

## IMDB Movie Analysis

### Project Description –

Movie Analysis is one of the most important functions for a film distributing company in order to broadcast appropriate movies and gain huge number of profits and minimal loss.

In this project, I have used the 5 why approach to find solutions for various problems.

One of them (for eg) is:

**A VOD platform company incurs a huge amount of loss due to broadcasting some of the movies.**

#### 5 why approach:

Q. Why did the VOD company incur huge loss?

A. Because the movies broadcasted were not seen / subscribed by many people.

Q. Why were the movies not seen by many people?

A. Because the movies did not meet the expectations of the public.

Q. Why did the movies not meet the expectations of the public?

A. Because the director and actor of the movies were not liked by much people.

Q. Why did not company not know that those movies were not liked by much people?

A. Because the ratings and score of those movies were miscalculated.

Q. Why were the ratings and score miscalculated?

A. Because of the ineffectiveness of the data analyst team.

As we can see, wrong assumptions made by the data analyst team can lead to huge loss for a company which is difficult to oversee and overcome.

### Approach –

As a Data Analyst, it is my duty to perform analysis to find out the following which would be helpful for the IMDB Movie Analysis:

1. Movies with highest profit
2. Top movies (250)
3. Best directors
4. Popular genres
5. Critic favourite and audience favourite actors

For finding all these details, we must clean the collected dataset that is used for the analysis. Cleaning includes: deleting unwanted columns, removing rows with null values, removing duplicates etc.

## Tech-Stack used –

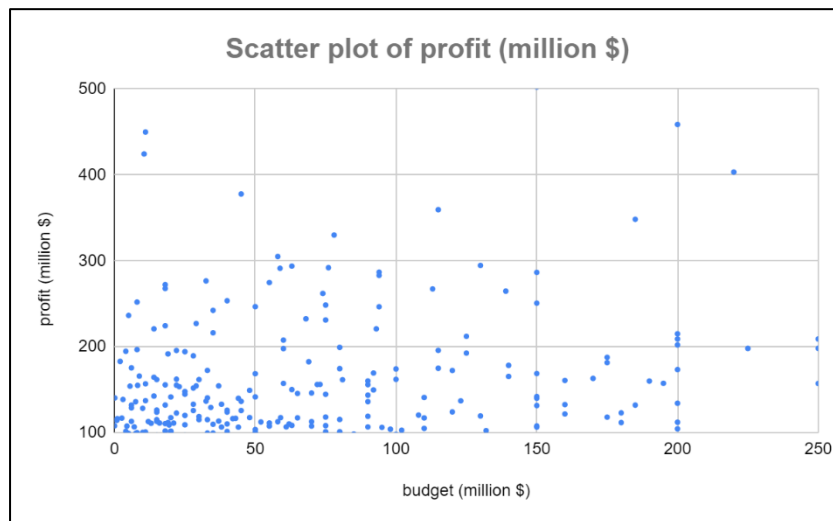
I have used **Google Spreadsheet** for this project.

Spreadsheet makes it easy to extract and visualize the data and interpret the results of the queries.

## Insights –

After cleaning the provided dataset,I have derived the following information/insights using Excel:

### 1. Movies with highest profit:



Here is a scatter plot that plots budget in million \$ against profit in million \$.

The top 10 movies with highest profit are given in the table:

| movie_title                               | gross (million \$) | budget (million \$) | profit (million \$) |
|---|--------------------|---------------------|---------------------|
| Avatar                                    | 760.5              | 237                 | 523.5               |
| Jurassic World                            | 652.2              | 150                 | 502.2               |
| Titanic                                   | 658.7              | 200                 | 458.7               |
| Star Wars: Episode IV - A New Hope        | 460.9              | 11                  | 449.9               |
| E.T. the Extra-Terrestrial                | 434.9              | 10.5                | 424.4               |
| The Avengers                              | 623.3              | 220                 | 403.3               |
| The Lion King                             | 422.8              | 45                  | 377.8               |
| Star Wars: Episode I - The Phantom Menace | 474.5              | 115                 | 359.5               |
| The Dark Knight                           | 533.3              | 185                 | 348.3               |
| The Hunger Games                          | 408                | 78                  | 330                 |

## 2. IMDB Top 250 Movies:

The top 250 movies are given in the table below:

| IMDb_Top_250                                      | Rank |
|---|------|
| The Shawshank Redemption                          | 1    |
| The Godfather                                     | 2    |
| The Dark Knight                                   | 3    |
| The Godfather: Part II                            | 4    |
| The Lord of the Rings: The Return of the King     | 5    |
| Schindler's List                                  | 6    |
| Pulp Fiction                                      | 7    |
| The Good, the Bad and the Ugly                    | 8    |
| Inception   | 9    |
| The Lord of the Rings: The Fellowship of the Ring | 10   |
| Fight Club  | 11   |
| Forrest Gump                                      | 12   |
| Star Wars: Episode V - The Empire Strikes Back    | 13   |
| The Lord of the Rings: The Two Towers             | 14   |
| The Matrix  | 15   |
| Goodfellas  | 16   |
| Star Wars: Episode IV - A New Hope                | 17   |
| One Flew Over the Cuckoo's Nest                   | 18   |
| City of God                                       | 19   |
| Seven Samurai                                     | 20   |
| Interstellar                                      | 21   |
| Saving Private Ryan                               | 22   |
| Se7en   | 23   |
| The Silence of the Lambs                          | 24   |
| Spirited Away                                     | 25   |
| American History X                                | 26   |
| The Usual Suspects                                | 27   |
| Modern Times                                      | 28   |
| The Dark Knight Rises                             | 29   |
| Gladiator   | 30   |
| Terminator 2: Judgment Day                        | 31   |

|  |    |
|--|----|
| Django Unchained                           | 32 |
| The Departed                               | 33 |
| The Lion King                              | 34 |
| The Green Mile                             | 35 |
| The Prestige                               | 36 |
| The Pianist                                | 37 |
| Apocalypse Now                             | 38 |
| Raiders of the Lost Ark                    | 39 |
| Psycho                                     | 40 |
| Back to the Future                         | 41 |
| Alien                                      | 42 |
| Memento                                    | 43 |
| Whiplash                                   | 44 |
| The Lives of Others                        | 45 |
| Children of Heaven                         | 46 |
| WALL·E                                     | 47 |
| Braveheart                                 | 48 |
| Amélie                                     | 49 |
| Baahubali: The Beginning                   | 50 |
| Star Wars: Episode VI - Return of the Jedi | 51 |
| Once Upon a Time in America                | 52 |
| Princess Mononoke                          | 53 |
| Aliens                                     | 54 |
| American Beauty                            | 55 |
| Lawrence of Arabia                         | 56 |
| Das Boot                                   | 57 |
| Requiem for a Dream                        | 58 |
| Oldboy                                     | 59 |
| Reservoir Dogs                             | 60 |
| A Separation                               | 61 |
| Toy Story 3                                | 62 |
| Up   | 63 |
| Inside Out                                 | 64 |
| Batman Begins                              | 65 |
| Inglourious Basterds                       | 66 |

|                                       |     |
|---------------------------------------|-----|
| Indiana Jones and the Last Crusade    | 67  |
| L.A. Confidential                     | 68  |
| Toy Story                             | 69  |
| Scarface                              | 70  |
| Eternal Sunshine of the Spotless Mind | 71  |
| Amadeus                               | 72  |
| Raging Bull                           | 73  |
| Metropolis                            | 74  |
| Unforgiven                            | 75  |
| Downfall                              | 76  |
| Room                                  | 77  |
| Snatch                                | 78  |
| 2001: A Space Odyssey                 | 79  |
| Good Will Hunting                     | 80  |
| Snatch                                | 81  |
| The Sting                             | 82  |
| The Hunt                              | 83  |
| Some Like It Hot                      | 84  |
| Monty Python and the Holy Grail       | 85  |
| Captain America: Civil War            | 86  |
| How to Train Your Dragon              | 87  |
| The Wolf of Wall Street               | 88  |
| Finding Nemo                          | 89  |
| A Beautiful Mind                      | 90  |
| Casino                                | 91  |
| V for Vendetta                        | 92  |
| The Thing                             | 93  |
| Die Hard                              | 94  |
| Blade Runner                          | 95  |
| Gran Torino                           | 96  |
| Warrior                               | 97  |
| Howl's Moving Castle                  | 98  |
| Into the Wild                         | 99  |
| Pan's Labyrinth                       | 100 |
| The Big Lebowski                      | 101 |

|  |     |
|--|-----|
| Incendies  | 102 |
| Gone with the Wind                                     | 103 |
| The Secret in Their Eyes                               | 104 |
| Trainspotting  | 105 |
| The Bridge on the River Kwai                           | 106 |
| Lock, Stock and Two Smoking Barrels                    | 107 |
| On the Waterfront                                      | 108 |
| The Avengers   | 109 |
| Guardians of the Galaxy                                | 110 |
| Mad Max: Fury Road                                     | 111 |
| The Revenant   | 112 |
| The Bourne Ultimatum                                   | 113 |
| Pirates of the Caribbean: The Curse of the Black Pearl | 114 |
| Monsters, Inc.   | 115 |
| The Martian  | 116 |
| Shutter Island   | 117 |
| Jurassic Park  | 118 |
| Gone Girl  | 119 |
| The Truman Show  | 120 |
| The Avengers   | 121 |
| Deadpool   | 122 |
| Kill Bill: Vol. 1                                      | 123 |
| Prisoners  | 124 |
| The Sixth Sense  | 125 |
| Sin City   | 126 |
| Rush   | 127 |
| The Grand Budapest Hotel                               | 128 |
| Million Dollar Baby                                    | 129 |
| The Help   | 130 |
| No Country for Old Men                                 | 131 |
| There Will Be Blood                                    | 132 |
| 12 Years a Slave                                       | 133 |
| Spotlight  | 134 |
| Hotel Rwanda   | 135 |
| The Imitation Game                                     | 136 |

|                                    |     |
|------------------------------------|-----|
| The Princess Bride                 | 137 |
| Groundhog Day                      | 138 |
| The Sea Inside                     | 139 |
| Tae Guk Gi: The Brotherhood of War | 140 |
| Stand by Me                        | 141 |
| Akira                              | 142 |
| Elite Squad                        | 143 |
| The Terminator                     | 144 |
| Platoon                            | 145 |
| Butch Cassidy and the Sundance Kid | 146 |
| Donnie Darko                       | 147 |
| Annie Hall                         | 148 |
| The Wizard of Oz                   | 149 |
| Before Sunrise                     | 150 |
| The Best Years of Our Lives        | 151 |
| Amores Perros                      | 152 |
| The Celebration                    | 153 |
| Rocky                              | 154 |
| X-Men: Days of Future Past         | 155 |
| Ratatouille                        | 156 |
| Star Trek                          | 157 |
| Life of Pi                         | 158 |
| Casino Royale                      | 159 |
| Blood Diamond                      | 160 |
| The Incredibles                    | 161 |
| Cinderella Man                     | 162 |
| Big Fish                           | 163 |
| The Pursuit of Happyness           | 164 |
| Kill Bill: Vol. 2                  | 165 |
| Catch Me If You Can                | 166 |
| The Iron Giant                     | 167 |
| JFK                                | 168 |
| Serenity                           | 169 |
| Magnolia                           | 170 |
| District 9                         | 171 |

|                                 |     |
|---------------------------------|-----|
| Mystic River                    | 172 |
| Aladdin                         | 173 |
| Rain Man                        | 174 |
| Her                             | 175 |
| Dances with Wolves              | 176 |
| Dead Poets Society              | 177 |
| The Artist                      | 178 |
| The King's Speech               | 179 |
| Brazil                          | 180 |
| In Bruges                       | 181 |
| Mulholland Drive                | 182 |
| Slumdog Millionaire             | 183 |
| Black Swan                      | 184 |
| The Perks of Being a Wallflower | 185 |
| True Romance                    | 186 |
| Dancer in the Dark              | 187 |
| The Exorcist                    | 188 |
| Jaws                            | 189 |
| Casino Royale                   | 190 |
| Doctor Zhivago                  | 191 |
| The Straight Story              | 192 |
| Fiddler on the Roof             | 193 |
| Sicko                           | 194 |
| My Name Is Khan                 | 195 |
| The Sound of Music              | 196 |
| Persepolis                      | 197 |
| Dallas Buyers Club              | 198 |
| Shaun of the Dead               | 199 |
| Sling Blade                     | 200 |
| Boyhood                         | 201 |
| Bowling for Columbine           | 202 |
| Central Station                 | 203 |
| Young Frankenstein              | 204 |
| Before Sunset                   | 205 |
| Waltz with Bashir               | 206 |



|                                     |     |
|-------------------------------------|-----|
| A Fistful of Dollars                | 207 |
| Avatar                              | 208 |
| The Hobbit: The Desolation of Smaug | 209 |
| Iron Man                            | 210 |
| Edge of Tomorrow                    | 211 |
| Big Hero 6                          | 212 |
| The Hobbit: An Unexpected Journey   | 213 |
| How to Train Your Dragon 2          | 214 |
| Toy Story 2                         | 215 |
| Children of Men                     | 216 |
| The Insider                         | 217 |
| The Hateful Eight                   | 218 |
| The Bourne Identity                 | 219 |
| Almost Famous                       | 220 |
| Captain Phillips                    | 221 |
| Shrek                               | 222 |
| Hero                                | 223 |
| The Notebook                        | 224 |
| Glory                               | 225 |
| Walk the Line                       | 226 |
| Straight Outta Compton              | 227 |
| The Blues Brothers                  | 228 |
| The Right Stuff                     | 229 |
| Taken                               | 230 |
| The Untouchables                    | 231 |
| The World's Fastest Indian          | 232 |
| Edward Scissorhands                 | 233 |
| Glory                               | 234 |
| Ed Wood                             | 235 |
| My Fair Lady                        | 236 |
| Halloween                           | 237 |
| Hero                                | 238 |
| Hot Fuzz                            | 239 |
| Crouching Tiger, Hidden Dragon      | 240 |
| Halloween                           | 241 |

|                            |     |
|----------------------------|-----|
| The Remains of the Day     | 242 |
| Boogie Nights              | 243 |
| Letters from Iwo Jima      | 244 |
| The Fighter                | 245 |
| E.T. the Extra-Terrestrial | 246 |
| Crash                      | 247 |
| Amour                      | 248 |
| Nightcrawler               | 249 |
| Little Miss Sunshine       | 250 |

The movies under top 250 that are not in English are:

| Top_Foreign_Lang_Film              | language   |
|------------------------------------|------------|
| The Good, the Bad and the Ugly     | Italian    |
| City of God                        | Portuguese |
| Seven Samurai                      | Japanese   |
| Spirited Away                      | Japanese   |
| The Lives of Others                | German     |
| Children of Heaven                 | Persian    |
| Amélie                             | French     |
| Baahubali: The Beginning           | Telugu     |
| Princess Mononoke                  | Japanese   |
| Das Boot                           | German     |
| Oldboy                             | Korean     |
| A Separation                       | Persian    |
| Metropolis                         | German     |
| Downfall                           | German     |
| The Hunt                           | Danish     |
| Howl's Moving Castle               | Japanese   |
| Pan's Labyrinth                    | Spanish    |
| Incendies                          | French     |
| The Secret in Their Eyes           | Spanish    |
| The Sea Inside                     | Spanish    |
| Tae Guk Gi: The Brotherhood of War | Korean     |
| Akira                              | Japanese   |

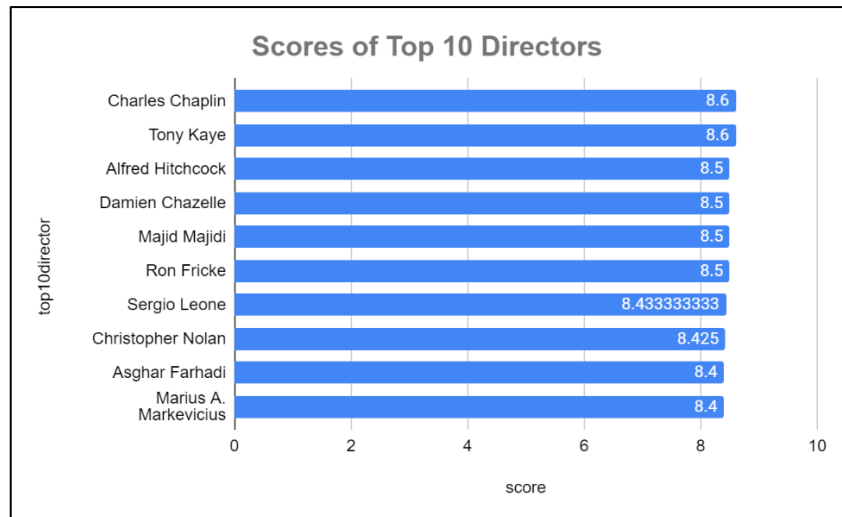
|                                |            |
|--------------------------------|------------|
| Elite Squad                    | Portuguese |
| Amores Perros                  | Spanish    |
| The Celebration                | Danish     |
| My Name Is Khan                | Hindi      |
| Persepolis                     | French     |
| Central Station                | Portuguese |
| Waltz with Bashir              | Hebrew     |
| A Fistful of Dollars           | Italian    |
| Hero                           | Mandarin   |
| Hero                           | Mandarin   |
| Crouching Tiger, Hidden Dragon | Mandarin   |
| Letters from Iwo Jima          | Japanese   |
| Amour                          | French     |
| Veer-Zaara                     | Hindi      |
| The Chorus                     | French     |
| Nine Queens                    | Spanish    |
| 4 Months, 3 Weeks and 2 Days   | Romanian   |
| Apocalypto                     | Maya       |
| Black Book                     | Dutch      |
| Kung Fu Hustle                 | Cantonese  |
| The White Ribbon               | German     |
| The Triplets of Belleville     | French     |
| Run Lola Run                   | German     |
| The Lunchbox                   | Hindi      |
| A Very Long Engagement         | French     |
| Ponyo                          | Japanese   |
| The Great Beauty               | Italian    |
| The Red Violin                 | French     |
| Good Bye Lenin!                | German     |
| Y Tu Mamá También              | Spanish    |
| Lady Vengeance                 | Korean     |
| The Flowers of War             | Mandarin   |
| The Kite Runner                | Dari       |
| House of Flying Daggers        | Mandarin   |
| Headhunters                    | Norwegian  |

|                                  |            |
|----------------------------------|------------|
| March of the Penguins            | French     |
| The Legend of Drunken Master     | Cantonese  |
| The Raid: Redemption             | Indonesian |
| The Painted Veil                 | Mandarin   |
| 2046                             | Cantonese  |
| The Orphanage                    | Spanish    |
| Maria Full of Grace              | Spanish    |
| Red Cliff                        | Mandarin   |
| The Baader Meinhof Complex       | German     |
| Mongol: The Rise of Genghis Khan | Mongolian  |
| Paris, je t'aime                 | French     |
| Shaolin Soccer                   | Cantonese  |
| The Good, the Bad, the Weird     | Korean     |
| L'auberge espagnole              | French     |
| Soul Kitchen                     | German     |
| Tsotsi                           | Zulu       |
| District B13                     | French     |
| Timecrimes                       | Spanish    |
| The Passion of the Christ        | Aramaic    |
| The Protector                    | Thai       |
| Curse of the Golden Flower       | Mandarin   |
| The Host                         | Korean     |
| Trollhunter                      | Norwegian  |
| Jab Tak Hai Jaan                 | Hindi      |
| El Mariachi                      | Spanish    |
| High Tension                     | French     |
| Coco Before Chanel               | French     |
| Rumble in the Bronx              | Cantonese  |
| Wasabi                           | French     |
| [Rec] 2                          | Spanish    |
| Night Watch                      | Russian    |
| The Interpreter                  | Aboriginal |
| Dead Snow                        | Norwegian  |
| The Legend of Zorro              | Spanish    |
| In the Land of Blood and Honey   | Bosnian    |

### 3. Best Directors:

The top 10 directors along with their IMDB scores are:

| top10director         | 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         |



### 4. Popular genres:

The top 5 genres with their average gross values (in million \$) is given below:

| genre_1   | genre_2   | Average_gross |
|-----------|-----------|---------------|
| Family    | Sci-Fi    | 434.9         |
| Adventure | Sci-Fi    | 228.6375      |
| Adventure | Animation | 116.9974359   |
| Adventure | Family    | 116.5314286   |
| Action    | Adventure | 110.3844125   |

The most viewed movies belong to family and Sci-Fi genres.

### 5. Critic favourite and audience favourite actors:

| actor_1_name      | AVERAGE of num_critic_for_reviews | AVERAGE of num_user_for_reviews |
|-------------------|-----------------------------------|---------------------------------|
| Brad Pitt         | 245                               | 742.3529412                     |
| Leonardo DiCaprio | 330.1904762                       | 914.4761905                     |
| Meryl Streep      | 181.4545455                       | 297.1818182                     |

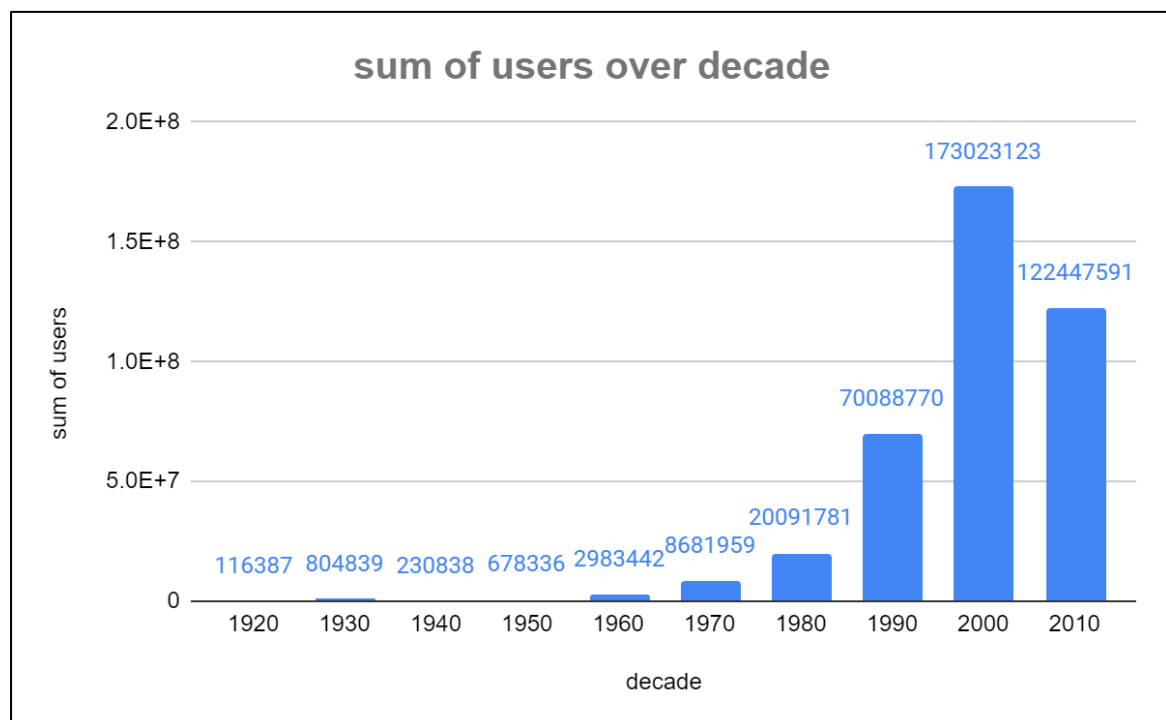
Leonardo DiCaprio is the most famous actor among the audience as well as the critics.

### 6. Change in voted users over decades:

| decade | df_by_decade |
|--------|--------------|
| 1920   | 116387       |
| 1930   | 804839       |
| 1940   | 230838       |
| 1950   | 678336       |
| 1960   | 2983442      |
| 1970   | 8681959      |
| 1980   | 20091781     |
| 1990   | 70088770     |
| 2000   | 173023123    |
| 2010   | 122447591    |

df\_by\_decade is the total number of voted users in that particular decade.

Maximum number of votes by users was obtained in the decade 2000.



**Result –**

IMDB Movie Analysis helps us to analyse various properties of movies which includes their ratings, number of likes, top movies, best actors and directors and so on.

This project has helped me learn how to analyse and make decisions during broadcasting of a movie which will lead to the growth of the company (VOD distributor).