PSYLIQ

TASK - 3

# HR DATA ANALYSIS USING EXCEL & POWER BI

SHUBHAM KUMAR

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

| Age | Attrition | BusinessTravel | Department | DistanceFromHome | Education | EducationField | EmployeeCount | EmployeeID | Gender | JobLevel | JobRole | MaritalStatus | MonthlyIncome |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 51 | No | Travel_Rarely | Sales | 6 | 2 | Life Sciences | 1 | 1 | Female | 1 | Healthcare Representative | Married | 131160 |
| 31 | Yes | Travel_Frequently | Research & Development | 10 | 1 | Life Sciences | 1 | 2 | Female | 1 | Research Scientist | Single | 41890 |
| 32 | No | Travel_Frequently | Research & Development | 17 | 4 | Other | 1 | 3 | Male | 4 | Sales Executive | Married | 193280 |
| 38 | No | Non-Travel | Research & Development | 2 | 5 | Life Sciences | 1 | 4 | Male | 3 | Human Resources | Married | 83210 |
| 32 | No | Travel_Rarely | Research & Development | 10 | 1 | Medical | 1 | 5 | Male | 1 | Sales Executive | Single | 23420 |
| 46 | No | Travel_Rarely | Research & Development | 8 | 3 | Life Sciences | 1 | 6 | Female | 4 | Research Director | Married | 40710 |
| 31 | No | Travel_Rarely | Research & Development | 1 | 3 | Life Sciences | 1 | 9 | Male | 3 | Laboratory Technician | Married | 20440 |
| 45 | No | Travel_Rarely | Research & Development | 17 | 2 | Medical | 1 | 11 | Male | 2 | Laboratory Technician | Married | 79910 |
| 36 | No | Travel_Rarely | Research & Development | 28 | 1 | Life Sciences | 1 | 12 | Male | 1 | Laboratory Technician | Married | 33770 |
| 55 | No | Travel_Rarely | Research & Development | 14 | 4 | Life Sciences | 1 | 13 | Female | 1 | Sales Executive | Single | 55380 |
| 47 | Yes | Non-Travel | Research & Development | 1 | 1 | Medical | 1 | 14 | Male | 1 | Research Scientist | Married | 57620 |
| 37 | No | Travel_Rarely | Research & Development | 1 | 3 | Life Sciences | 1 | 16 | Male | 2 | Healthcare Representative | Married | 53460 |
| 37 | No | Non-Travel | Research & Development | 1 | 3 | Medical | 1 | 18 | Male | 2 | Sales Executive | Divorced | 41270 |
| 35 | No | Travel_Rarely | Sales | 7 | 4 | Life Sciences | 1 | 19 | Male | 1 | Sales Representative | Divorced | 24380 |
| 38 | No | Travel_Rarely | Research & Development | 8 | 3 | Life Sciences | 1 | 20 | Female | 1 | Manager | Divorced | 68700 |
| 50 | No | Travel_Rarely | Sales | 8 | 4 | Life Sciences | 1 | 22 | Male | 1 | Research Scientist | Divorced | 96670 |
| 53 | No | Travel_Rarely | Research & Development | 11 | 4 | Life Sciences | 1 | 23 | Female | 2 | Research Scientist | Married | 21480 |
| 42 | No | Travel_Rarely | Research & Development | 4 | 4 | Life Sciences | 1 | 24 | Male | 1 | Manufacturing Director | Married | 89260 |
| 55 | No | Travel_Rarely | Research & Development | 1 | 4 | Other | 1 | 26 | Female | 1 | Research Scientist | Married | 67990 |
| 37 | No | Travel_Rarely | Sales | 5 | 1 | Marketing | 1 | 28 | Male | 1 | Research Scientist | Single | 27050 |
| 44 | Yes | Travel_Frequently | Research & Development | 1 | 2 | Medical | 1 | 29 | Male | 2 | Research Scientist | Divorced | 103330 |
| 38 | No | Travel_Rarely | Sales | 2 | 3 | Marketing | 1 | 30 | Female | 1 | Manager | Divorced | 44480 |
| 49 | No | Travel_Frequently | Research & Development | 1 | 1 | Medical | 1 | 33 | Female | 2 | Research Scientist | Single | 35910 |
| 36 | No | Travel_Rarely | Sales | 5 | 3 | Technical Degree | 1 | 34 | Male | 3 | Sales Executive | Single | 54050 |
| 31 | No | Travel_Frequently | Research & Development | 9 | 4 | Medical | 1 | 35 | Male | 1 | Sales Executive | Divorced | 46840 |
| 37 | No | Travel_Frequently | Sales | 9 | 1 | Marketing | 1 | 37 | Male | 1 | Laboratory Technician | Married | 15140 |
| 42 | No | Travel_Frequently | Sales | 2 | 2 | Marketing | 1 | 38 | Male | 3 | Research Director | Married | 29560 |
| 35 | No | Travel_Rarely | Sales | 20 | 2 | Life Sciences | 1 | 40 | Male | 1 | Laboratory Technician | Married | 51540 |

general_data | manager_survey_data | employee_survey_data | out_time | in_time

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

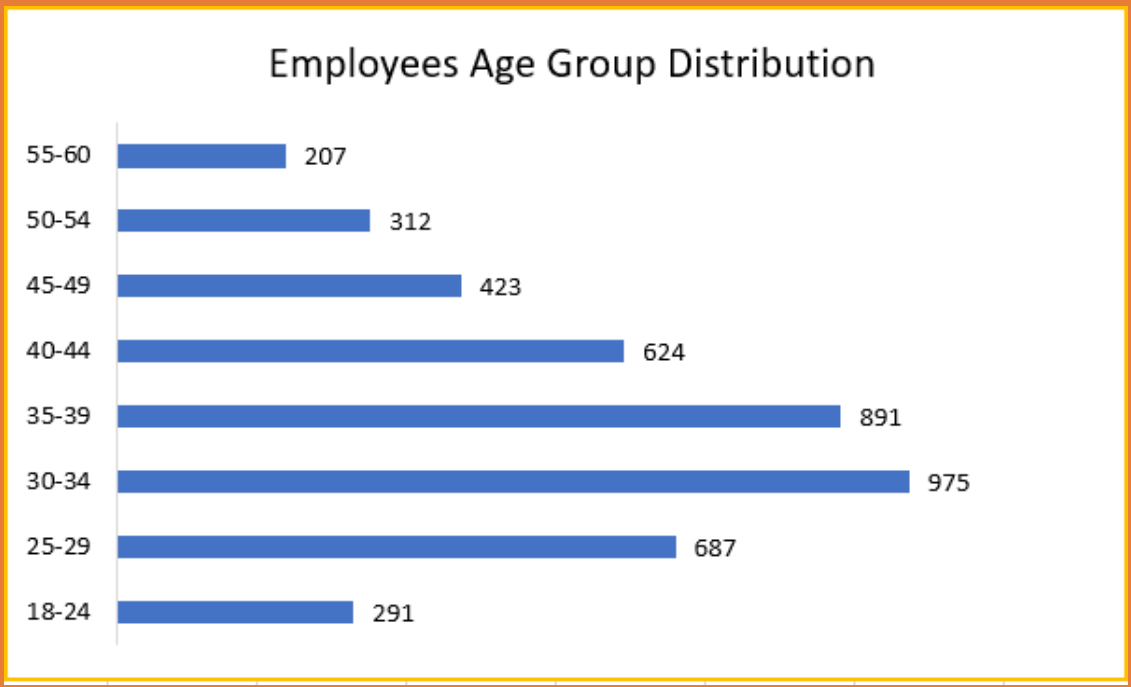
| Job Role | Average Monthly Income |
|---|---|
| Healthcare Representative | 60983.74 |
| Human Resources | 58528.08 |
| Laboratory Technician | 66314.05 |
| Manager | 63395.88 |
| Manufacturing Director | 69183.72 |
| Research Director | 65473.13 |
| Research Scientist | 64975.68 |
| Sales Executive | 65186.69 |
| Sales Representative | 65370.96 |
| **Grand Total** | **65029.31** |

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

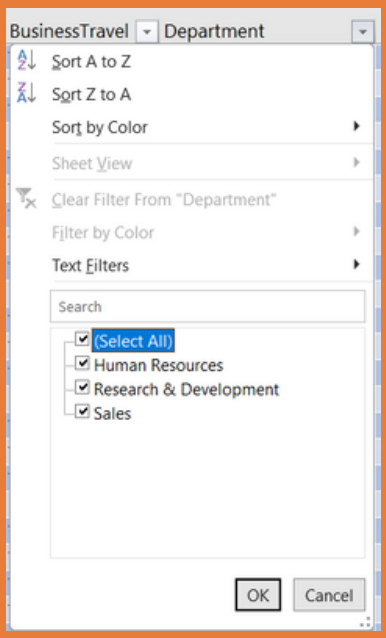
| BusinessTravel | Department | DistanceFromHome | Education | EducationField | EmployeeCount | EmployeeID | Gender | JobLevel | JobRole | MaritalStatus | MonthlyIncome | NumCompaniesWor |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Travel_Rarely | Sales | 6 | 2 | Life Sciences | 1 | 1 | Female | 1 | Healthcare Representative | Married | 131160 | |
| Travel_Frequently | Research & Development | 10 | 1 | Life Sciences | 1 | 2 | Female | 1 | Research Scientist | Single | 41890 | |
| Travel_Frequently | Research & Development | 17 | 4 | Other | 1 | 3 | Male | 4 | Sales Executive | Married | 193280 | |
| Non-Travel | Research & Development | 2 | 5 | Life Sciences | 1 | 4 | Male | 3 | Human Resources | Married | 83210 | |
| Travel_Rarely | Research & Development | 10 | 1 | Medical | 1 | 5 | Male | 1 | Sales Executive | Single | 23420 | |
| Travel_Rarely | Research & Development | 8 | 3 | Life Sciences | 1 | 6 | Female | 4 | Research Director | Married | 40710 | |
| Travel_Rarely | Research & Development | 11 | 2 | Medical | 1 | 7 | Male | 2 | Sales Executive | Single | 58130 | |
| Travel_Rarely | Research & Development | 18 | 3 | Life Sciences | 1 | 8 | Male | 2 | Sales Executive | Married | 31430 | |
| Travel_Rarely | Research & Development | 1 | 3 | Life Sciences | 1 | 9 | Male | 3 | Laboratory Technician | Married | 20440 | |
| Non-Travel | Research & Development | 7 | 4 | Medical | 1 | 10 | Female | 4 | Laboratory Technician | Divorced | 134640 | |
| Travel_Rarely | Research & Development | 17 | 2 | Medical | 1 | 11 | Male | 2 | Laboratory Technician | Married | 79910 | |
| Travel_Rarely | Research & Development | 28 | 1 | Life Sciences | 1 | 12 | Male | 1 | Laboratory Technician | Married | 33770 | |
| Travel_Rarely | Research & Development | 14 | 4 | Life Sciences | 1 | 13 | Female | 1 | Sales Executive | Single | 55380 | |
| Non-Travel | Research & Development | 1 | 1 | Medical | 1 | 14 | Male | 1 | Research Scientist | Married | 57620 | |
| Travel_Rarely | Research & Development | 1 | 3 | Life Sciences | 1 | 15 | Male | 1 | Manufacturing Director | Married | 25920 | |
| Travel_Rarely | Research & Development | 1 | 3 | Life Sciences | 1 | 16 | Male | 2 | Healthcare Representative | Married | 53460 | |
| Travel_Rarely | Research & Development | 3 | 2 | Life Sciences | 1 | 17 | Male | 1 | Laboratory Technician | Single | 42130 | |
| Non-Travel | Research & Development | 1 | 3 | Medical | 1 | 18 | Male | 2 | Sales Executive | Divorced | 41270 | |
| Travel_Rarely | Sales | 7 | 4 | Life Sciences | 1 | 19 | Male | 1 | Sales Representative | Divorced | 24380 | |
| Travel_Rarely | Research & Development | 8 | 3 | Life Sciences | 1 | 20 | Female | 1 | Manager | Divorced | 68700 | |
| Travel_Frequently | Research & Development | 1 | 4 | Other | 1 | 21 | Male | 2 | Laboratory Technician | Divorced | 104470 | |
| Travel_Rarely | Sales | 8 | 4 | Life Sciences | 1 | 22 | Male | 1 | Research Scientist | Divorced | 96670 | |
| Travel_Rarely | Research & Development | 11 | 4 | Life Sciences | 1 | 23 | Female | 2 | Research Scientist | Married | 21480 | |
| Travel_Rarely | Research & Development | 4 | 4 | Life Sciences | 1 | 24 | Male | 1 | Manufacturing Director | Married | 89260 | |
| Travel_Frequently | Research & Development | 16 | 4 | Medical | 1 | 25 | Male | 1 | Laboratory Technician | Single | 65130 | |
| Travel_Rarely | Research & Development | 1 | 4 | Other | 1 | 26 | Female | 1 | Research Scientist | Married | 67990 | |
| Travel_Frequently | Research & Development | 9 | 3 | Life Sciences | 1 | 27 | Female | 1 | Manager | Married | 162910 | |
| Travel_Rarely | Sales | 5 | 1 | Marketing | 1 | 28 | Male | 1 | Research Scientist | Single | 27050 | |

Q-2 | general_data_worksheet | employee_survey_data | manager_survey_data | in_time | out_time

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

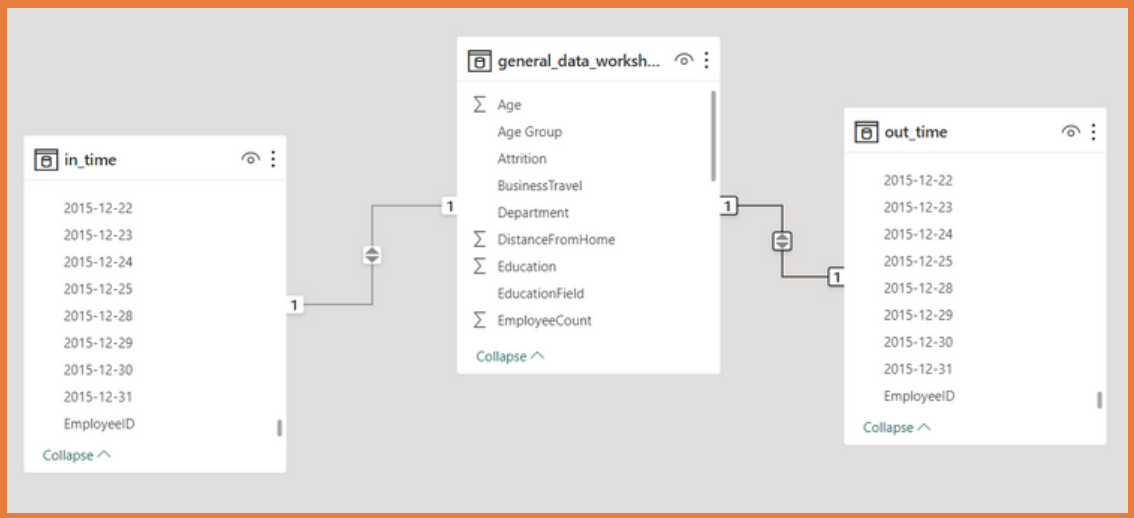


# 5. Identify and clean any missing or inconsistent data in the "Department" column.

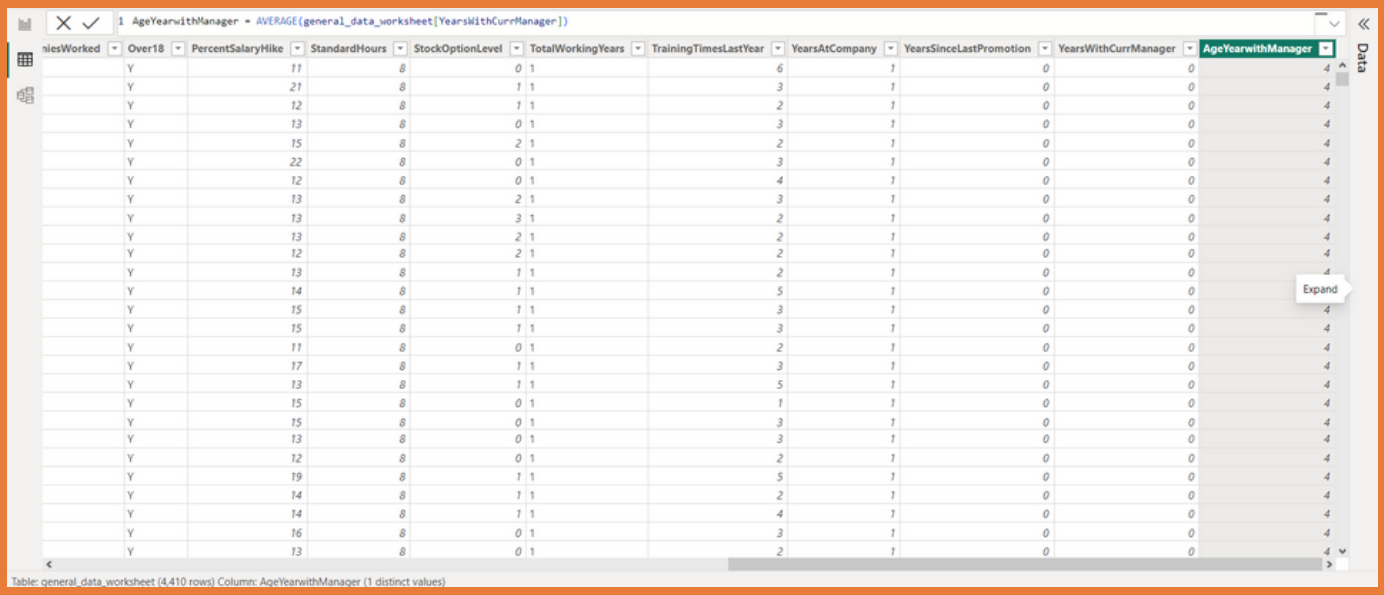



**Note:- There is not any missing or inconsistent data in the “Department” Column**

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



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



```
AgeYearwithManager = AVERAGE(general_data_worksheet[YearsWithCurrManager])
```

| niesWorked | Over18 | PercentSalaryHike | StandardHours | StockOptionLevel | TotalWorkingYears | TrainingTimesLastYear | YearsAtCompany | YearsSinceLastPromotion | YearsWithCurrManager | AgeYearwithManager |
|---|---|---|---|---|---|---|---|---|---|---|
| | Y | 11 | 8 | 0 | 1 | 6 | 1 | 0 | 0 | 4 |
| | Y | 21 | 8 | 1 | 1 | 3 | 1 | 0 | 0 | 4 |
| | Y | 12 | 8 | 1 | 1 | 2 | 1 | 0 | 0 | 4 |
| | Y | 13 | 8 | 0 | 1 | 3 | 1 | 0 | 0 | 4 |
| | Y | 15 | 8 | 2 | 1 | 2 | 1 | 0 | 0 | 4 |
| | Y | 22 | 8 | 0 | 1 | 3 | 1 | 0 | 0 | 4 |
| | Y | 12 | 8 | 0 | 1 | 4 | 1 | 0 | 0 | 4 |
| | Y | 13 | 8 | 2 | 1 | 3 | 1 | 0 | 0 | 4 |
| | Y | 13 | 8 | 3 | 1 | 2 | 1 | 0 | 0 | 4 |
| | Y | 13 | 8 | 2 | 1 | 2 | 1 | 0 | 0 | 4 |
| | Y | 12 | 8 | 2 | 1 | 2 | 1 | 0 | 0 | 4 |
| | Y | 13 | 8 | 1 | 1 | 2 | 1 | 0 | 0 | 4 |
| | Y | 14 | 8 | 1 | 1 | 5 | 1 | 0 | 0 | 4 |
| | Y | 15 | 8 | 1 | 1 | 3 | 1 | 0 | 0 | 4 |
| | Y | 15 | 8 | 1 | 1 | 3 | 1 | 0 | 0 | 4 |
| | Y | 11 | 8 | 0 | 1 | 2 | 1 | 0 | 0 | 4 |
| | Y | 17 | 8 | 1 | 1 | 3 | 1 | 0 | 0 | 4 |
| | Y | 13 | 8 | 1 | 1 | 5 | 1 | 0 | 0 | 4 |
| | Y | 15 | 8 | 0 | 1 | 1 | 1 | 0 | 0 | 4 |
| | Y | 15 | 8 | 0 | 1 | 3 | 1 | 0 | 0 | 4 |
| | Y | 13 | 8 | 0 | 1 | 3 | 1 | 0 | 0 | 4 |
| | Y | 12 | 8 | 0 | 1 | 2 | 1 | 0 | 0 | 4 |
| | Y | 19 | 8 | 1 | 1 | 5 | 1 | 0 | 0 | 4 |
| | Y | 14 | 8 | 1 | 1 | 2 | 1 | 0 | 0 | 4 |
| | Y | 14 | 8 | 1 | 1 | 4 | 1 | 0 | 0 | 4 |
| | Y | 16 | 8 | 0 | 1 | 3 | 1 | 0 | 0 | 4 |
| | Y | 13 | 8 | 0 | 1 | 2 | 1 | 0 | 0 | 4 |

Table: general_data_worksheet (4,410 rows) Column: AgeYearwithManager (1 distinct values)

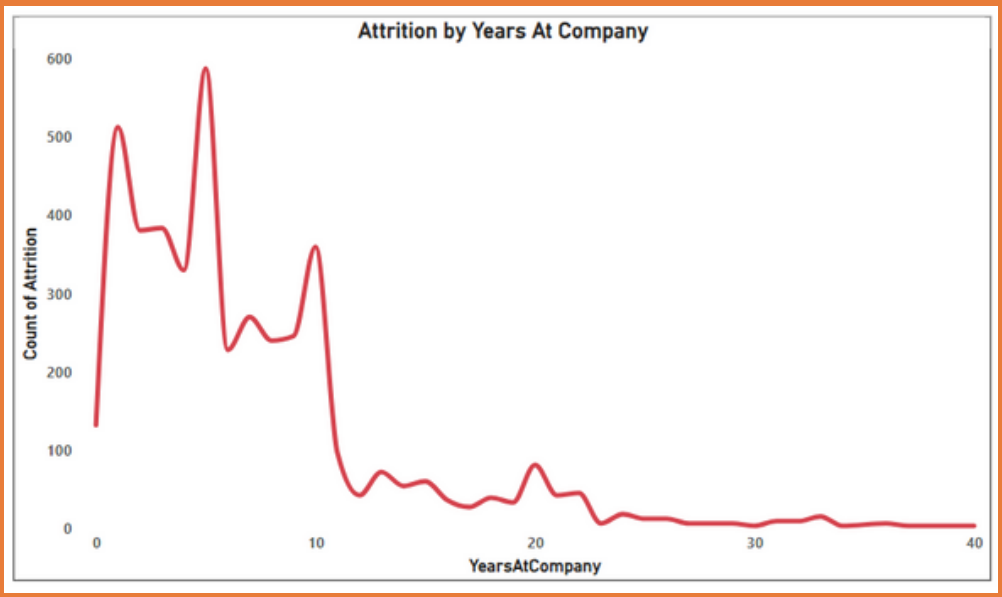
# 8. Using Excel, create a pivot table that displays the count of employees in each Marital Status category, segmented by Department.

| Marital Status By Department | Count of Employees |
|---|---|
| **Human Resources** | **189** |
| Divorced | 21 |
| Married | 96 |
| Single | 72 |
| **Research & Development** | **2883** |
| Divorced | 621 |
| Married | 1350 |
| Single | 912 |
| **Sales** | **1338** |
| Divorced | 339 |
| Married | 573 |
| Single | 426 |
| **Grand Total** | **4410** |

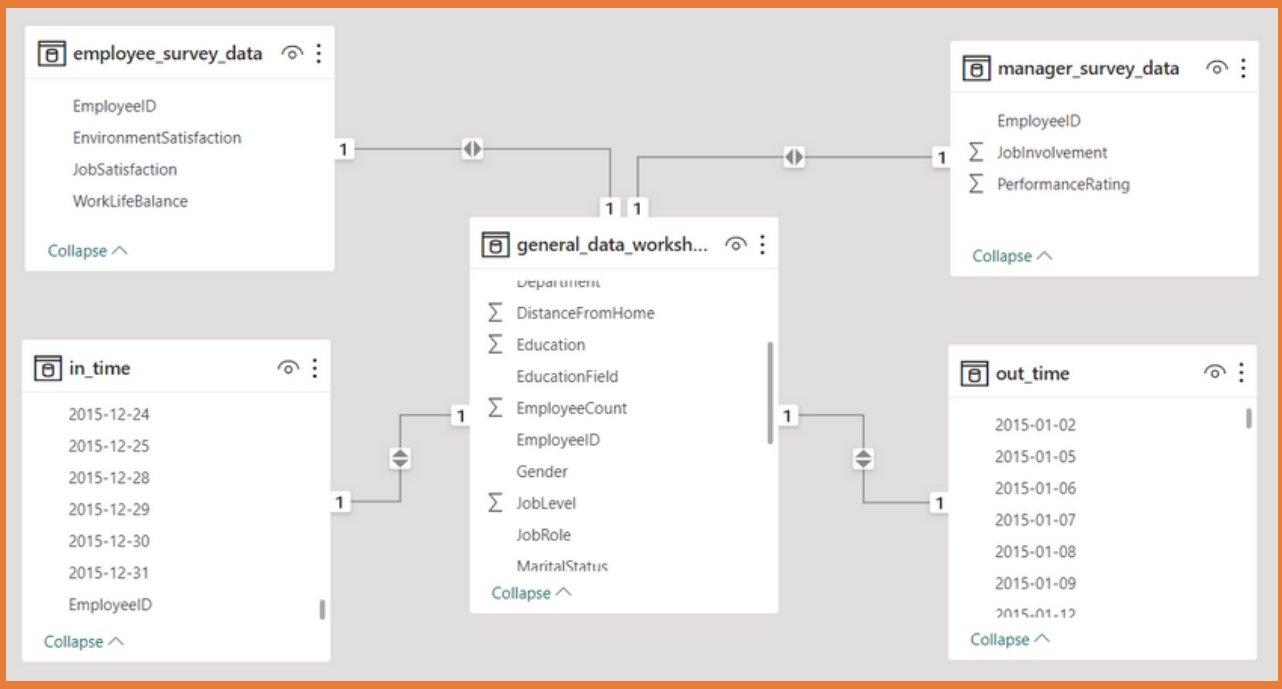
## 9. Apply conditional formatting to highlight employees with both above-average Monthly Income and above-average Job Satisfaction.

| EmployeeID | MonthlyIncome | Job Satisfacti |
|---|---|---|
| 1 | 131160 | 4 |
| 2 | 41890 | 2 |
| 3 | 193280 | 2 |
| 4 | 83210 | 4 |
| 5 | 23420 | 1 |
| 6 | 40710 | 2 |
| 7 | 58130 | 3 |
| 8 | 31430 | 2 |
| 9 | 20440 | 4 |
| 10 | 134640 | 1 |
| 11 | 79910 | 4 |
| 12 | 33770 | 4 |
| 13 | 55380 | 1 |
| 14 | 57620 | 2 |
| 15 | 25920 | 4 |
| 16 | 53460 | 4 |
| 17 | 42130 | 3 |
| 18 | 41270 | 4 |
| 19 | 24380 | 2 |
| 20 | 68700 | 1 |
| 21 | 104470 | 2 |

## 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.




## Steps:-

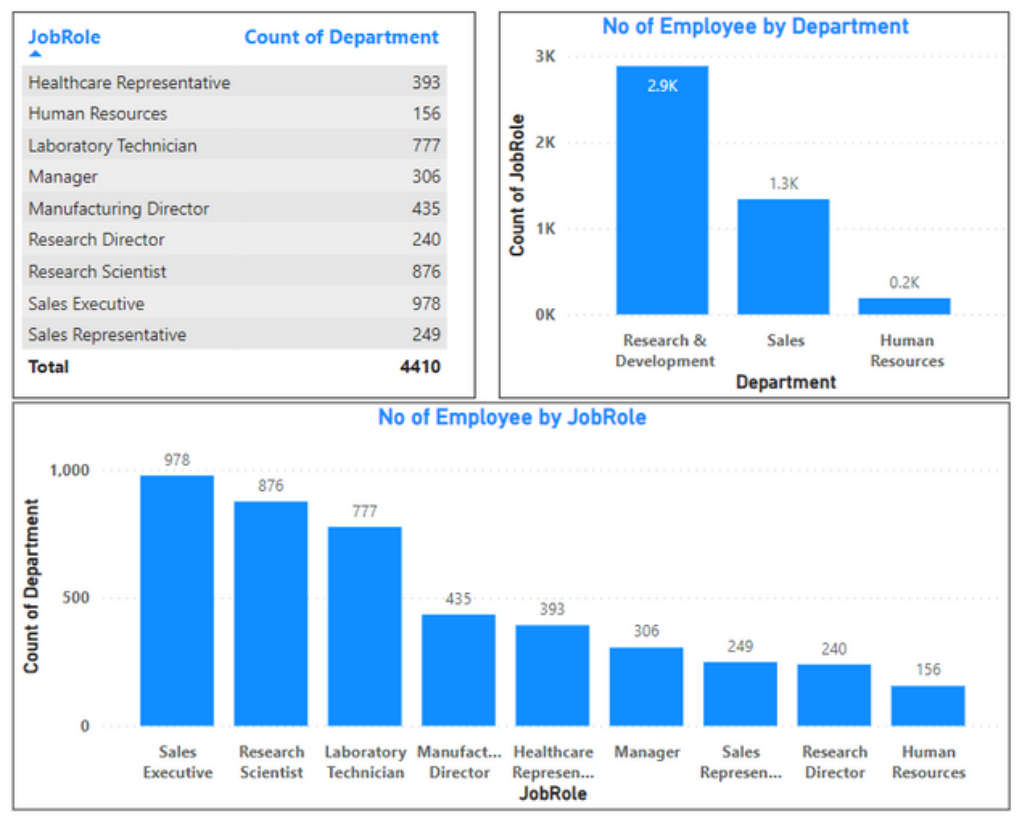
- Normalize the table e by dividing it into multiple tables. These tables serve as the main focus point of the star schema.
- Create a dimension table and ensure that all the dimension tables have a primary key and are connected to the foreign key.
- This relationship structure ensures accurate data correlation, enabling Power BI to perform precise analysis and reporting.

## Benefits:-

- Star schema is easy to understand and query, making it an ideal choice for data warehousing and business intelligence applications.
- it is a scalable and flexible solution for data warehousing.
- The denormalized structure of the star schema reduces the need for complex joins, resulting in faster query performance.
- Star schema is widely supported by business intelligence tools, making it easier to integrateion

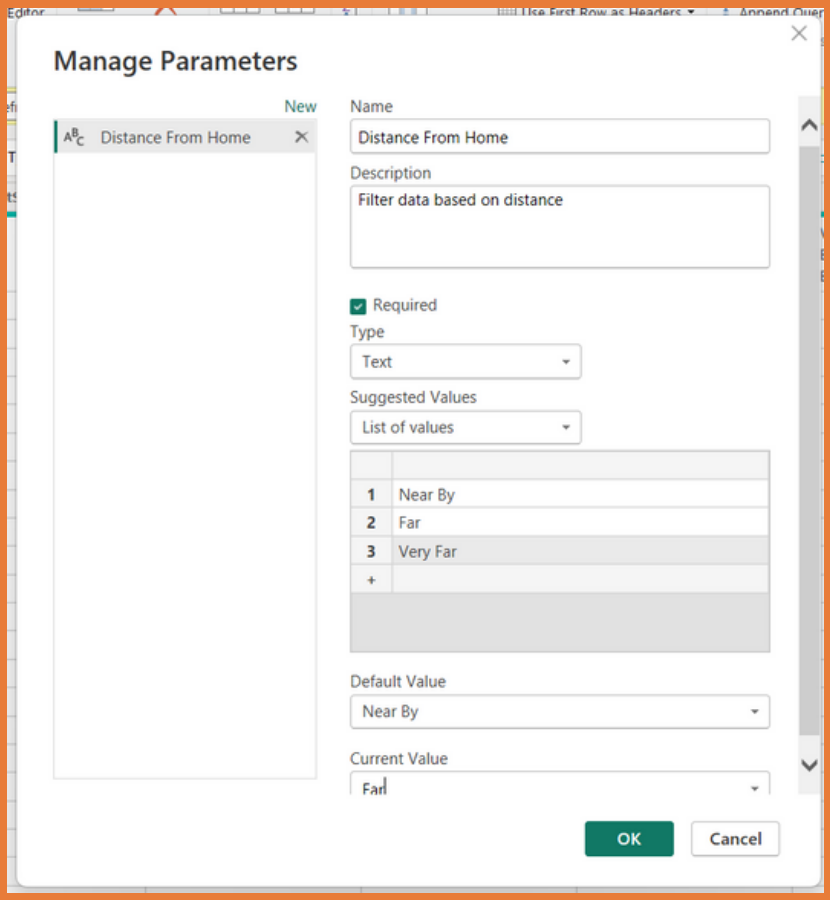
## 12. Using DAX, calculate the rolling 3-month average of Monthly Income for each employee.

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



| JobRole | Count of Department |
|---|---|
| Healthcare Representative | 393 |
| Human Resources | 156 |
| Laboratory Technician | 777 |
| Manager | 306 |
| Manufacturing Director | 435 |
| Research Director | 240 |
| Research Scientist | 876 |
| Sales Executive | 978 |
| Sales Representative | 249 |
| **Total** | **4410** |

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




- **Open Power BI Desktop**
- **Click Get Data**
- **Select “Blank Query” from “Other” then click “Connect”**
- **In the Query Editor window click “Manage Parameters” from the ribbon**
- ***Click “New”***
- ***Type a name for the parameter***
- ***You can also write a description***
- ***Select Type as Text***
- ***From “Allowed Values” select “List of Values”.***
- ***Fill the list with some valid values.***
- ***Select a “Default Value” and “Current Value” then click OK***

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

| Sum of MonthlyIncome | Column Labels | | | |
|---|---|---|---|---|
| Row Labels | Human Resources | Research & Development | Sales | Grand Total |
| 3 | ₹ 16,48,500 | ₹ 2,81,17,740 | ₹ 1,17,92,400 | ₹ 4,15,58,640 |
| 4 | ₹ 7,54,800 | ₹ 1,52,77,290 | ₹ 87,53,070 | ₹ 2,47,85,160 |
| 5 | ₹ 8,55,840 | ₹ 1,01,07,870 | ₹ 24,28,860 | ₹ 1,33,92,570 |
| **Grand Total** | **₹ 32,59,140** | **₹ 5,35,02,900** | **₹ 2,29,74,330** | **₹ 7,97,36,370** |

## 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.

| A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|
| Monthly Income | Salary Hike | Hike_amount | New_amount | | | |
| 131160 | 10% | 13116 | 144276 | | | |
| 41890 | 10% | 4189 | 46079 | | | |
| 193280 | 10% | 19328 | 212608 | | | |
| 83210 | 10% | 8321 | 91531 | | | |
| 23420 | 10% | 2342 | 25762 | | | |
| 40710 | 10% | 4071 | 44781 | | | |
| 58130 | 10% | 5813 | 63943 | | | |
| 31430 | 10% | 3143 | 34573 | | | |
| 20440 | 10% | 2044 | 22484 | | | |
| 134640 | 10% | 13464 | 148104 | | | |
| 79910 | 10% | 7991 | 87901 | | | |
| 33770 | 10% | 3377 | 37147 | | | |
| 55380 | 10% | 5538 | 60918 | | | |
| 57620 | 10% | 5762 | 63382 | | | |
| 25920 | 10% | 2592 | 28512 | | | |
| 53460 | 10% | 5346 | 58806 | | | |
| 42130 | 10% | 4213 | 46343 | | | |
| 41270 | 10% | 4127 | 45397 | | | |
| 24380 | 10% | 2438 | 26818 | | | |
| 68700 | 10% | 6870 | 75570 | | | |
| 104470 | 10% | 10447 | 114917 | | | |
| 96670 | 10% | 9667 | 106337 | | | |
| 21480 | 10% | 2148 | 23628 | | | |
| 89260 | 10% | 8926 | 98186 | | | |
| 65130 | 10% | 6513 | 71643 | | | |
| 67990 | 10% | 6799 | 74789 | | | |
| 162910 | 10% | 16291 | 179201 | | | |
| 27050 | 10% | 2705 | 29755 | | | |

Q-2 | Q-4 | Q-8 | Q-15 | general_data_worksheet | Q-16 | employee

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

Data Integrity and Consistency:

- Address inconsistencies in data files, ensuring meticulous data profiling and accurate validation.

Label Correction

- Correct missing columns labels in the data sets.

Improved Data Quality:

- It helps ensure that data meets specific criteria, such as valid values, proper data types, or required fields.

Reduced Integration Issues:

- By validating the data against a predefined schema, you can catch compatibility issues, data mismatches, or inconsistencies early on.

Data Adherence Verification:

- Verify if the dataset aligns with the predefined schema by checking if the actual data meets the expected structure and rules.

THANK YOU!

SHUBHAM KUMAR