



IMDB MOVIE ANALYSIS

Analysis of
movies and
directors with
repect to
budget and
IMDB rating



PROJECT AGENDA

1

PROJECT DESCRIPTION

2

APPROACH

3

TECH-STACK USED

4

INSIGHTS

5

RESULT

PROJECT DESCRIPTION

The project is about finding the **insights and graphical visualization** needed for the IMDB Movie Analysis.

The tools used to answer the question asked by the stakeholders in **Excel**.



APPROACH

First, the dataset has been loaded in Excel then the dataset has been thoroughly studied to understand the attributes then **used multiple functions, pivot tables, and then a few graphs** to find out the **answers** to the questions asked by the stakeholders.



TECH-STACK USED

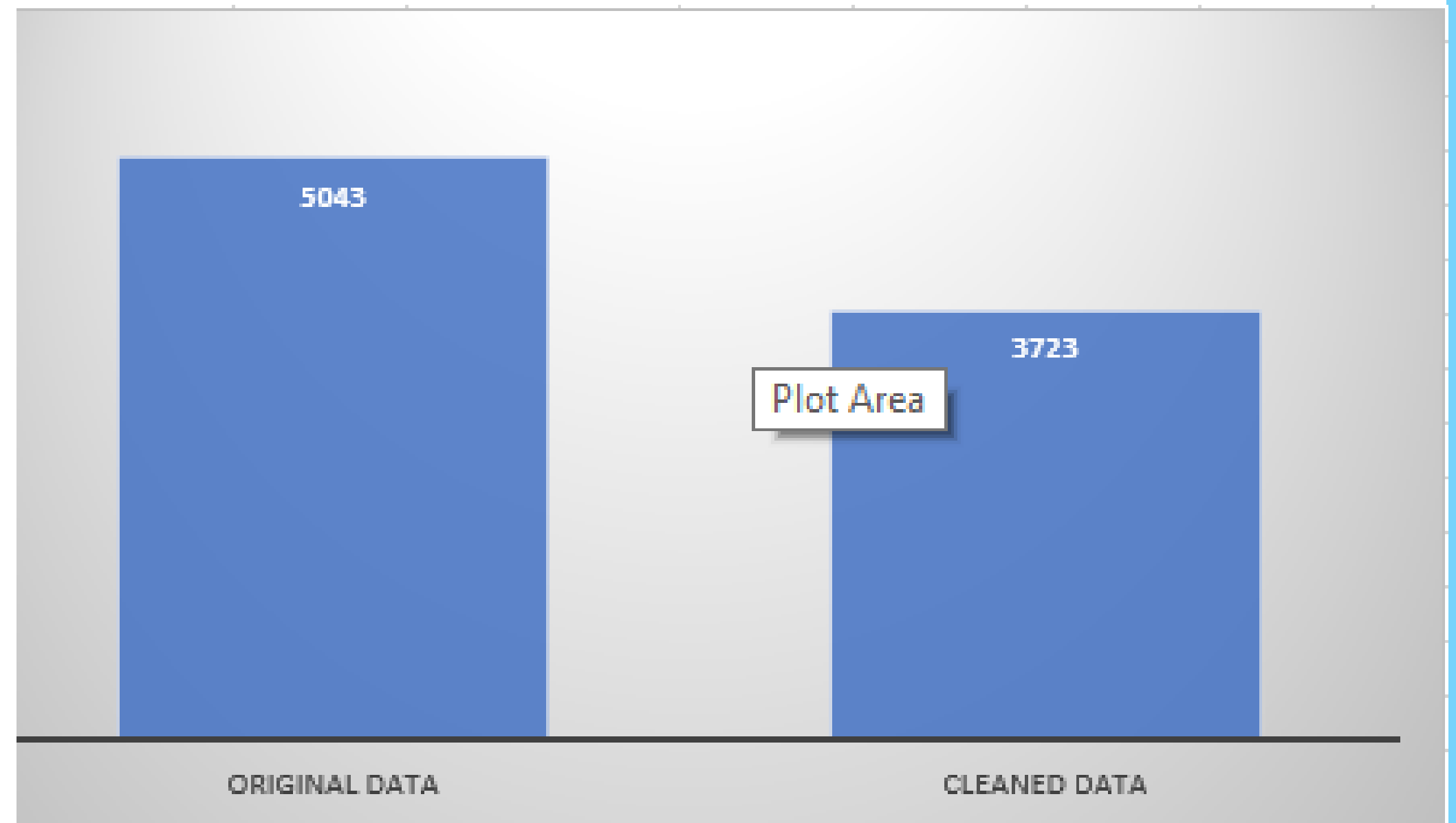
The software used to solve the problem the statement is in **Excel** which is a powerful tool to do **EDA** on this data and create **data visualization**



INSIGHTS

A. CLEANING THE DATA:

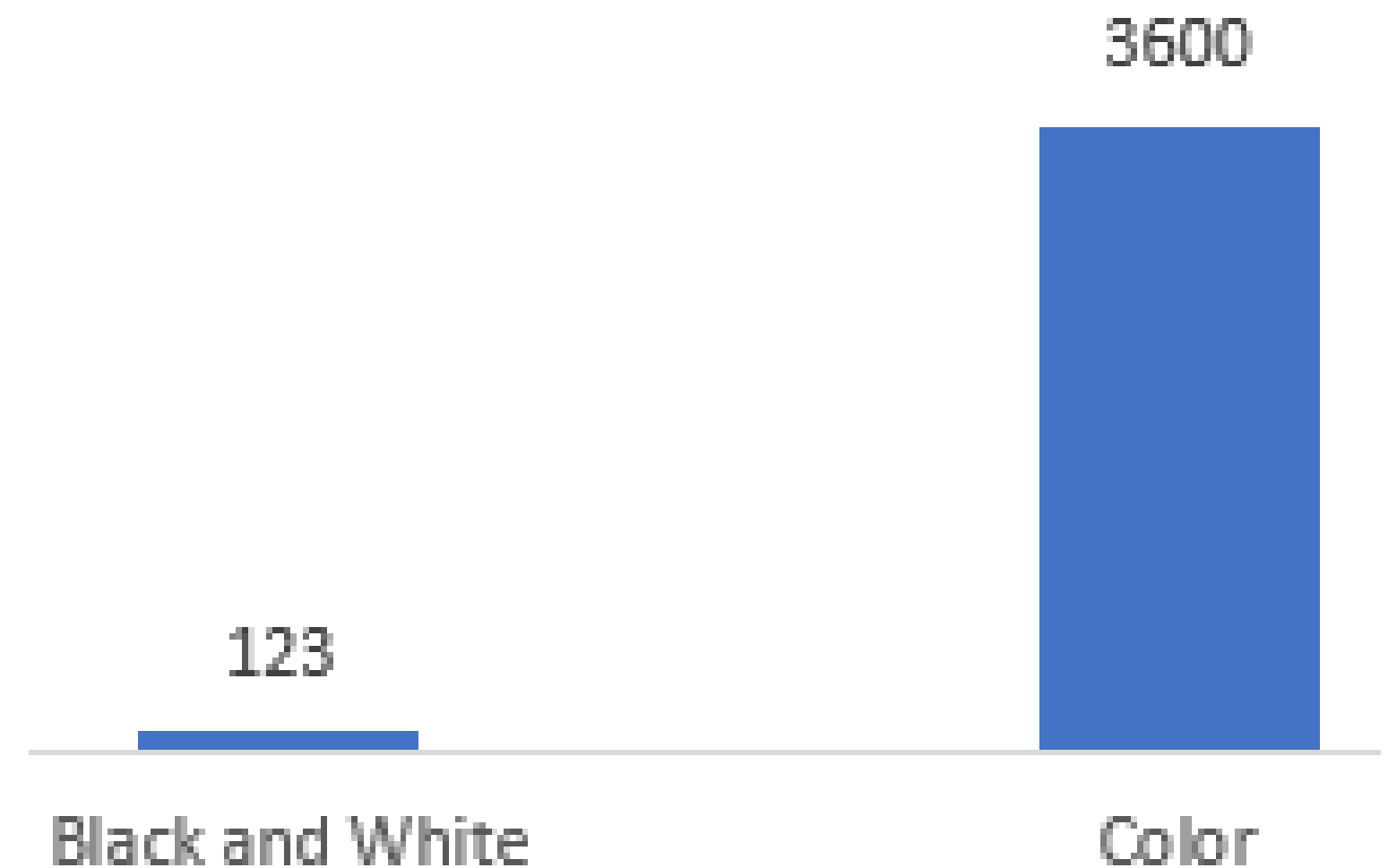
Since the dataset had a lot of **null values** I needed to **remove** the data with null values which is data cleaning to make use of the data for analysis



INSIGHTS

SOME INTERESTING INSIGHTS:

There are **123 black and white** movies released and **3600 color** movies released in the given dataset i.e movies in color are released **29 times** more than movies in black and white



INSIGHTS

SOME INTERESTING INSIGHTS:

Steven Spielberg has released a **total of 25 movies** which is the highest in the given dataset which is a period **between the 1920s and 2010s** and also has the most **sum of Facebook likes** which is **350000**

INSIGHTS

SOME INTERESTING INSIGHTS:

Interstellar is the movie which has the **highest facebook likes** which is **349000**

INSIGHTS

SOME INTERESTING INSIGHTS:

Sholem Aleichem: Laughing in the Darkness is the movie that was released with the **highest aspect ratio** which is **16**

INSIGHTS

SOME INTERESTING INSIGHTS:

Tarnation was the movie with the **lowest budget** which is **\$218**

The Host is the movie with the **highest Budget** which is **\$12255500000**

Insights

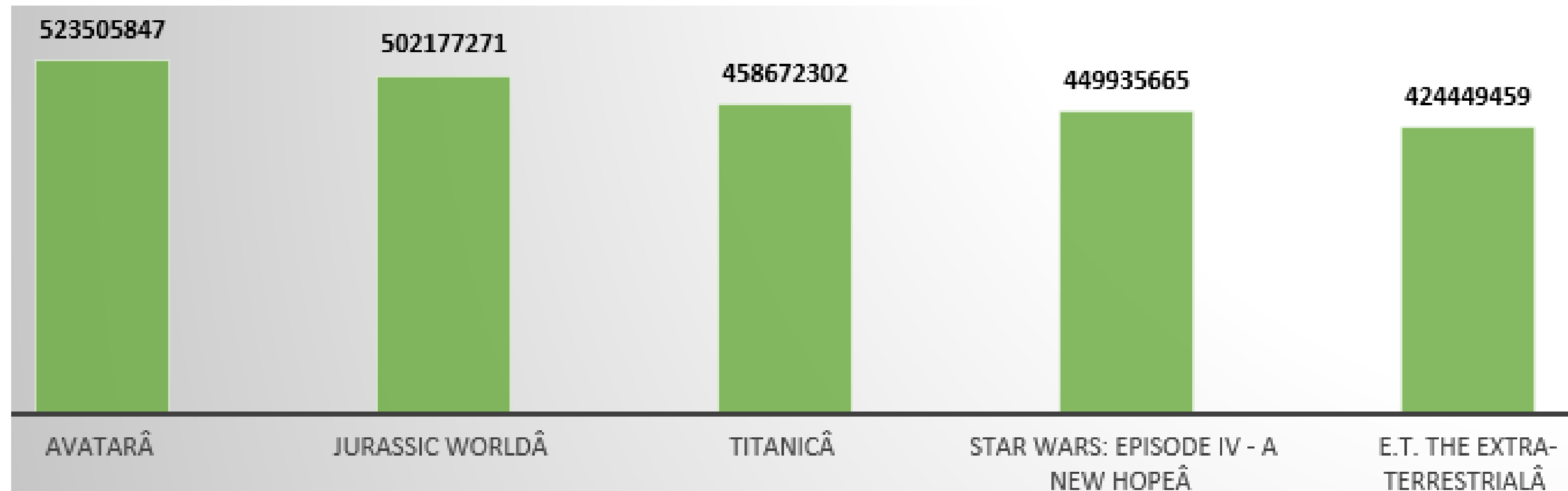
B. MOVIE WITH HIGHEST PROFIT:

The **movie** with the
highest profit is **Avatar**
with a profit of
\$523505847

Movie title	Sum of Profit
Avatar	523505847
Jurassic World	502177271
Titanic	458672302
Star Wars: Episode IV - A New Hope	449935665
E.T. the Extra-Terrestrial	424449459
The Avengers	403279547
The Lion King	377783777
The Jungle Book	375290282
Star Wars: Episode I - The Phantom Menace	359544677
The Dark Knight	348316061

INSIGHTS

B. TOP 5 MOVIE WITH THE HIGHEST PROFIT:



INSIGHTS

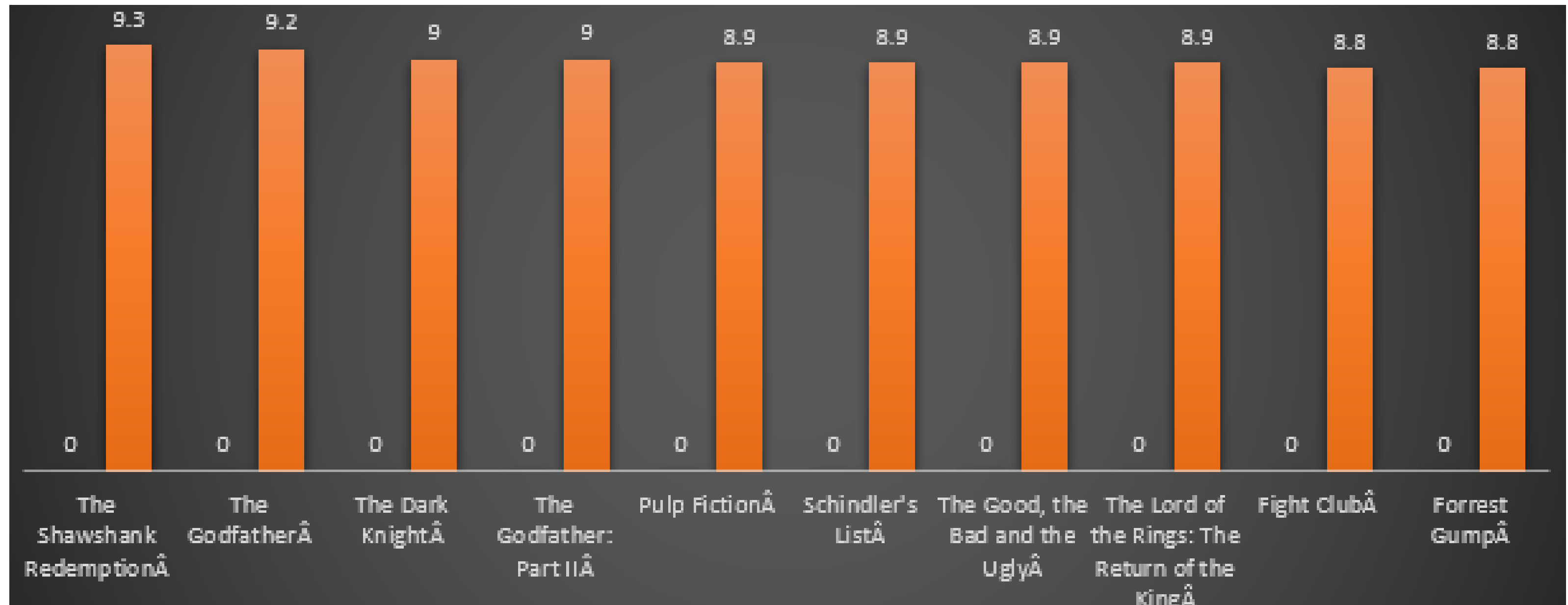
C. TOP 250 IMDB MOVIES :

The List of top 250 movies is filtered according to the IMDB scores and num of voted users is **greater than 25000**

Since the list is long, it is **available in Excel** and there will be the **top 10** in the **next slide**

INSIGHTS

C. TOP 10 IN TOP 250 IMDB MOVIES :



INSIGHTS

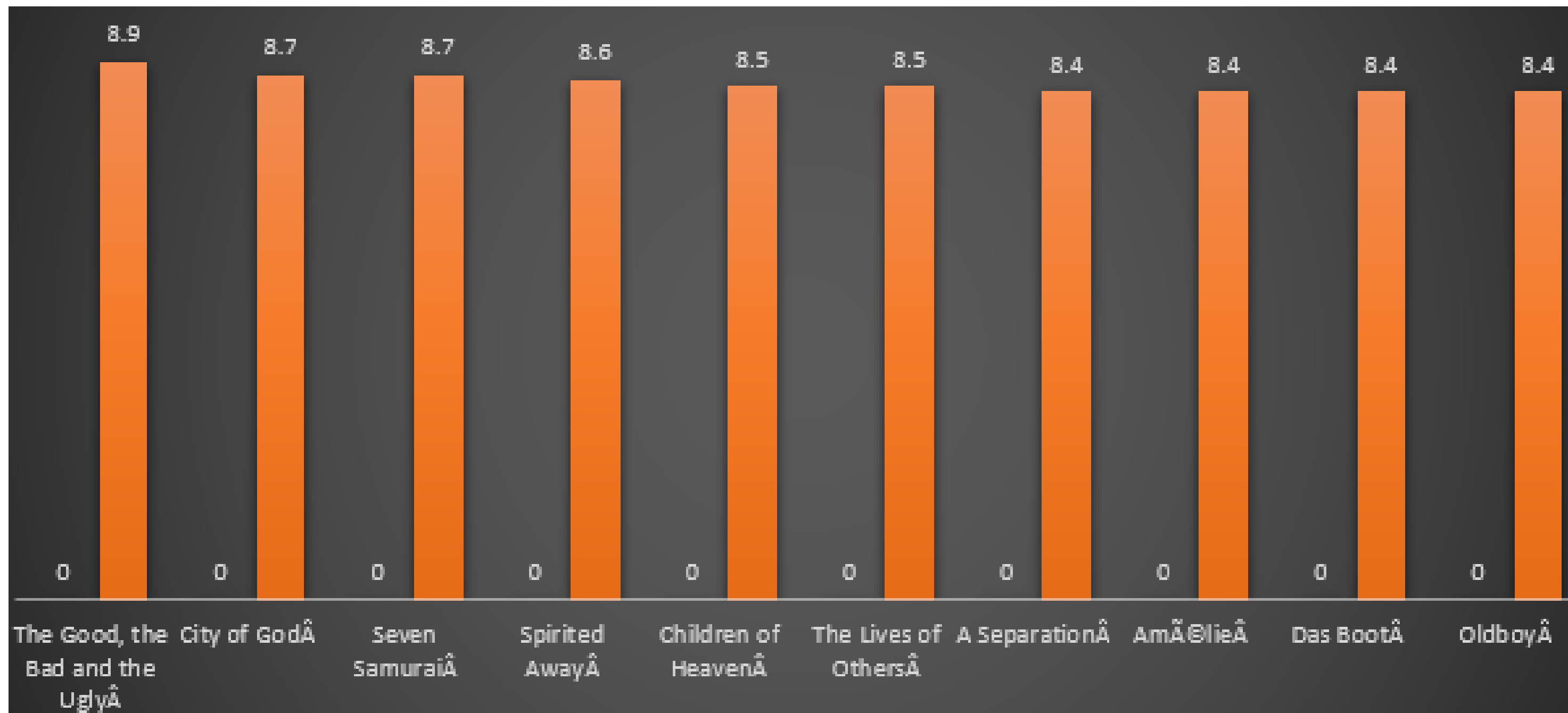
C. TOP FOREIGN MOVIES IN TOP 250 IMDB MOVIES :

The List of **top foreign movies** from the **list of 250 movies** is filtered where the condition is the movie language is not equal to english

Since the list is long, it is **available in Excel** and there will be the **top 10** in the **next slide**

INSIGHTS

C. TOP 10 FOREIGN IN TOP 250 IMDB MOVIES :



INSIGHTS

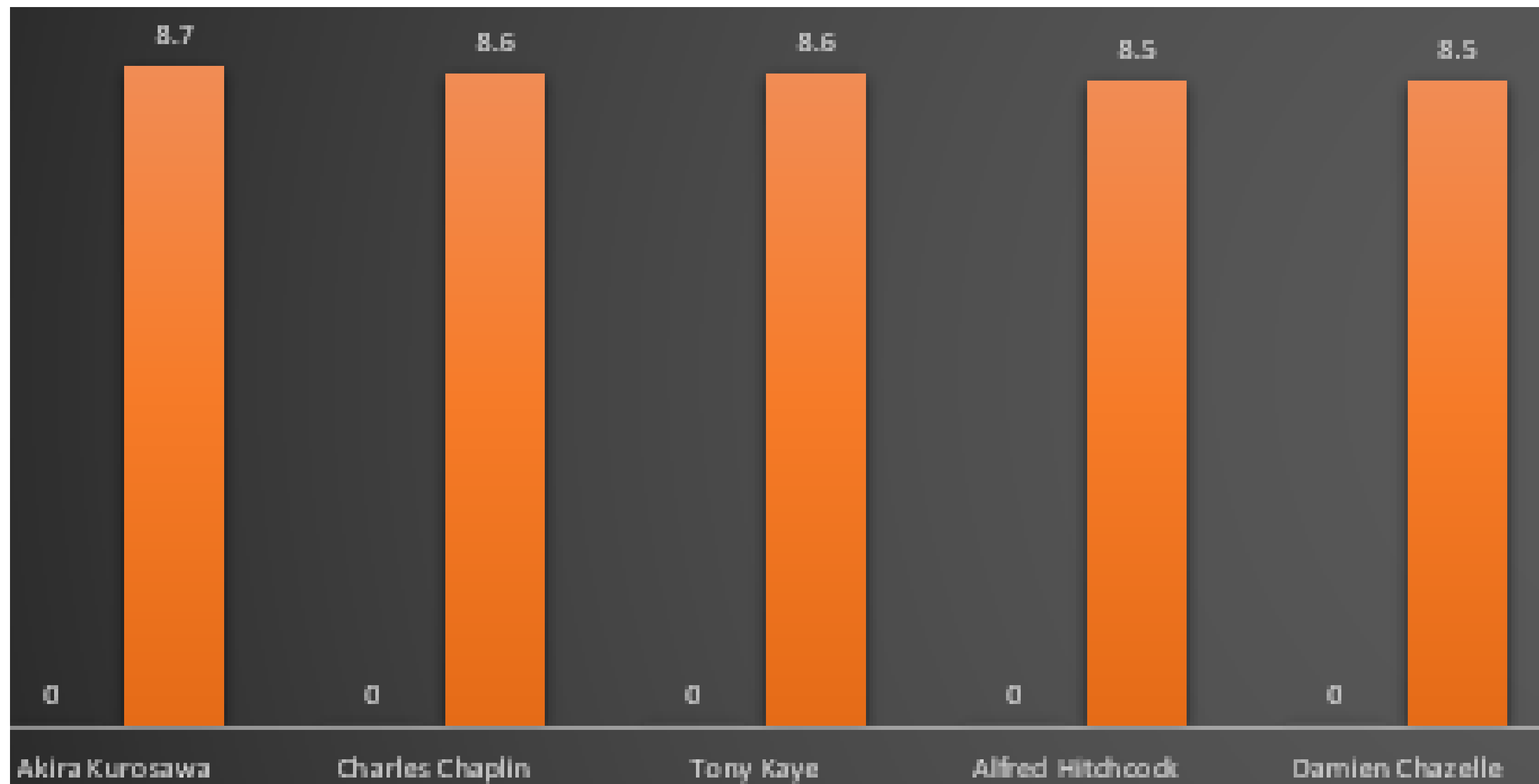
D. BEST DIRECTORS :

The List of Best Directors based on **IMDB score**

Director Names	Average of imdb_score
Akira Kurosawa	8.7
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	8.4333333333
Christopher Nolan	8.425
Asghar Farhadi	8.4
Richard Marquand	8.4

INSIGHTS

D. TOP 5 BEST DIRECTORS:



INSIGHTS

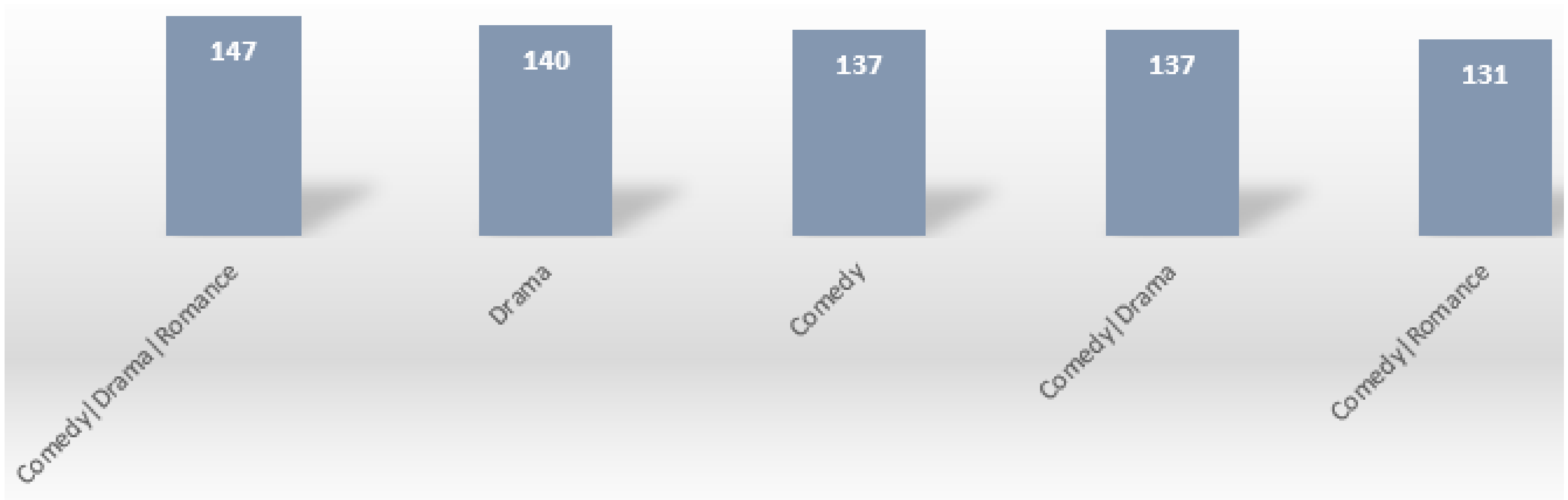
E. POPULAR GENRES :

The popular genres are based on the **frequency**

Genres	Frequency
Comedy Drama Romance	147
Drama	140
Comedy	137
Comedy Drama	137
Comedy Romance	131
Drama Romance	114
Crime Drama Thriller	81
Action Crime Thriller	55
Action Crime Drama Thriller	50
Action Adventure Sci-Fi	46

INSIGHTS

E. TOP 5 POPULAR GENRES :



INSIGHTS

F. CRITIC-FAVORITE AND AUDIENCE FAVORITE ACTORS:

Three columns have been made with the names **Meryl_Streep**, **Leo_Caprio**, and **brad_pitt** as required, and made another column with name **combined** where I combined all the rows in the above three columns which **in the Excel sheet**

INSIGHTS

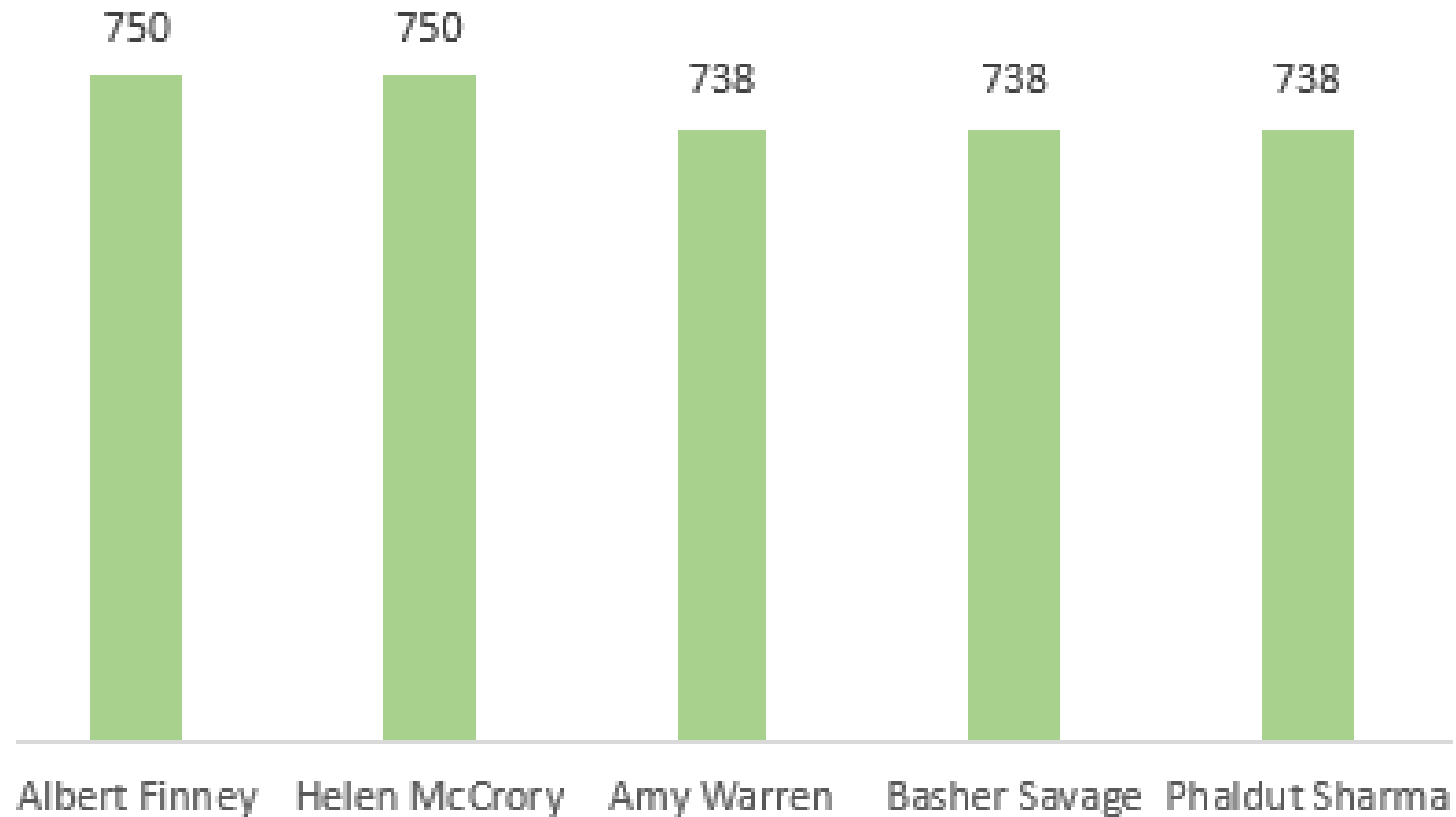
F. CTITIC-FAVORITE ACTORS:

Critic favorite actors are the actors who have the **most num critic reviews** from all the three columns such as from **actor_name_1 and 2 and 3**

actor_names	Average of num_critic_for_reviews
Albert Finney	750
Helen McCrory	750
Amy Warren	738
Basher Savage	738
Phaldut Sharma	738
Joseph Gordon-Levitt	727.5
Robert Downey Jr.	680.3333333
Russ Tamblyn	676
Alan D. Purwin	673
Dwight Henry	663
Gina Montana	663

INSIGHTS

F. TOP 5 CRITIC-FAVORITE ACTORS:



INSIGHTS

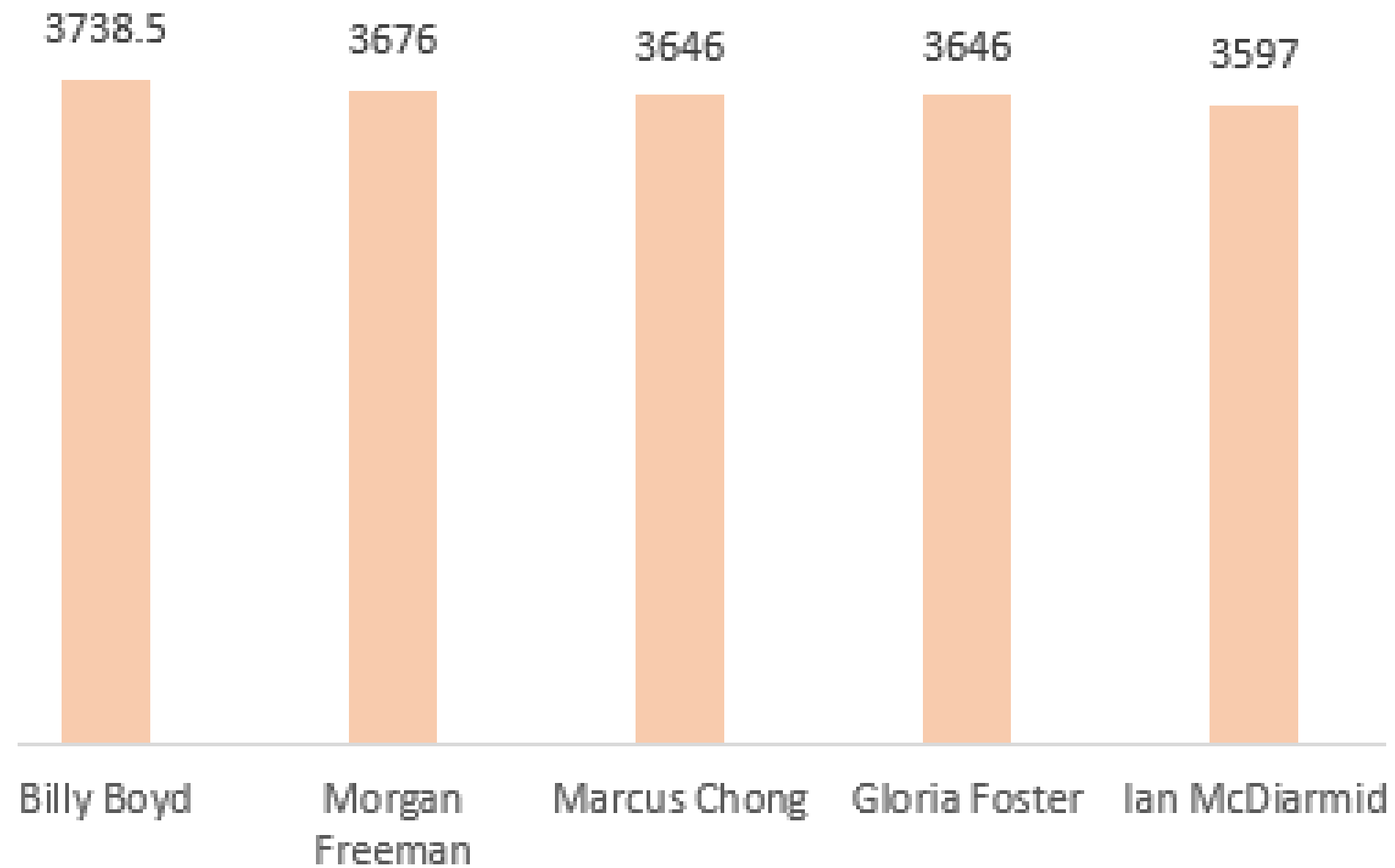
F. AUDIENCE-FAVORITE ACTORS:

Audience favorite actors are the actors who have the **most num of voted users** from all the three columns such as from **actor_name_1** and 2 and 3

actor names	Average of num_user_for_reviews
Billy Boyd	3738.5
Morgan Freeman	3676
Marcus Chong	3646
Gloria Foster	3646
Ian McDiarmid	3597
Hayden Christensen	3401
Heather Donahue	3400
Joshua Leonard	3400
Michael C. Williams	3400
Lauren Cohan	3018
Alan D. Purwin	3018

INSIGHTS

F. AUDIENCE-FAVORITE ACTORS:



INSIGHTS

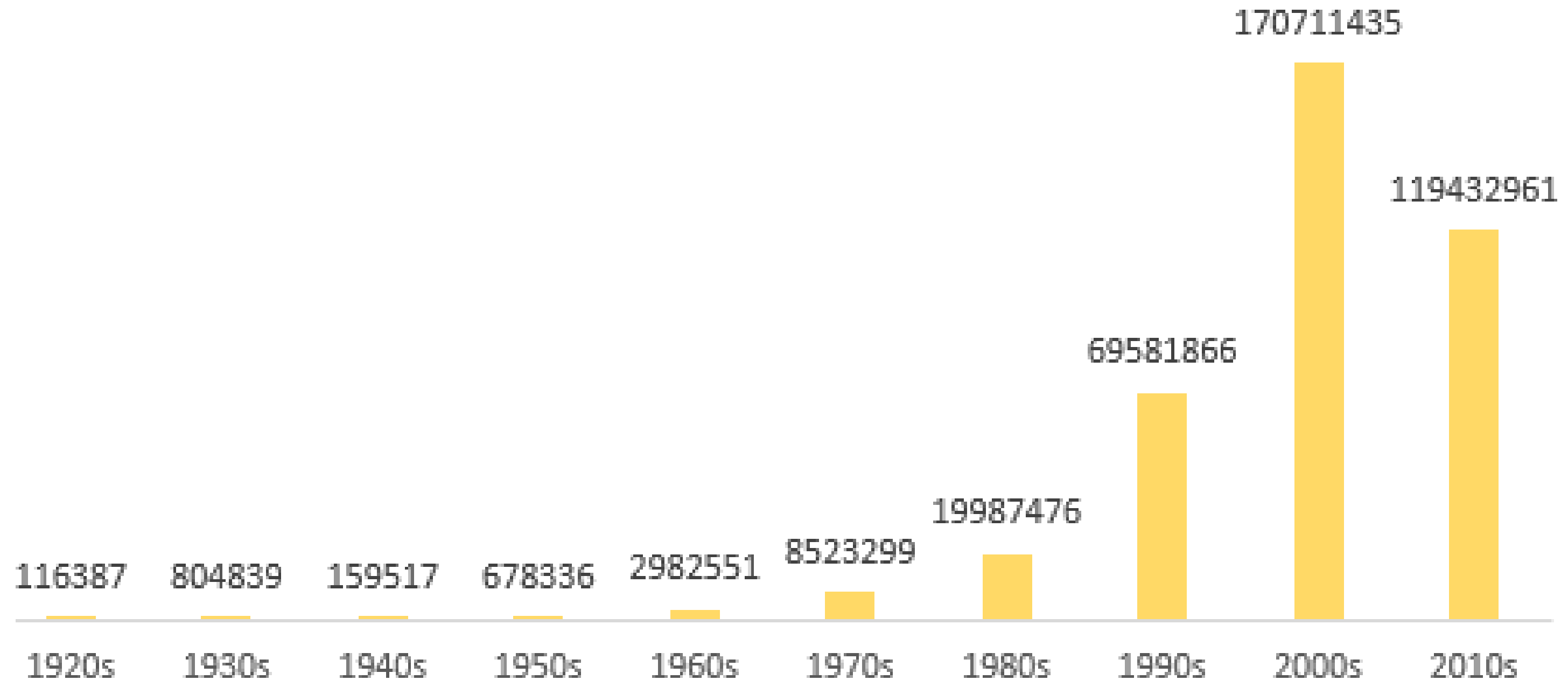
F. NUMBER OF USERS VOTED VS DECADES:

Every year is made to fit in bins of 10 years of gap starting from **1920** to **2010** and the **total of the users voted** is calculated

Decades	Sum of num_voted_users
1920s	116387
1930s	804839
1940s	159517
1950s	678336
1960s	2982551
1970s	8523299
1980s	19987476
1990s	69581866
2000s	170711435
2010s	119432961

INSIGHTS

F. NUMBER OF USERS VOTED VS DECADES:



RESULTS

In this project, I have learned to **combine unique data from multiple columns** and found out there are multiple ways but I used my own and the rest of the things which pivot table and the graphs which I learned in previous projects so that part was not that difficult and I learned how to easily **create decades column** which I have learned through this project which was challenging and also I tried to find **my own insights** in the project which was interesting and I learned how to **make the problem statement and then answered the question** created by me.



**THANK
YOU!**