# IMDB Movie Analysis

## Project Description:

The project imdb movie analysis is about finding the required insights by going through the raw data provided and first performing clean techniques to clean the raw data and transform it into a decent data through which accurate insights and knowledge can be extracted . We can use 5 “Why’s” Technique which is also called as Root Cause Analysis developed by Sakichi Toyoda to perform the analysis

In this project ,

1. Clean the raw data provided using various cleaning techniques
2. Perform analysis and visualization to get valuable insights

## Approach :

The main approach towards this project is to first understand the dataset provided. Then using various cleaning techniques we can clean the data and get rid of null values and duplicates to prepare data for analysis stage . For analysis depending upon the insights required we can use various charts , pivot tables ,functions ,etc . We can use “Why’s” technique to get to root of the problem and reach the desired solution . At the end we will display the insights extracted using various tables and charts to make it more easy to quickly understand the insights.

**Tech used :** The main software used during the project is Microsoft Excel

## Insights :

### A) Cleaning the data:

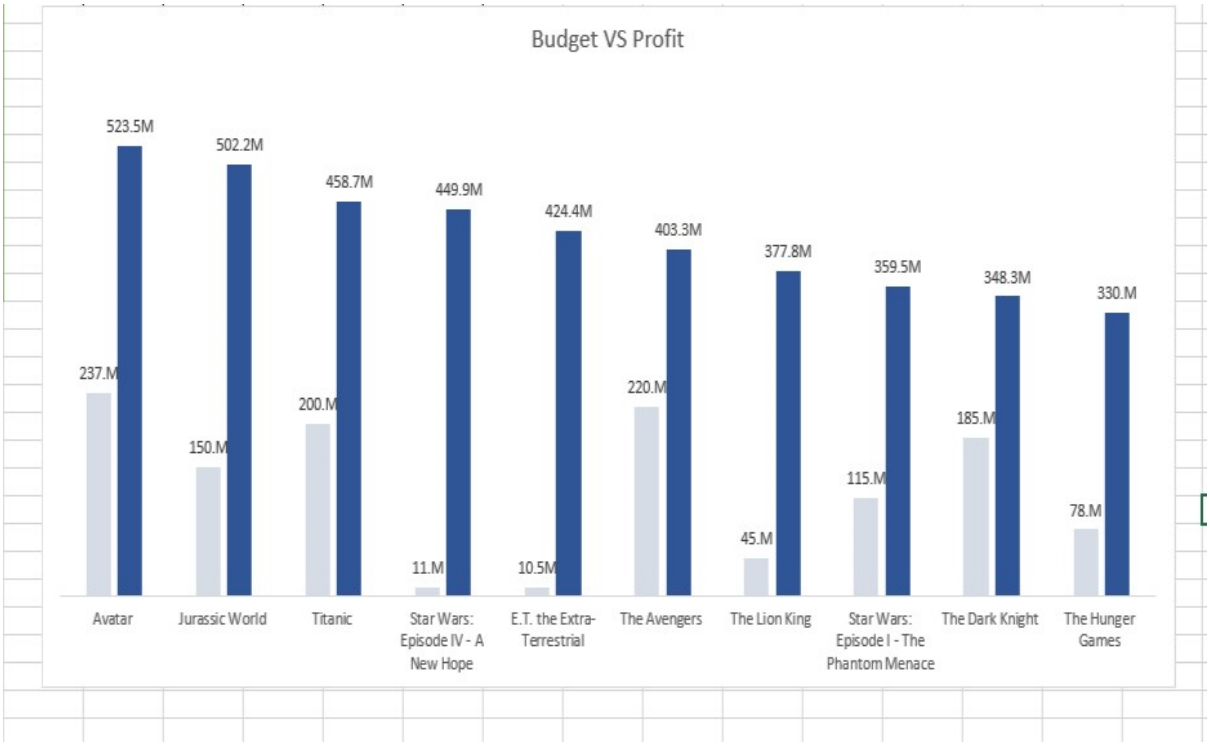
- Dropping duplicates, null values
- Dropping columns like color,director_facebook_likes, actor_1_facebook_likes , actor_2_facebook_likes , duration, actor_3_facebook_Likes,cast_total_facebook_likes,actor_2_name,director_facebook_likes, actor_1_facebook_likes, actor_2_facebook_likes, cast_total_facebook_likes,actor_2_name,actor_3_name,duration,facenumber_in_poster,content_rating,country,movie_imdb_link,aspect_ratio,plot_keywords

This is the most important and very first step analysis to make the data ready for analysis . Various cleaning techniques can be used .

### B) Movies with highest profit:

First we find the profit for every movie by calculating difference between their gross income and budget while filming . Then with the help of filter we adjusted the data with movies gaining highest profits to find out some of the movies with highest profit. Another task in this was to observe the outliers by plotting graph . We use a bar column for that . After carefully observing the bar chart we find out two outliers (Star Wars Episode IV and E.T) which had lowest budget of all but had significant higher profit margin .

| | director_name | gross | movie_title | language | imdb_score | budget | Profit |
|---|---|---|---|---|---|---|---|
| 2 | James Cameron | 760.5M | Avatar | English | 7.9 | 237.M | 523.5M |
| 3 | Colin Trevorrow | 652.2M | Jurassic World | English | 7 | 150.M | 502.2M |
| 4 | James Cameron | 658.7M | Titanic | English | 7.7 | 200.M | 458.7M |
| 5 | George Lucas | 460.9M | Star Wars: Episode IV - A New Hope | English | 8.7 | 11.M | 449.9M |
| 6 | Steven Spielberg | 434.9M | E.T. the Extra-Terrestrial | English | 7.9 | 10.5M | 424.4M |
| 7 | Joss Whedon | 623.3M | The Avengers | English | 8.1 | 220.M | 403.3M |
| 8 | Roger Allers | 422.8M | The Lion King | English | 8.5 | 45.M | 377.8M |
| 9 | George Lucas | 474.5M | Star Wars: Episode I - The Phantom Menace | English | 6.5 | 115.M | 359.5M |
| 10 | Christopher Nolan | 533.3M | The Dark Knight | English | 9 | 185.M | 348.3M |
| 11 | Gary Ross | 408.M | The Hunger Games | English | 7.3 | 78.M | 330.M |



## C) TOP 250 :

To get Top 250 movies from the data we used filter to sort the data according to top imdb score and num_voted_users > 250000. Then create a new column with name rank and find out rank of each film using rank.eq()+countif()-1 to get top 250 imdb rated films

After getting top 250 films we again sort the data using language where non English films are displayed to find top non English films in top 250 imdb list

| movie_title | num_voted_u | Colum | langua | country | imdb_sc | Rank |
|---|---|---|---|---|---|---|
| The Shawshank Redemption | 1689764 | | English | USA | 9.3 | 1 |
| The Godfather | 1155770 | | English | USA | 9.2 | 2 |
| The Dark Knight | 1676169 | | English | USA | 9 | 3 |
| The Godfather: Part II | 790926 | | English | USA | 9 | 4 |
| The Lord of the Rings: The Return of the King | 1215718 | | English | USA | 8.9 | 5 |
| Pulp Fiction | 1324680 | | English | USA | 8.9 | 6 |
| Schindler's List | 865020 | | English | USA | 8.9 | 7 |
| The Good, the Bad and the Ugly | 503509 | | Italian | Italy | 8.9 | 8 |
| Forrest Gump | 1251222 | | English | USA | 8.8 | 9 |
| Star Wars: Episode V - The Empire Strikes Back | 837759 | | English | USA | 8.8 | 10 |
| The Lord of the Rings: The Fellowship of the Ring | 1238746 | | English | New Zealan | 8.8 | 11 |
| Inception | 1468200 | | English | USA | 8.8 | 12 |
| Fight Club | 1347461 | | English | USA | 8.8 | 13 |
| Star Wars: Episode IV - A New Hope | 911097 | | English | USA | 8.7 | 14 |
| The Lord of the Rings: The Two Towers | 1100446 | | English | USA | 8.7 | 15 |
| The Matrix | 1217752 | | English | USA | 8.7 | 16 |
| One Flew Over the Cuckoo's Nest | 680041 | | English | USA | 8.7 | 17 |
| Goodfellas | 728685 | | English | USA | 8.7 | 18 |
| City of God | 533200 | | Portugues | Brazil | 8.7 | 19 |
| Seven Samurai | 229012 | | Japanese | Japan | 8.7 | 20 |
| Saving Private Ryan | 881236 | | English | USA | 8.6 | 21 |
| The Silence of the Lambs | 887467 | | English | USA | 8.6 | 22 |
| Se7en | 1023511 | | English | USA | 8.6 | 23 |
| Interstellar | 928227 | | English | USA | 8.6 | 24 |
| The Usual Suspects | 740918 | | English | USA | 8.6 | 25 |
| American History X | 782437 | | English | USA | 8.6 | 26 |
| Modern Times | 143086 | | English | USA | 8.6 | 27 |
| Spirited Away | 417971 | | Japanese | Japan | 8.6 | 28 |
| The Lion King | 644348 | | English | USA | 8.5 | 29 |
| Raiders of the Lost Ark | 661017 | | English | USA | 8.5 | 30 |
| The Dark Knight Rises | 1144337 | | English | USA | 8.5 | 31 |
| Back to the Future | 732212 | | English | USA | 8.5 | 32 |
| Terminator 2: Judgment Day | 744891 | | English | USA | 8.5 | 33 |
| Gladiator | 982637 | | English | USA | 8.5 | 34 |
| The Green Mile | 782610 | | English | USA | 8.5 | 35 |
| Alien | 563827 | | English | UK | 8.5 | 36 |
| Django Unchained | 955174 | | English | USA | 8.5 | 37 |
| Apocalypse Now | 450676 | | English | USA | 8.5 | 38 |
| The Departed | 873649 | | English | USA | 8.5 | 39 |
| Psycho | 422432 | | English | USA | 8.5 | 40 |
| Memento | 845580 | | English | USA | 8.5 | 41 |
| The Prestige | 844052 | | English | USA | 8.5 | 42 |
| Whiplash | 399138 | | English | USA | 8.5 | 43 |
| The Lives of Others | 259379 | | German | Germany | 8.5 | 44 |
| Children of Heaven | 27882 | | Persian | Iran | 8.5 | 45 |
| Samsara | 22457 | | None | USA | 8.5 | 46 |

| G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|
| The Pianist | 497946 | | English | France | 8.5 | 47 |
| Star Wars: Episode VI - Return of the Jedi | 681857 | | English | USA | 8.4 | 48 |
| American Beauty | 822500 | | English | USA | 8.4 | 49 |
| Aliens | 488537 | | English | USA | 8.4 | 50 |
| WALL·E | 718837 | | English | USA | 8.4 | 51 |
| A Separation | 151812 | | Persian | Iran | 8.4 | 52 |
| Braveheart | 736638 | | English | USA | 8.4 | 53 |
| Reservoir Dogs | 664719 | | English | USA | 8.4 | 54 |
| Oldboy | 356181 | | Korean | South Korea | 8.4 | 55 |
| Requiem for a Dream | 573541 | | English | USA | 8.4 | 56 |
| Das Boot | 168203 | | German | West Germa | 8.4 | 57 |
| Lawrence of Arabia | 192775 | | English | UK | 8.4 | 58 |
| Once Upon a Time in America | 221000 | | English | Italy | 8.4 | 59 |
| Amélie | 534262 | | French | France | 8.4 | 60 |
| Princess Mononoke | 221552 | | Japanese | Japan | 8.4 | 61 |
| Toy Story 3 | 544984 | | English | USA | 8.3 | 62 |
| Inside Out | 345198 | | English | USA | 8.3 | 63 |
| Toy Story | 623757 | | English | USA | 8.3 | 64 |
| The Sting | 175607 | | English | USA | 8.3 | 65 |
| Indiana Jones and the Last Crusade | 515306 | | English | USA | 8.3 | 66 |
| Good Will Hunting | 604904 | | English | USA | 8.3 | 67 |
| Up | 665575 | | English | USA | 8.3 | 68 |
| Unforgiven | 277505 | | English | USA | 8.3 | 69 |
| Batman Begins | 980946 | | English | USA | 8.3 | 70 |
| Inglourious Basterds | 885175 | | English | USA | 8.3 | 71 |
| 2001: A Space Odyssey | 427357 | | English | UK | 8.3 | 72 |
| Amadeus | 270790 | | English | USA | 8.3 | 73 |
| L.A. Confidential | 414219 | | English | USA | 8.3 | 74 |
| Snatch | 600996 | | English | UK | 8.3 | 75 |
| Some Like It Hot | 175196 | | English | USA | 8.3 | 76 |
| Scarface | 537442 | | English | USA | 8.3 | 77 |
| Eternal Sunshine of the Spotless Mind | 666937 | | English | USA | 8.3 | 78 |
| Hoop Dreams | 18980 | | English | USA | 8.3 | 79 |
| Room | 161288 | | English | Ireland | 8.3 | 80 |
| Monty Python and the Holy Grail | 382240 | | English | UK | 8.3 | 81 |
| The Hunt | 170155 | | Danish | Denmark | 8.3 | 82 |
| Metropolis | 111841 | | German | Germany | 8.3 | 83 |
| Downfall | 248354 | | German | Germany | 8.3 | 84 |
| Raging Bull | 235133 | | English | USA | 8.3 | 85 |
| Finding Nemo | 692482 | | English | USA | 8.2 | 86 |
| Gone with the Wind | 215340 | | English | USA | 8.2 | 87 |
| Captain America: Civil War | 272670 | | English | USA | 8.2 | 88 |
| Gran Torino | 561773 | | English | USA | 8.2 | 89 |
| A Beautiful Mind | 610568 | | English | USA | 8.2 | 90 |
| Die Hard | 592582 | | English | USA | 8.2 | 91 |
| How to Train Your Dragon | 485430 | | English | USA | 8.2 | 92 |
| The Bridge on the River Kwai | 149444 | | English | UK | 8.2 | 93 |

| B | H | I | J | K | L | M |
|---|---|---|---|---|---|---|
| Pan's Labyrinth | 467234 | | Spanish | Spain | 8.2 | 94 |
| The Secret in Their Eyes | 131831 | | Spanish | Argentina | 8.2 | 95 |
| The Wolf of Wall Street | 780588 | | English | USA | 8.2 | 96 |
| V for Vendetta | 791783 | | English | USA | 8.2 | 97 |
| Trainspotting | 469561 | | English | UK | 8.2 | 98 |
| On the Waterfront | 100890 | | English | USA | 8.2 | 99 |
| Into the Wild | 426359 | | English | USA | 8.2 | 100 |
| Lock, Stock and Two Smoking Barrels | 414976 | | English | UK | 8.2 | 101 |
| The Big Lebowski | 537419 | | English | USA | 8.2 | 102 |
| Incendies | 80429 | | French | Canada | 8.2 | 103 |
| Blade Runner | 461609 | | English | USA | 8.2 | 105 |
| The Thing | 258078 | | English | USA | 8.2 | 106 |
| Casino | 333542 | | English | USA | 8.2 | 107 |
| Warrior | 332276 | | English | USA | 8.2 | 108 |
| Howl's Moving Castle | 214091 | | Japanese | Japan | 8.2 | 109 |
| The Avengers | 995415 | | English | USA | 8.1 | 110 |
| Deadpool | 479047 | | English | USA | 8.1 | 111 |
| Jurassic Park | 613473 | | English | USA | 8.1 | 112 |
| The Sixth Sense | 704766 | | English | USA | 8.1 | 113 |
| Monsters, Inc. | 585659 | | English | USA | 8.1 | 114 |
| Pirates of the Caribbean: The Curse of the Black Pearl | 809474 | | English | USA | 8.1 | 115 |
| Guardians of the Galaxy | 682155 | | English | USA | 8.1 | 116 |
| The Help | 318955 | | English | USA | 8.1 | 117 |
| Platoon | 291603 | | English | UK | 8.1 | 118 |
| The Martian | 472488 | | English | USA | 8.1 | 119 |
| The Bourne Ultimatum | 491077 | | English | USA | 8.1 | 120 |
| Rocky | 375240 | | English | USA | 8.1 | 121 |
| Gone Girl | 569841 | | English | USA | 8.1 | 122 |
| Butch Cassidy and the Sundance Kid | 152089 | | English | USA | 8.1 | 123 |
| The Imitation Game | 467613 | | English | UK | 8.1 | 124 |
| Million Dollar Baby | 482064 | | English | USA | 8.1 | 125 |
| The Truman Show | 667983 | | English | USA | 8.1 | 126 |
| Groundhog Day | 437418 | | English | USA | 8.1 | 127 |
| No Country for Old Men | 612060 | | English | USA | 8.1 | 128 |
| The Revenant | 406020 | | English | USA | 8.1 | 129 |
| Shutter Island | 786092 | | English | USA | 8.1 | 130 |
| Stand by Me | 271794 | | English | USA | 8.1 | 131 |
| Kill Bill: Vol. 1 | 735784 | | English | USA | 8.1 | 132 |
| 12 Years a Slave | 439176 | | English | USA | 8.1 | 133 |
| Annie Hall | 192940 | | English | USA | 8.1 | 134 |
| Sin City | 656640 | | English | USA | 8.1 | 135 |
| The Grand Budapest Hotel | 475518 | | English | USA | 8.1 | 136 |
| The Terminator | 600266 | | English | UK | 8.1 | 137 |
| Spotlight | 195333 | | English | USA | 8.1 | 138 |
| The Best Years of Our Lives | 40359 | | English | USA | 8.1 | 139 |
| The Wizard of Oz | 291875 | | English | USA | 8.1 | 140 |
| There Will Be Blood | 372990 | | English | USA | 8.1 | 141 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Prisoners | 383591 | | English | USA | 8.1 | 142 |
| The Princess Bride | 294163 | | English | USA | 8.1 | 143 |
| Woodstock | 12631 | | English | USA | 8.1 | 144 |
| Hotel Rwanda | 264533 | | English | UK | 8.1 | 145 |
| Mad Max: Fury Road | 552503 | | English | Australia | 8.1 | 146 |
| Amores Perros | 173551 | | Spanish | Mexico | 8.1 | 147 |
| Before Sunrise | 183288 | | English | USA | 8.1 | 148 |
| The Celebration | 65951 | | Danish | Denmark | 8.1 | 149 |
| Donnie Darko | 580999 | | English | USA | 8.1 | 151 |
| Elite Squad | 81644 | | Portugues | Brazil | 8.1 | 152 |
| The Sea Inside | 64556 | | Spanish | Spain | 8.1 | 153 |
| Rush | 312629 | | English | UK | 8.1 | 154 |
| Tae Guk Gi: The Brotherhood of War | 31943 | | Korean | South Korea | 8.1 | 155 |
| Akira | 106160 | | Japanese | Japan | 8.1 | 156 |
| Jaws | 412454 | | English | USA | 8 | 157 |
| The Exorcist | 284252 | | English | USA | 8 | 158 |
| Aladdin | 260939 | | English | USA | 8 | 159 |
| The Incredibles | 479166 | | English | USA | 8 | 160 |
| Dances with Wolves | 186485 | | English | USA | 8 | 161 |
| The Sound of Music | 148172 | | English | USA | 8 | 162 |
| Rain Man | 383784 | | English | USA | 8 | 163 |
| Slumdog Millionaire | 641997 | | English | UK | 8 | 164 |
| The King's Speech | 503631 | | English | UK | 8 | 165 |
| Catch Me If You Can | 525801 | | English | USA | 8 | 166 |
| Star Trek | 504419 | | English | USA | 8 | 167 |
| The Pursuit of Happyness | 338383 | | English | USA | 8 | 168 |
| Doctor Zhivago | 55816 | | English | USA | 8 | 169 |
| Black Swan | 551363 | | English | USA | 8 | 170 |
| District 9 | 531737 | | English | South Africa | 8 | 171 |
| Young Frankenstein | 112671 | | English | USA | 8 | 172 |
| Dead Poets Society | 277451 | | English | USA | 8 | 173 |
| Mystic River | 338415 | | English | USA | 8 | 174 |
| Ratatouille | 473887 | | English | USA | 8 | 175 |
| Fiddler on the Roof | 29839 | | English | USA | 8 | 176 |
| Kill Bill: Vol. 2 | 512749 | | English | USA | 8 | 177 |
| X-Men: Days of Future Past | 514125 | | English | USA | 8 | 178 |
| JFK | 113472 | | English | France | 8 | 179 |
| The Artist | 190030 | | English | France | 8 | 180 |
| Sling Blade | 72443 | | English | USA | 8 | 181 |
| Dallas Buyers Club | 326494 | | English | USA | 8 | 182 |
| Boyhood | 266020 | | English | USA | 8 | 183 |
| Bowling for Columbine | 123090 | | English | Germany | 8 | 184 |
| Casino Royale | 470483 | | English | UK | 8 | 185 |
| Casino Royale | 470501 | | English | UK | 8 | 186 |
| Sicko | 66610 | | English | USA | 8 | 187 |
| Shaun of the Dead | 395921 | | English | UK | 8 | 188 |
| Life of Pi | 440084 | | English | USA | 8 | 189 |

| | | | | | | |
|---|---|---|---|---|---|---|
| The Perks of Being a Wallflower | 351274 | | English | USA | 8 | 190 |
| A Fistful of Dollars | 147566 | | Italian | Italy | 8 | 191 |
| Before Sunset | 168398 | | English | USA | 8 | 192 |
| Central Station | 28951 | | Portugues | Brazil | 8 | 193 |
| Her | 355126 | | English | USA | 8 | 194 |
| Waltz with Bashir | 46107 | | Hebrew | Israel | 8 | 195 |
| True Romance | 163492 | | English | USA | 8 | 196 |
| Persepolis | 70194 | | French | France | 8 | 197 |
| Big Fish | 350698 | | English | USA | 8 | 198 |
| The Straight Story | 63733 | | English | France | 8 | 199 |
| Brazil | 152306 | | English | UK | 8 | 200 |
| In Bruges | 307639 | | English | UK | 8 | 201 |
| Mulholland Drive | 235992 | | English | France | 8 | 202 |
| My Name Is Khan | 69759 | | Hindi | India | 8 | 203 |
| Dancer in the Dark | 79330 | | English | Denmark | 8 | 204 |
| Magnolia | 241030 | | English | USA | 8 | 205 |
| Serenity | 242599 | | English | USA | 8 | 206 |
| Cinderella Man | 148238 | | English | USA | 8 | 207 |
| Blood In, Blood Out | 23181 | | English | USA | 8 | 208 |
| Blood Diamond | 400292 | | English | Germany | 8 | 209 |
| The Iron Giant | 128455 | | English | USA | 8 | 210 |
| Avatar | 886204 | | English | USA | 7.9 | 211 |
| E.T. the Extra-Terrestrial | 281842 | | English | USA | 7.9 | 212 |
| Shrek | 467113 | | English | USA | 7.9 | 213 |
| Iron Man | 696338 | | English | USA | 7.9 | 214 |
| Toy Story 2 | 385871 | | English | USA | 7.9 | 215 |
| Straight Outta Compton | 119928 | | English | USA | 7.9 | 216 |
| Taken | 483756 | | English | France | 7.9 | 217 |
| Crouching Tiger, Hidden Dragon | 217740 | | Mandarin | Taiwan | 7.9 | 218 |
| Walk the Line | 188637 | | English | USA | 7.9 | 219 |
| The Fighter | 275869 | | English | USA | 7.9 | 220 |
| The Bourne Identity | 407601 | | English | USA | 7.9 | 221 |
| Big Hero 6 | 279093 | | English | USA | 7.9 | 222 |
| My Fair Lady | 66959 | | English | USA | 7.9 | 223 |
| Captain Phillips | 323353 | | English | USA | 7.9 | 224 |
| Little Miss Sunshine | 355810 | | English | USA | 7.9 | 225 |
| The Untouchables | 219008 | | English | USA | 7.9 | 226 |
| Crash | 361169 | | English | Germany | 7.9 | 227 |
| Halloween | 157857 | | English | USA | 7.9 | 228 |
| Halloween | 157863 | | English | USA | 7.9 | 229 |
| Edward Scissorhands | 357581 | | English | USA | 7.9 | 230 |
| The Hobbit: The Desolation of Smaug | 483540 | | English | USA | 7.9 | 231 |
| How to Train Your Dragon 2 | 221128 | | English | USA | 7.9 | 232 |
| The Blues Brothers | 142448 | | English | USA | 7.9 | 233 |
| Nightcrawler | 293304 | | English | USA | 7.9 | 234 |
| Do the Right Thing | 59524 | | English | USA | 7.9 | 235 |
| The Wrestler | 251349 | | English | USA | 7.9 | 236 |

| | | | | | | |
|---|---|---|---|---|---|---|
| The Wrestler | 251349 | | English | USA | 7.9 | 236 |
| Hot Fuzz | 352695 | | English | UK | 7.9 | 237 |
| The Remains of the Day | 45703 | | English | UK | 7.9 | 238 |
| Boogie Nights | 189032 | | English | USA | 7.9 | 239 |
| The Hateful Eight | 272839 | | English | USA | 7.9 | 240 |
| Once | 90827 | | English | Ireland | 7.9 | 241 |
| Glory | 101888 | | English | USA | 7.9 | 242 |
| Glory | 101889 | | English | USA | 7.9 | 243 |
| Before Midnight | 95362 | | English | USA | 7.9 | 244 |
| 4 Months, 3 Weeks and 2 Days | 44763 | | Romanian | Romania | 7.9 | 245 |
| Moon | 260607 | | English | UK | 7.9 | 246 |
| Nine Queens | 38215 | | Spanish | Argentina | 7.9 | 247 |
| The Chorus | 44151 | | French | France | 7.9 | 248 |
| The Second Mother | 7025 | | Portugues | Brazil | 7.9 | 249 |
| Letters from Iwo Jima | 132149 | | Japanese | USA | 7.9 | 250 |

| Column1 | num_voted_user | language | country | imdb_score | Rank |
|---|---|---|---|---|---|
| The Good, the Bad and the Ugly | 503509 | Italian | Italy | 8.9 | 8 |
| City of God | 533200 | Portuguese | Brazil | 8.7 | 19 |
| Seven Samurai | 229012 | Japanese | Japan | 8.7 | 20 |
| Spirited Away | 417971 | Japanese | Japan | 8.6 | 28 |
| The Lives of Others | 259379 | German | Germany | 8.5 | 44 |
| Children of Heaven | 27882 | Persian | Iran | 8.5 | 45 |
| A Separation | 151812 | Persian | Iran | 8.4 | 52 |
| Oldboy | 356181 | Korean | South Korea | 8.4 | 55 |
| Das Boot | 168203 | German | West Germany | 8.4 | 57 |
| Amélie | 534262 | French | France | 8.4 | 60 |
| Princess Mononoke | 221552 | Japanese | Japan | 8.4 | 61 |
| The Hunt | 170155 | Danish | Denmark | 8.3 | 82 |
| Metropolis | 111841 | German | Germany | 8.3 | 83 |
| Downfall | 248354 | German | Germany | 8.3 | 84 |
| Pan's Labyrinth | 467234 | Spanish | Spain | 8.2 | 94 |
| The Secret in Their Eyes | 131831 | Spanish | Argentina | 8.2 | 95 |
| Incendies | 80429 | French | Canada | 8.2 | 103 |
| Howl's Moving Castle | 214091 | Japanese | Japan | 8.2 | 109 |
| Amores Perros | 173551 | Spanish | Mexico | 8.1 | 147 |
| The Celebration | 65951 | Danish | Denmark | 8.1 | 149 |
| Elite Squad | 81644 | Portuguese | Brazil | 8.1 | 152 |
| The Sea Inside | 64556 | Spanish | Spain | 8.1 | 153 |
| Tae Guk Gi: The Brotherhood of War | 31943 | Korean | South Korea | 8.1 | 155 |
| Akira | 106160 | Japanese | Japan | 8.1 | 156 |
| A Fistful of Dollars | 147566 | Italian | Italy | 8 | 191 |
| Central Station | 28951 | Portuguese | Brazil | 8 | 193 |
| Waltz with Bashir | 46107 | Hebrew | Israel | 8 | 195 |
| Persepolis | 70194 | French | France | 8 | 197 |
| My Name Is Khan | 69759 | Hindi | India | 8 | 203 |
| Crouching Tiger, Hidden Dragon | 217740 | Mandarin | Taiwan | 7.9 | 218 |
| 4 Months, 3 Weeks and 2 Days | 44763 | Romanian | Romania | 7.9 | 245 |
| Nine Queens | 38215 | Spanish | Argentina | 7.9 | 247 |
| The Chorus | 44151 | French | France | 7.9 | 248 |
| Letters from Iwo Jima | 132149 | Japanese | USA | 7.9 | 250 |

## D) BEST DIRECTORS:

To find the best directors we first created a pivot table using Director name and imdb score. We assigned director names to display in row and for summation we gave imdb score instructing it to give average imdb score for every director . Finally we sorted the pivoted table using Top 10 sorting option to get Top 10 Best Directors from imdb data

| 3 | Top Directors | Average of imdb_score |
|---|---|---|
| 4 | Akira Kurosawa | 8.7 |
| 5 | Alfred Hitchcock | 8.5 |
| 6 | Asghar Farhadi | 8.4 |
| 7 | Charles Chaplin | 8.6 |
| 8 | Christopher Nolan | 8.425 |
| 9 | Damien Chazelle | 8.5 |
| 10 | Majid Majidi | 8.5 |
| 11 | Richard Marquand | 8.4 |
| 12 | Ron Fricke | 8.5 |
| 13 | Sergio Leone | 8.433333333 |
| 14 | Tony Kaye | 8.6 |
| 15 | **Grand Total** | **8.47** |

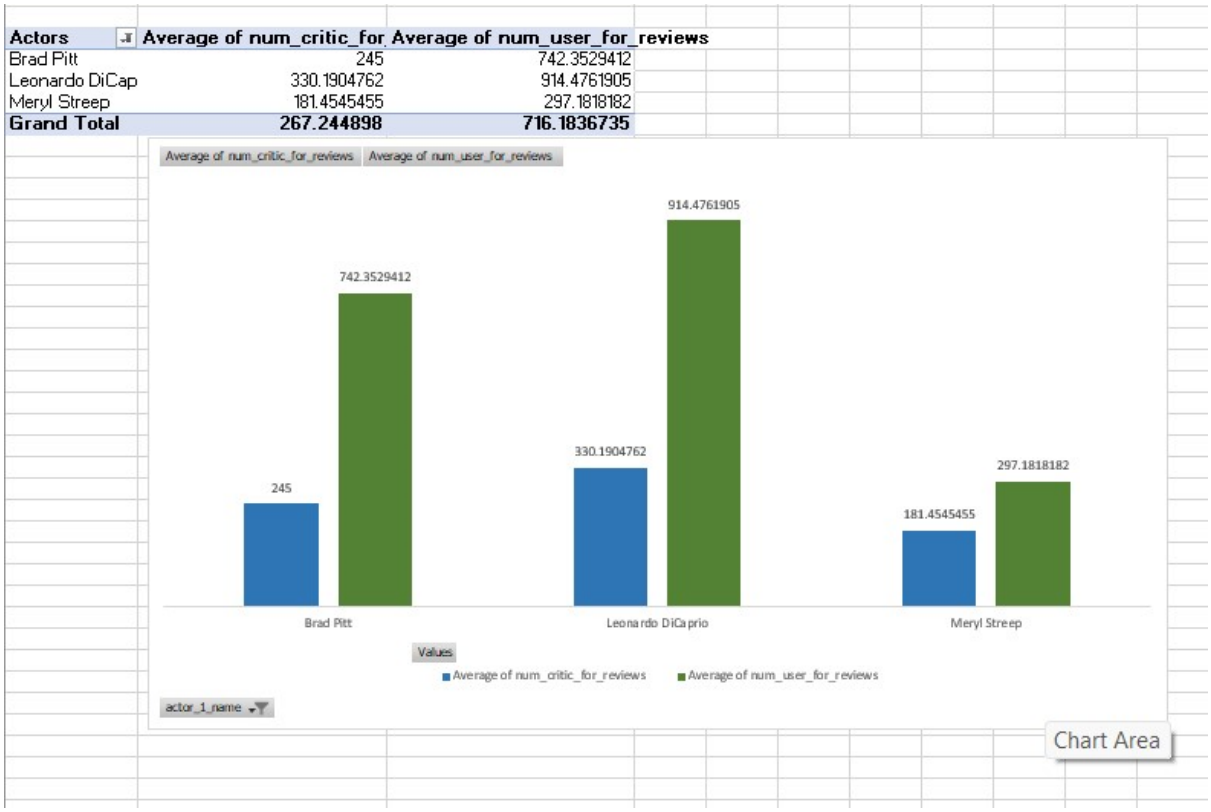
## E) POPULAR GENRES:

In the data every movie has multiple genres associated with it . It is inappropriate to directly use pivot table . We first separate generes using in built text to columns function in excel which generates multiple genre in different columns . We give each genre name as genre 1 , genre 2 ,etc. For analysis purpose we use genre 1 as our main genre . Using genre(main) we build pivot table where we give Genre(main) to rows and summation of imdb scores to find out popular genre . We find out that Comedy , Action and Drama are popular genres based on imdb scores .

| Genres | Sum of imdb_score |
|---|---|
| Action | 5981.7 |
| Adventure | 2400.4 |
| Animation | 303.3 |
| Biography | 1460.4 |
| Comedy | 6064.2 |
| Crime | 1754.9 |
| Documentary | 176.7 |
| Drama | 4502.2 |
| Family | 19.5 |
| Fantasy | 232.4 |
| Horror | 928.9 |
| Musical | 13.5 |
| Mystery | 153 |
| Romance | 7.1 |
| Sci-Fi | 46.4 |
| Thriller | 4.8 |
| Western | 16.2 |
| **Grand Total** | **24065.6** |

## F) CHARTS:

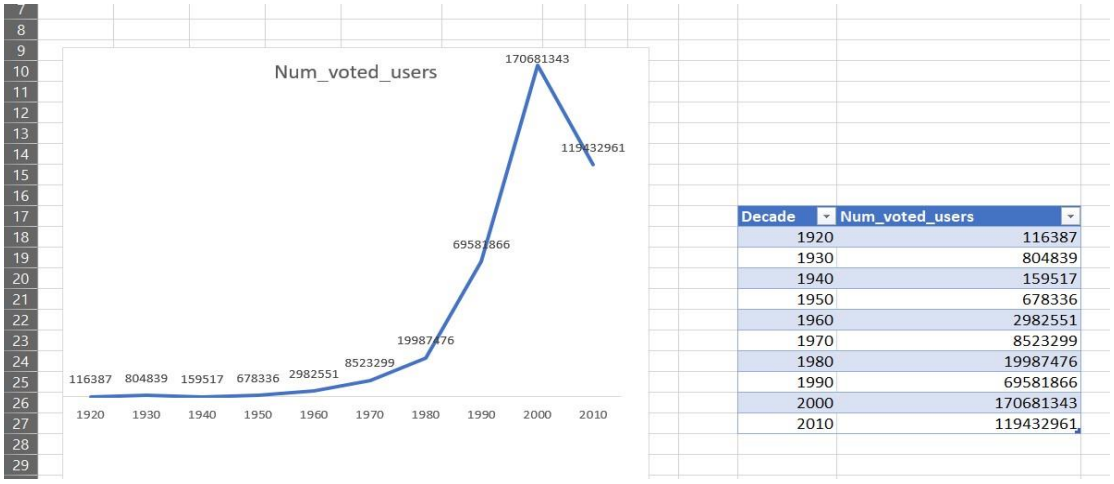
To find critic and audience favourite actor we consider means of critic reviews and user reviews . For this analysis we first create pivot table with fields like actor_1_name , num_crirtic_review and num_user_reviews . As required in the insights we find out means of critic and user reviews for Brad Pitt, Merly Stepp and Leonardo Di Caprio by filter the actor_name field and prepare a bar chart to get better visualization to determine who is critic and audience favourite actor . We find out that Leonardo Di Caprio is the Critic and Audience Favourite Actor out of all.



| Actors | Average of num_critic_for | Average of num_user_for_reviews |
|---|---|---|
| Brad Pitt | 245 | 742.3529412 |
| Leonardo DiCap | 330.1904762 | 914.4761905 |
| Meryl Streep | 181.4545455 | 297.1818182 |
| **Grand Total** | **267.244898** | **716.1836735** |



Another task in this was to prepare a chart for num_voted_users for decades starting from 1920 – 2010 . We achieved this task by using pivot tables with fileds like title_year and num_voted_users .We added using sum function num_voted users separately for every decade and stored them in different table for analysis.





| Decade | Num_voted_users |
|---|---|
| 1920 | 116387 |
| 1930 | 804839 |
| 1940 | 159517 |
| 1950 | 678336 |
| 1960 | 2982551 |
| 1970 | 8523299 |
| 1980 | 19987476 |
| 1990 | 69581866 |
| 2000 | 170681343 |
| 2010 | 119432961 |

From the chart we can visualize that over the passing decades from 1920 more people started watching and voting for films with Decade 2000 being the decade with most num_voted _users .

**Dataset attached :**

https://docs.google.com/spreadsheets/d/1UZYvjRvoc0T4jexBQipWtv5B-pomEKHm/edit?usp=sharing&ouid=101908371018515181983&rtpof=true&sd=true

**Results :**

During this project I have learned how a data analyst works in real life . I also got experience to handle vast amounts of data and perform desired analysis