

IMDB MOVIE ANALYSIS

The IMDb logo, featuring the text "IMDb" in a bold, black, sans-serif font, centered within a yellow rounded square.

IMDb

Project by-
Surendra upadhyay
DATA ANALYST

- Dataset having 28 columns, 5044 Rows of different IMDB Movies.
- Defined a problem, clean the data as necessary remove null and duplicate values.
- Explore the data set and derive insights
- Use root cause analysis Five 'Whys' approach.
- Create chart pivot table and graphs to answer given questions
- Give answer to asked questions and create reports for data-driven decision

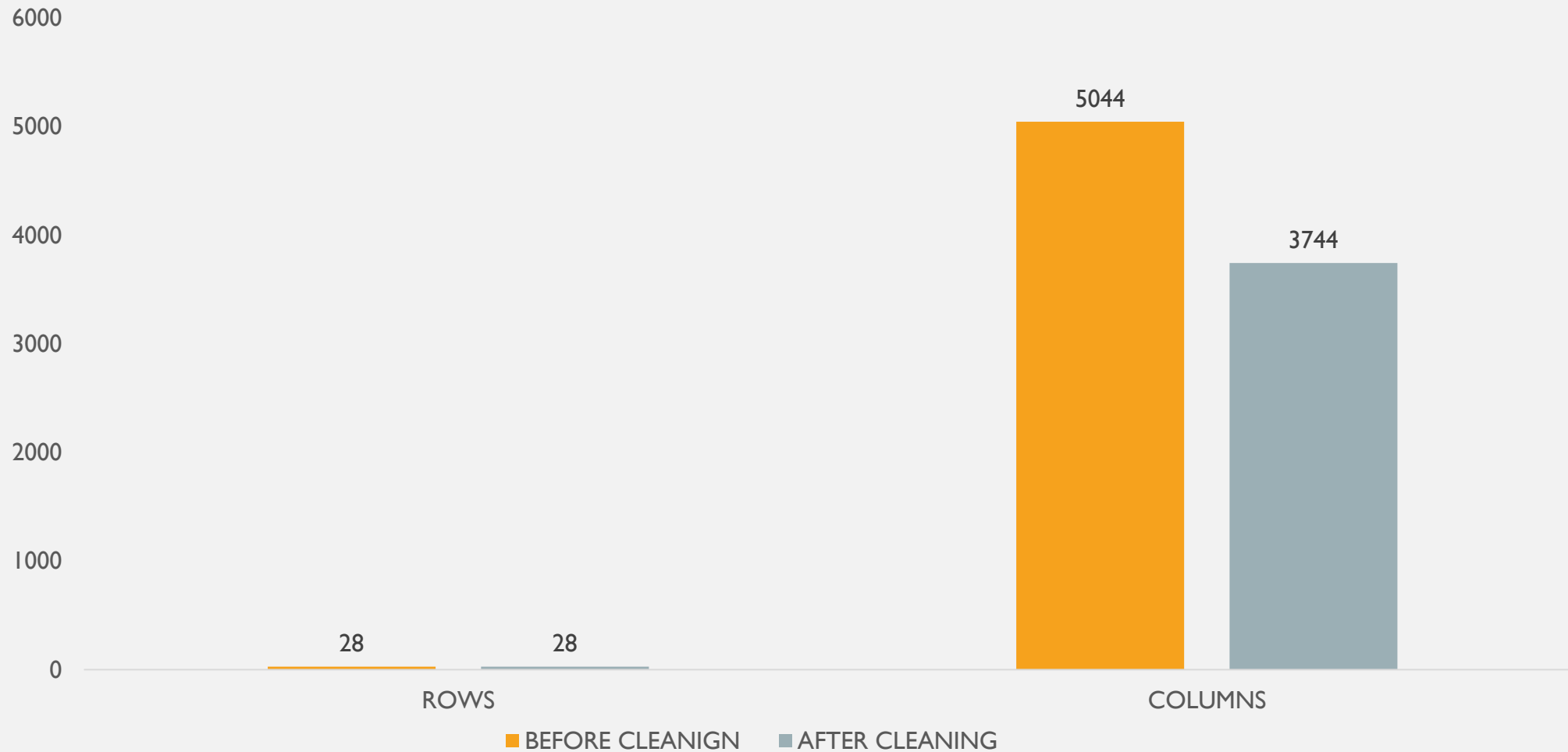
- Download and Data open in excel
- Find null and duplicate values and clean all data
- Data processing and solved as per asked problems.
- Use filter ,pivot table, sum if, average if, count if and other functions to give answer of asked questions
- Create charts to for easy and meaningful data representation
- Create report in ppt format and submit the project



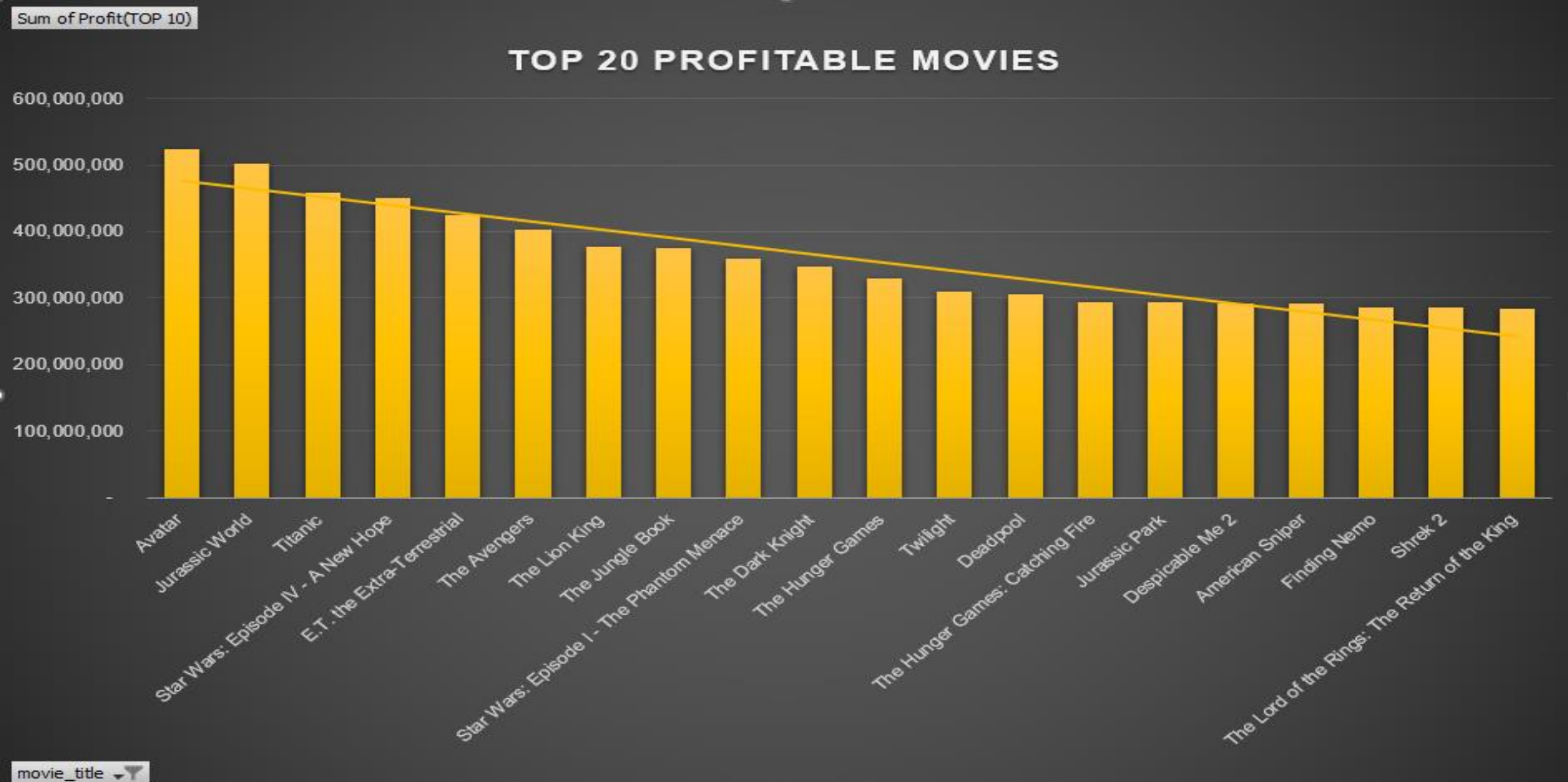
Google Drive



- 1. Clean the data-



2. Find the movies with the highest profit?



3.Find IMDB Top 20.

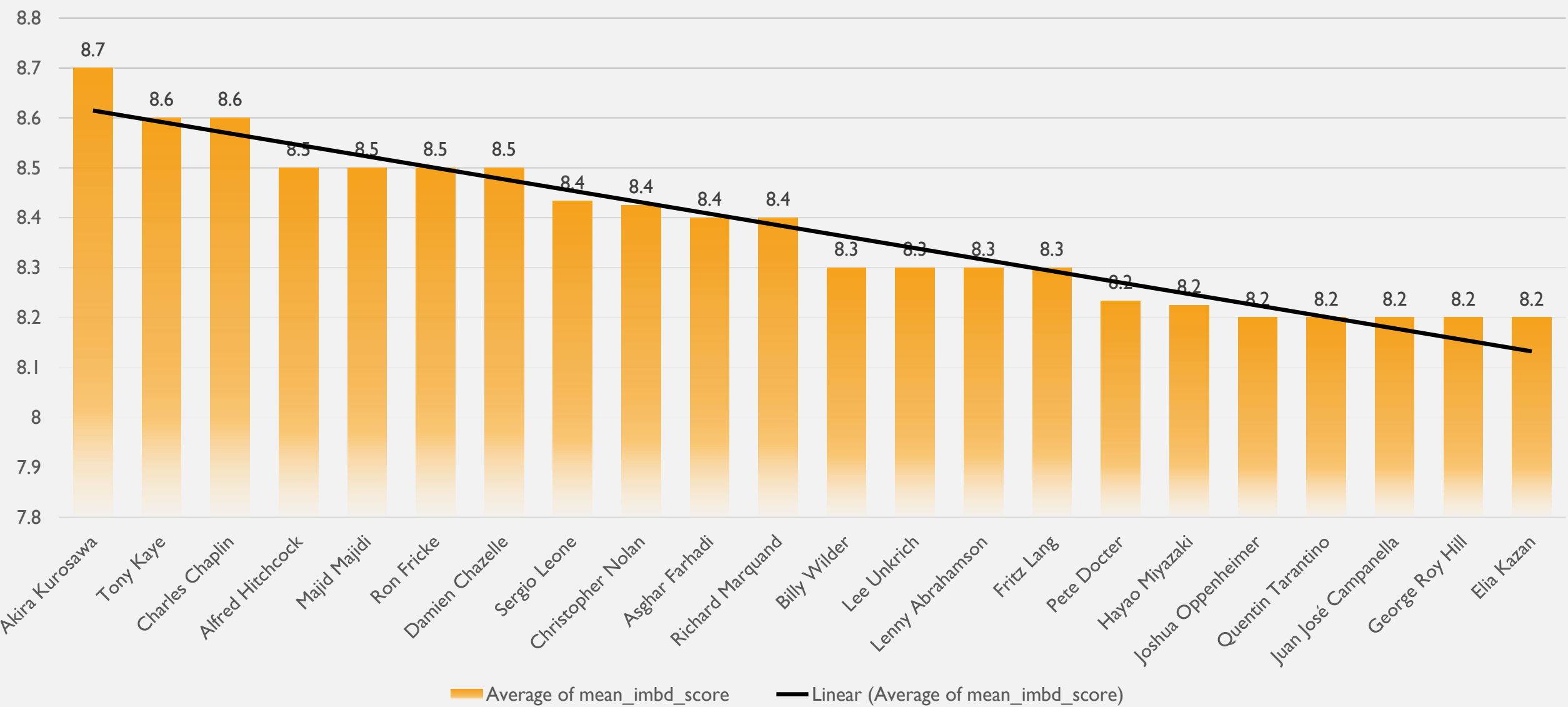


Movie_Title	IMDb_Top_250	Num_Voted_Users > 25,000
The Shawshank Redemption	9.3	1689764
The Godfather	9.2	1155770
The Dark Knight	9	1676169
The Godfather: Part II	9	790926
The Lord of the Rings:The Return of the King	8.9	1215718
Pulp Fiction	8.9	1324680
Schindler's List	8.9	865020
The Good, the Bad and 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

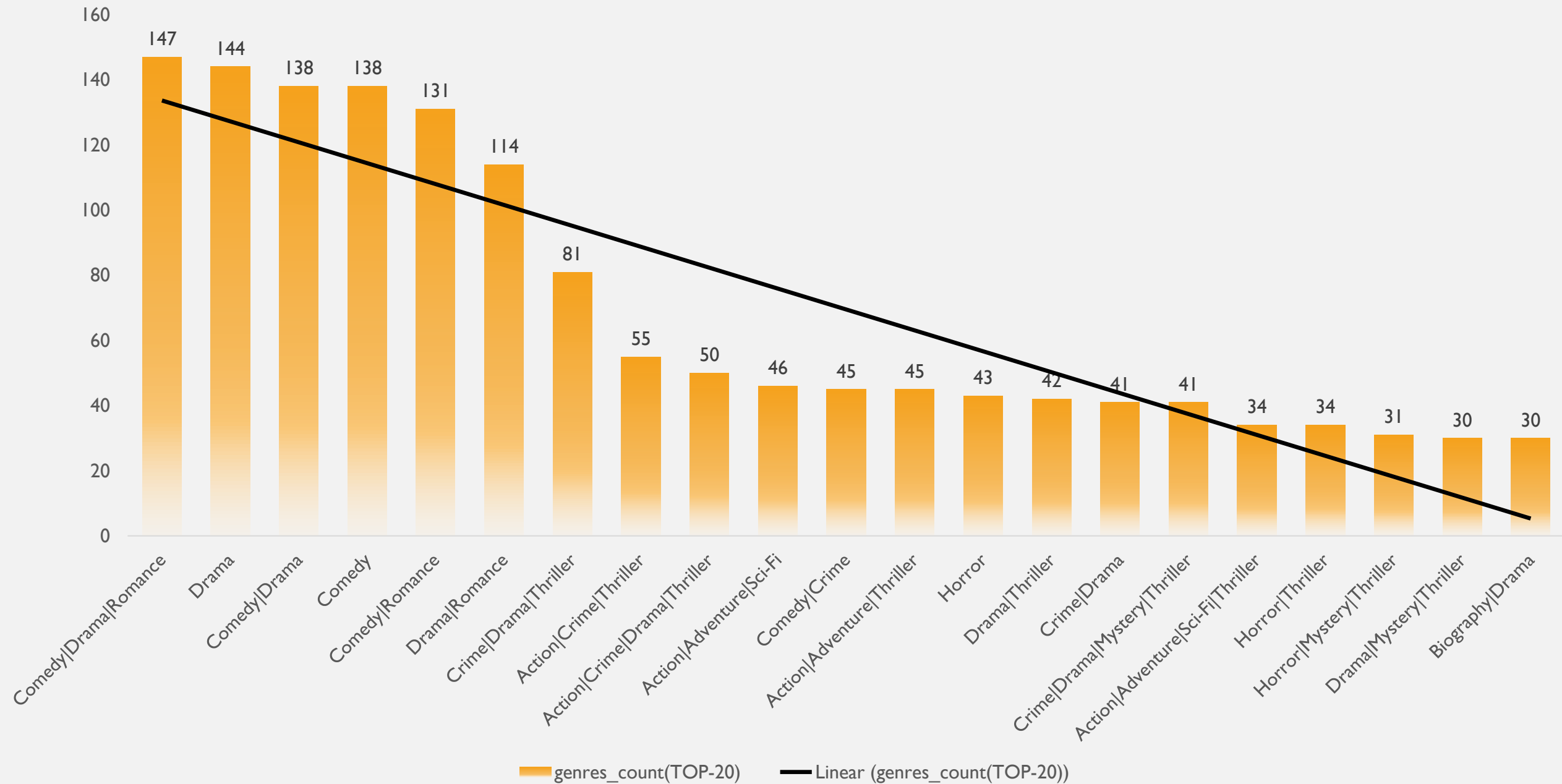
4.Find the best directors



AVERAGE OF MEAN_IMBD_SCORE

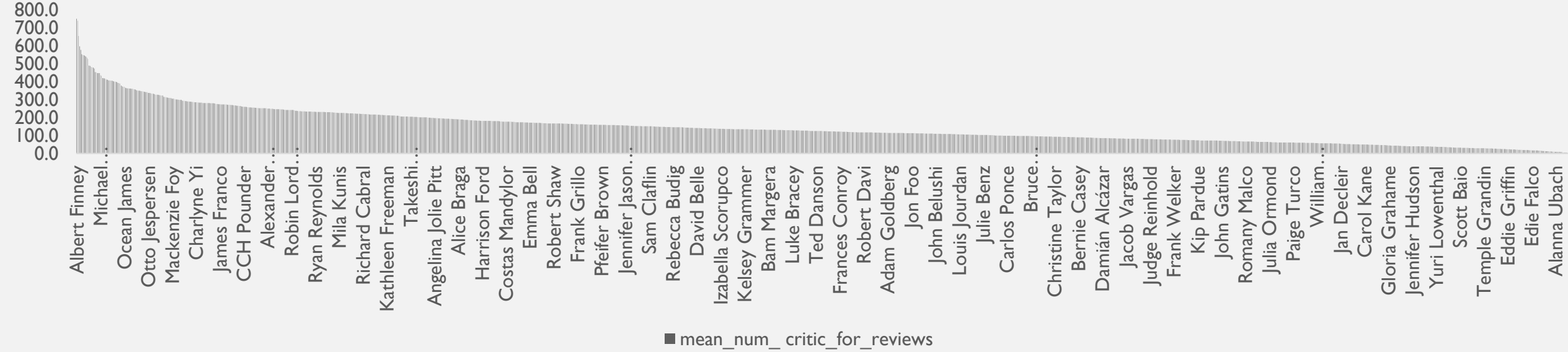


5.Find popular genres

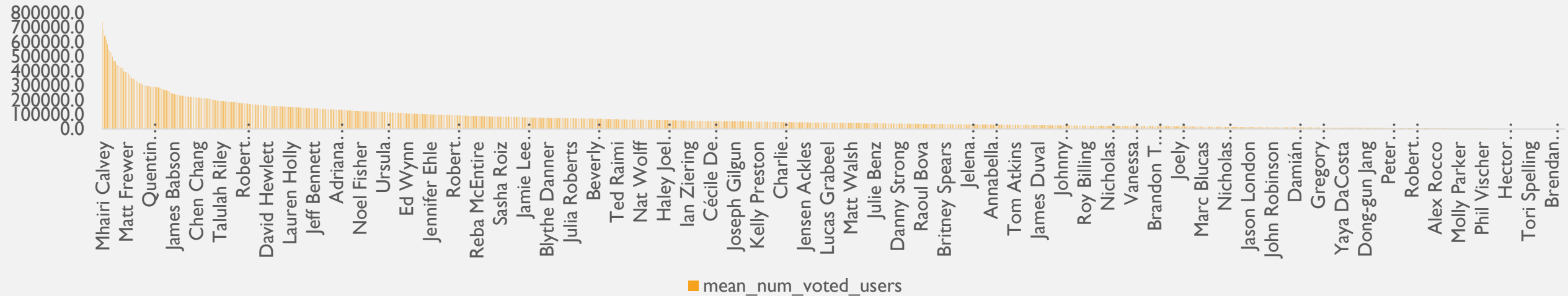


6.Find the critic-favorite and audience-favorite actors

mean_num_critic_for_reviews



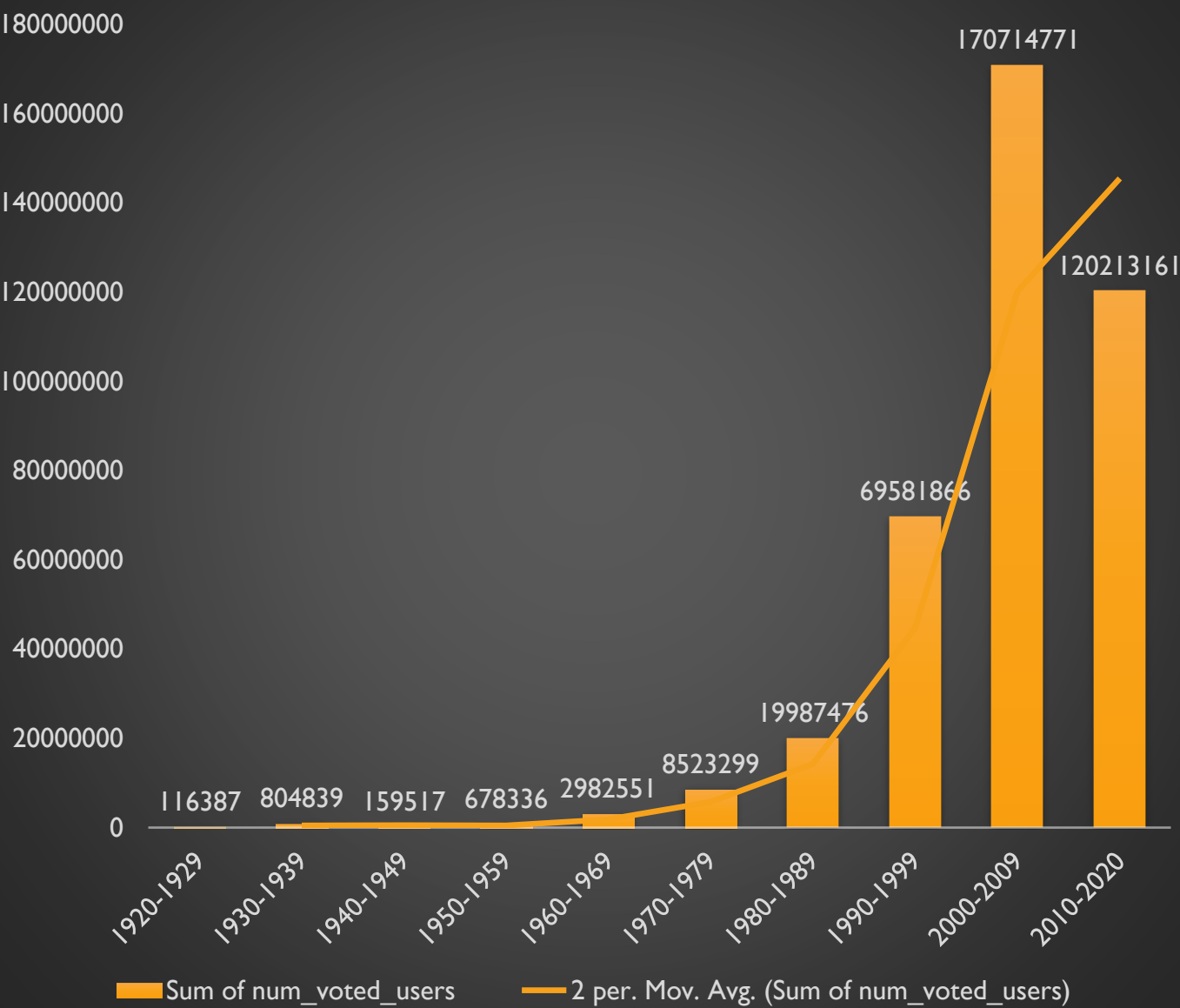
mean_num_voted_users



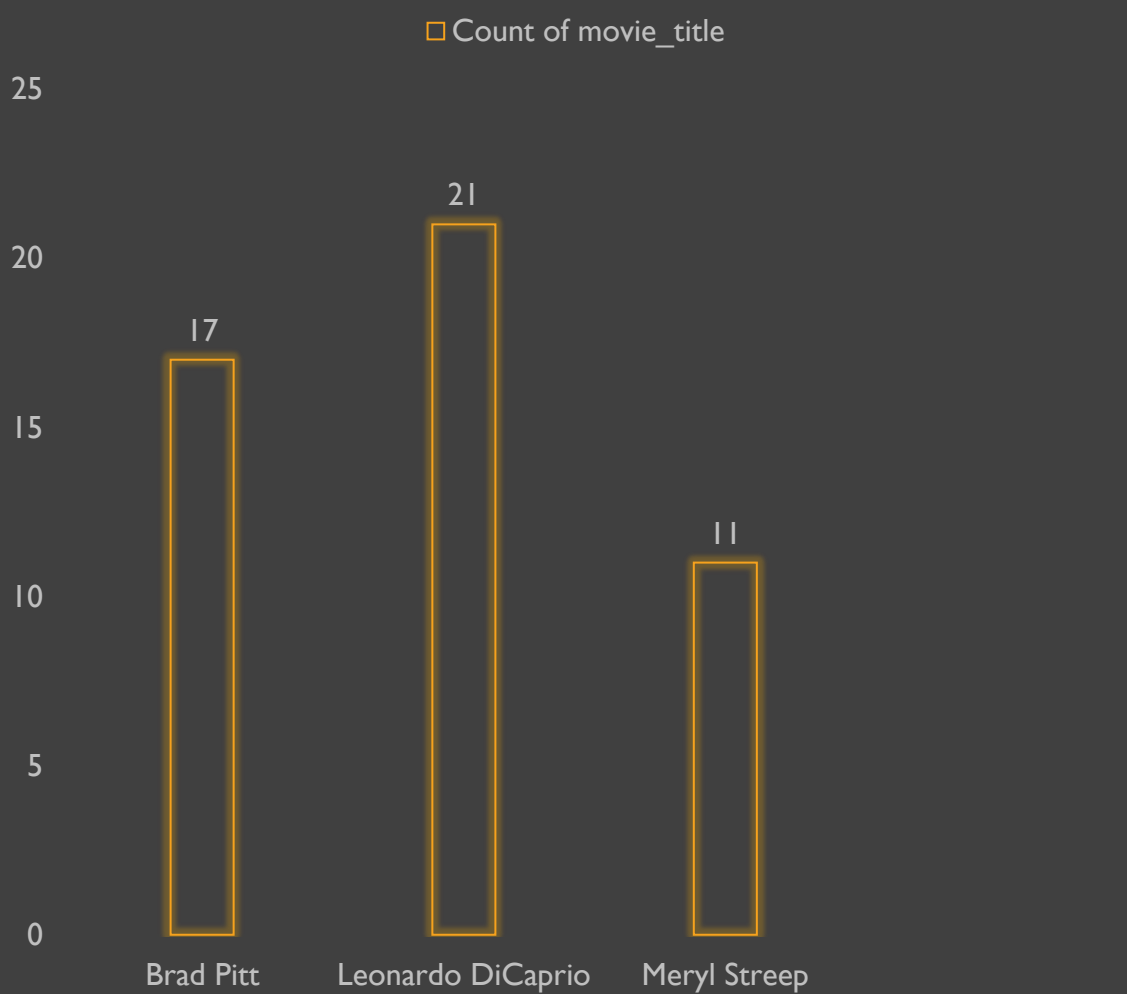
SUM OF NUM_VOTED_USERS DECADE WISE



Sum of num_voted_users



Count of Movie_Title
ACTOR_I_NAME_
MOVIE_COUNT('Meryl Streep',
'Leonardo DiCaprio', and 'Brad Pitt')



THE END



THANK YOU