

# IMDB Movie Analysis

By Sudhansu



# AGENDA

- ❑ Project Description
- ❑ Approach
- ❑ Tech-Stack Used
- ❑ Insights
- ❑ Result

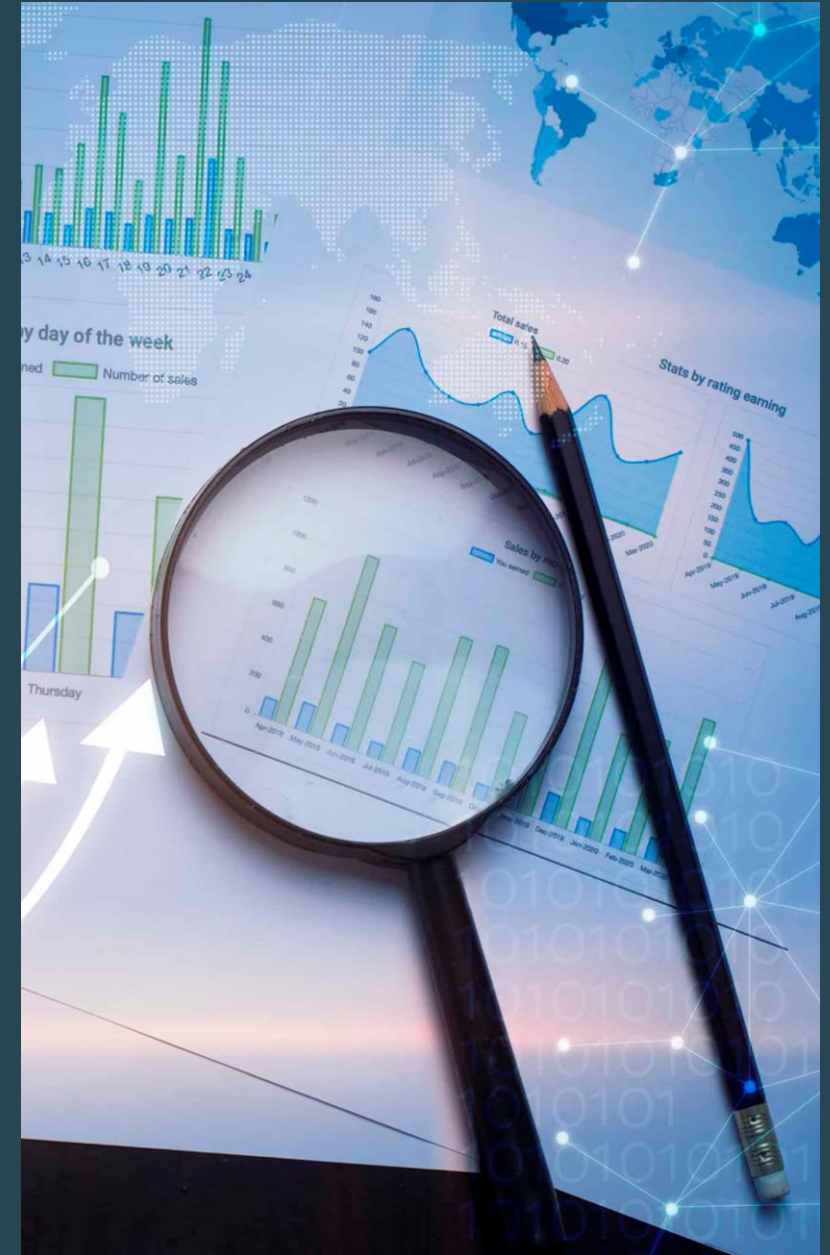
Excel workbook hyperlink : click on this

[Cleaned dataset](#)    [data cleaning steps & visualisations](#)

# Project Description

Dharma production is facing losses due to flop of last few movies. They have usually released movies for the Indian audience but for their next project they are planning to release a movie for the global audience.

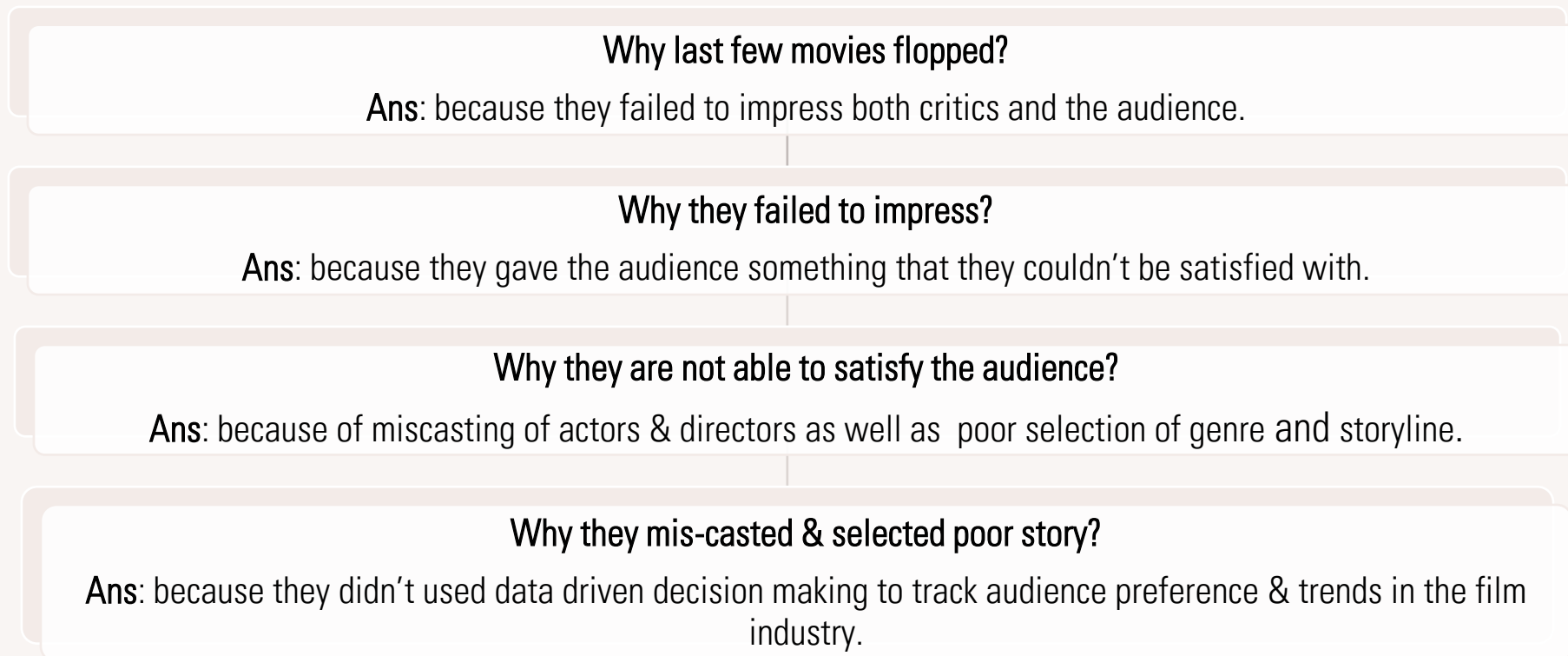
Thus helping the company to plan their every move analytically based on data. Generating meaningful insights from the IMDB dataset that can help them to make movie that appeals to both Indian and global audience while minimizing the risk of losses.



# Approach

**Downloaded** and **imported** the dataset in the excel, performed **data cleaning** i.e. removing missing & invalid rows and handling outliers. Used **pivot table** and **excel chart** to draw graphs and finding insights. And finally drawn conclusions from insights and made recommendations

## Root cause analysis : 5 Whys approach



# Impact

Meaningful insights from data analysis will help in :

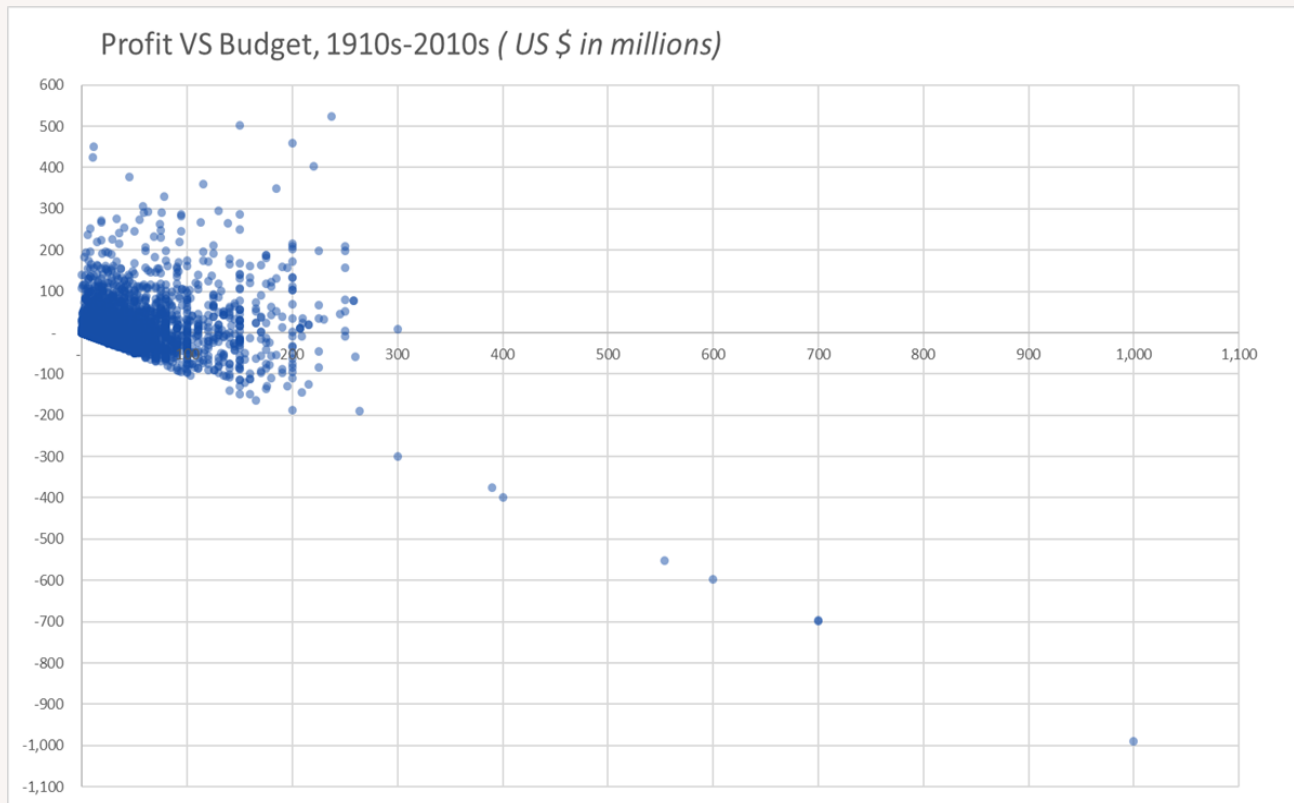
- Tracking audience preferences and identifying target audience
- Trends in film industry, trending genres and story
- Determining budget and minimizing the risk of losses

# Tech-Stack Used

Used Microsoft Excel version 2305, for data cleaning and visualizations. As Excel offers several powerful tools for data visualization that can help you analyze and understand your data better. With features such as charts, graphs, pivot tables, and other formatting options.

# Insights

There is no strong correlation between budget and profit except that the huge loss making movies are very high budget movies



- Top 10 movies by profit

movie_title	budget	profit
Avatar	\$ 23,70,00,000	\$ 52,35,05,847
Jurassic World	\$ 15,00,00,000	\$ 50,21,77,271
Titanic	\$ 20,00,00,000	\$ 45,86,72,302
Star Wars: Episode IV - A New Hope	\$ 1,10,00,000	\$ 44,99,35,665
E.T. the Extra-Terrestrial	\$ 1,05,00,000	\$ 42,44,49,459
The Avengers	\$ 22,00,00,000	\$ 40,32,79,547
The Lion King	\$ 4,50,00,000	\$ 37,77,83,777
Star Wars: Episode I - The Phantom Menace	\$ 11,50,00,000	\$ 35,95,44,677
The Dark Knight	\$ 18,50,00,000	\$ 34,83,16,061
The Hunger Games	\$ 7,80,00,000	\$ 32,99,99,255

- Avatar is the most profitable movie followed by Jurassic world which are high budget movies
- Star Wars and E. T. the Extra-Terrestrial are high profit making movies despite being quite low budget

# The Shawshank Redemption is the highest rated movie around the globe whereas Airlift is the highest rated Indian movie among the Top 250 IMDB rating movies

*IMDB top 250 rated movies, users vote greater than 25k*

rank	movie_title	imdb_score
1	The Shawshank Redemption	9.3
2	The Godfather	9.2
3	The Dark Knight	9
4	The Godfather: Part II	9
5	Fargo	9
6	Pulp Fiction	8.9
7	The Lord of the Rings: The Return of the King	8.9
8	Schindler's List	8.9
9	The Good, the Bad and the Ugly	8.9
10	12 Angry Men	8.9
11	Inception	8.8
12	Fight Club	8.8
13	Forrest Gump	8.8
14	The Lord of the Rings: The Fellowship of the Ring	8.8
15	Star Wars: Episode V - The Empire Strikes Back	8.8
16	Daredevil	8.8
17	It's Always Sunny in Philadelphia	8.8
18	The Matrix	8.7
19	The Lord of the Rings: The Two Towers	8.7
20	Star Wars: Episode IV - A New Hope	8.7

*IMDB top 10 rated Indian movies, users vote greater than 25k*

Rank	movie_title	imdb_score
1	Airlift	8.5
2	Rang De Basanti	8.4
3	Baahubali: The Beginning	8.4
4	Lage Raho Munna Bhai	8.2
5	My Name Is Khan	8
6	Veer-Zaara	7.9
7	The Lunchbox	7.8
8	Jab Tak Hai Jaan	6.9
9	Yeh Jawaani Hai Deewani	6.9
10	Krrish	6.3



# Sergio Leone and Christopher Nolan are the top 2 directors in the world

---

*Top 10 directors by avg. IMDB score*

Rank	Director Name	Average imdb_score
1	Sergio Leone	8.48
2	Christopher Nolan	8.43
3	Hayao Miyazaki	8.23
4	Quentin Tarantino	8.20
5	Stanley Kubrick	8.00
6	Sidney Lumet	7.98
7	Frank Darabont	7.98
8	Alfred Hitchcock	7.95
9	James Cameron	7.91
10	Francis Ford Coppola	7.81

- These are the names of top 10 directors calculated by sorting their name by mean of imdb\_score of those directors with movies greater than 3 in top 250 imdb movies
- And number of voted users greater than 25000



# Biography is the genre with highest average IMDB ratings around the globe

---

*Top 10 Genres by avg. IMDB score*





Rank	Genres	Average imdb_score
1	Biography	7.38
2	History	7.36
3	War	7.29
4	Drama	7.07
5	Animation	7.02
6	Sport	6.85
7	Crime	6.81
8	Mystery	6.73
9	Romance	6.69
10	Adventure	6.67

- These are the names of top 10 genres calculated by sorting their name by mean of imdb\_score of those movies which has number of voted users greater than 25000





# Leonardo DiCaprio is the most favourite actor among the critics as well as the audience

---

*Top 3 critic – favourite actors*

actors name		Average num_critic_for_reviews
Leonardo DiCaprio		330
Brad Pitt		232
Meryl Streep		163

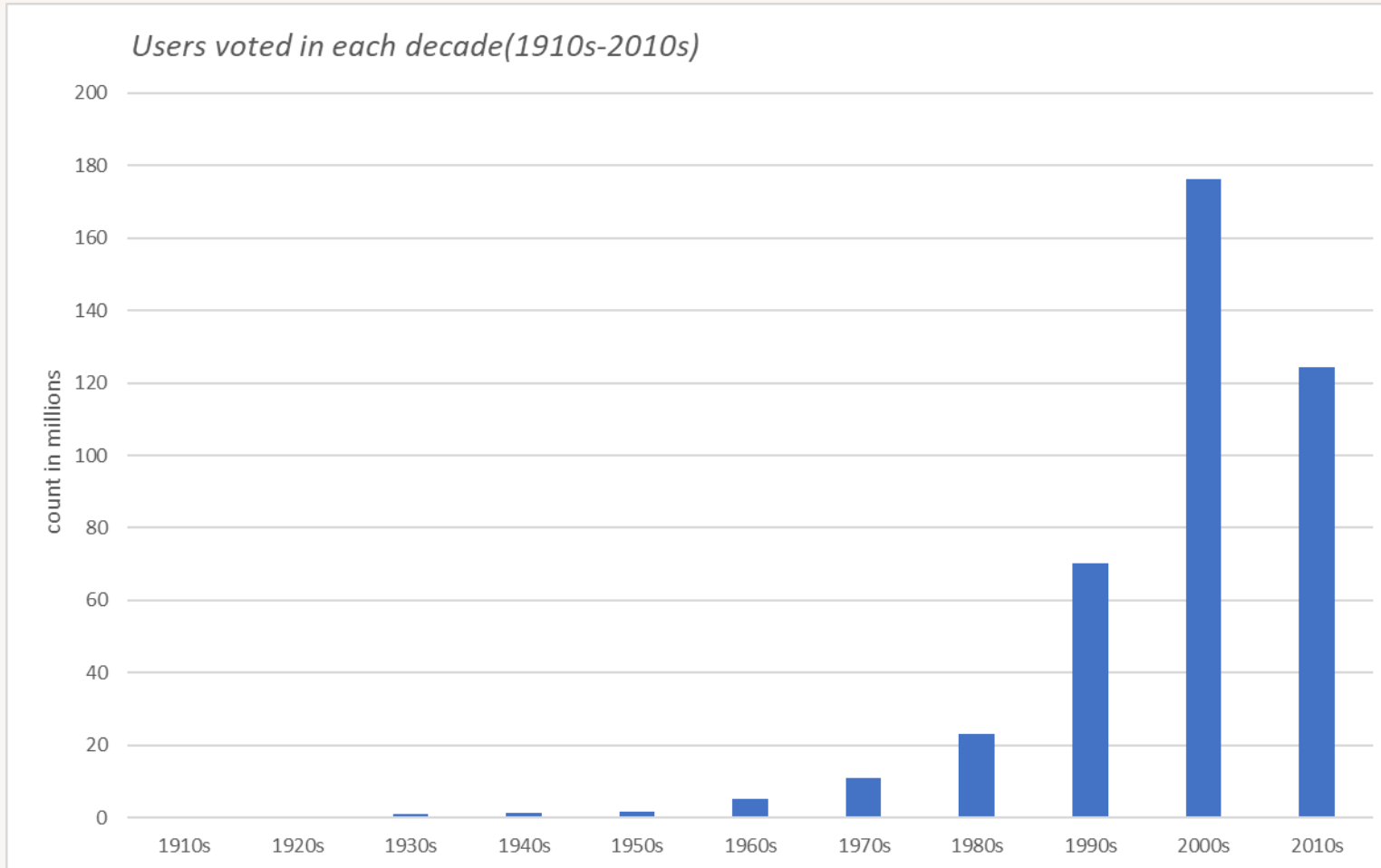
*Top 3 audience – favourite actors*

actors name		Average num_user_for_reviews
Leonardo DiCaprio		914
Brad Pitt		702
Meryl Streep		257

- These top actors are ranked by mean of number of critic reviews and user reviews
- Only those movies of actors are taken in consideration for mean review calculation in which they are in lead role

# Users vote count increased from ~10k in 1910s to ~180 million in 2000s

---



- Sum of users vote increased exponentially to all time high in 2000s
- Bars are calculated by taking sum total votes of users for all movies in each decade

# Results

## ❑ Conclusions and recommendations :

- Most **profitable** movies are in the **range** of \$100million to \$300million, the company can keep their **budget** between this range to minimize the risk of loss.
- **Shawshank Redemption** followed by **The Godfather** and **The Dark Knight** are the highest IMDB rated movies around the world, the company can take **inspiration** from **plot** of these movies.
- **Sergio Leone, Christopher Nolan** and **Hayao Miyazaki** are the **top 3** directors with highest mean IMDB rating movies in the world, the company can **work** in **collaboration** with these directors.
- **Biography, History** and **War** are the **top 3** genres with highest mean IMDB rating movies in the world, the company can choose genre for their next movies from these.
- **Leonardo DiCaprio** is the **most favourite** actor among the critics as well as the audience around the world, the company can **cast** him or can add a **cameo role** in the movie to **create hype** among the fans for the movie.

Thank you