

**Task 6**

# **IMDB MOVIE ANALYSIS**

**By Yash Rajput**  
**yashrajput.social@gmail.com**  
**9175034021**

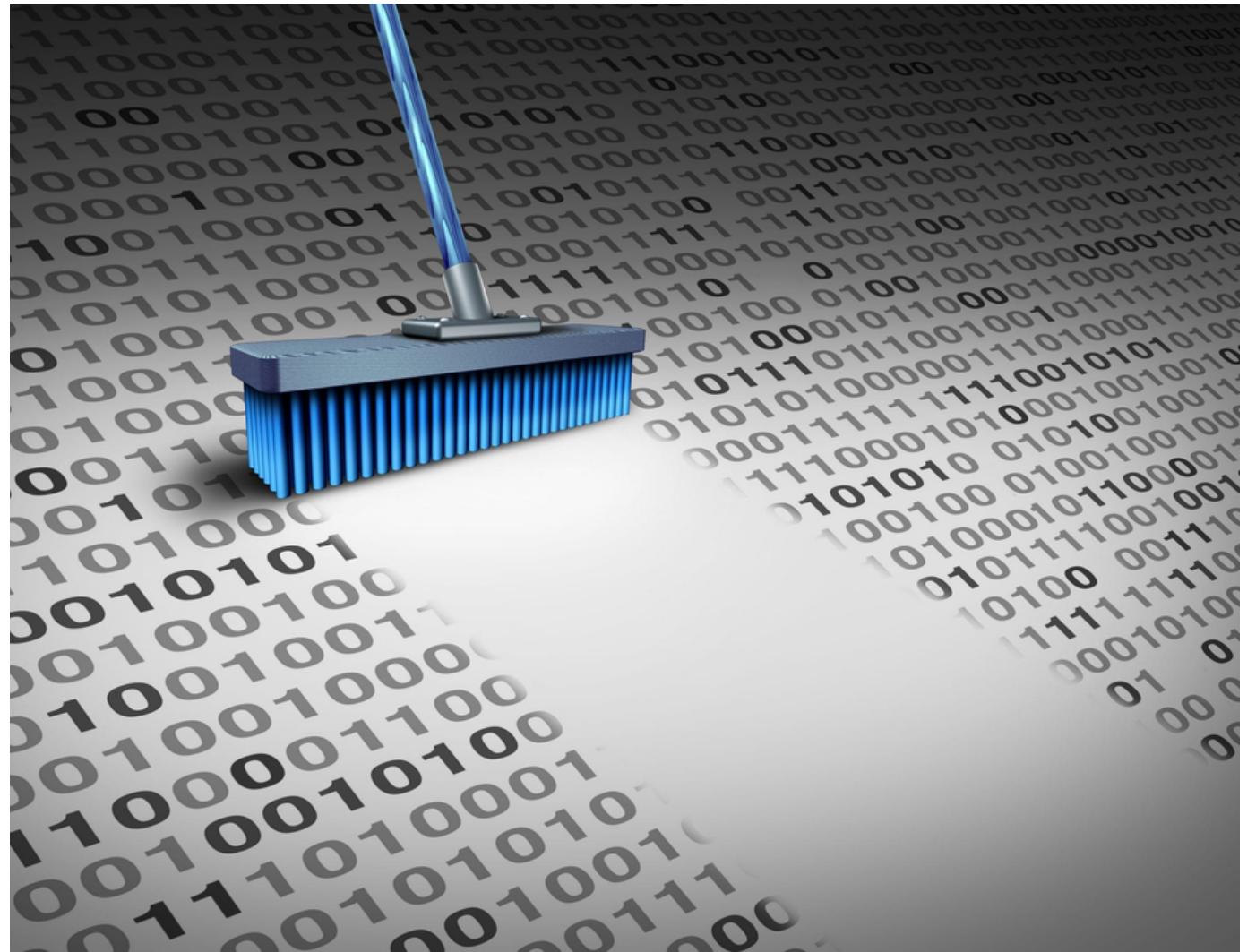
# **LIST OF CONTENTS**

- 1    ABOUT DATASET**
- 2    CLEANING OF THE DATA**
- A    MOVIE GENRE ANALYSIS**
- B    MOVIE DURATION ANALYSIS**
- C    LANGUAGE ANALYSIS**
- D    DIRECTOR ANALYSIS**
- E    BUDGET ANALYSIS**

# ABOUT DATASET

This data set contains the data of movies, data like director's name, genres, actors name, country, budget etc





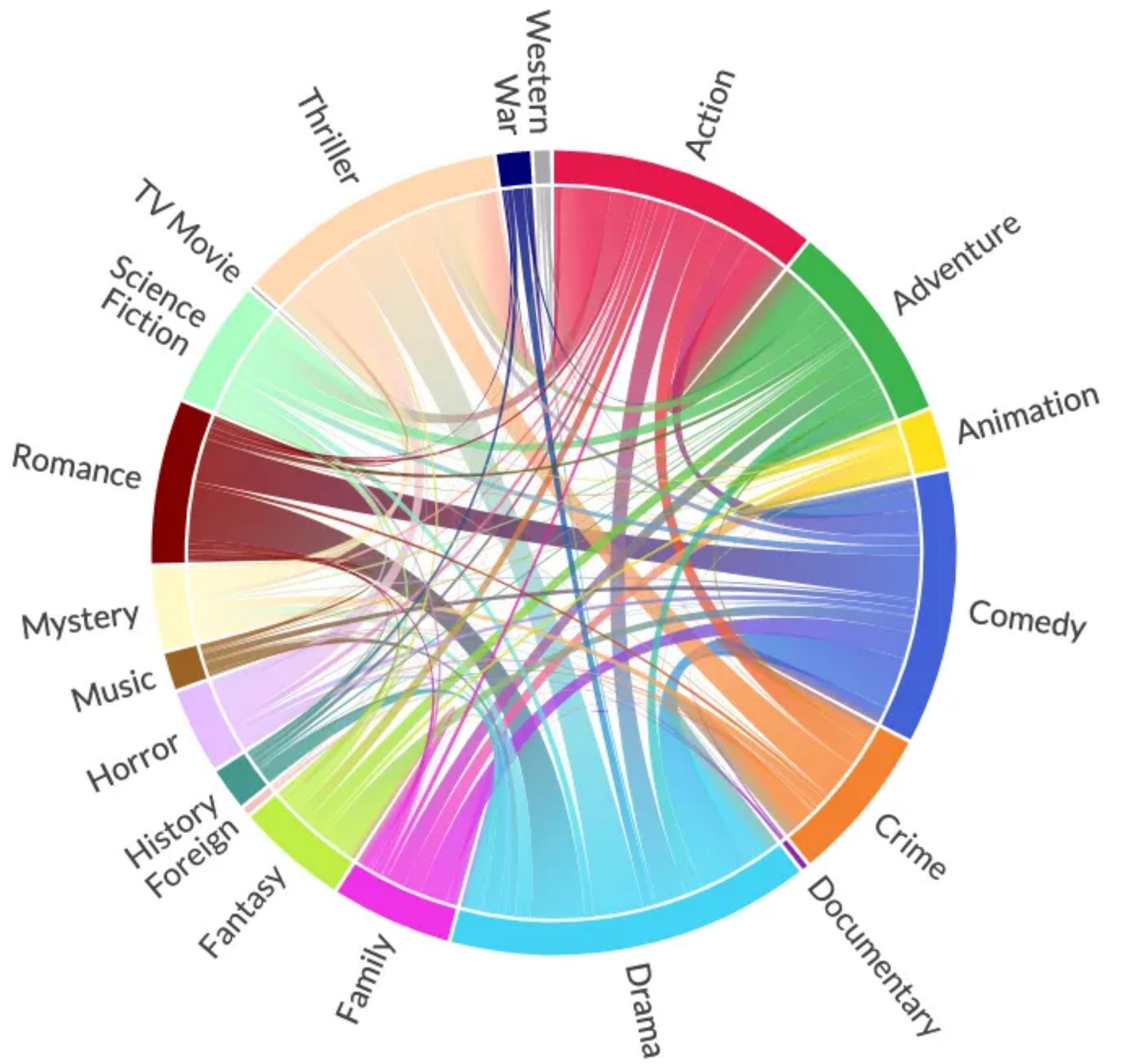
# CLEANING OF THE DATA

# DATA CLEANING

1. First dropping the columns which have no use for the analysis that we will be doing
2. Columns like 'Color', 'director\_facebook\_likes', 'actor\_3\_facebook\_likes', 'actor\_2\_name', 'actor\_1\_facebook\_likes', 'cast\_total\_facebook\_likes', 'actor\_3\_name', 'facenumber\_in\_posts', 'plot\_keywords', 'movie\_imdb\_link', 'content\_rating', 'actor\_2\_facebook\_likes', 'aspect\_ratio', 'movie\_facebook\_likes' are the columns containing irrelevant data for the analysis tasks provided. So, these columns needs to be dropped.
3. After dropping the irrelevant columns now we need to remove the rows from the dataset having anyone of its column value as blank/NULL
4. Then we need to get rid off the duplicate values in the dataset which can be achieved by using the 'Remove Duplicate Values/Cells' available in the 'Data' tab



[GOOGLE SHEET](#)

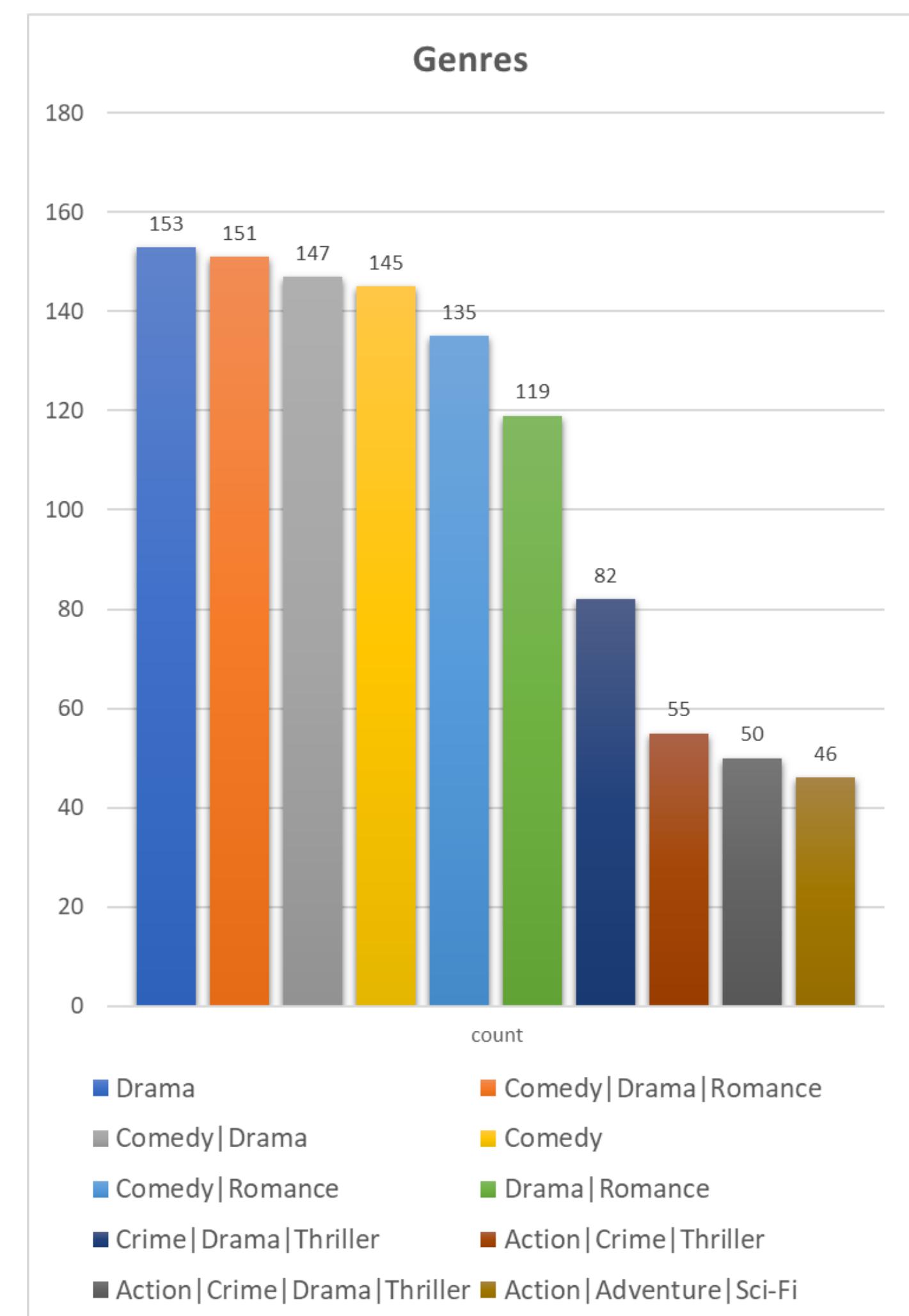


# MOVIE GENRE ANALYSIS

# GENRES ANALYSIS

Finding the count of the movies of a genres using a  
Pivot Table

## MOVIE GENRES LINK



# GENRES VS IMDB

Finding the count of the movies of a genres using a  
Pivot Table

## GENRES VS IMDB SCORE



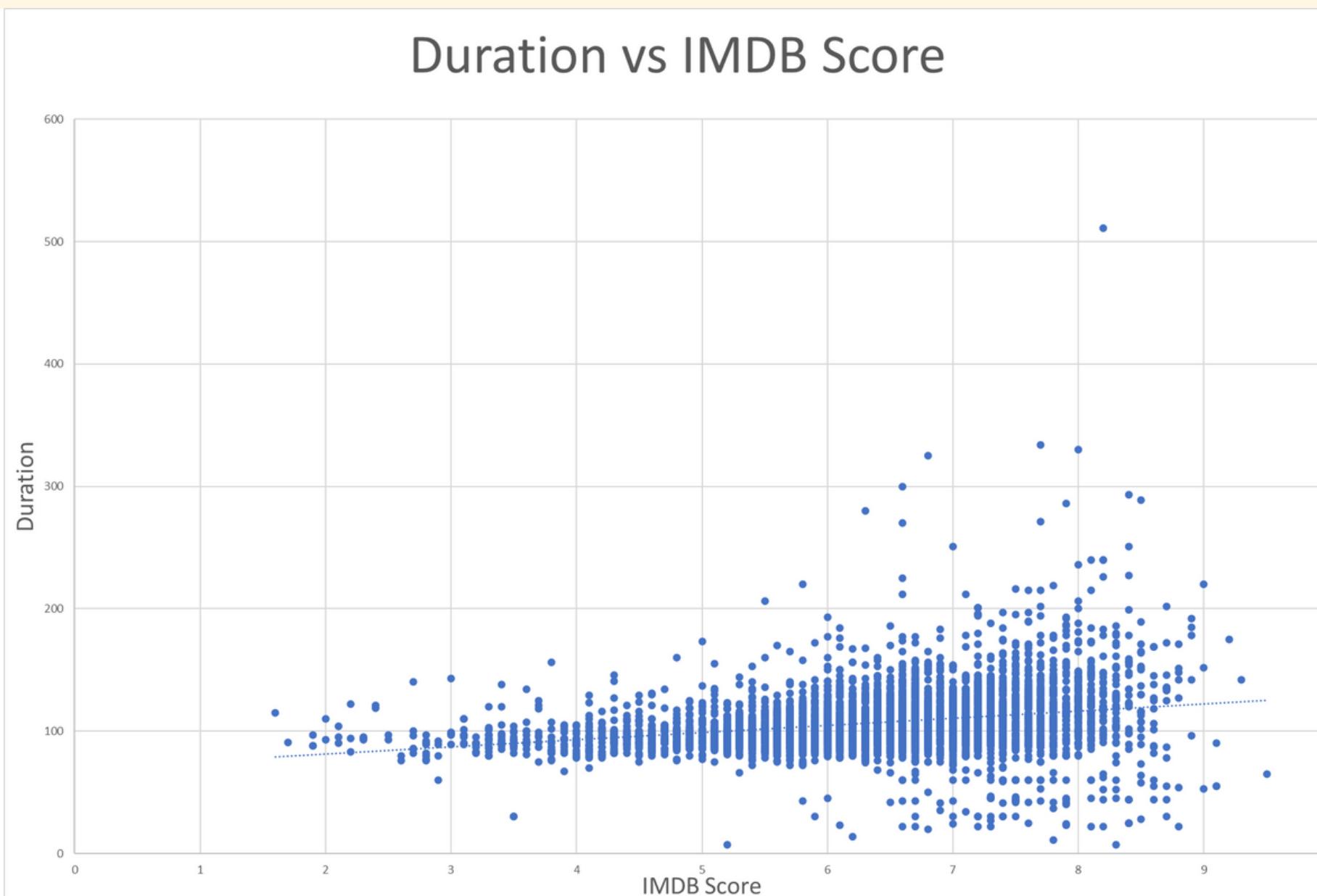
Top Genres	IMDB Score
Adventure   Animation   Drama   Family   Musical	8.5
Crime   Drama   Fantasy   Mystery	8.5
Action   Adventure   Drama   Fantasy   War	8.4
Adventure   Animation   Fantasy	8.4
Adventure   Drama   Thriller   War	8.4
Adventure   Animation   Comedy   Drama   Family   Fantasy	8.3
Adventure   Drama   War	8.3
Biography   Drama   History   Music	8.3
Documentary   Drama   Sport	8.3
Documentary   War	8.3
Biography   Crime   Documentary   History	8.2
Drama   Fantasy   War	8.2
Drama   Mystery   War	8.2

# MOVIE DURATION ANALYSIS



# DURATION ANALYSIS

## DURATION ANALYSIS



## KEY INSIGHTS

1. Average Duration: 107.2 minutes
2. Average Earning per min: 38519.6 USD
3. Shortest Movie: Shaun The Sheep and The Touch
4. Longest Movie: Trapped

Duration	Avg IMDB Score	Number of Movies
<120	6.26	3884
120-240	7.4	1132
240-360	7.62	15
>360	8.2	1

## REGRESSION ANALYSIS

Regression Statistics	
Multiple R	0.969290389
R Square	0.939523858
Adjusted R Square	0.939324933
Standard Error	1.608455107
Observations	5028

Duration as input(X)  
IMDB Score as output(Y)

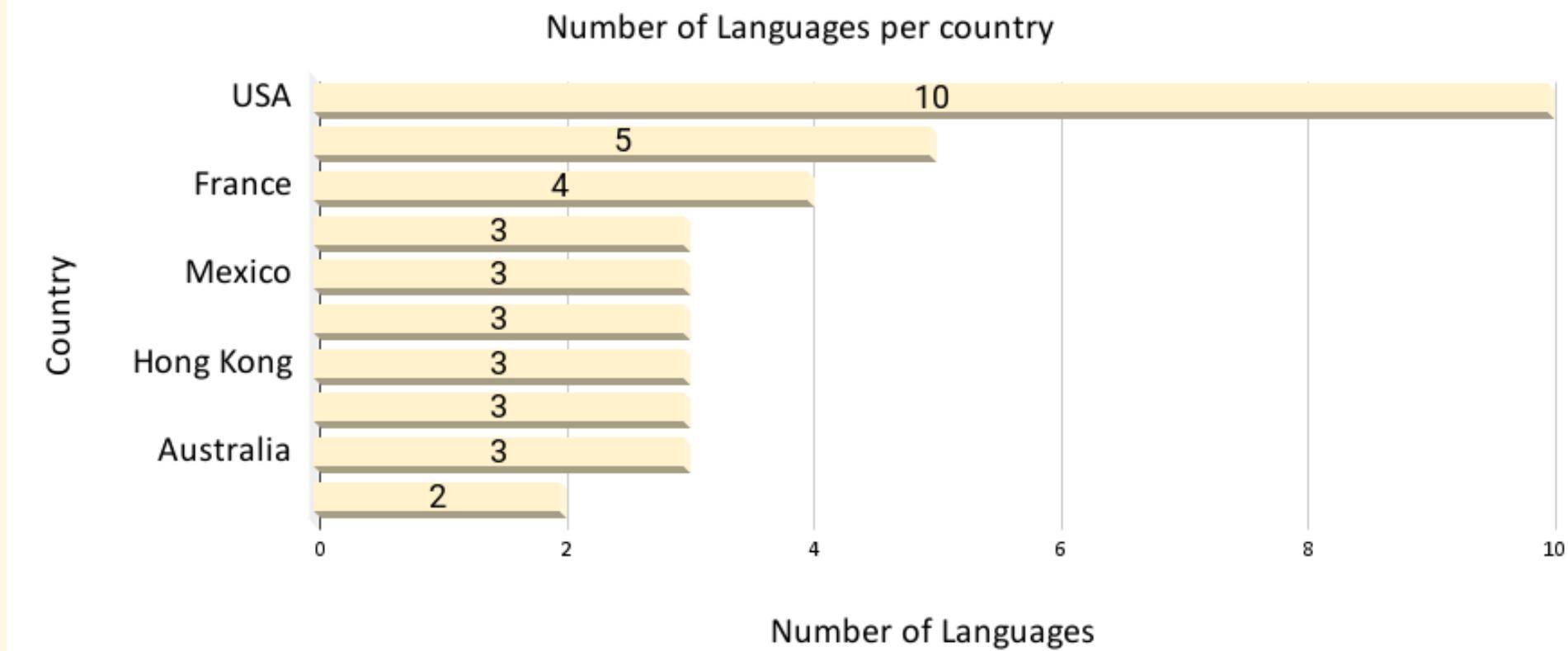
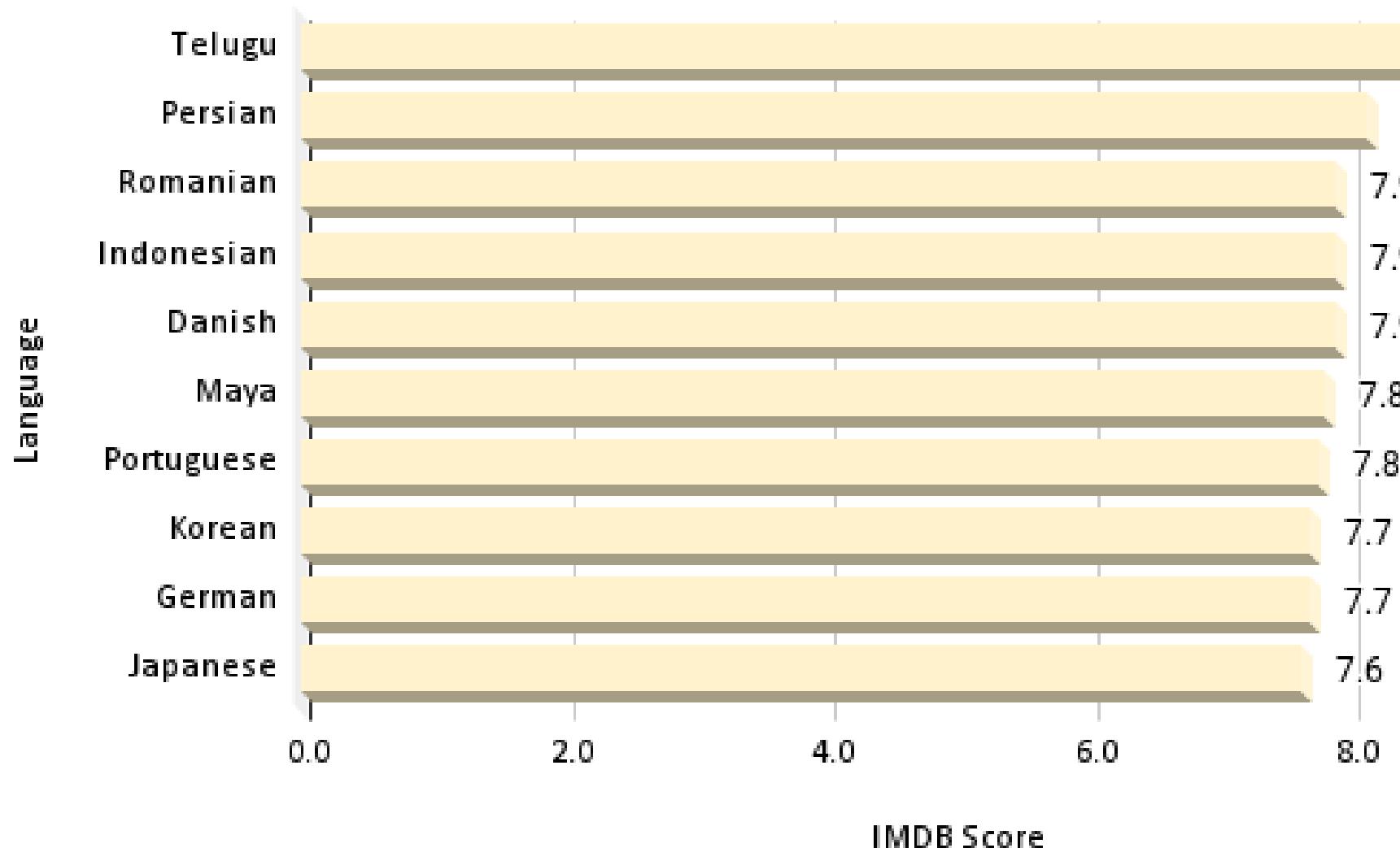
# LANGUAGE ANALYSIS



# LANGUAGE ANALYSIS

## LANGUAGE ANALYSIS

Language vs Avg IMDB Score



Top 10 Languages		Top 10 Languages	
Language	Count	Language	Earning
English	3669	English	195277417291
French	37	Spanish	225331009
Spanish	26	French	218613624
Mandarin	14	Mandarin	162767817
German	13	Aboriginal	78680789
Japanese	12	Cantonese	64040464
Hindi	10	Japanese	57216469
Cantonese	8	Maya	50859889
Italian	7	German	43678911
Portuguese	5	Hindi	33103751

KEY INSIGHT: ABORIGINAL FILMS ARE LESS BUT HAVE PERFORMED FAIRLY WELL

# DIRECTOR ANALYSIS



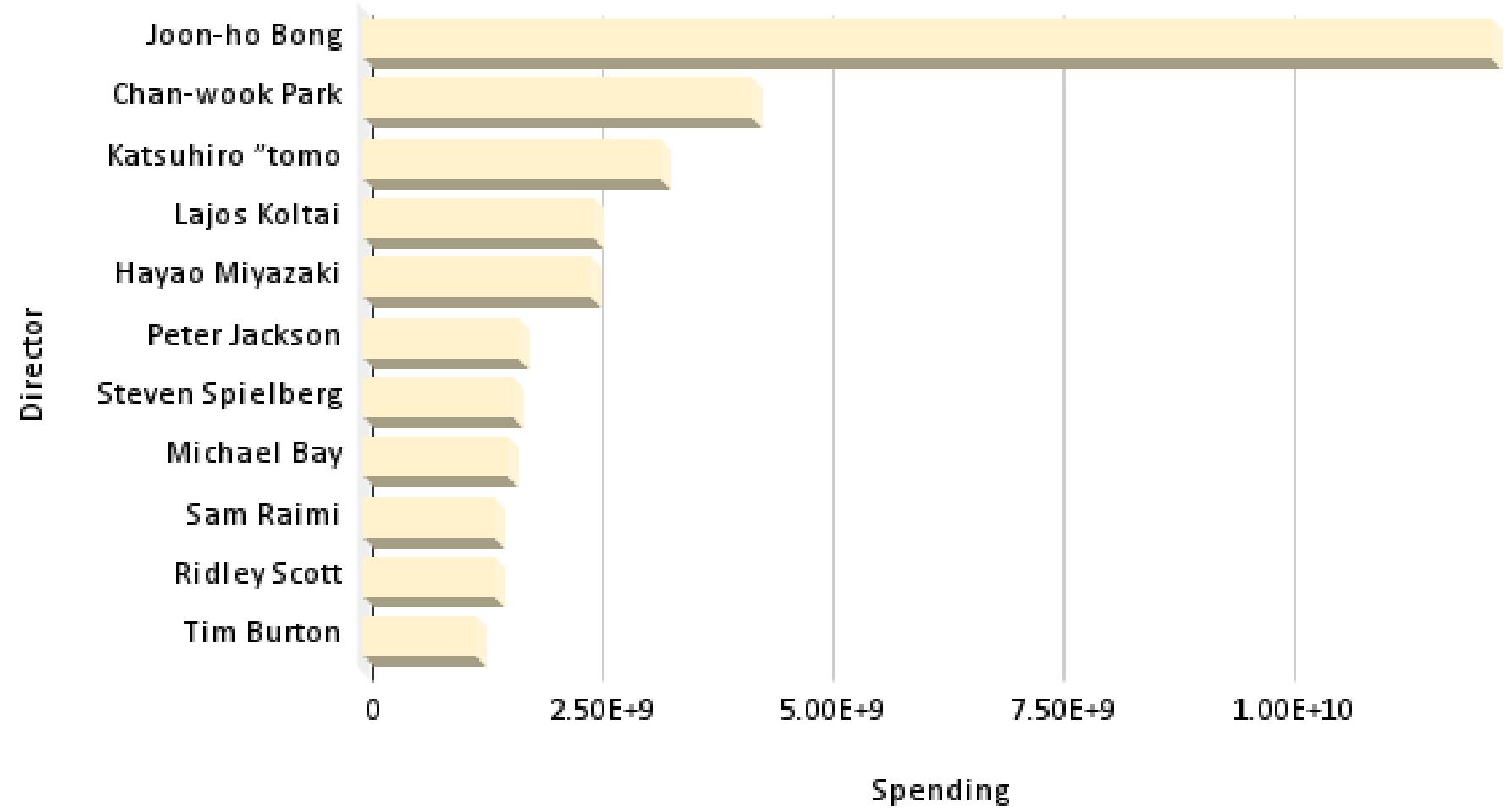
# DIRECTOR ANALYSIS

Based on Mean IMDB Scores

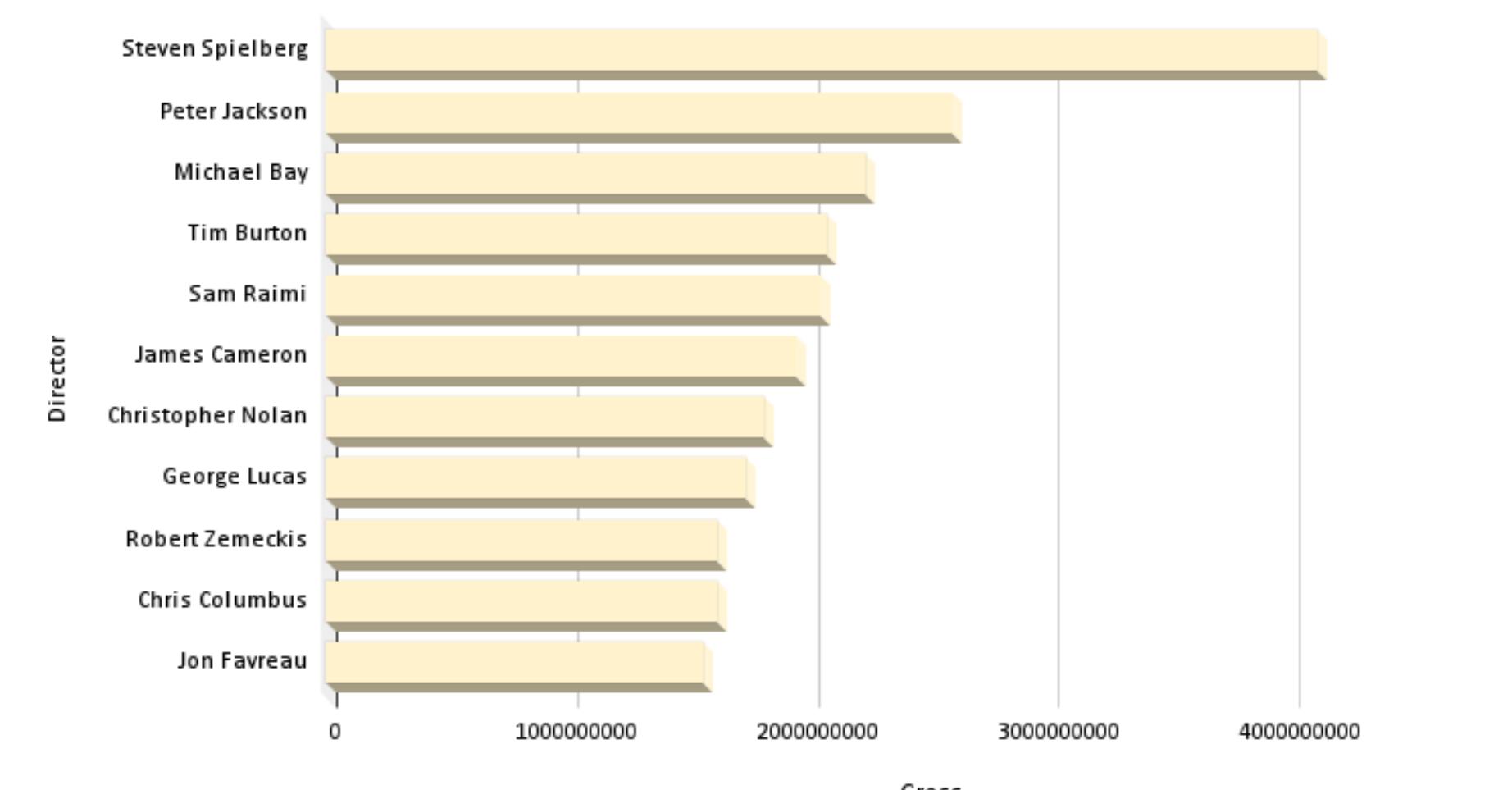
## BEST DIRECTORS

Director Name	Mean of Imdb Score	Percentile
Tony Kaye	8.60	99.94279176
Charles Chaplin	8.60	99.94279176
Ron Fricke	8.50	99.71395881
Majid Majidi	8.50	99.71395881
Damien Chazelle	8.50	99.71395881
Alfred Hitchcock	8.50	99.71395881
Sergio Leone	8.43	99.65675057
Christopher Nolan	8.43	99.59954233
S.S. Rajamouli	8.40	99.37070938
Richard Marquand	8.40	99.37070938

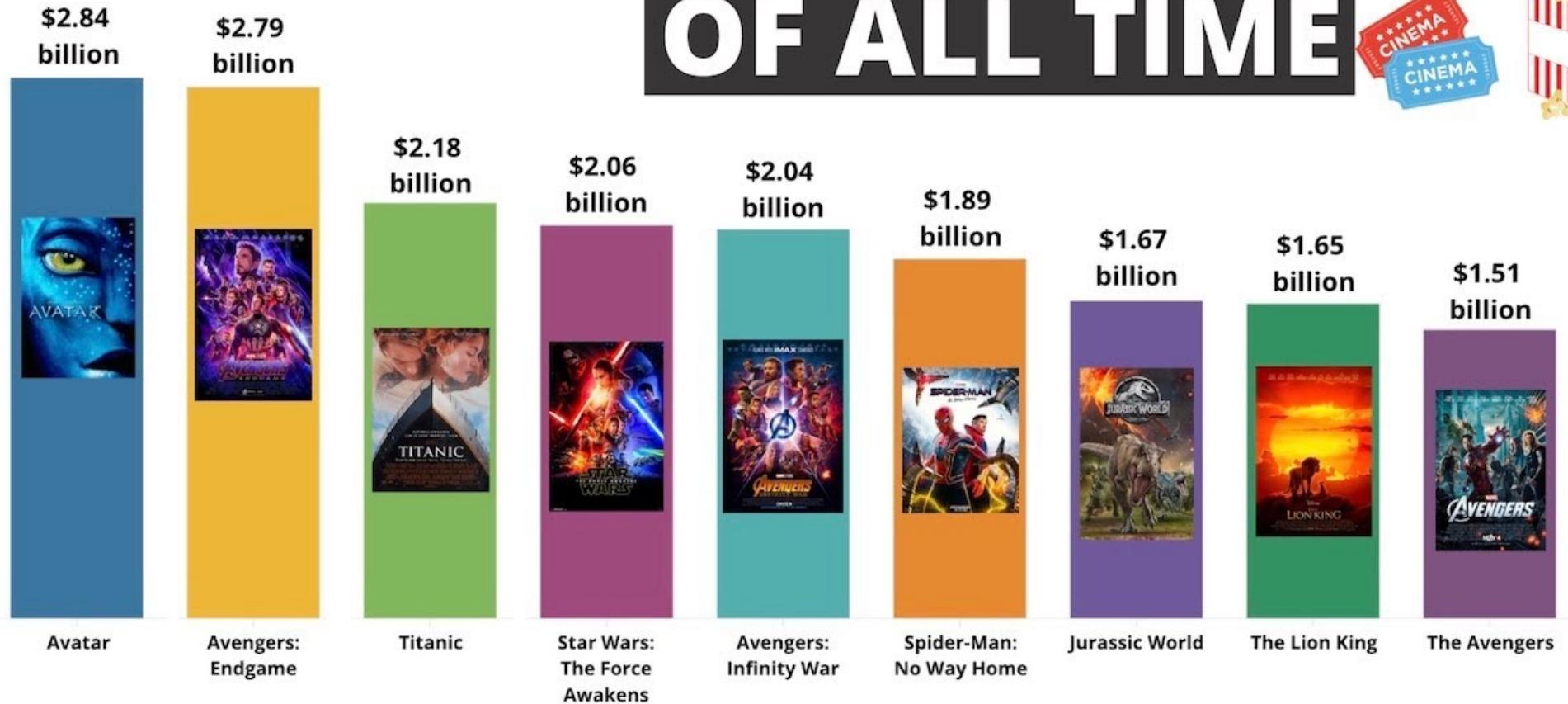
Spending vs Director



Gross vs Director



# TOP GROSSING MOVIES OF ALL TIME



# BUDGET ANALYSIS

# MOVIES WITH HIGHEST PROFIT

Highest Profit Is found by Gross - Budget

**Highest Grossing Movie: Avatar**

**Movie with Highest Budget: The Host**

**Movie with Highest Profit: Avatar**

## BUDGET ANALYSIS

Movie Title	SUM of Gross
Avatar	760505847
The Jungle Book	725290282
Spider-Man 3	673060606
Alice in Wonderland	668370412
Titanic	658672302
King Kong	654153780
Jurassic World	652177271
The Avengers	623279547
Skyfall	608720554
The Dark Knight	533316061
Home	532031025

Movie Title	Budget
The Host	12255500000
Lady Vengeance	42000000000
Fateless	25000000000
Princess Mononoke	24000000000
Steamboy	2127519898
Akira	11000000000
Godzilla 2000	10000000000
Tango	7000000000
Kabhi Alvida Naa Kehna	7000000000
King Kong	6210000000

Movie Title	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
The Hunger Games	329999255

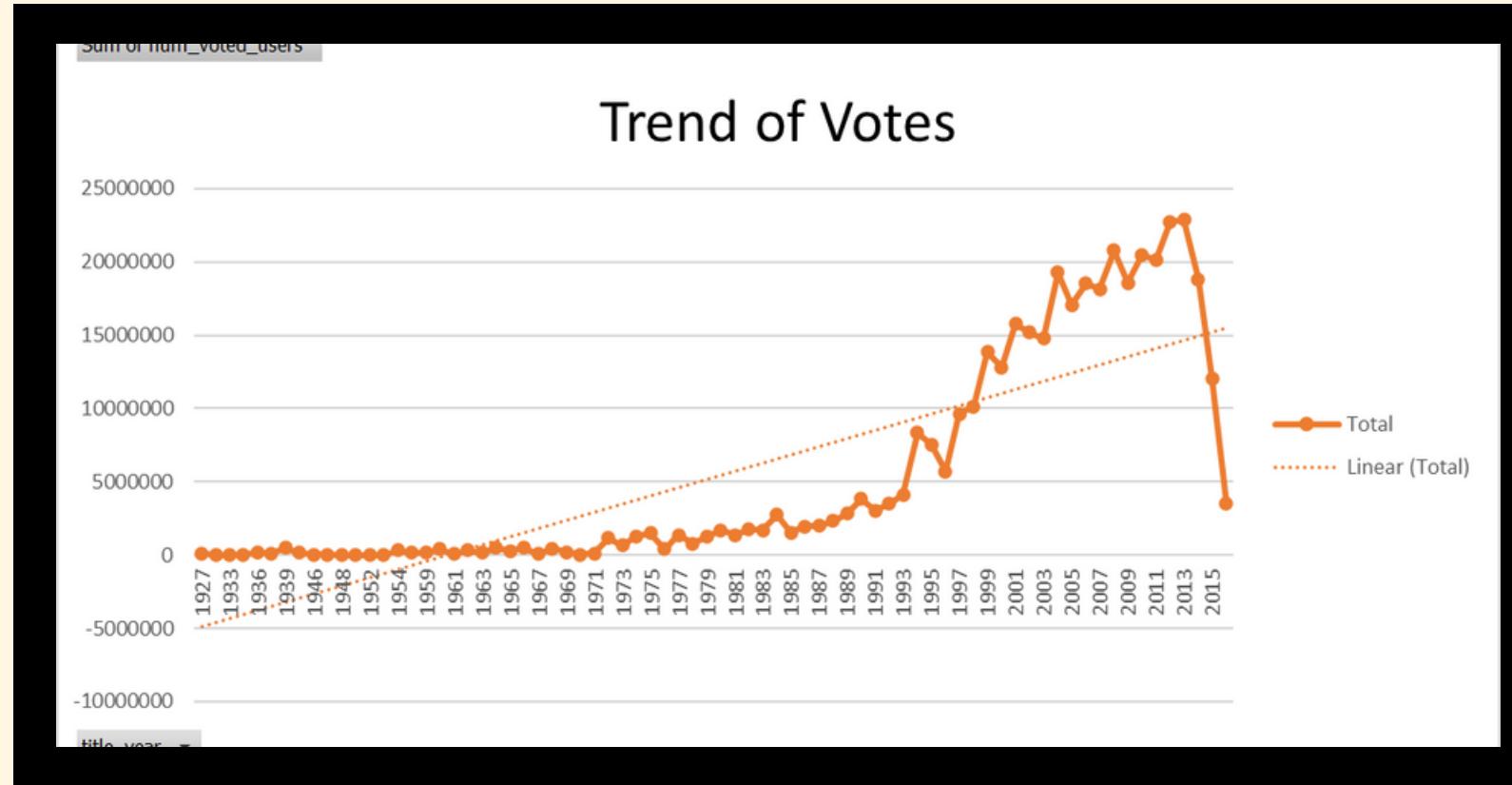
**Top 10 with Highest Grossing Movies**

**Top 10 with Highest Budget**

**Top 10 movies with Highest Profit**

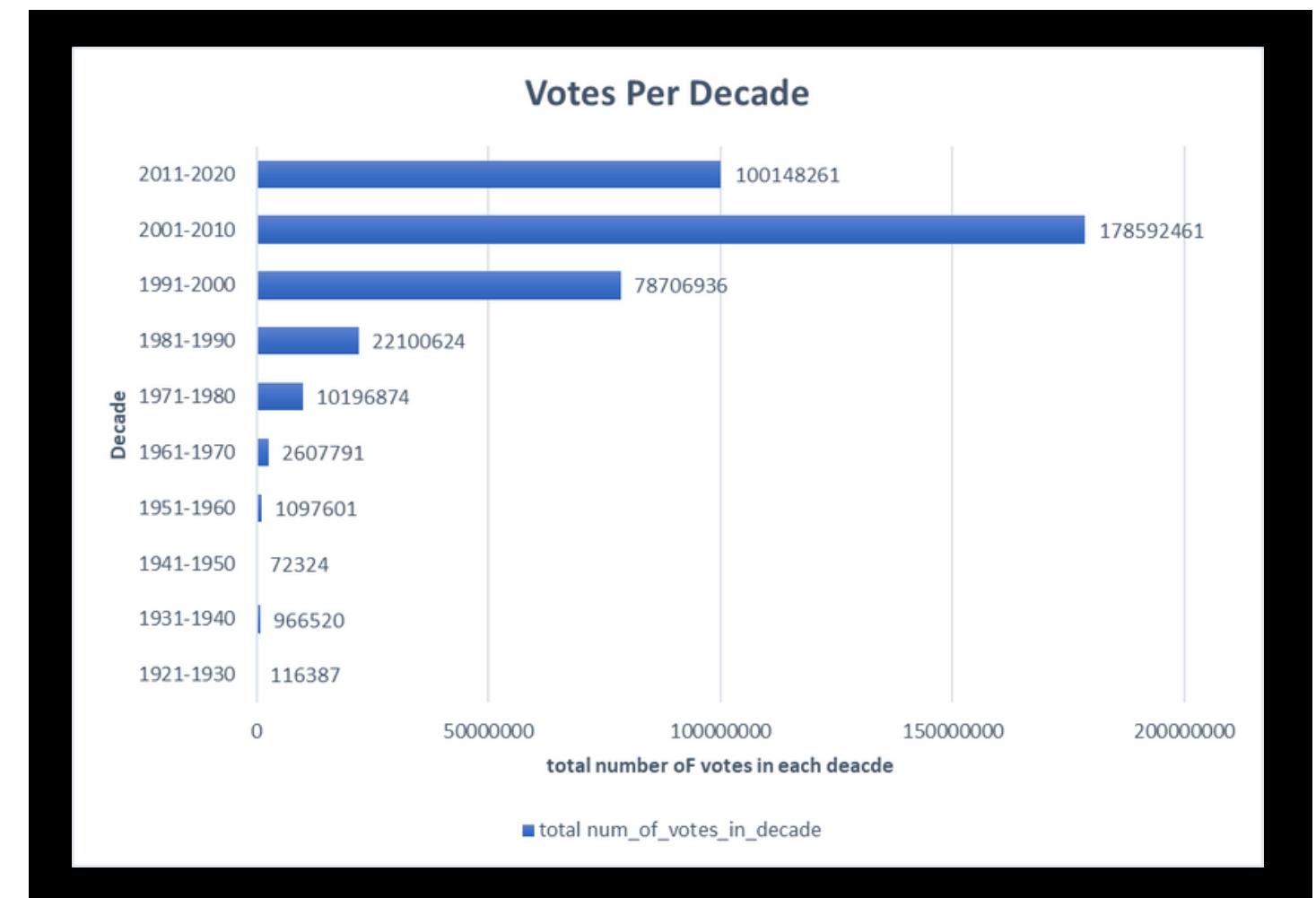
# OTHER CHARTS





Here we can see the trend of votes and can conclude that there was a downfall in the year 2016

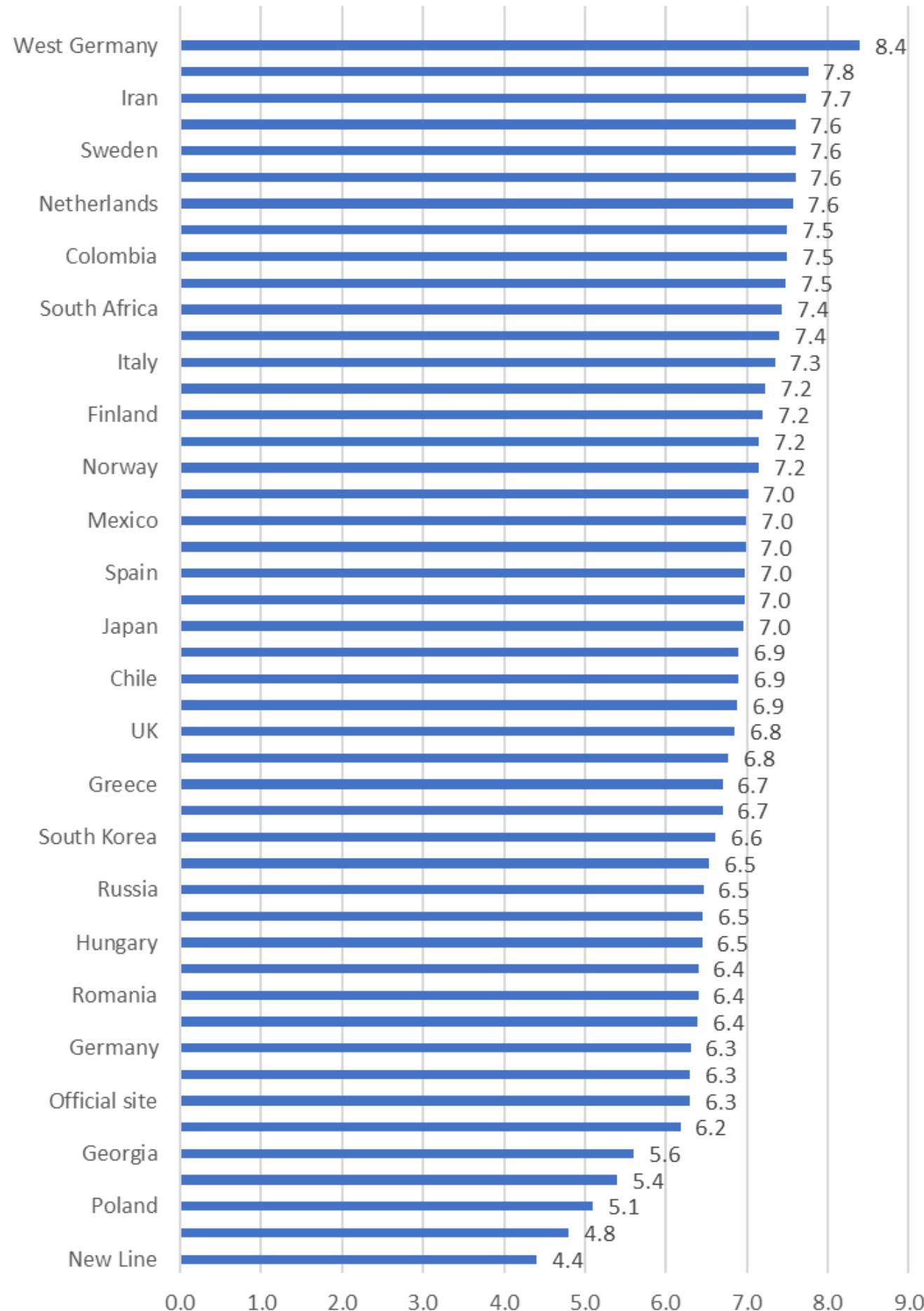
Here we can see that the decade 2001-2010 was the most active decade in the history with the maximum votings done by public

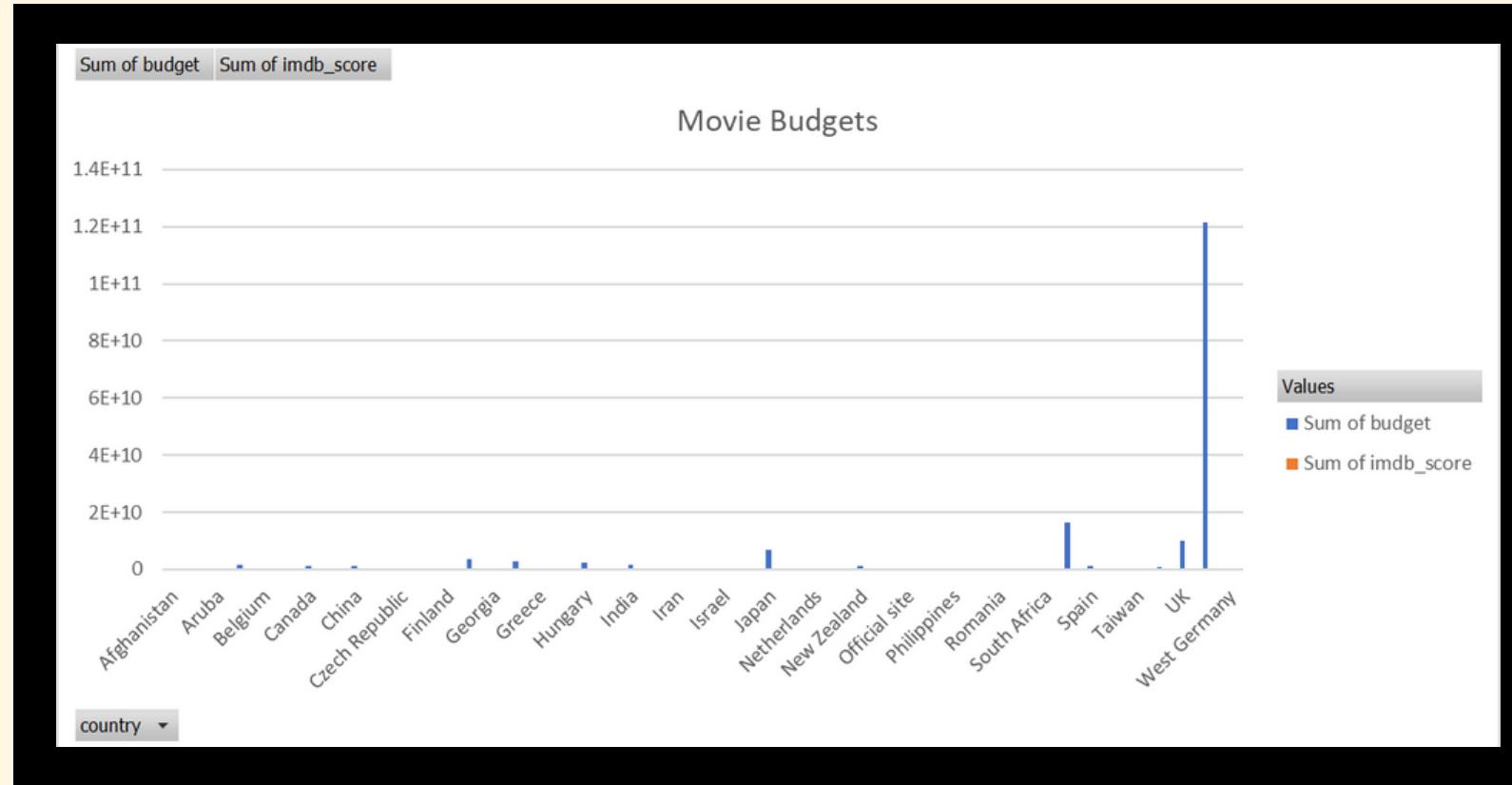


# COUNTRIES WITH BEST MOVIES

West Germany and Iran are leaders in making the best movies

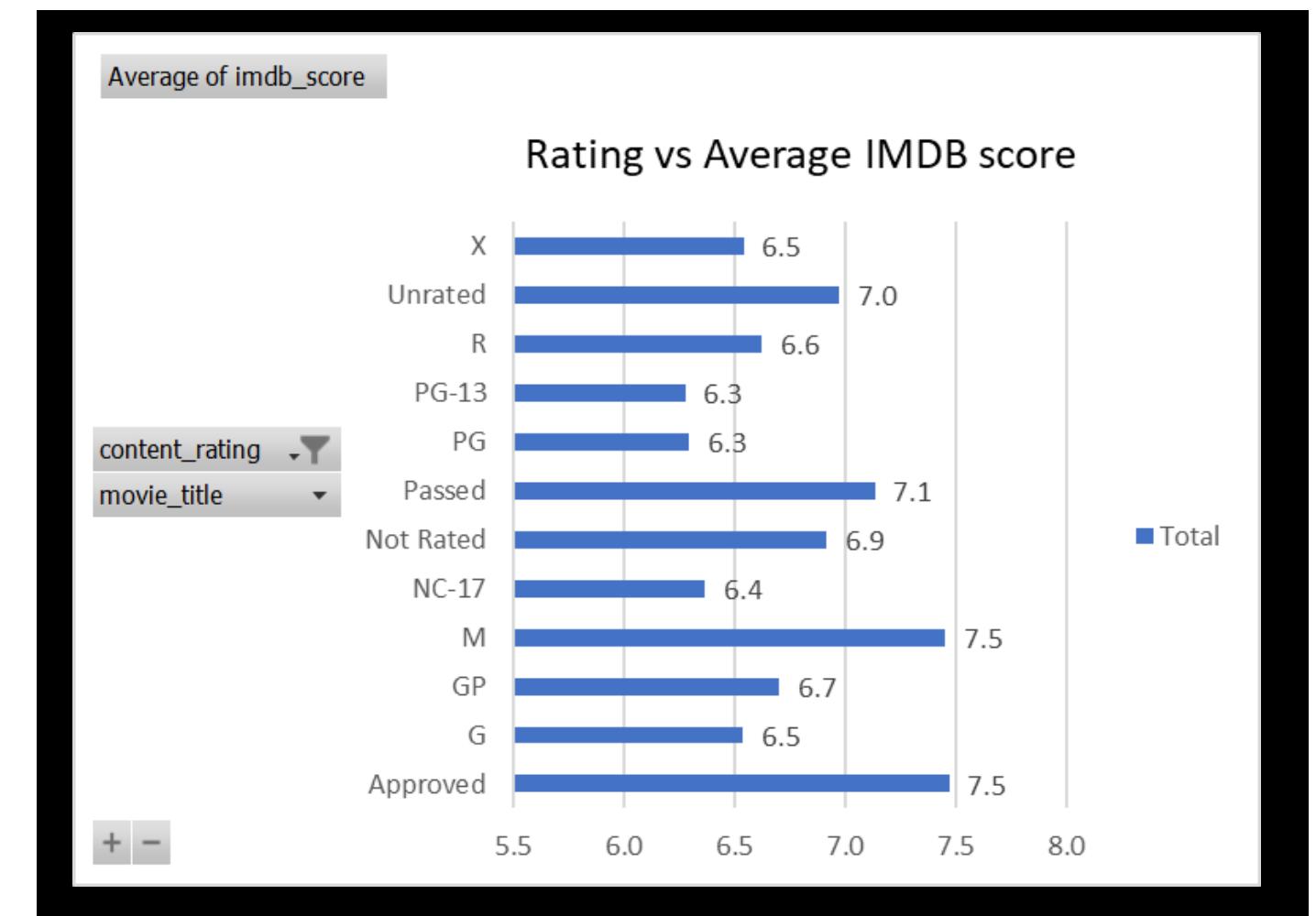
Average of IMDB score





The expense of USA is extremely high compared to any other country

It is seen that M rated and Approved movies do well compared to others



# THANK YOU

Yash Rajput  
Yashrajput.social@gmail.com