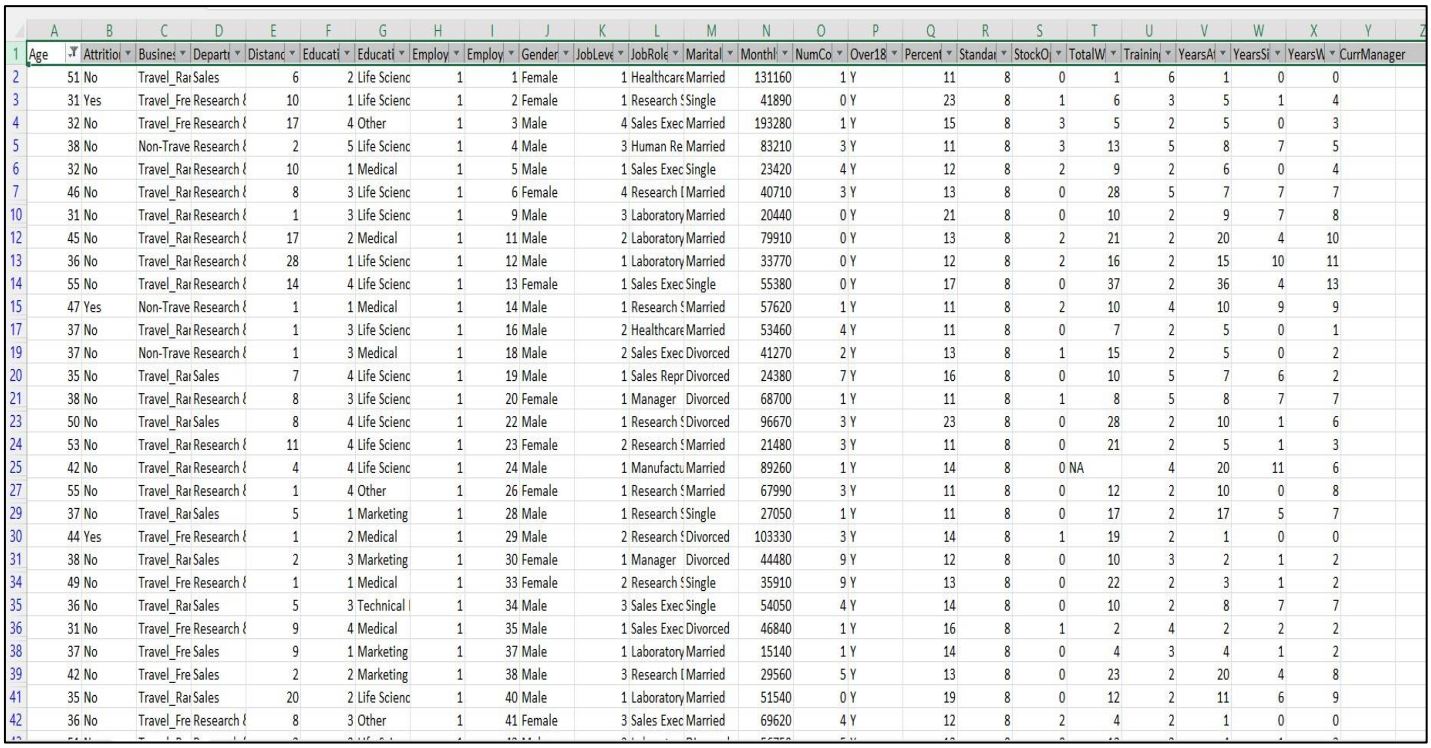
Name: **Jitha P Nair**
Email: jithapnair20@gmail.com
LinkedIn: https://www.linkedin.com/in/jithapnair/

HR Analysis Insights – Task 3

Abstract

HR analysis offers a robust framework for decision-makers, HR professionals, and organizational leaders to make informed, data-driven decisions. By harnessing the power of analytics, organizations can elevate their HR practices, foster a culture of continuous improvement, and strategically position themselves for sustained success in an ever-evolving business landscape.

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

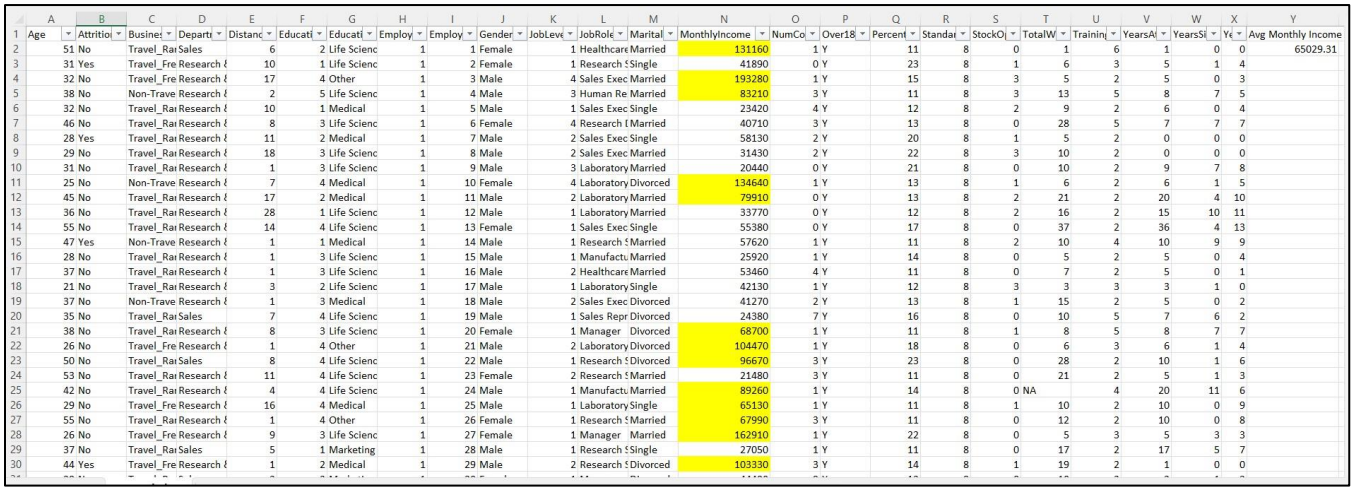


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	Age	Attritio	Busines	Departi	Distanc	Educati	Educati	Employ	Employ	Gender	JobLeve	JobRole	Marital	Monthl	NumCo	Over18	Percent	Standar	StockO	TotalW	Trainin	YearsA	YearsSi	YearsW	CurrManager
2	51	No	Travel_Rar	Sales	6	2	Life Scienc	1	1	Female	1	Healthcare	Married	131160	1	Y	11	8	0	1	6	1	0	0	
3	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	4	
4	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	3	
5	38	No	Non-Trave	Research &	2	5	Life Scienc	1	4	Male	3	Human Re	Married	83210	3	Y	11	8	3	13	5	8	7	5	
6	32	No	Travel_Rar	Research &	10	1	Medical	1	5	Male	1	Sales Exec	Single	23420	4	Y	12	8	2	9	2	6	0	4	
7	46	No	Travel_Rar	Research &	8	3	Life Scienc	1	6	Female	4	Research I	Married	40710	3	Y	13	8	0	28	5	7	7	7	
10	31	No	Travel_Rar	Research &	1	3	Life Scienc	1	9	Male	3	Laboratory	Married	20440	0	Y	21	8	0	10	2	9	7	8	
12	45	No	Travel_Rar	Research &	17	2	Medical	1	11	Male	2	Laboratory	Married	79910	0	Y	13	8	2	21	2	20	4	10	
13	36	No	Travel_Rar	Research &	28	1	Life Scienc	1	12	Male	1	Laboratory	Married	33770	0	Y	12	8	2	16	2	15	10	11	
14	55	No	Travel_Rar	Research &	14	4	Life Scienc	1	13	Female	1	Sales Exec	Single	55380	0	Y	17	8	0	37	2	36	4	13	
15	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	9	
17	37	No	Travel_Rar	Research &	1	3	Life Scienc	1	16	Male	2	Healthcare	Married	53460	4	Y	11	8	0	7	2	5	0	1	
19	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	2	
20	35	No	Travel_Rar	Sales	7	4	Life Scienc	1	19	Male	1	Sales Repr	Divorced	24380	7	Y	16	8	0	10	5	7	6	2	
21	38	No	Travel_Rar	Research &	8	3	Life Scienc	1	20	Female	1	Manager	Divorced	68700	1	Y	11	8	1	8	5	8	7	7	
23	50	No	Travel_Rar	Sales	8	4	Life Scienc	1	22	Male	1	Research S	Divorced	96670	3	Y	23	8	0	28	2	10	1	6	
24	53	No	Travel_Rar	Research &	11	4	Life Scienc	1	23	Female	2	Research S	Married	21480	3	Y	11	8	0	21	2	5	1	3	
25	42	No	Travel_Rar	Research &	4	4	Life Scienc	1	24	Male	1	Manufactu	Married	89260	1	Y	14	8	0	NA	4	20	11	6	
27	55	No	Travel_Rar	Research &	1	4	Other	1	26	Female	1	Research S	Married	67990	3	Y	11	8	0	12	2	10	0	8	
29	37	No	Travel_Rar	Sales	5	1	Marketing	1	28	Male	1	Research S	Single	27050	1	Y	11	8	0	17	2	17	5	7	
30	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	0	
31	38	No	Travel_Rar	Sales	2	3	Marketing	1	30	Female	1	Manager	Divorced	44480	9	Y	12	8	0	10	3	2	1	2	
34	49	No	Travel_Fre	Research &	1	1	Medical	1	33	Female	2	Research S	Single	35910	9	Y	13	8	0	22	2	3	1	2	
35	36	No	Travel_Rar	Sales	5	3	Technical	1	34	Male	3	Sales Exec	Single	54050	4	Y	14	8	0	10	2	8	7	7	
36	31	No	Travel_Fre	Research &	9	4	Medical	1	35	Male	1	Sales Exec	Divorced	46840	1	Y	16	8	1	2	4	2	2	2	
38	37	No	Travel_Fre	Sales	9	1	Marketing	1	37	Male	1	Laboratory	Married	15140	1	Y	14	8	0	4	3	4	1	2	
39	42	No	Travel_Fre	Sales	2	2	Marketing	1	38	Male	3	Research I	Married	29560	5	Y	13	8	0	23	2	20	4	8	
41	35	No	Travel_Rar	Sales	20	2	Life Scienc	1	40	Male	1	Laboratory	Married	51540	0	Y	19	8	0	12	2	11	6	9	
42	36	No	Travel_Fre	Research &	8	3	Other	1	41	Female	3	Sales Exec	Married	69620	4	Y	12	8	2	4	2	1	0	0	

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

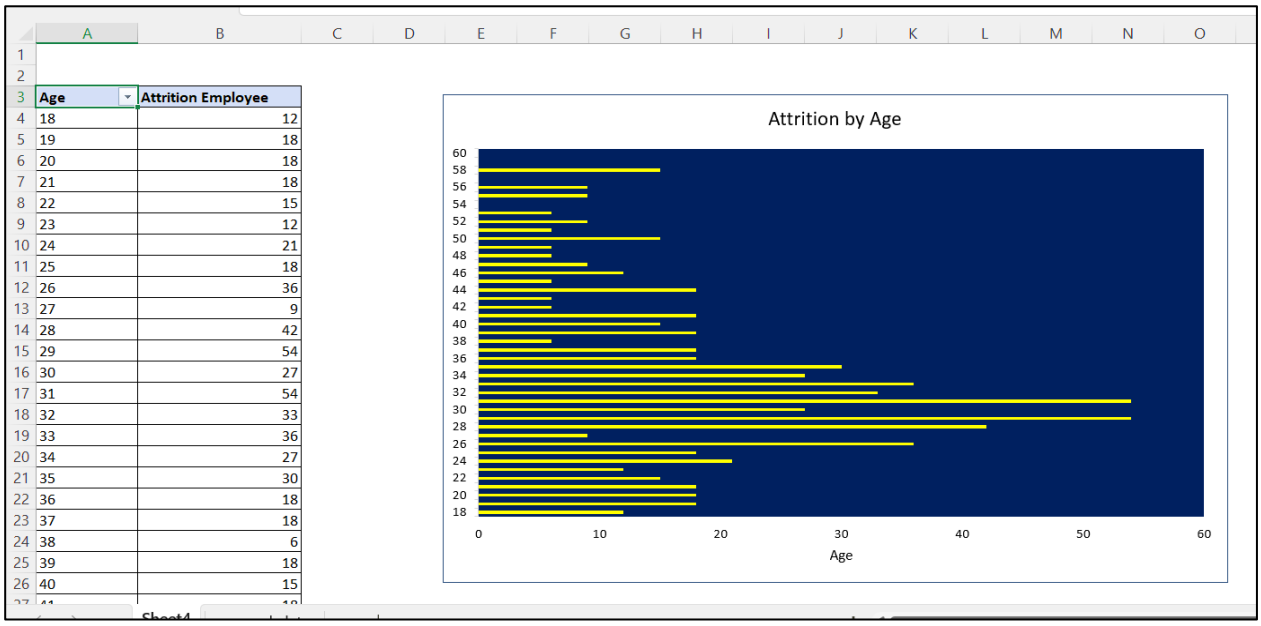
JobRole	Avg Monthly Income
Manufacturing Director	$69,183.72
Laboratory Technician	$66,314.05
Research Director	$65,473.13
Sales Representative	$65,370.96
Sales Executive	$65,186.69
Research Scientist	$64,975.68
Manager	$63,395.88
Healthcare Representative	$60,983.74
Human Resources	$58,528.08
Total	**$65,029.31**

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



Age	Attrition	Business	Department	Distance	Education	Education	Employee	Employee	Gender	JobLevel	JobRole	Marital	MonthlyIncome	NumCo	Over18	Percent	Standard	StockO	TotalW	Training	YearsA	YearsSi	Ye	Avg Monthly Income
51	No	Travel_Ra	Sales	6	2	Life Scienc	1	1	Female	1	Healthcare	Married	131160	1	Y	11	8	0	1	6	1	0	0	65029.31
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	4	
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	3	
38	No	Non-Trave	Research &	2	5	Life Scienc	1	4	Male	3	Human Re	Married	83210	3	Y	11	8	3	13	5	8	7	5	
32	No	Travel_Ra	Research &	10	1	Medical	1	5	Male	1	Sales Exec	Single	23420	4	Y	12	8	2	9	2	6	0	4	
46	No	Travel_Ra	Research &	8	3	Life Scienc	1	6	Female	4	Research I	Married	40710	3	Y	13	8	0	28	5	7	7	7	
28	Yes	Travel_Ra	Research &	11	2	Medical	1	7	Male	2	Sales Exec	Single	58130	2	Y	20	8	1	5	2	0	0	0	
29	No	Travel_Ra	Research &	18	3	Life Scienc	1	8	Male	2	Sales Exec	Married	31430	2	Y	22	8	3	10	2	0	0	0	
31	No	Travel_Ra	Research &	1	3	Life Scienc	1	9	Male	3	Laboratory	Married	20440	0	Y	21	8	0	10	2	9	7	8	
25	No	Non-Trave	Research &	7	4	Medical	1	10	Female	4	Laboratory	Divorced	134640	1	Y	13	8	1	6	2	6	1	5	
45	No	Travel_Ra	Research &	17	2	Medical	1	11	Male	2	Laboratory	Married	79910	0	Y	13	8	2	21	2	20	4	10	
36	No	Travel_Ra	Research &	28	1	Life Scienc	1	12	Male	1	Laboratory	Married	33770	0	Y	12	8	2	16	2	15	10	11	
55	No	Travel_Ra	Research &	14	4	Life Scienc	1	13	Female	1	Sales Exec	Single	55380	0	Y	17	8	0	37	2	36	4	13	
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	9	
28	No	Travel_Ra	Research &	1	3	Life Scienc	1	15	Male	1	Manufactu	Married	25920	1	Y	14	8	0	5	2	5	0	4	
37	No	Travel_Ra	Research &	1	3	Life Scienc	1	16	Male	2	Healthcare	Married	53460	4	Y	11	8	0	7	2	5	0	1	
21	No	Travel_Ra	Research &	3	2	Life Scienc	1	17	Male	1	Laboratory	Single	42130	1	Y	12	8	3	3	3	3	1	0	
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	2	
35	No	Travel_Ra	Sales	7	4	Life Scienc	1	19	Male	1	Sales Repr	Divorced	24380	7	Y	16	8	0	10	5	7	6	2	
38	No	Travel_Ra	Research &	8	3	Life Scienc	1	20	Female	1	Manager	Divorced	68700	1	Y	11	8	1	8	5	8	7	7	
26	No	Travel_Fre	Research &	1	4	Other	1	21	Male	2	Laboratory	Divorced	104470	1	Y	18	8	0	6	3	6	1	4	
50	No	Travel_Ra	Sales	8	4	Life Scienc	1	22	Male	1	Research S	Divorced	96670	3	Y	23	8	0	28	2	10	1	6	
53	No	Travel_Ra	Research &	11	4	Life Scienc	1	23	Female	2	Research S	Married	21480	3	Y	11	8	0	21	2	5	1	3	
42	No	Travel_Ra	Research &	4	4	Life Scienc	1	24	Male	1	Manufactu	Married	89260	1	Y	14	8	0	NA	4	20	11	6	
29	No	Travel_Fre	Research &	16	4	Medical	1	25	Male	1	Laboratory	Single	65130	1	Y	11	8	1	10	2	10	0	9	
55	No	Travel_Ra	Research &	1	4	Other	1	26	Female	1	Research S	Married	67990	3	Y	11	8	0	12	2	10	0	8	
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	3	
37	No	Travel_Ra	Sales	5	1	Marketing	1	28	Male	1	Research S	Single	27050	1	Y	11	8	0	17	2	17	5	7	
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	0	

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

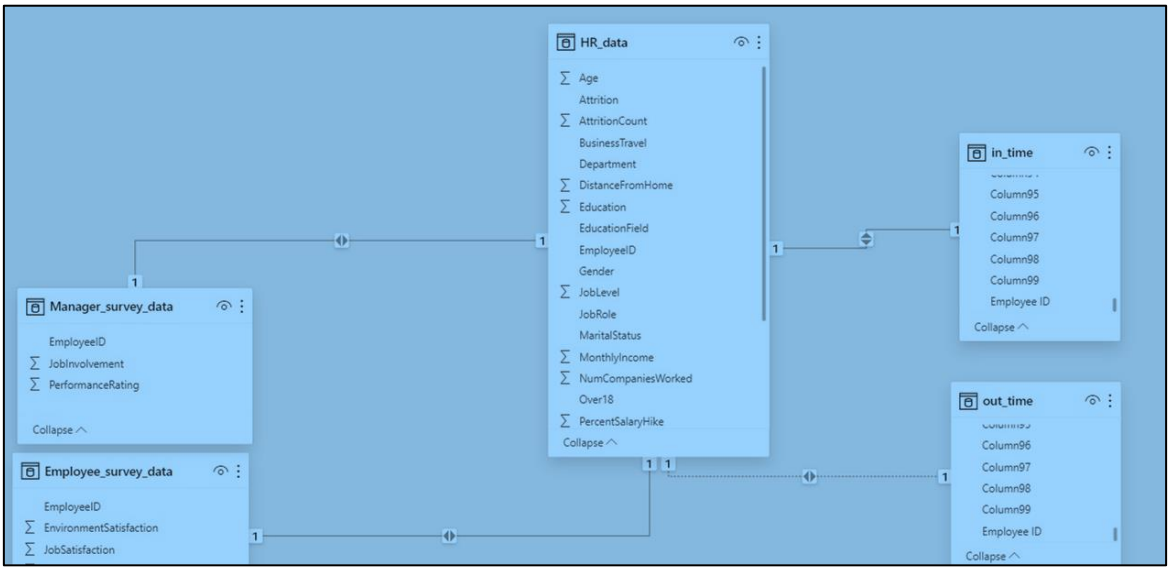


Age	Attrition Employee
18	12
19	18
20	18
21	18
22	15
23	12
24	21
25	18
26	36
27	9
28	42
29	54
30	27
31	54
32	33
33	36
34	27
35	30
36	18
37	18
38	6
39	18
40	15

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

Answer: No Missing and Inconsistent data found in the Column

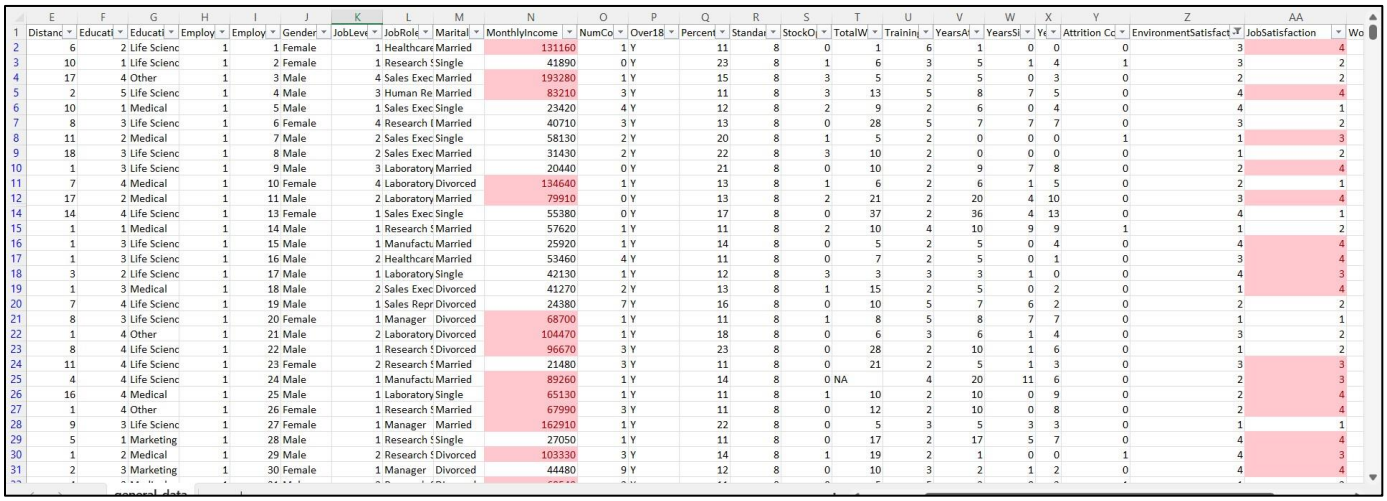
6. **In Power BI, establish a relationship between the "EmployeeID" in the employee data and the "EmployeeID" in the time tracking data.**



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

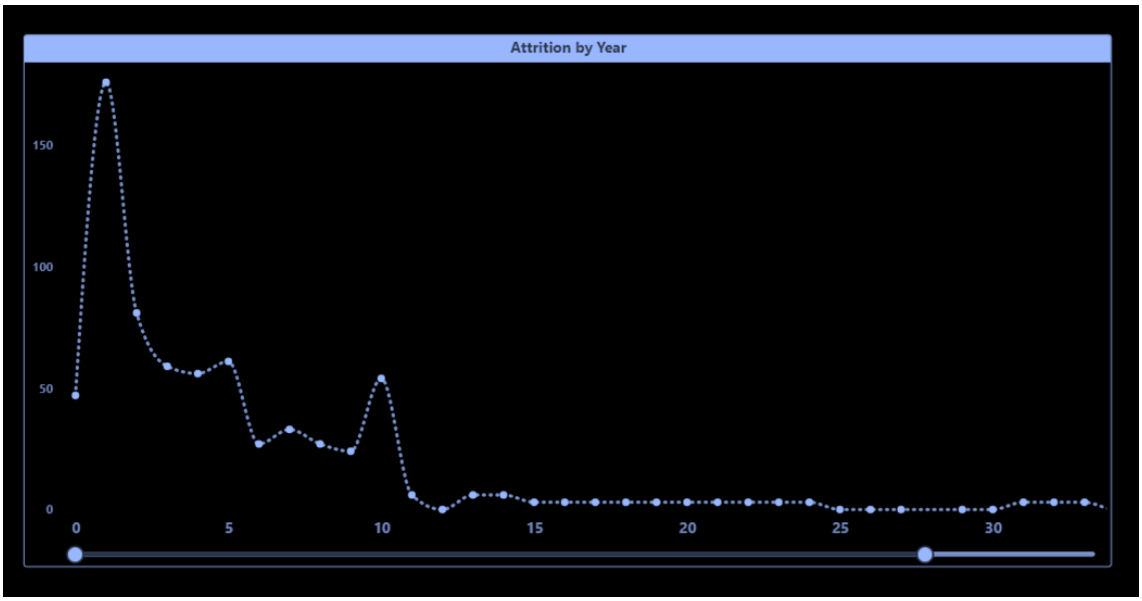
8. **Using Excel, create a pivot table that displays the count of employees in each Marital Status category, segmented by Department.**

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



	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA
1	Distanc	Educati	Educati	Employ	Employ	Gender	JobLeve	JobRole	Marital	MonthlyIncome	NumCo	Over18	Percent	Standar	StockO	TotalW	Trainin	YearsA	YearsSi	Ye	Attrition Co	EnvironmentSatisfact	JobSatisfaction	Wo
2	6	2	Life Scienc	1	1	Female	1	Healthcare	Married	131160	1	Y	11	8	0	1	6	1	0	0	0	3	4	
3	10	1	Life Scienc	1	2	Female	1	Research S	Single	41890	0	Y	23	8	1	6	3	5	1	4	1	3	2	
4	17	4	Other	1	3	Male	4	Sales Exec	Married	193280	1	Y	15	8	3	5	2	5	0	3	0	2	2	
5	2	5	Life Scienc	1	4	Male	3	Human Re	Married	83210	3	Y	11	8	3	13	5	8	7	5	0	4	4	
6	10	1	Medical	1	5	Male	1	Sales Exec	Single	23420	4	Y	12	8	2	9	2	6	0	4	0	4	1	
7	8	3	Life Scienc	1	6	Female	4	Research I	Married	40710	3	Y	13	8	0	28	5	7	7	7	0	3	2	
8	11	2	Medical	1	7	Male	2	Sales Exec	Single	58130	2	Y	20	8	1	5	2	0	0	0	1	1	3	
9	18	3	Life Scienc	1	8	Male	2	Sales Exec	Married	31430	2	Y	22	8	3	10	2	0	0	0	0	1	2	
10	1	3	Life Scienc	1	9	Male	3	Laboratory	Married	20440	0	Y	21	8	0	10	2	9	7	8	0	2	4	
11	7	4	Medical	1	10	Female	4	Laboratory	Divorced	134640	1	Y	13	8	1	6	2	6	1	5	0	2	1	
12	17	2	Medical	1	11	Male	2	Laboratory	Married	79910	0	Y	13	8	2	21	2	20	4	10	0	3	4	
13	14	4	Life Scienc	1	13	Female	1	Sales Exec	Single	55380	0	Y	17	8	0	37	2	36	4	13	0	4	1	
14	1	1	Medical	1	14	Male	1	Research S	Married	57620	1	Y	11	8	2	10	4	10	9	9	1	1	2	
15	1	3	Life Scienc	1	15	Male	1	Manufactu	Married	25920	1	Y	14	8	0	5	2	5	0	4	0	4	4	
16	1	3	Life Scienc	1	16	Male	2	Healthcare	Married	53460	4	Y	11	8	0	7	2	5	0	1	0	3	4	
17	3	2	Life Scienc	1	17	Male	1	Laboratory	Single	42130	1	Y	12	8	3	3	3	3	1	0	0	4	3	
18	1	3	Medical	1	18	Male	2	Sales Exec	Divorced	41270	2	Y	13	8	1	15	2	5	0	2	0	1	4	
19	7	4	Life Scienc	1	19	Male	1	Sales Repr	Divorced	24380	7	Y	16	8	0	10	5	7	6	2	0	2	2	
20	8	3	Life Scienc	1	20	Female	1	Manager	Divorced	68700	1	Y	11	8	1	8	5	8	7	7	0	1	1	
21	1	4	Other	1	21	Male	2	Laboratory	Divorced	104470	1	Y	18	8	0	6	3	6	1	4	0	3	2	
22	8	4	Life Scienc	1	22	Male	1	Research S	Divorced	96670	3	Y	23	8	0	28	2	10	1	6	0	1	2	
23	11	4	Life Scienc	1	23	Female	2	Research S	Married	21480	3	Y	11	8	0	21	2	5	1	3	0	3	3	
24	4	4	Life Scienc	1	24	Male	1	Manufactu	Married	89260	1	Y	14	8	0	NA	4	20	11	6	0	2	3	
25	16	4	Medical	1	25	Male	1	Laboratory	Single	65130	1	Y	11	8	1	10	2	10	0	9	0	2	4	
26	1	4	Other	1	26	Female	1	Research S	Married	67990	3	Y	11	8	0	12	2	10	0	8	0	2	4	
27	9	3	Life Scienc	1	27	Female	1	Manager	Married	162910	1	Y	22	8	0	5	3	5	3	3	0	1	1	
28	5	1	Marketing	1	28	Male	1	Research S	Single	27050	1	Y	11	8	0	17	2	17	5	7	0	4	4	
29	1	2	Medical	1	29	Male	2	Research S	Divorced	103330	3	Y	14	8	1	19	2	1	0	0	1	4	3	
30	2	3	Marketing	1	30	Female	1	Manager	Divorced	44480	9	Y	12	8	0	10	3	2	1	2	0	4	4	

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.

Creating a star schema involves organizing a database into a central fact table surrounded by dimension tables. This structure is commonly used in data warehouses and business intelligence systems.

Hypothetical Dataset: Sales and Product Information

1. Identify the Fact Table:
   - Fact Table: `Sales`
   - Columns: `OrderID` (Primary Key), `ProductID` (Foreign Key), `CustomerID` (Foreign Key), `OrderDate`, `QuantitySold`, `TotalAmount`, etc.

2. Identify Dimension Tables:
   - Dimension Table 1: `Products`
   - Columns: `ProductID` (Primary Key), `ProductName`, `Category`, `Brand`, `Price`, etc.

   - Dimension Table 2: `Customers`

- Columns: `CustomerID` (Primary Key), `CustomerName`, `Address`, `PhoneNumber`, etc.

- Dimension Table 3: `Dates`
  - Columns: `DateID` (Primary Key), `Date`, `DayOfWeek`, `Month`, `Quarter`, `Year`, etc.

- Additional Dimension Tables as Needed: Depending on the dataset, you might include additional dimension tables like `Locations`, `Salespersons`, etc.

3. Create Relationships:
   - Establish relationships between the primary key in the fact table and the foreign keys in the dimension tables.
   - For example, connect `ProductID` in the `Sales` table to `ProductID` in the `Products` table, and similarly for other foreign keys.

Benefits of Star Schema:

1. Simplicity and Understandability:
   - The star schema is simple and intuitive, making it easy for users to understand the database structure. This simplicity enhances user adoption and query performance.
2. Query Performance:
   - Queries involving specific dimensions can be executed more efficiently, as data related to each dimension is stored in separate dimension tables. This reduces the need for complex joins and improves query speed.
3. Flexibility and Maintainability:
   - Changes to the database structure are more manageable. Adding new dimensions or modifying existing ones doesn't require extensive modifications to the entire schema.
4. Scalability:
   - The star schema is scalable for large datasets. It's designed to handle substantial amounts of data and still provide good query

performance.

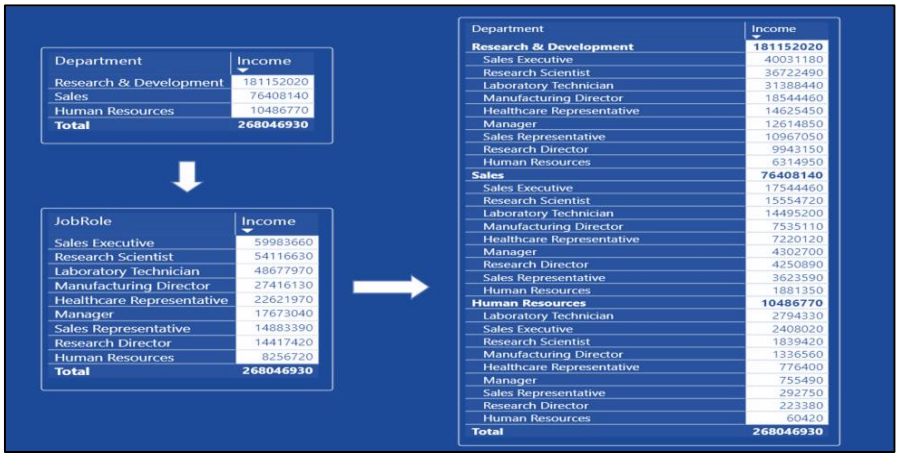
5. Better Support for Business Intelligence (BI) Tools:
   - BI tools are often designed to work well with star schemas. The structure aligns with the way users think about their data, facilitating reporting and analysis.
6. Normalized vs. Denormalized Data:
   - Star schemas strike a balance between normalized and denormalized data. While dimensions are typically normalized, the fact table contains denormalized data for efficient querying.

Creating a star schema for a dataset provides a structured and efficient way to organize and access data, offering benefits in terms of simplicity, query performance, flexibility, and support for business intelligence activities.

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

```
Rolling 3-Month Average =
VAR CurrentMonth = 'EmployeeIncome'[Month]
RETURN
  CALCULATE(
    AVERAGE('EmployeeIncome'[MonthlyIncome]),
    FILTER(
      ALL('EmployeeIncome'),
      'EmployeeIncome'[Month] >= CurrentMonth - 2 &&
      'EmployeeIncome'[Month] <= CurrentMonth
    )
  )
```

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



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

In Power BI, you can set up parameterized queries using parameters to allow users to filter data based on the "Distance from Home" column. Parameters provide a way to dynamically modify the behavior of your queries and allow users to input values at runtime. Here are the steps to set up parameterized queries in Power BI:

Step 1: Create a Parameter

- In Power BI Desktop, go to the "Home" tab.
- Click on "Manage Parameters" in the External Data group.
- In the Manage Parameters window, click on "New" to create a new parameter.
- Name the parameter (e.g., "DistanceParameter") and set the data type to Decimal Number or Text based on the data type of your "Distance from Home" column.
- Set the default value and any other relevant properties.
- Click "OK" to create the parameter.

Step 2: Modify the Query

- In the Power Query Editor, locate the query related to your dataset.
- Edit the formula for the filter in the "Distance from Home" column to use the parameter. For example, if your original filter is:

```M
#"Filtered Rows" = Table.SelectRows(Source, each [Distance from Home] <= 10)
```

You would modify it to:

```M
#"Filtered Rows" = Table.SelectRows(Source, each [Distance from Home] <= DistanceParameter)
```

Step 3: Apply Changes

- Click "Close & Apply" to apply the changes and close the Power Query Editor.

Step 4: Use the Parameter in Visualizations

- In your report, create a visual (e.g., a table, chart, or slicer) that displays data related to the "Distance from Home" column.
- In the visualization, drag and drop the "DistanceParameter" parameter onto the visual, or use it in the filter pane.
- Users can now interactively change the parameter value to filter data based on the "Distance from Home" column.

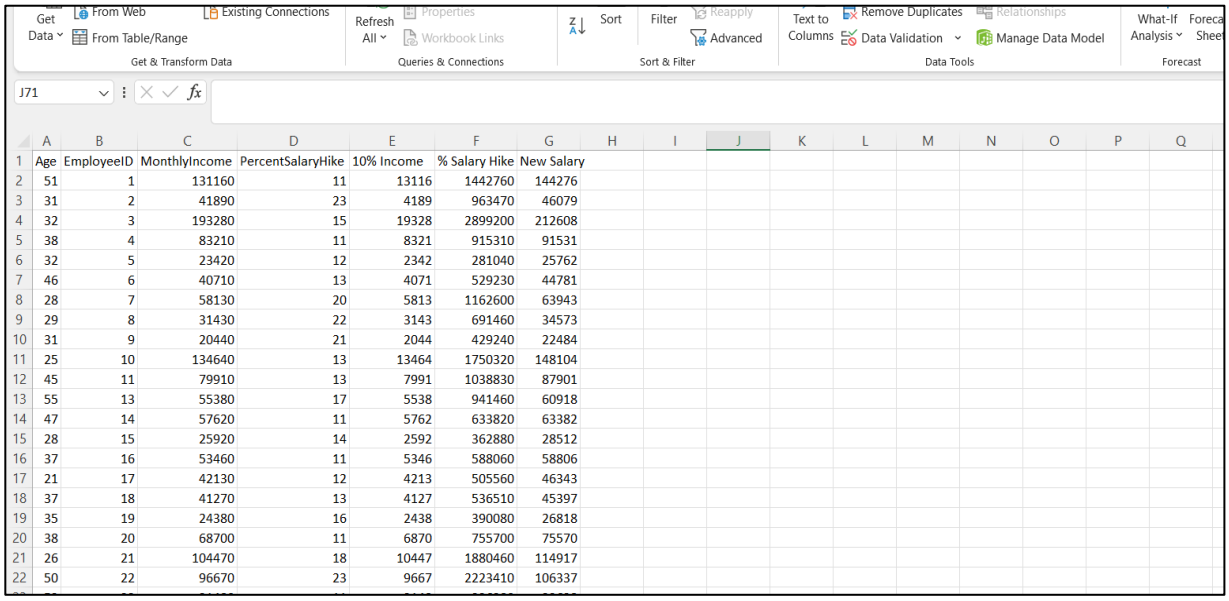
Step 5: Publish and Share

- Save your Power BI file and publish it to the Power BI service.
- Users accessing the report in the Power BI service or Power BI Report Server can also interact with the parameterized query.

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.

Total MonthlyIncome	Department			
JobLevel	Human Resources	Research & Development	Sales	Total
3	1648500	28117740	11792400	41558640
4	754800	15277290	8753070	24785160
5	855840	10107870	2428860	13392570
Total	**3259140**	**53502900**	**22974330**	**79736370**

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	H	I	J	K	L	M	N	O	P	Q
1	Age	EmployeeID	MonthlyIncome	PercentSalaryHike	10% Income	% Salary Hike	New Salary										
2	51	1	131160	11	13116	1442760	144276										
3	31	2	41890	23	4189	963470	46079										
4	32	3	193280	15	19328	2899200	212608										
5	38	4	83210	11	8321	915310	91531										
6	32	5	23420	12	2342	281040	25762										
7	46	6	40710	13	4071	529230	44781										
8	28	7	58130	20	5813	1162600	63943										
9	29	8	31430	22	3143	691460	34573										
10	31	9	20440	21	2044	429240	22484										
11	25	10	134640	13	13464	1750320	148104										
12	45	11	79910	13	7991	1038830	87901										
13	55	13	55380	17	5538	941460	60918										
14	47	14	57620	11	5762	633820	63382										
15	28	15	25920	14	2592	362880	28512										
16	37	16	53460	11	5346	588060	58806										
17	21	17	42130	12	4213	505560	46343										
18	37	18	41270	13	4127	536510	45397										
19	35	19	24380	16	2438	390080	26818										
20	38	20	68700	11	6870	755700	75570										
21	26	21	104470	18	10447	1880460	114917										
22	50	22	96670	23	9667	2223410	106337										

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

Steps to Verify Data Adherence to Predefined Schema:

1.Define the Schema:

- Clearly define the expected schema for your data. This includes specifying the expected data types, constraints, and

relationships for each field in your dataset.

2. Use Data Profiling Tools:
   - Utilize data profiling tools or features in data analysis tools to automatically analyze the structure and characteristics of your data.
3. Review Metadata:
   - Examine metadata associated with your data source to understand the expected structure and data types.
4. Perform Descriptive Statistics:
   - Calculate descriptive statistics (e.g., mean, median, standard deviation) for numeric fields and explore the distribution of categorical fields to identify outliers or unexpected values.

Actions to Take if Inconsistencies are Found:

1. Document Inconsistencies:
   - Document the inconsistencies, including details such as the affected columns, unexpected values, and any patterns observed.
2. Communicate with Data Providers:
   - If the data is received from external sources, communicate with the data providers to understand the reasons for the inconsistencies and work towards resolving them.
3. Data Cleansing:
   - Implement data cleansing procedures to address inconsistencies. This may involve correcting or removing erroneous data points, filling in missing values, or transforming data to adhere to the predefined schema.
4. Implement Validation Rules:
   - Set up validation rules or constraints in your data processing pipeline to prevent similar inconsistencies from occurring in the future.
5. Update Documentation:
   - Update your documentation to reflect any changes made to the schema or data cleansing procedures. This ensures that stakeholders are aware of the data quality improvements.
6. Create Alerts or Monitoring:

- Implement monitoring mechanisms to receive alerts when new data does not conform to the predefined schema. This allows for early detection and quick resolution of issues.

7. Iterative Process:

- Data validation is often an iterative process. Regularly revisit the data validation steps as new data is acquired or as business requirements evolve.

8. Collaborate with Data Governance Team:

- If your organization has a data governance team, collaborate with them to ensure that data quality standards and processes are aligned with organizational goals.