Project-5
IMDb Movie
Analysis

Description:

This project is about the IMDB movie analysis based on different criteria. We have IMBD's dataset for different movies, which contains many information and we have to analyze the dataset using the questions that where asked.

Analysis done for the following questions:

A. Cleaning the data
B. Movies with highest profit
C. Top 250
D. Best Directors
E. Popular Genres
F. Charts

Approach:

We have to clean the given datasets which contains tables that contain information regarding various fields related to movies to be rated by IMDB. We have to go through the dataset closely and then executed various queries to extract the required insights from the dataset that was cleaned.

Tech/Tools used:

PPT – To prepare a detailed report.
EXCEL – Excel was used to perform entire analysis.

Insights:

A. Cleaning the data:

1. Dropping unnecessary columns.

(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)

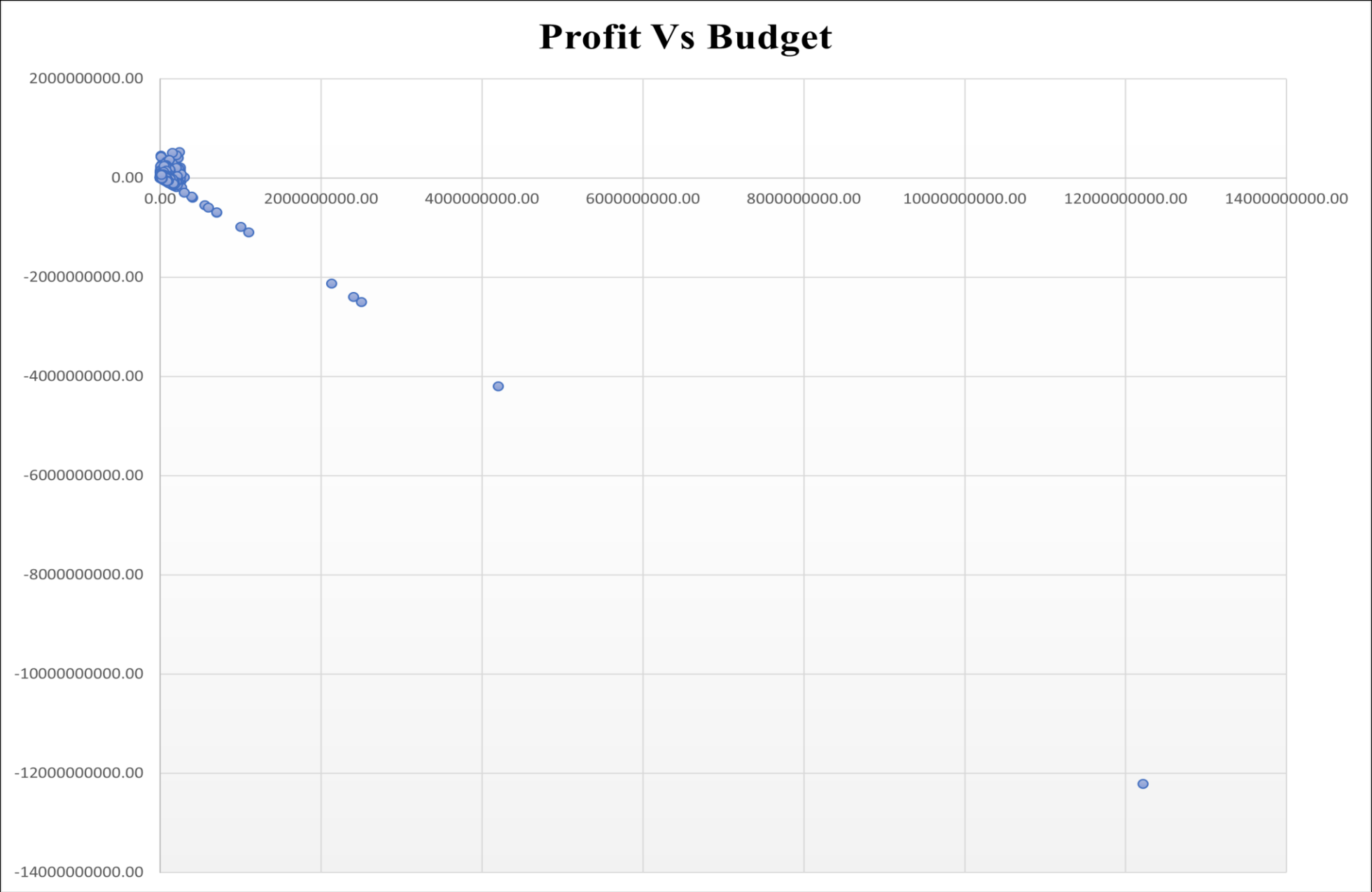2. Remove Blank Cell / Null Value.

3. Removing Duplicate.

Project-5 IMDB
Movie Cleaned Dat

B. Movies with highest profit:

Profit Vs Budget

OUTLIERS:

-12213298588

-4199788333

-2499804112

-2397701809

-2127109510

Top 5 Highest Profit:

director_name	actor_1_name	movie_title	title_year	imdb_score	Profit
James Cameron	CCH Pounder	Avatar	2009	7.9	523505847
Colin Trevorrow	Bryce Dallas Howa	Jurassic World	2015	7	502177271
James Cameron	Leonardo DiCapric	Titanic	1997	7.7	458672302
George Lucas	Harrison Ford	Star Wars: Epis	1977	8.7	449935665
Steven Spielberg	Henry Thomas	E.T. the Extra-T	1982	7.9	424449459

C. Top 250:

1. Filtering the data where num_voted_users > 25,000.
2. Sort the data using imbd_score column in largest to smallest score.
3. Create column 'Rank' for top 250 movies.
4. Use Sequence Formula to Rank.
5. Formula =SEQUENCE(COUNTA(G2:G251),1,1,1)
6. Filter out language by unselecting English to extract the top foreign language movie.

Top 250 Movies

	director_name	num_critic	duration	gross	genres	actor_1_name	movie_title	num_voted_users	num_use	language	country	budget	title_year	imdb_sc	Profit	rank
1																
2	Frank Darabont	199	142	28341469	Crime|Dra	Morgan Freema	The Shawshank RedemptionÂ	1689764	4144	English	USA	25000000	1994	9.3	3341469	1
3	Francis Ford Coppe	208	175	134821952	Crime|Dra	Al Pacino	The GodfatherÂ	1155770	2238	English	USA	6000000	1972	9.2	128821952	2
4	Christopher Nolan	645	152	533316061	Action|Cri	Christian Bale	The Dark KnightÂ	1676169	4667	English	USA	185000000	2008	9	348316061	3
5	Francis Ford Coppe	149	220	57300000	Crime|Dra	Robert De Niro	The Godfather: Part IIÂ	790926	650	English	USA	13000000	1974	9	44300000	4
6	Quentin Tarantino	215	178	107930000	Crime|Dra	Bruce Willis	Pulp FictionÂ	1324680	2195	English	USA	8000000	1994	8.9	99930000	5
7	Peter Jackson	328	192	377019252	Action|Ad	Orlando Bloom	The Lord of the Rings: The Retu	1215718	3189	English	USA	94000000	2003	8.9	283019252	6
8	Steven Spielberg	174	185	96067179	Biography	Liam Neeson	Schindler's ListÂ	865020	1273	English	USA	22000000	1993	8.9	74067179	7
9	Sergio Leone	181	142	6100000	Western	Clint Eastwood	The Good, the Bad and the Ugl	503509	780	Italian	Italy	1200000	1966	8.9	4900000	8
10	Christopher Nolan	642	148	292568851	Action|Ad	Leonardo DiCa	InceptionÂ	1468200	2803	English	USA	160000000	2010	8.8	132568851	9
11	David Fincher	315	151	37023395	Drama	Brad Pitt	Fight ClubÂ	1347461	2968	English	USA	63000000	1999	8.8	-25976605	10
12	Robert Zemeckis	149	142	329691196	Comedy|D	Tom Hanks	Forrest GumpÂ	1251222	1398	English	USA	55000000	1994	8.8	274691196	11
13	Peter Jackson	297	171	313837577	Action|Ad	Christopher Lee	The Lord of the Rings: The Fello	1238746	5060	English	New Zealar	93000000	2001	8.8	220837577	12
14	Irvin Kershner	223	127	290158751	Action|Ad	Harrison Ford	Star Wars: Episode V - The Emp	837759	900	English	USA	18000000	1980	8.8	272158751	13
15	Lana Wachowski	313	136	171383253	Action|Sci	Keanu Reeves	The MatrixÂ	1217752	3646	English	USA	63000000	1999	8.7	108383253	14
16	Peter Jackson	294	172	340478898	Action|Ad	Christopher Lee	The Lord of the Rings: The Two	1100446	2417	English	USA	94000000	2002	8.7	246478898	15
17	George Lucas	282	125	460935665	Action|Ad	Harrison Ford	Star Wars: Episode IV - A New	911097	1470	English	USA	11000000	1977	8.7	449935665	16
18	Martin Scorsese	192	146	46836394	Biography	Robert De Niro	GoodfellasÂ	728685	989	English	USA	25000000	1990	8.7	21836394	17
19	Milos Forman	149	133	112000000	Drama	Scatman Croth	One Flew Over the Cuckoo's Ne	680041	760	English	USA	4400000	1975	8.7	107600000	18
20	Fernando Meirelle	214	135	7563397	Crime|Dra	Alice Braga	City of GodÂ	533200	749	Portuguese	Brazil	3300000	2002	8.7	4263397	19
21	Akira Kurosawa	153	202	269061	Action|Ad	Takashi Shimur	Seven SamuraiÂ	229012	596	Japanese	Japan	2000000	1954	8.7	-1730939	20
22	David Fincher	216	127	100125340	Crime|Dra	Morgan Freema	Se7enÂ	1023511	1080	English	USA	33000000	1995	8.6	67125340	21
23	Christopher Nolan	712	169	187991439	Adventure	Matthew McCc	InterstellarÂ	928227	2725	English	USA	165000000	2014	8.6	22991439	22
24	Jonathan Demme	185	138	130727000	Crime|Dra	Anthony Hopkir	The Silence of the LambsÂ	887467	916	English	USA	19000000	1991	8.6	111727000	23
25	Steven Spielberg	219	169	216119491	Action|Dra	Tom Hanks	Saving Private RyanÂ	881236	2277	English	USA	70000000	1998	8.6	146119491	24
26	Tony Kaye	162	101	6712241	Crime|Dra	Ethan Suplee	American History XÂ	782437	1420	English	USA	7500000	1998	8.6	-787759	25
27	Bryan Singer	162	106	23272306	Crime|Dra	Kevin Spacey	The Usual SuspectsÂ	740918	1182	English	USA	6000000	1995	8.6	17272306	26
28	Hayao Miyazaki	246	125	10049886	Adventure	Rumi Hiiragi	Spirited AwayÂ	417971	993	Japanese	Japan	19000000	2001	8.6	-8950114	27

Top Foreign Language Movies

	director_name	num_critic_fo	duration	gross	genres	actor_1_name	movie_title	num_voted	num_use	language	country	budget	title_year	imdb_score	Profit	rank
2	Sergio Leone	181	142	6100000	Western	Clint Eastwood	The Good, the Bad and the Ugly	503509	780	Italian	Italy	1200000	1966	8.9	4900000	8
3	Fernando Meirelle	214	135	7563397	Crime|Drama	Alice Braga	City of God	533200	749	Portuguese	Brazil	3300000	2002	8.7	4263397	19
4	Akira Kurosawa	153	202	269061	Action|Adventure|	Takashi Shimura	Seven Samurai	229012	596	Japanese	Japan	2000000	1954	8.7	-1730939	20
5	Hayao Miyazaki	246	125	10049886	Adventure|Animati	Bunta Sugawara	Spirited Away	417971	902	Japanese	Japan	19000000	2001	8.6	-8950114	27
6	Florian Henckel vo	215	137	11284657	Drama|Thriller	Sebastian Koch	The Lives of Others	259379	407	German	Germany	2000000	2006	8.5	9284657	46
7	Majid Majidi	46	89	925402	Drama|Family	Bahare Seddiqi	Children of Heaven	27882	130	Persian	Iran	180000	1997	8.5	745402	47
8	Jean-Pierre Jeunet	242	122	33201661	Comedy|Romance	Mathieu Kassovitz	Amélie	534262	1314	French	France	77000000	2001	8.4	-43798339	54
9	Chan-wook Park	305	120	2181290	Drama|Mystery|Th	Min-sik Choi	Oldboy	356181	809	Korean	South Kore	3000000	2003	8.4	-818710	56
10	Hayao Miyazaki	174	134	2298191	Adventure|Animati	Minnie Driver	Princess Mononoke	221552	570	Japanese	Japan	2.4E+09	1997	8.4	-2397701809	57
11	Wolfgang Petersen	96	293	11433134	Adventure|Drama	Jürgen Prochnow	Das Boot	168203	426	German	West Germ	14000000	1981	8.4	-2566866	60
12	Asghar Farhadi	354	123	7098492	Drama|Mystery	Shahab Hosseini	A Separation	151812	264	Persian	Iran	500000	2011	8.4	6598492	61
13	S.S. Rajamouli	44	159	6498000	Action|Adventure|	Tamannaah Bhatia	Baahubali: The Beginning	62756	410	Telugu	India	18026148	2015	8.4	-11528148	62
14	Oliver Hirschbiegel	192	178	5501940	Biography|Drama	Thomas Kretschmann	Downfall	248354	564	German	Germany	13500000	2004	8.3	-7998060	80
15	Thomas Vinterberg	349	115	610968	Drama	Thomas Bo Larsen	The Hunt	170155	249	Danish	Denmark	3800000	2012	8.3	-3189032	84
16	Fritz Lang	260	145	26435	Drama|Sci-Fi	Brigitte Helm	Metropolis	111841	413	German	Germany	6000000	1927	8.3	-5973565	86
17	Guillermo del Toro	406	112	37623143	Drama|Fantasy|W	Ivana Baquero	Pan's Labyrinth	467234	1083	Spanish	Spain	13500000	2006	8.2	24123143	96
18	Hayao Miyazaki	212	119	4710455	Adventure|Animati	Christian Bale	Howl's Moving Castle	214091	330	Japanese	Japan	24000000	2004	8.2	-19289545	105
19	Juan José Camp	262	129	20167424	Drama|Mystery|Th	Ricardo Darín	The Secret in Their Eyes	131831	231	Spanish	Argentina	2000000	2009	8.2	18167424	107
20	Denis Villeneuve	226	139	6857096	Drama|Mystery|W	Lubna Azabal	Incendies	80429	156	French	Canada	6800000	2010	8.2	57096	109
21	Alejandro G. Iñá	157	115	5383834	Drama|Thriller	Adriana Barraza	Amores Perros	173551	361	Spanish	Mexico	2000000	2000	8.1	3383834	148
22	Katsuhiro Ōtomo	150	124	439162	Action|Animation|	Mitsuo Iwata	Akira	106160	430	Japanese	Japan	1.1E+09	1988	8.1	-1099560838	150
23	José Padilha	142	115	8060	Action|Crime|Dran	Wagner Moura	Elite Squad	81644	107	Portuguese	Brazil	4000000	2007	8.1	-3991940	151
24	Thomas Vinterberg	98	105	1647780	Drama	Ulrich Thomsen	The Celebration	65951	258	Danish	Denmark	1300000	1998	8.1	347780	152
25	Alejandro Amená	157	125	2086345	Biography|Drama	Belén Rueda	The Sea Inside	64556	140	Spanish	Spain	10000000	2004	8.1	-7913655	153
26	Je-kyu Kang	86	148	1110186	Action|Drama|Wa	Min-sik Choi	Tae Guk Gi: The Brotherhood of W	31943	224	Korean	South Kore	12800000	2004	8.1	-11689814	155
27	Sergio Leone	122	99	3500000	Action|Drama|We	Clint Eastwood	A Fistful of Dollars	147566	235	Italian	Italy	200000	1964	8	3300000	194

D. Best Directors:

Top 10 Directors	Average of imdb_score
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	8.4
Christopher Nolan	8.4
Asghar Farhadi	8.4
Marius A. Markevicius	8.4

E. Popular Genres

Popular Genres	Average of imdb_score
Crime|Drama|Fantasy|Mystery	8.5
Adventure|Animation|Drama|Family|Musical	8.5
Adventure|Drama|Thriller|War	8.4
Adventure|Animation|Fantasy	8.4
Action|Adventure|Drama|Fantasy|War	8.4
Documentary|War	8.3
Documentary|Drama|Sport	8.3
Biography|Drama|History|Music	8.3
Adventure|Animation|Comedy|Drama|Family|Fantasy	8.3
Adventure|Drama|War	8.3

F. Charts:

1.Highest Mean

actor_1_name 🔽	Mean of num_user_for_reviews	Mean of num_critic_for_reviews
Brad Pitt	742.35	245
Leonardo DiCaprio	914.48	330.19
Meryl Streep	297.18	181.45
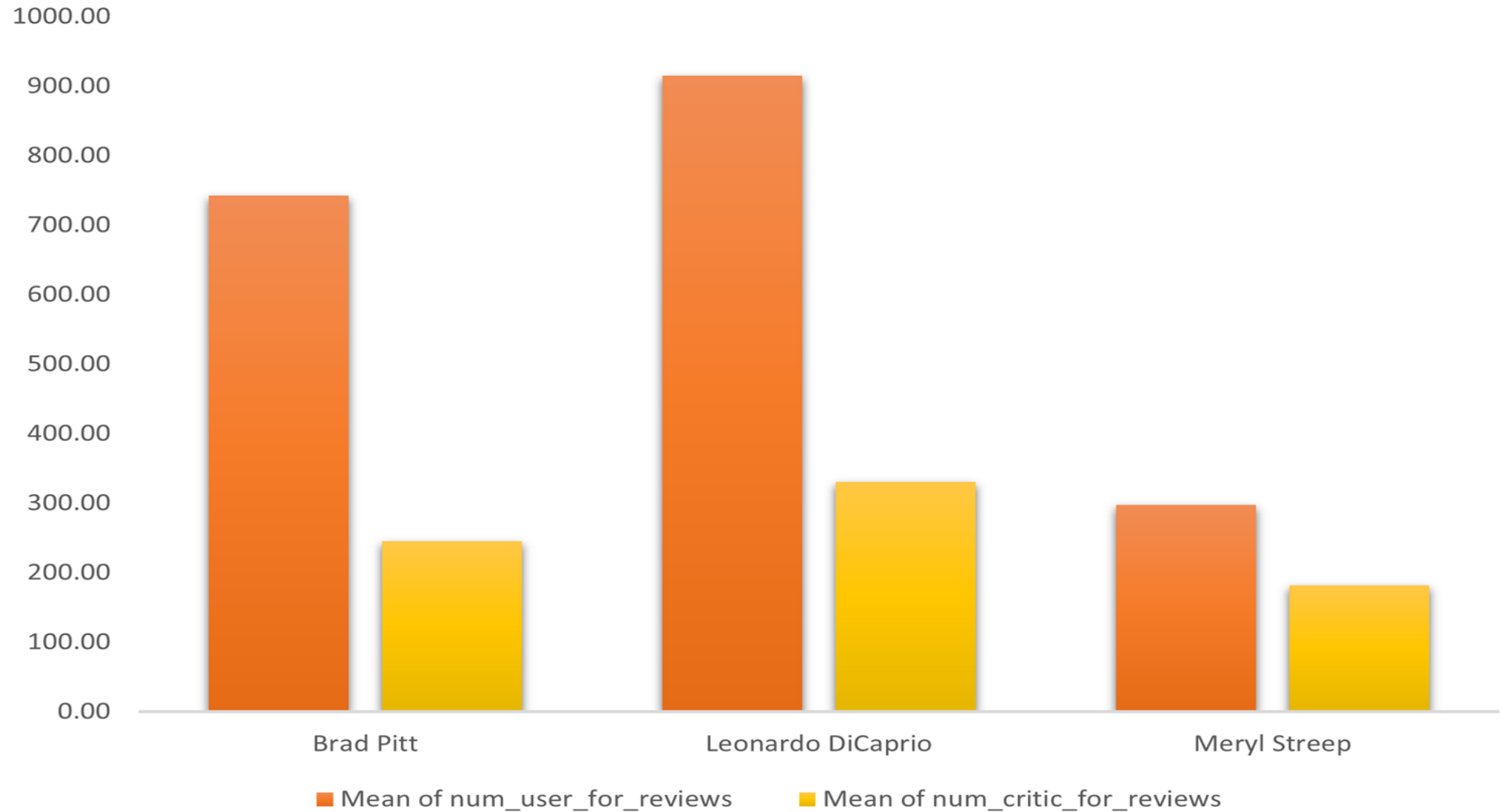
2. Voted users over decades

Decade	Number of Users Voted
1921-1930	116387
1931-1940	966520
1941-1950	72324
1951-1960	1097601
1961-1970	2609930
1971-1980	10377495
1981-1990	22455094
1991-2000	79427470
2001-2010	180142087
2011-2020	101974167

Number of Users Voted over Decades
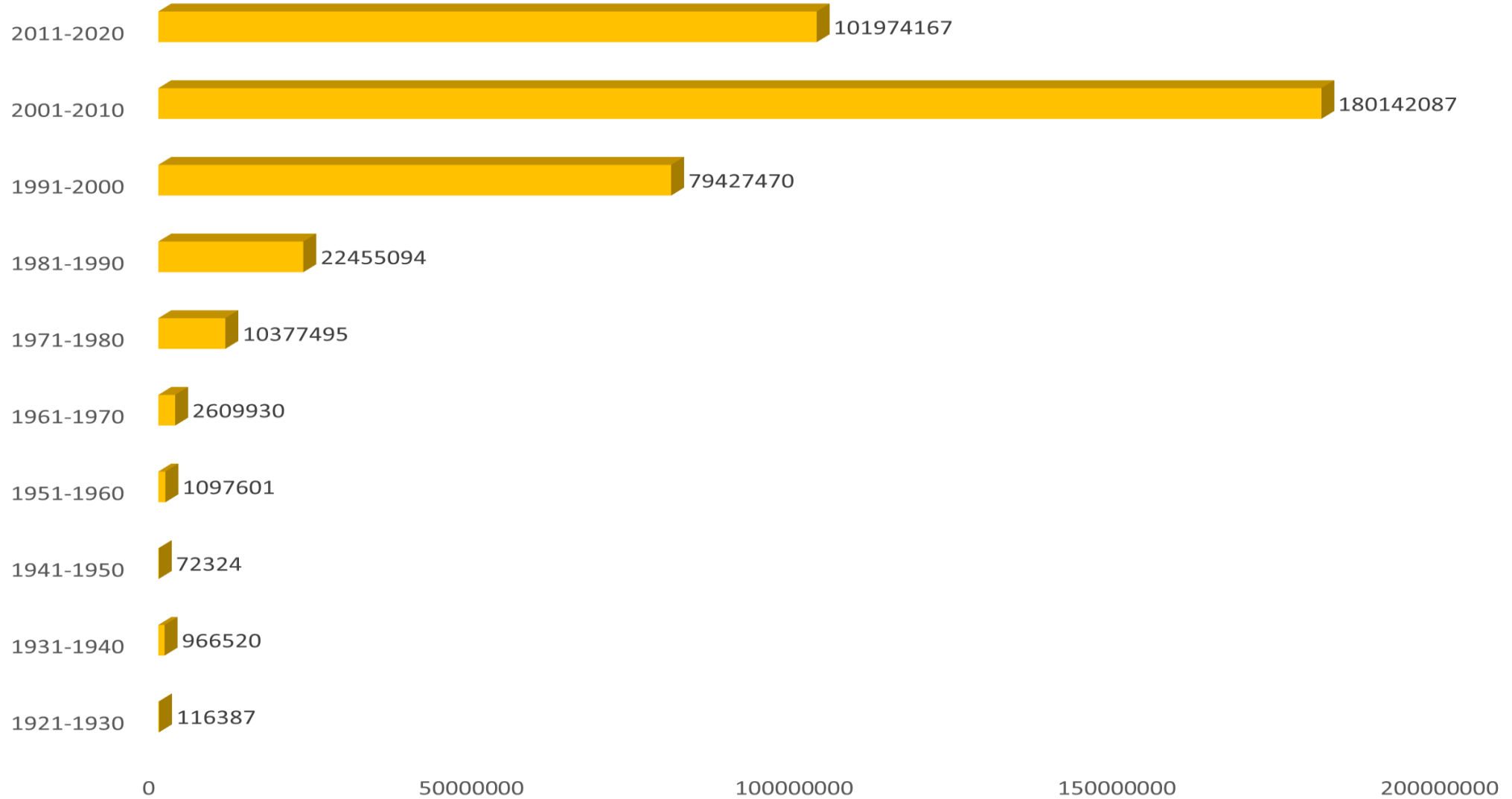
Results:

Thus, all the questions that where asked in the project have been provided with answer, data table and along with graphs.

From this project, I have learned how to clean data, manipulate data and about pivot tables and charts & graphs and techniques like sort, filter, pivot table, etc …

Poject-5 Results