

IMDB

Movie Analysis

By- ANJALI RAJPUT

Project Description

- In this project the Datasets was provided and we were required to frame the problem.
- I was expected to define my own problems and expected to shed light on it.
- We are going to analyse the data by EDA(Exploratory Data Analysis) which is used to summarize ,visualize and understand the characteristics of the given dataset.

The Root Cause Analysis approach or 5 Why's Analysis approach developed b SAKICHI TOYODA, founder of Toyota was used to create the report which convey the data story.

Approach

Approach towards the project means Overall strategy that outlines how the project will be executed or done to achieve its Objectives.

- In this project first we need to import the dataset that provided and load that into Excel.
- After loading the data use specific functions and formulae in excel in order to achieve the clean dataset to drive required insights.
- Using 5 why's approach for the deep understanding of the data and root cause problems.
- Analysing the data for the various problem statements.
- Using statistical and visualization tools to shed lights on results.

Tech-stack Used:

- Ms excel (for analysis)
- MS Power-point(for making ppt of project)



Excel

Cleaning data

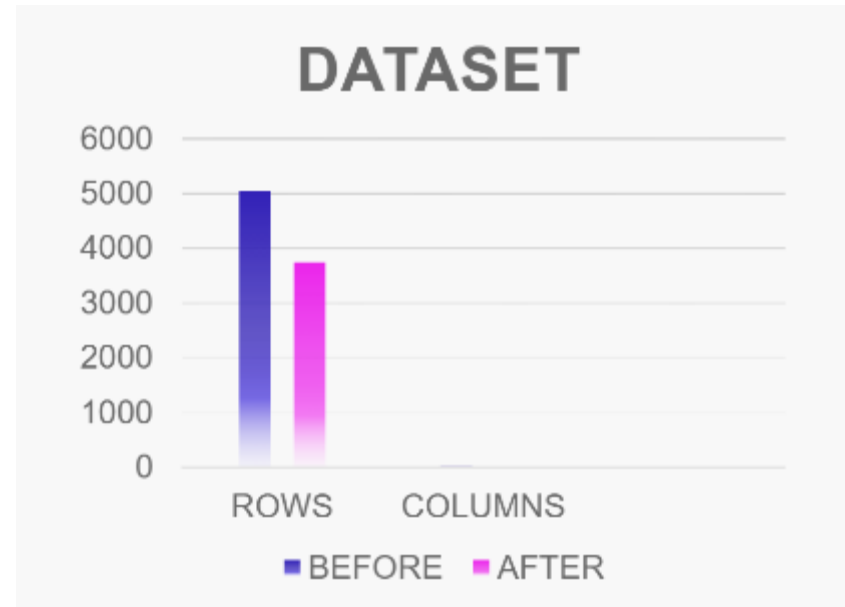
Before Cleaning --

- Total number of ROWS- 5044
- Total number of COLUMNS- 28

After Cleaning --

- Total number of ROWS- 3724
- Total number of COLUMNS- 13

In the process of cleaning, 45 duplicates were removed and 3724 rows were obtained after removing the rows having blank cell and the unnecessary columns





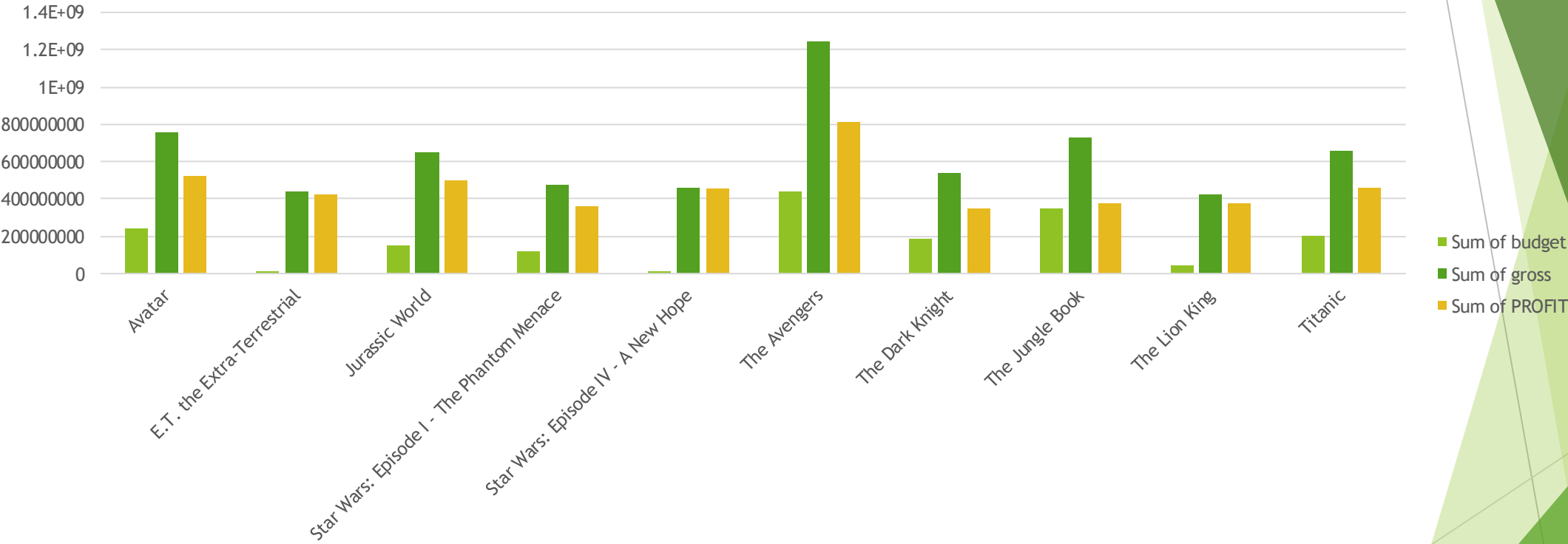
Insights

MOST PROFITABLE MOVIES

TOP 10 MOVIES IN PROFIT			
	Sum of budget	Sum of gross	Sum of PROFIT
Avatar	237000000	760505847	523505847
E.T. the Extra-Terrestrial	10500000	434949459	424449459
Jurassic World	150000000	652177271	502177271
Star Wars: Episode I - The Phantom Menace	115000000	474544677	359544677
Star Wars: Episode IV - A New Hope	11000000	460935665	449935665
The Avengers	440000000	1246559094	806559094
The Dark Knight	185000000	533316061	348316061
The Jungle Book	350000000	725290282	375290282
The Lion King	45000000	422783777	377783777
Titanic	200000000	658672302	458672302
Grand Total	1743500000	6369734435	4626234435

BASED ON THE ANALYSIS OF GROSS V/S BUDGET, AVTAR WAS THE MOST PROFITABLE MOVIE.

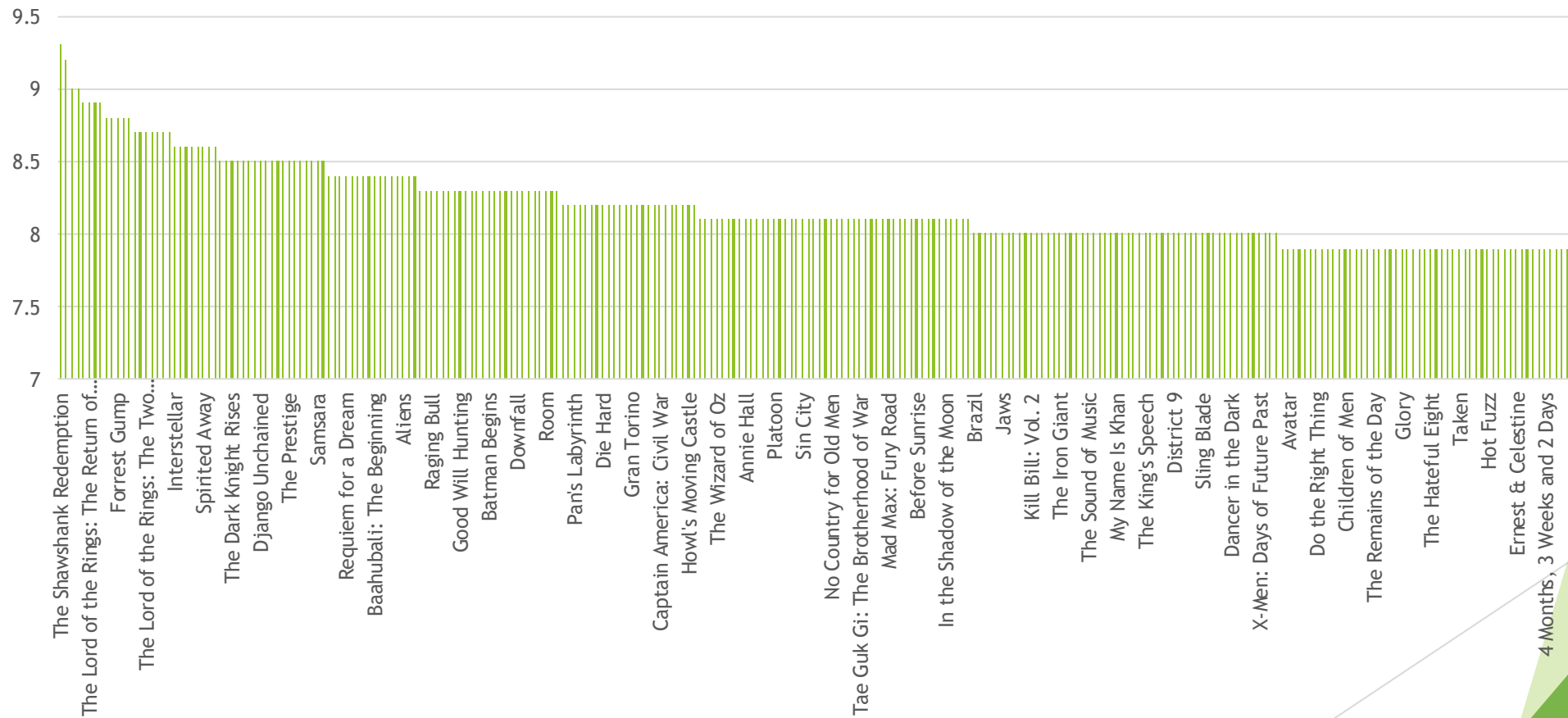
Most Profitable Movies



TOP 250 ENGLISH MOVIES

The Shawshank Redemption is the best English movie with average IMDB score of 9.3

TOP 250 English Movies

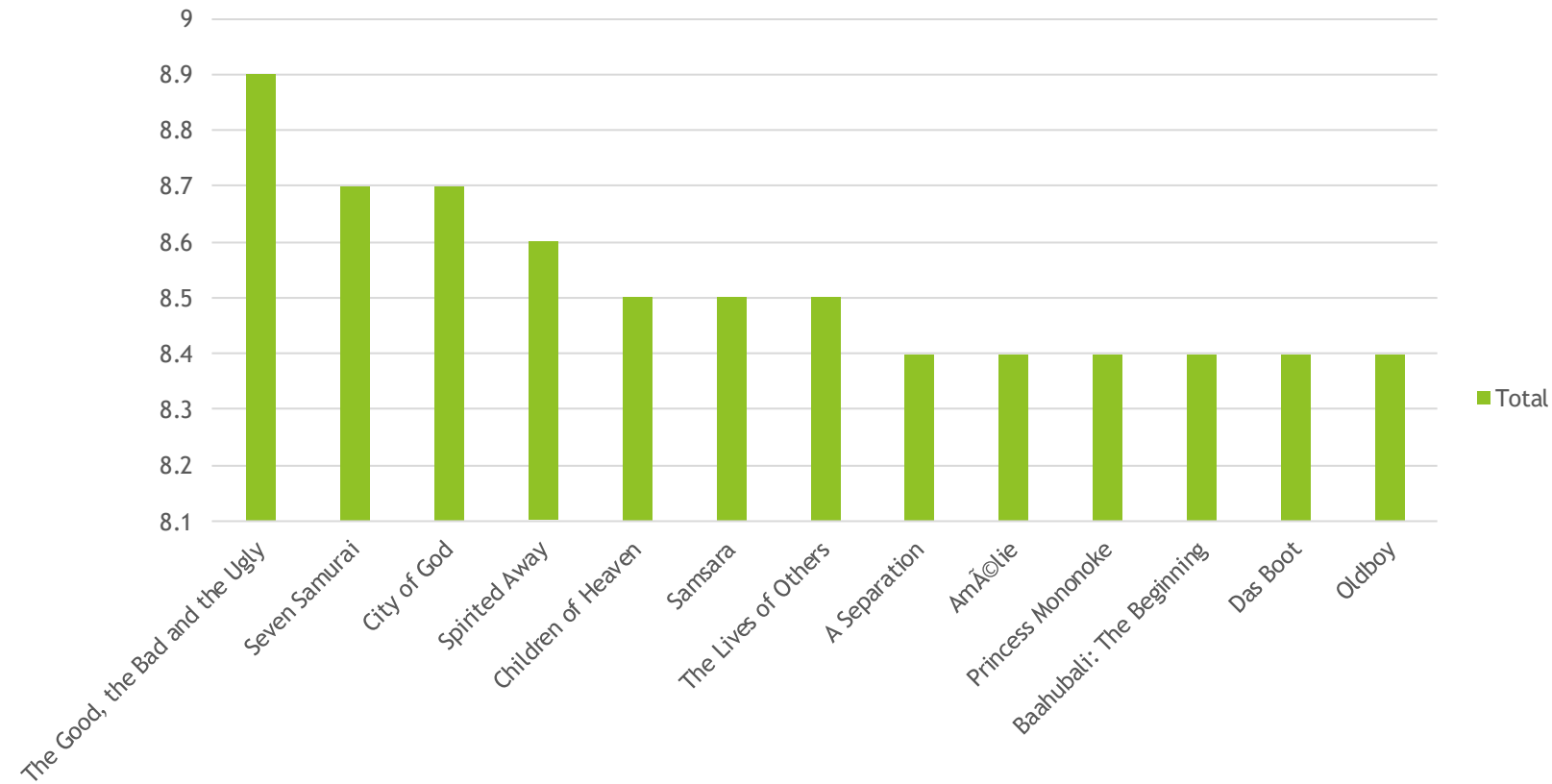


TOP 10 FOREIGN LANGUAGE MOVIES

Row Labels	Average of imdb_score
The Good, the Bad and the Ugly	8.9
Seven Samurai	8.7
City of God	8.7
Spirited Away	8.6
Children of Heaven	8.5
Samsara	8.5
The Lives of Others	8.5
A Separation	8.4
Am��lie	8.4
Princess Mononoke	8.4
Baahubali: The Beginning	8.4
Das Boot	8.4
Oldboy	8.4
Grand Total	8.523076923

**THE GOOD,THE BAD AND THE UGLY
is the best Foreign movies(Italian) with highest
IMDB ratings.**

TOP 10 FOREIGN LANGUAGE MOVIES

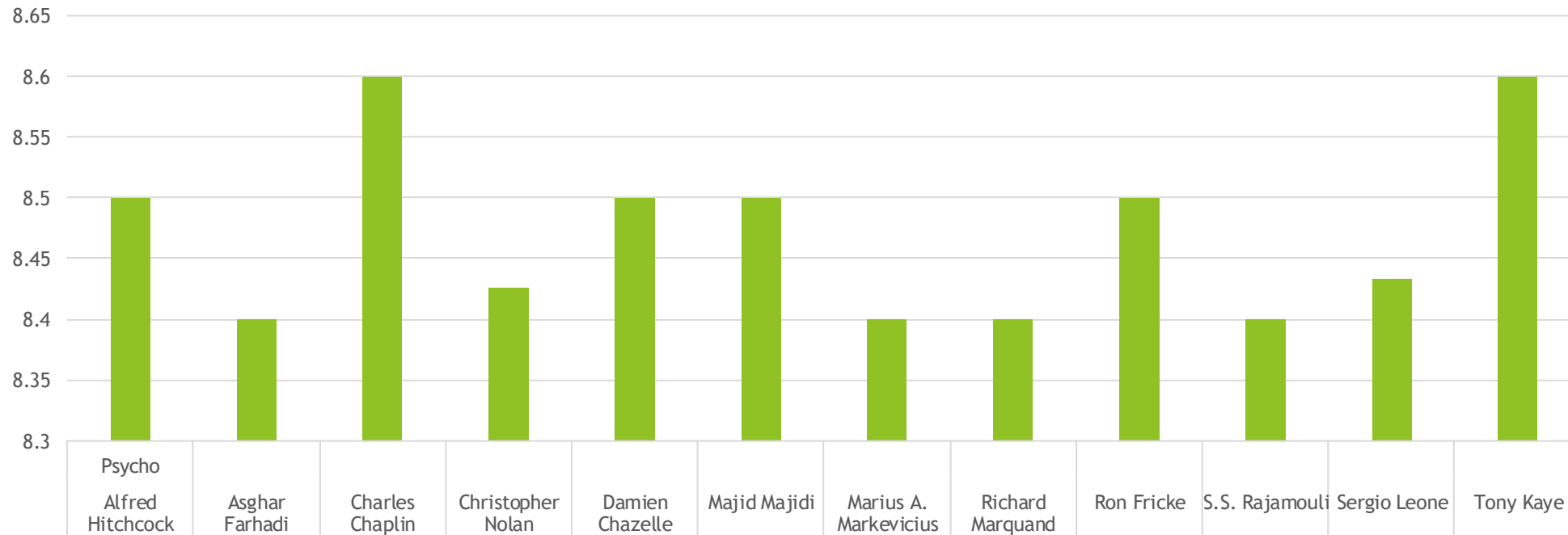


BEST DIRECTORS

Row Labels	Average of imdb_score
Alfred Hitchcock	8.5
Psycho	8.5
Asghar Farhadi	8.4
Charles Chaplin	8.6
Christopher Nolan	8.425
Damien Chazelle	8.5
Majid Majidi	8.5
Marius A. Markevicius	8.4
Richard Marquand	8.4
Ron Fricke	8.5
S.S. Rajamouli	8.4
Sergio Leone	8.433333333
Tony Kaye	8.6
Grand Total	8.452380952

AKIRA KUROSAWA is the Best Director with the Average IMDB rating of 8.7

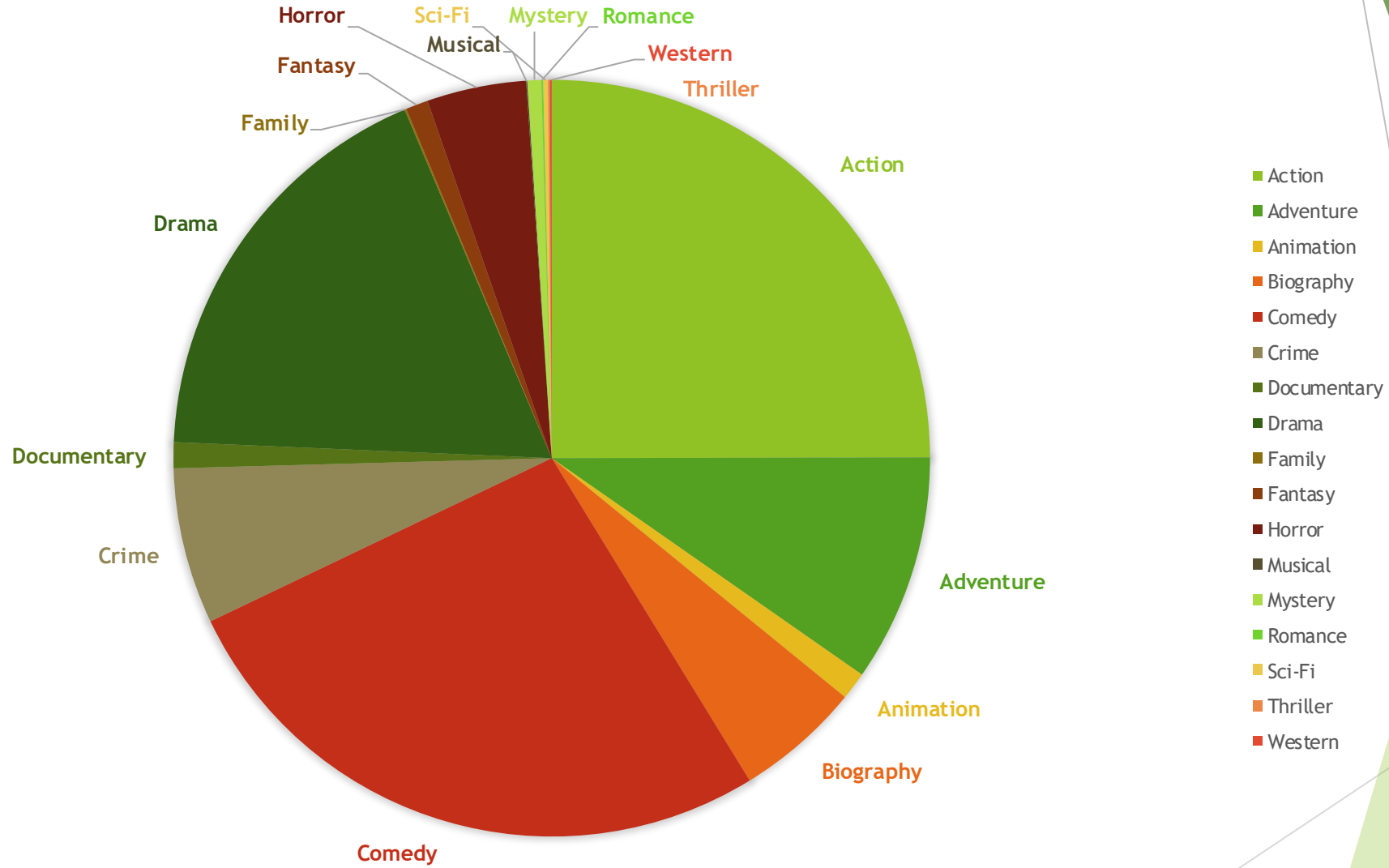
Best Directors



TOP GENRE

GENRE	Count of genres
Action	970
Adventure	378
Animation	46
Biography	208
Comedy	1036
Crime	259
Documentary	43
Drama	698
Family	3
Fantasy	37
Horror	165
Musical	2
Mystery	23
Romance	3
Sci-Fi	8
Thriller	3
Western	3
Grand Total	3885

DRAMA, COMEDY, AND THRILLER WERE TOP THREE GENRE BASED ON TGE MOVIES RELEASED BY BEST DIRECTORS.

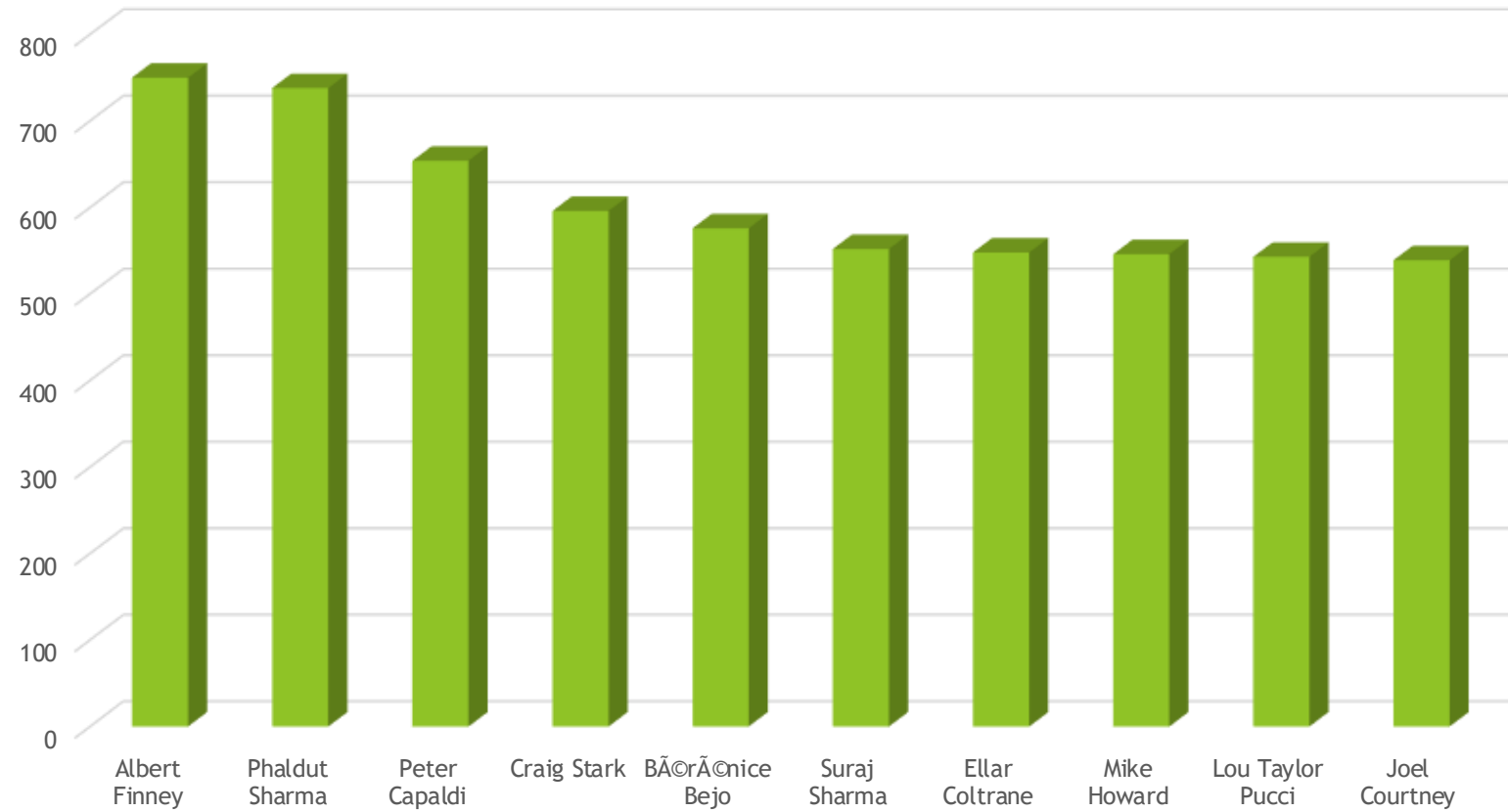


ACTOR ANALYSIS

Row Labels	Average of num_critic_for_reviews
Albert Finney	750
Phaldut Sharma	738
Peter Capaldi	654
Craig Stark	596
BÃ©rÃ©nice Bejo	576
Suraj Sharma	552
Ellar Coltrane	548
Mike Howard	546
Lou Taylor Pucci	543
Joel Courtney	539
Grand Total	617.4545455

- **MERYL STREEP - 11**
- **LEONARD DICAPRIO - 21**
- **BRAD PITT - 17**

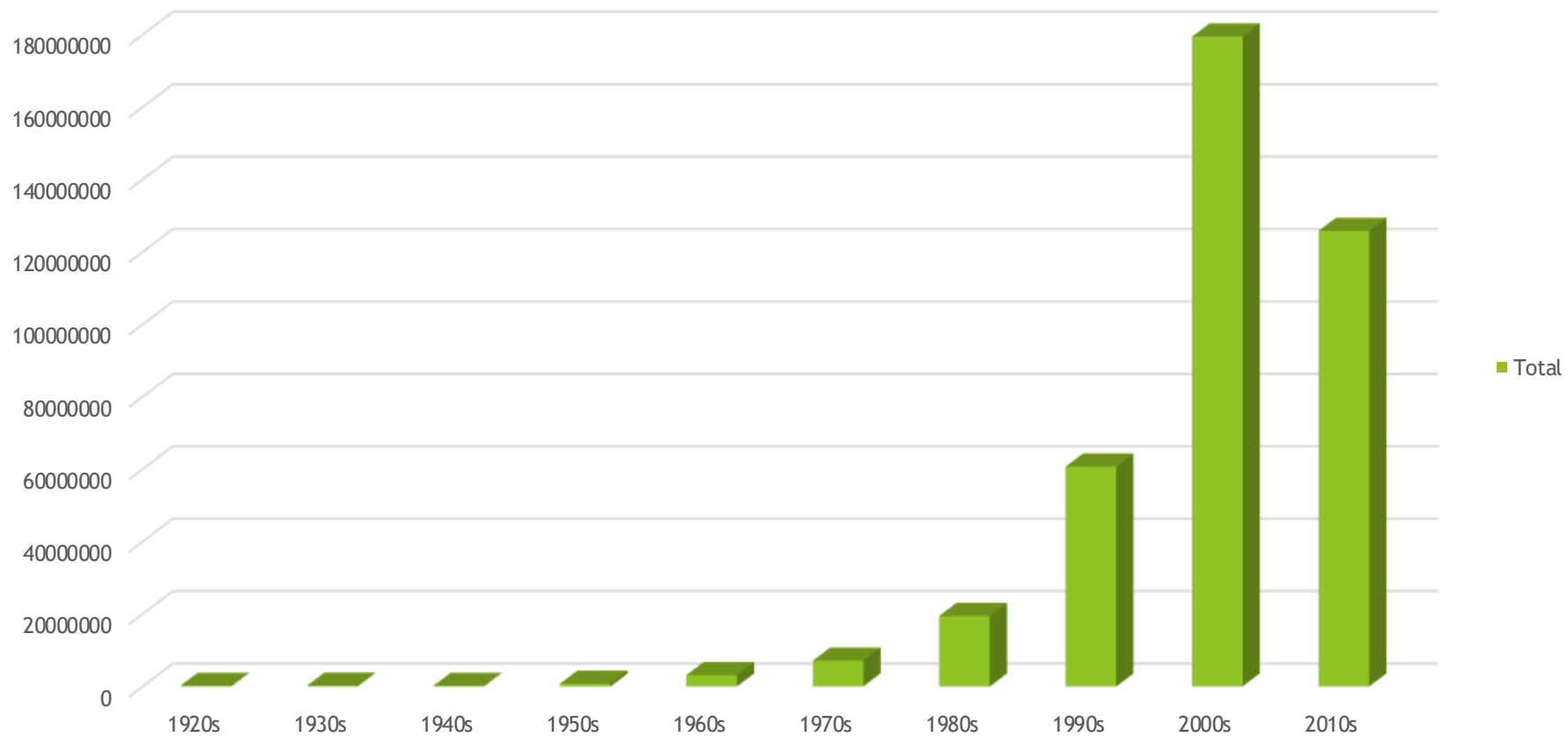
TOP ACTORS



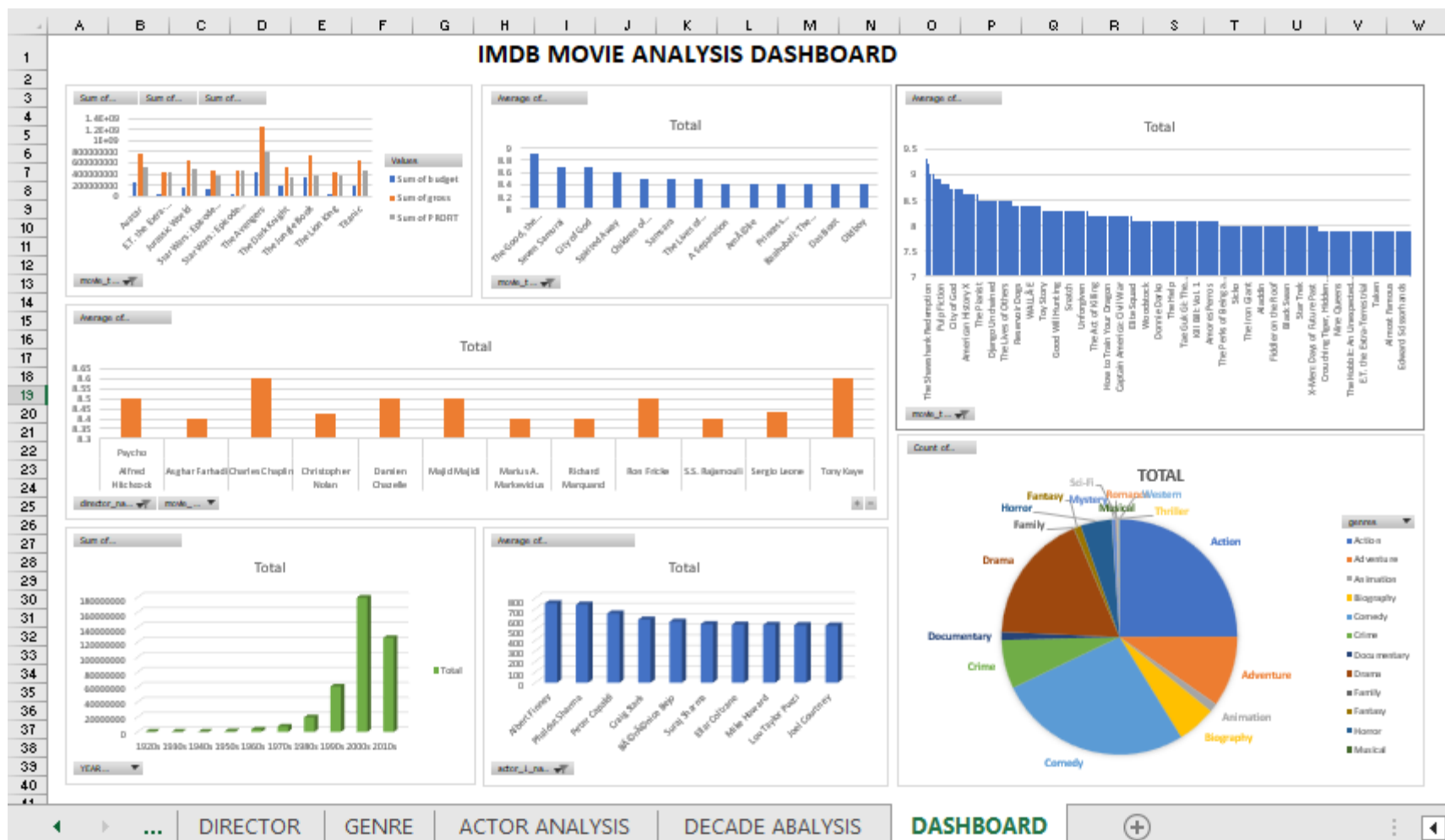
DECADE ANALYSIS

helper-year	Row Labels	Sum of num_voted_users
	1920 1920s	196422
	1930 1930s	207002
	1940 1940s	146351
	1950 1950s	742720
	1960 1960s	3121043
	1970 1970s	7117872
	1980 1980s	19463438
	1990 1990s	60725295
	2000 2000s	179469850
	2010 2010s	125794186
	2020 2020s	Grand Total
		396984179

**HELPER-YEAR COUNM WAS CREATED AND
VLOOK-UP FUNCTION WAS USED TO EXTRACT
THE YEAR RANGE.**



DASHBOARD



RESULT

Profitable Movies

Avatar was the most
Profitable movie.
\$52,35,05,847

Best Director

Akira Kurosawa with
average rating of 8.7

Genre

Based on the best
directors the top 3
genre were DRAMA,
COMEDY and THRILLER

Top 250 English Movies

Shawshank Redemption in English
with average rating of 9.2 and
The Good ,The Bad and The Ugly
in Italian with average voting of
8.9

Actor

Critics favorite is Albert
Finney (750)audience
favorite is health
Donahue (3400 votes)

Thank you

anjalirajput2005@gmail.com