IMDB Movie Analysis

Trainity Task-5

About The Project

- In this project, we will analyze the movie data from the excel data set in project description.

- We have the movies data from the past decade along with various pieces of information about the movie, its actors, and the voters who have rated these movies online. In this analysis, we will try to find some interesting insights into these movies and their voters
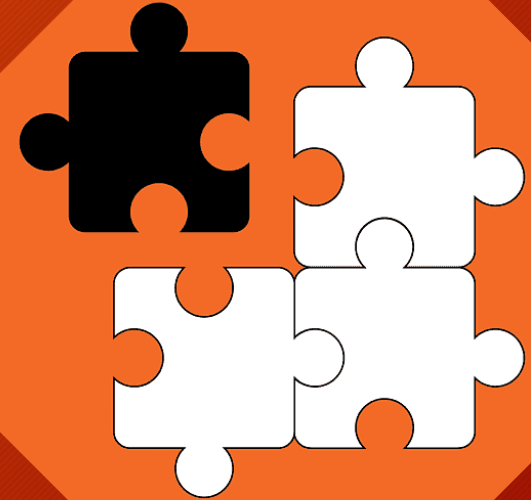
IMDb

Approach

IMDb

- Get the information from given description from data and understood problem.

- Use MS excel to analyze the data through filter ,pivot table, sum if, average if, count if and other functions to answer the question.

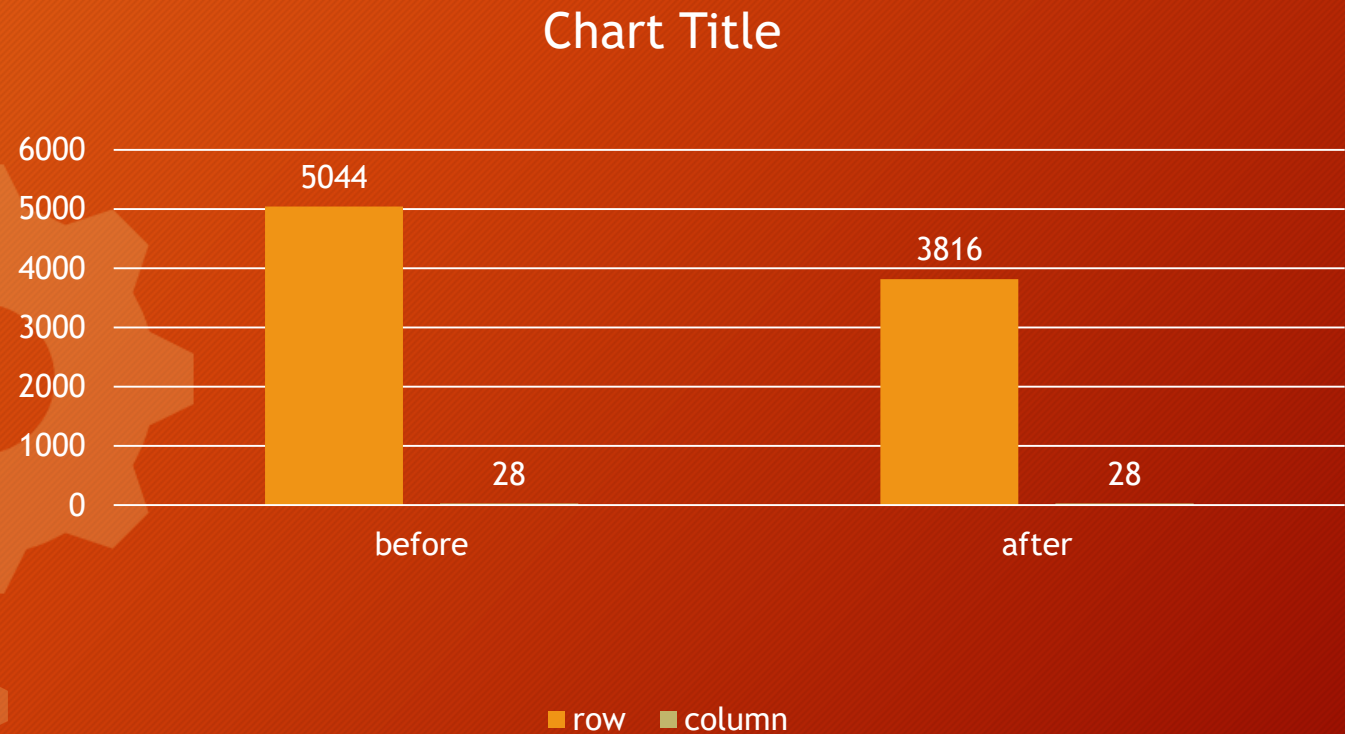- Create charts to for easy and meaningful data representation.

Tech used

IMDb

Insights

IMDb

- **Cleaning the data:**

Chart Title

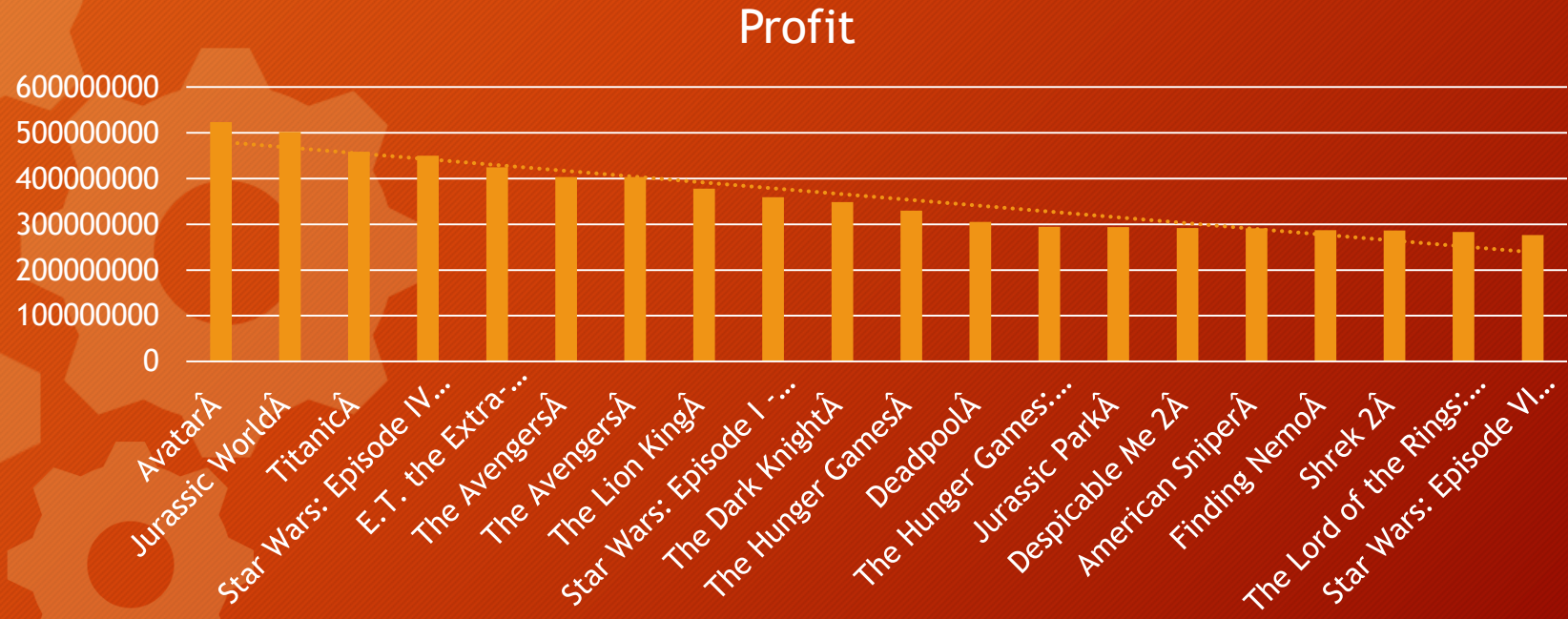| Data Clean		
	Before	After
Row	5044	3816
Column	28	28

IMDb

- **Movies with highest profit:**

  Find the movies with the highest profit?

Profit

Insights

IMDb

- **Top 250:**

  Find Top IMDB:

Here is snippet of output data.

Movie Title	IMDB	num_voted_users > 25,000.
The Shawshank Redemption	9.3	1689764
The Godfather	9.2	1155770
The Dark Knight	9	1676169
The Godfather: Part II	9	790926
The Lord of the Rings: The Return of the King	8.9	1215718
Pulp Fiction	8.9	1324680
Schindler's List	8.9	865020
The Good, the Bad and the Ugly	8.9	503509
Forrest Gump	8.8	1251222
Star Wars: Episode V - The Empire Strikes Back	8.8	837759
The Lord of the Rings: The Fellowship of the Ring	8.8	1238746
Inception	8.8	1468200
Fight Club	8.8	1347461
Star Wars: Episode IV - A New Hope	8.7	911097
The Lord of the Rings: The Two Towers	8.7	1100446
The Matrix	8.7	1217752
One Flew Over the Cuckoo's Nest	8.7	680041
Goodfellas	8.7	728685
City of God	8.7	533200
Seven Samurai	8.7	229012

Insights

IMDb

- **Best Directors:**

Find the best directors

From the data table, we have find Akira kurosawa as best director.

Director name	Average of IMDB SCORE
Akira Kurosawa	8.7
Asghar Farhadi	8.4
Charles Chaplin	8.6
Christopher Nolan	8.425
Damien Chazelle	8.5
Majid Majidi	8.5
Michael Curtiz	8.6
Richard Marquand	8.4
Ron Fricke	8.5
Sergio Leone	8.433333333
Tony Kaye	8.6

Insights

- **Popular Genres:**

  Find popular genres:

Genre	Average FB Likes
Action	12979.44223
Adventure	14816.11224
Animation	5617.194444
Biography	12588.55474
Comedy	4521.691867
Crime	9312.768473
Documentary	3796.285714
Drama	8919.568932
Fantasy	11682.77419
Horror	7538.913043

Genre Popularity on Facebook
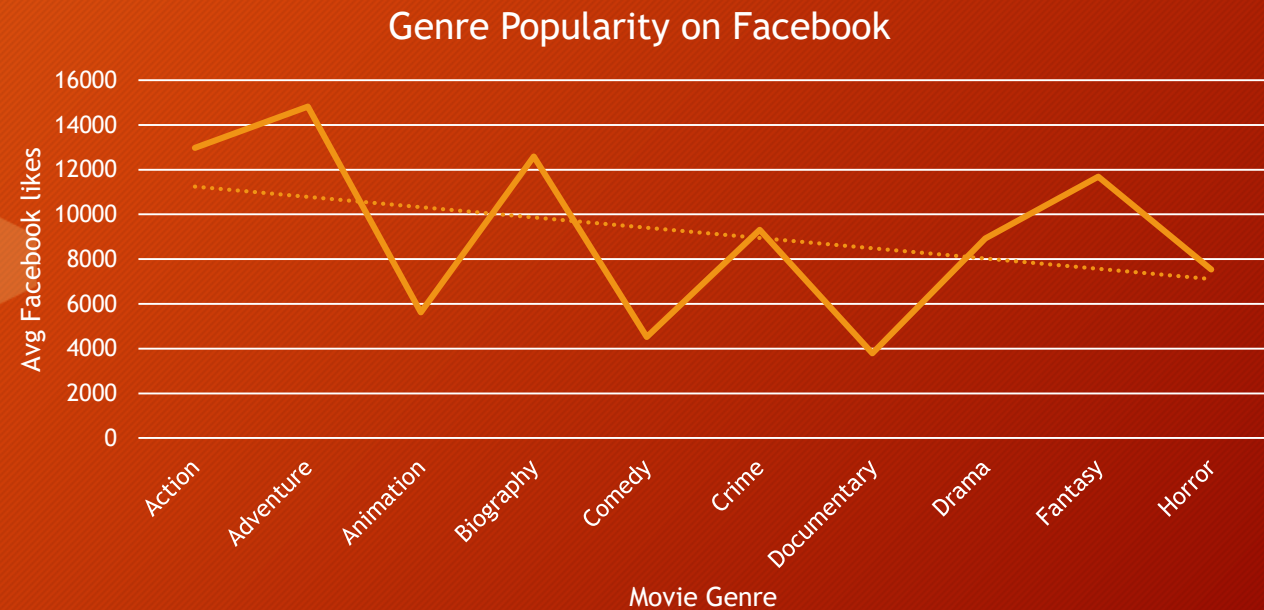
Insights

- **Charts**

Find the mean of the num_critic_for_reviews and num_users_for_review and identify the actors which have the highest mean.

User	Mean critic review
Meryl Streep	181.4545
Leonardo DiCaprio	330.19
Brad Pitt	245

Leonardo DiCaprio has highest num_critic_for_reviews means of 330.19

User	Mean user review
Meryl Streep	297.1818
Leonardo DiCaprio	914.4761
Brad Pitt	742.3529

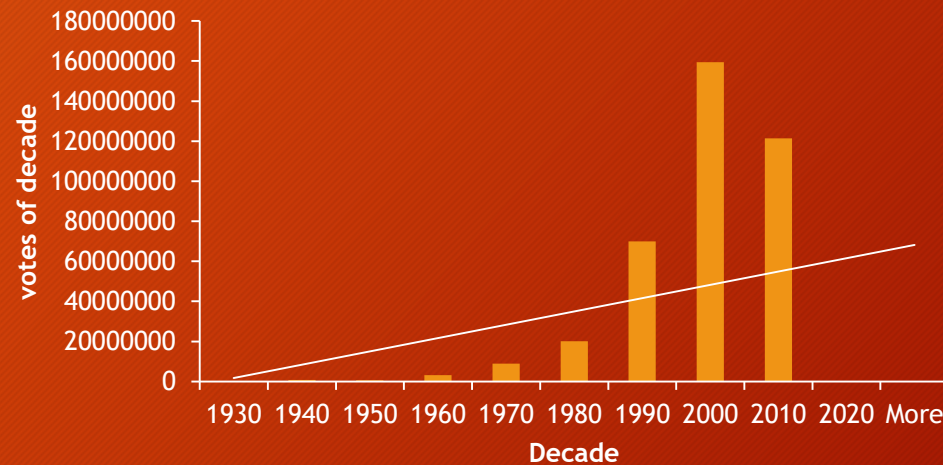Leonardo DiCcaprio has highest num_user_for_reviews mean of 914.47

Insights

Change in number of voted users over decades using a bar chart

Decades	Voted user of decade
1930	116387
1940	804839
1950	636002
1960	3120472
1970	8854391
1980	20092021
1990	69935395
2000	159445066
2010	121375798
2020	0

Vote Data Over Decades

Result

- Through this project, I came to know about , how analytics help in getting sights related from movie data.
- It helped me to brush up my concepts related chart, formula and pivot table which are helpful in data analysis.
- It helped me to understand dataset through charts and various other methods.

Thank You

By : Rohit Bahuguna
You can find me at rohitbahuguna14@gmail.com