PROJECT 5: IMDB Movie Analytics-I

Final Project – I

NAME: K. Yaswanth
GMAIL: kunayaswanth123@gmail.com

To provide the best report for the Internet Movie Database with the help of MS Excel knowledge.

SUBMITTING TO TRAINITY :)

SPECIAL THANKS TO TRAINITY TEAM

Project Description:

1. IMDB (Internet Movie Database) is a popular online platform where users can rate and review movies.
2. In this project, we will be analysing a dataset containing various features of movies, such as budget, genre, runtime, and cast, along with their corresponding IMDB ratings.
3. The objective of the project is to explore the IMDb dataset and perform exploratory data analysis

Approach:

1. Downloading IMDB_movies provided. Understanding the data.
2. Clean the data, like removing outliers, removing irrelevant columns, handling missing values, and formatting the data for better analysis.
3. Analyse the dataset using Excel functions and formulas to find meaningful insights such as movies with highest profit, best directors, popular movies etc.

Tech – Stack Used:

1. Microsoft Excel
2. Dataset statistics (given)

Insights:

Through this project, it became easy to learn the real-life project experiences and learned how to use function in large amount of data and help the imdb company.

The below are questions and answers for this imdb movie analytics are as follows:

A. Cleaning the data: This is one of the most important steps to perform before moving forward with the analysis. Use your knowledge learned till now to do this. (Dropping columns, removing null values, etc.)
Your task: Clean the data

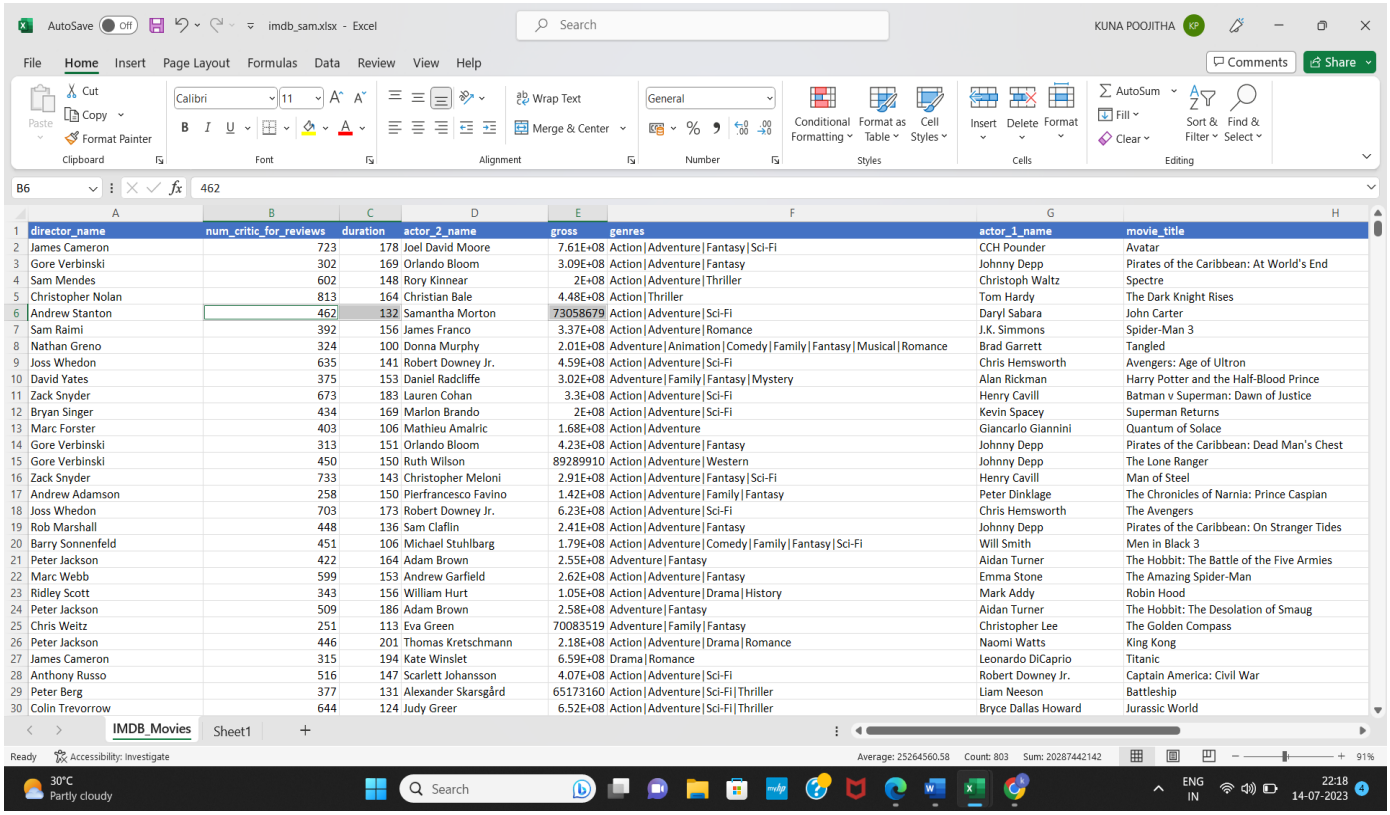
Ans: importing the csv file to MS Excel sheet.

1. Remove the duplicates (By going to data tab and clicking on remove duplicates). All duplicates are removed.
2. Remove unwanted columns
3. Remove the null values which contain the rows.

Removing unwanted columns by selecting columns and press ctrl + "- "

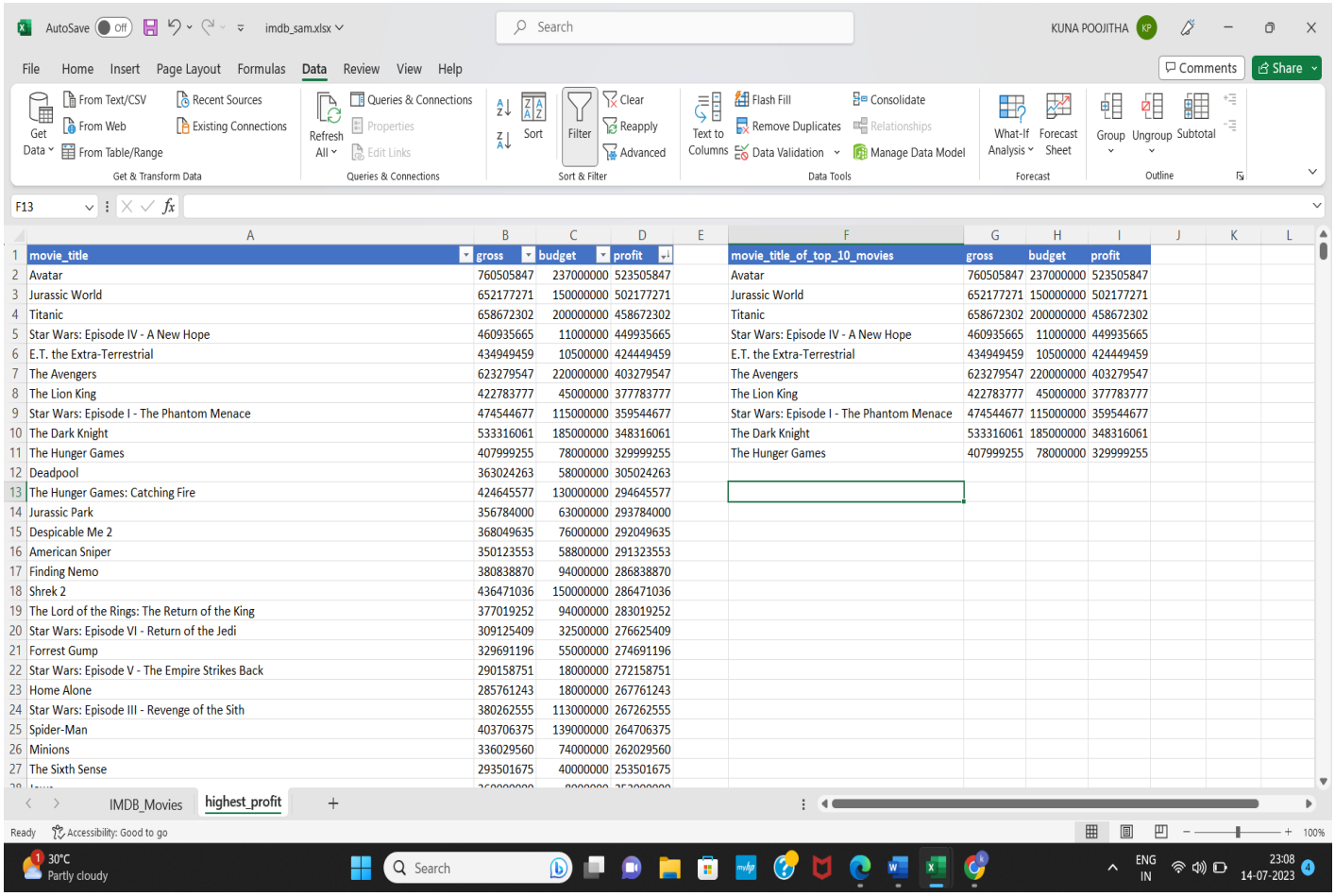
To remove null valued row, we go to special and select blank then press ctrl + "- "

In this way we clean the data.



	director_name	num_critic_for_reviews	duration	actor_2_name	gross	genres	actor_1_name	movie_title
2	James Cameron	723	178	Joel David Moore	7.61E+08	Action|Adventure|Fantasy|Sci-Fi	CCH Pounder	Avatar
3	Gore Verbinski	302	169	Orlando Bloom	3.09E+08	Action|Adventure|Fantasy	Johnny Depp	Pirates of the Caribbean: At World's End
4	Sam Mendes	602	148	Rory Kinnear	2E+08	Action|Adventure|Thriller	Christoph Waltz	Spectre
5	Christopher Nolan	813	164	Christian Bale	4.48E+08	Action|Thriller	Tom Hardy	The Dark Knight Rises
6	Andrew Stanton	462	132	Samantha Morton	73058679	Action|Adventure|Sci-Fi	Daryl Sabara	John Carter
7	Sam Raimi	392	156	James Franco	3.37E+08	Action|Adventure|Romance	J.K. Simmons	Spider-Man 3
8	Nathan Greno	324	100	Donna Murphy	2.01E+08	Adventure|Animation|Comedy|Family|Fantasy|Musical|Romance	Brad Garrett	Tangled
9	Joss Whedon	635	141	Robert Downey Jr.	4.59E+08	Action|Adventure|Sci-Fi	Chris Hemsworth	Avengers: Age of Ultron
10	David Yates	375	153	Daniel Radcliffe	3.02E+08	Adventure|Family|Fantasy|Mystery	Alan Rickman	Harry Potter and the Half-Blood Prince
11	Zack Snyder	673	183	Lauren Cohan	3.3E+08	Action|Adventure|Sci-Fi	Henry Cavill	Batman v Superman: Dawn of Justice
12	Bryan Singer	434	169	Marlon Brando	2E+08	Action|Adventure|Sci-Fi	Kevin Spacey	Superman Returns
13	Marc Forster	403	106	Mathieu Amalric	1.68E+08	Action|Adventure	Giancarlo Giannini	Quantum of Solace
14	Gore Verbinski	313	151	Orlando Bloom	4.23E+08	Action|Adventure|Fantasy	Johnny Depp	Pirates of the Caribbean: Dead Man's Chest
15	Gore Verbinski	450	150	Ruth Wilson	89289910	Action|Adventure|Western	Johnny Depp	The Lone Ranger
16	Zack Snyder	733	143	Christopher Meloni	2.91E+08	Action|Adventure|Fantasy|Sci-Fi	Henry Cavill	Man of Steel
17	Andrew Adamson	258	150	Pierfrancesco Favino	1.42E+08	Action|Adventure|Family|Fantasy	Peter Dinklage	The Chronicles of Narnia: Prince Caspian
18	Joss Whedon	703	173	Robert Downey Jr.	6.23E+08	Action|Adventure|Sci-Fi	Chris Hemsworth	The Avengers
19	Rob Marshall	448	136	Sam Claflin	2.41E+08	Action|Adventure|Fantasy	Johnny Depp	Pirates of the Caribbean: On Stranger Tides
20	Barry Sonnenfeld	451	106	Michael Stuhlbarg	1.79E+08	Action|Adventure|Comedy|Family|Fantasy|Sci-Fi	Will Smith	Men in Black 3
21	Peter Jackson	422	164	Adam Brown	2.55E+08	Adventure|Fantasy	Aidan Turner	The Hobbit: The Battle of the Five Armies
22	Marc Webb	599	153	Andrew Garfield	2.62E+08	Action|Adventure|Fantasy	Emma Stone	The Amazing Spider-Man
23	Ridley Scott	343	156	William Hurt	1.05E+08	Action|Adventure|Drama|History	Mark Addy	Robin Hood
24	Peter Jackson	509	186	Adam Brown	2.58E+08	Adventure|Fantasy	Aidan Turner	The Hobbit: The Desolation of Smaug
25	Chris Weitz	251	113	Eva Green	70083519	Adventure|Family|Fantasy	Christopher Lee	The Golden Compass
26	Peter Jackson	446	201	Thomas Kretschmann	2.18E+08	Action|Adventure|Drama|Romance	Naomi Watts	King Kong
27	James Cameron	315	194	Kate Winslet	6.59E+08	Drama|Romance	Leonardo DiCaprio	Titanic
28	Anthony Russo	516	147	Scarlett Johansson	4.07E+08	Action|Adventure|Sci-Fi	Robert Downey Jr.	Captain America: Civil War
29	Peter Berg	377	131	Alexander Skarsgård	65173160	Action|Adventure|Sci-Fi|Thriller	Liam Neeson	Battleship
30	Colin Trevorrow	644	124	Judy Greer	6.52E+08	Action|Adventure|Sci-Fi|Thriller	Bryce Dallas Howard	Jurassic World

B. Movies with highest profit: Create a new column called profit which contains the difference of the two columns: gross and budget. Sort the column using the profit column as reference. Plot profit (y-axis) vs budget (x- axis) and observe the outliers using the appropriate chart type.
Your task: Find the movies with the highest profit?



movie_title	gross	budget	profit	movie_title_of_top_10_movies	gross	budget	profit
Avatar	760505847	237000000	523505847	Avatar	760505847	237000000	523505847
Jurassic World	652177271	150000000	502177271	Jurassic World	652177271	150000000	502177271
Titanic	658672302	200000000	458672302	Titanic	658672302	200000000	458672302
Star Wars: Episode IV - A New Hope	460935665	11000000	449935665	Star Wars: Episode IV - A New Hope	460935665	11000000	449935665
E.T. the Extra-Terrestrial	434949459	10500000	424449459	E.T. the Extra-Terrestrial	434949459	10500000	424449459
The Avengers	623279547	220000000	403279547	The Avengers	623279547	220000000	403279547
The Lion King	422783777	45000000	377783777	The Lion King	422783777	45000000	377783777
Star Wars: Episode I - The Phantom Menace	474544677	115000000	359544677	Star Wars: Episode I - The Phantom Menace	474544677	115000000	359544677
The Dark Knight	533316061	185000000	348316061	The Dark Knight	533316061	185000000	348316061
The Hunger Games	407999255	78000000	329999255	The Hunger Games	407999255	78000000	329999255
Deadpool	363024263	58000000	305024263				
The Hunger Games: Catching Fire	424645577	130000000	294645577				
Jurassic Park	356784000	63000000	293784000				
Despicable Me 2	368049635	76000000	292049635				
American Sniper	350123553	58800000	291323553				
Finding Nemo	380838870	94000000	286838870				
Shrek 2	436471036	150000000	286471036				
The Lord of the Rings: The Return of the King	377019252	94000000	283019252				
Star Wars: Episode VI - Return of the Jedi	309125409	32500000	276625409				
Forrest Gump	329691196	55000000	274691196				
Star Wars: Episode V - The Empire Strikes Back	290158751	18000000	272158751				
Home Alone	285761243	18000000	267761243				
Star Wars: Episode III - Revenge of the Sith	380262555	113000000	267262555				
Spider-Man	403706375	139000000	264706375				
Minions	336029560	74000000	262029560				
The Sixth Sense	293501675	40000000	253501675				

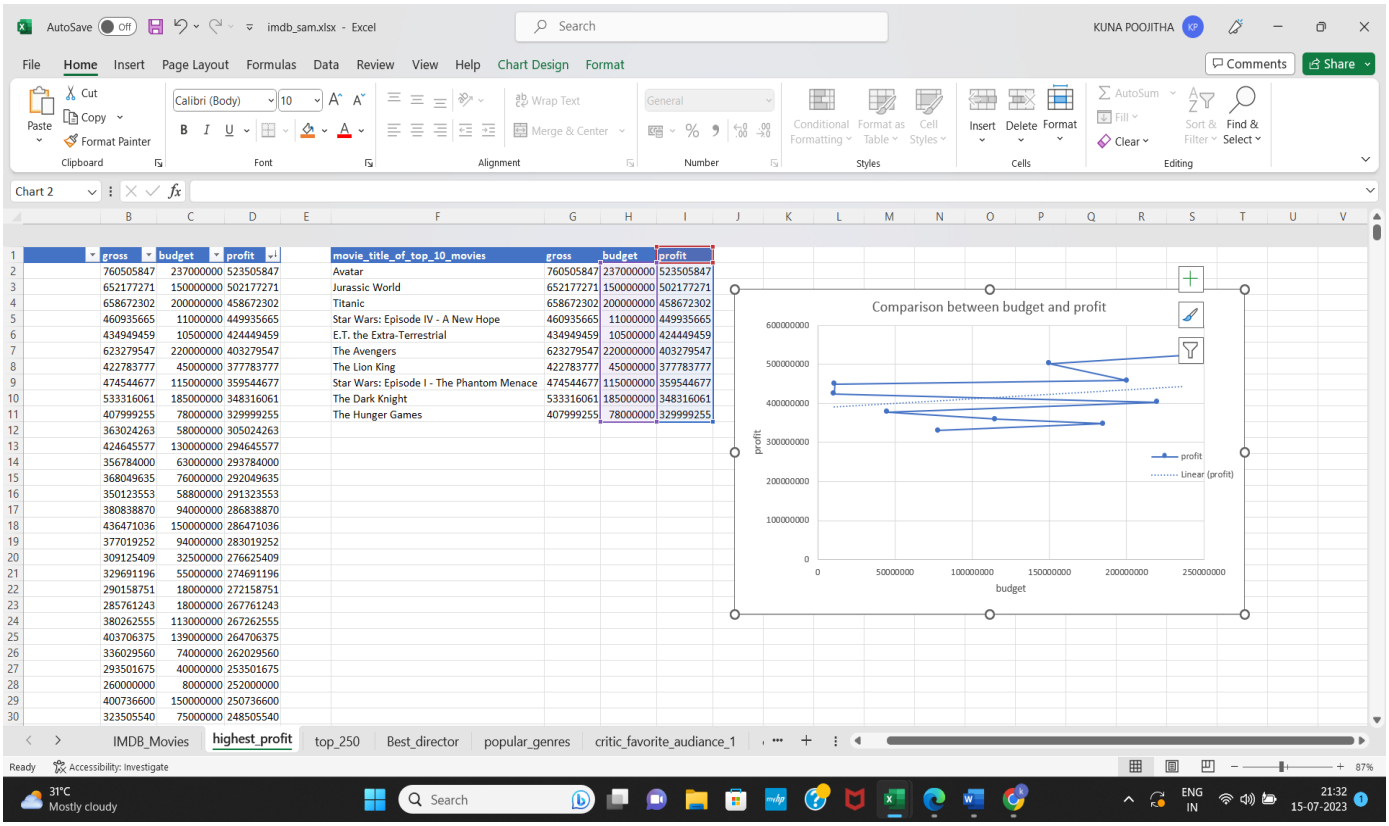
Top 10 movie titles by using filter option in data and we copy the movie_title column, gross column, budget column and paste it in another worksheet.

Steps:

1. Create new column profit which is difference between gross and budget.
2. To find top 10 movies we use resize the table and apply filter operation in data tab in descending order, here we find top 10 columns.

In this way we get the top 10 movies. In descending order.

Chart:



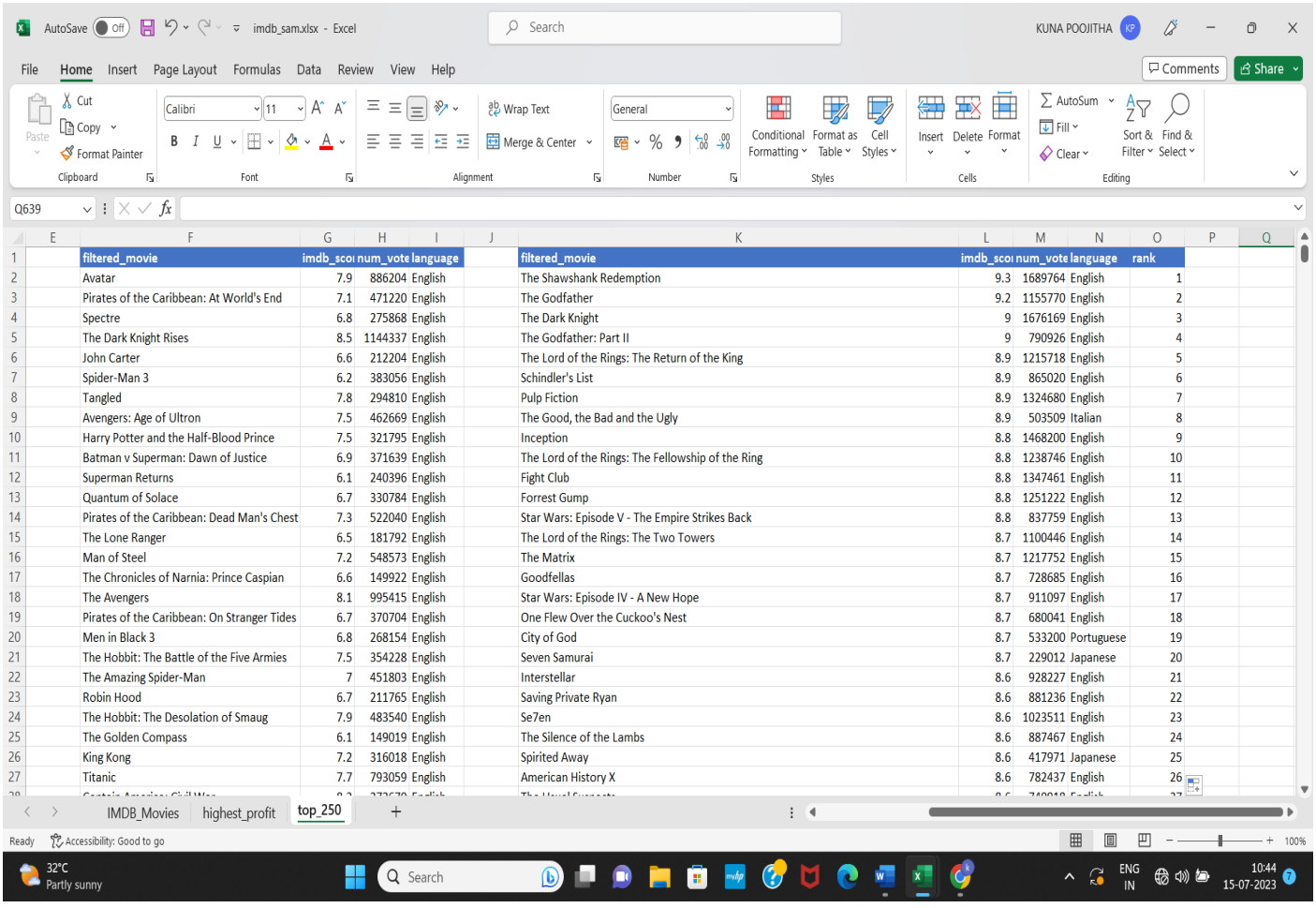
	gross	budget	profit	movie_title_of_top_10_movies	gross	budget	profit
2	760505847	237000000	523505847	Avatar	760505847	237000000	523505847
3	652177271	150000000	502177271	Jurassic World	652177271	150000000	502177271
4	658672302	200000000	458672302	Titanic	658672302	200000000	458672302
5	460935665	11000000	449935665	Star Wars: Episode IV - A New Hope	460935665	11000000	449935665
6	434949459	10500000	424449459	E.T. the Extra-Terrestrial	434949459	10500000	424449459
7	623279547	220000000	403279547	The Avengers	623279547	220000000	403279547
8	422783777	45000000	377783777	The Lion King	422783777	45000000	377783777
9	474544677	115000000	359544677	Star Wars: Episode I - The Phantom Menace	474544677	115000000	359544677
10	533316061	185000000	348316061	The Dark Knight	533316061	185000000	348316061
11	407999255	78000000	329999255	The Hunger Games	407999255	78000000	329999255
12	363024263	58000000	305024263				
13	424645577	130000000	294645577				
14	356784000	63000000	293784000				
15	368049635	76000000	292049635				
16	350123553	58800000	291323553				
17	380838870	94000000	286838870				
18	436471036	150000000	286471036				
19	377019252	94000000	283019252				
20	309125409	32500000	276625409				
21	329691196	55000000	274691196				
22	290158751	18000000	272158751				
23	285761243	18000000	267761243				
24	380262555	113000000	267262555				
25	403706375	139000000	264706375				
26	336029560	74000000	262029560				
27	293501675	40000000	253501675				
28	260000000	8000000	252000000				
29	400736600	150000000	250736600				
30	323505540	75000000	248505540				

C. **Top 250:** Create a new column IMDb_Top_250 and store the top 250 movies with the highest IMDb Rating (corresponding to the column: imdb_score). Also make sure that for all of these movies, the num_voted_users is greater than 25,000. Also add a Rank column containing the values 1 to 250 indicating the ranks of the corresponding films.

Extract all the movies in the IMDb_Top_250 column which are not in the English language and store them in a new column named Top_Foreign_Lang_Film. You can use your own imagination also!
Your task: Find IMDB Top 250

Ans:

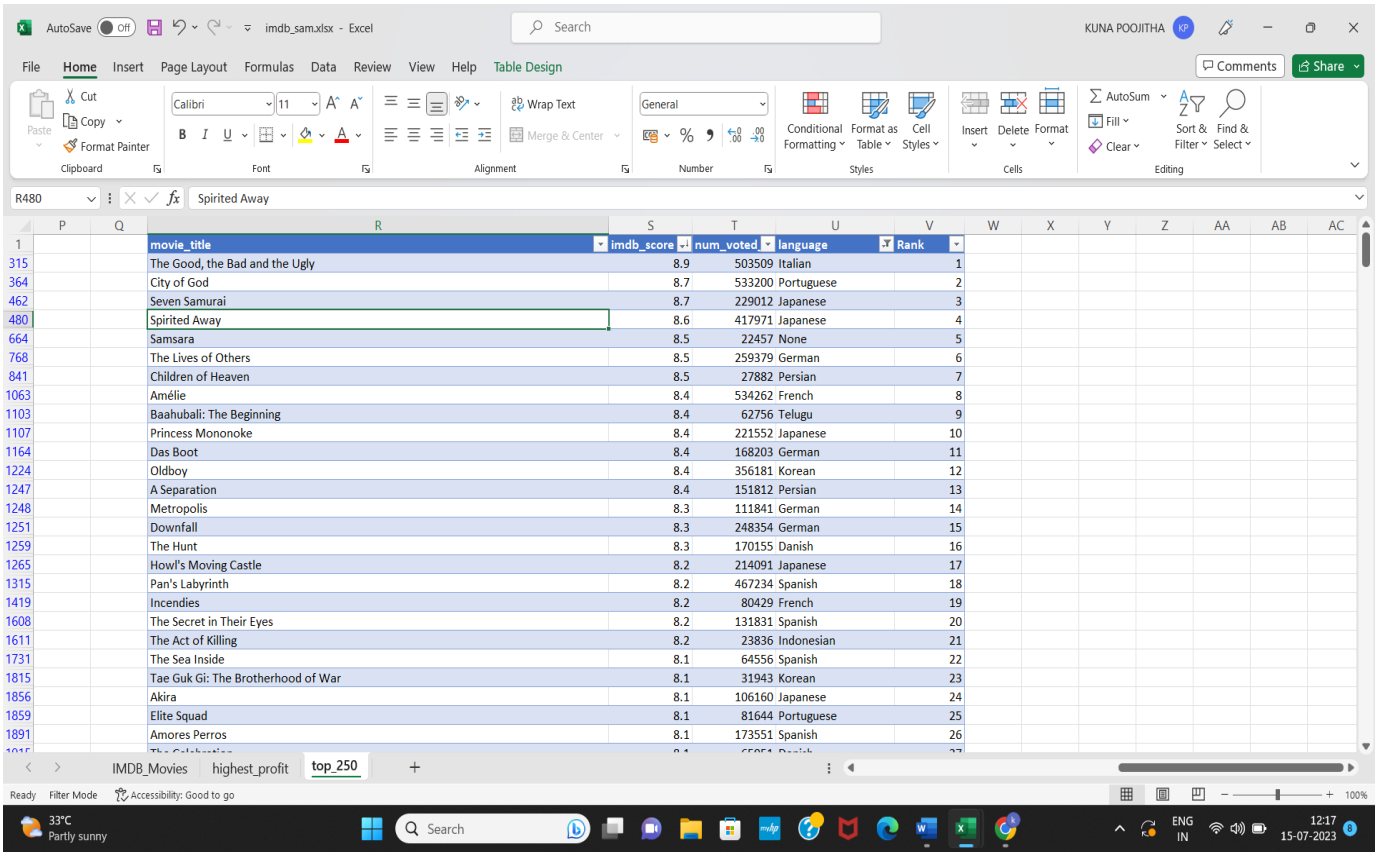
Here we filter the num_voted_users which is greater than 25,000 using filter function and by using sortby function we use descending order of imdb score.



filtered_movie	imdb_sco	num_vote	language	filtered_movie	imdb_sco	num_vote	language	rank
Avatar	7.9	886204	English	The Shawshank Redemption	9.3	1689764	English	1
Pirates of the Caribbean: At World's End	7.1	471220	English	The Godfather	9.2	1155770	English	2
Spectre	6.8	275868	English	The Dark Knight	9	1676169	English	3
The Dark Knight Rises	8.5	1144337	English	The Godfather: Part II	9	790926	English	4
John Carter	6.6	212204	English	The Lord of the Rings: The Return of the King	8.9	1215718	English	5
Spider-Man 3	6.2	383056	English	Schindler's List	8.9	865020	English	6
Tangled	7.8	294810	English	Pulp Fiction	8.9	1324680	English	7
Avengers: Age of Ultron	7.5	462669	English	The Good, the Bad and the Ugly	8.9	503509	Italian	8
Harry Potter and the Half-Blood Prince	7.5	321795	English	Inception	8.8	1468200	English	9
Batman v Superman: Dawn of Justice	6.9	371639	English	The Lord of the Rings: The Fellowship of the Ring	8.8	1238746	English	10
Superman Returns	6.1	240396	English	Fight Club	8.8	1347461	English	11
Quantum of Solace	6.7	330784	English	Forrest Gump	8.8	1251222	English	12
Pirates of the Caribbean: Dead Man's Chest	7.3	522040	English	Star Wars: Episode V - The Empire Strikes Back	8.8	837759	English	13
The Lone Ranger	6.5	181792	English	The Lord of the Rings: The Two Towers	8.7	1100446	English	14
Man of Steel	7.2	548573	English	The Matrix	8.7	1217752	English	15
The Chronicles of Narnia: Prince Caspian	6.6	149922	English	Goodfellas	8.7	728685	English	16
The Avengers	8.1	995415	English	Star Wars: Episode IV - A New Hope	8.7	911097	English	17
Pirates of the Caribbean: On Stranger Tides	6.7	370704	English	One Flew Over the Cuckoo's Nest	8.7	680041	English	18
Men in Black 3	6.8	268154	English	City of God	8.7	533200	Portuguese	19
The Hobbit: The Battle of the Five Armies	7.5	354228	English	Seven Samurai	8.7	229012	Japanese	20
The Amazing Spider-Man	7	451803	English	Interstellar	8.6	928227	English	21
Robin Hood	6.7	211765	English	Saving Private Ryan	8.6	881236	English	22
The Hobbit: The Desolation of Smaug	7.9	483540	English	Se7en	8.6	1023511	English	23
The Golden Compass	6.1	149019	English	The Silence of the Lambs	8.6	887467	English	24
King Kong	7.2	316018	English	Spirited Away	8.6	417971	Japanese	25
Titanic	7.7	793059	English	American History X	8.6	782437	English	26

Steps: -
1. First, filter movies with votes >25000 were found using the FILTER
2. =FILTER(F2:I3856,num_voted_users>=25000)
3. To find the top 250 movies
4. =SORT(FILTER(F2:H2633,F2:F2633>=LARGE(F2:F2633,250)),{1,3},{-1,-1})
5. We get top 250 columns

Ans2: To find top 250 in foreign language



movie_title	imdb_score	num_voted	language	Rank
The Good, the Bad and the Ugly	8.9	503509	Italian	1
City of God	8.7	533200	Portuguese	2
Seven Samurai	8.7	229012	Japanese	3
Spirited Away	8.6	417971	Japanese	4
Samsara	8.5	22457	None	5
The Lives of Others	8.5	259379	German	6
Children of Heaven	8.5	27882	Persian	7
Amélie	8.4	534262	French	8
Baahubali: The Beginning	8.4	62756	Telugu	9
Princess Mononoke	8.4	221552	Japanese	10
Das Boot	8.4	168203	German	11
Oldboy	8.4	356181	Korean	12
A Separation	8.4	151812	Persian	13
Metropolis	8.3	111841	German	14
Downfall	8.3	248354	German	15
The Hunt	8.3	170155	Danish	16
Howl's Moving Castle	8.2	214091	Japanese	17
Pan's Labyrinth	8.2	467234	Spanish	18
Incendies	8.2	80429	French	19
The Secret in Their Eyes	8.2	131831	Spanish	20
The Act of Killing	8.2	23836	Indonesian	21
The Sea Inside	8.1	64556	Spanish	22
Tae Guk Gi: The Brotherhood of War	8.1	31943	Korean	23
Akira	8.1	106160	Japanese	24
Elite Squad	8.1	81644	Portuguese	25
Amores Perros	8.1	173551	Spanish	26

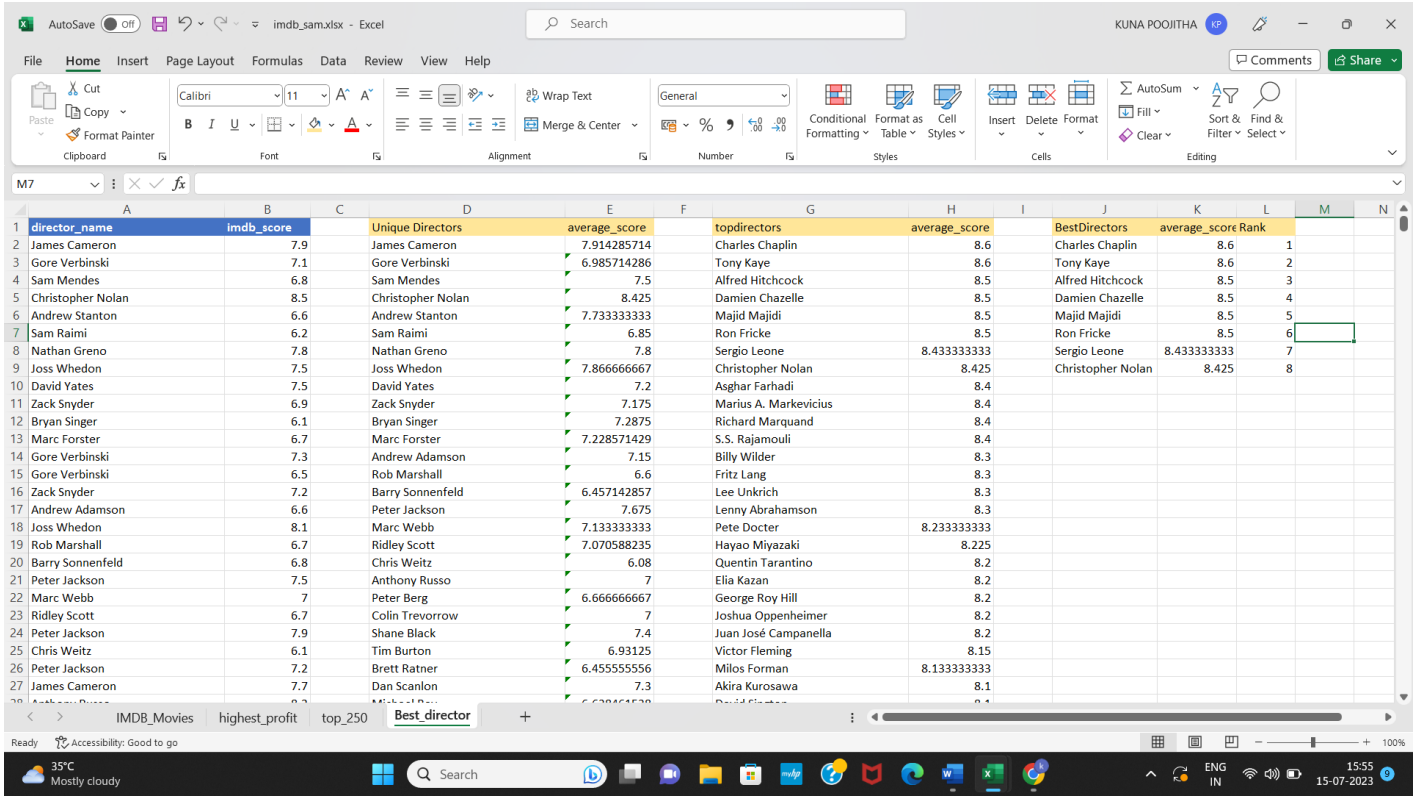
Steps:

1. We create the same columns by copying the cells.
2. We create the table, then apply the filter
3. Sort by descending order of imdb_score
4. Now, we rank the order as 1,2,3…

D. **Best Directors:** To Group the column using the director_name column.

Find out the top 10 directors for whom the mean of imdb_score is the highest and store them in a new column top10director. In case of a tie in IMDb score between two directors, sort them alphabetically.
Your task: Find the best directors

Ans: -



director_name	imdb_score	Unique Directors	average_score	topdirectors	average_score	BestDirectors	average_score	Rank
James Cameron	7.9	James Cameron	7.914285714	Charles Chaplin	8.6	Charles Chaplin	8.6	1
Gore Verbinski	7.1	Gore Verbinski	6.985714286	Tony Kaye	8.6	Tony Kaye	8.6	2
Sam Mendes	6.8	Sam Mendes	7.5	Alfred Hitchcock	8.5	Alfred Hitchcock	8.5	3
Christopher Nolan	8.5	Christopher Nolan	8.425	Damien Chazelle	8.5	Damien Chazelle	8.5	4
Andrew Stanton	6.6	Andrew Stanton	7.733333333	Majid Majidi	8.5	Majid Majidi	8.5	5
Sam Raimi	6.2	Sam Raimi	6.85	Ron Fricke	8.5	Ron Fricke	8.5	6
Nathan Greno	7.8	Nathan Greno	7.8	Sergio Leone	8.433333333	Sergio Leone	8.433333333	7
Joss Whedon	7.5	Joss Whedon	7.866666667	Christopher Nolan	8.425	Christopher Nolan	8.425	8
David Yates	7.5	David Yates	7.2	Asghar Farhadi	8.4			
Zack Snyder	6.9	Zack Snyder	7.175	Marius A. Markevicius	8.4			
Bryan Singer	6.1	Bryan Singer	7.2875	Richard Marquand	8.4			
Marc Forster	6.7	Marc Forster	7.228571429	S.S. Rajamouli	8.4			
Gore Verbinski	7.3	Andrew Adamson	7.15	Billy Wilder	8.3			
Gore Verbinski	6.5	Rob Marshall	6.6	Fritz Lang	8.3			
Zack Snyder	7.2	Barry Sonnenfeld	6.457142857	Lee Unkrich	8.3			
Andrew Adamson	6.6	Peter Jackson	7.675	Lenny Abrahamson	8.3			
Joss Whedon	8.1	Marc Webb	7.133333333	Pete Docter	8.233333333			
Rob Marshall	6.7	Ridley Scott	7.070588235	Hayao Miyazaki	8.225			
Barry Sonnenfeld	6.8	Chris Weitz	6.08	Quentin Tarantino	8.2			
Peter Jackson	7.5	Anthony Russo	7	Elia Kazan	8.2			
Marc Webb	7	Peter Berg	6.666666667	George Roy Hill	8.2			
Ridley Scott	6.7	Colin Trevorrow	7	Joshua Oppenheimer	8.2			
Peter Jackson	7.9	Shane Black	7.4	Juan José Campanella	8.2			
Chris Weitz	6.1	Tim Burton	6.93125	Victor Fleming	8.15			
Peter Jackson	7.2	Brett Ratner	6.455555556	Milos Forman	8.133333333			
James Cameron	7.7	Dan Scanlon	7.3	Akira Kurosawa	8.1			

Steps: -

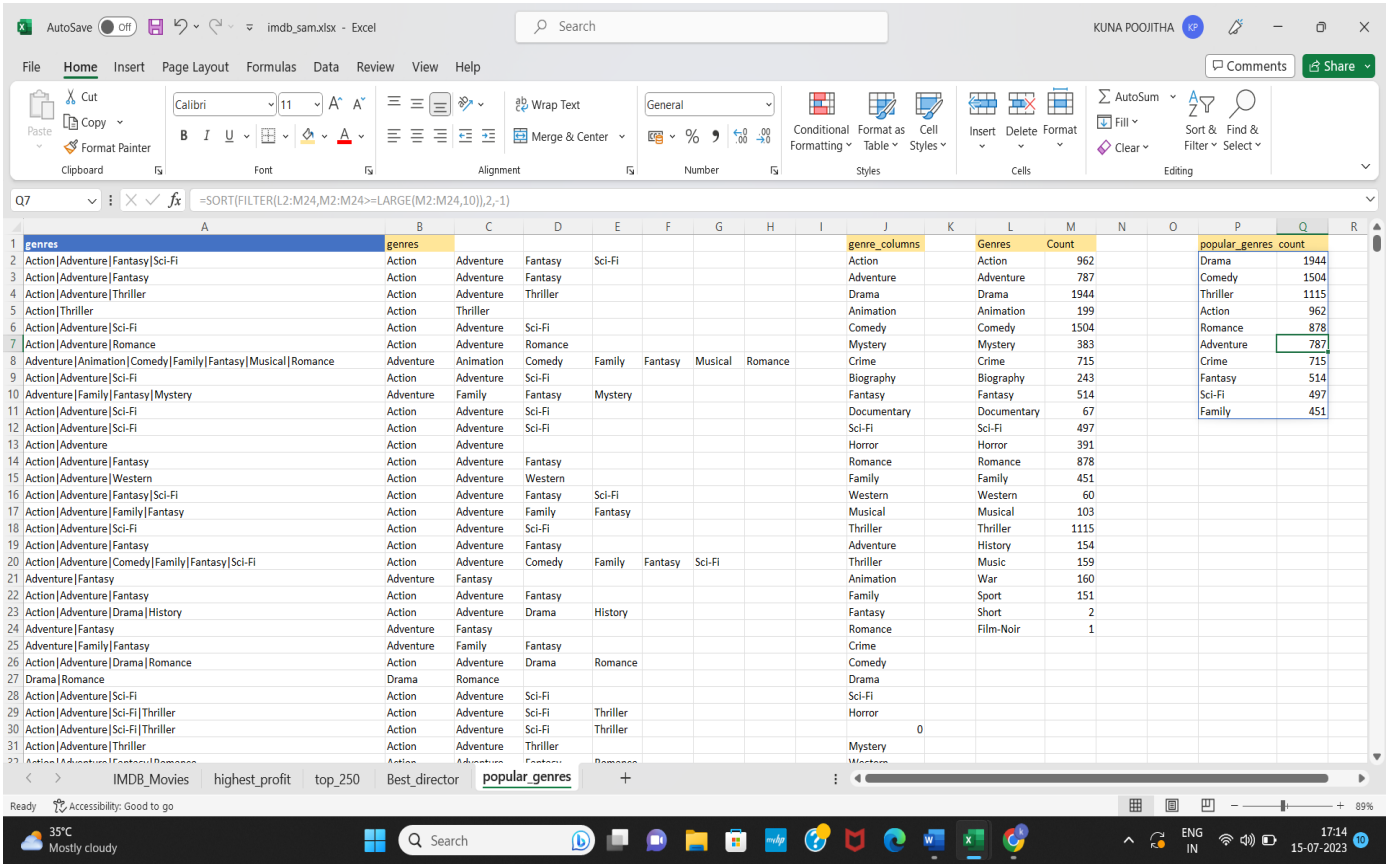
1. We copy the columns director_name, imdb_score from imdb_movies
2. We need to select all the unique rows. So, that we apply function =UNIQUE(A2:A3856) to get unique values.
3. We create other column and name as average_score, then we apply =AVERAGEIF(A2:A3856,D2,B2:B3856) to get all the average values of each director.
4. We sort the directors by applying the function based on requirements is =SORT(D2:E1752,{2,1},{-1,1})

5. We filter the data by applying the filter sort and filter operation for first 10 rows =SORT(FILTER(D2:E1752,E2:E1752>LARGE(E2:E1752,10)),{2,1},{-1,1})

6. We get desired result.

E. **Popular Genres:** Perform this step using the knowledge gained while performing previous steps.
Your task: Find popular genres

Ans:



Steps: -

1. Copying the genres column in imdb_movies, then go to data tab and apply text to columns
2. We set all the data to one column
3. Creating other column as genres and apply unique function to get all the rows and exclude 0
   =UNIQUE(FILTER(J2:J107,J2:J107<>0))
4. Applying countif function on each entity in genres to get the count
   =COUNTIFS($B$2:$H$3856,L2)
   Here $ is fixed.
5. To, find top 10 rows then, by applying sort
   =SORT(FILTER(L2:M24,M2:M24>=LARGE(M2:M24,10)),2,-1)
6. We, get desired result.

F. **Charts:** Create three new columns namely, Meryl_Streep, Leo_Caprio, and Brad_Pitt which contain the movies in which the actors: 'Meryl Streep', 'Leonardo DiCaprio', and 'Brad Pitt' are the lead actors. Use only the actor_1_name column for extraction. Also, make sure that you use the names 'Meryl Streep', 'Leonardo DiCaprio', and 'Brad Pitt' for the said extraction.

Append the rows of all these columns and store them in a new column named Combined.

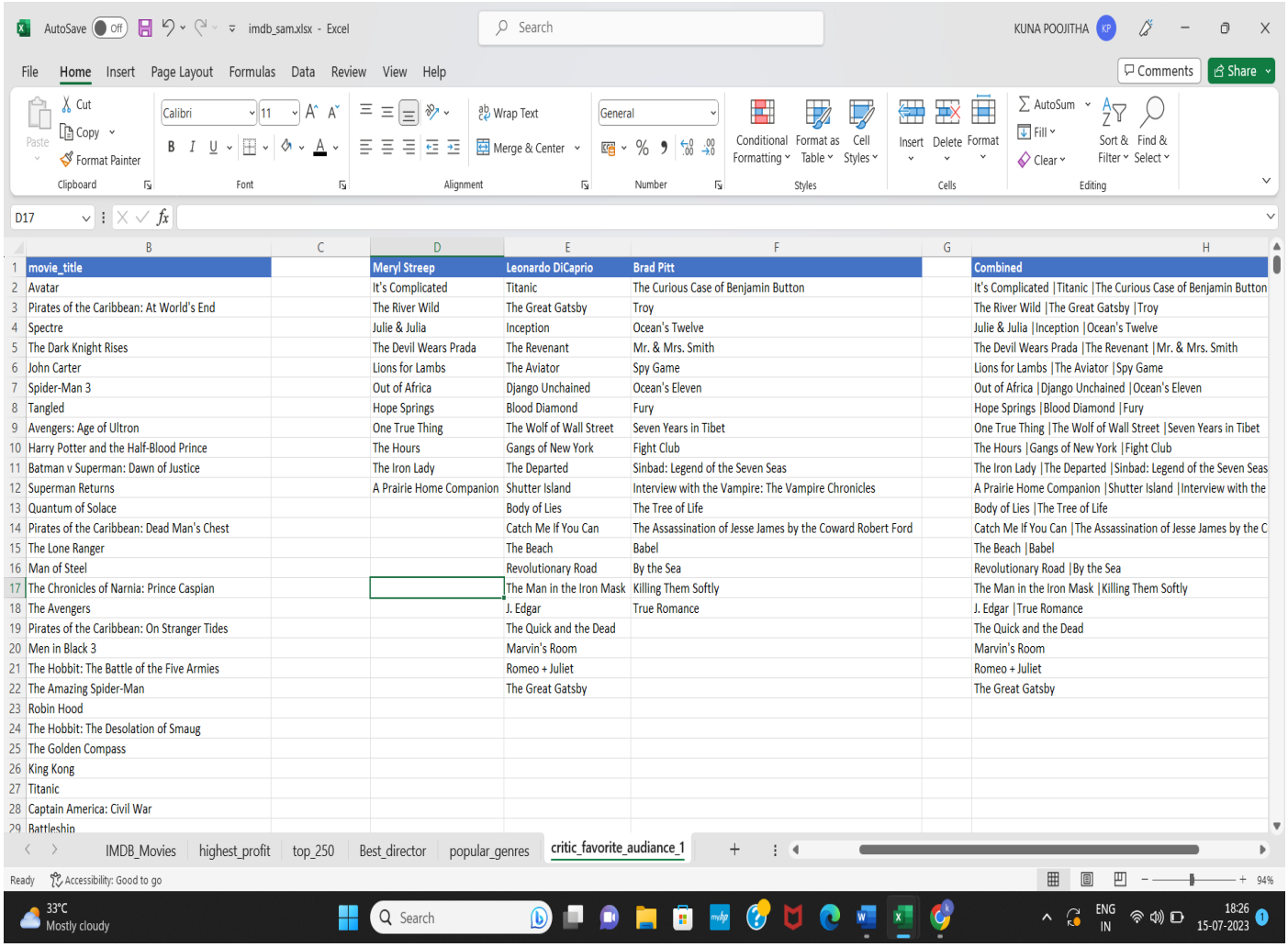
Group the combined column using the actor_1_name column.

Find the mean of the num_critic_for_reviews and num_users_for_review and identify the actors which have the highest mean.

Observe the change in number of voted users over decades using a bar chart. Create a column called decade which represents the decade to which every movie belongs to. For example, the title_year year 1923, 1925 should be stored as 1920s. Sort the column based on the column decade, group it by decade and find the sum of users voted in each decade. Store this in a new data frame called df_by_decade.

Your task: Find the critic-favorite and audience-favorite actors

Ans 1: -

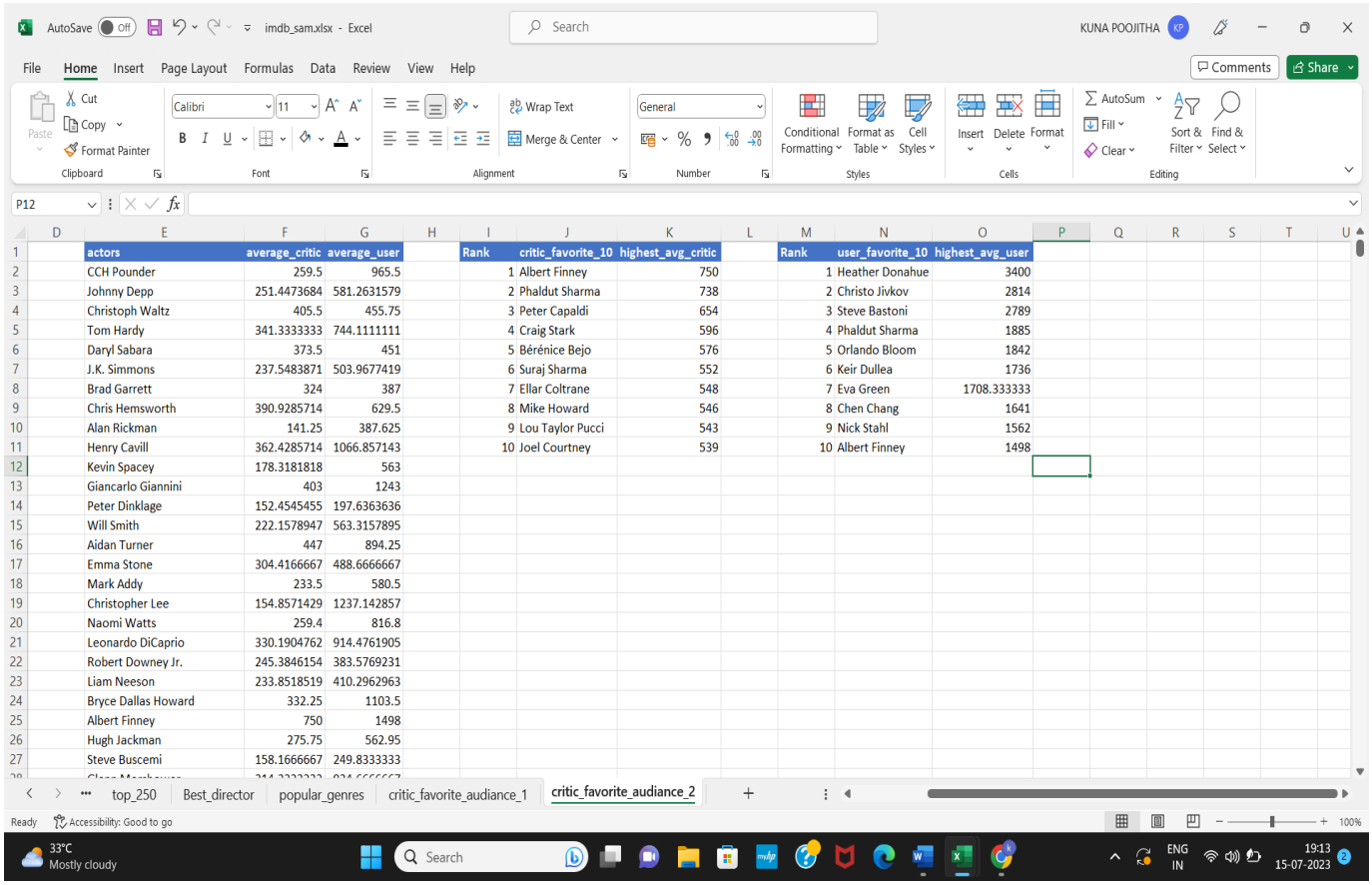


	movie_title	Meryl Streep	Leonardo DiCaprio	Brad Pitt	Combined
2	Avatar	It's Complicated	Titanic	The Curious Case of Benjamin Button	It's Complicated |Titanic |The Curious Case of Benjamin Button
3	Pirates of the Caribbean: At World's End	The River Wild	The Great Gatsby	Troy	The River Wild |The Great Gatsby |Troy
4	Spectre	Julie & Julia	Inception	Ocean's Twelve	Julie & Julia |Inception |Ocean's Twelve
5	The Dark Knight Rises	The Devil Wears Prada	The Revenant	Mr. & Mrs. Smith	The Devil Wears Prada |The Revenant |Mr. & Mrs. Smith
6	John Carter	Lions for Lambs	The Aviator	Spy Game	Lions for Lambs |The Aviator |Spy Game
7	Spider-Man 3	Out of Africa	Django Unchained	Ocean's Eleven	Out of Africa |Django Unchained |Ocean's Eleven
8	Tangled	Hope Springs	Blood Diamond	Fury	Hope Springs |Blood Diamond |Fury
9	Avengers: Age of Ultron	One True Thing	The Wolf of Wall Street	Seven Years in Tibet	One True Thing |The Wolf of Wall Street |Seven Years in Tibet
10	Harry Potter and the Half-Blood Prince	The Hours	Gangs of New York	Fight Club	The Hours |Gangs of New York |Fight Club
11	Batman v Superman: Dawn of Justice	The Iron Lady	The Departed	Sinbad: Legend of the Seven Seas	The Iron Lady |The Departed |Sinbad: Legend of the Seven Seas
12	Superman Returns	A Prairie Home Companion	Shutter Island	Interview with the Vampire: The Vampire Chronicles	A Prairie Home Companion |Shutter Island |Interview with the
13	Quantum of Solace		Body of Lies	The Tree of Life	Body of Lies |The Tree of Life
14	Pirates of the Caribbean: Dead Man's Chest		Catch Me If You Can	The Assassination of Jesse James by the Coward Robert Ford	Catch Me If You Can |The Assassination of Jesse James by the C
15	The Lone Ranger		The Beach	Babel	The Beach |Babel
16	Man of Steel		Revolutionary Road	By the Sea	Revolutionary Road |By the Sea
17	The Chronicles of Narnia: Prince Caspian		The Man in the Iron Mask	Killing Them Softly	The Man in the Iron Mask |Killing Them Softly
18	The Avengers		J. Edgar	True Romance	J. Edgar |True Romance
19	Pirates of the Caribbean: On Stranger Tides		The Quick and the Dead		The Quick and the Dead
20	Men in Black 3		Marvin's Room		Marvin's Room
21	The Hobbit: The Battle of the Five Armies		Romeo + Juliet		Romeo + Juliet
22	The Amazing Spider-Man		The Great Gatsby		The Great Gatsby
23	Robin Hood				
24	The Hobbit: The Desolation of Smaug				
25	The Golden Compass				
26	King Kong				
27	Titanic				
28	Captain America: Civil War				
29	Battleship				

Steps: -

1. Copy movie_title, actor_1_name from imdb_movies
2. To filter movies of Meryl Streep,Leonardo DiCaprioand BradPitt
   Meryl Streep =FILTER(B2:B3856,A2:A3856=D1)
   Leonardo Dicaprioand =FILTER(B2:B3856,A2:A3856=E1)
   Brad Pitt =FILTER(B2:B3856,A2:A3856=F1)
3. We get all the movies of the actors which they acted.
4. We join all the movies by using text join using delimiter “|”
   =TEXTJOIN("|",TRUE,D2,E2,F2)

Ans_2:

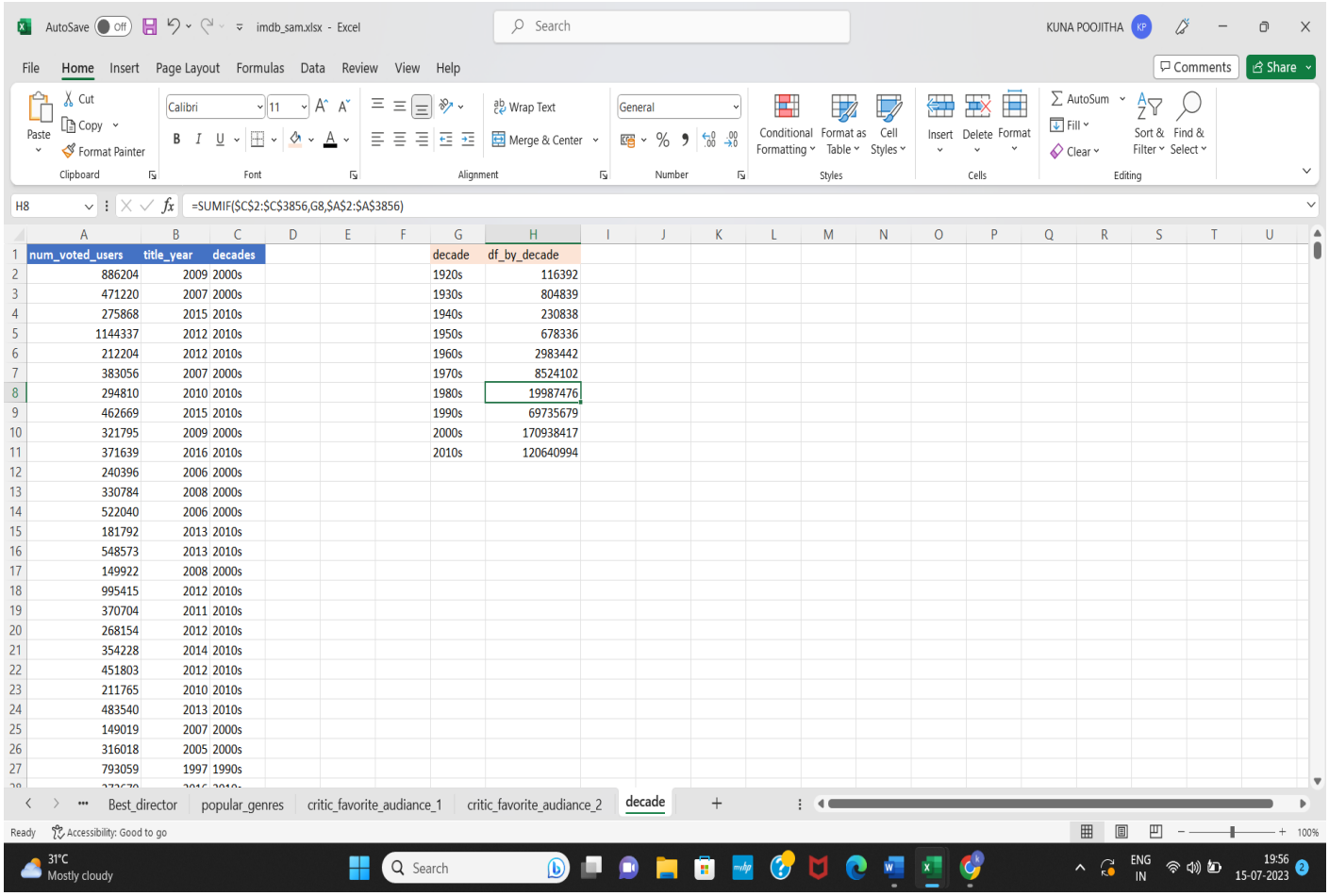


actors	average_critic	average_user	Rank	critic_favorite_10	highest_avg_critic	Rank	user_favorite_10	highest_avg_user
CCH Pounder	259.5	965.5	1	Albert Finney	750	1	Heather Donahue	3400
Johnny Depp	251.4473684	581.2631579	2	Phaldut Sharma	738	2	Christo Jivkov	2814
Christoph Waltz	405.5	455.75	3	Peter Capaldi	654	3	Steve Bastoni	2789
Tom Hardy	341.3333333	744.1111111	4	Craig Stark	596	4	Phaldut Sharma	1885
Daryl Sabara	373.5	451	5	Bérénice Bejo	576	5	Orlando Bloom	1842
J.K. Simmons	237.5483871	503.9677419	6	Suraj Sharma	552	6	Keir Dullea	1736
Brad Garrett	324	387	7	Ellar Coltrane	548	7	Eva Green	1708.333333
Chris Hemsworth	390.9285714	629.5	8	Mike Howard	546	8	Chen Chang	1641
Alan Rickman	141.25	387.625	9	Lou Taylor Pucci	543	9	Nick Stahl	1562
Henry Cavill	362.4285714	1066.857143	10	Joel Courtney	539	10	Albert Finney	1498
Kevin Spacey	178.3181818	563						
Giancarlo Giannini	403	1243						
Peter Dinklage	152.4545455	197.6363636						
Will Smith	222.1578947	563.3157895						
Aidan Turner	447	894.25						
Emma Stone	304.4166667	488.6666667						
Mark Addy	233.5	580.5						
Christopher Lee	154.8571429	1237.142857						
Naomi Watts	259.4	816.8						
Leonardo DiCaprio	330.1904762	914.4761905						
Robert Downey Jr.	245.3846154	383.5769231						
Liam Neeson	233.8518519	410.2962963						
Bryce Dallas Howard	332.25	1103.5						
Albert Finney	750	1498						
Hugh Jackman	275.75	562.95						
Steve Buscemi	158.1666667	249.8333333						

Steps: -

1. We copy the rows of num_critic_for_reviews , num_user_for_reviews, actor_1_name.
2. Filter unique actors using =UNIQUE(C2:C3856)
3. Then we find the average critcs and average users, of each actors.
   =AVERAGEIF($C$2:$C$3856,E2,$A$2:$A$3856)
   =AVERAGEIF($C$2:$C$3856,E2,$B$2:$B$3856)
4. From this created columns, we create critic_favourite_10 and highest_avg_critic using Sort method, getting top 10 critic actors and their average
   =SORT(FILTER(E2:F1509,F2:F1509>=LARGE(F2:F1509,10)),2,-1)
5. Similarly, we apply filter operation finding best actor accord to user using
   =FILTER(SORT(FILTER(E2:G1509,G2:G1509>=LARGE(G2:G1509,10)),3,-1),{1,0,1})
   Filters the top 10 rows according to user, sorting in descending order.

Ans_3:



num_voted_users	title_year	decades
886204	2009	2000s
471220	2007	2000s
275868	2015	2010s
1144337	2012	2010s
212204	2012	2010s
383056	2007	2000s
294810	2010	2010s
462669	2015	2010s
321795	2009	2000s
371639	2016	2010s
240396	2006	2000s
330784	2008	2000s
522040	2006	2000s
181792	2013	2010s
548573	2013	2010s
149922	2008	2000s
995415	2012	2010s
370704	2011	2010s
268154	2012	2010s
354228	2014	2010s
451803	2012	2010s
211765	2010	2010s
483540	2013	2010s
149019	2007	2000s
316018	2005	2000s
793059	1997	1990s

decade	df_by_decade
1920s	116392
1930s	804839
1940s	230838
1950s	678336
1960s	2983442
1970s	8524102
1980s	19987476
1990s	69735679
2000s	170938417
2010s	120640994

To find in a decade total number of voted users.

Steps:

1. Copy the columns num_voted_users, title_year
2. Create decades column and apply operation =LEFT(B2,3)&”0s” to get decades
3. Now, Create decade column and df_by_decade
4. In decade we apply =SORT(UNIQUE(C2:C3856),1,1) to get sorted and unique value
5. In df_by_decade column we apply =SUMIF($C$2:$C$3856,G2,$A$2:$A$3856)
   To get sum of all the num_voted_users in each decade.

   In this way, we get the critic_favourite and audience_favourite actors and corresponding results.

Excel Drive Link: -
https://docs.google.com/spreadsheets/d/1TS1fGRx0MdZlIWz34VAHE5uo3Eafqja6/edit?usp=sharing&ouid=103431433403818797511&rtpof=true&sd=true

Results: With the help of the project learned how to use functions in formula bar and learned how to use the functions in appropriate manner. Here, we find the best answers for all the questions will definitely benefit the company.

Thank you trainity