I. Project Summary

This project attempts to answer three questions using data from the TMDB movie database and the Oscars Awards dataset using Excel techniques to combine and calculate the value needed for them. The Excel file showing how the two datasets are combined for section II can be found here (The file is too large and need to download). The Excel file showing the steps to produce the answers for section III can be found here.

The first question is, "Does winning or being nominated for Oscars affect the ratings of movies on TMDB?". And the data illustrates that Oscar-winning or nominated movies receive higher TMDB ratings. More Oscar wins mean higher TMDB scores. Movies with a certain number of Oscars have higher ratings than those with nominations only. This trend applies to various award categories. This may be due to the Oscar awards being determined by professionals and experts in the film industry with their own criteria, and the mainstream audience will also have a chance of having the same standards as the experts when rating movies.

The second question is, "Does having casts or crews who won or are nominated for the Oscars awards affect the score on TMDB." The results found that there is a connection between the presence of Oscar-winning cast and crew members and the audience's score on TMDB. The reasons may stem from the TMDB score being the overall quality of the films that the mainstream audience thinks, and the quality and performance of actors, actresses and crews significantly affect the overall quality of the movie.

The Last question is, "Do the Oscars' Best Picture awards reflect a film's financial success?". The data clearly indicates that Best Picture Awards do not reflect a film's financial success, as some movies were not nominated or won the awards that received a significant profit. Even though the first question suggests that they can reflect the overall audience belief of the film's quality; however, having only high quality does not mean most people will watch the movies in cinemas since most of the revenue is calculated from the box office. There are a lot of factors that affect a film's commercial success; for instance, the period when the movies were released in the cinemas (they should not be released at the same time as another famous sequel movie, during the financial crisis or any incidents that decrease a chance of people buying a ticket for the film). How successful are the marketing campaigns that the studio makes which generate awareness and encourage people to watch movies in cinemas?

II. Wrangling Details

A. Dataset details

1. Dataset 1 (Provided) - TMDB 5000 Movie Dataset

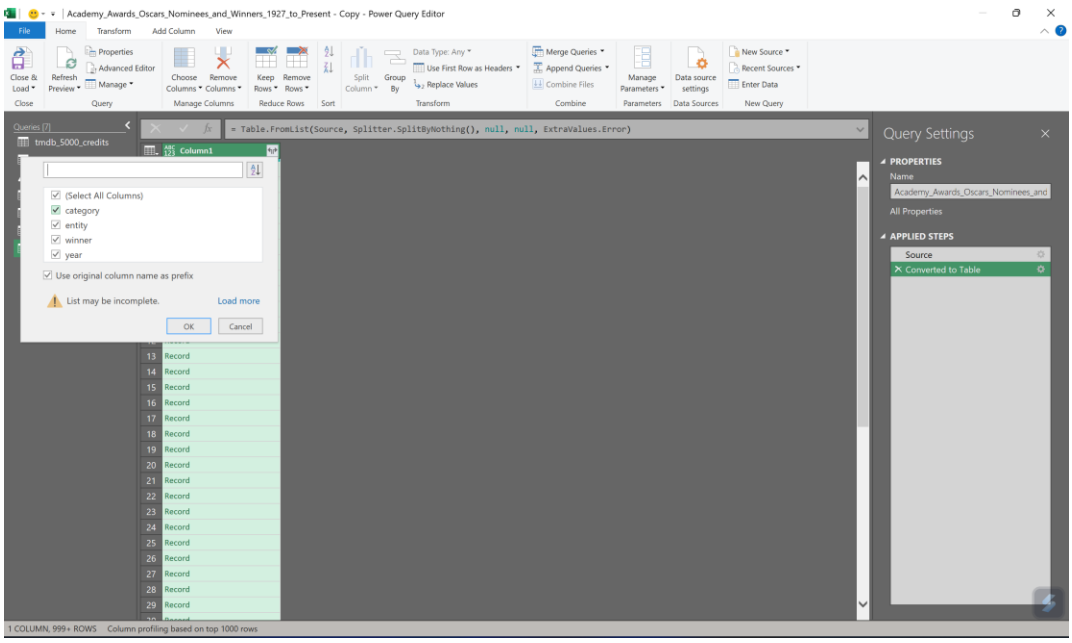
   The data in this dataset is extracted from the movie database website: TMDB, which is a website that provides information related to movies, TV shows, and actors. The data is in a CSV format and comprises two files: credits.csv and movies.csv, with 4 and 20 columns, respectively. The credit.csv has 19,212 blocks, and movies.csv has 96,120 blocks. However, there are some problems with this data. Firstly, the data in 7 columns are in JSON format, and two datasets are separated, so a combination of both must be done first before combining with dataset 2, plus two of them do not have the same number of rows with movies.csv has three rows more so we can assume that every rows are not directly connected to each other for two datasets. The movie.csv's columns include budget, genres, homepage, id, keywords, original_language, original_title, overview, popularity, production_companies, production_countries, release_date, revenue, runtime, spoken_languages, status, tagline, title, vote_average and vote_count and credits.csv includes movie_id, title and cast, crew while they are connected with id and movie_id

2. Dataset 2 - Academy Awards Oscars: Nominees and Winners 1927 to Present.

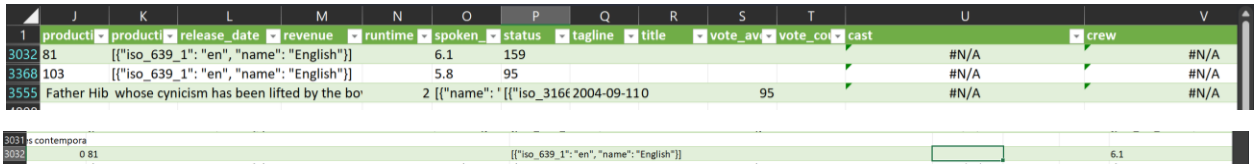
   The data is extracted from the Oscars website about Academy Awards, honouring excellence in the film industry. The format of this data is JSON, and the data has a very clean format which is objects inside an array because there are a total of 11,058 objects; every object has exactly four properties which are an entity, category, winner and year and values which all of them are single values which includes string, numeric and Boolean.

B. Step to combine the datasets

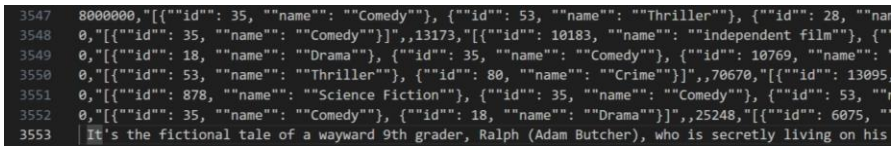
1. Before combining data sets, the two CSV files are imported into Excel files using From Text/CSV and JSON data is imported using Power Query.



2. A VlookUP with an exact match is used to combine movie.csv and credit.csv with this syntax: =VLOOKUP([@id],tmdb_5000_credits[#All],3,FALSE). This syntax returns the cast for each movie by matching the movie ID. As mentioned previously, movies.csv has three more rows, so after matching, it is found that all of the data in these rows are entered incorrectly and are supposed to belong to the row above (half of row 3031 is empty, and the format of each block is suggesting). Therefore a fix in the CSV files is needed, which can be simply done by merging lines 3031 & 3032, 3367 & 3368, 3554 & 3555 using VSCode (see 2nd Picture)

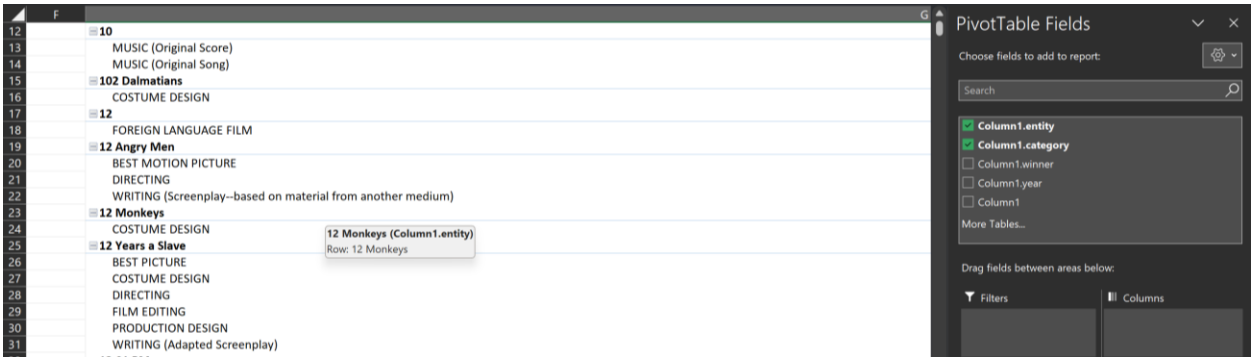


The error example (it shows 3553 because first two have already been fixed)

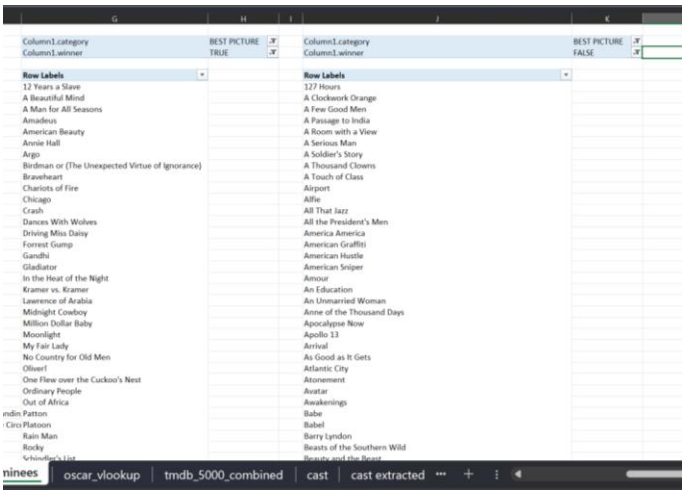


3. Redo the second step into TMDB_combined and still use the VlookUp function to ensure that all data matches are correct. Every row is combined successfully, as there is no #N/A value.

4. In Oscar.JSON, switching the column 1 category and column 2 entity because the second column will be used to match with TMDB, and Vlookup can only match the leftmost column.
5. This step will combine TMDB_combined with Oscar.JSON with the movie title. According to the pivot table of data in the Oscars dataset, a movie receives one or more awards, and there are more than 100 distinct award types. Therefore the best way to combine them is to only merge those that received the best award(best Picture) and the total number of the award received except the best Picture.

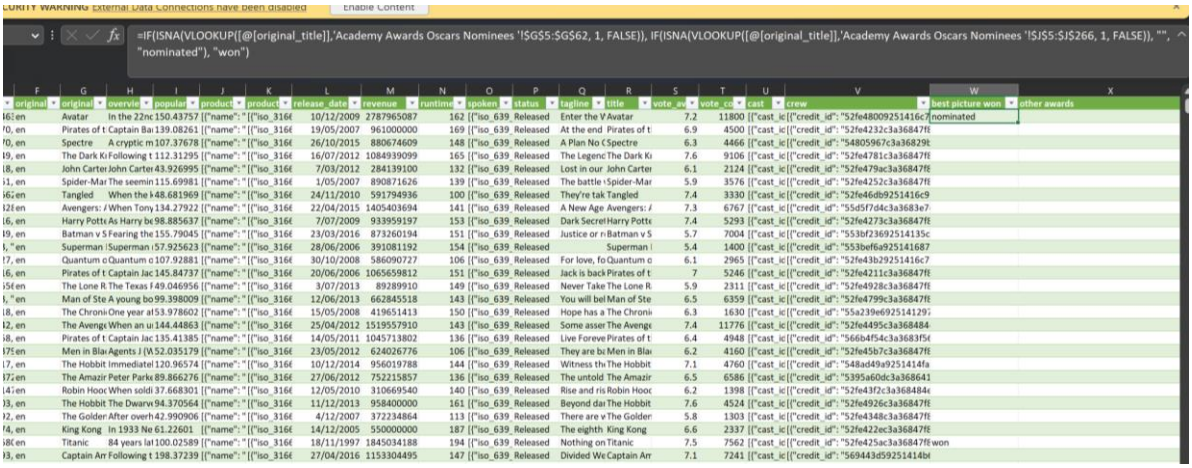


a. Use pivot table to get a list of movies that were nominated and won the best picture

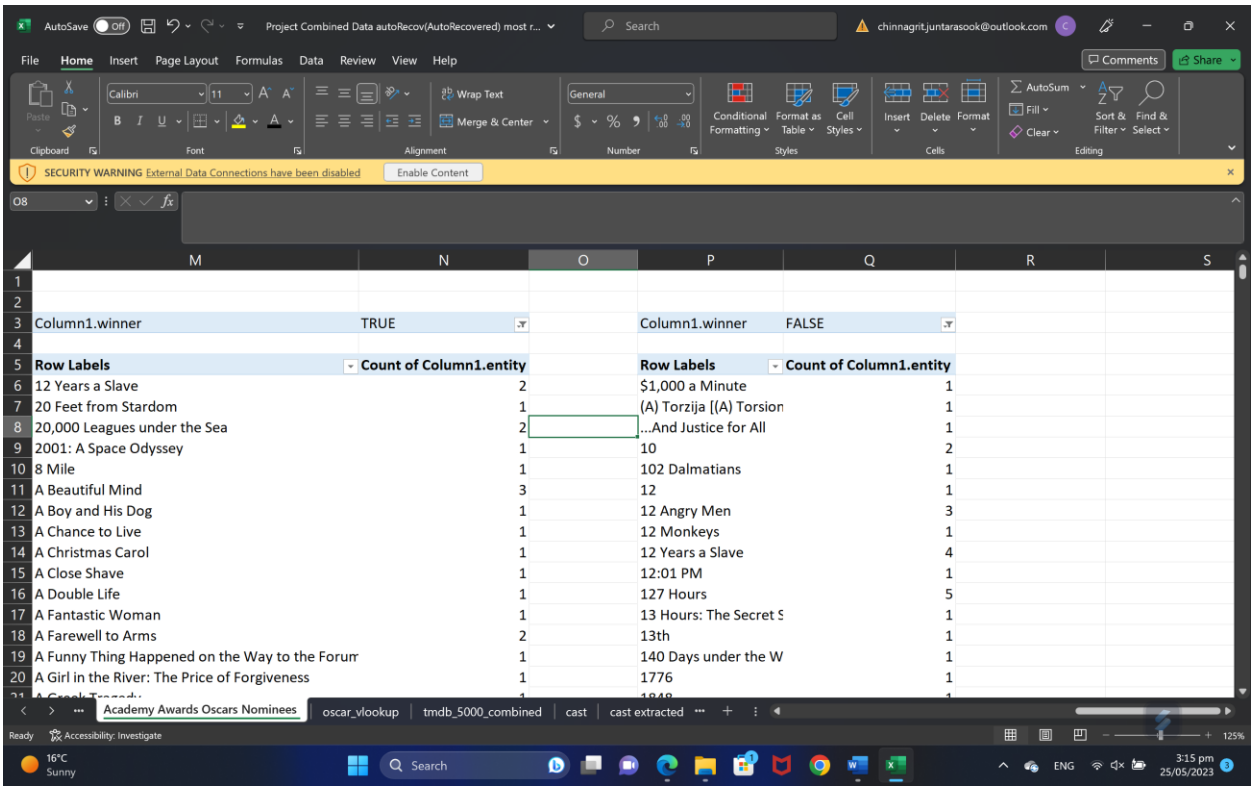


b. Use IF, ISNA and Vlookup function to check if a movie in that row is in the pivot table for best Picture & true or not. If the lookup returns #N/A (not found), then the ISNA function will return true and proceed to the second IF function, but if it returns a value (found), then that cell is equal to 'won'. The second if function has the same meaning but for checking the best Picture & false pivot table (nominated) and then returns 'nominated' if

found. A blank indicates that a movie has not won or nominated Oscars Best Picture award.



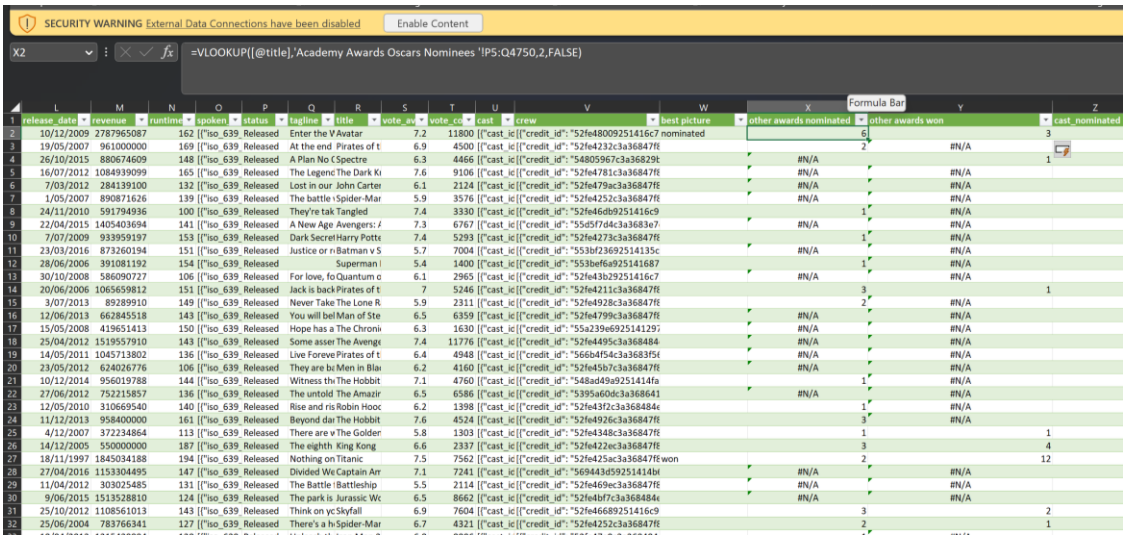
c. Create another two pivot tables to count every time a movie receive a nomination or is won in any award simply by counting how many time they appear in the datasets for TRUE(won) and FALSE(nominated) separately.



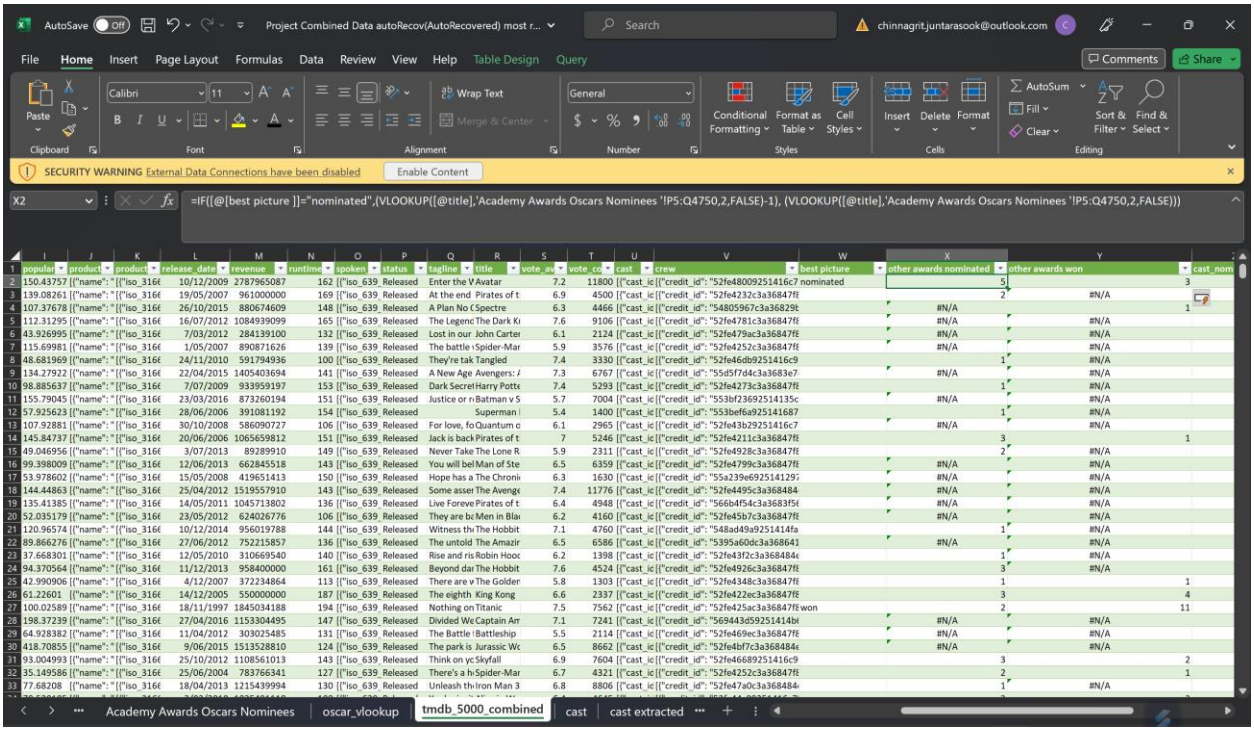
Column1.winner	TRUE
Row Labels	Count of Column1.entity
12 Years a Slave	2
20 Feet from Stardom	1
20,000 Leagues under the Sea	2
2001: A Space Odyssey	1
8 Mile	1
A Beautiful Mind	3
A Boy and His Dog	1
A Chance to Live	1
A Christmas Carol	1
A Close Shave	1
A Double Life	1
A Fantastic Woman	1
A Farewell to Arms	2
A Funny Thing Happened on the Way to the Forum	1
A Girl in the River: The Price of Forgiveness	1

Column1.winner	FALSE
Row Labels	Count of Column1.entity
$1,000 a Minute	1
(A) Torzija [(A) Torsion	1
...And Justice for All	1
10	2
102 Dalmatians	1
12	1
12 Angry Men	3
12 Monkeys	1
12 Years a Slave	4
12:01 PM	1
127 Hours	5
13 Hours: The Secret S	1
13th	1
140 Days under the W	1
1776	1

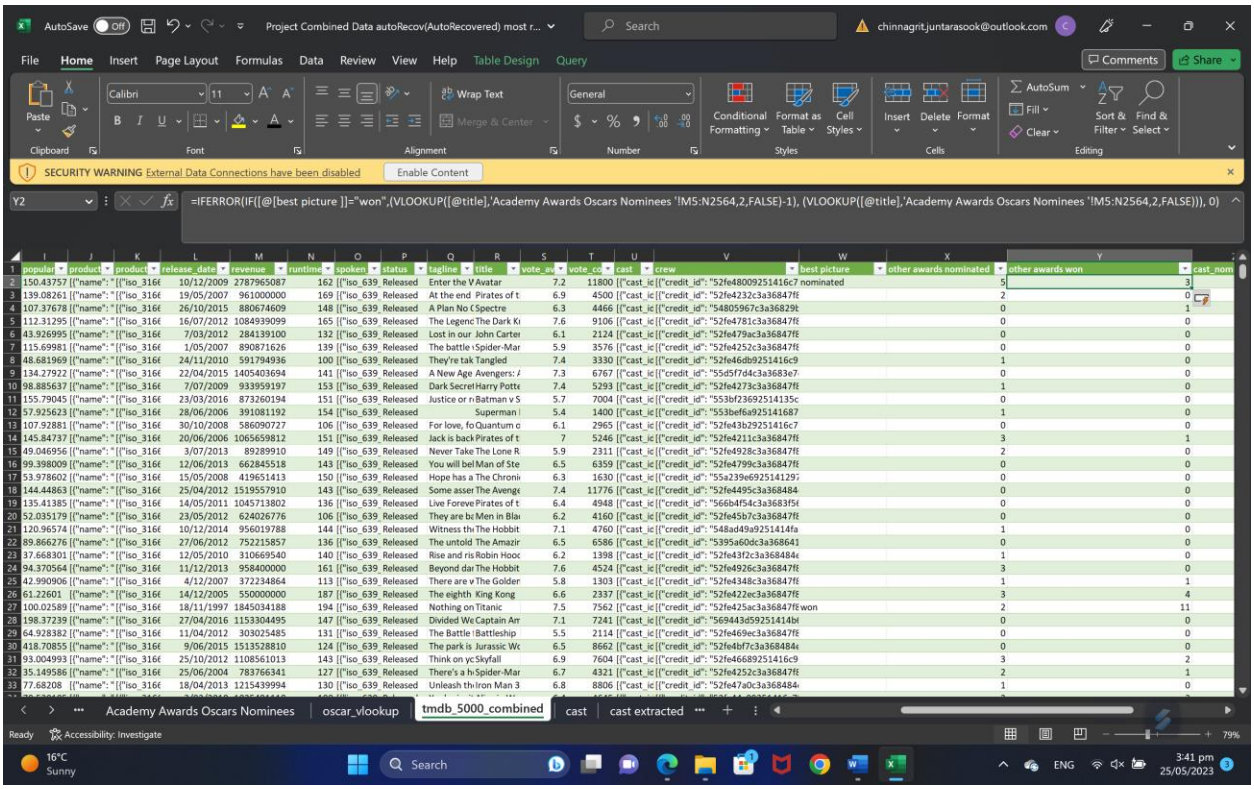
d. Use Vlookup function to merge these pivot tables with the tmdb_5000_combined



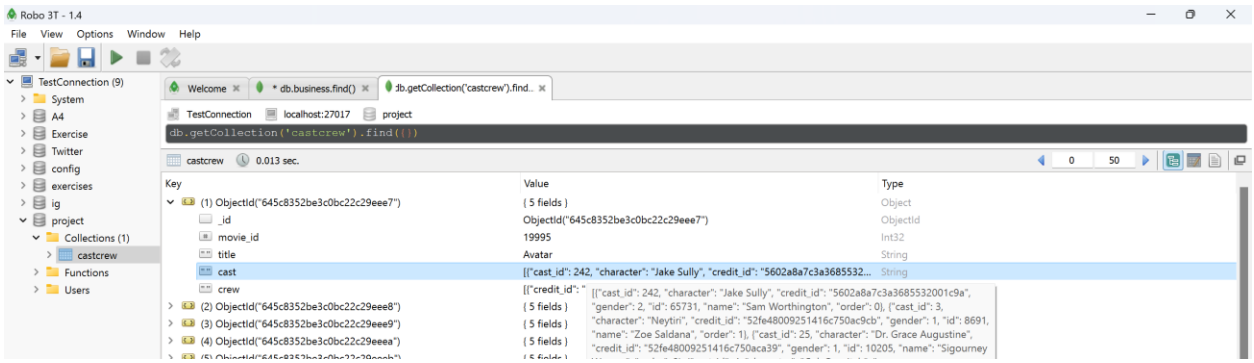
e. Modify the syntax so that they do not count the Best Picture by minus one in other awards nominated if the Best Picture cell is nominated and the same for won.



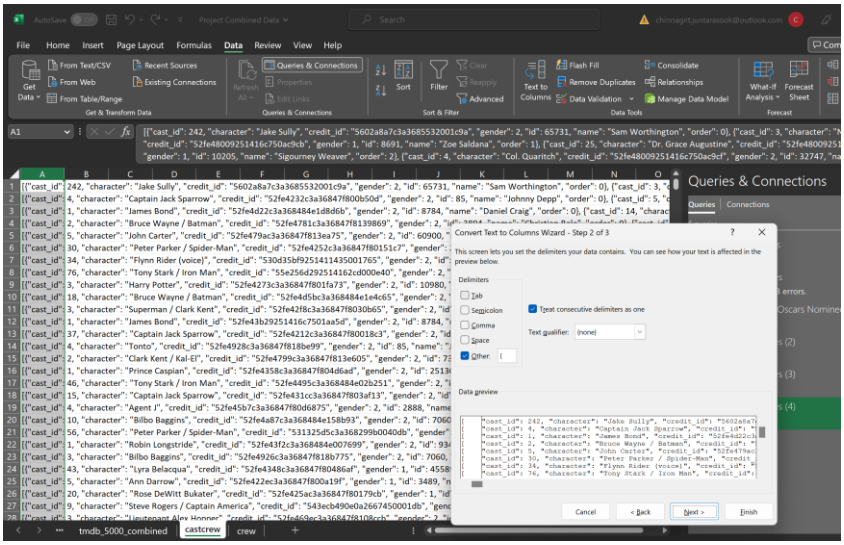
f. Modify the syntax so that it returns 0 instead of #N/A by using the IFERROR function

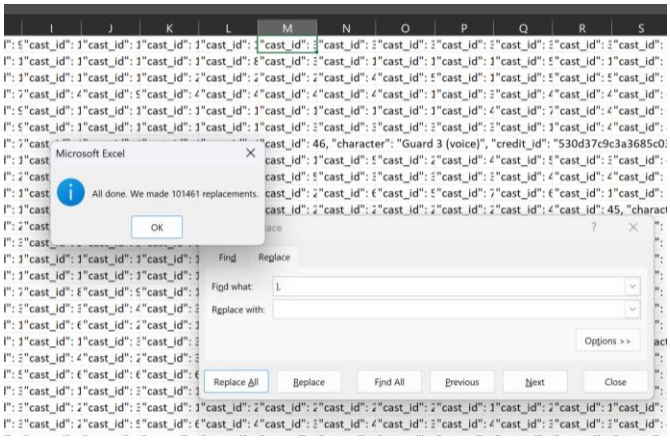


6. Assuming there will be no increase in the number of rows (4804 rows, including headers), importing credit.csv to MongoDB using this query: mongoimport --type csv --headerline --db project --collection castcrew --file tmdb_5000_credits.csv, for separating value in array of object purposes. However, MongoDB does not detect arrays of objects because there are double quotation marks around them. Therefore it is not possible to use Mongodb to separate each object.

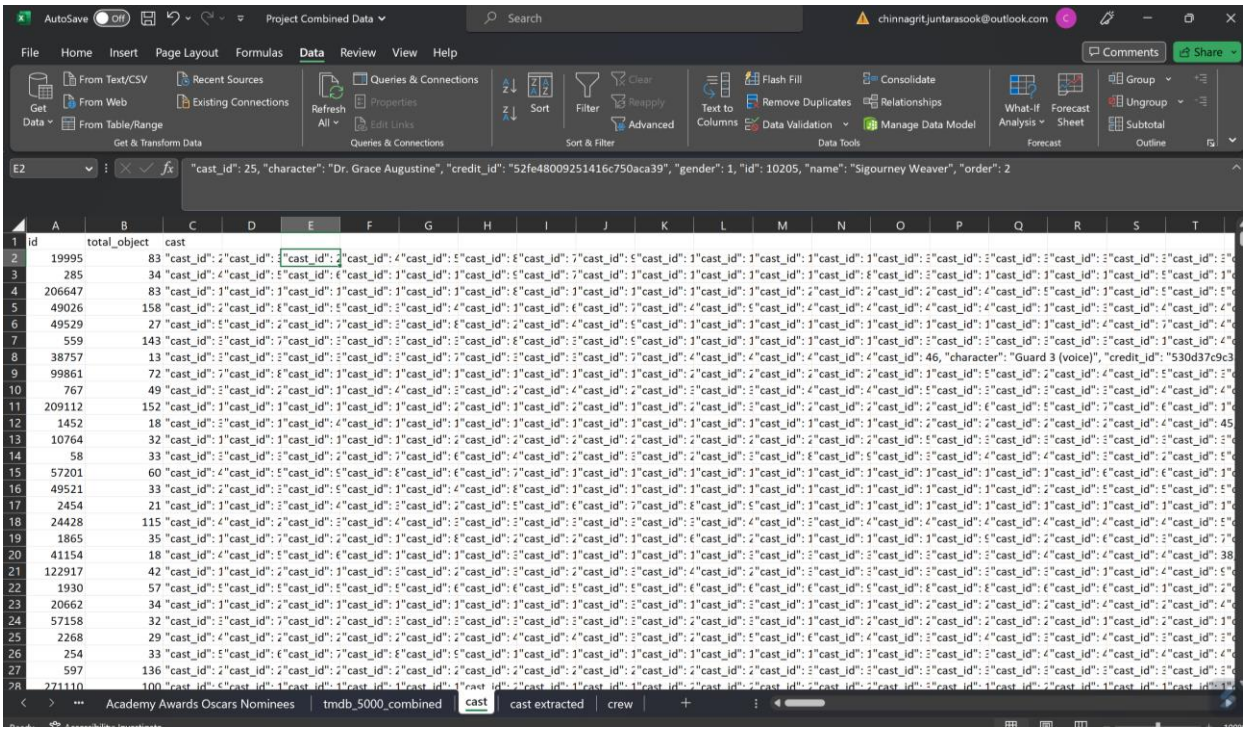


7. Therefore, we have to separate each object by using { signs so that every cast is separated into single columns and then use find and replace to delete all of }, and }] signs so that every block have the same format, which is "cast_id": value, "character": value, "credit_id": value, "gender": value, "id": value, "name": value, "order": value

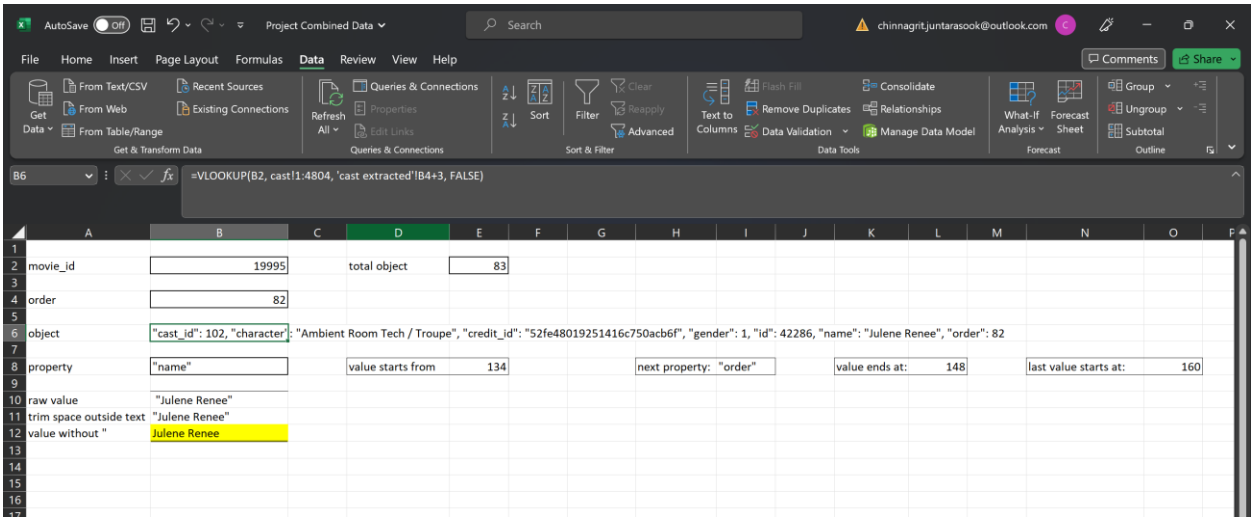


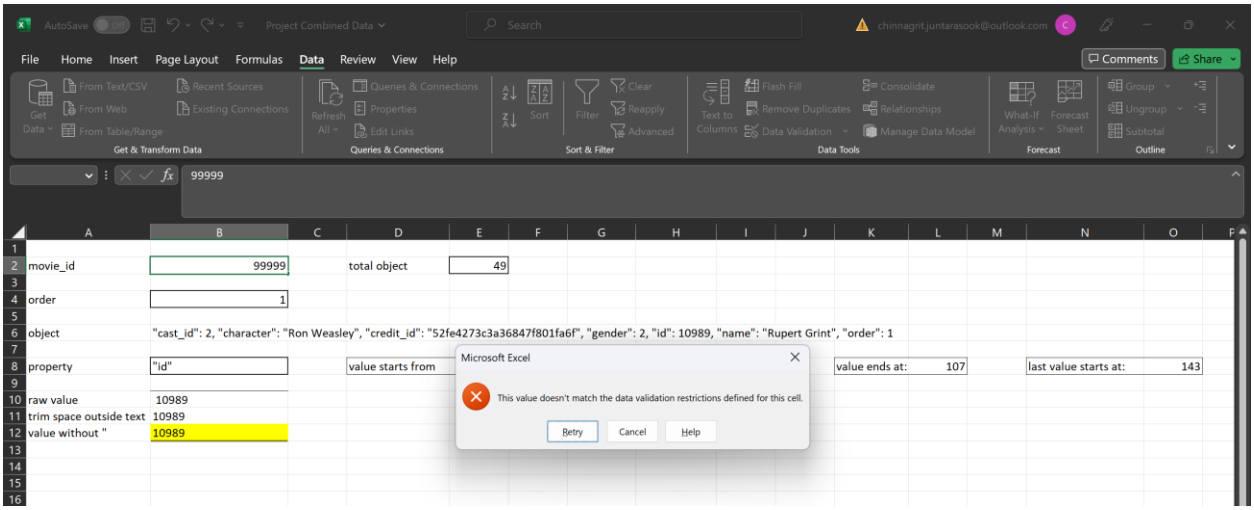


8. The next step is to create a value extractor from the object by using the data validation function and vlookup to answer the question and further exploration; the steps are
    a. Create an id and the total number of the cast for each movie

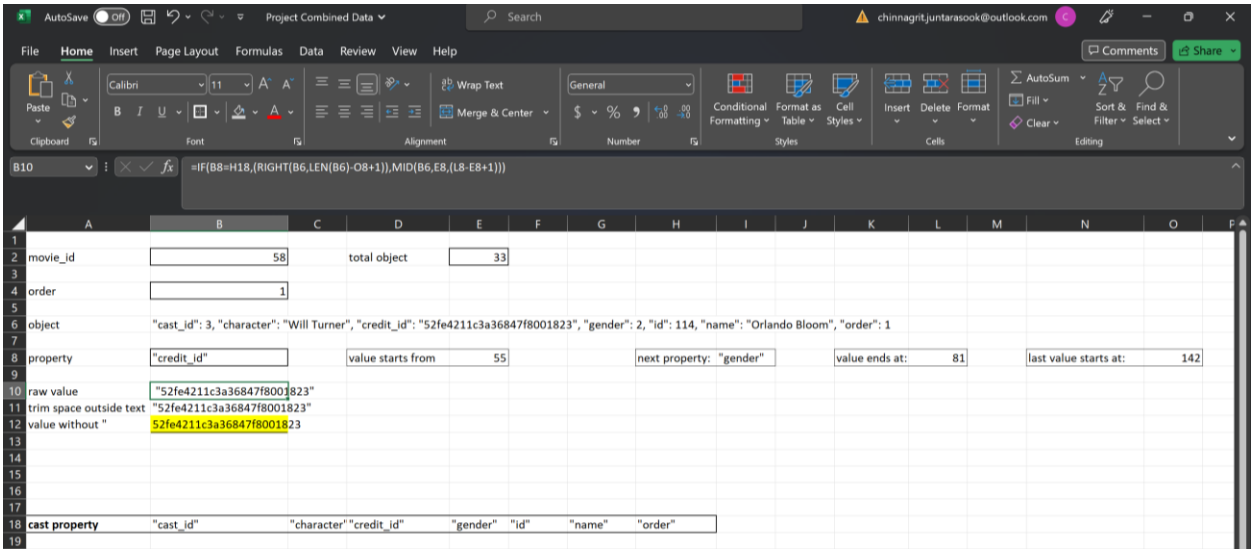


    b. Use Vlookup to get an object based on the given criteria, which are movie_id and order of object. The movie_id is validated using a list of id, and order is validated using the total object, so when an unidentified movie id is entered, the error message will appear

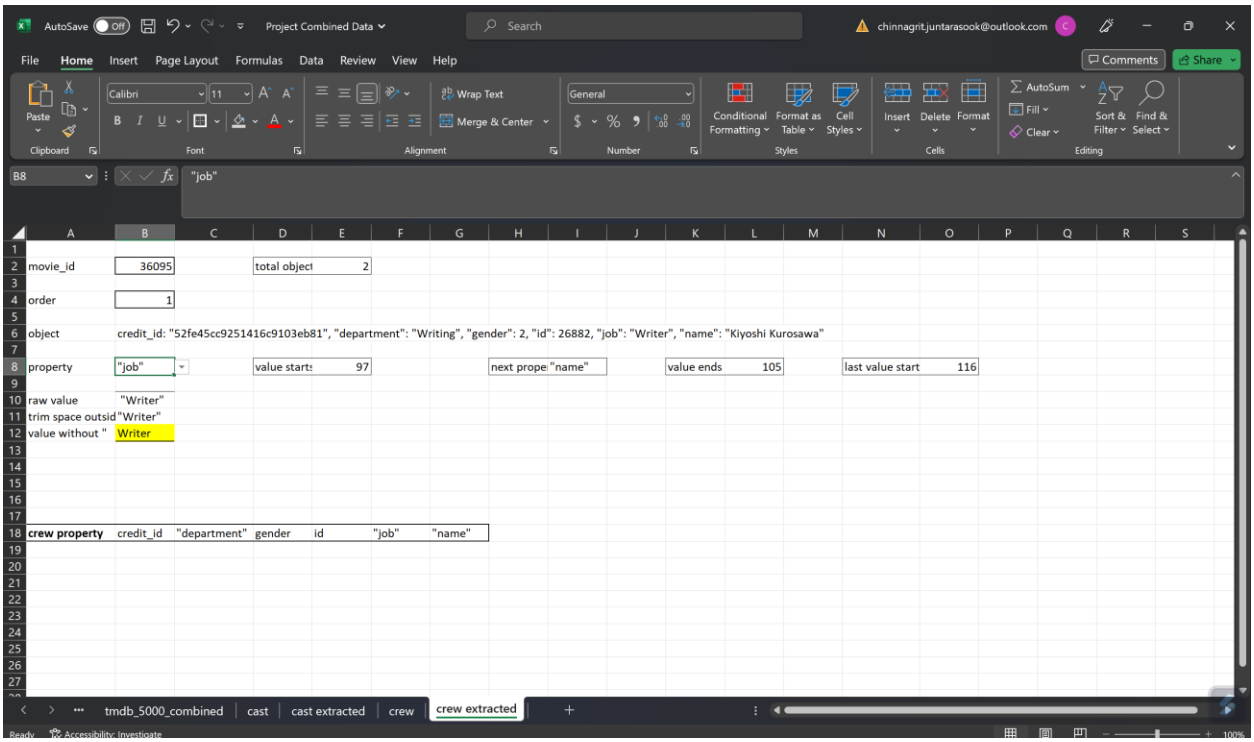




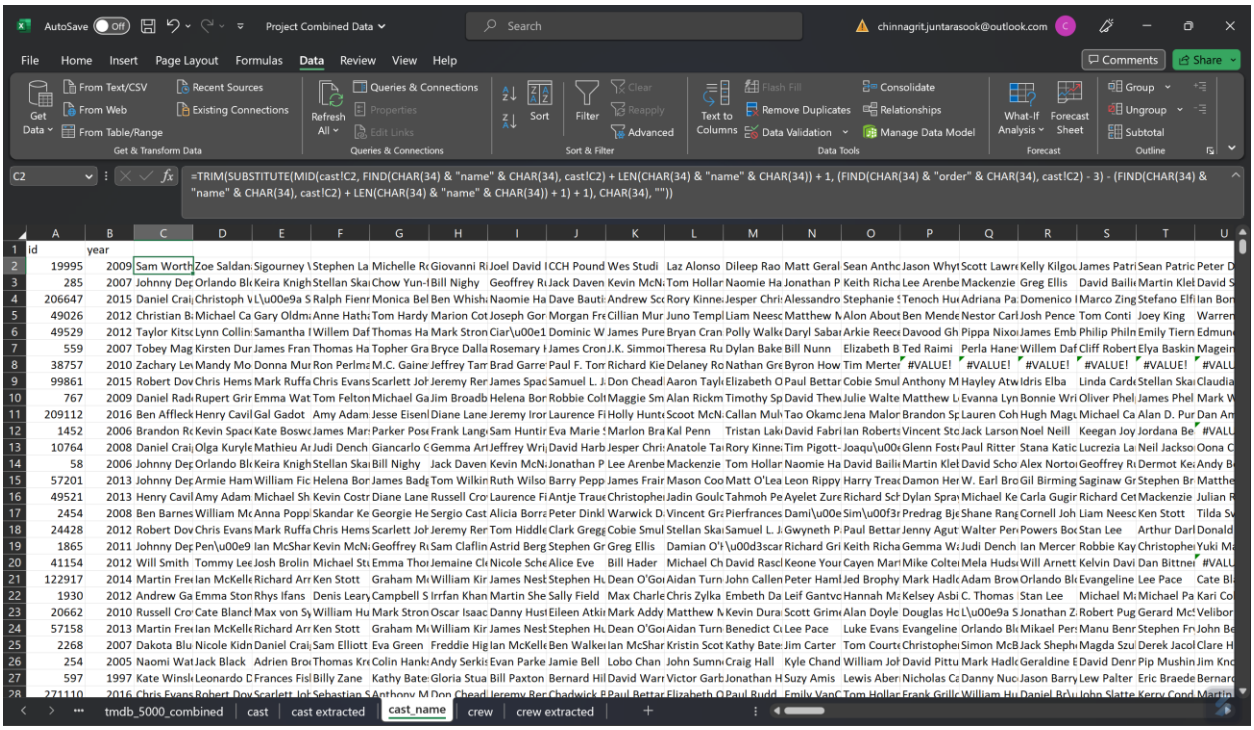
c. Next is that the raw values are extracted based on the given property, validated by the cast property list below with the formula shown. Then the extra space before and after the value is removed, and the quotation marks are removed if there are any.



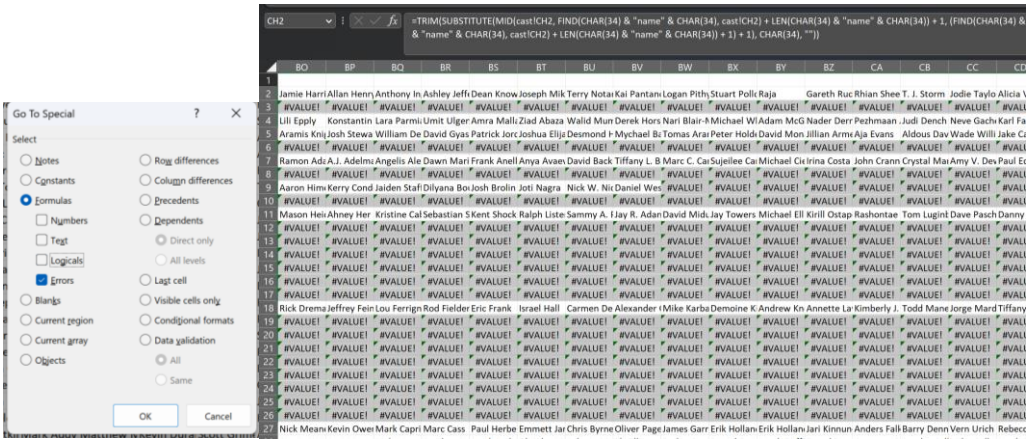
9. Repeat the 7th step for the crew object so the format is consistent, which is credit_id: "value", "department": "value", "gender": "value", "id": "value", "job": "value", "name": "value". Then repeat the 8th step.



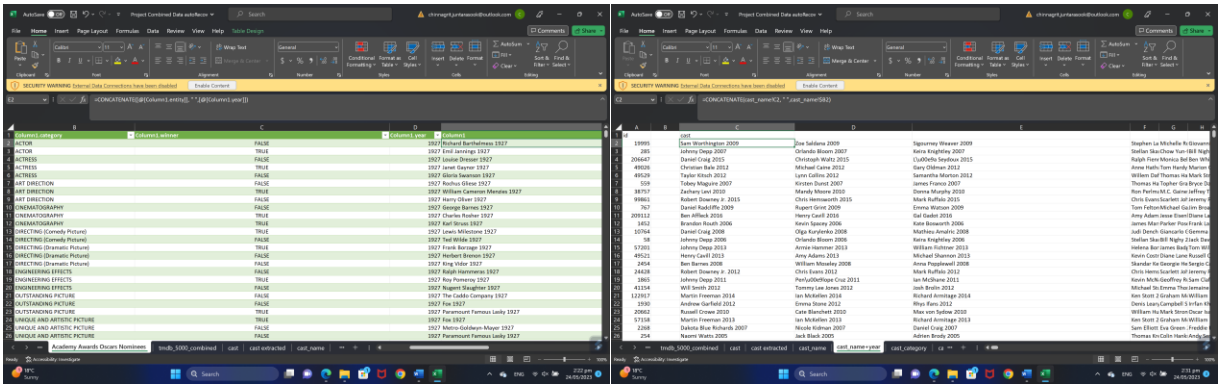
10. Nesting all the formulas done in the 8th step to create a new sheet specifically for storing all the cast names for each movie to prepare to match with the Oscars dataset



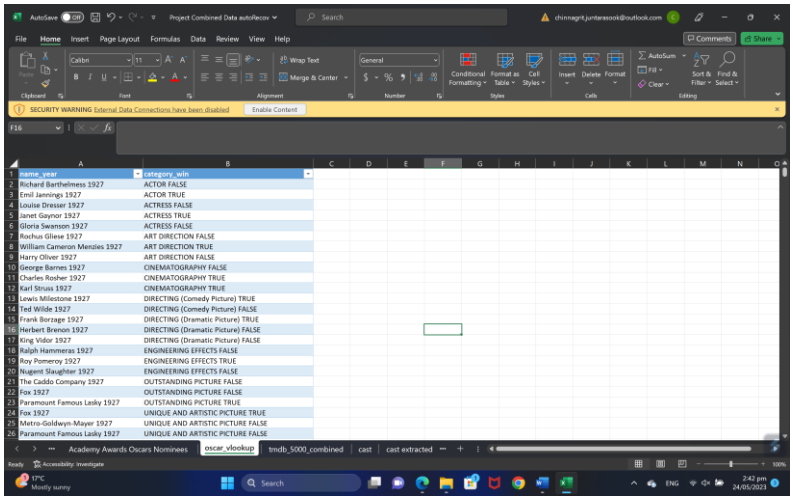
11. Delete all of the #VALUE! Result for the formula because there is no object in that cell, using the Go To Special function to select all of the error results and delete them



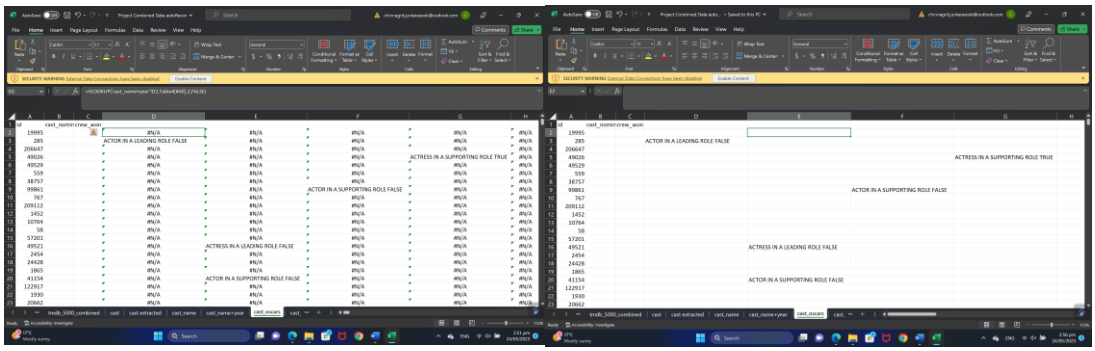
12. The next steps will start to combine the Oscars.JSON with all of the cast names. Unfortunately, the Oscars dataset does not contain information about what movies people were nominated for or won the Oscars belong to. So we need to use the year column to assume that if cast A won Oscars in 2000, the movie released in 2000 that has cast A will be the movie that has cast who won the Oscars. Therefore, for simple matching, we combine the name and year together using concatenate in one cell for both Oscars.JSON and the cast sheet.



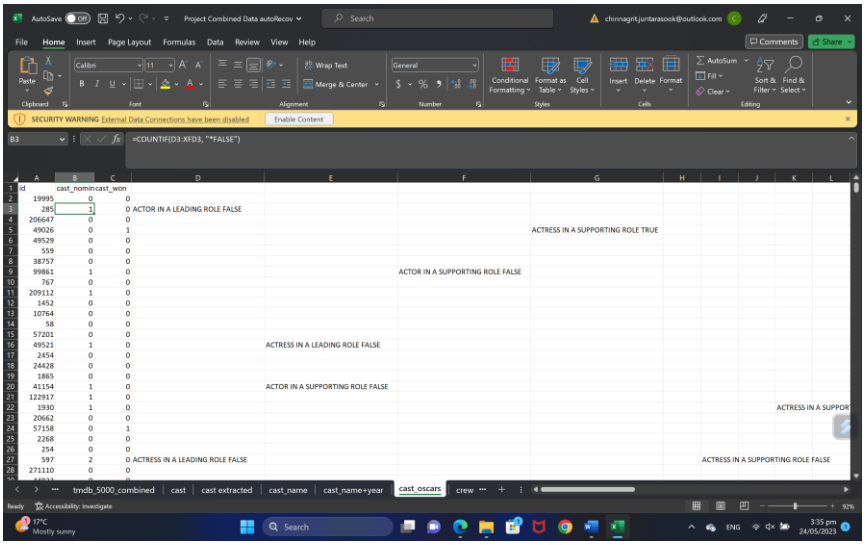
13. Create a new sheet specifically for matching, which contains some data merged together



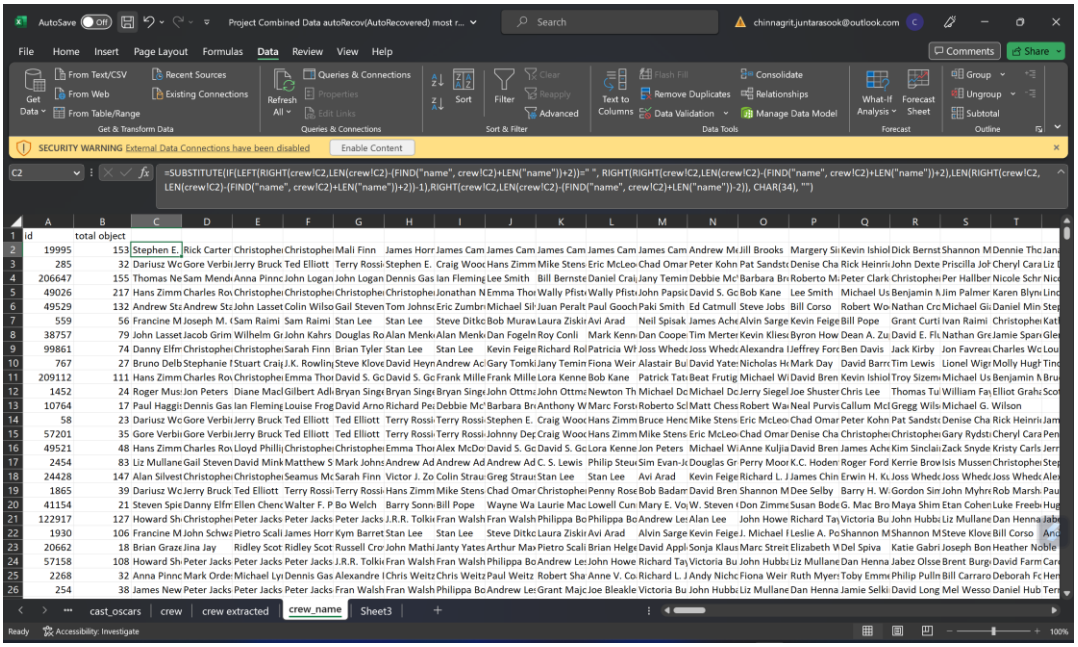
14. Use Vlookup with an exact match to find if the cast+ year matches with the data in Oscars.json. Then use go to special to delete all #N/A

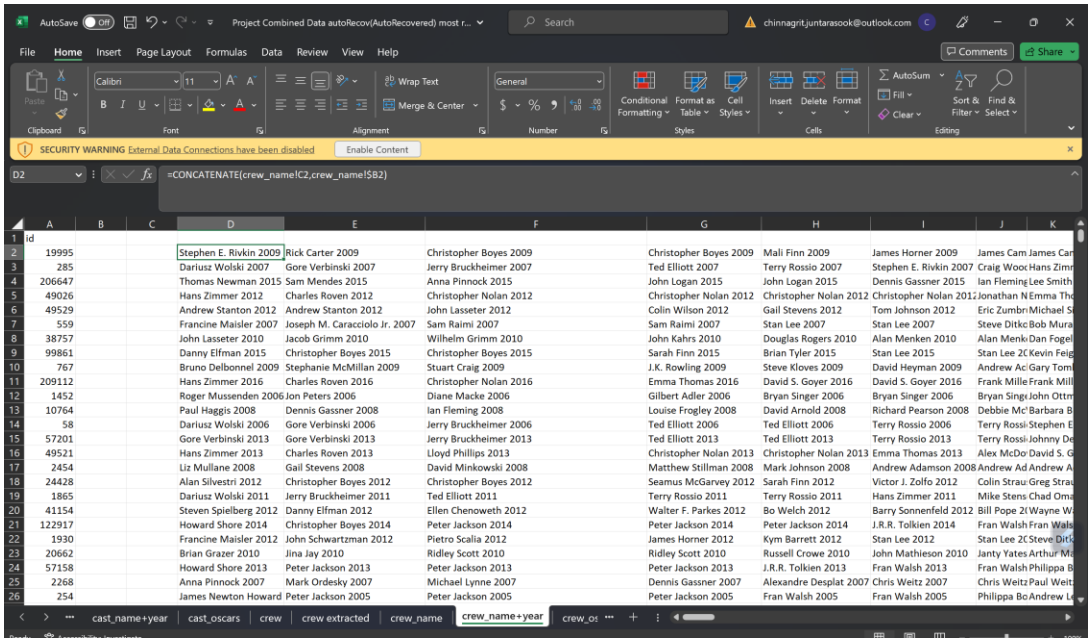


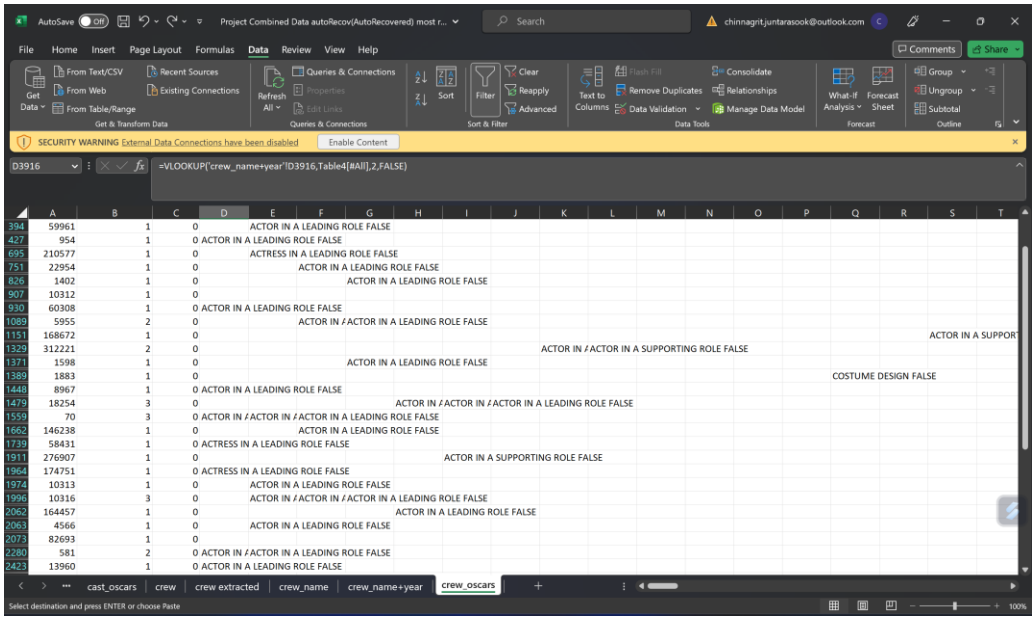
15. Count the total number of cast nominated and won for each movie by using countif, if it ends with a false count for nominated and ends with a true count for won



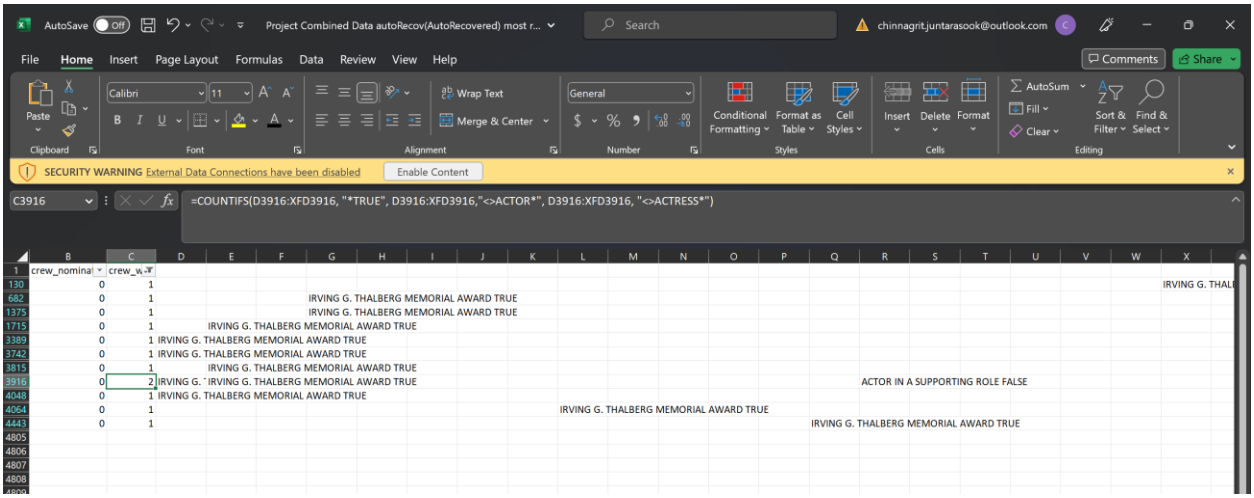
16. Replicating since step 10 for the crew. Firstly, extracting the name from the array of the object of the crew by nesting function, done in step 9. Then, concatenating crew name and year for matching. Then, use Vlookup to match the category and if win for crew

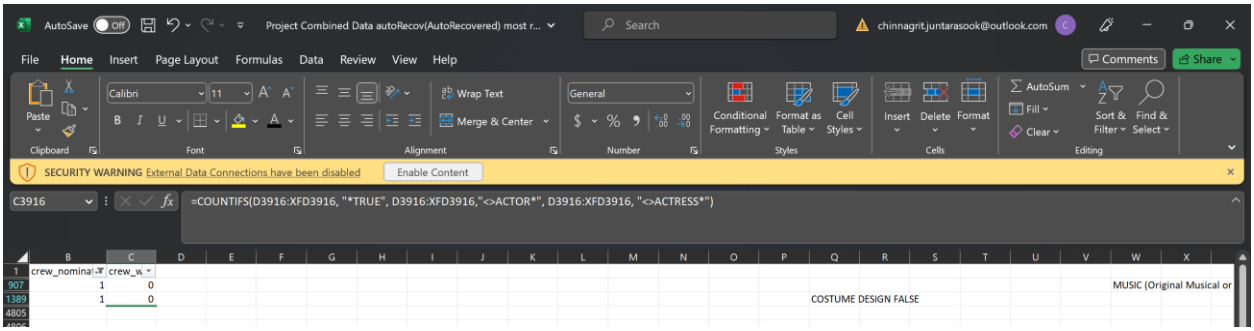




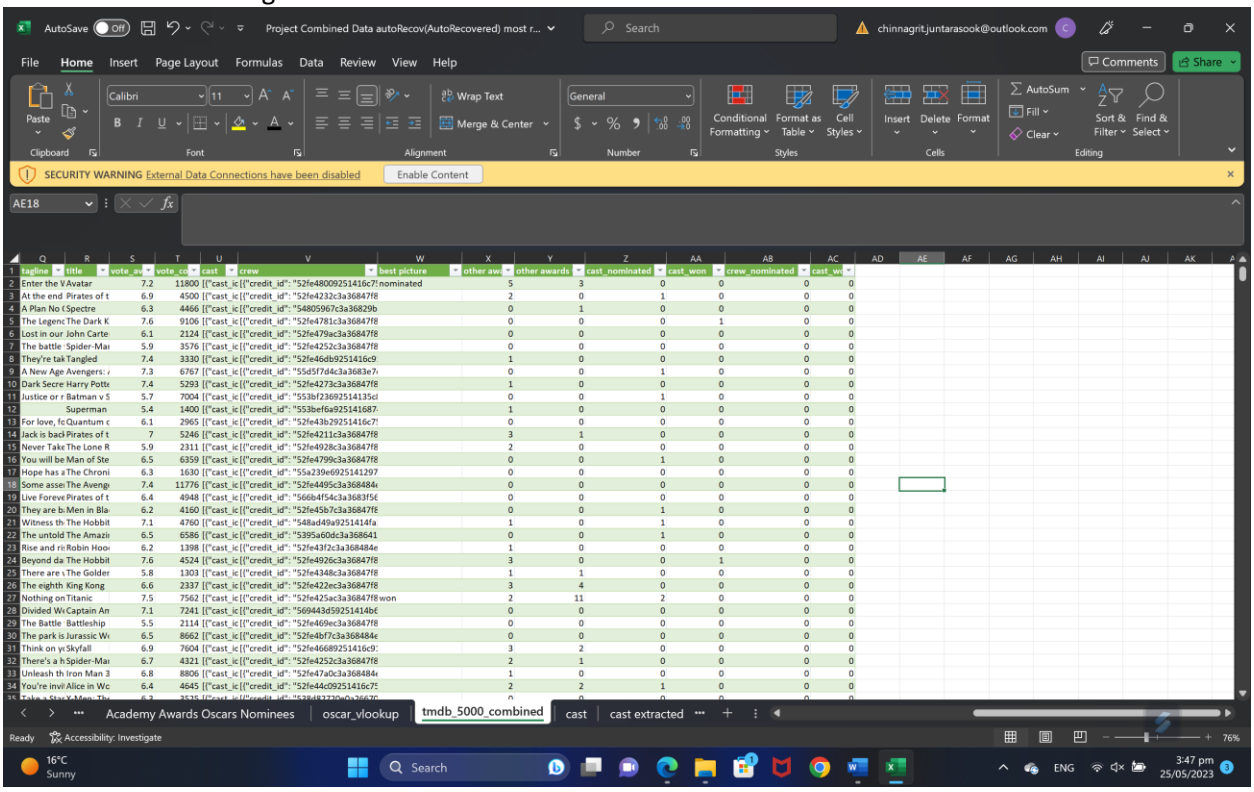


17. However, most of the categories are actors and actresses, which may suggest that some casts are also part of the crew; therefore, deletion of these cast categories is required to not skew the result by modifying countif function to not contain actress or actor





18. Then merge the total number of award cast and crew received with the combined sheet.



19. Copy the whole sheet into the new file as a final storage format to perform tasks for the next section

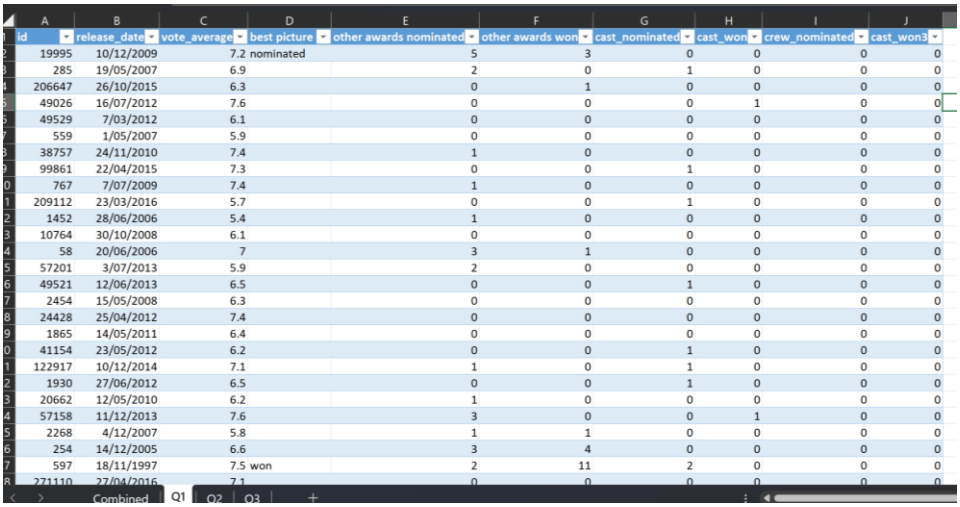
III. Question and Answer

1. Question:

Is there a discernible correlation between the attainment of Academy Awards (Oscars, movie awards only) victories or nominations and the attainment of higher overall ratings on TMDB from mainstream audiences for movies released between the years 1927 and 2017, considering that the movie data in both TMDB and Oscars datasets correspond to the year of release? (using year outside this range can skew result as Oscars datasets do not have a movie before 1927 and after 2017)

Steps performed to produce the answers:

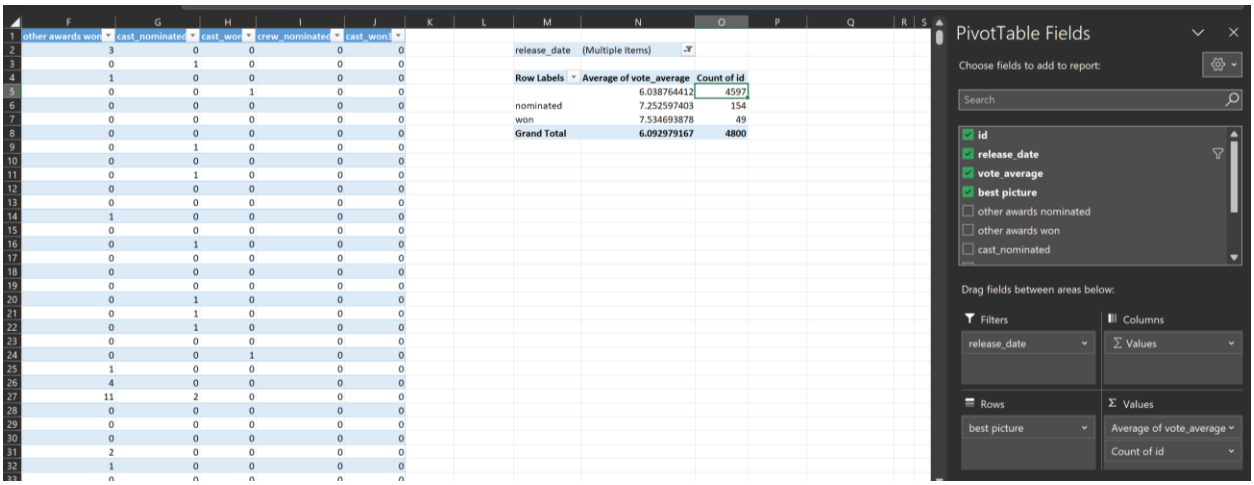
a. Only imported the needed column into a new sheet, including id, release date, vote_average, best Picture, other awards nominated, other awards won, cast_nominated, cast_won, crew_nominated, and crew_won.



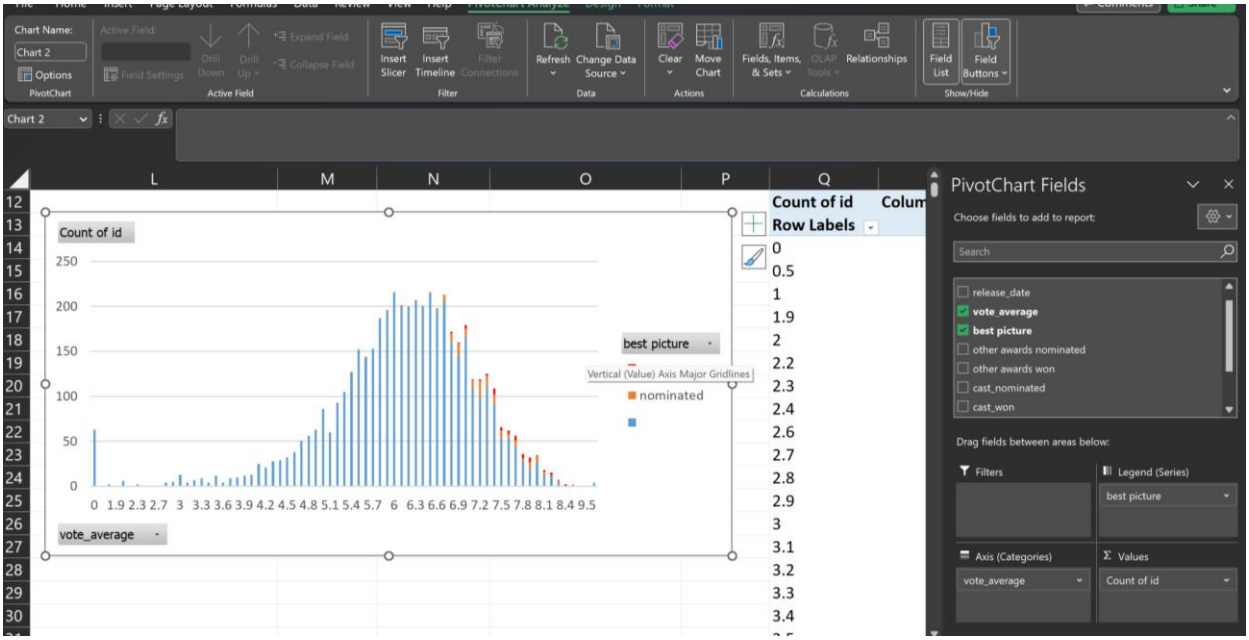
id	release_date	vote_average	best picture	other awards nominated	other awards won	cast_nominated	cast_won	crew_nominated	cast_won3
19995	10/12/2009	7.2	nominated	5	3	0	0	0	0
285	19/05/2007	6.9		2	0	1	0	0	0
206647	26/10/2015	6.3		0	1	0	0	0	0
49026	16/07/2012	7.6		0	0	0	1	0	0
49529	7/03/2012	6.1		0	0	0	0	0	0
559	1/05/2007	5.9		0	0	0	0	0	0
38757	24/11/2010	7.4		1	0	0	0	0	0
99861	22/04/2015	7.3		0	0	1	0	0	0
767	7/07/2009	7.4		1	0	0	0	0	0
209112	23/03/2016	5.7		0	0	1	0	0	0
1452	28/06/2006	5.4		1	0	0	0	0	0
10764	30/10/2008	6.1		0	0	0	0	0	0
58	20/06/2006	7		3	1	0	0	0	0
57201	3/07/2013	5.9		2	0	0	0	0	0
49521	12/06/2013	6.5		0	0	1	0	0	0
2454	15/05/2008	6.3		0	0	0	0	0	0
24428	25/04/2012	7.4		0	0	0	0	0	0
1865	14/05/2011	6.4		0	0	0	0	0	0
41154	23/05/2012	6.2		0	0	1	0	0	0
122917	10/12/2014	7.1		1	0	1	0	0	0
1930	27/06/2012	6.5		0	0	1	0	0	0
20662	12/05/2010	6.2		1	0	0	0	0	0
57158	11/12/2013	7.6		3	0	0	1	0	0
2268	4/12/2007	5.8		1	1	0	0	0	0
254	14/12/2005	6.6		3	4	0	0	0	0
597	18/11/1997	7.5	won	2	11	2	0	0	0
271110	27/04/2016	7.1		0	0	0	0	0	0

Combined | Q1 | Q2 | Q3

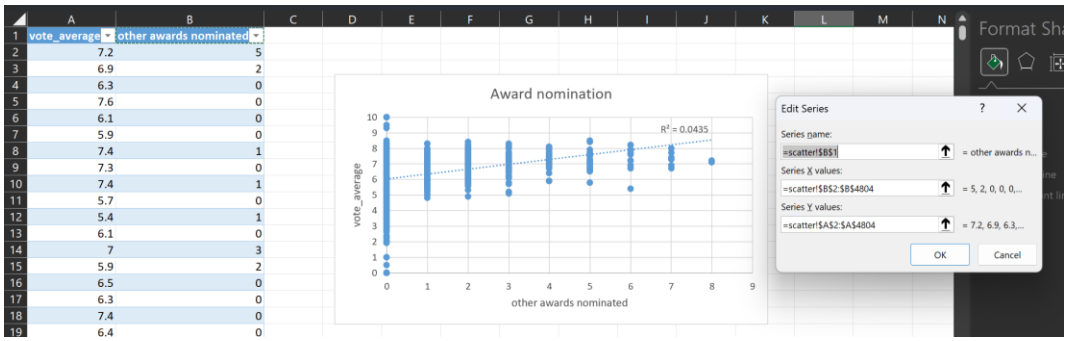
b. Create a pivot table that filters out movies outside the 1927 and 2017 range, and calculate average ratings for movies that have not been nominated or won the best Picture, movies that are nominated and movies that won the Oscars.

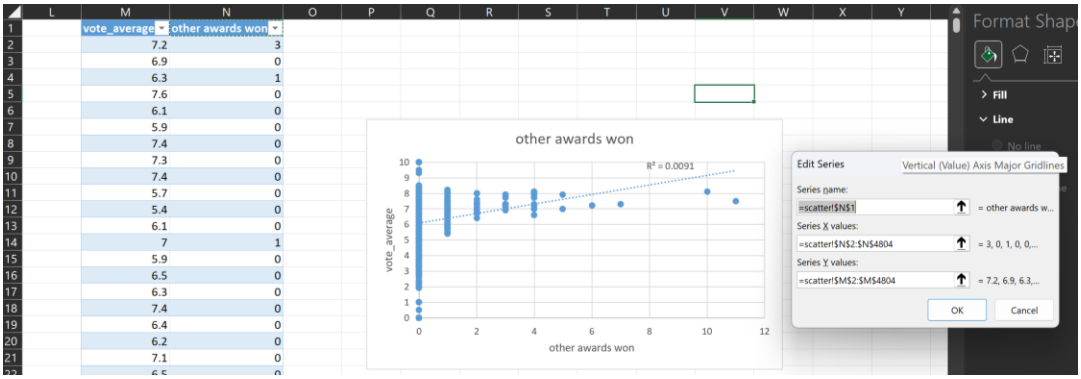


c. Create a pivot chart that shows every rating movie received based on the best picture category.



d. Create two scatter plots to find the correlation between the vote average and the number of awards won and nominated. In the scatter plot, y-axis is the average rating, and x-axis is the number of awards nominated and awards won, respectively,





Answer:

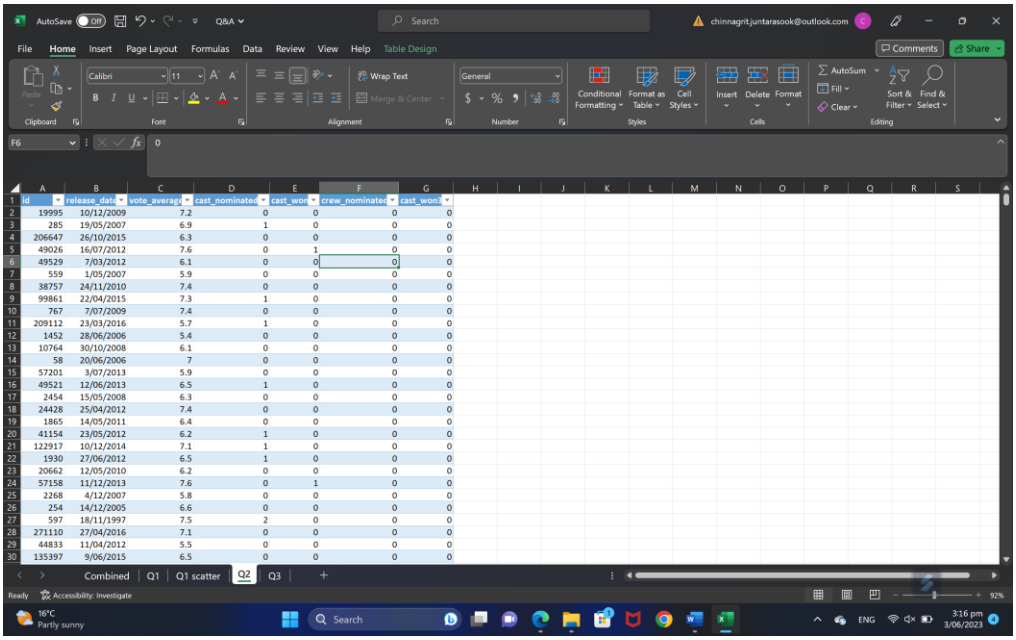
Based on the available data analysis, there appears to be a positive correlation between the number of Oscars won or nominated by movies and the user ratings they receive on TMDB. Specifically, the data indicate that movies with a higher number of Oscar wins tend to receive higher scores on TMDB compared to movies with fewer wins. Furthermore, on average, movies that have won a certain number of Oscars tend to have higher TMDB scores than those that have received Oscar nominations. For example, the pivot table in b indicates that a movie that has not won any Oscars and was only nominated for Best Picture has an average rating of 6.04. In contrast, movies that were nominated for Best Picture have an average rating of 7.25, while movies that actually won the Best Picture award have an average rating of 7.53. Additionally, this trend is also observed in other award categories. Both the scatter plot of other award nominations and TMDB scores, as well as the scatter plot of other award wins and TMDB scores, show a weak positive relationship. However, it's worth noting that the correlation between nominations and TMDB scores is stronger, but this may be due to the small size of the data.

2. Question:

To what extent does the recognition of cast and crew with Oscar awards in the same year as the release of their respective movies influence the audience score on TMDB, for movies released between the years 1927 and 2017?

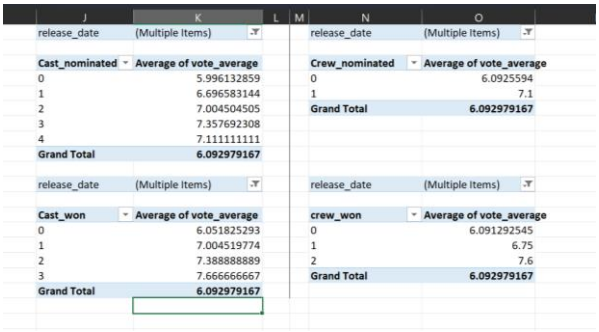
Steps performed to produce the answers:

a. Copy the columns that are required for this question: id, release_date, vote_average, cast_nominated, cast_won, crew_nominated, crew_won



id	release_date	vote_average	cast_nominated	cast_won	crew_nominated	cast_won2
19995	10/12/2009	7.2	0	0	0	0
285	19/05/2007	6.9	1	0	0	0
206647	26/10/2015	6.3	0	0	0	0
49026	16/07/2012	7.6	0	1	0	0
49529	7/03/2012	6.1	0	0	0	0
559	1/05/2007	5.9	0	0	0	0
38757	24/11/2010	7.4	0	0	0	0
99861	22/04/2015	7.3	1	0	0	0
767	7/07/2009	7.4	0	0	0	0
209112	23/03/2016	5.7	1	0	0	0
1452	28/06/2006	5.4	0	0	0	0
10764	30/10/2008	6.1	0	0	0	0
58	20/06/2006	7	0	0	0	0
57201	3/07/2013	5.9	0	0	0	0
49521	12/06/2013	6.5	1	0	0	0
2454	15/05/2008	6.3	0	0	0	0
24428	25/04/2012	7.4	0	0	0	0
1865	14/05/2011	6.4	0	0	0	0
41154	23/05/2012	6.2	1	0	0	0
122917	10/12/2014	7.1	1	0	0	0
1930	27/06/2012	6.5	1	0	0	0
20662	12/05/2010	6.2	0	0	0	0
57158	11/12/2013	7.6	0	1	0	0
2268	4/12/2007	5.8	0	0	0	0
254	14/12/2005	6.6	0	0	0	0
597	18/11/1997	7.5	2	0	0	0
271110	27/04/2016	7.1	0	0	0	0
44833	11/04/2012	5.5	0	0	0	0
135397	9/06/2015	6.5	0	0	0	0

b. Create the pivot table that shows an average of vote_average for every number of cast and crew nominated and won an Oscars award



release_date	(Multiple Items)
Cast_nominated	Average of vote_average
0	5.996132859
1	6.696583144
2	7.004504505
3	7.357692308
4	7.111111111
Grand Total	6.092979167

release_date	(Multiple Items)
Crew_nominated	Average of vote_average
0	6.0925594
1	7.1
Grand Total	6.092979167

release_date	(Multiple Items)
Cast_won	Average of vote_average
0	6.051825293
1	7.004519774
2	7.388888889
3	7.666666667
Grand Total	6.092979167

release_date	(Multiple Items)
crew_won	Average of vote_average
0	6.091292545
1	6.75
2	7.6
Grand Total	6.092979167

c. For a visualisation, copy the result from the pivot table in a separate row and create a line chart based on the information.

Answer:

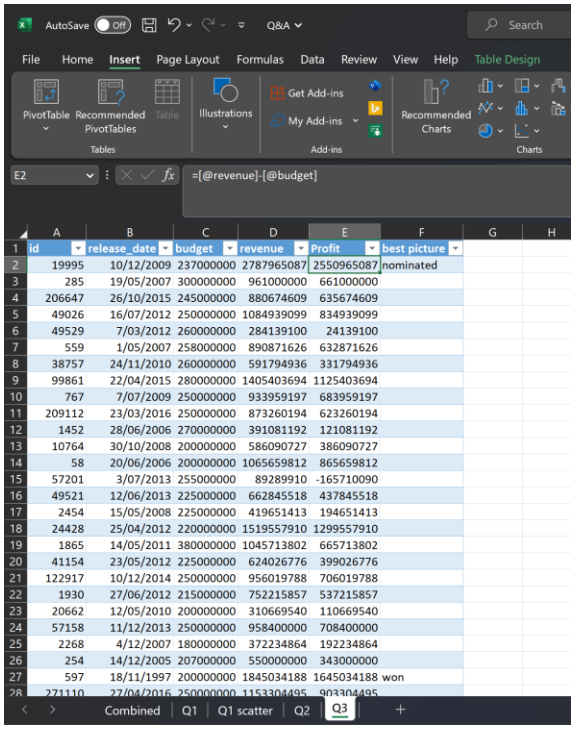
The data suggests that there is a clear connection between the number of cast and crew members who receive Oscar nominations or awards in the same year as a movie's release and the audience's score on TMDB. This conclusion is supported by evidence showing the average scores for different levels of cast and crew recognition at the Oscars. Therefore, we can confidently say that there is a correlation between the presence of Oscar-winning cast and crew members in a film's release year and the audience's score on TMDB.

3. Question:

   To what extent can the Oscars best picture awards be considered indicative of a film's financial success for those released between the years 1927 and 2017?

   Steps performed to produce the answers:

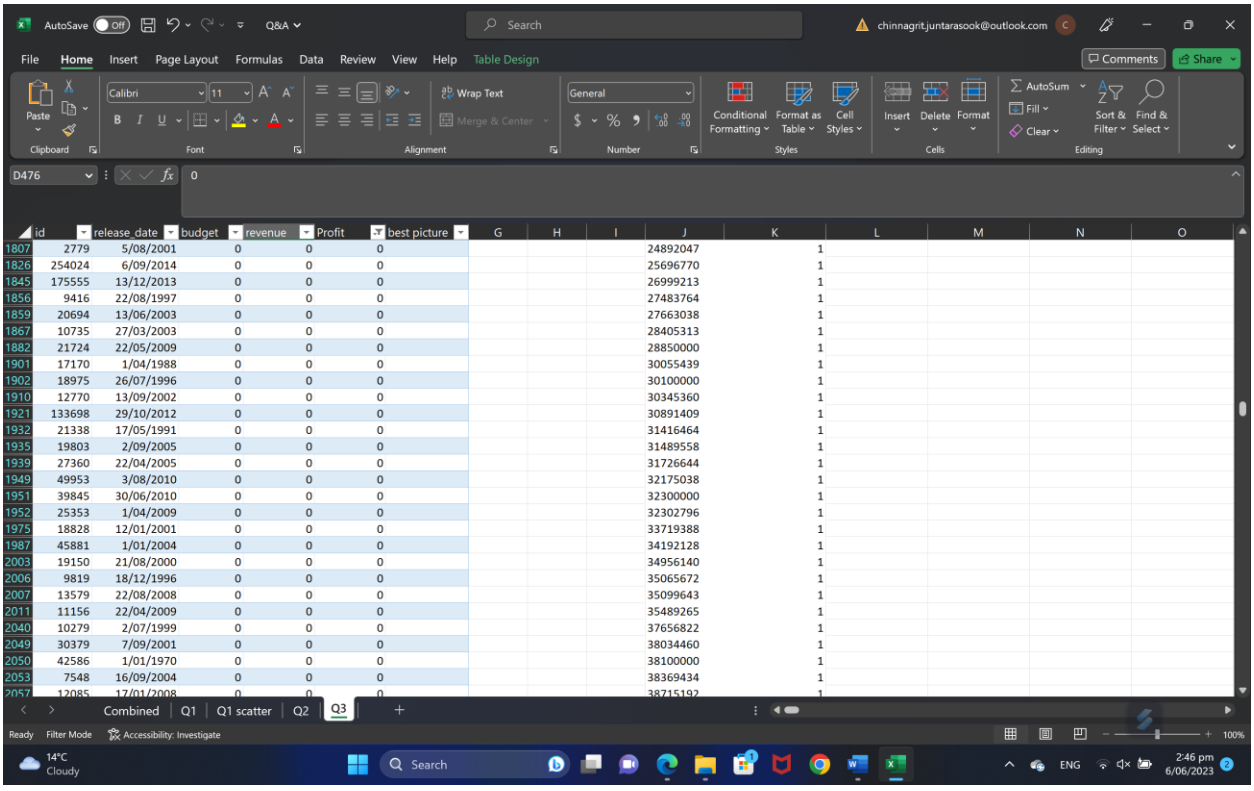
   a. Copy only the columns needed, which are id, release_date, budget, revenue, and best Picture, then Calculate the profit by assuming that it is revenue minus budget



	A	B	C	D	E	F	G	H
1	id	release_date	budget	revenue	Profit	best picture		
2	19995	10/12/2009	237000000	2787965087	2550965087	nominated		
3	285	19/05/2007	300000000	961000000	661000000			
4	206647	26/10/2015	245000000	880674609	635674609			
5	49026	16/07/2012	250000000	1084939099	834939099			
6	49529	7/03/2012	260000000	284139100	24139100			
7	559	1/05/2007	258000000	890871626	632871626			
8	38757	24/11/2010	260000000	591794936	331794936			
9	99861	22/04/2015	280000000	1405403694	1125403694			
10	767	7/07/2009	250000000	933959197	683959197			
11	209112	23/03/2016	250000000	873260194	623260194			
12	1452	28/06/2006	270000000	391081192	121081192			
13	10764	30/10/2008	200000000	586090727	386090727			
14	58	20/06/2006	200000000	1065659812	865659812			
15	57201	3/07/2013	255000000	89289910	-165710090			
16	49521	12/06/2013	225000000	662845518	437845518			
17	2454	15/05/2008	225000000	419651413	194651413			
18	24428	25/04/2012	220000000	1519557910	1299557910			
19	1865	14/05/2011	380000000	1045713802	665713802			
20	41154	23/05/2012	225000000	624026776	399026776			
21	122917	10/12/2014	250000000	956019788	706019788			
22	1930	27/06/2012	215000000	752215857	537215857			
23	20662	12/05/2010	200000000	310669540	110669540			
24	57158	11/12/2013	250000000	958400000	708400000			
25	2268	4/12/2007	180000000	372234864	192234864			
26	254	14/12/2005	207000000	550000000	343000000			
27	597	18/11/1997	200000000	1845034188	1645034188	won		
28	271110	27/04/2016	250000000	1153304495	903304495			

Combined | Q1 | Q1 scatter | Q2 | Q3

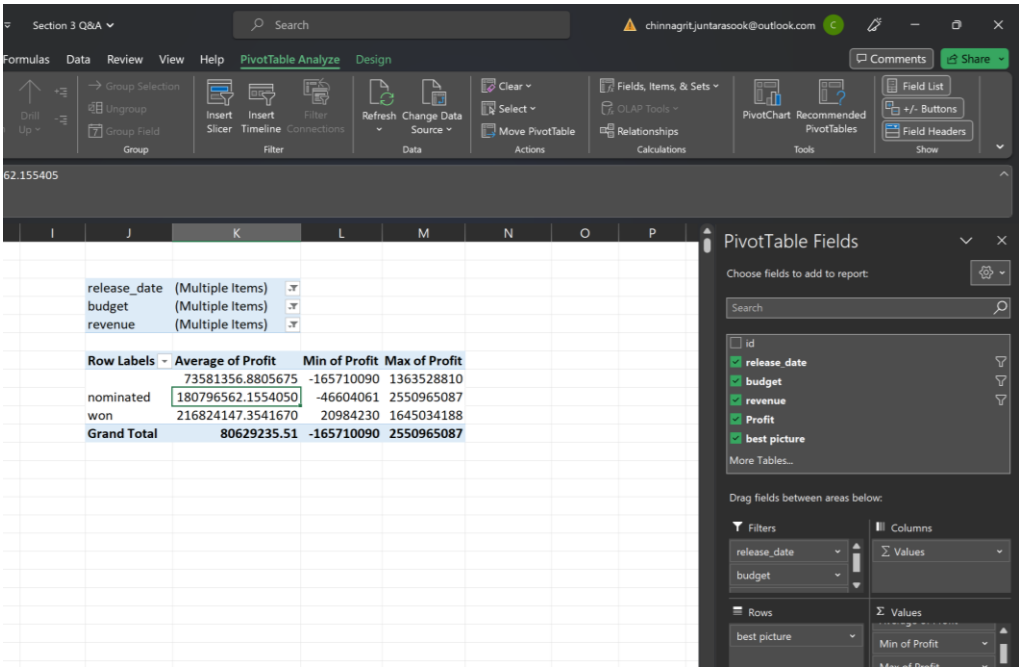
b. While creating a pivot table, it is found that 891 movies have 0 profit, most of which results from having both 0 budget and revenue; thus, excluding this movie is necessary not to skew the result.



	id	release_date	budget	revenue	Profit	best picture	G	H	I	J	K	L	M	N	O
1807	2779	5/08/2001	0	0	0					24892047	1				
1826	254024	6/09/2014	0	0	0					25696770	1				
1845	175555	13/12/2013	0	0	0					26999213	1				
1856	9416	22/08/1997	0	0	0					27483764	1				
1859	20694	13/06/2003	0	0	0					27663038	1				
1867	10735	27/03/2003	0	0	0					28405313	1				
1882	21724	22/05/2009	0	0	0					28850000	1				
1901	17170	1/04/1988	0	0	0					30055439	1				
1902	18975	26/07/1996	0	0	0					30100000	1				
1910	12770	13/09/2002	0	0	0					30345360	1				
1921	133698	29/10/2012	0	0	0					30891409	1				
1932	21338	17/05/1991	0	0	0					31416464	1				
1935	19803	2/09/2005	0	0	0					31489558	1				
1939	27360	22/04/2005	0	0	0					31726644	1				
1949	49953	3/08/2010	0	0	0					32175038	1				
1951	39845	30/06/2010	0	0	0					32300000	1				
1952	25353	1/04/2009	0	0	0					32302796	1				
1975	18828	12/01/2001	0	0	0					33719388	1				
1987	45881	1/01/2004	0	0	0					34192128	1				
2003	19150	21/08/2000	0	0	0					34956140	1				
2006	9819	18/12/1996	0	0	0					35065672	1				
2007	13579	22/08/2008	0	0	0					35099643	1				
2011	11156	22/04/2009	0	0	0					35489265	1				
2040	10279	2/07/1999	0	0	0					37656822	1				
2049	30379	7/09/2001	0	0	0					38034460	1				
2050	42586	1/01/1970	0	0	0					38100000	1				
2053	7548	16/09/2004	0	0	0					38369434	1				
2057	12085	17/01/2008	0	0	0					38715192	1				

Combined | Q1 | Q1 scatter | Q2 | Q3

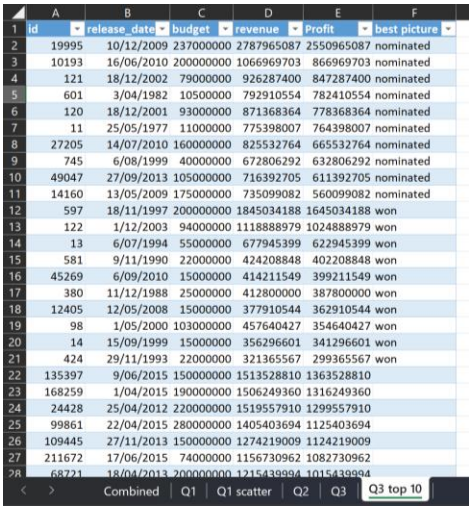
c. Create a pivot table that filters out movies outside the year range, 0 budget and revenue, which indicate the average, max and min profit for each best picture type movie has achieved.



release_date	(Multiple Items)
budget	(Multiple Items)
revenue	(Multiple Items)

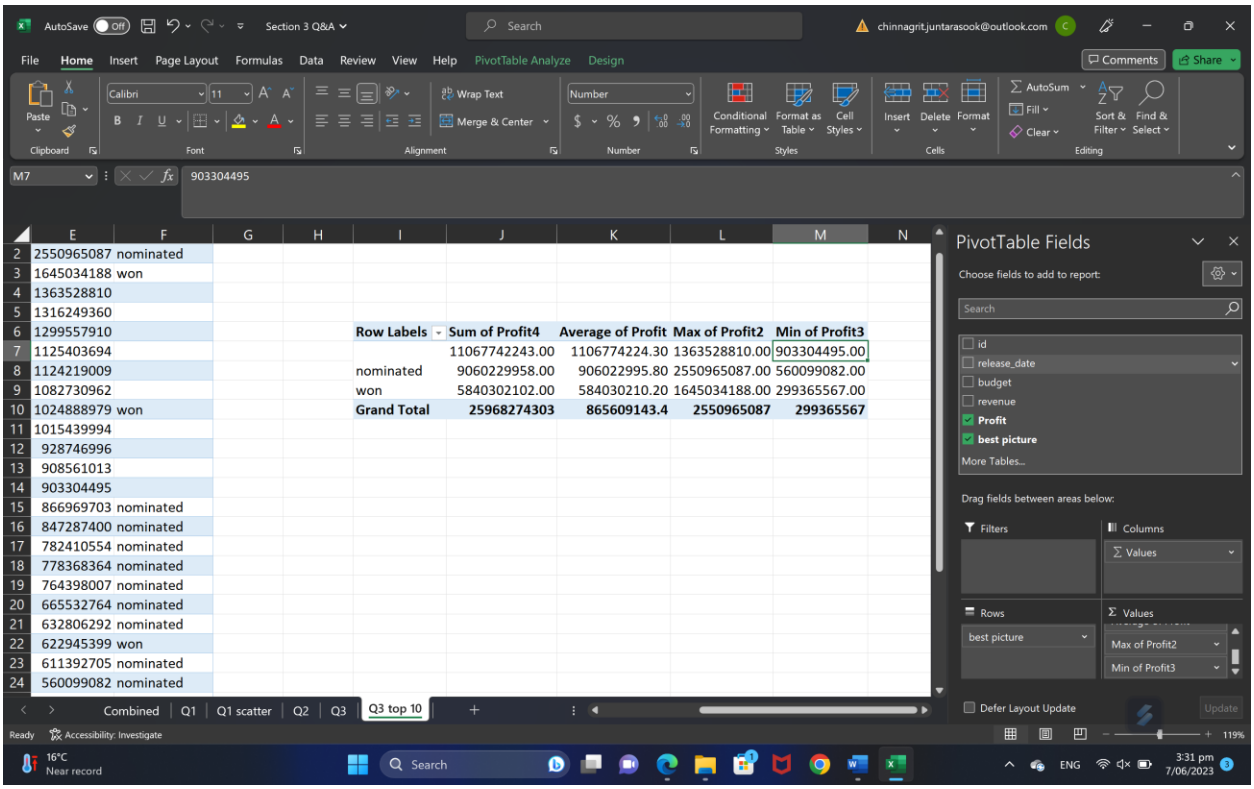
Row Labels	Average of Profit	Min of Profit	Max of Profit
	73581356.8805675	-165710090	1363528810
nominated	180796562.1554050	-46604061	2550965087
won	216824147.3541670	20984230	1645034188
Grand Total	**80629235.51**	**-165710090**	**2550965087**

d. The pivot table in (c.) may not be enough to conclude that Oscar's best Picture can be a financial success indicator for films. Therefore, it is crucial to consider if the best movies not in the nominated or won category are not as financially successful as those in. Firstly, get a list of the top 10 highest-profit movies for each best picture category using the filter function and copy the value into a new sheet.



id	release_date	budget	revenue	Profit	best picture
19995	10/12/2009	237000000	2787965087	2550965087	nominated
10193	16/06/2010	200000000	1066969703	866969703	nominated
121	18/12/2002	79000000	926287400	847287400	nominated
601	3/04/1982	10500000	792910554	782410554	nominated
120	18/12/2001	93000000	871368364	778368364	nominated
11	25/05/1977	11000000	775398007	764398007	nominated
27205	14/07/2010	160000000	825532764	665532764	nominated
745	6/08/1999	40000000	672806292	632806292	nominated
49047	27/09/2013	105000000	716392705	611392705	nominated
14160	13/05/2009	175000000	735099082	560099082	nominated
597	18/11/1997	200000000	1845034188	1645034188	won
122	1/12/2003	94000000	1118888979	1024888979	won
13	6/07/1994	55000000	677945399	622945399	won
581	9/11/1990	22000000	424208848	402208848	won
45269	6/09/2010	15000000	414211549	399211549	won
380	11/12/1988	25000000	412800000	387800000	won
12405	12/05/2008	15000000	377910544	362910544	won
98	1/05/2000	103000000	457640427	354640427	won
14	15/09/1999	15000000	356296601	341296601	won
424	29/11/1993	22000000	321365567	299365567	won
135397	9/06/2015	150000000	1513528810	1363528810	
168259	1/04/2015	190000000	1506249360	1316249360	
24428	25/04/2012	220000000	1519557910	1299557910	
99861	22/04/2015	280000000	1405403694	1125403694	
109445	27/11/2013	150000000	1274219009	1124219009	
211672	17/06/2015	74000000	1156730962	1082730962	
68721	18/04/2013	200000000	1215439994	1015439994	

e. Create a pivot table the same way as (c.)



Answer:

The data in (c.) shows that the average profit for movies nominated or won Oscars Best Picture awards is significantly higher than those not. However, the data in (d.) suggests that the average profit from the top 10 highest-profit films in won and nominated categories is less than movies that are not in these categories, which indicates that films that were not nominated or won Oscar Best Picture could also be financially successful and even more than those that achieved the award. To conclude, the Oscar Best Picture cannot be a movie's financial success indicator.

IV. Individual Reflection

A situation when you applied what you had already learned in Data Management to reason, think, and decide

(I did not do the Data Management paper as I took Database Systems as a prerequisite for this course.)

I use some of the concepts about ERD relationships. For example, when I combined TMDB movies with the Oscars, I initially used Vlookup to return the type of awards films achieved. After reviewing the result, I found that the movie I'm sure won Best Picture returned other awards instead. So I re-audited the datasets and saw that the TMDB dataset has an optional one-to-many relationship with the Oscar dataset, where a movie can have zero to many Oscars. Then I had to change the method. I combined them to count how many Awards each film achieved.

A situation when you applied self-management strategies to complete the project tasks.

The first thing I did for this assignment was planned what section to be done first, which is Section II → III → I → IV. Then I chose to prioritise what section based on the mark allocated, which in this project, I spent most of my time doing Section II. Then, I set goals for starting and completing each task based on my timetable, workload of each section and assignment due dates from other courses to not fall behind or be too late to change something that I need to inform the teaching staff.

A situation when you applied resilience strategies to complete the project tasks.

One task that I found too complicated is writing an Excel function that extracts just the name of cast&crew from the object where a cell can only contain one name. Therefore, I overcame the challenge by, instead of writing from scratch, I broke down and simplified the task by doing a function one by one (8^{th} and 9^{th} steps), such as finding when the value starts and ends and how long it is, then nested all the functions together in one cell so that I can apply it to other cells without do it one by one for every cell.