IMDB Movie Analysis



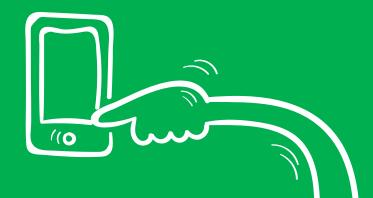


Project Description

This project is about performing data analysis on the given data set. This project mainly requires Microsoft excel to group, filter and visualize the data.



Project Link- IMDB.xlsx



Tech-Stack

Microsoft Excel 2019

Table of Contents

03 05 01 Cleaning the data **Top 250 IMDB Movies Popular Genres** 04 02 06 **Best Directors Charts Movies with Highest Profit**

Cleaning of Data

Use your knowledge learned till now and clean the data provided. (Dropping columns, removing null values, etc.) **Your task:** Clean the data

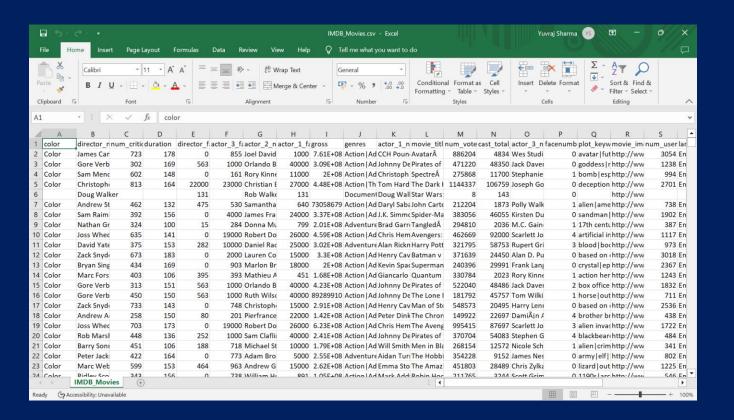


Cleaning the data involves the following steps-

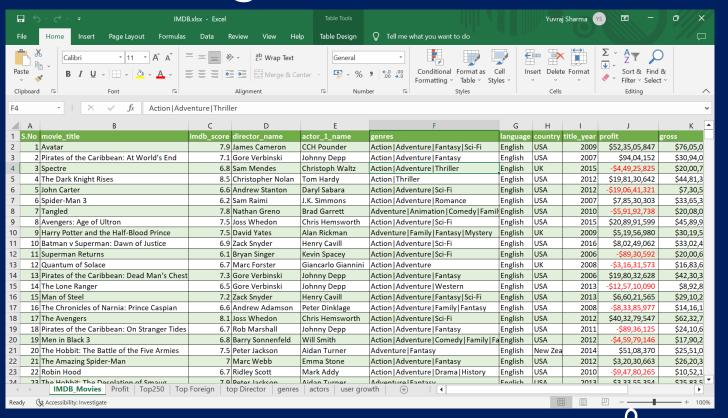
- 1. Delete the unwanted column.
- 2. Arrange the column & adjust their width.
- 3. Remove the rows with blank cells.
- 4. Remove the duplicate row using "Delete Duplicate Row Function".
- 5. Format them and add serial number.



Raw data



After Cleaning the data



Movies With Highest Profit

Create a new column called profit which contains the difference of the two columns: gross and budget. Sort the column using the profit column as reference and observe the outliers using the appropriate chart type.

Your task: Find the movies with the highest profit?



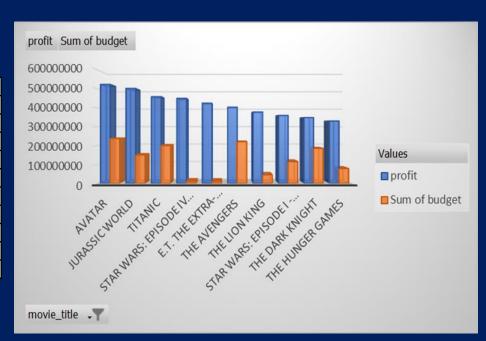
Steps finding Movies with highest profit-

- 1. Create a new column 'profit', which is the difference between Gross and Budget column.
- 2. Make a pivot table using movie title, profit and budget.
- 3. Use value filter to get top 10 highest profit movies.
- 4. Plot graph.



Movies	profit	Sum of budget
Avatar	523505847	237000000
Jurassic World	502177271	150000000
Titanic	458672302	200000000
Star Wars: Episode IV - A New Hope	449935665	11000000
E.T. the Extra-Terrestrial	424449459	10500000
The Avengers	403279547	220000000
The Lion King	377783777	45000000
Star Wars: Episode I - The Phantom Menace	359544677	115000000
The Dark Knight	348316061	185000000
The Hunger Games	329999255	78000000

Star Wars: Episode IV – A New Hope & E.T. the Extra Terrestrial are the outliers because their profit is huge as compared to their budget.





Top 250 IMDB Movies

Create a new column IMDB_Top_250 and store the top 250 movies with the highest IMDb Rating (corresponding to the column: IMDB score. Also Extract all the movies in the IMDB_Top_250 column which are not in the English language.



Steps to find top 250 IMDB Movies-

- 1. Remove the column which are not required.
- 2. Sort the movies with num_voted_users and remove the movies having votes less than 25000.
- 3. Sort Movies according to IMDB score and remove all the movies having rank below 250.

Top IMDB Foreign Movies-Table 1- Filter all movies which are not in English Table 2 – Filter all movies which are not from USA.



Ran		imdb_s num_vot core ed users				
	The Shawshank Redemption	9.3 1689764				
	The Godfather	9.2 1155770				101 The Bridge on the River Kwai
	The Dark Knight	9 1676169	51 Baahubali: The Beginning	8.4	62756	102 The Secret in Their Eyes
	The Godfather: Part II	9 790926	52Braveheart	8.4	736638	103 The Thing
	Pulp Fiction	8.9 1324680	53Das Boot	8.4	168203	104The Wolf of Wall Street
	Schindler's List	8.9 865020	54Lawrence of Arabia	8.4	192775	105 Trainspotting
	The Good, the Bad and the Ugly	8.9 503509	55 <mark>Oldboy</mark>	8.4		106V for Vendetta
	The Lord of the Rings: The Return of the		56Once Upon a Time in America	8.4		107Warrior
	King	8.9 1215718	57 Princess Mononoke	8.4		10812 Years a Slave
	Fight Club	8.8 1347461	58 Requiem for a Dream	8.4		109Akira
	Forrest Gump	8.8 1251222	59 Reservoir Dogs	8.4		110Amores Perros
	Inception	8.8 1468200	60 Star Wars: Episode VI - Return of the Jedi	8.4		111 Annie Hall
	Star Wars: Episode V - The Empire Strikes		61WALL-E	8.4		112Before Sunrise
	Back	8.8 837759	622001: A Space Odyssey	8.3		113Butch Cassidy and the Sundance Kid
	The Lord of the Rings: The Fellowship of		63Amadeus		270790	114 Deadpool
	the Ring	8.8 1238746	64Batman Begins	8.3		115 Donnie Darko
	City of God	8.7 533200	65 Downfall	8.3		116 Elite Squad
	Goodfellas	8.7 728685	66 Eternal Sunshine of the Spotless Mind	8.3		117 Gone Girl
	One Flew Over the Cuckoo's Nest	8.7 680041	67Good Will Hunting	8.3		118 Groundhog Day
	Seven Samurai	8.7 229012	68 Indiana Jones and the Last Crusade		515306	119Guardians of the Galaxy
183	Star Wars: Episode IV - A New Hope	8.7 911097	69 Inglourious Basterds	8.3		120 Hotel Rwanda
	The Lord of the Rings: The Two Towers	8.7 1100446	70 Inside Out	8.3		121 Jurassic Park
	The Matrix	8.7 1217752	71 L.A. Confidential		414219	122Kill Bill: Vol. 1
	American History X	8.6 782437	72 Metropolis 73 Monty Python and the Holy Grail	8.3	111841 382240	123Mad Max: Fury Road
	Interstellar	8.6 928227			235133	124Million Dollar Baby
	Modern Times	8.6 143086	74Raging Bull 75Room		161288	125Monsters, Inc. 126No Country for Old Men
	Saving Private Ryan	8.6 881236	76Scarface	8.3		Pirates of the Caribbean: The Curse of the
	Se7en	8.6 1023511 8.6 417971	77 Snatch		600996	127 Black Pearl
	Spirited Away		78 Some Like It Hot		175196	128 Platoon
	The Silence of the Lambs The Usual Suspects	8.6 887467 8.6 740918	79 The Hunt		170155	129 Prisoners
	Alien	8.5 563827	80 The Sting		175607	130 Rocky
	Apocalypse Now	8.5 450676	81 Toy Story 3	8.3		131 Rush
	Back to the Future	8.5 732212	82 Toy Story	8.3		132 Shutter Island
	Children of Heaven	8.5 27882	83 Unforgiven		277505	133 Sin City
	Django Unchained	8.5 955174	84 Up		665575	134Spotlight
	Gladiator	8.5 982637	85 A Beautiful Mind	8.2		135 Stand by Me
	Memento	8.5 845580	86 Blade Runner		461609	136 Tae Guk Gi: The Brotherhood of War
	Psycho	8.5 422432	87 Captain America: Civil War		272670	137 The Avengers
	Raiders of the Lost Ark	8.5 661017	88Casino	8.2		138 The Best Years of Our Lives
	Terminator 2: Judgment Day	8.5 744891	89 Die Hard		592582	139 The Bourne Ultimatum
	The Dark Knight Rises	8.5 1144337	90 Finding Nemo	8.2		140 The Celebration
	The Departed	8.5 873649	91 Gone with the Wind	8.2	215340	141 The Grand Budapest Hotel
	The Green Mile	8.5 782610	92 Gran Torino	8.2	561773	142The Help
	The Lion King	8.5 644348	93 How to Train Your Dragon	8.2		143 The Imitation Game
43	The Lives of Others	8.5 259379	94 Howl's Moving Castle	8.2	214091	144 The Martian
	The Pianist	8.5 497946	95 Incendies	8.2		145 The Princess Bride
45	The Prestige	8.5 844052	96 Into the Wild	8.2		146 The Revenant
	Whiplash	8.5 399138	97 Lock, Stock and Two Smoking Barrels	8.2		147 The Sea Inside
	A Separation	8.4 151812	98 On the Waterfront		100890	148 The Sixth Sense
	Aliens	8.4 488537	99 Pan's Labyrinth	8.2		149 The Terminator
	Amelie	8.4 534262	100 The Big Lebowski	8.2	537419	150 The Truman Show
50/	American Beauty	8.4 822500				

8.2 149444

8.2 131831 8.2 258078 8.2 780588 8.2 469561 8.2 791783 8.2 332276 8.1 439176 8.1 173551 8.1 192940

8.1 152089

8.1

151 The Wizard of Oz	8.1 291875	201 True Romance	8 163492
152 There Will Be Blood	8.1 372990	202Waltz with Bashir	8 46107
153A Fistful of Dollars	8 147566	203 X-Men: Davs of Future Past	8 514125
154Aladdin	8 260939	204 Young Frankenstein	8 112671
155Before Sunset	8 168398	2054 Months, 3 Weeks and 2 Days	7.9 44763
156 Big Fish	8 350698	206 Almost Famous	7.9 207287
157 Black Swan	8 551363	207 Amour	7.9 70382
158 Blood Diamond	8 400292	208Avatar	7.9 886204
159 Bowling for Columbine	8 123090	209 Before Midnight	7.9 95362
160 Boyhood	8 266020	210 Big Hero 6	7.9 279093
161 Brazil	8 152306	211 Boogie Nights	7.9 189032
162 Casino Royale	8 470483	212 Captain Phillips	7.9 323353
163 Catch Me If You Can	8 525801	213 Children of Men	7.9 361767
164Central Station	8 28951	214 Crash	7.9 361169
165Cinderella Man	8 148238	215 Crouching Tiger, Hidden Dragon	7.9 217740
166 Dallas Buvers Club	8 326494	216Do the Right Thing	7.9 59524
167 Dancer in the Dark	8 79330	217 E.T. the Extra-Terrestrial	7.9 59524
168 Dances with Wolves	8 186485	218 Ed Wood	7.9 261642
169 Dead Poets Society	8 277451	219Edge of Tomorrow	7.9 431620
170 District 9		220 Edward Scissorhands	7.9 431620
	8 531737		
171 Doctor Zhivago 172 Fiddler on the Roof	8 55816 8 29839	221 Glory 222 Halloween	7.9 101888
			7.9 157857
173Her	8 355126	223Hero	7.9 149414
174In Bruges	8 307639	224 Hot Fuzz	7.9 352695
175 Jaws	8 412454	225 How to Train Your Dragon 2	7.9 221128
176JFK	8 113472	226 Iron Man	7.9 696338
177Kill Bill: Vol. 2	8 512749	227 Letters from Iwo Jima	7.9 132149
178Life of Pi	8 440084	228 Little Miss Sunshine	7.9 355810
179Magnolia	8 241030	229 Moon	7.9 260607
180 Mulholland Drive	8 235992	230 My Fair Lady	7.9 66959
181 My Name Is Khan	8 69759	231 Nightcrawler	7.9 293304
182 Mystic River	8 338415	232 Nine Queens	7.9 38215
183 Persepolis	8 70194	233 Once	7.9 90827
184Rain Man	8 383784	234 Shrek	7.9 467113
185 Ratatouille	8 473887	235 Straight Outta Compton	7.9 119928
186 Serenity	8 242599	236 Taken	7.9 483756
187 Shaun of the Dead	8 395921	237 The Blues Brothers	7.9 142448
188Sicko	8 66610	238 The Bourne Identity	7.9 407601
189 Sling Blade	8 72443	239 The Chorus	7.9 44151
190 Slumdog Millionaire	8 641997	240 The Fighter	7.9 275869
191 Star Trek	8 504419	241 The Hateful Eight	7.9 272839
192 The Artist	8 190030	242 The Hobbit: An Unexpected Journey	7.9 637246
193 The Exorcist	8 284252	243 The Hobbit: The Desolation of Smaug	7.9 483540
194 The Incredibles	8 479166	244 The Insider	7.9 133526
195 The Iron Giant	8 128455	245 The Notebook	7.9 396396
196The King's Speech	8 503631	246 The Remains of the Day	7.9 45703
197 The Perks of Being a Wallflower	8 351274	247 The Right Stuff	7.9 45271
198 The Pursuit of Happyness	8 338383	248 The Untouchables	7.9 219008
199The Sound of Music	8 148172	249 The World's Fastest Indian	7.9 44198

Top IMDB movies which are not in English language

Rankmovie_title	imdb_score	language	country
The Good, the Bad and the			
1 Ugly	8.9	Italian	Italy
		Portugues	
2City of God	8.7		Brazil
3Seven Samurai		Japanese	
4Spirited Away		Japanese	
5Children of Heaven		Persian	Iran
6The Lives of Others		German	Germany
7A Separation		Persian	Iran -
8 Amelie		French	France
9Baahubali: The Beginning	8.4	Telugu	India
10 Das Boot	8.4	German	West Germany
11 Oldboy	8.4	Korean	South Kore
12 Princess Mononoke	8.4	Japanese	Japan
13 Downfall	8.3	German	Germany
14 Metropolis	8.3	German	Germany
15 The Hunt	8.3	Danish	Denmark
16 Howl's Moving Castle	8.2	Japanese	Japan
17 Incendies		French	Canada
18 Pan's Labyrinth	8.2	Spanish	Spain
19 The Secret in Their Eyes	8.2	Spanish	Argentina
20 Akira	8.1	Japanese	Japan
21 Amores Perros		Spanish	Mexico
		Portugues	
22 Elite Squad	8.1	e	Brazil
Tae Guk Gi: The			
23Brotherhood of War		Korean	South Kore
24 The Celebration		Danish	Denmark
25 The Sea Inside		Spanish	Spain
26A Fistful of Dollars	8	Italian	Italy
27 Control Station		Portugues	Deseil
27 Central Station		e Hindi	Brazil India
28 My Name Is Khan 29 Persepolis		French	India France
30Waltz with Bashir		Hebrew	France Israel
4 Months, 3 Weeks and 2	8	nebrew	israei
31 Days	7.0	Romanian	Romania
32 Amour		French	France
Crouching Tiger, Hidden	7.0		
33 Dragon	7.9	Mandarin	Taiwan
34Hero	7.9	Mandarin	China
35 Letters from Iwo Jima	7.9	Japanese	USA
36 Nine Queens		Spanish	Argentina
37 The Chorus		French	France

Top foreign IMDB movies which are not from USA

8.1 English

8.1 Korean

8.1 Danish

8.1 English

8.1 English

8 Italian 8 English

8 English

8 English

8 English 8 Portuguese Brazil 8 English

8 English

8 English 8 English

8 English

8 French

8 English

8 English 8 English

8 English

8 English

8 Hebrew

7.9 English

7.9 English

7.9 English 7.9 Spanish

7.9 English

7.9 English

7.9 French

7.9 English

7.9 English

7.9 Mandarin

7.9 Mandarin

8 Hindi

8.1 Spanish

UK

UK

UK Italy

UK UK

Spain

South Korea

Denmark

Germany

Germany

Denmark

France

France

India

France

France

France

France

Taiwan

China

UK

Germany

Argentina

Ireland

France

France

New Zealand

UK

Israel 7.9 Romanian Romania 7.9 French

UK UK

UK

South Africa UK

Ran		imdb_sc			
k	movie_title	ore	language	country	
1	The Good, the Bad and the Ugly	8.9	Italian	Italy	38Rush
	The Lord of the Rings: The Fellowship of				39 Tae Guk Gi: The Brotherhood of War
2	the Ring	8.8	English	New Zealand	40 The Celebration
3	City of God	8.7	Portuguese	Brazil	41 The Imitation Game
4	Seven Samurai	8.7	Japanese	Japan	42 The Sea Inside
5	Spirited Away	8.6	Japanese	Japan	43 The Terminator
	Alien	8.5	English	UK	44A Fistful of Dollars
7	Children of Heaven	8.5	Persian	Iran	45 Blood Diamond
8	The Lives of Others	8.5	German	Germany	46 Bowling for Columbine
9	The Pianist	8.5	English	France	47Brazil
10	A Separation	8.4	Persian	Iran	48 Casino Royale
11	Amelie	8.4	French	France	49 Central Station
12	Baahubali: The Beginning	8.4	Telugu	India	50 Dancer in the Dark
				West	51 District 9
13	Das Boot	8.4	German	Germany	52 In Bruges
14	Lawrence of Arabia	8.4	English	UK	53JFK
15	Oldboy	8.4	Korean	South Korea	54 Mulholland Drive
16	Once Upon a Time in America	8.4	English	Italy	55 My Name Is Khan
17	Princess Mononoke	8.4	Japanese	Japan	56 Persepolis
18	2001: A Space Odyssey	8.3	English	UK	57 Shaun of the Dead
19	Downfall	8.3	German	Germany	58 Slumdog Millionaire
20	Metropolis	8.3	German	Germany	59 The Artist
21	Monty Python and the Holy Grail	8.3	English	UK	60 The King's Speech
22	Room	8.3	English	Ireland	61 The Straight Story
23	Snatch	8.3	English	UK	62 Waltz with Bashir
24	The Hunt	8.3	Danish	Denmark	634 Months, 3 Weeks and 2 Days
25	Howl's Moving Castle	8.2	Japanese	Japan	64Amour
26	Incendies	8.2	French	Canada	65Crash
27	Lock, Stock and Two Smoking Barrels	8.2	English	UK	66 Crouching Tiger, Hidden Dragon
	Pan's Labyrinth	8.2	Spanish	Spain	67Hero
29	The Bridge on the River Kwai	8.2	English	UK	68 Hot Fuzz
30	The Secret in Their Eyes	8.2	Spanish	Argentina	69Moon
31	Trainspotting	8.2	English	UK	70 Nine Queens
32	Akira	8.1	Japanese	Japan	71 Once
33	Amores Perros	8.1	Spanish	Mexico	72 Taken
34	Elite Squad	8.1	Portuguese	Brazil	73 The Chorus
35	Hotel Rwanda	8.1	English	UK	74 The Remains of the Day
36	Mad Max: Fury Road	8.1	English	Australia	75 The World's Fastest Indian
	Platoon		English	UK	

Top Directors

Find out the top 10 directors for whom the mean of imdb_score is the highest and store them in a new column top10director. In case of a tie in IMDb score between two directors, sort them alphabetically.

Your task: Find the best directors



Steps to get Top Directors-

- 1. Use pivot table to get directors and their mean IMDB score.
- 2. Sort them with their mean IMDB score.
- 3. Use filter value to get top 10 IMDB directors.
- 4. Manually sort them alphabetically.



top 10 Directors	Average of imdb_score
Charles Chaplin	8.6
Tony Kaye	8.6
Alfred Hitchcock	8.5
Damien Chazelle	8.5
Majid Majidi	8.5
Ron Fricke	8.5
Sergio Leone	8.433333333
Christopher Nolan	8.425
Asghar Farhadi	8.4
Marius A. Markevicius	8.4
Richard Marquand	8.4
S.S. Rajamouli	8.4



Top Genres

Popular Genres: Perform this step using the knowledge gained while performing previous steps. **Your task:** Find popular genres.

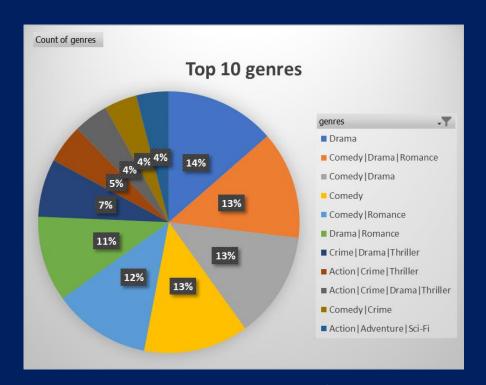


Steps to get top 10 genres-

- 1. Use Pivot Table to group the genres with highest count.
- 2. Use value filter to get top 10 genres.3. Plot the pie chart.



Genres	Count of genres
Drama	152
Comedy Drama Romance	149
Comedy Drama	147
Comedy	145
Comedy Romance	135
Drama Romance	118
Crime Drama Thriller	80
Action Crime Thriller	54
Action Crime Drama Thriller	48
Comedy Crime	45
Action Adventure Sci-Fi	45





Charts

Who has the most reviews of critics and users among Leonardo DiCaprio, Brad Pitt and Meryl Streep. Also plot the graph for growth of users votes for past 10 decades.



Steps to get the charts-

- 1. Use pivot table to get all three actors in column with their critics and user reviews in rows.
- 2. Plot the chart.
- 3. Now use title year and sum of user votes to plot the graph of user growth.
- 4. Use group function to group the years in the interval of 10 years.
- 5. Plot the chart.



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

Combined	Group by
Babel	Brad Pitt
By the Sea	Brad Pitt
Fight Club	Brad Pitt
Fury	Brad Pitt
Interview with the Vampire: The	
Vampire Chronicles	Brad Pitt
Killing Them Softly	Brad Pitt
Mr. & Mrs. Smith	Brad Pitt
Ocean's Eleven	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
	Didd i itt
The Assassination of Jesse James	
by the Coward Robert Ford	Brad Pitt
The Curious Case of Benjamin	0 0.11
Button	Brad Pitt Brad Pitt
The Tree of Life Troy	Brad Pitt
True Romance	Brad Pitt
Blood Diamond	Leonardo DiCaprio
Body of Lies	Leonardo DiCaprio
	Leonardo DiCaprio
Catch Me If You Can	· '
Django Unchained	Leonardo DiCaprio Leonardo DiCaprio
Gangs of New York Inception	Leonardo DiCaprio
J. Edgar	Leonardo DiCaprio
Marvin's Room	Leonardo DiCaprio
Revolutionary Road	Leonardo DiCaprio
Romeo + Juliet	Leonardo DiCaprio
Shutter Island	Leonardo DiCaprio
The Aviator	Leonardo DiCaprio
The Beach	Leonardo DiCaprio
The Departed	Leonardo DiCaprio
The Great Gatsby	Leonardo DiCaprio
The Man in the Iron Mask	Leonardo DiCaprio
The Quick and the Dead	Leonardo DiCaprio
The Revenant	Leonardo DiCaprio
The Wolf of Wall Street	Leonardo DiCaprio
A Prairie Home Companion	Meryl Streep
Hope Springs	Meryl Streep
It's Complicated	Meryl Streep
Julie & Julia	Meryl Streep
Lions for Lambs	Meryl Streep
One True Thing	Meryl Streep
Out of Africa	Meryl Streep
The Devil Wears Prada	Meryl Streep
The Hours	Meryl Streep
The Iron Lady	Meryl Streep
The River Wild	Meryl Streep
me niver wild	ivici yi bu eep



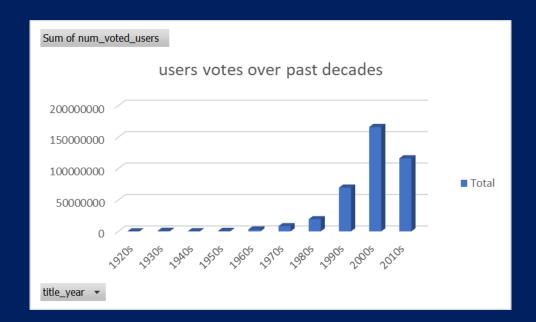
Values	Brad Pitt	Leonardo DiCaprio	Meryl Streep
Average of critic reviews	245	322.2	181.4545455
Average of user reviews	742.3529412	922.55	297.1818182





df_by_decades	Sum of num_voted_users
1920s	116387
1930s	804839
1940s	230838
1950s	678336
1960s	2983442
1970s	8269828
1980s	19344369
1990s	69635863
2000s	166041149
2010s	116239604

Exponential Growth in user votes since past 10 decades





Insights & Results-

This project helped me understand different types of techniques and functions of MS excel.



Thanks

