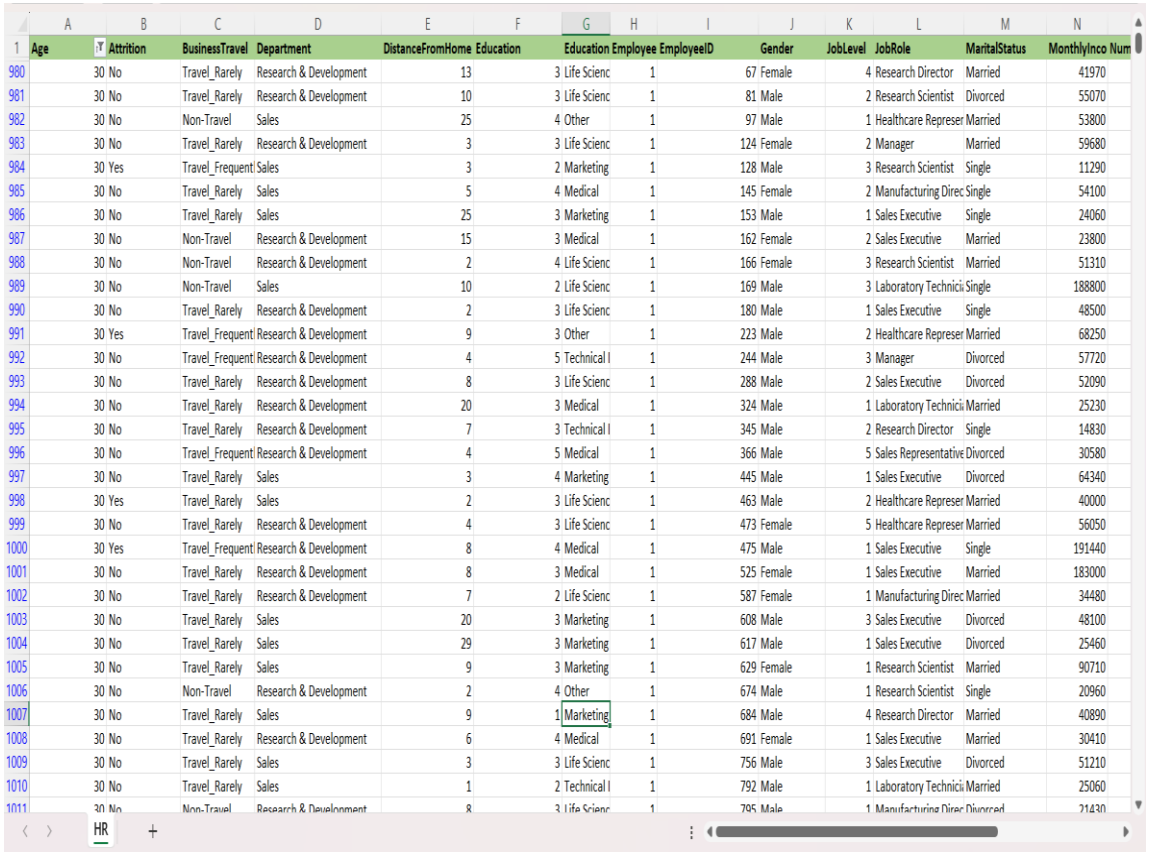
PSYLIQ

HR DATA ANALYSIS ASSESSMENT

Name- Debashree Priya Sahoo

Q1. Using Excel, how would you filter the dataset to only show employees aged 30 and above?



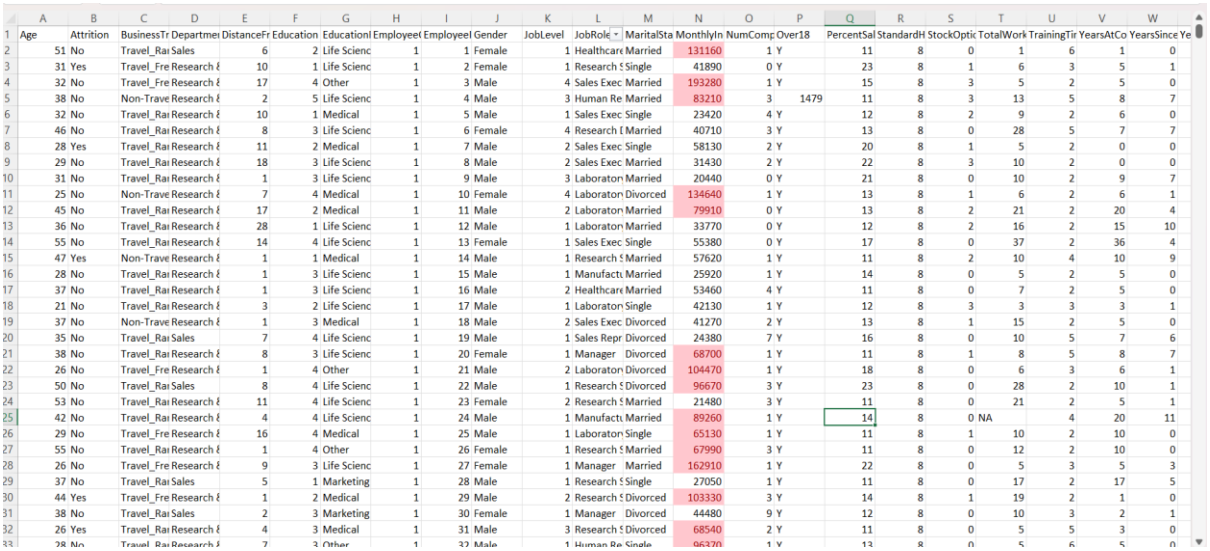
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Age	Attrition	BusinessTravel	Department	DistanceFromHome	Education	Education	Employee	EmployeeID	Gender	JobLevel	JobRole	MaritalStatus	MonthlyInco	Num
980	30	No	Travel_Rarely	Research & Development	13	3	Life Scienc	1	67	Female	4	Research Director	Married	41970	
981	30	No	Travel_Rarely	Research & Development	10	3	Life Scienc	1	81	Male	2	Research Scientist	Divorced	55070	
982	30	No	Non-Travel	Sales	25	4	Other	1	97	Male	1	Healthcare Represer	Married	53800	
983	30	No	Travel_Rarely	Research & Development	3	3	Life Scienc	1	124	Female	2	Manager	Married	59680	
984	30	Yes	Travel_Frequent	Sales	3	2	Marketing	1	128	Male	3	Research Scientist	Single	11290	
985	30	No	Travel_Rarely	Sales	5	4	Medical	1	145	Female	2	Manufacturing Direc	Single	54100	
986	30	No	Travel_Rarely	Sales	25	3	Marketing	1	153	Male	1	Sales Executive	Single	24060	
987	30	No	Non-Travel	Research & Development	15	3	Medical	1	162	Female	2	Sales Executive	Married	23800	
988	30	No	Non-Travel	Research & Development	2	4	Life Scienc	1	166	Female	3	Research Scientist	Married	51310	
989	30	No	Non-Travel	Sales	10	2	Life Scienc	1	169	Male	3	Laboratory Technici	Single	188800	
990	30	No	Travel_Rarely	Research & Development	2	3	Life Scienc	1	180	Male	1	Sales Executive	Single	48500	
991	30	Yes	Travel_Frequent	Research & Development	9	3	Other	1	223	Male	2	Healthcare Represer	Married	68250	
992	30	No	Travel_Frequent	Research & Development	4	5	Technical I	1	244	Male	3	Manager	Divorced	57720	
993	30	No	Travel_Rarely	Research & Development	8	3	Life Scienc	1	288	Male	2	Sales Executive	Divorced	52090	
994	30	No	Travel_Rarely	Research & Development	20	3	Medical	1	324	Male	1	Laboratory Technici	Married	25230	
995	30	No	Travel_Rarely	Research & Development	7	3	Technical I	1	345	Male	2	Research Director	Single	14830	
996	30	No	Travel_Frequent	Research & Development	4	5	Medical	1	366	Male	5	Sales Representative	Divorced	30580	
997	30	No	Travel_Rarely	Sales	3	4	Marketing	1	445	Male	1	Sales Executive	Divorced	64340	
998	30	Yes	Travel_Rarely	Sales	2	3	Life Scienc	1	463	Male	2	Healthcare Represer	Married	40000	
999	30	No	Travel_Rarely	Research & Development	4	3	Life Scienc	1	473	Female	5	Healthcare Represer	Married	56050	
1000	30	Yes	Travel_Frequent	Research & Development	8	4	Medical	1	475	Male	1	Sales Executive	Single	191440	
1001	30	No	Travel_Rarely	Research & Development	8	3	Medical	1	525	Female	1	Sales Executive	Married	183000	
1002	30	No	Travel_Rarely	Research & Development	7	2	Life Scienc	1	587	Female	1	Manufacturing Direc	Married	34480	
1003	30	No	Travel_Rarely	Sales	20	3	Marketing	1	608	Male	3	Sales Executive	Divorced	48100	
1004	30	No	Travel_Rarely	Sales	29	3	Marketing	1	617	Male	1	Sales Executive	Divorced	25460	
1005	30	No	Travel_Rarely	Sales	9	3	Marketing	1	629	Female	1	Research Scientist	Married	90710	
1006	30	No	Non-Travel	Research & Development	2	4	Other	1	674	Male	1	Research Scientist	Single	20960	
1007	30	No	Travel_Rarely	Sales	9	1	Marketing	1	684	Male	4	Research Director	Married	40890	
1008	30	No	Travel_Rarely	Research & Development	6	4	Medical	1	691	Female	1	Sales Executive	Married	30410	
1009	30	No	Travel_Rarely	Sales	3	3	Life Scienc	1	756	Male	3	Sales Executive	Divorced	51210	
1010	30	No	Travel_Rarely	Sales	1	2	Technical I	1	792	Male	1	Laboratory Technici	Married	25060	
1011	30	No	Non-Travel	Research & Development	8	3	Life Scienc	1	795	Male	1	Manufacturing Direc	Divorced	21430	

HR +

Q2. Create a pivot table to summarize the average Monthly Income by Job Role.

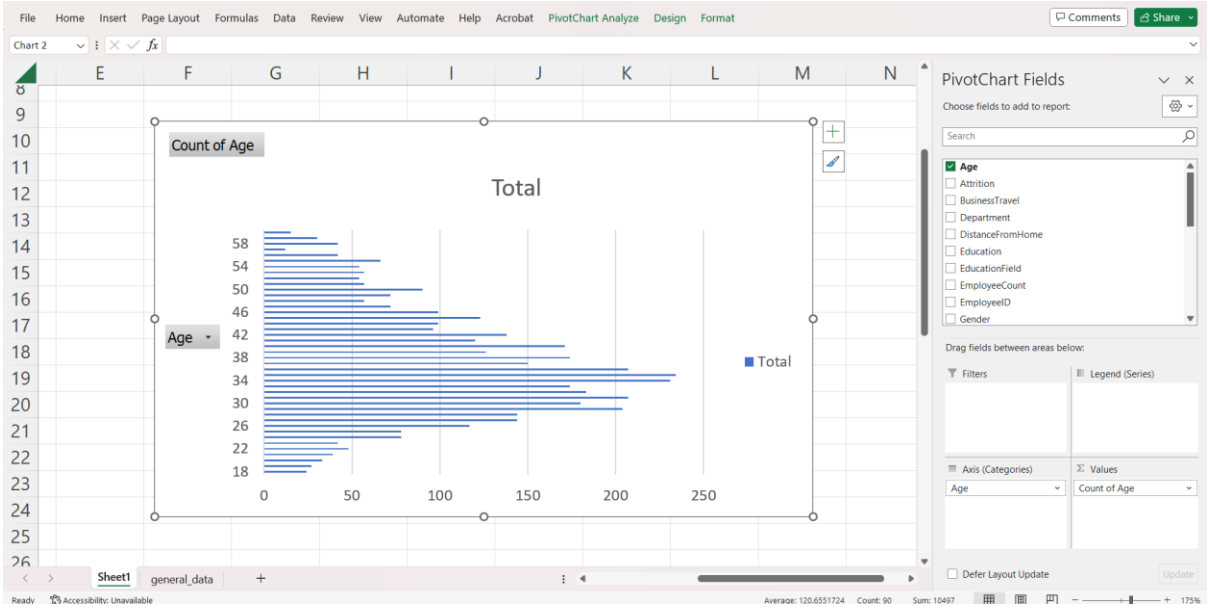
Age	(All)
Row Labels	**Average of MonthlyIncome**
Healthcare Representative	60983.74046
Human Resources	58528.07692
Laboratory Technician	66314.05405
Manager	63395.88235
Manufacturing Director	69183.72414
Research Director	65473.125
Research Scientist	64975.68493
Sales Executive	65186.68712
Sales Representative	65370.96386
Grand Total	**65029.31293**

Q3. Apply conditional formatting to highlight employees with Monthly Income above the company's average income.



Age	Attrition	BusinessTr	Departmer	DistanceFr	Education	EducationI	EmployeeC	EmployeeI	Gender	JobLevel	JobRole	MaritalSta	MonthlyIn	NumComp	Over18	PercentSal	StandardH	StockOptic	TotalWork	TrainingTir	YearsAtCo	YearsSince	Ye
51	No	Travel_Rai	Sales	6	2	Life Scienc	1	1	Female	1	Healthcare	Married	131160	1	Y	11	8	0	1	6	1	0	
31	Yes	Travel_Fre	Research &	10	1	Life Scienc	1	2	Female	1	Research S	Single	41890	0	Y	23	8	1	6	3	5	1	
32	No	Travel_Fre	Research &	17	4	Other	1	3	Male	4	Sales Exec	Married	193280	1	Y	15	8	3	5	2	5	0	
38	No	Non-Trave	Research &	2	5	Life Scienc	1	4	Male	3	Human Re	Married	83210	3	1479	11	8	3	13	5	8	7	
32	No	Travel_Rai	Research &	10	1	Medical	1	5	Male	1	Sales Exec	Single	23420	4	Y	12	8	2	9	2	6	0	
46	No	Travel_Rai	Research &	8	3	Life Scienc	1	6	Female	4	Research I	Married	40710	3	Y	13	8	0	28	5	7	7	
28	Yes	Travel_Rai	Research &	11	2	Medical	1	7	Male	2	Sales Exec	Single	58130	2	Y	20	8	1	5	2	0	0	
29	No	Travel_Rai	Research &	18	3	Life Scienc	1	8	Male	2	Sales Exec	Married	31430	2	Y	22	8	3	10	2	0	0	
31	No	Travel_Rai	Research &	1	3	Life Scienc	1	9	Male	3	Laboratori	Married	20440	0	Y	21	8	0	10	2	9	7	
25	No	Non-Trave	Research &	7	4	Medical	1	10	Female	4	Laboratori	Divorced	134640	1	Y	13	8	1	6	2	6	1	
45	No	Travel_Rai	Research &	17	2	Medical	1	11	Male	2	Laboratori	Married	79910	0	Y	13	8	2	21	2	20	4	
36	No	Travel_Rai	Research &	28	1	Life Scienc	1	12	Male	1	Laboratori	Married	33770	0	Y	12	8	2	16	2	15	10	
55	No	Travel_Rai	Research &	14	4	Life Scienc	1	13	Female	1	Sales Exec	Single	55380	0	Y	17	8	0	37	2	36	4	
47	Yes	Non-Trave	Research &	1	1	Medical	1	14	Male	1	Research S	Married	57620	1	Y	11	8	2	10	4	10	9	
28	No	Travel_Rai	Research &	1	3	Life Scienc	1	15	Male	1	Manufactu	Married	25920	1	Y	14	8	0	5	2	5	0	
37	No	Travel_Rai	Research &	1	3	Life Scienc	1	16	Male	2	Healthcare	Married	53460	4	Y	11	8	0	7	2	5	0	
21	No	Travel_Rai	Research &	3	2	Life Scienc	1	17	Male	1	Laboratori	Single	42130	1	Y	12	8	3	3	3	3	1	
37	No	Non-Trave	Research &	1	3	Medical	1	18	Male	2	Sales Exec	Divorced	41270	2	Y	13	8	1	15	2	5	0	
35	No	Travel_Rai	Sales	7	4	Life Scienc	1	19	Male	1	Sales Repr	Divorced	24380	7	Y	16	8	0	10	5	7	6	
38	No	Travel_Rai	Research &	8	3	Life Scienc	1	20	Female	1	Manager	Divorced	68700	1	Y	11	8	1	8	5	8	7	
26	No	Travel_Fre	Research &	1	4	Other	1	21	Male	2	Laboratori	Divorced	104470	1	Y	18	8	0	6	3	6	1	
50	No	Travel_Rai	Sales	8	4	Life Scienc	1	22	Male	1	Research S	Divorced	96670	3	Y	23	8	0	28	2	10	1	
53	No	Travel_Rai	Research &	11	4	Life Scienc	1	23	Female	2	Research S	Married	21480	3	Y	11	8	0	21	2	5	1	
42	No	Travel_Rai	Research &	4	4	Life Scienc	1	24	Male	1	Manufactu	Married	89260	1	Y	14	8	0	NA	4	20	11	
29	No	Travel_Fre	Research &	16	4	Medical	1	25	Male	1	Laboratori	Single	65130	1	Y	11	8	1	10	2	10	0	
55	No	Travel_Rai	Research &	1	4	Other	1	26	Female	1	Research S	Married	67990	3	Y	11	8	0	12	2	10	0	
26	No	Travel_Fre	Research &	9	3	Life Scienc	1	27	Female	1	Manager	Married	162910	1	Y	22	8	0	5	3	5	3	
37	No	Travel_Rai	Sales	5	1	Marketing	1	28	Male	1	Research S	Single	27050	1	Y	11	8	0	17	2	17	5	
44	Yes	Travel_Fre	Research &	1	2	Medical	1	29	Male	2	Research S	Divorced	103330	3	Y	14	8	1	19	2	1	0	
38	No	Travel_Rai	Sales	2	3	Marketing	1	30	Female	1	Manager	Divorced	44480	9	Y	12	8	0	10	3	2	1	
26	Yes	Travel_Rai	Research &	4	3	Medical	1	31	Male	3	Research S	Divorced	68540	2	Y	11	8	0	5	5	3	0	
28	No	Travel_Rai	Research &	7	3	Other	1	32	Male	1	Human Re	Single	96370	1	Y	13	8	0	5	6	5	0	

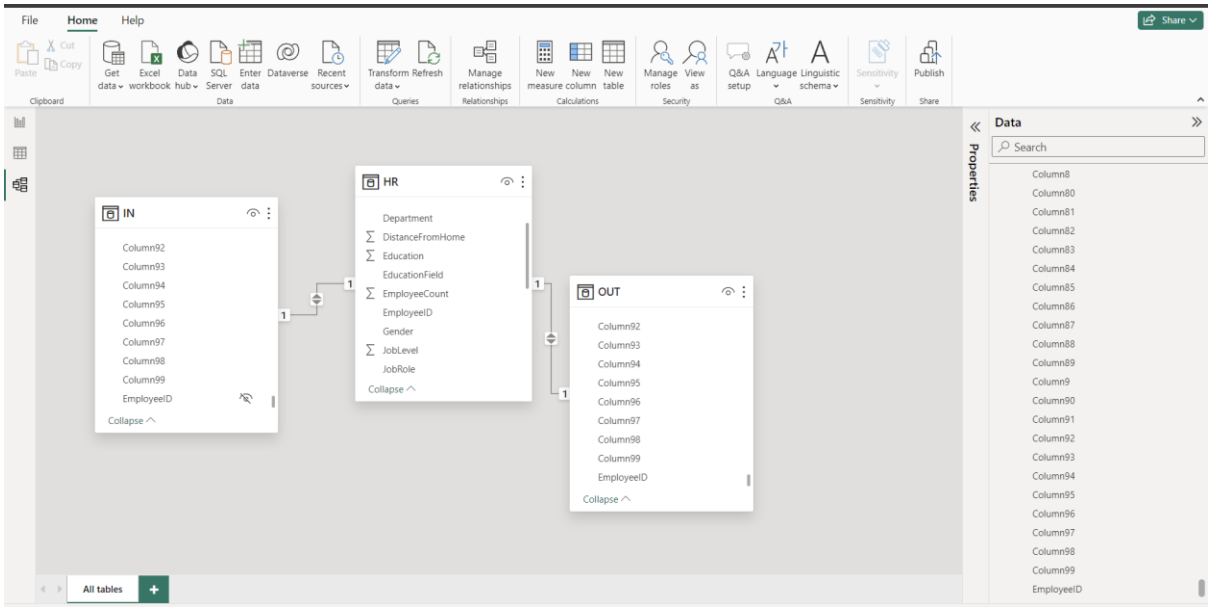
Q4. Create a bar chart in Excel to visualize the distribution of employee ages.



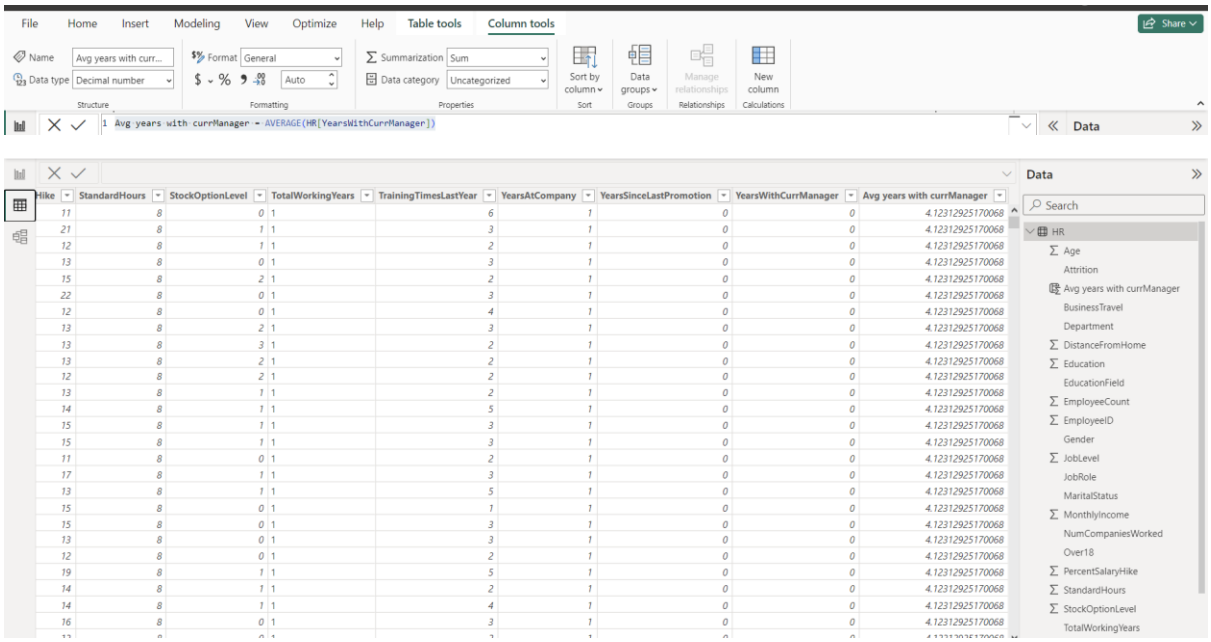
Q5. Identify and clean any missing or inconsistent data in the "Department" column

- None(NO MISSING OR INCONSISTENT DATA)

Q6. In Power BI, establish a relationship between the "EmployeeID" in the employee data and the "EmployeeID" in the time tracking data.



Q7. Using DAX, create a calculated column that calculates the average years an employee has spent with their current manager



Q8. Using Excel, create a pivot table that displays the count of employees in each Marital Status category, segmented by Department.

Row Labels	Sum of EmployeeCount
Divorced	**981**
Human Resources	21
Research & Development	621
Sales	339
Married	**2019**
Human Resources	96
Research & Development	1350
Sales	573
Single	**1410**
Human Resources	72
Research & Development	912
Sales	426
Grand Total	**4410**

Q9. Apply conditional formatting to highlight employees with both above-average Monthly Income and above-average Job Satisfaction

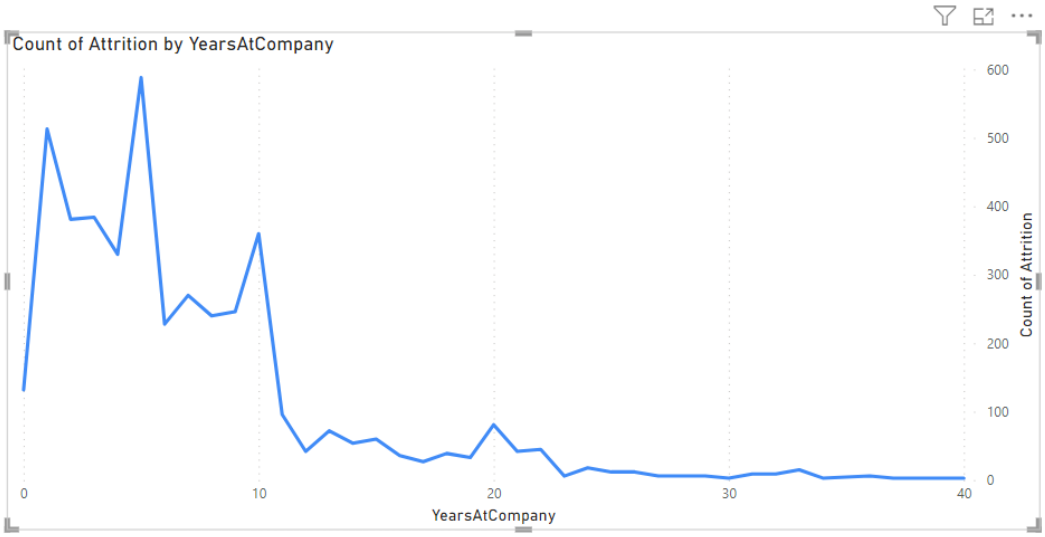
	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	EmployeeID	Gender	JobLevel	JobRole	MaritalStatus	MonthlyIn	NumComp	Over18	PercentSal	StandardH	StockOptic	TotalWork	TrainingTir	YearsAtCo	YearsSince	YearsWithCurrManager	
2	1	1 Female	1	Healthcare	Married	131160	1	Y	11	8	0	1	6	1	0	0	
3	1	2 Female	1	Research S	Single	41890	0	Y	23	8	1	6	3	5	1	4	
4	1	3 Male	4	Sales Exec	Married	193280	1	Y	15	8	3	5	2	5	0	3	
5	1	4 Male	3	Human Re	Married	83210	3	Y	11	8	3	13	5	8	7	5	
6	1	5 Male	1	Sales Exec	Single	23420	4	Y	12	8	2	9	2	6	0	4	
7	1	6 Female	4	Research I	Married	40710	3	Y	13	8	0	28	5	7	7	7	
8	1	7 Male	2	Sales Exec	Single	58130	2	Y	20	8	1	5	2	0	0	0	
9	1	8 Male	2	Sales Exec	Married	31430	2	Y	22	8	3	10	2	0	0	0	
10	1	9 Male	3	Laboraton	Married	20440	0	Y	21	8	0	10	2	9	7	8	
11	1	10 Female	4	Laboraton	Divorced	134640	1	Y	13	8	1	6	2	6	1	5	
12	1	11 Male	2	Laboraton	Married	79910	0	Y	13	8	2	21	2	20	4	10	
13	1	12 Male	1	Laboraton	Married	33770	0	Y	12	8	2	16	2	15	10	11	
14	1	13 Female	1	Sales Exec	Single	55380	0	Y	17	8	0	37	2	36	4	13	
15	1	14 Male	1	Research S	Married	57620	1	Y	11	8	2	10	4	10	9	9	
16	1	15 Male	1	Manufactu	Married	25920	1	Y	14	8	0	5	2	5	0	4	
17	1	16 Male	2	Healthcare	Married	53460	4	Y	11	8	0	7	2	5	0	1	
18	1	17 Male	1	Laboraton	Single	42130	1	Y	12	8	3	3	3	3	1	0	
19	1	18 Male	2	Sales Exec	Divorced	41270	2	Y	13	8	1	15	2	5	0	2	
20	1	19 Male	1	Sales Repr	Divorced	24380	7	Y	16	8	0	10	5	7	6	2	
21	1	20 Female	1	Manager	Divorced	68700	1	Y	11	8	1	8	5	8	7	7	
22	1	21 Male	2	Laboraton	Divorced	104470	1	Y	18	8	0	6	3	6	1	4	
23	1	22 Male	1	Research S	Divorced	96670	3	Y	23	8	0	28	2	10	1	6	
24	1	23 Female	2	Research S	Married	21480	3	Y	11	8	0	21	2	5	1	3	
25	1	24 Male	1	Manufactu	Married	89260	1	Y	14	8	0	NA	4	20	11	6	
26	1	25 Male	1	Laboraton	Single	65130	1	Y	11	8	1	10	2	10	0	9	
27	1	26 Female	1	Research S	Married	67990	3	Y	11	8	0	12	2	10	0	8	
28	1	27 Female	1	Manager	Married	162910	1	Y	22	8	0	5	3	5	3	3	
29	1	28 Male	1	Research S	Single	27050	1	Y	11	8	0	17	2	17	5	7	
30	1	29 Male	2	Research S	Divorced	103330	3	Y	14	8	1	19	2	1	0	0	
31	1	30 Female	1	Manager	Divorced	44480	9	Y	12	8	0	10	3	2	1	2	
32	1	31 Male	3	Research S	Divorced	68540	2	Y	11	8	0	5	5	3	0	2	
33	1	32 Male	1	Human Re	Single	96370	1	Y	13	8	0	5	6	5	0	2	

Sheet1 general_data +

	A	B	C	D	E
	EmployeeI	Environme	JobSatisfaction	WorkLifeBalance	
	1	3	4	2	
	2	3	2	4	
	3	2	2	1	
	4	4	4	3	
	5	4	1	3	
	6	3	2	2	
	7	1	3	1	
	8	1	2	3	
	9	2	4	3	
	10	2	1	3	
	11	3	4	3	
	12	NA	4	3	
	13	4	1	3	
	14	1	2	2	
	15	4	4	2	
	16	3	4	4	
	17	4	3	4	
	18	1	4	3	
	19	2	2	2	
	20	1	1	3	
	21	3	2	1	
	22	1	2	2	
	23	3	3	2	
	24	2	3	3	
	25	2	4	2	
	26	2	4	3	
	27	1	1	3	
	28	4	4	3	
	29	4	3	1	
	30	4	4	3	
	31	1	2	3	
	32	4	4	3	

employee_survey_data

Q10.In Power BI, create a line chart that visualizes the trend of Employee Attrition over the years.



Q11. Describe how you would create a star schema for this dataset, explaining the benefits of doing so.

A star schema involves designing a database structure that separates dimensional and fact data, promoting efficient querying and analysis. The Star Schema for this dataset can include

Fact Table:

- HR Fact table

Dimension table

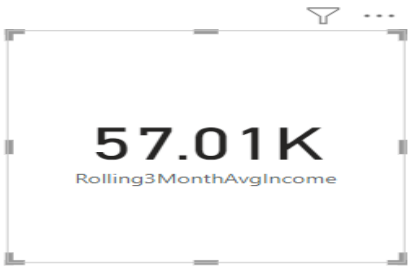
- Employee Dimension
- Job Dimension
- Location Dimension

Benefits of Star Schema

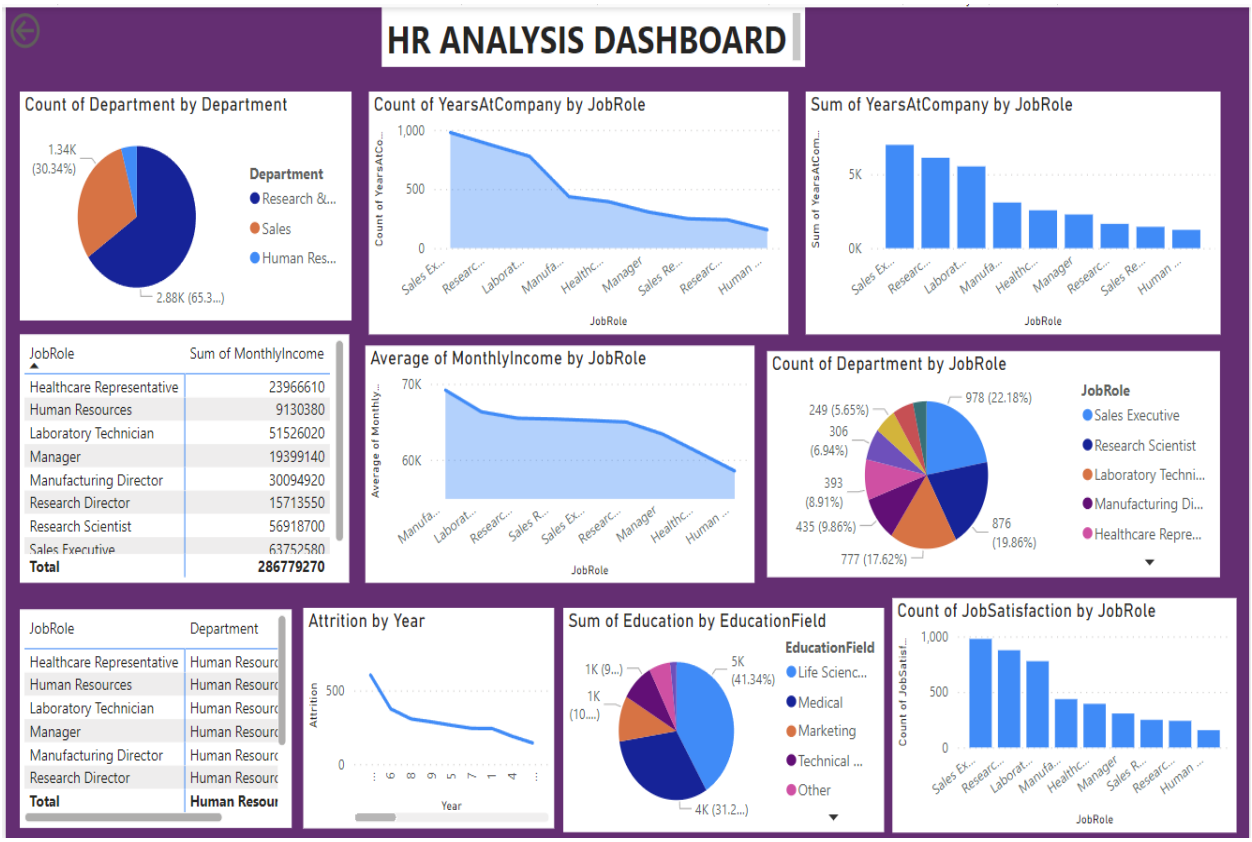
- Easier to understand and maintain, especially for end-users who are performing analyses.
- New dimensions can be added easily without affecting existing structures, providing scalability.
- Reduces data redundancy by normalizing dimension tables, minimizing storage requirements.

Q12. Using DAX, calculate the rolling 3-month average of Monthly Income for each employee.

```
Rolling3MonthAvgIncome =
CALCULATE(
    AVERAGE('generL_DATA'[MonthlyIncome]),
    FILTER(
        ALL('generL_DATA'),
        'generL_DATA'[Age] >= MAX('generL_DATA'[Age]) - 3 && 'generL_DATA'[Age] <= MAX('generL_DATA'[Age])
    )
)
```



Q13. Create a hierarchy in Power BI that allows users to drill down from Department to Job Role to further narrow their analysis.



Q14. How can you set up parameterized queries in Power BI to allow users to filter data based 2 of 2 on the Distance from Home column?

In Power BI, one can set up parameterized queries using Power Query and Power BI's query parameters to allow users to filter data based on the Distance from Home column.

- Load dataset in PowerBI that includes the "Distance from Home" column.
- Go to the "Home" tab. Click on "Transform Data" to open the Power Query Editor.

Create a Parameter

- In the Power Query Editor, go to the "Home" tab. Click on "Manage Parameters." In the "Manage Parameters" window, click on "New."
- Name your parameter (e.g., DistanceParameter). Set the data type to Decimal Number (or Text, depending on your column data type). Set the Suggested Values to "Any value."

Filter Data Based on Parameter

- In the Power Query Editor, select the column (e.g., "Distance from Home") that you want to filter.
- In the filter window, set up the filter condition based on the parameter.

  [Distance from Home] <= DistanceParameter

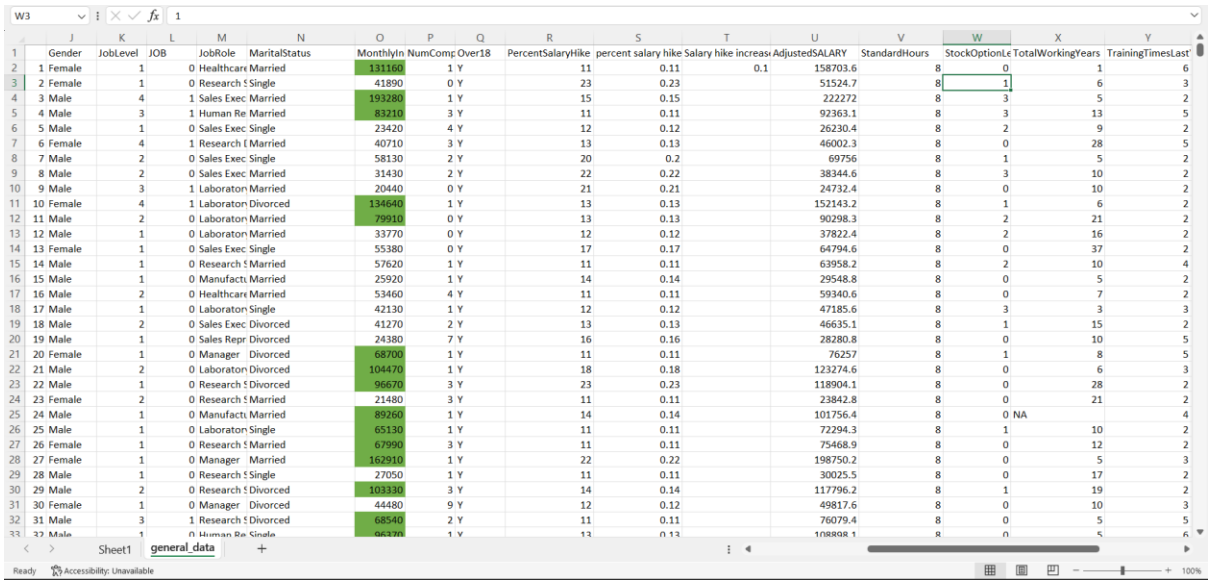
- Click on "Close & Apply" in the Home tab to apply the changes to your Power BI model.

- **Create a Parameterized Visual**
- In the Report view, create a visual (e.g., a table or a chart) that uses the filtered data. Use the "DistanceParameter" as a slicer or filter in your visualizations. it will now dynamically filter the data based on the selected distance value.

Q15. In Excel, calculate the total Monthly Income for each Department, considering only the employees with a Job Level greater than or equal to 3

JobLevel	(Multiple Items)
Row Labels	**Sum of MonthlyIncome**
Human Resources	3259140
Research & Development	53502900
Sales	22974330
Grand Total	**79736370**

Q16. Explain how to perform a What-If analysis in Excel to understand the impact of a 10% increase in Percent Salary Hike on Monthly Income.



	Gender	JobLevel	JOB	JobRole	MaritalStatus	MonthlyIn	NumComp	Over18	PercentSalaryHike	percent salary hike	Salary hike increas	AdjustedSALARY	StandardHours	StockOptionLe	TotalWorkingYears	TrainingTimesLast
1	Female	1	0	Healthcare	Married	131160	1	Y	11	0.11	0.1	158703.6	8	0	1	6
2	Female	1	0	Research S	Single	41890	0	Y	23	0.23		51524.7	8	1	6	3
3	Male	4	1	Sales Exec	Married	193280	1	Y	15	0.15		222272	8	3	5	2
4	Male	3	1	Human Re	Married	83210	3	Y	11	0.11		92363.1	8	3	13	5
5	Male	1	0	Sales Exec	Single	23420	4	Y	12	0.12		26230.4	8	2	9	2
6	Female	4	1	Research I	Married	40710	3	Y	13	0.13		46002.3	8	0	28	5
7	Male	2	0	Sales Exec	Single	58130	2	Y	20	0.2		69756	8	1	5	2
8	Male	2	0	Sales Exec	Married	31430	2	Y	22	0.22		38344.6	8	3	10	2
9	Male	3	1	Laboraton	Married	20440	0	Y	21	0.21		24732.4	8	0	10	2
10	Female	4	1	Laboraton	Divorced	134640	1	Y	13	0.13		152143.2	8	1	6	2
11	Male	2	0	Laboraton	Married	79910	0	Y	13	0.13		90298.3	8	2	21	2
12	Male	1	0	Laboraton	Married	33770	0	Y	12	0.12		37822.4	8	2	16	2
13	Female	1	0	Sales Exec	Single	55380	0	Y	17	0.17		64794.6	8	0	37	2
14	Male	1	0	Research S	Married	57620	1	Y	11	0.11		63958.2	8	2	10	4
15	Male	1	0	Manufactu	Married	25920	1	Y	14	0.14		29548.8	8	0	5	2
16	Male	2	0	Healthcare	Married	53460	4	Y	11	0.11		59340.6	8	0	7	2
17	Male	1	0	Laboraton	Single	42130	1	Y	12	0.12		47185.6	8	3	3	3
18	Male	2	0	Sales Exec	Divorced	41270	2	Y	13	0.13		46635.1	8	1	15	2
19	Male	1	0	Sales Repr	Divorced	24380	7	Y	16	0.16		28280.8	8	0	10	5
20	Female	1	0	Manager	Divorced	68700	1	Y	11	0.11		76257	8	1	8	5
21	Male	2	0	Laboraton	Divorced	104470	1	Y	18	0.18		123274.6	8	0	6	3
22	Male	1	0	Research S	Divorced	96670	3	Y	23	0.23		118904.1	8	0	28	2
23	Female	2	0	Research S	Married	21480	3	Y	11	0.11		23842.8	8	0	21	2
24	Male	1	0	Manufactu	Married	89260	1	Y	14	0.14		101756.4	8	0	NA	4
25	Male	1	0	Laboraton	Single	65130	1	Y	11	0.11		72294.3	8	1	10	2
26	Female	1	0	Research S	Married	67990	3	Y	11	0.11		75468.9	8	0	12	2
27	Female	1	0	Manager	Married	162910	1	Y	22	0.22		198750.2	8	0	5	3
28	Male	1	0	Research S	Single	27050	1	Y	11	0.11		30025.5	8	0	17	2
29	Male	2	0	Research S	Divorced	103330	3	Y	14	0.14		117796.2	8	1	19	2
30	Female	1	0	Manager	Divorced	44480	9	Y	12	0.12		49817.6	8	0	10	3
31	Male	3	1	Research S	Divorced	68540	2	Y	11	0.11		76079.4	8	0	5	5

Q17. Verify if the data adheres to a predefined schema. What actions would you take if you find inconsistencies

- Ensure the dataset aligns with the predefined schema, considering structure, data types, and specified rules.
- Add the 'EmployeeID' label to the 'in-time' and 'out-time' datasets to maintain uniformity across data sources.
- There is a need to reorder EmployeeID column, and changing data type of ‘TotalWorkingHours’ column

- Address missing values (NA) and blank values in the 'general_data,' 'employee_survey_data,' and 'manager_survey_data' datasets. This might involve filling missing values, removing incomplete records, or applying other appropriate data cleaning techniques.
- Implement data validation checks to ensure that the changes made align with the predefined schema. This may involve running scripts or queries to validate data against schema rules.
- Confirm that data profiling was conducted meticulously.