

Project on

IMDB Movie Analysis

PROJECT DESCRIPTION- This project is about the IMDB Movie analysis.

During the project I gain some enormous knowledge and learned some statistical & analytical process for tracking to learned how we extract some meaningful insight from it.

Approach- I take Microsoft Excel platform for deriving some business insight to help my firm. Observe all the tables, columns, rows. Afterwards check all the tables' content carefully. Then one by one I have executed queries according to the questions asked. I used my analytical thinking and some own personal experience for making this project.

Tech -stack used-I used latest version of Microsoft Excel sheets to perform statistical analysis on this dataset because allows users to edit, organize, and analyze different types of information and also used NOTEPAD++ for this project.

Insight- This project provides me immense knowledge about the job role of data analytics, how a data analytics should work to derive the meaningful business insights for their company. I learn the responsibility and importance of this analytics process for a company for their business growth.

Result-This project give me a futuristics view about the responsibility and importance role of data analytics, definitely it will help me in future for my working career in a company for the role of data analytics.

A . Your task: Clean the data:

Solution: In the 'L' Column there is need for data cleaning so I used "FLASH FILL" Function.

B. Find the movies with the highest profit?

Formula: first I calculate the profit by =(gross cell num – budget cell num).

Then I use filter function by using sort function in it to extract top 10 movie with highest profit

Solution: Here are some top 10 movies with highest profit .

movie_title	profit
1.Avatar	523505847
2.Jurassic World	502177271
3.Titanic	458672302
4.Star Wars: Episode IV - A New Hope	449935665
5.E.T. the Extra-Terrestrial	424449459
6.The Avengers	403279547
7.The Avengers	403279547
8.The Lion King	377783777
9.Star Wars: Episode I - The Phantom Menace	359544677
10.The Dark Knight	348316061

C. Find IMDB Top 250

FORMULA: I used sort and filter function.

RANK IMDB_TOP_250

- 1 The Shawshank Redemption
- 2 The Godfather
- 3 The Dark Knight
- 4 The Godfather: Part
- 5 Fargo
- 6 The Lord of the Rings: The Return of the King
- 7 Schindler's List
- 8 Pulp Fiction
- 9 The Good, the Bad and the Ugly
- 10 12 Angry Men
- 11 Inception
- 12 The Lord of the Rings: The Fellowship of the Ring
- 13 Daredevil
- 14 Fight Club
- 15 Forrest Gump
- 16 It's Always Sunny in Philadelphia
- 17 Star Wars: Episode V - The Empire Strikes Back
- 18 The Lord of the Rings: The Two Towers
- 19 The Matrix
- 20 Friday Night Lights
- 21 Goodfellas
- 22 Star Wars: Episode IV - A New Hope
- 23 One Flew Over the Cuckoo's Nest
- 24 City of God
- 25 Seven Samurai
- 26 Interstellar
- 27 Hannibal
- 28 Saving Private Ryan
- 29 Luther
- 30 Se7en
- 31 The Silence of the Lambs
- 32 Spirited Away
- 33 Spartacus: War of the Damned
- 34 American History
- 35 The Usual Suspects
- 36 Once Upon a Time in the West
- 37 It's a Wonderful Life

- 38 Modern Times
- 39 Casablanca
- 40 The Dark Knight Rises
- 41 Gladiator
- 42 Terminator 2: Judgment Day
- 43 Django Unchained
- 44 The Departed
- 45 The Lion King
- 46 The Green Mile
- 47 Outlander
- 48 The Prestige
- 49 The Pianist
- 50 Apocalypse Now
- 51 Entourage
- 52 Raiders of the Lost Ark
- 53 Psycho
- 54 Back to the Future
- 55 Alien
- 56 Memento
- 57 Airlift
- 58 Whiplash
- 59 The Lives of Others
Dr. Strangelove or: How I Learned to Stop Worrying and Love the
- 60 Bomb
- 61 Children of Heaven
- 62 WALLÂ.
- 63 Braveheart
- 64 Stargate
- 65 AmÃ©lie
- 66 Baahubali: The Beginning
- 67 Star Wars: Episode VI - Return of the Jedi
- 68 Once Upon a Time in America
- 69 Princess Mononoke
- 70 The Shining
- 71 Aliens
- 72 American Beauty
- 73 Lawrence of Arabia
- 74 Das Boot

- 75 Veronica Mars
- 76 The Inbetweeners
- 77 Rang De Basanti
- 78 Psych
- 79 Requiem for a Dream
- 80 M*A*S*
- 81 Batman: The Dark Knight Returns, Part
- 82 Oldboy
- 83 To Kill a Mockingbird
- 84 Reservoir Dogs
- 85 A Separation
- 86 Toy Story
- 87 Up
- 88 Inside Out
- 89 Batman Begins
- 90 Life
- 91 Inglourious Basterds
- 92 Indiana Jones and the Last Crusade
- 93 L.A. Confidential
- 94 Toy Story
- 95 Scarface
- 96 Eternal Sunshine of the Spotless Mind
- 97 Amadeus
- 98 Raging Bull
- 99 Metropolis
- 100 Unforgiven
- 101 Downfall
- 102 Room
- 103 Snatch
- 104 2001: A Space Odyssey
- 105 Good Will Hunting
- 106 Snatch
- 107 The Sting
- 108 The Great Escape
- 109 The Hunt
- 110 The Apartment
- 111 Judgment at Nuremberg
- 112 Some Like It Hot

- 113 Singin' in the Rain
- 114 Inside Job
- 115 Taxi Driver
- 116 Monty Python and the Holy Grail
- 117 Captain America: Civil War
- 118 How to Train Your Dragon
- 119 The Wolf of Wall Street
- 120 Finding Nemo
- 121 A Beautiful Mind
- 122 Casino
- 123 V for Vendetta
- 124 The Thing
- 125 Die Hard
- 126 Blade Runner
- 127 Gran Torino
- 128 Warrior
- 129 Howl's Moving Castle
- 130 Into the Wild
- 131 Pan's Labyrinth
- 132 The Deer Hunter
- 133 The Big Lebowski
- 134 Buffy the Vampire Slayer
- 135 Incendies
- 136 The Elephant Man
- 137 Gone with the Wind
- 138 The Secret in Their Eyes
- 139 Trainspotting
- 140 The Bridge on the River Kwai
- 141 Lage Raho Munna Bhai
- 142 Mr. Smith Goes to Washington
- 143 Lock, Stock and Two Smoking Barrels
- 144 Rebecca
- 145 On the Waterfront
- 146 It Happened One Night
- 147 The Avengers
- 148 Guardians of the Galaxy
- 149 Mad Max: Fury Road
- 150 The Revenant

- 151 The Bourne Ultimatum
- 152 Pirates of the Caribbean: The Curse of the Black Pearl
- 153 Monsters, Inc
- 154 The Martian
- 155 Shutter Island
- 156 Jurassic Park
- 157 Gone Girl
- 158 The Truman Show
- 159 The Avengers
- 160 Deadpool
- 161 Kill Bill: Vol
- 162 Solaris
- 163 Prisoners
- 164 The Sixth Sense
- 165 Sin City
- 166 Rush
- 167 The Grand Budapest Hotel
- 168 Million Dollar Baby
- 169 The Help
- 170 No Country for Old Men
- 171 There Will Be Blood
- 172 Gandhi
- 173 12 Years a Slave
- 174 Spotlight
- 175 Hotel Rwanda
- 176 Hachi: A Dog's Tale
- 177 The Imitation Game
- 178 The Princess Bride
- 179 Groundhog Day
- 180 The Sea Inside
- 181 Tae Guk Gi: The Brotherhood of War
- 182 Barry Lyndon
- 183 Stand by Me
- 184 Akira
- 185 Elite Squad
- 186 The Terminator
- 187 Platoon
- 188 Butch Cassidy and the Sundance Kid

- 189 [Donnie Darko](#)
- 190 [Annie Hall](#)
- 191 [Network](#)
- 192 [A Christmas Story](#)
- 193 [The Man Who Shot Liberty Valance](#)
- 194 [Cat on a Hot Tin Roof](#)
- 195 [The Wizard of Oz](#)
- 196 [Before Sunrise](#)
- 197 [The Best Years of Our Lives](#)
- 198 [Amores Perros](#)
- 199 [Touching the Void](#)
- 200 [The Celebration](#)
- 201 [Rocky](#)
- 202 [High Noon](#)
- 203 [X-Men: Days of Future Past](#)
- 204 [Ratatouille](#)
- 205 [Star Trek](#)
- 206 [Life of Pi](#)
- 207 [Casino Royale](#)
- 208 [Blood Diamond](#)
- 209 [The Incredibles](#)
- 210 [Cinderella Man](#)
- 211 [Big Fish](#)
- 212 [The Pursuit of Happyness](#)
- 213 [Kill Bill: Vol](#)
- 214 [Catch Me If You Can](#)
- 215 [The Iron Giant](#)
- 216 [JFK](#)
- 217 [Serenity](#)
- 218 [Magnolia](#)
- 219 [District](#)
- 220 [Mystic River](#)
- 221 [Aladdin](#)
- 222 [Rain Man](#)
- 223 [Her](#)
- 224 [Dances with Wolves](#)
- 225 [Dead Poets Society](#)
- 226 [The Artist](#)

- 227 The King's Speech
- 228 Brazil
- 229 In Bruges
- 230 Mulholland Drive
- 231 The Return
- 232 Slumdog Millionaire
- 233 The Diving Bell and the Butterfly
- 234 Black Swan
- 235 The Perks of Being a Wallflower
- 236 True Romance
- 237 Dancer in the Dark
- 238 The Exorcist
- 239 Jaws
- 240 Patton
- 241 Casino Royale
- 242 Doctor Zhivago
- 243 The Straight Story
- 244 Fiddler on the Roof
- 245 Sicko
- 246 My Name Is Khan
- 247 The Sound of Music
- 248 Persepolis
- 249 The Wild Bunch
- 250.Dallas Buyers Club

Top_Foreign_Lang_Film.

IMDB_TOP_250

A Separation

Deadpool

language

Japanese

Japanese

D. Find the best directors

FORMULA: I USED SORT AND FILTER FUNCTION

NOTE: I ONLY CONSIDER THOSE DIRECTOR WHOSE MOVIES GOT MORE THAN 25000 USERS VOTING.BECAUSE NUM OF VOTING IS ALSO AN IMPORTANT PARAMETER.

TOP_TEN_DIRECTOR_ACCORDING_TO_IMDB_SCORE

1. Frank Darabont
2. Francis Ford Coppola
3. Christopher Nolan
4. Francis Ford Coppola
5. Peter Jackson
6. Steven Spielberg
7. Quentin Tarantino
8. Sergio Leone
9. Sidney Lumet
10. Christopher Nolan

E. Find popular genres:

NOTE: According to me the movies who got highest collection and also got highest IMDB rating with maximum voting by Public is most liked genres movies because in both the scenario it is totally depends on audience that how much they liked that particular movie which belong to that particular genre.

So I make two list one based on collection and another based on IMDB.

A. Top 10 most popular genres movies based on Collection.

Movie_title	Genres
1. Avatar	Action Adventure Fantasy Sci-Fi
2. Jurassic World	Action Adventure Sci-Fi Thriller

3.Titanic	Drama Romance
4.Star Wars: Episode IV - A New Hope	Action Adventure Fantasy Sci-Fi
5.E.T. the Extra-Terrestrial	Family Sci-Fi
6.The Avengers	Action Adventure Sci-Fi
7.The Avengers	Action Adventure Sci-Fi
8.The Lion King	Adventure Animation Drama Family Musical
9.Star Wars: Episode I - The Phantom Menace	Action Adventure Fantasy Sci-Fi
10.The Dark Knight	Action Crime Drama Thriller

B. Top 10 most popular genres movies based on IMDB Rating.

Movie_title	Genres
1.The Shawshank Redemption	Crime Drama
2.The Godfather	Crime Drama
3.The Dark Knight	Action Crime Drama Thriller
4.The Godfather: Part	Crime Drama
5.Fargo	Crime Drama Thriller
6.The Lord of the Rings: The Return of the King	Action Adventure Drama Fantasy
7.Pulp Fiction	Crime Drama
8.Schindler's List	Biography Drama History
9.The Good, the Bad and the Ugly	Western
10.12 Angry Men	Crime Drama

F.(1). Group the combined column using the actor_1_name column.

Actors	Count of Movies as a Lead actor
1.Brad Pitt	18
2.Leonardo DiCaprio	21
3.Meryl Streep	13
Grand Total	52

**F.(2). Find the Mean of
the num_critic_for_reviews and num_users_for_review and identify
the actors which have the highest mean.**

A. Mean for num_critic_for_review.

1. FOR BRAD PITT. = 231.94

2. LEONARDO DICAPRIO= 330.19

3. MERLY STREEP= 163.153

= Hence **LEONARDO DICAPRIO** has highest Mean for
num_critic_for_review.

B. Mean for num_users_for_review.

1. FOR BRAD PITT. = 702.44

2. LEONARDO DICAPRIO= 914.47

3. MERLY STREEP= 257.30

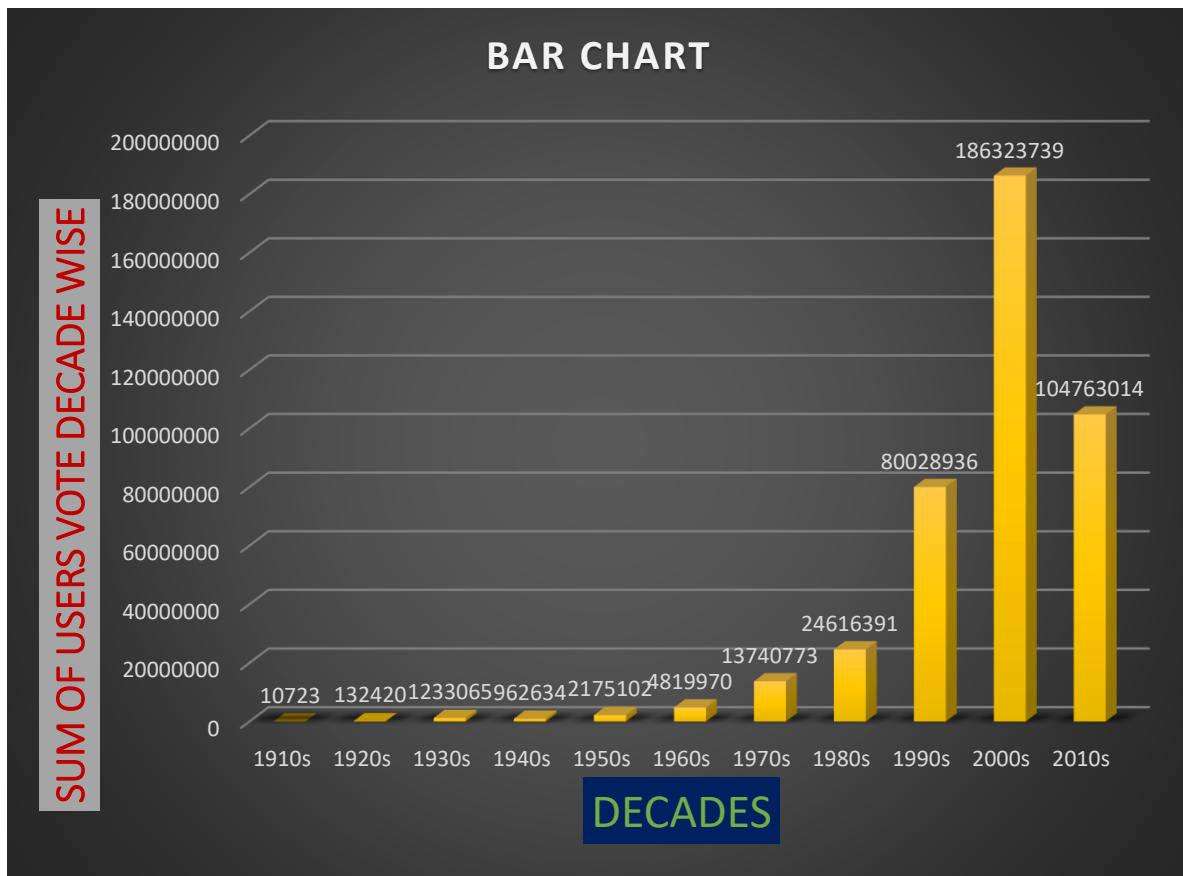
=Hence **LEONARDO DICAPRIO** has highest Mean for
num_user_for_review

FORMULA FOR MEAN- **1. =SUM(CELL NUM:CELL NUM)/COUNT(CELL
NUM:CELL NUM)**

2. AVERAGE(CELL NUM:CELL NUM)

**F(3). Observe the change in number of voted users over
decades using a bar chart.**

FORMULA: I USED FILTER FUNCTION AND THEN MAKE A
BAR CHART



F4. Find the critic-favorite and audience-favorite actors

critic-favorite actor : some critic fav actors who got maximum review by them for there movies are:

- 1.Tom Hardy
 - 2.Michael Fassbender
 - 3.Leonardo DiCaprio
 - 4.Albert Finney
 - 5.Albert Finney
 - 6.Tom Hardy
 - 7.Phaldut Sharma
 - 8.Henry Cavill
 - 9.CCH Pounder
 - 10.Matthew McConaughey
-

Audience fav actors are: who got max num of audience vote as well as max num of audience review.

- 1.Christopher Lee**
- 2.Christian Bale**
- 3.Morgan Freeman**
- 4.Keanu Reeves**
- 5.Natalie Portman**
- 6.Heather Donahue**
- 7.Orlando Bloom**
- 8.CCH Pounder**
- 9.Henry Cavil**
- 10.Brad Pitt**