8.1	383591
114	No Country for Old MenÂ	8.1	612060
115	Sin CityÂ	8.1	656640
116	Mad Max: Fury RoadÂ	8.1	552503
117	Butch Cassidy and the Sundance KidÂ	8.1	152089
118	Pirates of the Caribbean: The Curse of the Black PearlÂ	8.1	809474
119	Groundhog DayÂ	8.1	437418
120	The TerminatorÂ	8.1	600266
121	Guardians of the GalaxyÂ	8.1	682155
122	Tae Guk Gi: The Brotherhood of WarÂ	8.1	31943
123	RockyÂ	8.1	375240
124	Elite SquadÂ	8.1	81644
125	The AvengersÂ	8.1	995415
126	AkiraÂ	8.1	106160
127	The Sixth SenseÂ	8.1	704766
128	Shutter IslandÂ	8.1	786092
129	The Imitation GameÂ	8.1	467613



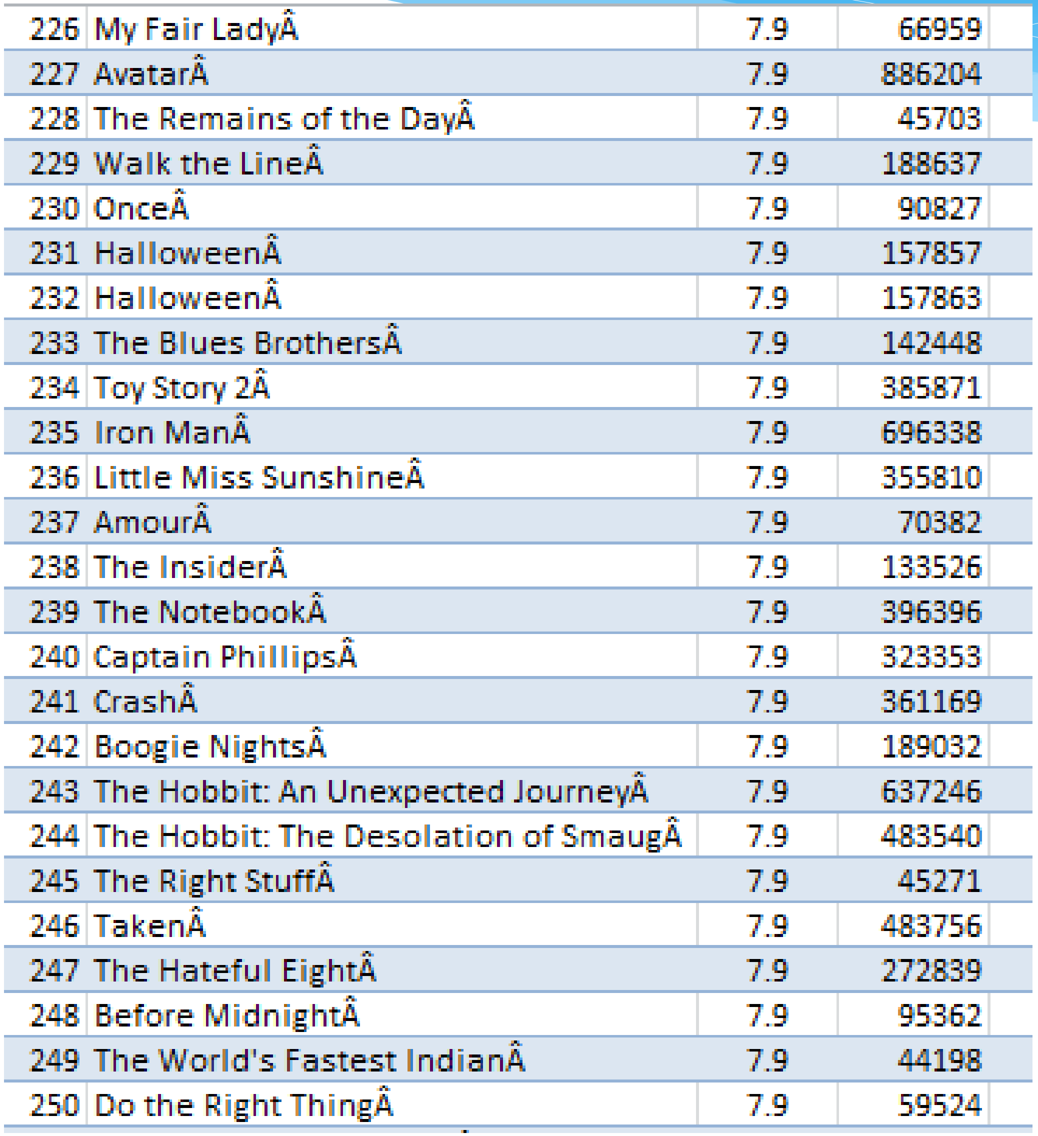
130	PlatoonÂ	8.1	291603
131	The Bourne UltimatumÂ	8.1	491077
132	There Will Be BloodÂ	8.1	372990
133	Monsters, Inc.Â	8.1	585659
134	The Truman ShowÂ	8.1	667983
135	Kill Bill: Vol. 1Â	8.1	735784
136	Donnie DarkoÂ	8.1	580999
137	Before SunriseÂ	8.1	183288
138	The MartianÂ	8.1	472488
139	Stand by MeÂ	8.1	271794
140	The Princess BrideÂ	8.1	294163
141	RushÂ	8.1	312629
142	12 Years a SlaveÂ	8.1	439176
143	Jurassic ParkÂ	8.1	613473
144	The HelpÂ	8.1	318955
145	Hotel RwandaÂ	8.1	264533
146	The CelebrationÂ	8.1	65951
147	DeadpoolÂ	8.1	479047
148	SpotlightÂ	8.1	195333
149	The Wizard of OzÂ	8.1	291875
150	The Grand Budapest HotelÂ	8.1	475518
151	The Best Years of Our LivesÂ	8.1	40359
152	Annie HallÂ	8.1	192940
153	Life of PiÂ	8	440084
154	Waltz with BashirÂ	8	46107
155	Rain ManÂ	8	383784
156	Sling BladeÂ	8	72443
157	The IncrediblesÂ	8	479166
158	RatatouilleÂ	8	473887
159	The Iron GiantÂ	8	128455
160	X-Men: Days of Future PastÂ	8	514125
161	Mystic RiverÂ	8	338415



162	Slumdog MillionaireÂ	8	641997
163	Black SwanÂ	8	551363
164	Doctor ZhivagoÂ	8	55816
165	The Straight StoryÂ	8	63733
166	Mulholland DriveÂ	8	235992
167	Shaun of the DeadÂ	8	395921
168	Blood DiamondÂ	8	400292
169	The Pursuit of HappynessÂ	8	338383
170	Star TrekÂ	8	504419
171	Dallas Buyers ClubÂ	8	326494
172	SerenityÂ	8	242599
173	My Name Is KhanÂ	8	69759
174	Dances with WolvesÂ	8	186485
175	Dancer in the DarkÂ	8	79330
176	Casino RoyaleÂ	8	470483
177	Casino RoyaleÂ	8	470501
178	In BrugesÂ	8	307639
179	Young FrankensteinÂ	8	112671
180	Bowling for ColumbineÂ	8	123090
181	SickoÂ	8	66610
182	The ArtistÂ	8	190030
183	District 9Â	8	531737
184	Fiddler on the RoofÂ	8	29839
185	JFKÂ	8	113472
186	MagnoliaÂ	8	241030
187	Dead Poets SocietyÂ	8	277451
188	Kill Bill: Vol. 2Â	8	512749
189	BoyhoodÂ	8	266020
190	Before SunsetÂ	8	168398
191	The Sound of MusicÂ	8	148172
192	AladdinÂ	8	260939
193	Cinderella ManÂ	8	148238

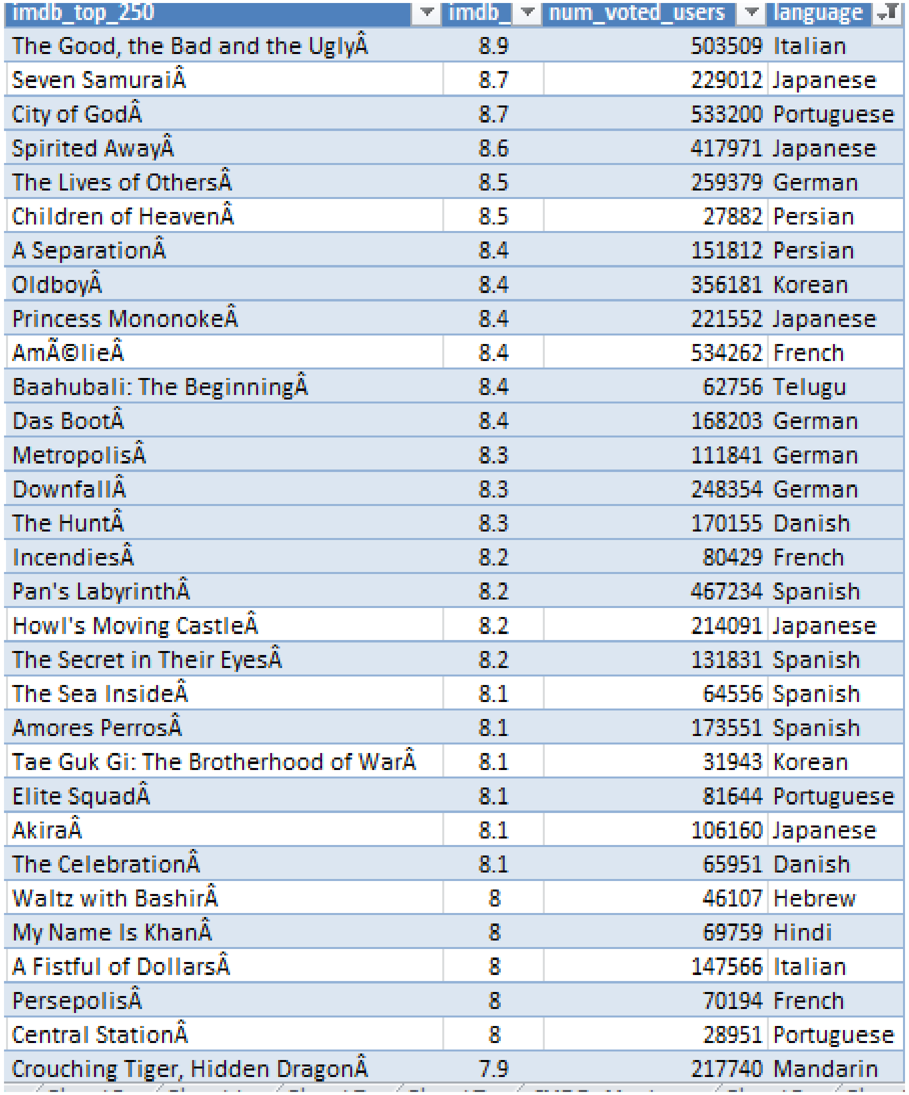
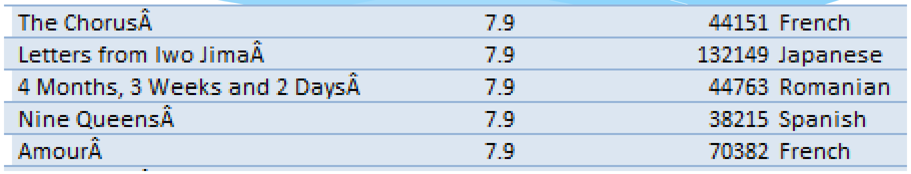


194	A Fistful of DollarsÂ	8	147566
195	HerÂ	8	355126
196	The Perks of Being a WallflowerÂ	8	351274
197	JawsÂ	8	412454
198	Catch Me If You CanÂ	8	525801
199	BrazilÂ	8	152306
200	Big FishÂ	8	350698
201	The King's SpeechÂ	8	503631
202	True RomanceÂ	8	163492
203	PersepolisÂ	8	70194
204	Central StationÂ	8	28951
205	The ExorcistÂ	8	284252
206	Children of MenÂ	7.9	361767
207	ShrekÂ	7.9	467113
208	Crouching Tiger, Hidden DragonÂ	7.9	217740
209	The UntouchablesÂ	7.9	219008
210	Almost FamousÂ	7.9	207287
211	The ChorusÂ	7.9	44151
212	Letters from Iwo JimaÂ	7.9	132149
213	4 Months, 3 Weeks and 2 DaysÂ	7.9	44763
214	NightcrawlerÂ	7.9	293304
215	The WrestlerÂ	7.9	251349
216	The FighterÂ	7.9	275869
217	How to Train Your Dragon 2Â	7.9	221128
218	Big Hero 6Â	7.9	279093
219	The Bourne IdentityÂ	7.9	407601
220	Edge of TomorrowÂ	7.9	431620
221	MoonÂ	7.9	260607
222	Hot FuzzÂ	7.9	352695
223	GloryÂ	7.9	101889
224	Straight Outta ComptonÂ	7.9	119928
225	Nine QueensÂ	7.9	38215



226	My Fair LadyÂ	7.9	66959
227	AvatarÂ	7.9	886204
228	The Remains of the DayÂ	7.9	45703
229	Walk the LineÂ	7.9	188637
230	OnceÂ	7.9	90827
231	HalloweenÂ	7.9	157857
232	HalloweenÂ	7.9	157863
233	The Blues BrothersÂ	7.9	142448
234	Toy Story 2Â	7.9	385871
235	Iron ManÂ	7.9	696338
236	Little Miss SunshineÂ	7.9	355810
237	AmourÂ	7.9	70382
238	The InsiderÂ	7.9	133526
239	The NotebookÂ	7.9	396396
240	Captain PhillipsÂ	7.9	323353
241	CrashÂ	7.9	361169
242	Boogie NightsÂ	7.9	189032
243	The Hobbit: An Unexpected JourneyÂ	7.9	637246
244	The Hobbit: The Desolation of SmaugÂ	7.9	483540
245	The Right StuffÂ	7.9	45271
246	TakenÂ	7.9	483756
247	The Hateful EightÂ	7.9	272839
248	Before MidnightÂ	7.9	95362
249	The World's Fastest IndianÂ	7.9	44198
250	Do the Right ThingÂ	7.9	59524

Top Foreign Language Movies

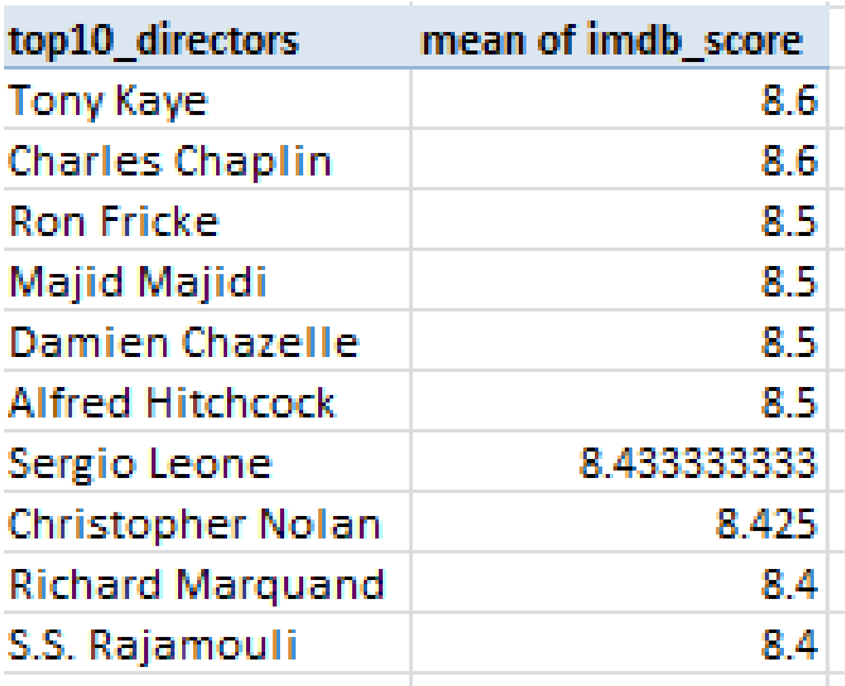



imdb_top_250	imdb_	num_voted_users	language
The Good, the Bad and the UglyÂ	8.9	503509	Italian
Seven SamuraiÂ	8.7	229012	Japanese
City of GodÂ	8.7	533200	Portuguese
Spirited AwayÂ	8.6	417971	Japanese
The Lives of OthersÂ	8.5	259379	German
Children of HeavenÂ	8.5	27882	Persian
A SeparationÂ	8.4	151812	Persian
OldboyÂ	8.4	356181	Korean
Princess MononokeÂ	8.4	221552	Japanese
AmÃ©lieÂ	8.4	534262	French
Baahubali: The BeginningÂ	8.4	62756	Telugu
Das BootÂ	8.4	168203	German
MetropolisÂ	8.3	111841	German
DownfallÂ	8.3	248354	German
The HuntÂ	8.3	170155	Danish
IncendiesÂ	8.2	80429	French
Pan's LabyrinthÂ	8.2	467234	Spanish
Howl's Moving CastleÂ	8.2	214091	Japanese
The Secret in Their EyesÂ	8.2	131831	Spanish
The Sea InsideÂ	8.1	64556	Spanish
Amores PerrosÂ	8.1	173551	Spanish
Tae Guk Gi: The Brotherhood of WarÂ	8.1	31943	Korean
Elite SquadÂ	8.1	81644	Portuguese
AkiraÂ	8.1	106160	Japanese
The CelebrationÂ	8.1	65951	Danish
Waltz with BashirÂ	8	46107	Hebrew
My Name Is KhanÂ	8	69759	Hindi
A Fistful of DollarsÂ	8	147566	Italian
PersepolisÂ	8	70194	French
Central StationÂ	8	28951	Portuguese
Crouching Tiger, Hidden DragonÂ	7.9	217740	Mandarin
The ChorusÂ	7.9	44151	French
Letters from Iwo JimaÂ	7.9	132149	Japanese
4 Months, 3 Weeks and 2 DaysÂ	7.9	44763	Romanian
Nine QueensÂ	7.9	38215	Spanish
AmourÂ	7.9	70382	French

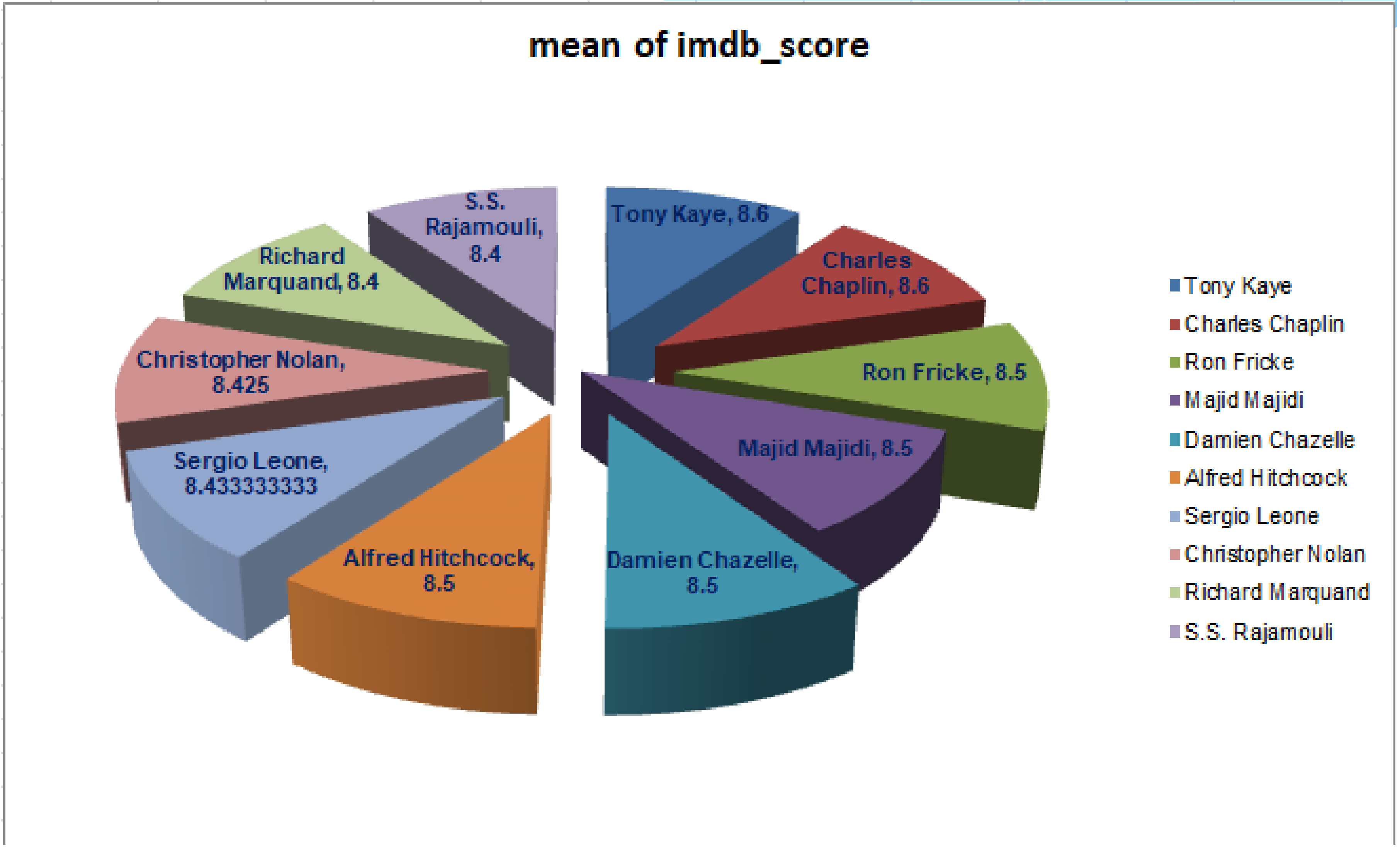
D.Your task: Find the best directors

Solution: Clicked on Pivot table its automatically selected the range of table after clicking ok. I have drag the director column to the row Label and imdb_score to the values. Rename the row label column to top10_directors and got the mean of imdb_score of every director and then sort the values column largest to smallest. Selected the top 10 director and transformed into chart.

Output:



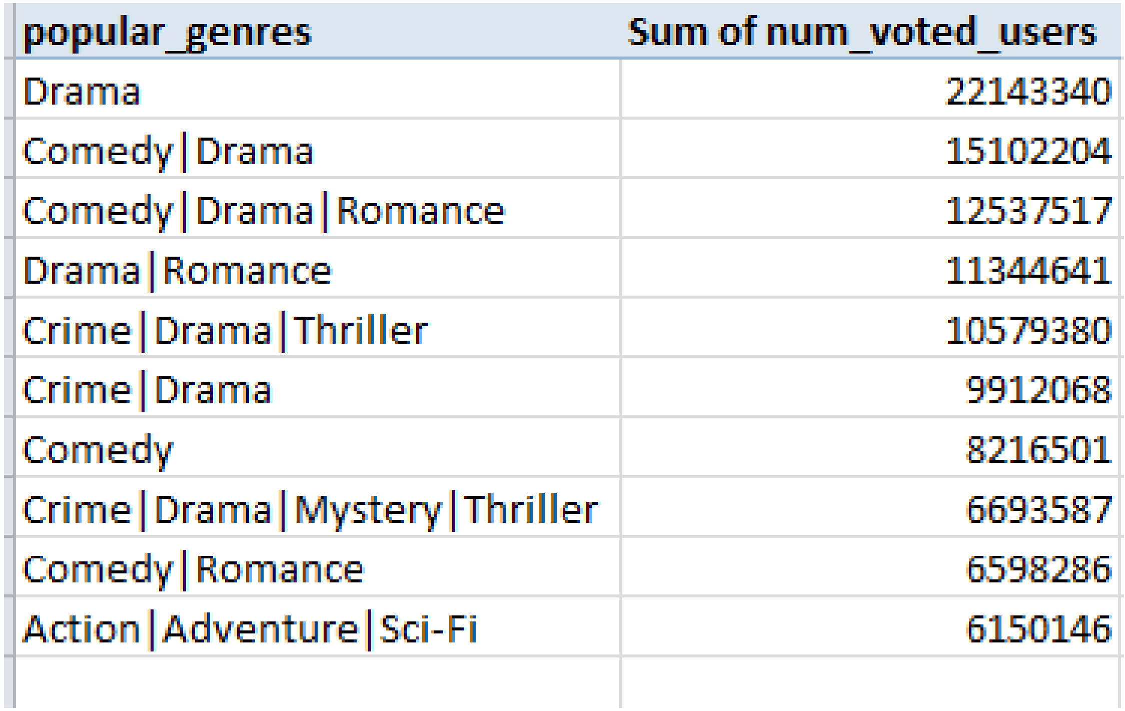
top10_directors	mean of imdb_score
Tony Kaye	8.6
Charles Chaplin	8.6
Ron Fricke	8.5
Majid Majidi	8.5
Damien Chazelle	8.5
Alfred Hitchcock	8.5
Sergio Leone	8.433333333
Christopher Nolan	8.425
Richard Marquand	8.4
S.S. Rajamouli	8.4


mean of imdb_score
S.S. Rajamouli, 8.4
Tony Kaye, 8.6
Richard Marquand, 8.4
Charles Chaplin, 8.6
Christopher Nolan, 8.425
Ron Fricke, 8.5
Majid Majidi, 8.5
Sergio Leone, 8.433333333
Alfred Hitchcock, 8.5
Damien Chazelle, 8.5
Tony Kaye
Charles Chaplin
Ron Fricke
Majid Majidi
Damien Chazelle
Alfred Hitchcock
Sergio Leone
Christopher Nolan
Richard Marquand
S.S. Rajamouli

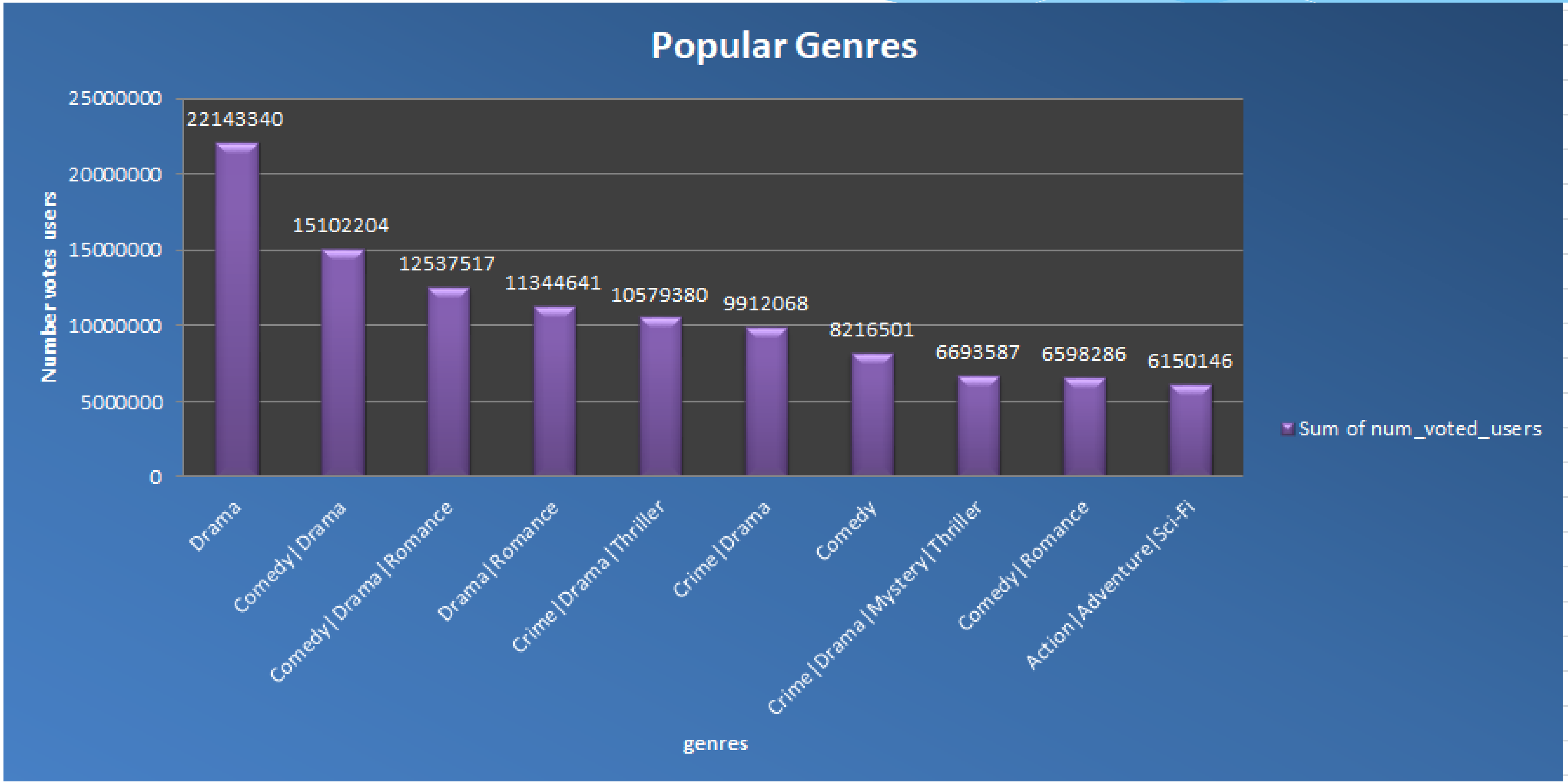
E. Your task: Find popular genres.

Solution: I have done the same thing just like task D. select the pivot table and drag genres column to row label and num_voted_user to the values field and then sort the sum of num_voted_user.

Output:



popular_genres	Sum of num_voted_users
Drama	22143340
Comedy|Drama	15102204
Comedy|Drama|Romance	12537517
Drama|Romance	11344641
Crime|Drama|Thriller	10579380
Crime|Drama	9912068
Comedy	8216501
Crime|Drama|Mystery|Thriller	6693587
Comedy|Romance	6598286
Action|Adventure|Sci-Fi	6150146


Popular Genres
Number votes users
25000000
20000000
15000000
10000000
5000000
0
22143340
15102204
12537517
11344641
10579380
9912068
8216501
6693587
6598286
6150146
Drama
Comedy|Drama
Comedy|Drama|Romance
Drama|Romance
Crime|Drama|Thriller
Crime|Drama
Comedy
Crime|Drama|Mystery|Thriller
Comedy|Romance
Action|Adventure|Sci-Fi
genres
Sum of num_voted_users

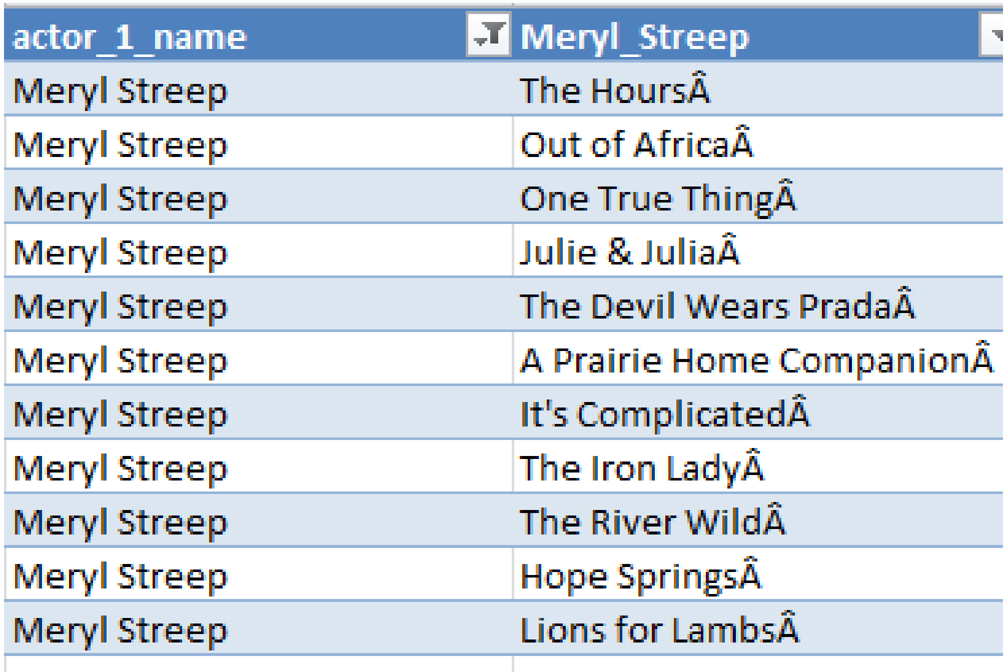
F. Your task: Find the critic-favorite and audience-favorite actors

Solution: Created three column named Meryl_Streep, Leo_Caprio, and Brad_Pitt and then used formula “=IF([@[actor_1_name]]= "Meryl Streep",[@[movie_title]],"") and for leo_Caprio and Brad_pitt same as well.

Created a “Combine” column and then use formula “ =(first_cell)&””&(second_cell)&””&(third_cell)

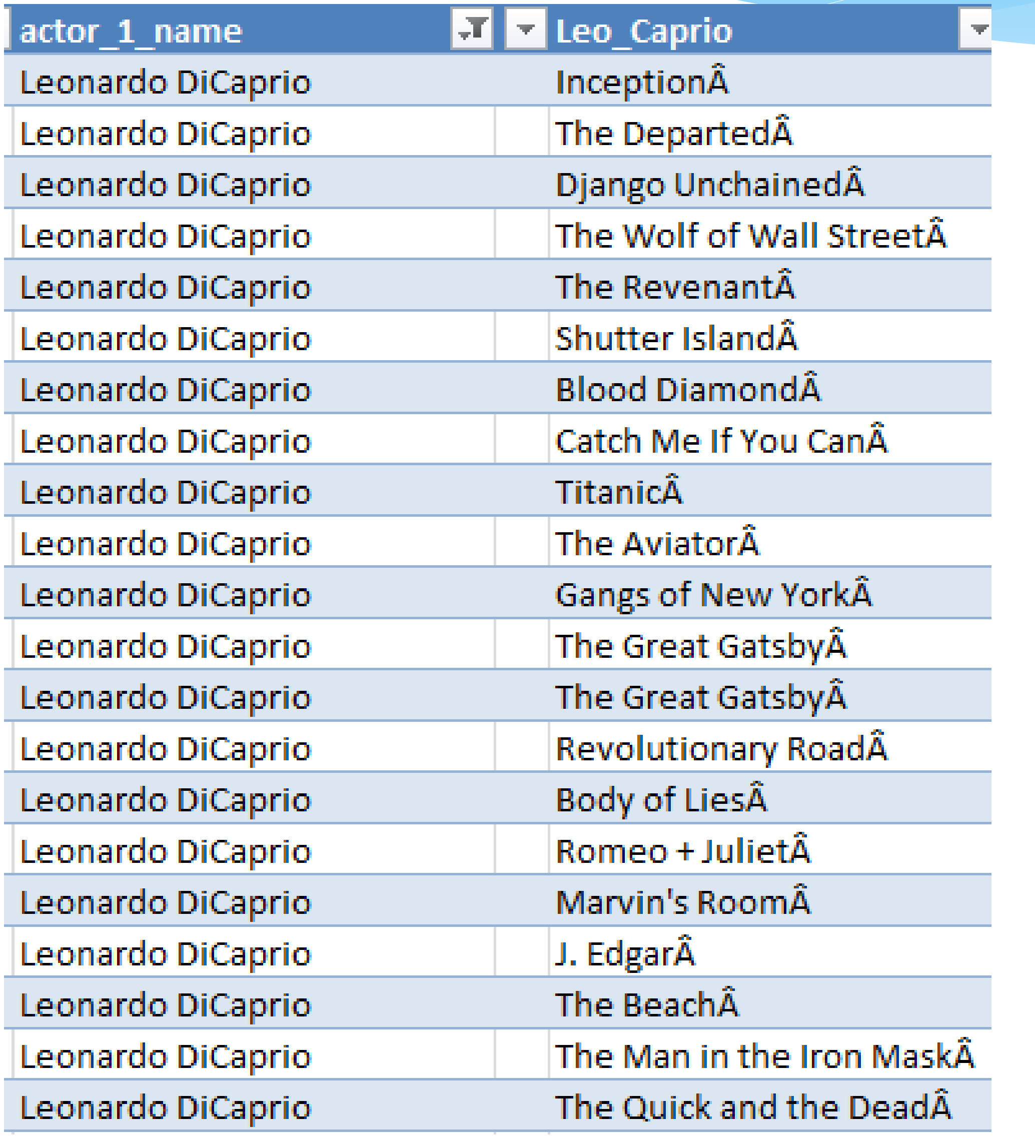
Output:

actor_1_name= Meryl Streep



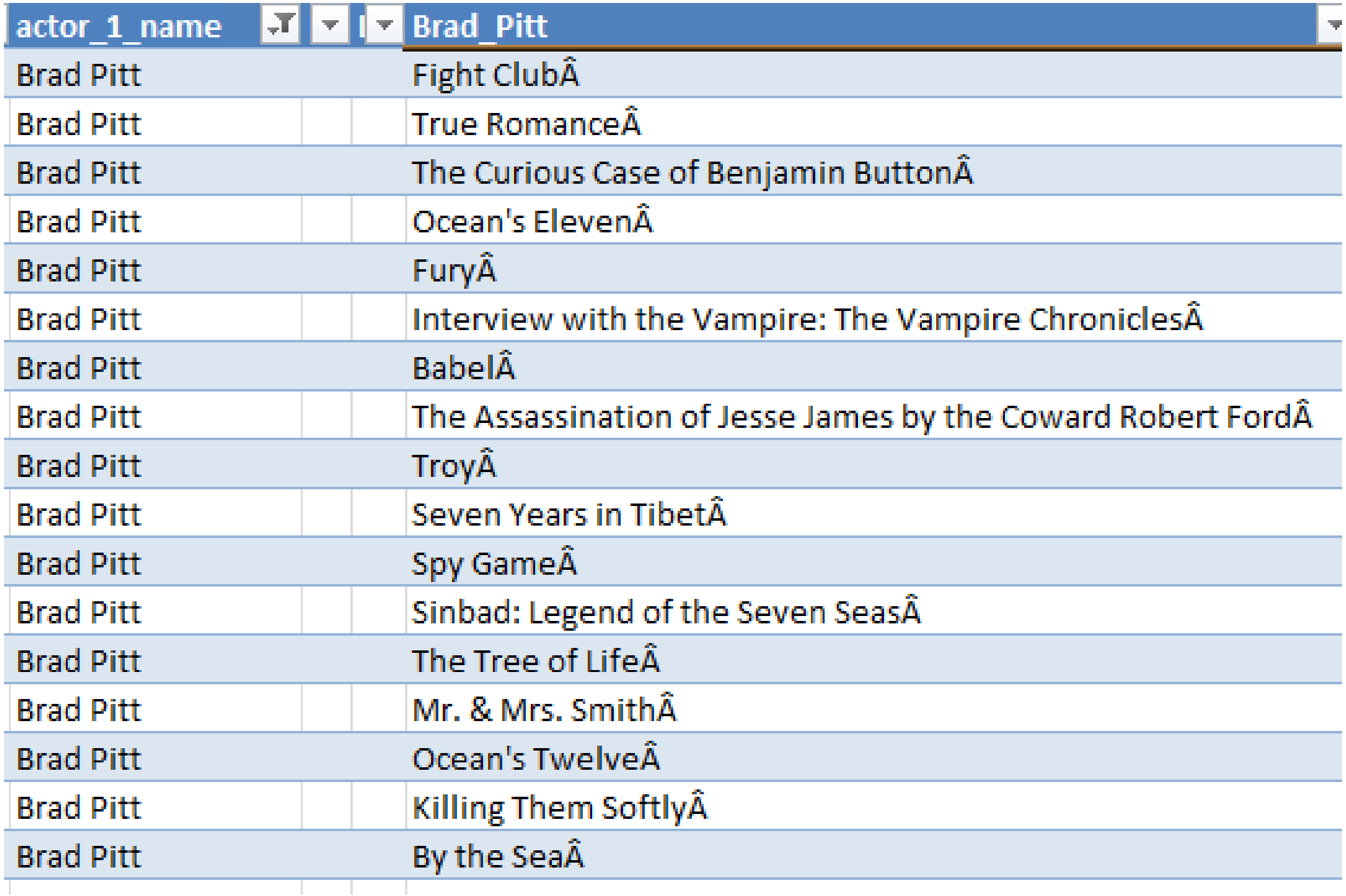
actor_1_name	Meryl_Streep
Meryl Streep	The HoursÂ
Meryl Streep	Out of AfricaÂ
Meryl Streep	One True ThingÂ
Meryl Streep	Julie & JuliaÂ
Meryl Streep	The Devil Wears PradaÂ
Meryl Streep	A Prairie Home CompanionÂ
Meryl Streep	It's ComplicatedÂ
Meryl Streep	The Iron LadyÂ
Meryl Streep	The River WildÂ
Meryl Streep	Hope SpringsÂ
Meryl Streep	Lions for LambsÂ

actor_1_name= Leonardo DiCaprio



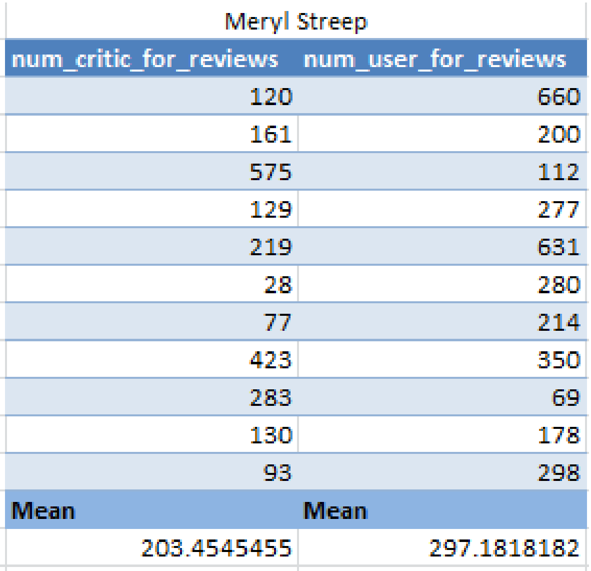
actor_1_name	Leo_Caprio
Leonardo DiCaprio	InceptionÂ
Leonardo DiCaprio	The DepartedÂ
Leonardo DiCaprio	Django UnchainedÂ
Leonardo DiCaprio	The Wolf of Wall StreetÂ
Leonardo DiCaprio	The RevenantÂ
Leonardo DiCaprio	Shutter IslandÂ
Leonardo DiCaprio	Blood DiamondÂ
Leonardo DiCaprio	Catch Me If You CanÂ
Leonardo DiCaprio	TitanicÂ
Leonardo DiCaprio	The AviatorÂ
Leonardo DiCaprio	Gangs of New YorkÂ
Leonardo DiCaprio	The Great GatsbyÂ
Leonardo DiCaprio	The Great GatsbyÂ
Leonardo DiCaprio	Revolutionary RoadÂ
Leonardo DiCaprio	Body of LiesÂ
Leonardo DiCaprio	Romeo + JulietÂ
Leonardo DiCaprio	Marvin's RoomÂ
Leonardo DiCaprio	J. EdgarÂ
Leonardo DiCaprio	The BeachÂ
Leonardo DiCaprio	The Man in the Iron MaskÂ
Leonardo DiCaprio	The Quick and the DeadÂ

actor_1_name= Brad Pitt



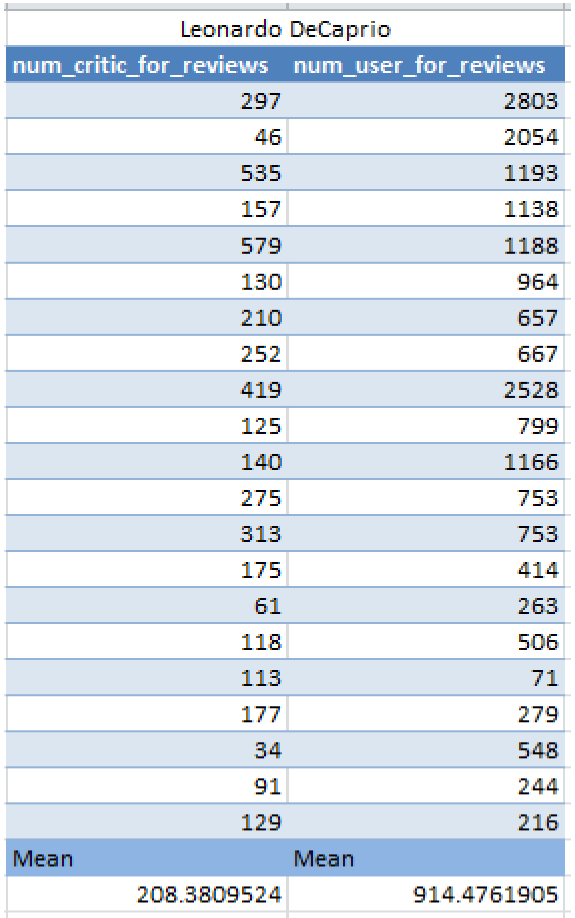
actor_1_name			Brad_Pitt
Brad Pitt			Fight ClubÂ
Brad Pitt			True RomanceÂ
Brad Pitt			The Curious Case of Benjamin ButtonÂ
Brad Pitt			Ocean's ElevenÂ
Brad Pitt			FuryÂ
Brad Pitt			Interview with the Vampire: The Vampire ChroniclesÂ
Brad Pitt			BabelÂ
Brad Pitt			The Assassination of Jesse James by the Coward Robert FordÂ
Brad Pitt			TroyÂ
Brad Pitt			Seven Years in TibetÂ
Brad Pitt			Spy GameÂ
Brad Pitt			Sinbad: Legend of the Seven SeasÂ
Brad Pitt			The Tree of LifeÂ
Brad Pitt			Mr. & Mrs. SmithÂ
Brad Pitt			Ocean's TwelveÂ
Brad Pitt			Killing Them SoftlyÂ
Brad Pitt			By the SeaÂ

Mean of num_critic_for_review and num_users_for_review
And the actor which have highest mean.
Highest Mean num_critic_for_reviews= Brad Pitt (218.82)
Highest Mean num_users_for_review= Leonardo DiCaprio(914.47)



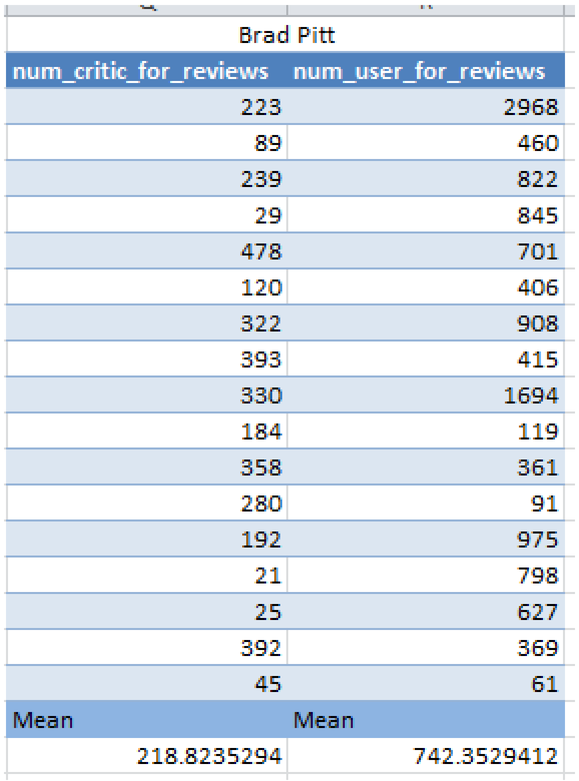
Meryl Streep

num_critic_for_reviews	num_user_for_reviews
120	660
161	200
575	112
129	277
219	631
28	280
77	214
423	350
283	69
130	178
93	298
Mean	**Mean**
203.4545455	297.1818182



Leonardo DeCaprio

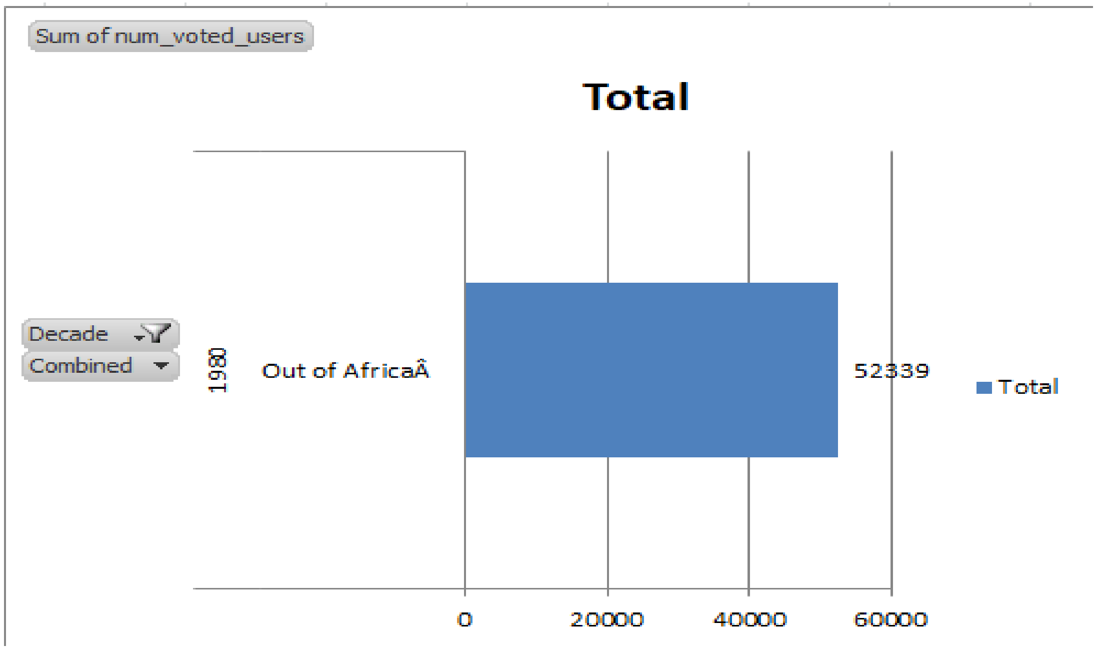
num_critic_for_reviews	num_user_for_reviews
297	2803
46	2054
535	1193
157	1138
579	1188
130	964
210	657
252	667
419	2528
125	799
140	1166
275	753
313	753
175	414
61	263
118	506
113	71
177	279
34	548
91	244
129	216
Mean	Mean
208.3809524	914.4761905

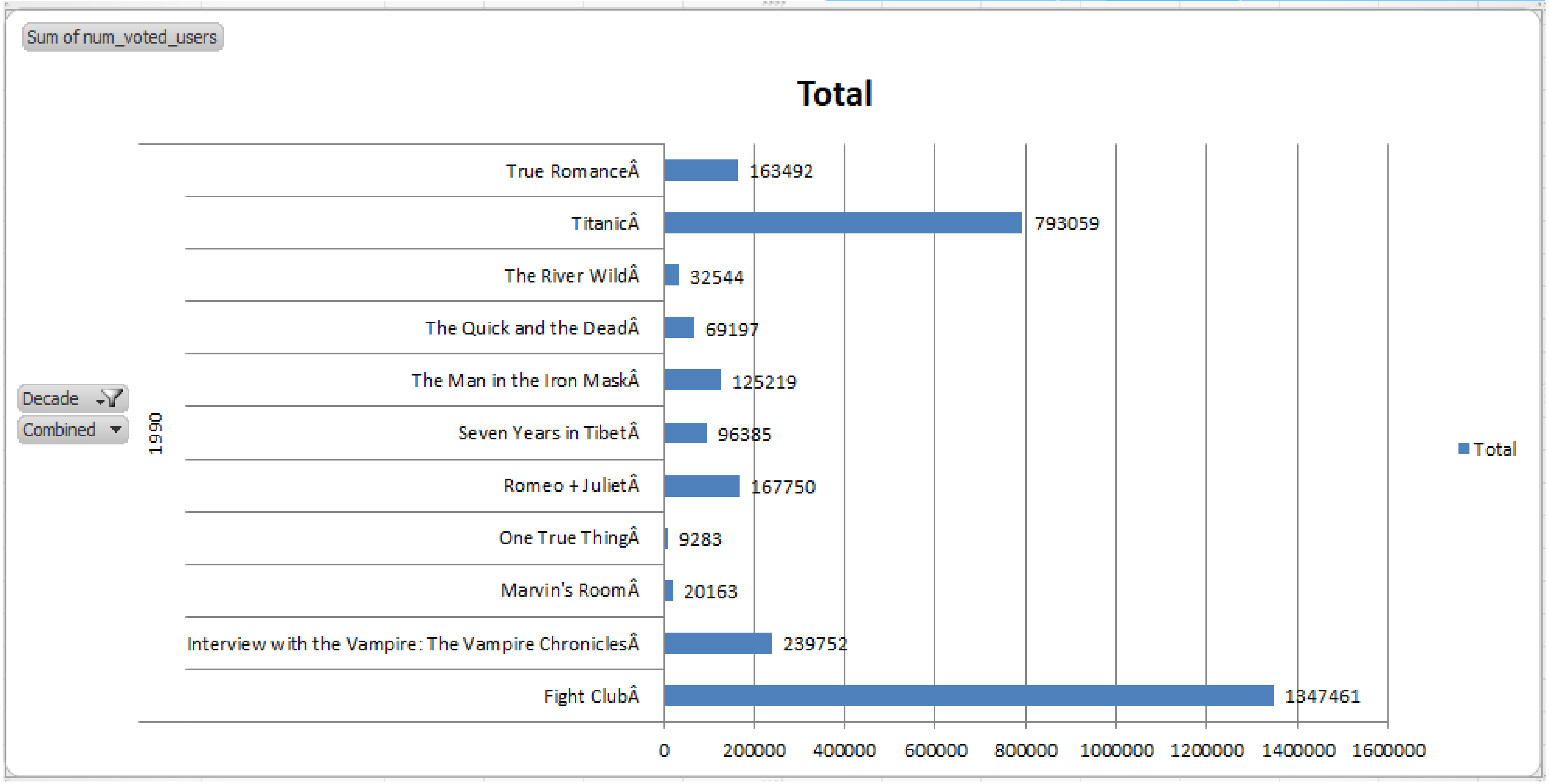


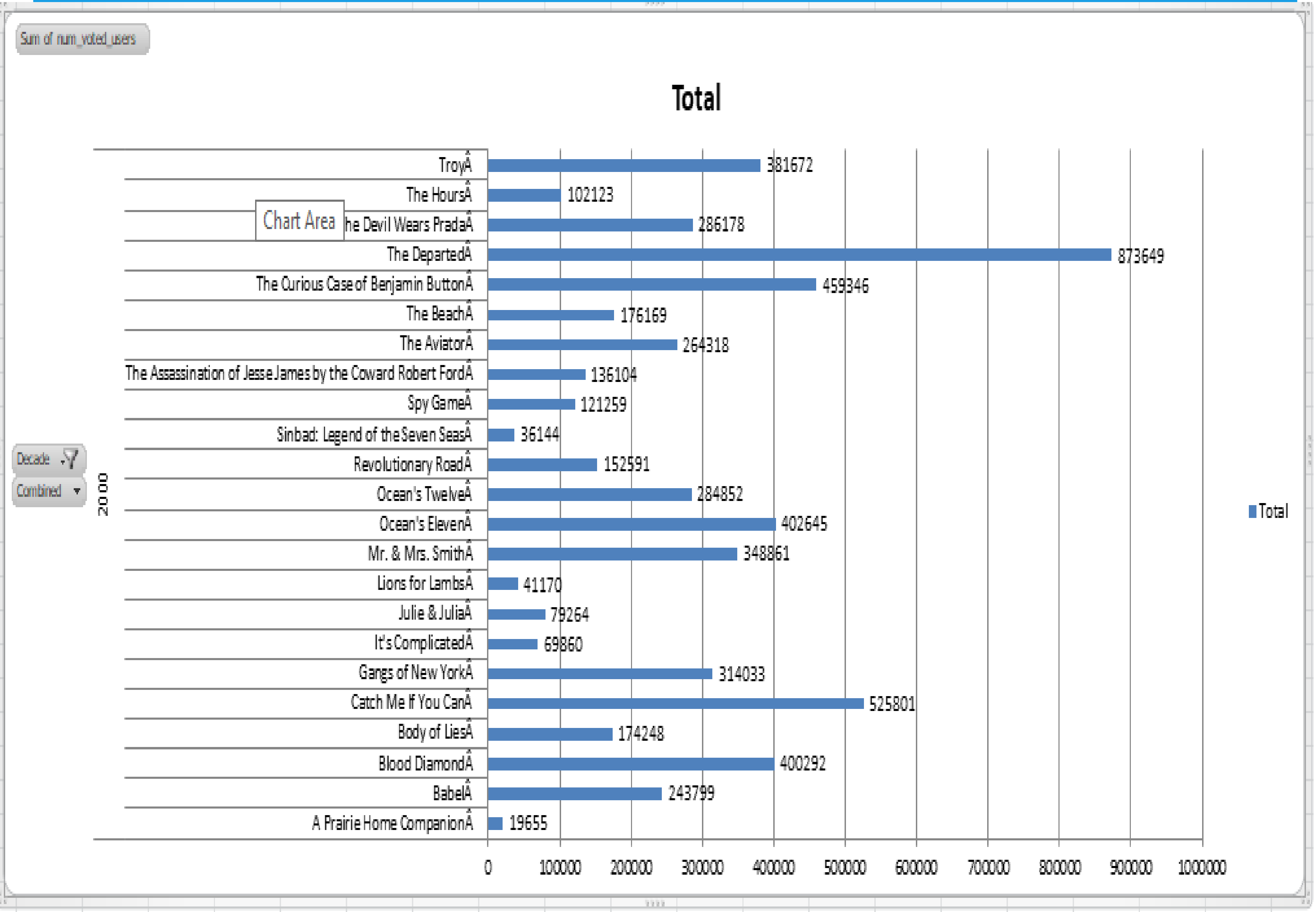
Brad Pitt

num_critic_for_reviews	num_user_for_reviews
223	2968
89	460
239	822
29	845
478	701
120	406
322	908
393	415
330	1694
184	119
358	361
280	91
192	975
21	798
25	627
392	369
45	61
Mean	Mean
218.8235294	742.3529412

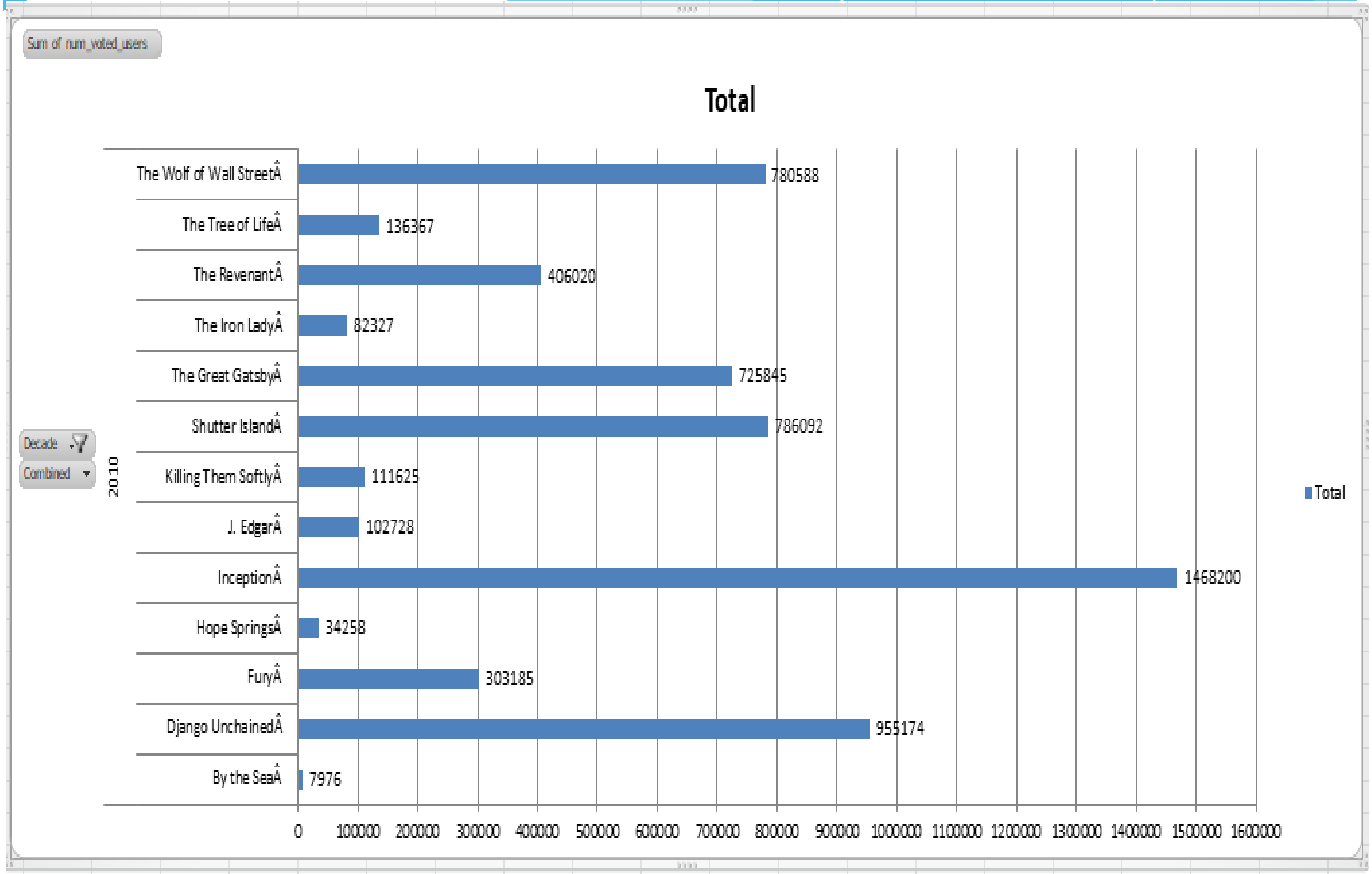
Created a decade column where we extract the value of movie year on which decade movies were released.



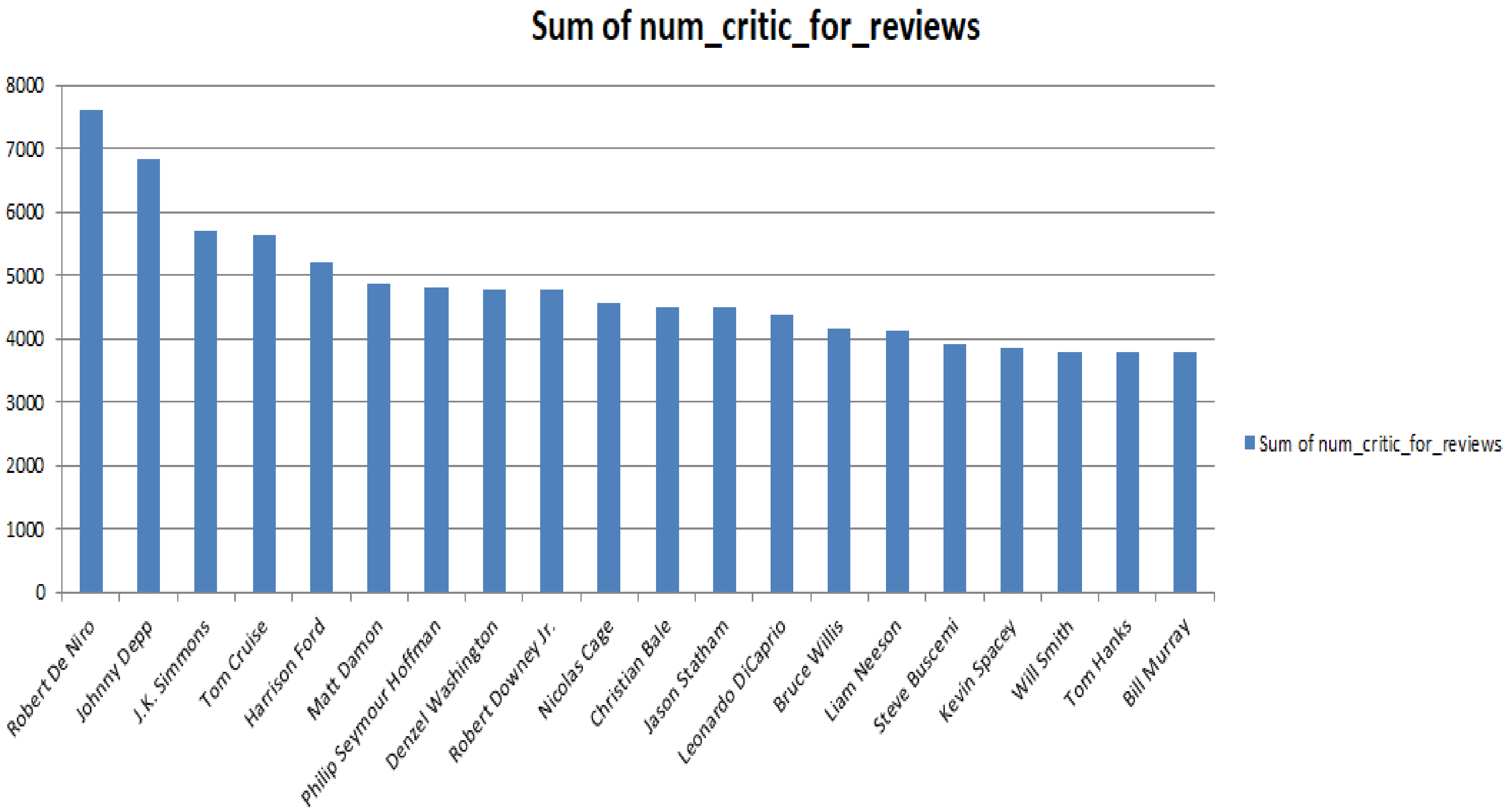

Sum of num_voted_users
Total
Decade
Combined
1990
True RomanceÂ
TitanicÂ
The River WildÂ
The Quick and the DeadÂ
The Man in the Iron MaskÂ
Seven Years in TibetÂ
Romeo + JulietÂ
One True ThingÂ
Marvin's RoomÂ
Interview with the Vampire: The Vampire ChroniclesÂ
Fight ClubÂ
163492
793059
32544
69197
125219
96385
167750
9283
20163
239752
1347461
0
200000
400000
600000
800000
1000000
1200000
1400000
1600000
Total

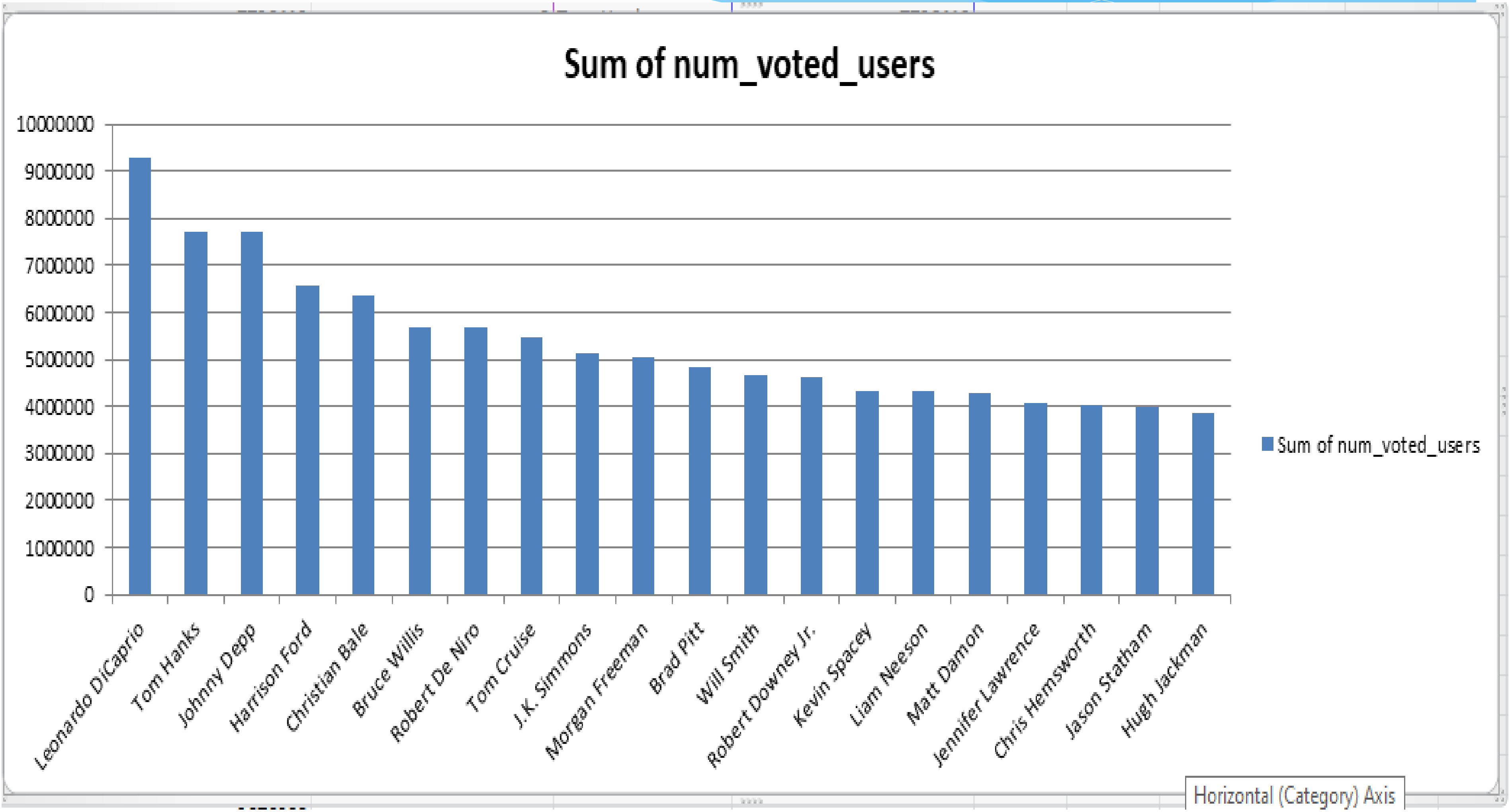


Sum of num_voted_users
Total
Chart Area
Decade
Combined
2000
TroyÂ 381672
The HoursÂ 102123
The Devil Wears PradaÂ 286178
The DepartedÂ 873649
The Curious Case of Benjamin ButtonÂ 459346
The BeachÂ 176169
The AviatorÂ 264318
The Assassination of Jesse James by the Coward Robert FordÂ 136104
Spy GameÂ 121259
Sinbad: Legend of the Seven SeasÂ 36144
Revolutionary RoadÂ 152591
Ocean's TwelveÂ 284852
Ocean's ElevenÂ 402645
Mr. & Mrs. SmithÂ 348861
Lions for LambsÂ 41170
Julie & JuliaÂ 79264
It's ComplicatedÂ 69860
Gangs of New YorkÂ 314033
Catch Me If You CanÂ 525801
Body of LiesÂ 174248
Blood DiamondÂ 400292
BabelÂ 243799
A Prairie Home CompanionÂ 19655
0
100000
200000
300000
400000
500000
600000
700000
800000
900000
1000000
Total


Sum of num_voted_users
Total
Decade
Combined
2010
The Wolf of Wall StreetÂ
The Tree of LifeÂ
The RevenantÂ
The Iron LadyÂ
The Great GatsbyÂ
Shutter IslandÂ
Killing Them SoftlyÂ
J. EdgarÂ
InceptionÂ
Hope SpringsÂ
FuryÂ
Django UnchainedÂ
By the SeaÂ
780588
136367
406020
82327
725845
786092
111625
102728
1468200
34258
303185
955174
7976
0
100000
200000
300000
400000
500000
600000
700000
800000
900000
1000000
1100000
1200000
1300000
1400000
1500000
1600000
Total

Find the critic-favorite and audience-favorite actors




Sum of num_voted_users
10000000
9000000
8000000
7000000
6000000
5000000
4000000
3000000
2000000
1000000
0
Leonardo DiCaprio
Tom Hanks
Johnny Depp
Harrison Ford
Christian Bale
Bruce Willis
Robert De Niro
Tom Cruise
J.K. Simmons
Morgan Freeman
Brad Pitt
Will Smith
Robert Downey Jr.
Kevin Spacey
Liam Neeson
Matt Damon
Jennifer Lawrence
Chris Hemsworth
Jason Statham
Hugh Jackman
Sum of num_voted_users
Horizontal (Category) Axis

Result

Through this project, we were able to gain insights into various aspects of the movie industry such as popular genres, best directors, and favorite actors. By analyzing the data, we were able to identify patterns and trends that can help movie studios make informed decisions. We also learned how to use Excel to perform data analysis and visualization. Overall, this project helped us develop our skills in data analysis and provided us with a deeper understanding of the movie industry.

The End