Prepared by :

Abhishek Arya

NET ID :axa220149

HOMEWORK 1

Descriptive/Diagnostic Analysis Case: Chilean College Admissions.

Introduction

This project focuses on analyzing the Chilean college admissions dataset and exploring various factors related to demographics, income levels, education, and test scores. The analysis involves data cleaning, transformation, and visualization to derive meaningful insights.

The Column Descriptions of the CSV:

•MRUN: Application number.

•DOB: date of birth.

•Gender: Male or Female

•Family size: number of people that live with the student.

•How many in your family work: number of people that work within the family.

•Income level: categorical variable representing monthly income, as detailed in the next

Table:

Category	From	To
1	0	350
2	350	1045
3	1045	1750
4	1750	2450
5	2450	3125
6	3125	more

•Level education father/mother: categorical variable representing the maximum level of education attained by the parents, as detailed in the next table:

Category	Highest Level of Education
0	No data
1	No education
2	Elementary incomplete
3	Elementary complete
4	Hish-school incomplete
5	High-school complete
6	Technical school incomplete
7	Technical school complete
8	University incomplete
9	University complete
10	Other

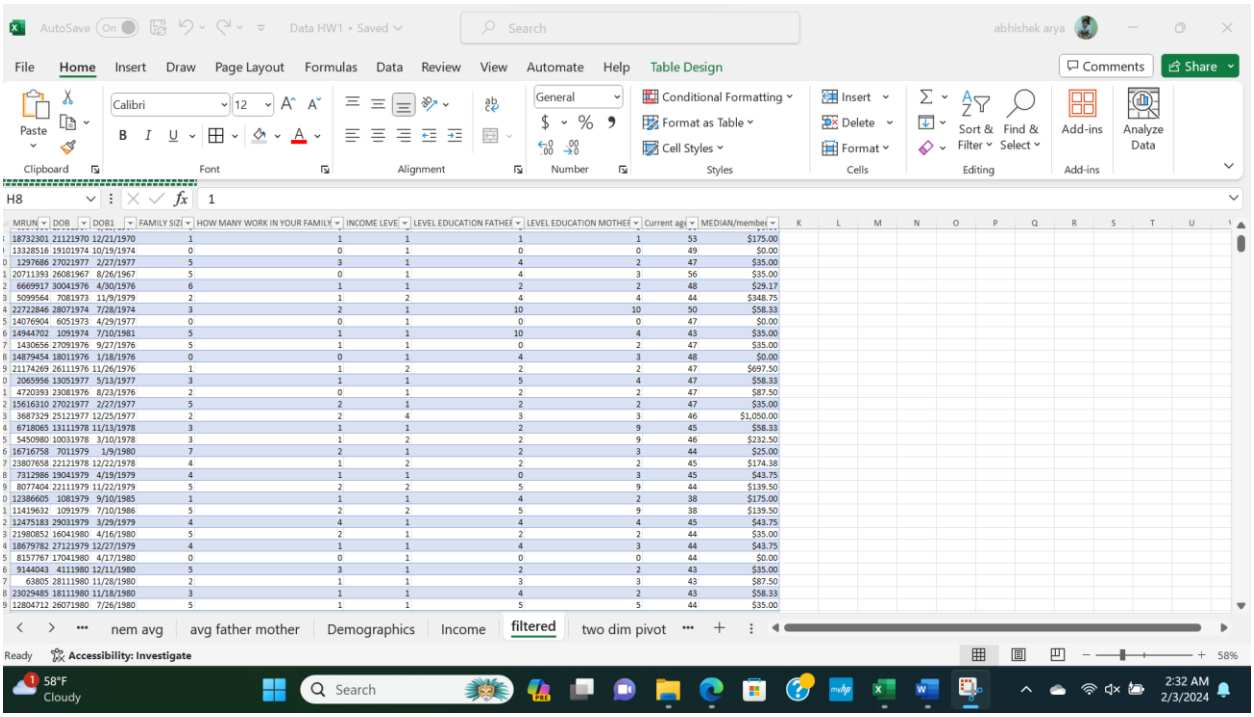
•High School Type: The type of Highschool student is from i.e. 2-Public, 3-voucher, 4-Private.

•Scores: NEM, LYC, MATE, HYCS, CIEN scores of applicants.

These are the steps that were followed to conduct the analysis and derive the insights:

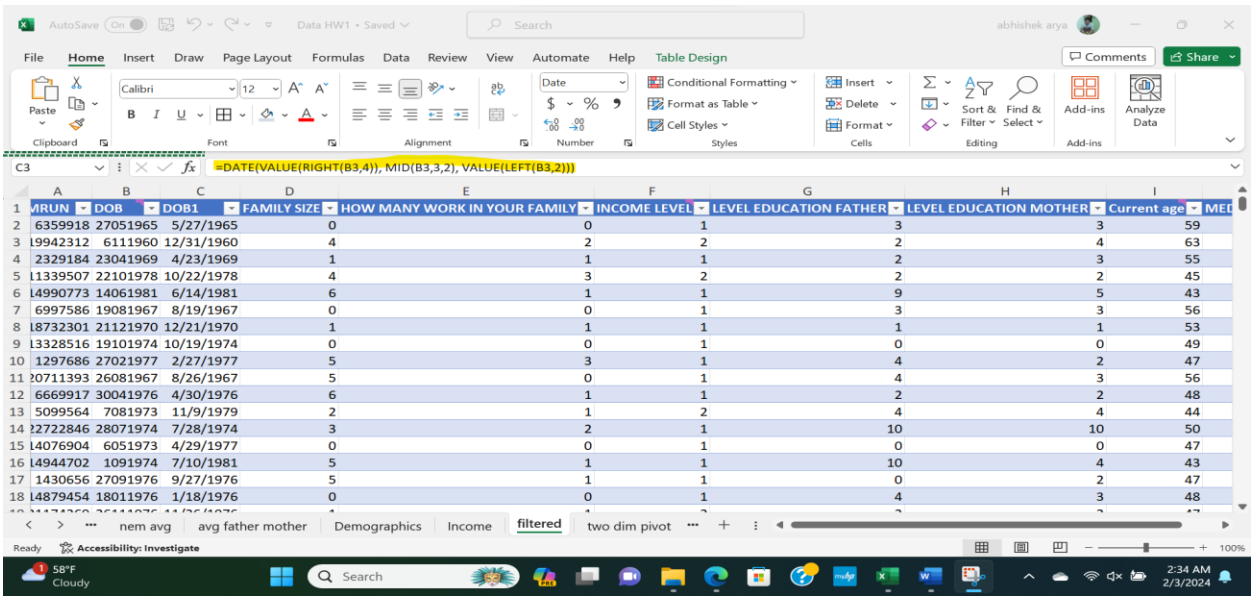
For question 1 , 2 and 3 I did the following processes:

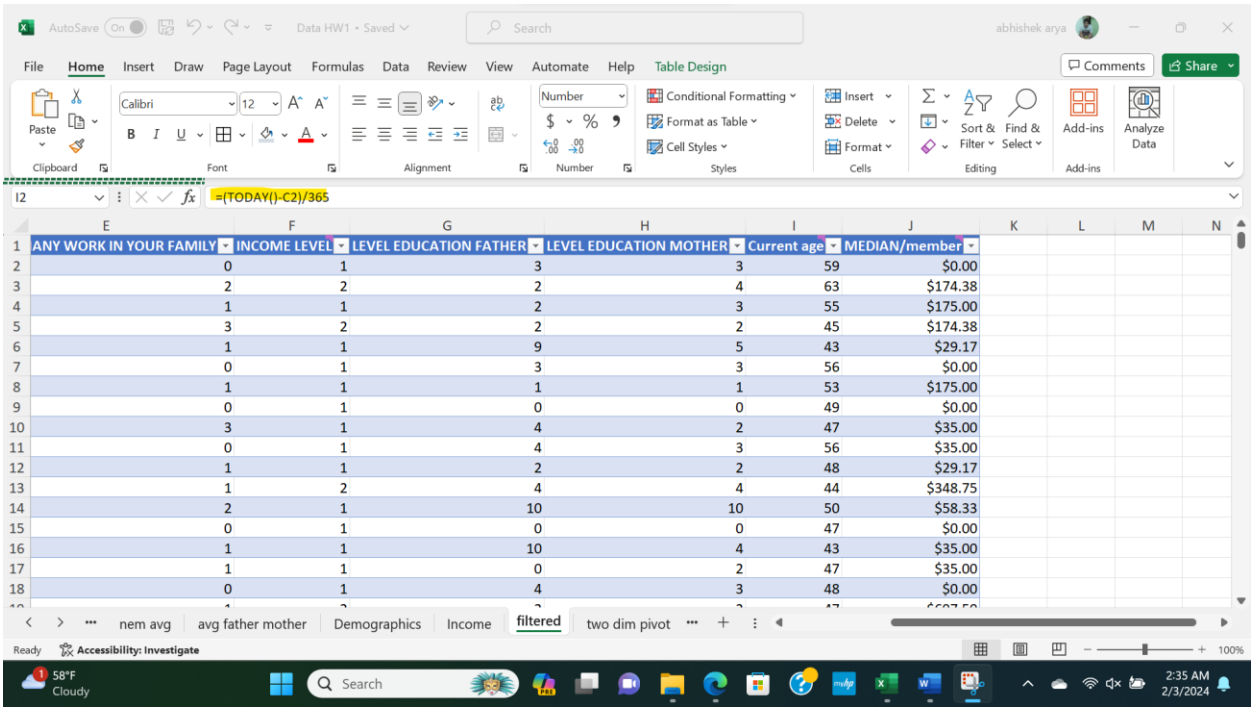
Separated Demographics data into columns using text-to-columns with a delimiter. Converted the data into a table for better organization and analysis. Filtered irrelevant, wrong, or duplicated data, **such as DOB with 0 and income level zero**.



4. Age Calculation

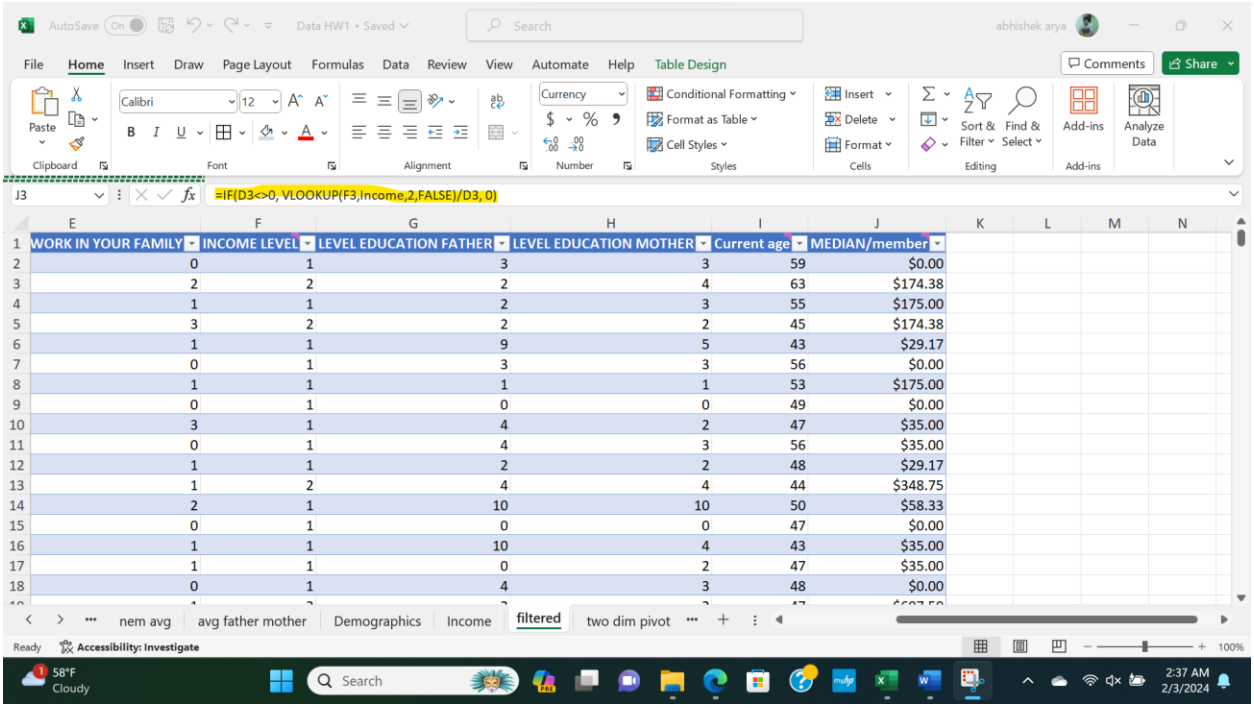
Computed the current age of each student using their date of birth (DOB) in the DMMYYYY format using the DATEDIF function.





5.Median Monthly Income Analysis

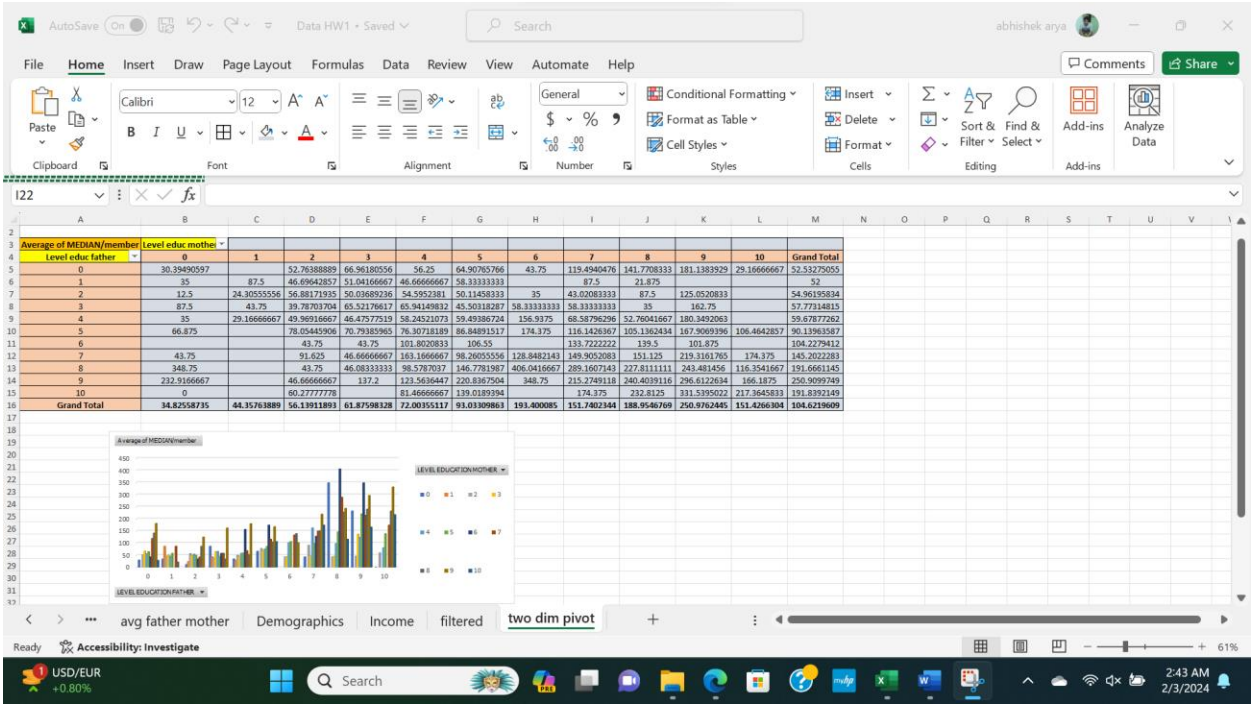
Calculated the median monthly income for each income level, considering the middle point for each range



	E	F	G	H	I	J
1	WORK IN YOUR FAMILY	INCOME LEVEL	LEVEL EDUCATION FATHER	LEVEL EDUCATION MOTHER	Current age	MEDIAN/member
2	0	1	3	3	59	$0.00
3	2	2	2	4	63	$174.38
4	1	1	2	3	55	$175.00
5	3	2	2	2	45	$174.38
6	1	1	9	5	43	$29.17
7	0	1	3	3	56	$0.00
8	1	1	1	1	53	$175.00
9	0	1	0	0	49	$0.00
10	3	1	4	2	47	$35.00
11	0	1	4	3	56	$35.00
12	1	1	2	2	48	$29.17
13	1	2	4	4	44	$348.75
14	2	1	10	10	50	$58.33
15	0	1	0	0	47	$0.00
16	1	1	10	4	43	$35.00
17	1	1	0	2	47	$35.00
18	0	1	4	3	48	$0.00

6. Relation between Income and Parent's Education Level

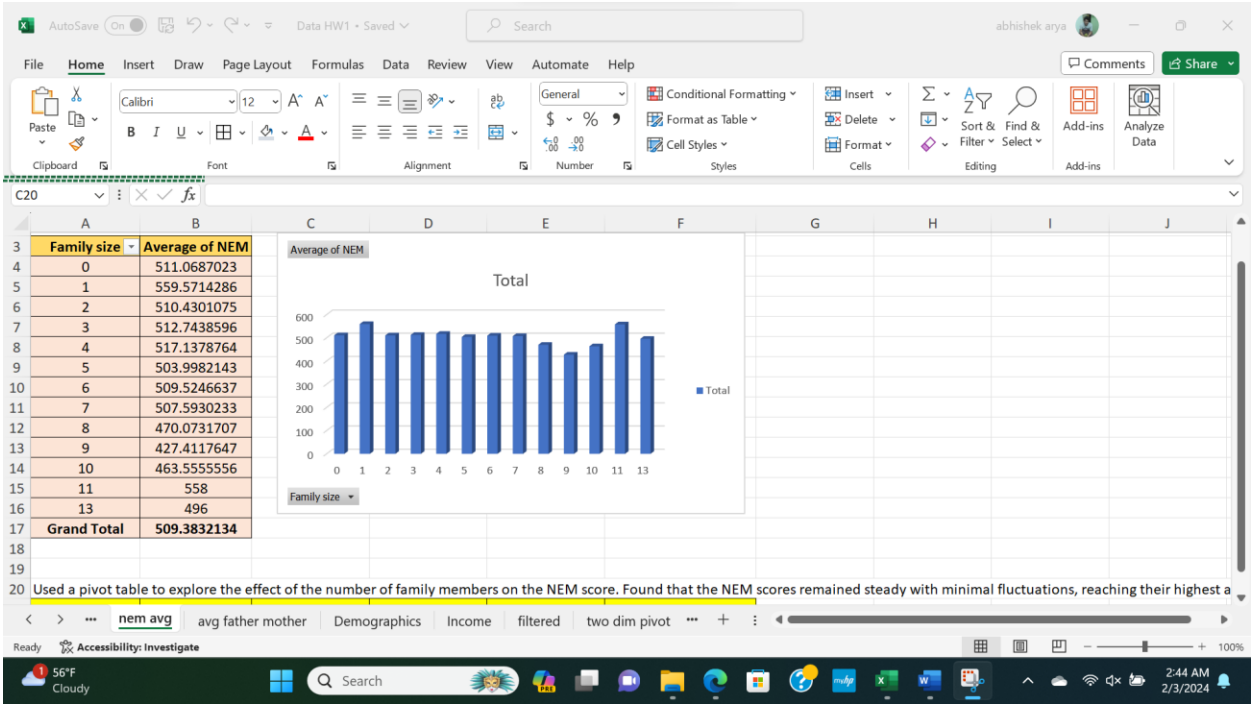
Analyzed the relationship between the median monthly income per family member and the level of education of the father and mother using a pivot table. Identified that education levels were proportional to median monthly income, with fluctuations and a peak at level 8.



Average of MEDAN/member	Level educ mother											
Level educ father	0	1	2	3	4	5	6	7	8	9	10	Grand Total
0	30.39490597		52.76388889	66.96180556	56.25	64.90765766	43.75	119.4940476	141.7708333	181.1383929	29.16666667	52.53275055
1	35	87.5	46.69642857	51.04166667	46.66666667	58.33333333		87.5	21.875			52
2	12.5	24.30555556	56.88171935	50.03689236	54.5952381	50.11458333	35	43.02083333	87.5	125.0520833		54.96195834
3	87.5	43.75	39.78703704	65.52176617	65.94149832	45.50318287	58.33333333	58.33333333	35	162.75		57.77314815
4	35	29.16666667	49.96916667	46.47577519	58.24521073	59.49386724	156.9375	68.58796296	52.76041667	180.3492063		59.67877262
5	66.875		78.05445906	70.79385965	76.30718189	86.84891517	174.375	116.1426367	105.1362434	167.9069396	106.4642857	90.13963587
6			43.75	43.75	101.8020833	106.55		133.7222222	139.5	101.875		104.2279412
7	43.75		91.625	46.66666667	163.1666667	98.26055556	128.8482143	149.9052083	151.125	219.3161765	174.375	145.2022283
8	348.75		43.75	46.08333333	98.5787037	146.7781987	406.0416667	289.1607143	227.8111111	243.481456	116.3541667	191.6661145
9	232.9166667		46.66666667	137.2	123.5636447	220.8367504	348.75	215.2749118	240.4039116	296.6122634	166.1875	250.9099749
10	0		60.27777778		81.46666667	139.0189394		174.375	232.8125	331.5395022	217.3645833	191.8392149
Grand Total	34.82558735	44.35763889	56.13911893	61.87598328	72.00355117	93.03309863	193.400085	151.7402344	188.9546769	250.9762445	151.4266304	104.6219609

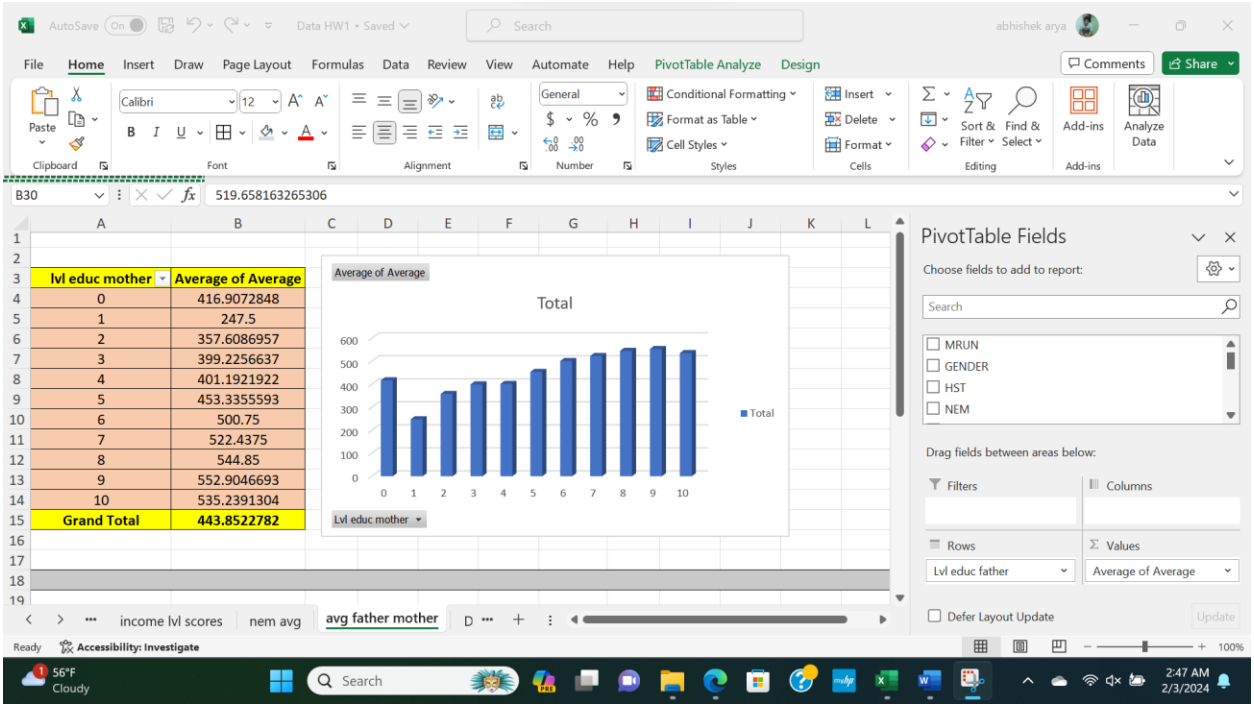
7. Effect of Family Size on NEM Score

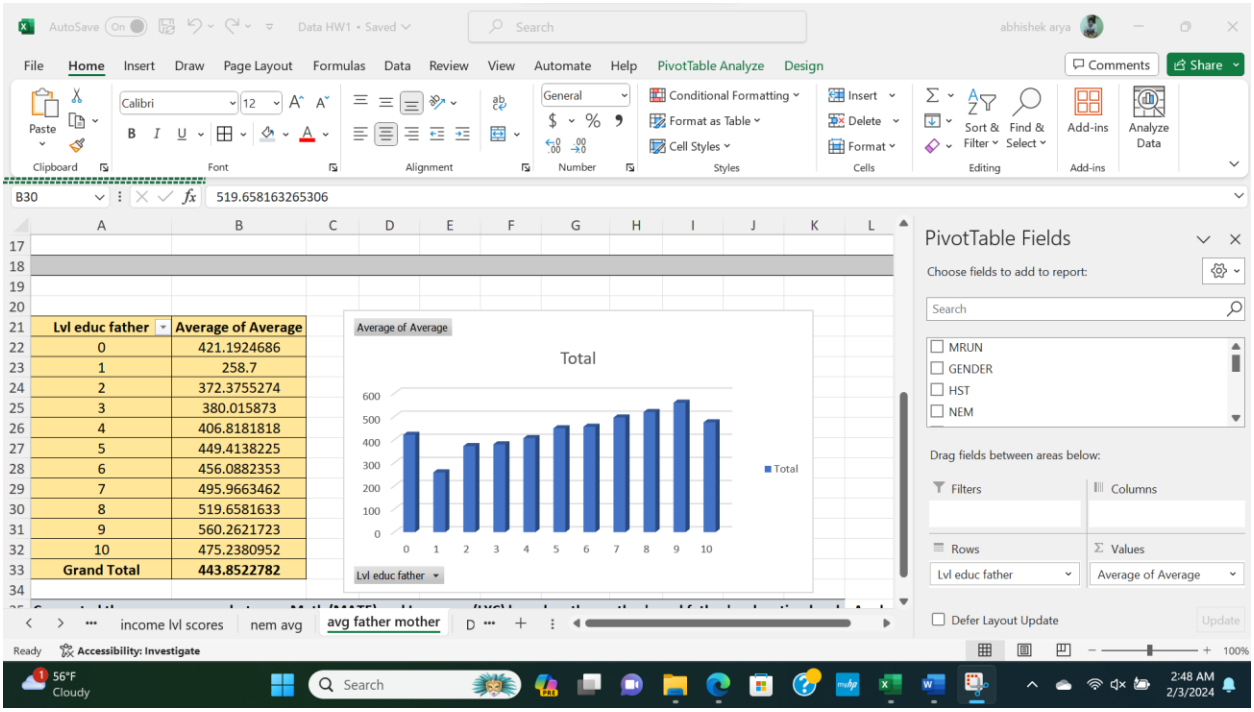
Used a pivot table to explore the effect of the number of family members on the NEM score. Found that the NEM scores remained steady with minimal fluctuations, reaching their highest average scores at 1 and 11 family members.



8. Impact of Parent's Education Level on Test Scores

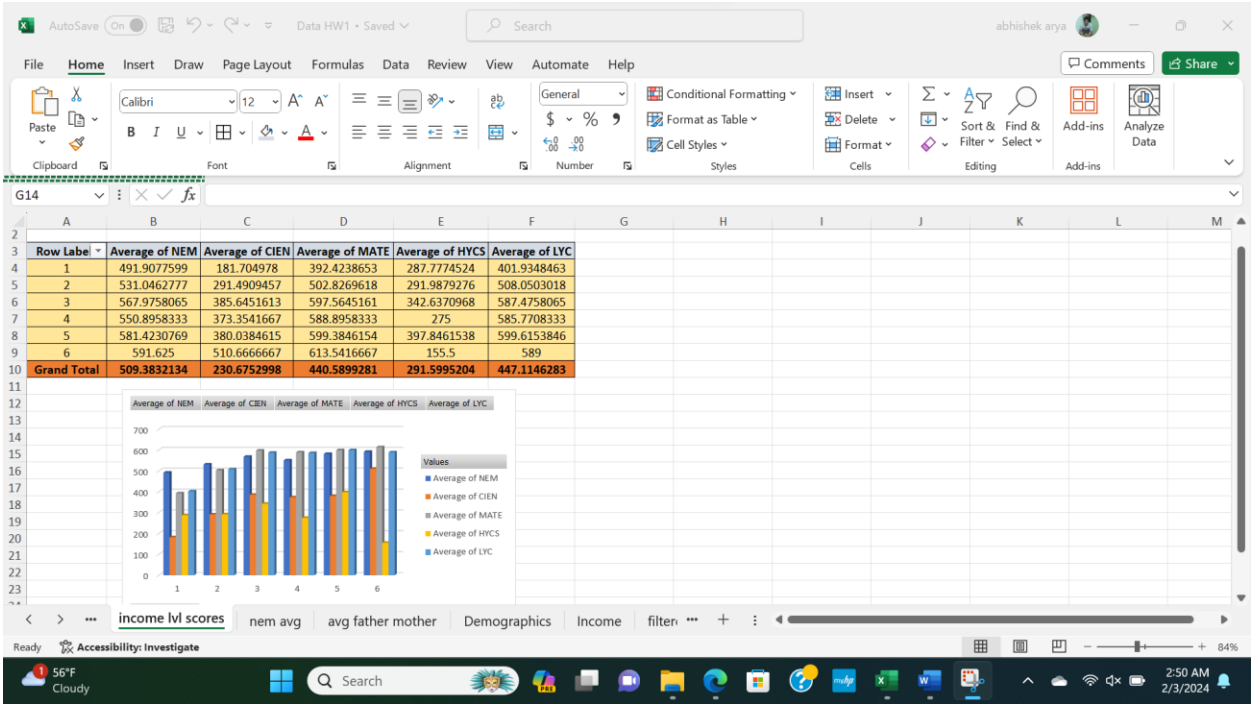
Computed the average scores between Math (MATE) and Language (LYC) based on the mother's and father's education levels. Analyzed that the mother's scores had a bigger effect than the father's on the average scores, particularly with the increase in income.





9. Effect of Income Level on Test Scores

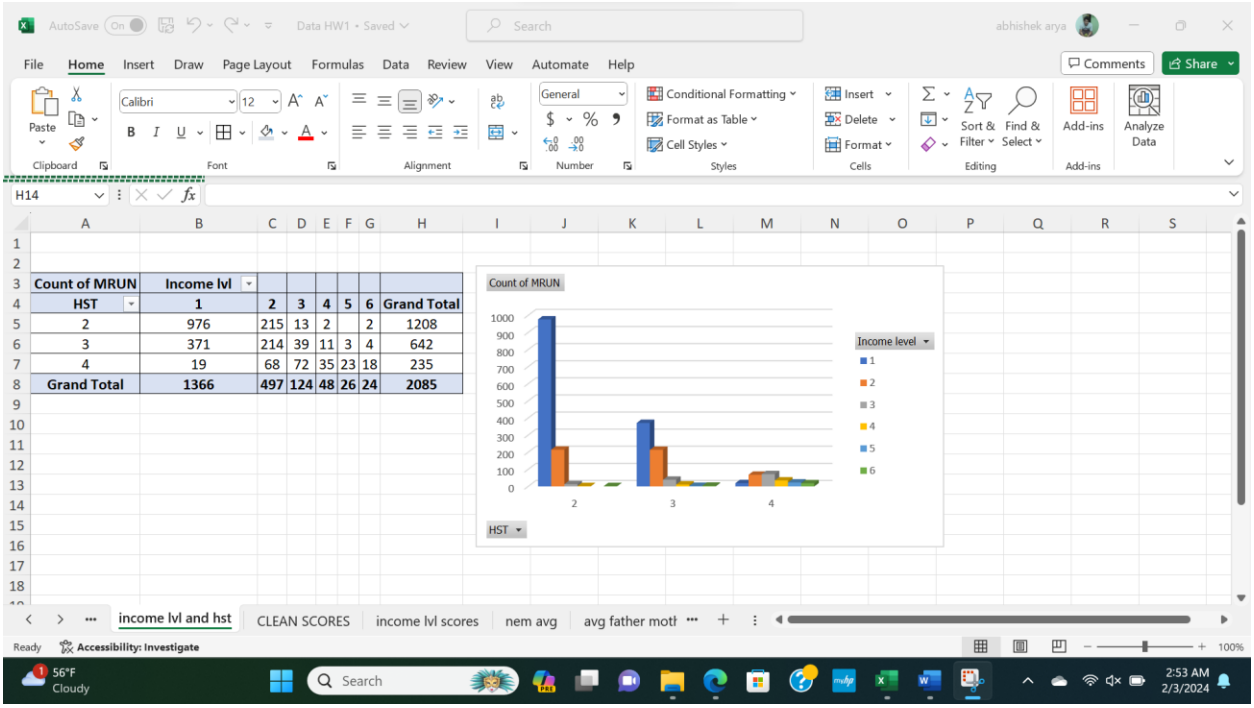
Examined the effect of the level of income on NEM, MATE, CIEN, HYC, and LYC average scores. Observed a steady increase with fluctuations, reaching their peak at specific income levels.



Row Labels	Average of NEM	Average of CIEN	Average of MATE	Average of HYCS	Average of LYC
1	491.9077599	181.704978	392.4238653	287.7774524	401.9348463
2	531.0462777	291.4909457	502.8269618	291.9879276	508.0503018
3	567.9758065	385.6451613	597.5645161	342.6370968	587.4758065
4	550.8958333	373.3541667	588.8958333	275	585.7708333
5	581.4230769	380.0384615	599.3846154	397.8461538	599.6153846
6	591.625	510.6666667	613.5416667	155.5	589
Grand Total	**509.3832134**	**230.6752998**	**440.5899281**	**291.5995204**	**447.1146283**

10. Relationship between High School Type and Income Level

Explored the relationship between high school types (HST) and income levels. Identified that students with income levels 1 and 2 predominantly opted for public schools (HST-2), while higher-income students preferred private schools (HST-4).



Count of MRUN	Income lvl					
HST	1	2	3	4	5	6	Grand Total
2	976	215	13	2		2	1208
3	371	214	39	11	3	4	642
4	19	68	72	35	23	18	235
Grand Total	1366	497	124	48	26	24	2085

Conclusion:

The analysis of the Chilean college admissions dataset revealed key insights. Higher education levels were correlated with increased median monthly income. Family size was found to impact NEM scores, with mothers' education levels showing a stronger influence on test scores. The study also identified a connection between income levels and test scores, as well as a preference for specific high school types based on income. Overall, these findings provide valuable understanding of the relationships among income, education, family dynamics, and academic performance.