



IMDB MOVIE ANALYSIS

EXCEL

PROJECT DESCRIPTION

This project will go through to find some of the insights on IMDB movies like most popular actor, most popular director etc.

PROJECT APPROACH

In order to find the insight, Excel was used. Using Excel, we performed Exploratory Analysis first like understanding data columns, checking for missing data, checking and removing outliers, etc. After that, we performed various operations like selecting, sorting, applying different formulas, etc. to get the insights we needed.

TECH STACK USED

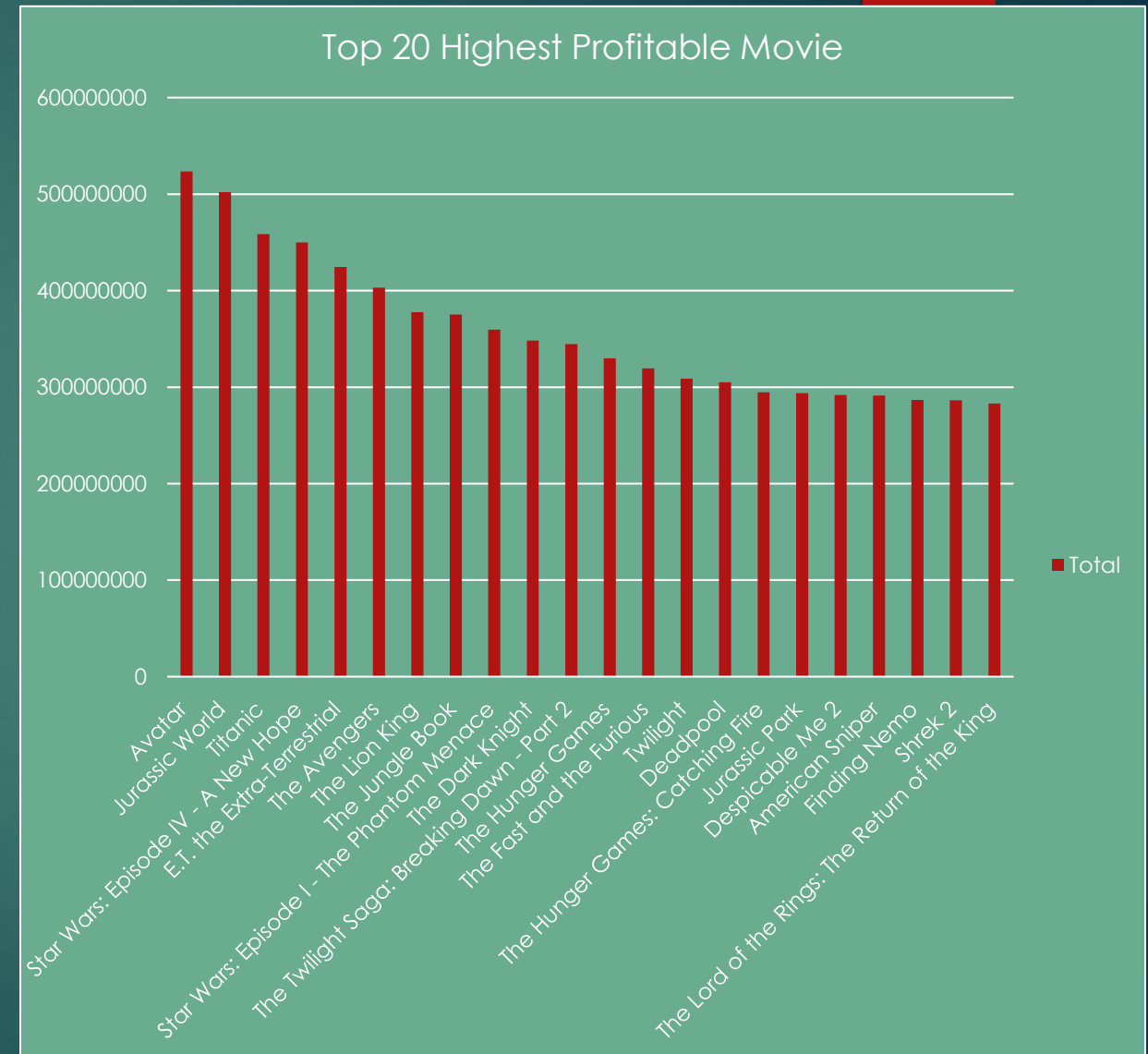
Microsoft Excel , Microsoft Word , and Google Drive.

PROJECT INSIGHT

1. **Cleaning the data:** At first, removed all the rows which were empty. Removed Â from all movie titles. Removed Duplicate values as well. Hide the columns which were not going to be used for finding insights.

2. Movies with the highest profit: Find the movies with the highest profit.

Row Labels	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
The Twilight Saga: Breaking Dawn - Part 2	344597846
The Hunger Games	329999255
The Fast and the Furious	319536930
Twilight	308898950
Deadpool	305024263
The Hunger Games: Catching Fire	294645577
Jurassic Park	293784000
Despicable Me 2	292049635
American Sniper	291323553
Finding Nemo	286838870
Shrek 2	286471036
The Lord of the Rings: The Return of the King	283019252



3. Top 250: Find IMDB top 250.

A new sheet has been created to show IMDB's top 250 movies with votes>25000

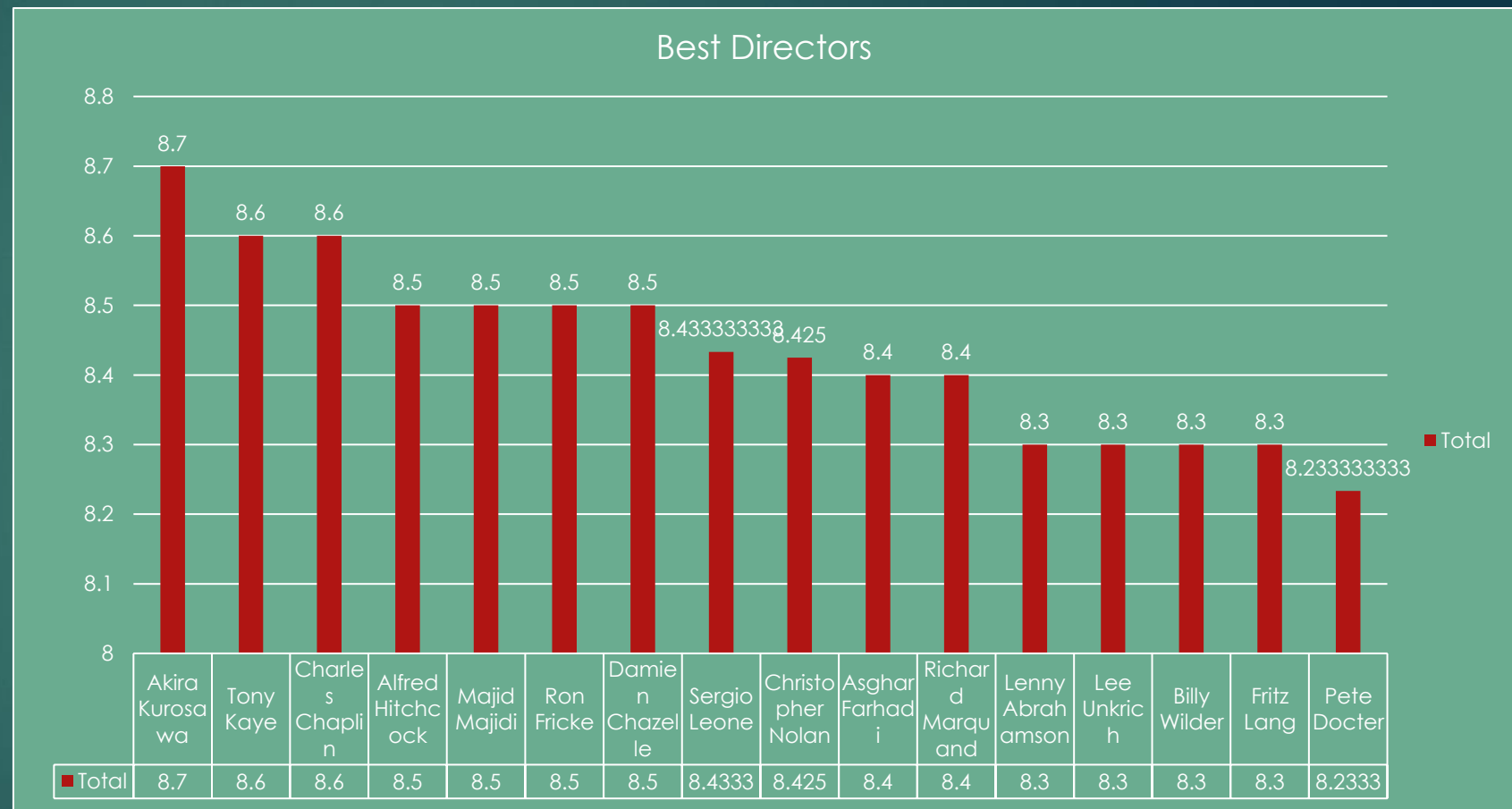
rank	movie_title	num_voted_users>25000	imdb_score	language
1	The Shawshank Redemption	1689764	9.3	English
2	The Godfather	1155770	9.2	English
3	The Dark Knight	1676169	9	English
4	The Godfather: Part II	790926	9	English
5	Pulp Fiction	1324680	8.9	English
6	The Lord of the Rings: The Return of the King	1215718	8.9	English
7	Schindler's List	865020	8.9	English
8	The Good, the Bad and the Ugly	503509	8.9	Italian
9	Inception	1468200	8.8	English
10	Fight Club	1347461	8.8	English
11	Forrest Gump	1251222	8.8	English
12	The Lord of the Rings: The Fellowship of the Ring	1238746	8.8	English
13	Star Wars: Episode V - The Empire Strikes Back	837759	8.8	English
14	The Matrix	1217752	8.7	English
15	The Lord of the Rings: The Two Towers	1100446	8.7	English
16	Star Wars: Episode IV - A New Hope	911097	8.7	English
17	Goodfellas	728685	8.7	English
18	One Flew Over the Cuckoo's Nest	680041	8.7	English
19	City of God	533200	8.7	Portugues
20	Seven Samurai	229012	8.7	Japanes
22	Se7en	1023511	8.6	English
23	Interstellar	928227	8.6	English
24	The Silence of the Lambs	887467	8.6	English
25	Saving Private Ryan	881236	8.6	English
26	American History X	782437	8.6	English
27	The Usual Suspects	740918	8.6	English
28	Spirited Away	417971	8.6	Japanes

IMDB's Top Foreign language film

rank	movie_title	num_voted_users	imdb_score	language
1	The Good, the Bad and the Ugly	503509	8.9	Italian
2	City of God	533200	8.7	Portuguese
3	Seven Samurai	229012	8.7	Japanese
4	Spirited Away	417971	8.6	Japanese
5	The Lives of Others	259379	8.5	German
6	Children of Heaven	27882	8.5	Persian
7	Samsara	22457	8.5	None
8	AmÃ©lie	534262	8.4	French
9	Oldboy	356181	8.4	Korean
10	Princess Mononoke	221552	8.4	Japanese
11	Das Boot	168203	8.4	German
12	A Separation	151812	8.4	Persian
13	Downfall	248354	8.3	German
14	The Hunt	170155	8.3	Danish
15	Metropolis	111841	8.3	German
16	Pan's Labyrinth	467234	8.2	Spanish
17	Howl's Moving Castle	214091	8.2	Japanese
18	The Secret in Their Eyes	131831	8.2	Spanish
19	Incendies	80429	8.2	French
20	The Act of Killing	23836	8.2	Indonesian
21	Amores Perros	173551	8.1	Spanish
22	Akira	106160	8.1	Japanese
23	Elite Squad	81644	8.1	Portuguese
24	The Celebration	65951	8.1	Danish
25	The Sea Inside	64556	8.1	Spanish
26	Tae Guk Gi: The Brotherhood of War	31943	8.1	Korean
27	A Fistful of Dollars	147566	8	Italian
28	Persepolis	70194	8	French

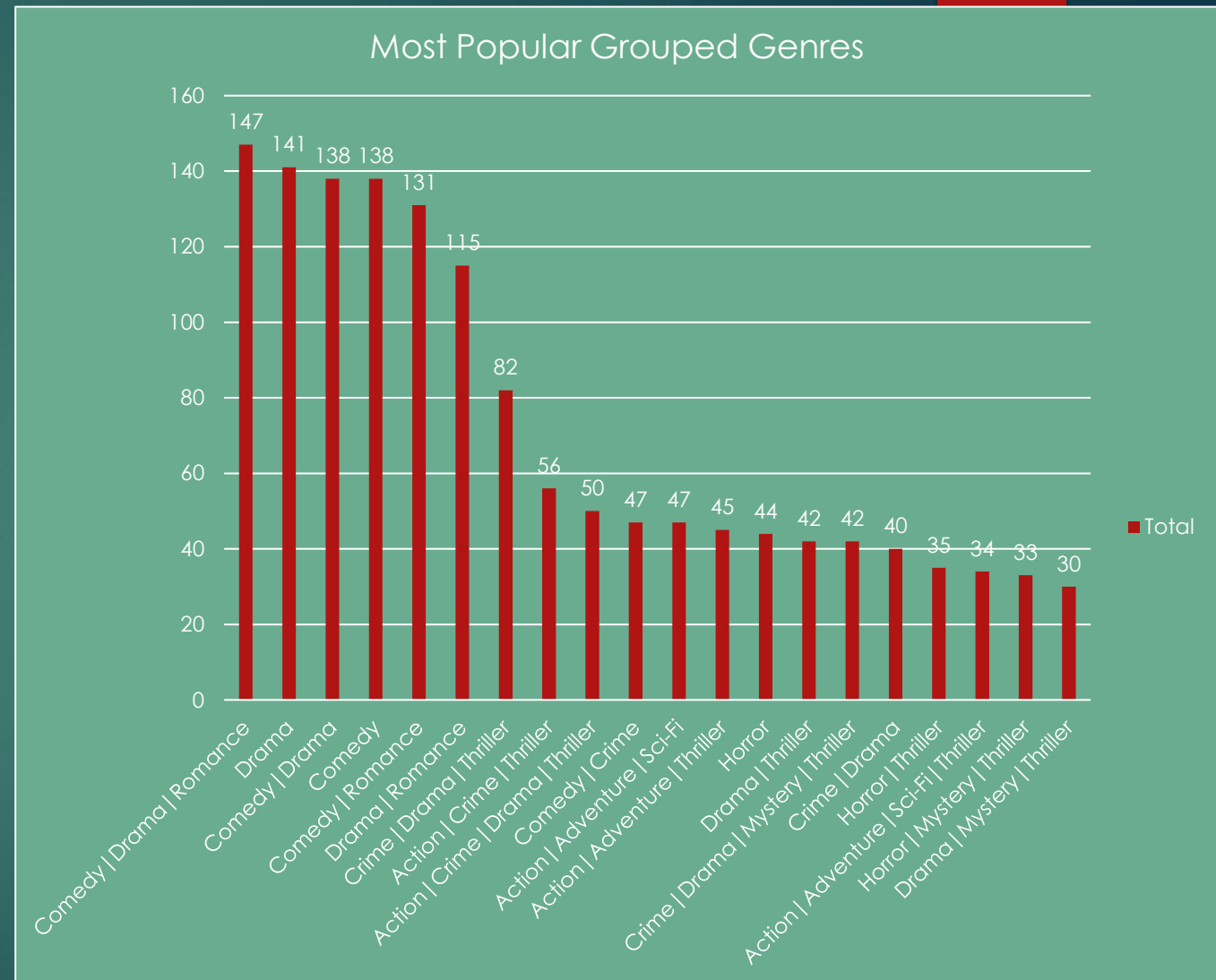
4. Best Directors: The top 10 directors for whom the mean of imdb_score is the highest. The best directors list is shown.

Row Labels	Average of imdb_score
Akira Kurosawa	8.7
Tony Kaye	8.6
Charles Chaplin	8.6
Alfred Hitchcock	8.5
Majid Majidi	8.5
Ron Fricke	8.5
Damien Chazelle	8.5
Sergio Leone	8.433333333
Christopher Nolan	8.425
Asghar Farhadi	8.4
Richard Marquand	8.4
Lenny Abrahamson	8.3
Lee Unkrich	8.3
Billy Wilder	8.3
Fritz Lang	8.3
Pete Docter	8.233333333



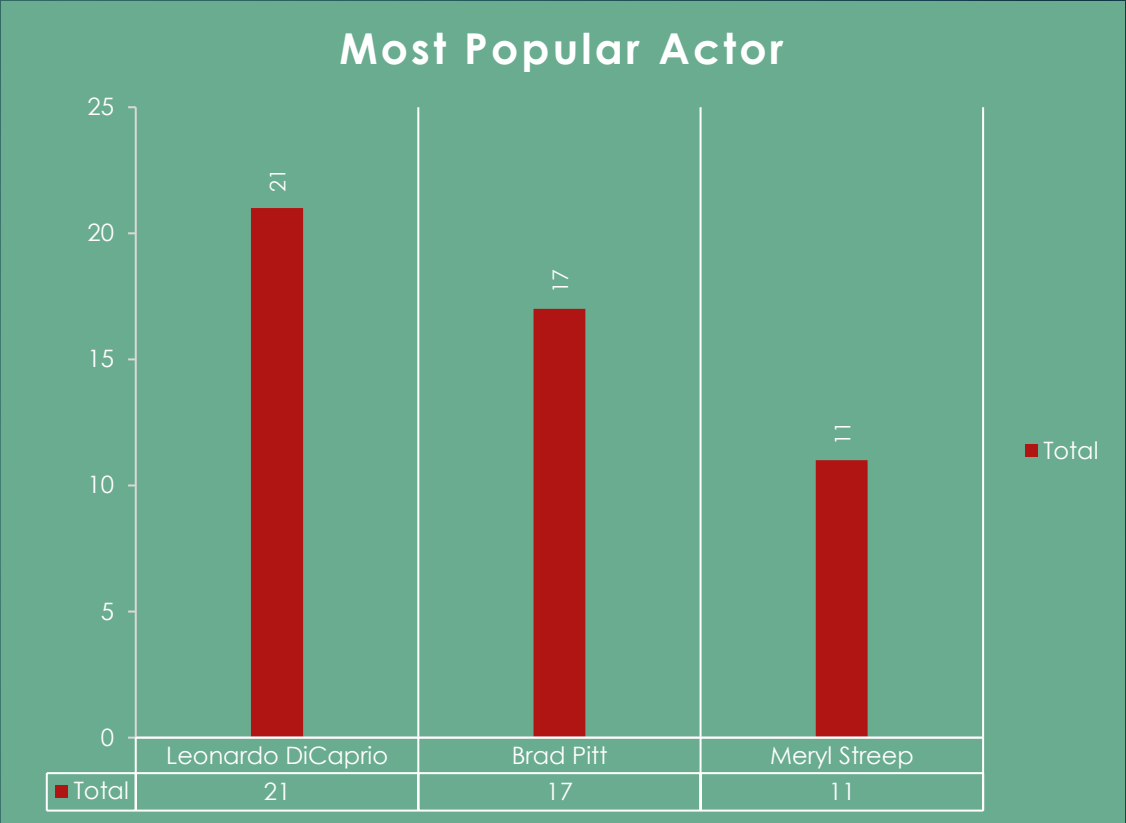
5. Popular Genres: Find the most popular genres.

Row Labels	Count of genres
Comedy Drama Romance	147
Drama	141
Comedy Drama	138
Comedy	138
Comedy Romance	131
Drama Romance	115
Crime Drama Thriller	82
Action Crime Thriller	56
Action Crime Drama Thriller	50
Comedy Crime	47
Action Adventure Sci-Fi	47
Action Adventure Thriller	45
Horror	44
Drama Thriller	42
Crime Drama Mystery Thriller	42
Crime Drama	40
Horror Thriller	35
Action Adventure Sci-Fi Thriller	34
Horror Mystery Thriller	33
Drama Mystery Thriller	30

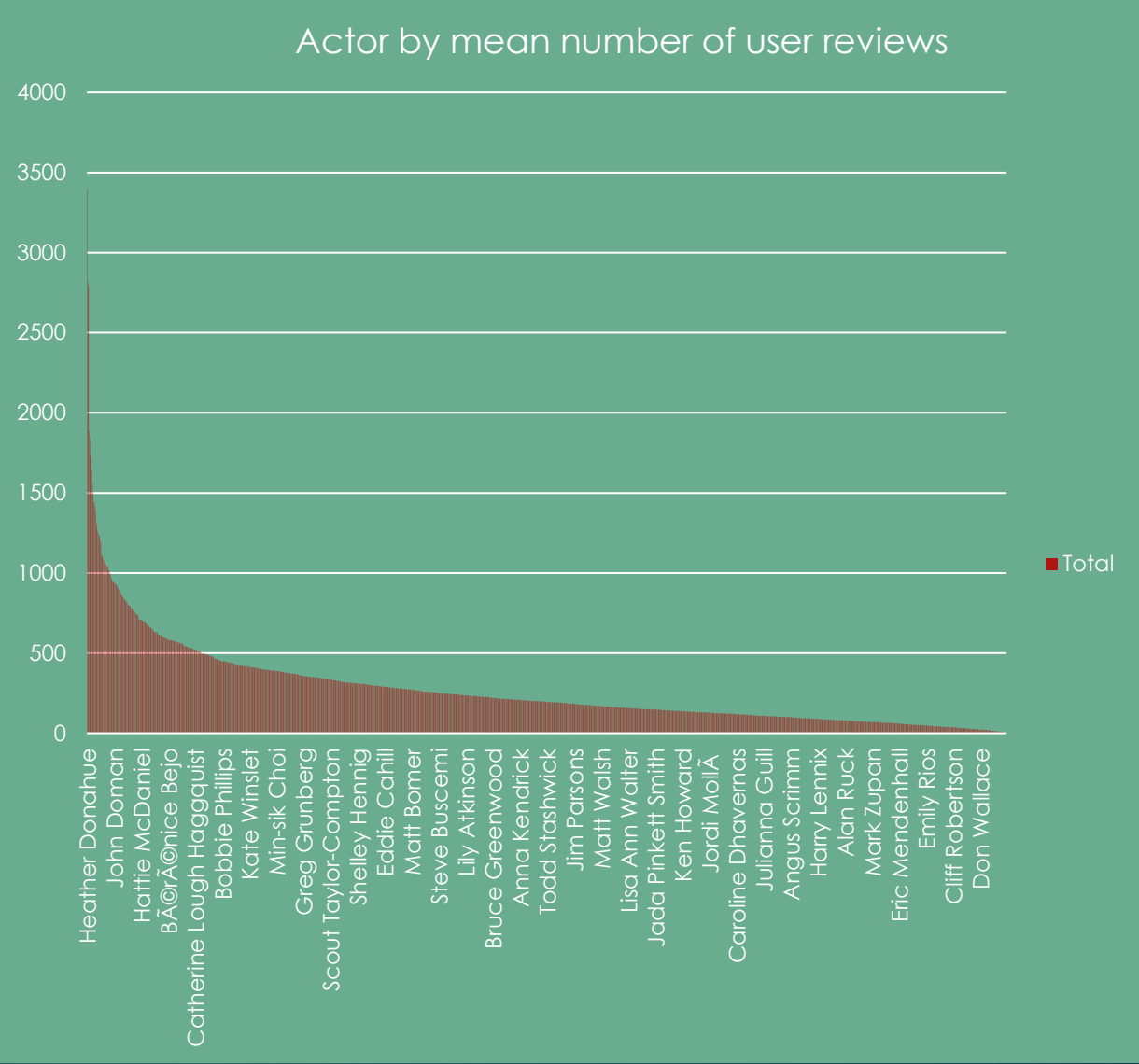


6. Charts: Used charts and graphs to find different-different insights.

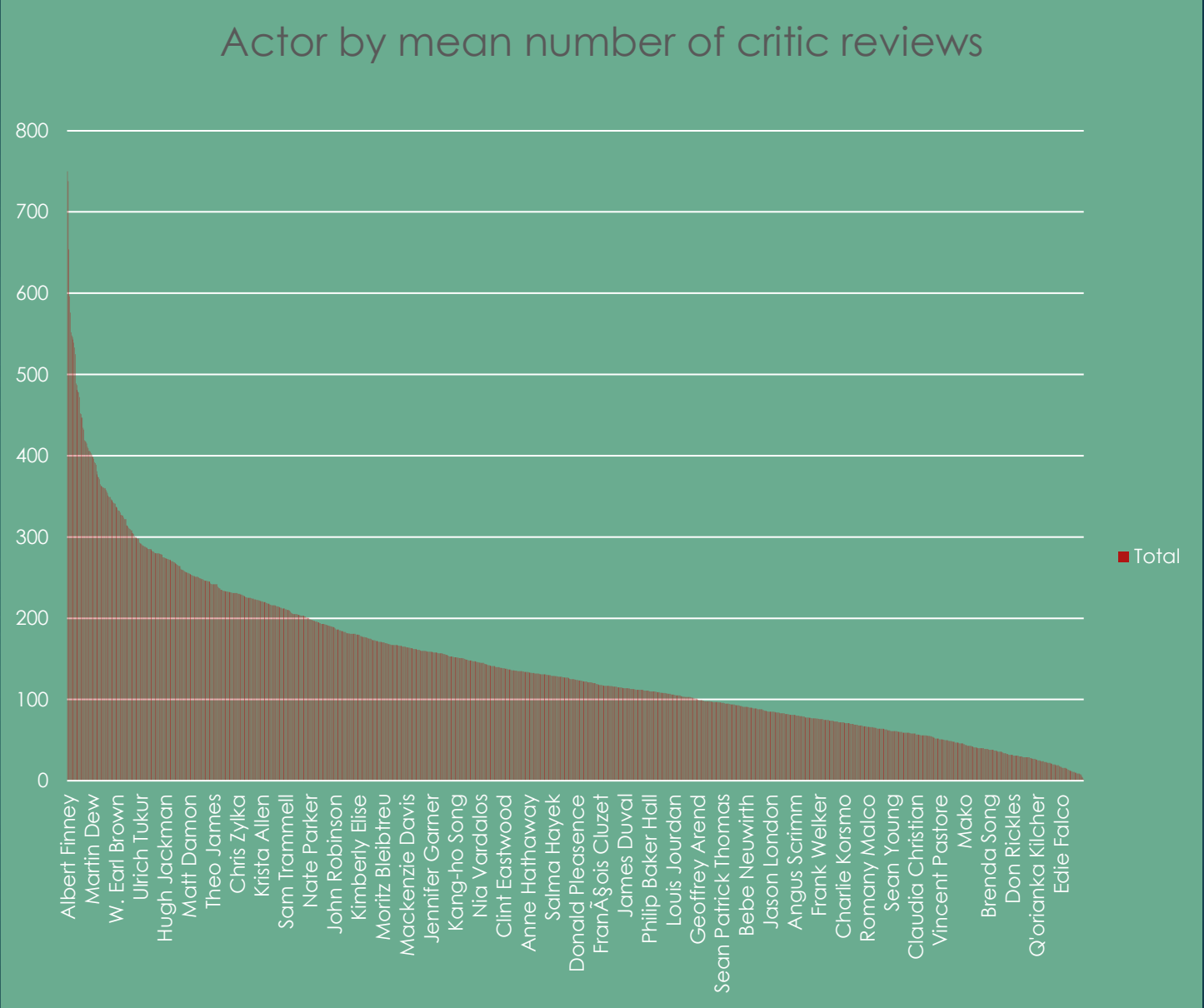
Row Labels	Count of actor_1_name
Leonardo DiCaprio	21
Brad Pitt	17
Meryl Streep	11
Grand Total	49



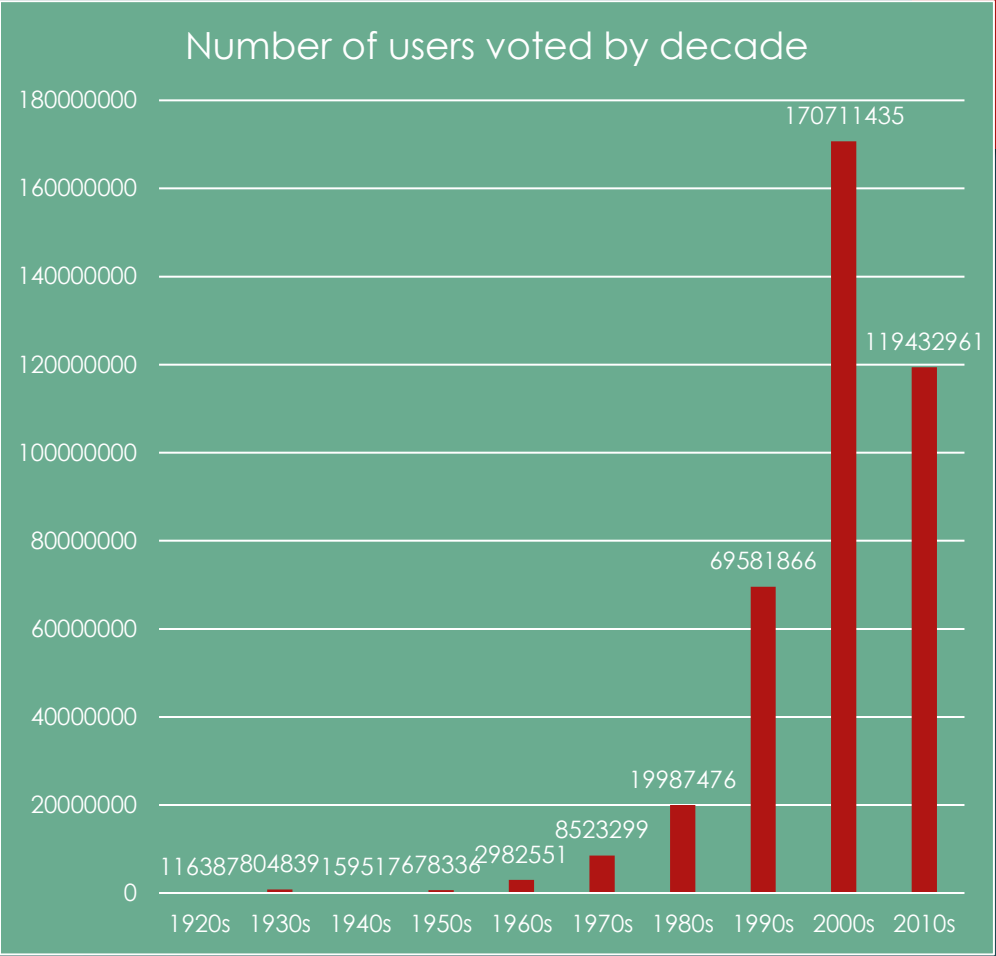
Row Labels	Average of num_user_for_reviews
Heather Donahue	3400
Christo Jivkov	2814
Steve Bastoni	2789
Phaldut Sharma	1885
Orlando Bloom	1842
Keir Dullea	1736
Eva Green	1708.333333
Chen Chang	1641
Nick Stahl	1562
Albert Finney	1498
Kevin Rankin	1445
Noah Huntley	1441
Osama bin Laden	1416
Seychelle Gabriel	1382
Mathieu Kassovitz	1314



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
Maika Monroe	533
Tim Holmes	525
Elina Alminas	489
Kurt Fuller	487
Iko Uwais	481
Quvenzhana Wallis	478.6666667
Edgar Arreola	478
Sharlto Copley	472
Cory Hardrict	452
Matt Frewer	451
Isabelle Huppert	447
Elizabeth McGovern	447



Row Labels		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
Grand Total		392978667



Result

1. Clean data.
2. Created Profit column by subtracting gross with budget using formula-
=(I2-W2)
Created Pivot Table and found out the top 20 movies. Created Chart for the bet.
3. A new sheet has been created to show IMDB's top 250 movies with votes>25000 and IMDB's Top Foreign language film.
4. Done using Pivot table. The grouping feature was used to find the mean of imdb score.
5. Done using Pivot table. The grouping feature was used to find the count of grouped genres.
6. Pivot Tables and Graphs were used to find these insights.
For finding the decade, I used the formula:
=LEFT(Y2,3)&"0s"



Thank YOU