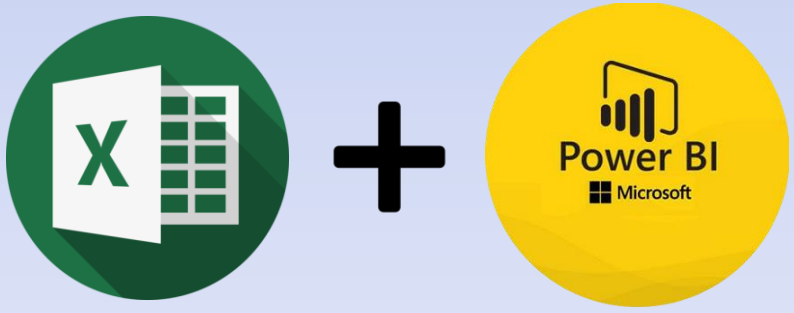
PSYLIQ

HR Data Analysis Using Power Bi & SQL



Author : Usha Kumari

HR DATA ANALYSIS (ADVANCE EXCEL & POWER BI)

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

Step 1: Start by filtering the data.

Step 2: Select the dropdown menu of the filter. Look for the option labeled "Number filters.“

Step 3: Within the number filters, locate the option labeled "Greater than or equal to." Insert the value 30 and press Enter.

Age	Attritio	BusinessTravel	Department
30	No	Travel_Rarely	Research & Development
30	No	Travel_Rarely	Research & Development
30	No	Non-Travel	Sales
30	No	Travel_Rarely	Research & Development
30	Yes	Travel_Frequently	Sales
30	No	Travel_Rarely	Sales
30	No	Travel_Rarely	Sales
30	No	Non-Travel	Research & Development
30	No	Non-Travel	Research & Development
30	No	Non-Travel	Sales
30	No	Travel_Rarely	Research & Development
30	Yes	Travel_Frequently	Research & Development
30	No	Travel_Frequently	Research & Development
30	No	Travel_Rarely	Research & Development
30	No	Travel_Rarely	Research & Development
30	No	Travel_Rarely	Research & Development
30	No	Travel_Frequently	Research & Development
30	No	Travel_Rarely	Sales
30	Yes	Travel_Rarely	Sales
30	No	Travel_Rarely	Research & Development
30	Yes	Travel_Frequently	Research & Development
30	No	Travel_Rarely	Research & Development
30	No	Travel_Rarely	Research & Development

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

Insert Pivot Table:

- Highlight your data and go to the "Insert" tab.
- Click "PivotTable" and choose where to place it (e.g., a new worksheet).

Design Pivot Table:

- Drag "Job Role" to "Rows" and "Monthly Income" to "Values."
- Set "Monthly Income" to display the average by selecting "Value Field Settings.

Job Role	Average Monthly Income
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

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

1.The Average Monthly income of the employees are 65029 ,Below are the Income of the employees Above Average.

EmployeeID	Gender	JobLevel	JobRole	MaritalStatus	MonthlyIncome
169	Male	3	Laboratory Technician	Single	188800
223	Male	2	Healthcare Representati	Married	68250
475	Male	1	Sales Executive	Single	191440
525	Female	1	Sales Executive	Married	183000
629	Female	1	Research Scientist	Married	90710
807	Male	3	Sales Executive	Divorced	102210
849	Male	3	Sales Executive	Married	179240
976	Male	1	Laboratory Technician	Married	87220
987	Male	1	Research Director	Single	95820
1002	Male	1	Sales Representative	Married	104450
1220	Male	2	Laboratory Technician	Divorced	171690
1305	Male	1	Manufacturing Director	Divorced	92500
1352	Male	2	Laboratory Technician	Married	104350
1358	Female	3	Sales Executive	Divorced	103060
1380	Female	1	Manager	Married	174260
1639	Male	3	Laboratory Technician	Single	188800
1693	Male	2	Healthcare Representati	Married	68250
1945	Male	1	Sales Executive	Single	191440
1995	Female	1	Sales Executive	Married	183000
2099	Female	1	Research Scientist	Married	90710
2277	Male	3	Sales Executive	Divorced	102210
2319	Male	3	Sales Executive	Married	179240
2446	Male	1	Laboratory Technician	Married	87220
2457	Male	1	Research Director	Single	95820
2472	Male	1	Sales Representative	Married	104450
2690	Male	2	Laboratory Technician	Divorced	171690
2775	Male	1	Manufacturing Director	Divorced	92500
2822	Male	2	Laboratory Technician	Married	104350
2828	Female	3	Sales Executive	Divorced	103060
2850	Female	1	Manager	Married	174260

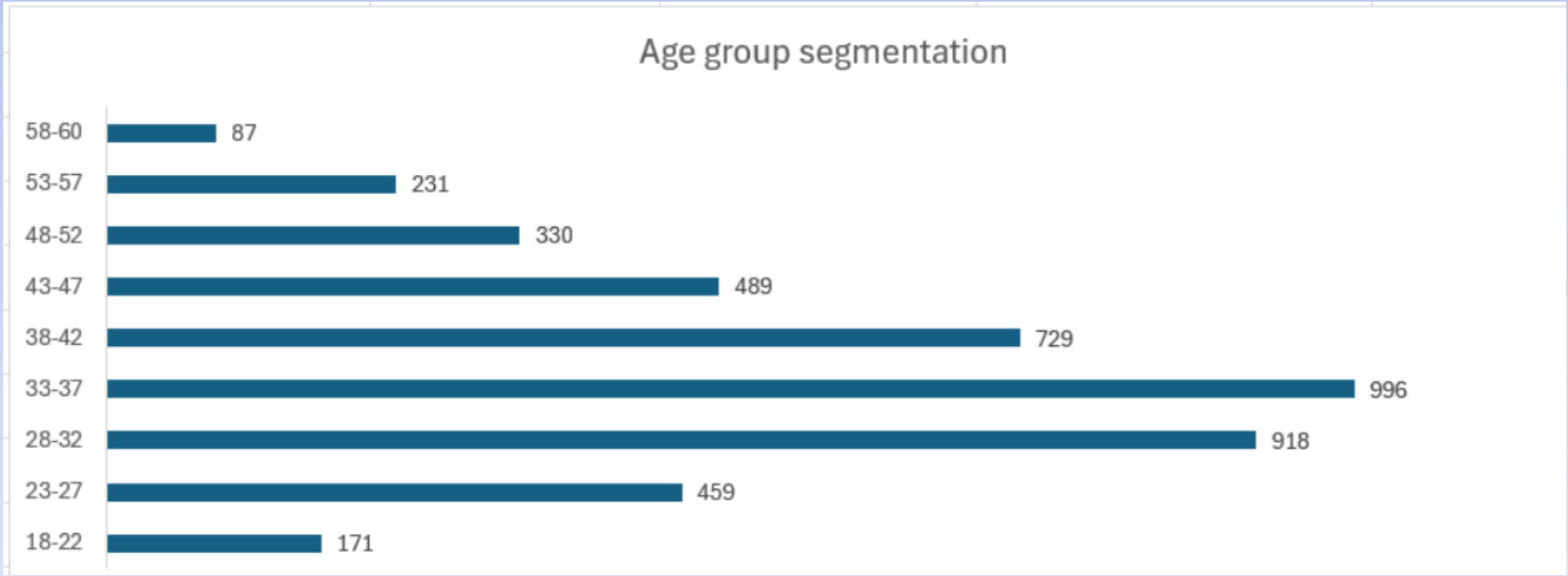
4. Create a bar chart in Excel to visualize the distribution of employee ages.

Select Age Data:

- Highlight the column containing the employee age data.

Insert Bar Chart:

- Group the data according to You
- Go to the "Insert" tab on the Excel ribbon.
- select the desired chart.



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

Step 1: Detect Missing Data

Filter the "Department" column to find blank cells, indicating missing data.

Step 2: Handle Missing Entries

Input correct departments or delete rows with missing data.

Step 3: Review and Standardize

Review unique values, correct inconsistencies, and standardize entries for consistency.

Age	Attrition	BusinessTravel	Department	DistanceFromHome
29	No	Travel_Rarely	Human Resources	8
36	No	Non-Travel	Human Resources	28
33	Yes	Travel_Rarely	Human Resources	28
21	Yes	Travel_Rarely	Human Resources	10
54	No	Travel_Rarely	Human Resources	9
25	No	Travel_Rarely	Human Resources	7
38	No	Travel_Rarely	Human Resources	2
55	Yes	Travel_Rarely	Human Resources	7
55	Yes	Travel_Rarely	Human Resources	16
36	No	Travel_Rarely	Human Resources	2
35	No	Travel_Rarely	Human Resources	1
29	Yes	Travel_Rarely	Human Resources	3
48	No	Travel_Rarely	Human Resources	2
50	No	Non-Travel	Human Resources	2
40	No	Travel_Rarely	Human Resources	2
28	No	Travel_Rarely	Human Resources	1
25	Yes	Travel_Rarely	Human Resources	25
28	No	Travel_Rarely	Human Resources	1
44	Yes	Travel_Rarely	Human Resources	6
24	Yes	Travel_Rarely	Human Resources	28
42	No	Travel_Frequently	Human Resources	8
44	No	Travel_Rarely	Human Resources	4
26	No	Travel_Rarely	Human Resources	9
45	Yes	Travel_Rarely	Human Resources	12
27	No	Travel_Rarely	Human Resources	11
33	No	Travel_Frequently	Human Resources	10
34	No	Travel_Rarely	Human Resources	2
42	No	Travel_Frequently	Human Resources	1
36	No	Travel_Rarely	Human Resources	1
28	Yes	Travel_Rarely	Human Resources	1

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

Insert Pivot Table:

- Select your data range, including headers.
- Go to the "Insert" tab, click "PivotTable," and choose the destination for the pivot table (e.g., a new worksheet).

Design Pivot Table:

- Drag "Department" to the "Columns" area. ▪ Drag "Marital Status" to the "Rows" area.
- Drag any field (e.g., "Employee ID") to the "Values" area, set it to show the count of values.

Departments	Count of Employees
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

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

Select Data:

- Highlight the columns containing "Monthly Income" and "Job Satisfaction" for all employees.
- Ctrl + Click: Use Ctrl + Click to select non-contiguous columns.

Apply Conditional Formatting:

- Navigate to the "Home" tab in Excel.
- Set Formatting: Define formatting options (e.g., color) for cells meeting the criteria.

EmployeeID	Gender	JobLevel	JobRole	MaritalStatus	MonthlyIncome	JobSatisfactio
1	Female	1	Healthcare Representa	Married	131160	4
2	Female	1	Research Scientist	Single	41890	2
3	Male	4	Sales Executive	Married	193280	2
4	Male	3	Human Resources	Married	83210	4
5	Male	1	Sales Executive	Single	23420	1
6	Female	4	Research Director	Married	40710	2
7	Male	2	Sales Executive	Single	58130	3
8	Male	2	Sales Executive	Married	31430	2
9	Male	3	Laboratory Technician	Married	20440	4
10	Female	4	Laboratory Technician	Divorced	134640	1
11	Male	2	Laboratory Technician	Married	79910	4
12	Male	1	Laboratory Technician	Married	33770	4
13	Female	1	Sales Executive	Single	55380	1
14	Male	1	Research Scientist	Married	57620	2
15	Male	1	Manufacturing Director	Married	25920	4
16	Male	2	Healthcare Representa	Married	53460	4
17	Male	1	Laboratory Technician	Single	42130	3
18	Male	2	Sales Executive	Divorced	41270	4
19	Male	1	Sales Representative	Divorced	24380	2
20	Female	1	Manager	Divorced	68700	1
21	Male	2	Laboratory Technician	Divorced	104470	2
22	Male	1	Research Scientist	Divorced	96670	2
23	Female	2	Research Scientist	Married	21480	3
24	Male	1	Manufacturing Director	Married	89260	3
25	Male	1	Laboratory Technician	Single	65130	4
26	Female	1	Research Scientist	Married	67990	4
27	Female	1	Manager	Married	162910	1
28	Male	1	Research Scientist	Single	27050	4
29	Male	2	Research Scientist	Divorced	103330	3
30	Female	1	Manager	Divorced	44480	4

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

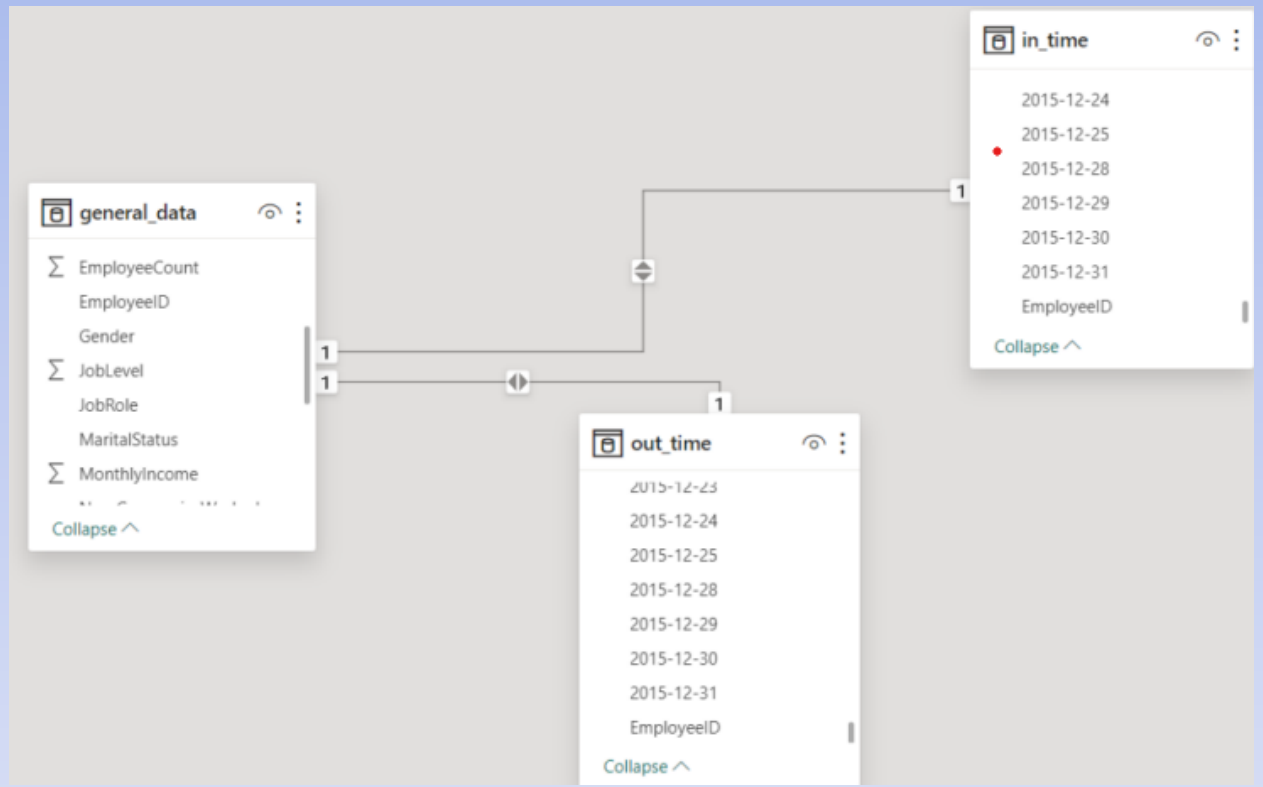
Step 1: Filter the "Job Level" column to display values greater than or equal to 3.

Step 2: Create a PivotTable with "Department" in the Rows and "Monthly Income" in the Values.

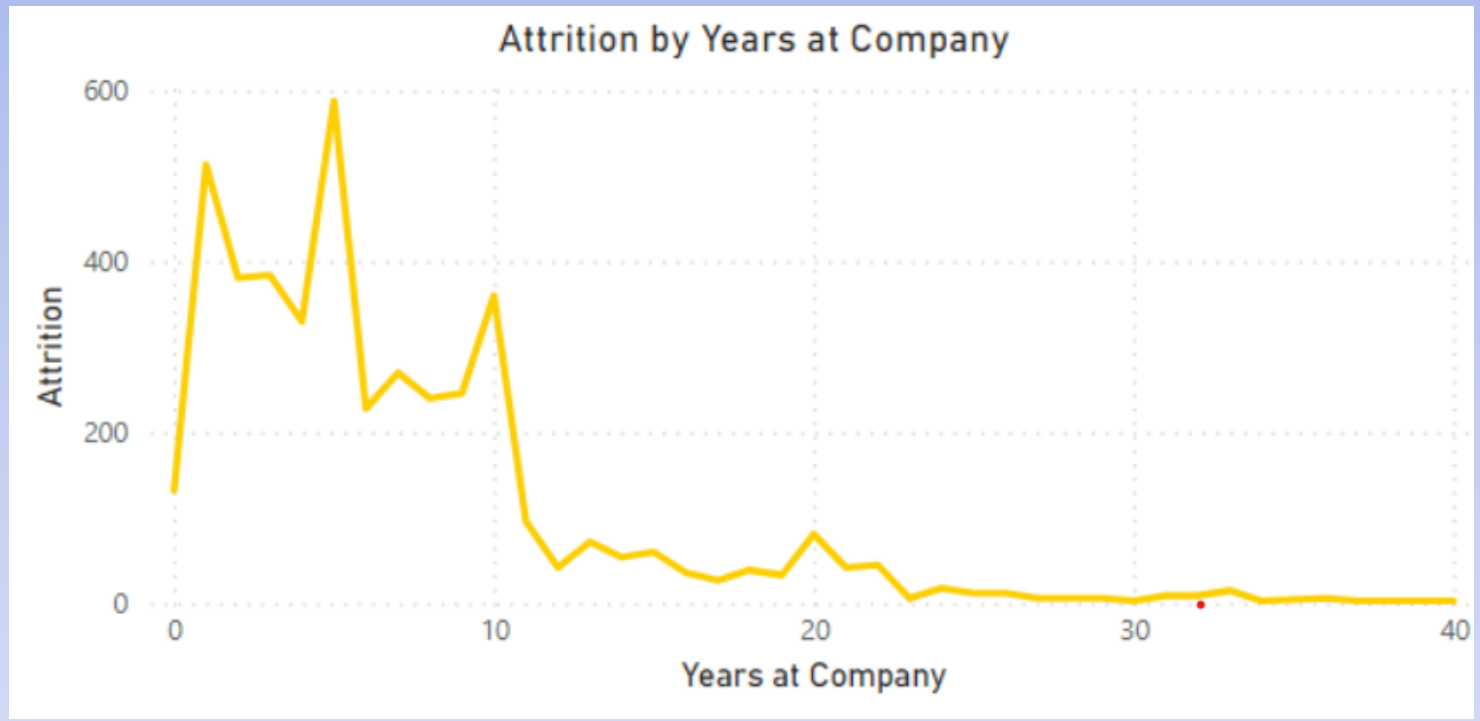
Step 3: Ensure "Monthly Income" is set to summarize by Sum.

Departments	Total Monthly Income
Human Resources	10943940
Research & Development	193702890
Sales	82132440
Grand Total	**286779270**

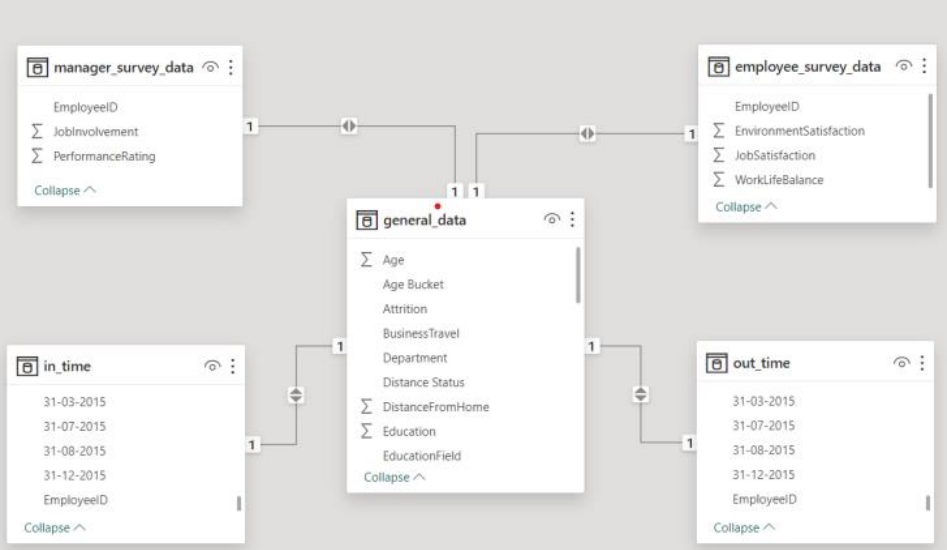
9. In Power BI, establish a relationship between the "EmployeeID" in the employee data and the "EmployeeID" in the time tracking data.



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



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



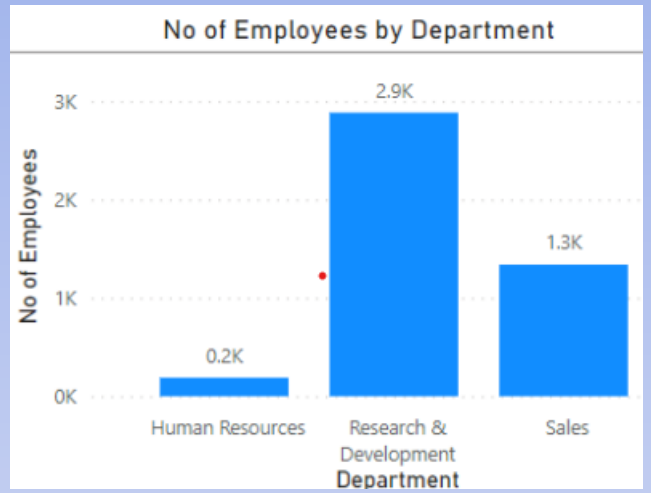
Dimension Tables:
Create 'Employee,' 'Satisfaction,' 'Job Performance,' and 'In-Out Time' Dimension Tables to categorize data.

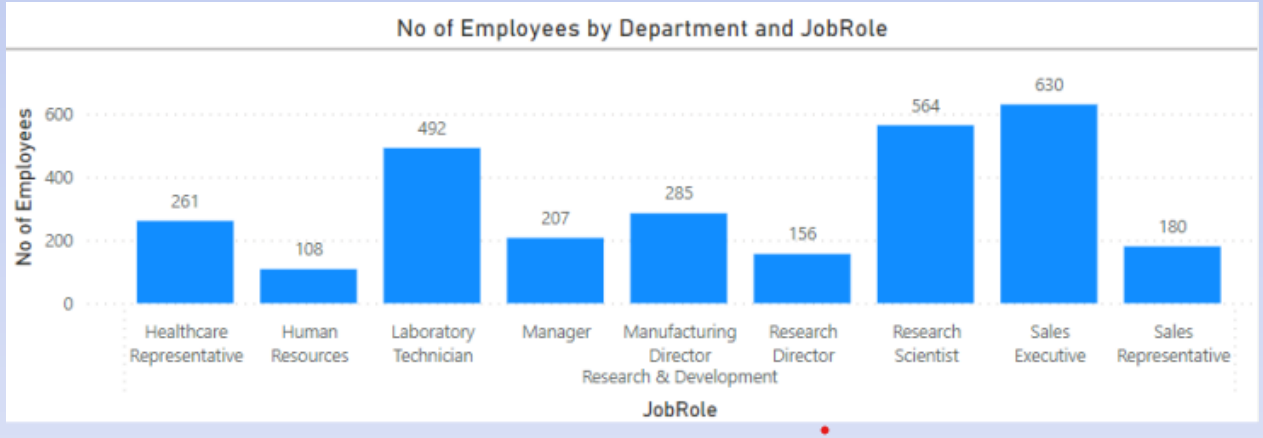
Key Connections:
Ensure each Dimension Table has 'Employee ID' as the primary key, connected to the Fact Tables' foreign key for data coherence.

Data Alignment:
This relational model guarantees precise data correlation, empowering Power BI for accurate analysis.

Benefits of Star Schema:
Offers a clear structure, enhances query performance, streamlines navigation, and enables complex analysis for deeper insights

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

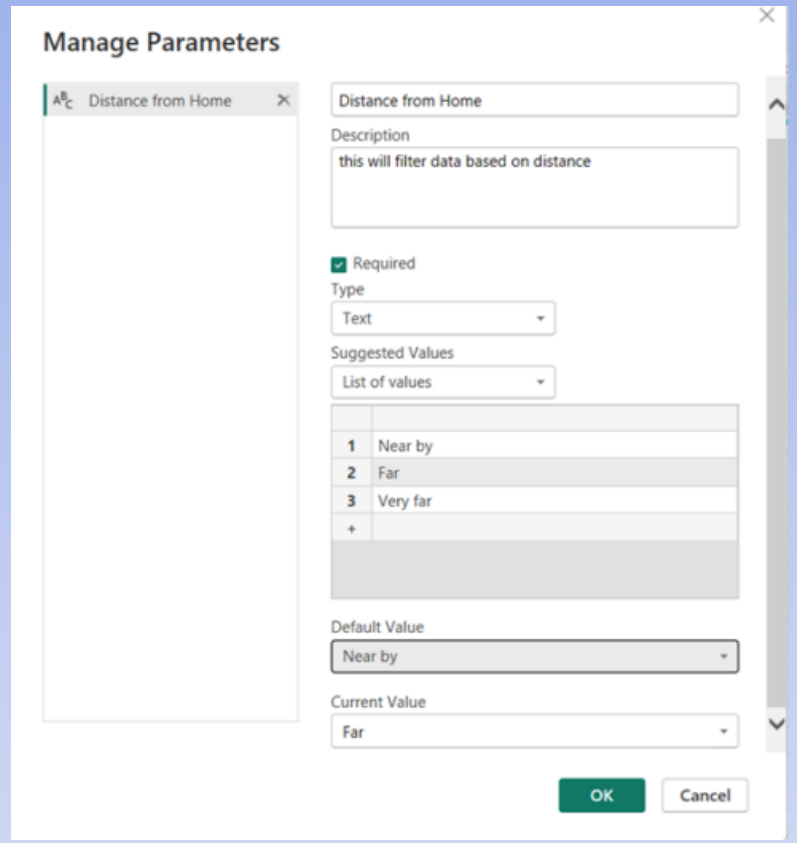




12. Using DAX, calculate the rolling 3-month average of Monthly Income for each employee.

```
Rolling 3-Month Average =
CALCULATE(
    AVERAGE('general_data'[MonthlyIncome]),
    FILTER(
        ALL('DateTable'[Value]),
        'DateTable'[Value] >= MAX('DateTable'[Value]) - 90 &&
        'DateTable'[Value] <= MAX('DateTable'[Value])
    )
)
```

14. How can you set up parameterized queries in Power BI to allow users to filter data based 1 of 2 on the Distance from Home column?



1.Create Parameter in Power Query Editor:

1. Load your data into Power Query Editor.
2. Go to the "Home" tab and click on "Manage Parameters."
3. Create a new parameter named "Distance from home" and set its data type to text.

2.Parameter Configuration:

1. In the parameter settings, configure suggested values to a list obtained from the drop-down.
2. Manually input desired values such as "0-5 miles," "5-10 miles," etc.
3. Establish a default value for user convenience, such as "0-5 miles."

3.Apply Changes:

1. After configuring the parameter, click on "Close & Apply" in the Home tab to implement the modifications.

4.User Interaction in Power BI Report:

1. In your Power BI report, add a slicer visual or an alternative method for users to input parameter values.
2. Manually update the filter condition in your visual to reference the parameter, ensuring seamless user interaction with the data based on their selected distance from home.

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

1.Data Schema Verification:

Ensure dataset conforms to predefined schema by validating actual data against expected structure and rules.

2.Data Refinement:

Reorder columns like "EmployeeID," adjust 'TotalWorkingHours' data type, and handle NA or blank values in 'general_data,' 'employee_survey_data,' and 'manager_survey_data' based on current data status.

]-/

3.Label Correction:

Rectify missing 'EmployeeID' labels in 'in-time' and 'out-time' data entries.

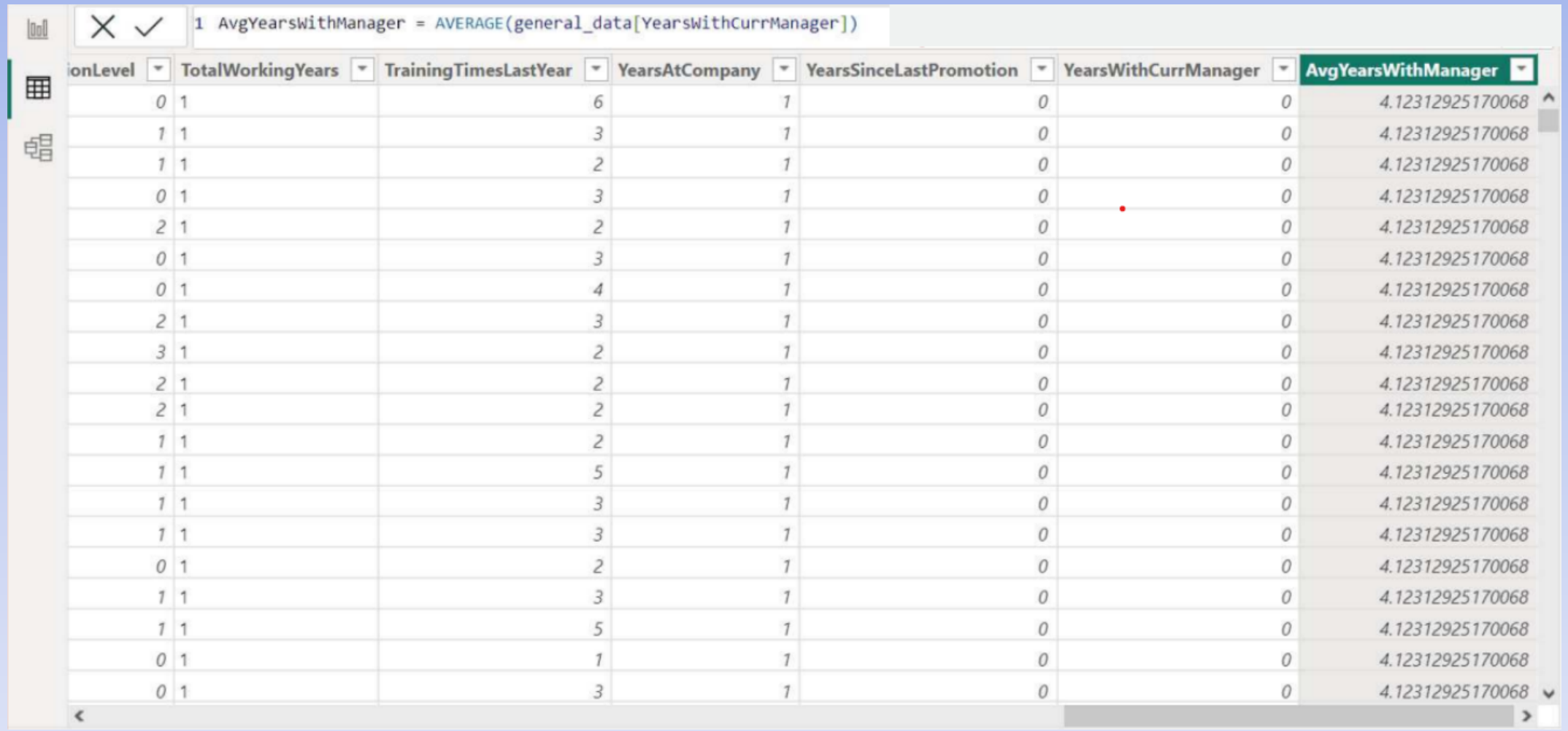
4.Consistency Assurance:

Address data inconsistencies across files, conducting thorough data profiling and validation.

5.Schema Conformance and Quality:

Ensure adherence to predefined schema by resolving identified inconsistencies and implementing necessary data quality measures.

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



```
1 AvgYearsWithManager = AVERAGE(general_data[YearsWithCurrManager])
```

onLevel	TotalWorkingYears	TrainingTimesLastYear	YearsAtCompany	YearsSinceLastPromotion	YearsWithCurrManager	AvgYearsWithManager
0	1	6	1	0	0	4.12312925170068
1	1	3	1	0	0	4.12312925170068
1	1	2	1	0	0	4.12312925170068
0	1	3	1	0	0	4.12312925170068
2	1	2	1	0	0	4.12312925170068
0	1	3	1	0	0	4.12312925170068
0	1	4	1	0	0	4.12312925170068
2	1	3	1	0	0	4.12312925170068
3	1	2	1	0	0	4.12312925170068
2	1	2	1	0	0	4.12312925170068
2	1	2	1	0	0	4.12312925170068
1	1	2	1	0	0	4.12312925170068
1	1	5	1	0	0	4.12312925170068
1	1	3	1	0	0	4.12312925170068
1	1	3	1	0	0	4.12312925170068
0	1	2	1	0	0	4.12312925170068
1	1	3	1	0	0	4.12312925170068
1	1	5	1	0	0	4.12312925170068
0	1	1	1	0	0	4.12312925170068
0	1	3	1	0	0	4.12312925170068

16. 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.

Scenario Summary			
	Current Values:	8% hike on monthly salary	10% hike on monthly salary
Changing Cells:			
E73	₹ 28,67,79,270	₹ 28,67,79,270	₹ 28,67,79,270
E74	2%	8%	10%
Result Cells:			
E76	₹ 29,25,14,855	₹ 30,97,21,612	₹ 31,54,57,197