**Employee Data: Performance of Employee**

**Introduction**

Human Resources (HR) is the set of people who make the workforce of an organization, business sector, industry, or economy. A human resources (HR) department performs human resource management functions, such as finding, hiring, and supporting new employees. Therefore, HR departments are responsible for such important tasks as reviewing resumes, keeping of employee information, and ensuring a company complies with labour laws and employment standards.

In this analysis, how employees are paid in relationship with department and job role is to be investigated.

**The Data**

The data contains different tables including general, survey data, in-time data, and out time data. The data contains exactly 4,410 sampled employees with unique ID numbers.

### Rating System for Job Involvement Level, Environment Satisfaction and Job Satisfaction Level:

4= Very High

3= High

2= Medium

1= Low

### Rating System for Work Life Balance:

4= Best

3= Better

2= Good

1= Bad

### Rating System for Education Level:

5 – Doctor

4= Master

3= Bachelor

2= College

1= Below College

**Project Tool Used: Excel, PowerBI**

**Analysis Questions**

The business questions to look into in this project are the following:

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

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

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

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.

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.

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.

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.

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

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.

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.

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

**Availability of Data**

- The provided data would be good enough for the analysis to answer the business questions asked.

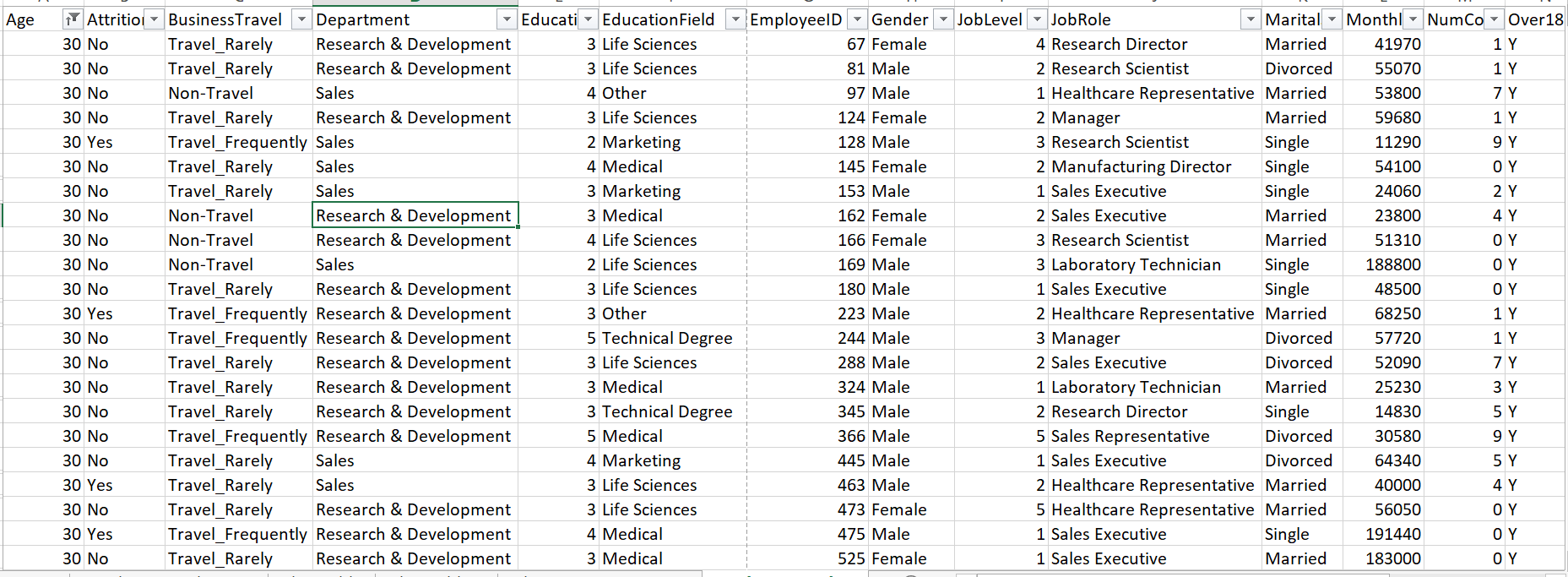
**Data Cleaning Process**

There were no duplicate in the dataset but unnecessary columns like “EmployeeCount”, and “StockOptionLevel” were removed due to their irrelevances to the business questions.

**Analysis and Insights**

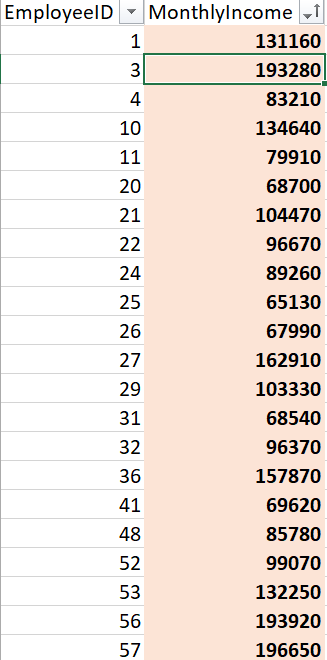
Using conditional formatting along with pivot tables and plotting different graphs, the following insights could be deduced from the analysis:

1. Employees aged 30 and above are shown below.



2. Manufacturing Director has the highest average Monthly Income followed by Laboratory Technician while Human Resources gets the least average Monthly Income.

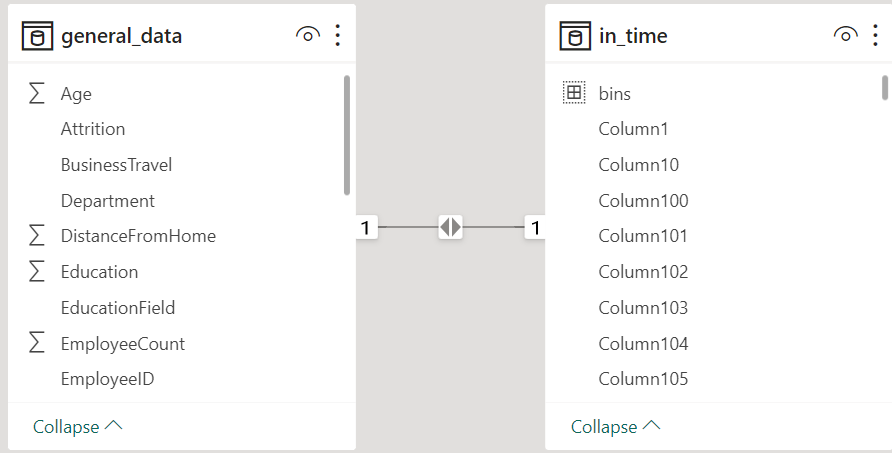
3. Employees with Monthly Income above the company's average income include 1, 3, 4, 10, 11, 20, 21, 22, 24, 25, 26 etc.



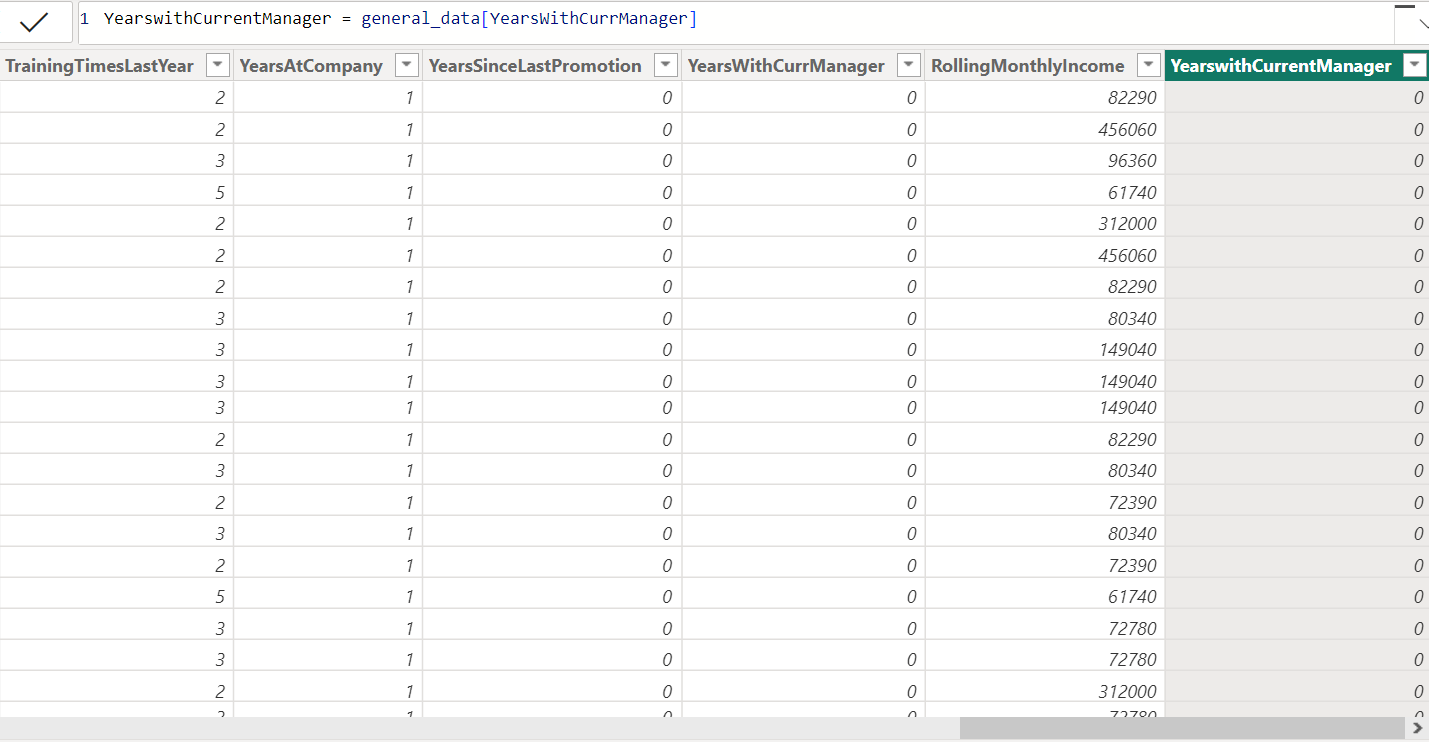
4. Employees with age 35 are the most in the company while those of 57 are the least. This means majority of the workers are mid-aged.

5. There is no any missing or inconsistent data in the "Department" column.

6. "EmployeeID" in the employee data and the "EmployeeID" in the time tracking data are related as shown below.

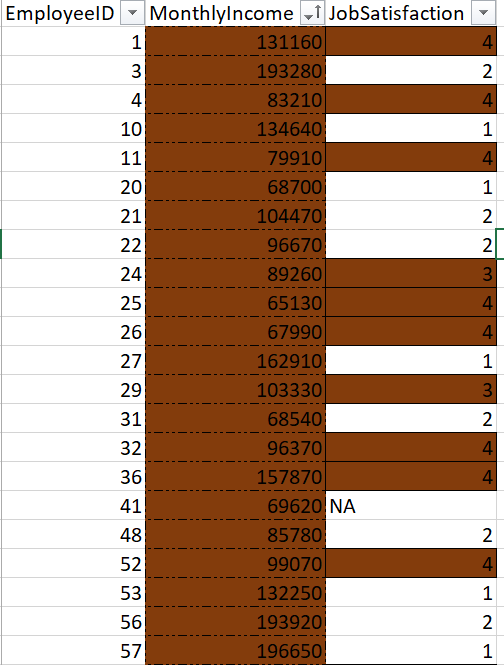


7. The average years an employee has spent with their current manager are calculated using the DAX below on powerBI.

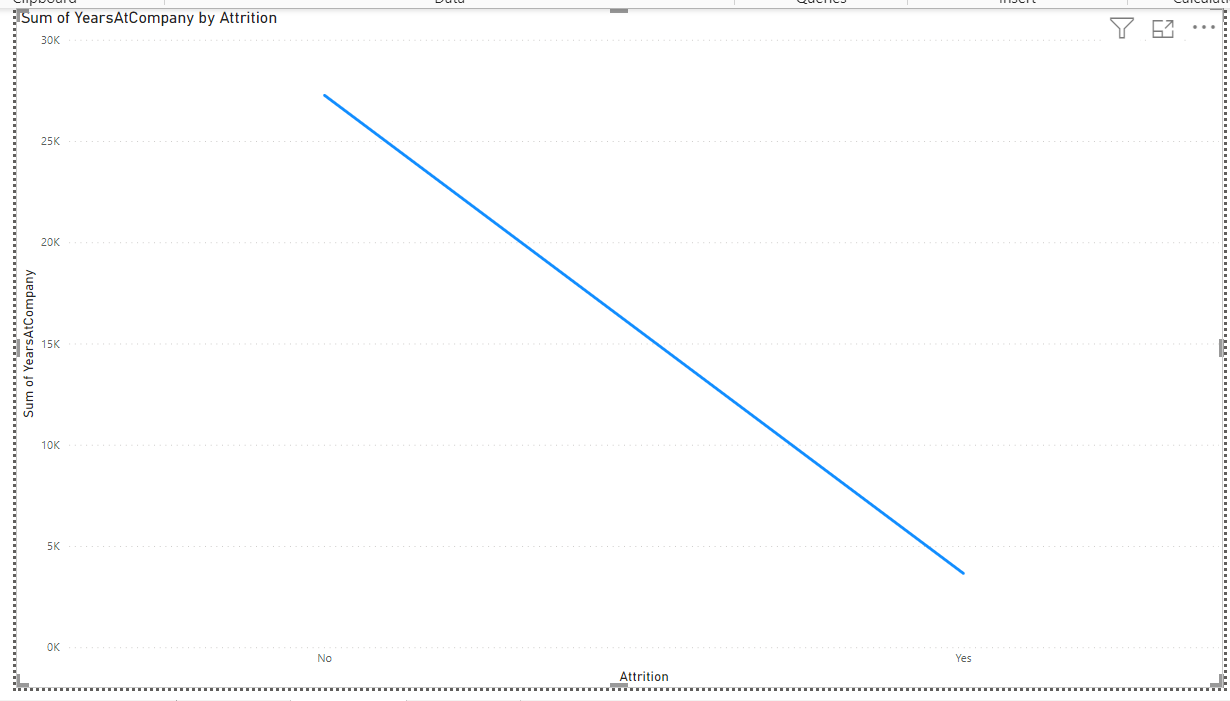


8. Majority of the employees are married followed by single ones while divorced workers are least in number. Also, majority of the employees are in Research & Development department with only few workers in Human Resources department.

9. Employees with both above-average Monthly Income and above-average Job Satisfaction include 1, 4, 11, 12, 24, 25, 26, 29, 32, 36, 52 among others as shown below



10. Employee with No Attrition are greatly more than those with Yes Attrition

  
11. An HR data model could be designed using a star schema to track the employee metrics, such as employee turnover, training, and performance. In this scenario, the fact table would contain the general data of the employees, while the dimension tables would provide details like time in/out and manager-survey data.

Out-time

Employee Id Integer

Day1 Date/time

Day 2 Date/time

Day 3 Date/time

In-time

Employee Id Integer

Day 1 Date/time

Day 2 Date/time

Day 3 Date/time

General Data

Employee Id Integer

Full Names Varchar (50)

Age Integer

Gender Varchar (50)

Department Varchar (50)

Job Role Varchar (50)

Manager-Survey Data

Employee Id Int

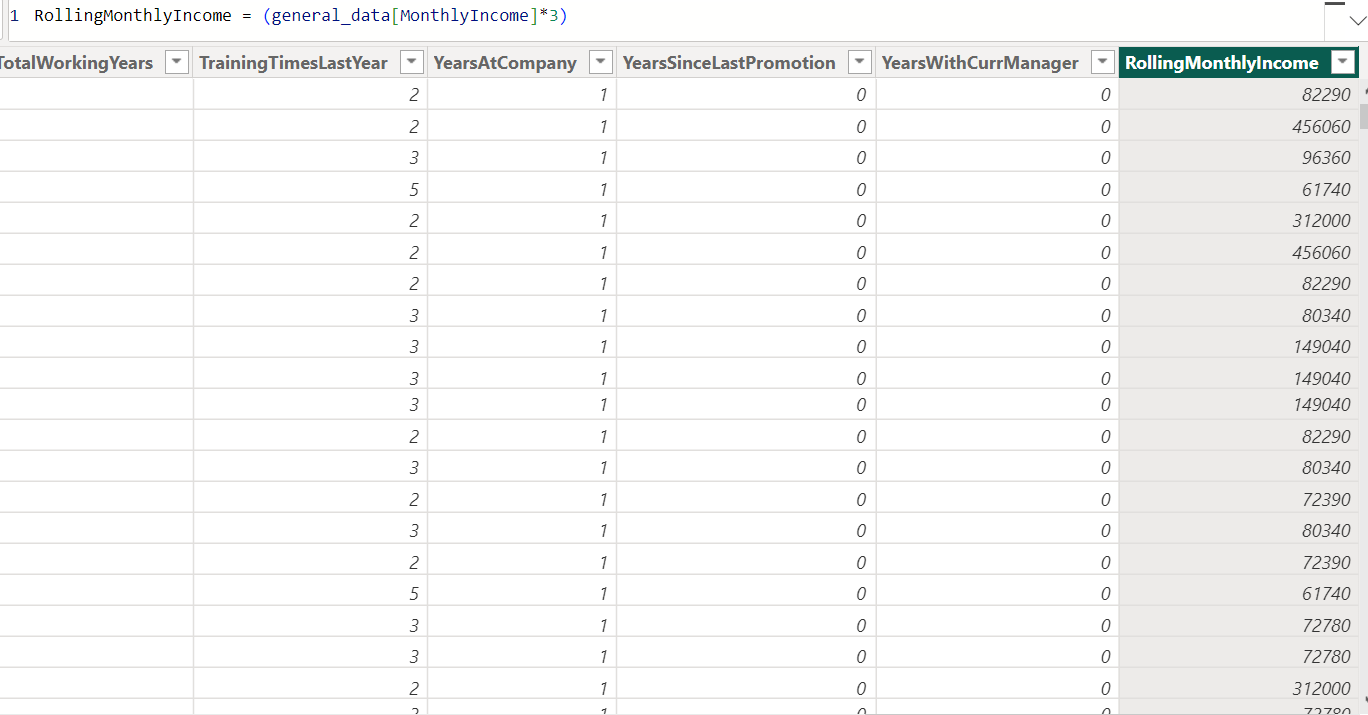
Job Involvement Int

Performance Rating Int

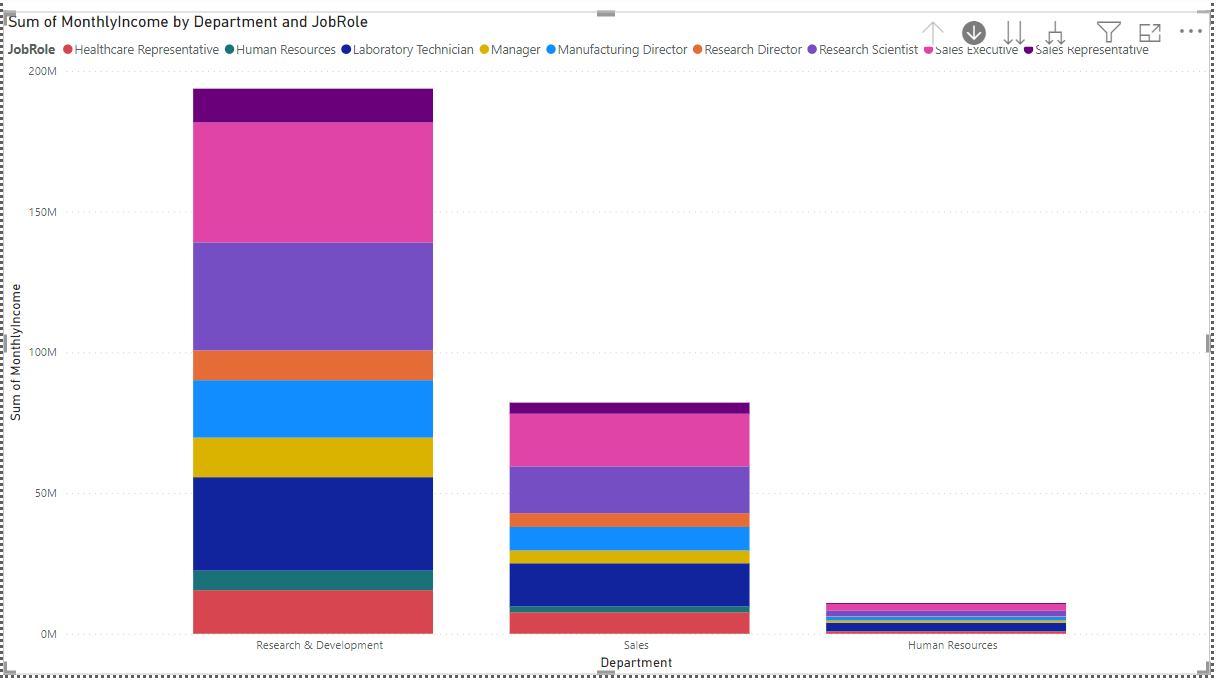
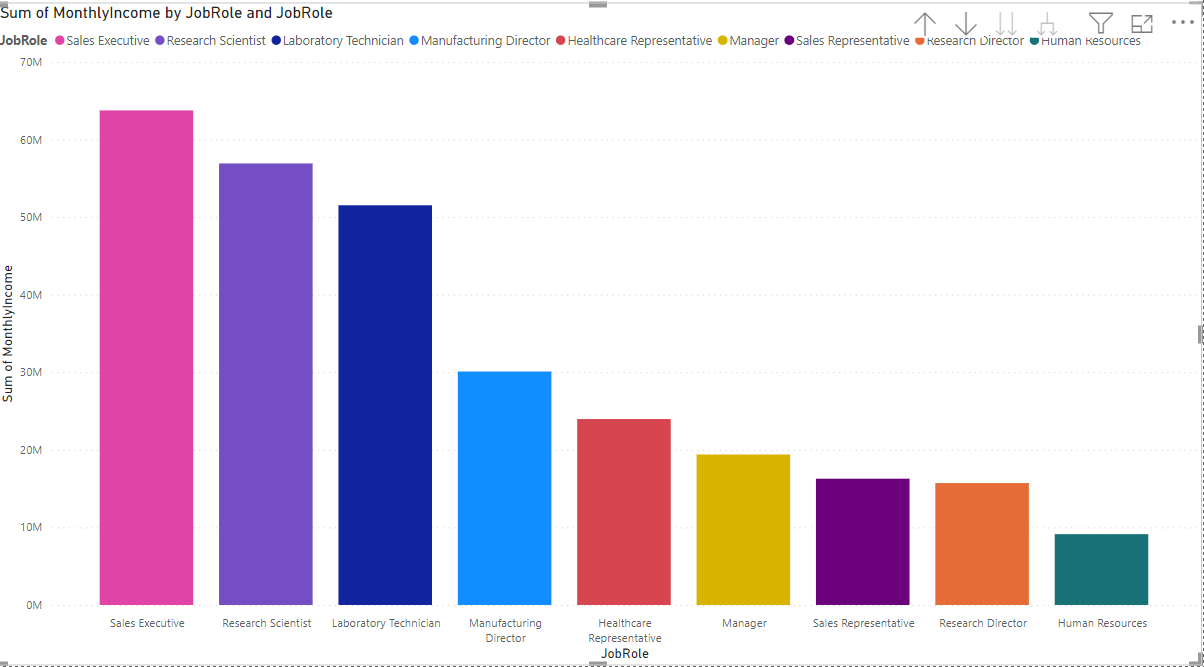
Benefits of a Start Schema

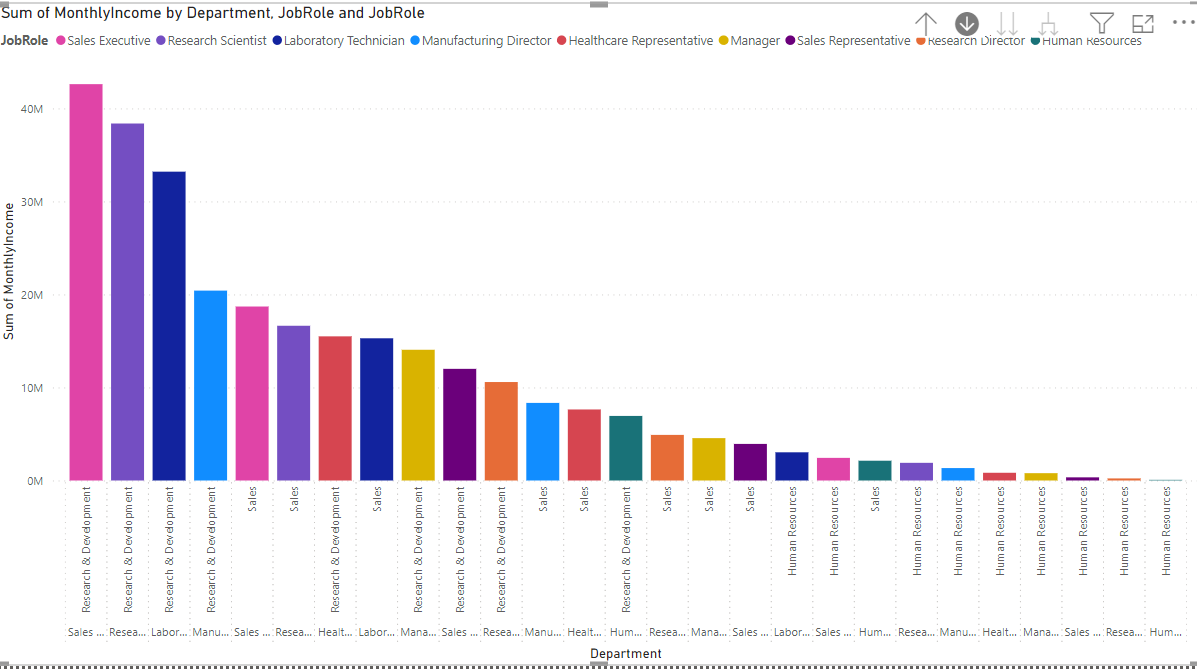
1. Improved performance: Queries will run faster on a star schema data model, which will enable analysts and business users to get answers to their questions quickly and efficiently.
2. Simplified analysis: Because a star schema is intuitive and clear, the organization of data will make analysis easy for business users, which means they can identify trends and patterns that might otherwise be missed.
3. Scalability: The star schema is able to handle large amounts of data, making it a great choice for businesses that deal with massive amounts of data that need to be analyzed quickly and combined in dynamic ways.
4. Easy maintenance: Because the star schema has a simple design, it makes it easy to maintain and update the model, all while reducing the risk of errors and downtime of the overall system.

12. The rolling 3-month average of Monthly Income for each employee can be calculated using the following code:

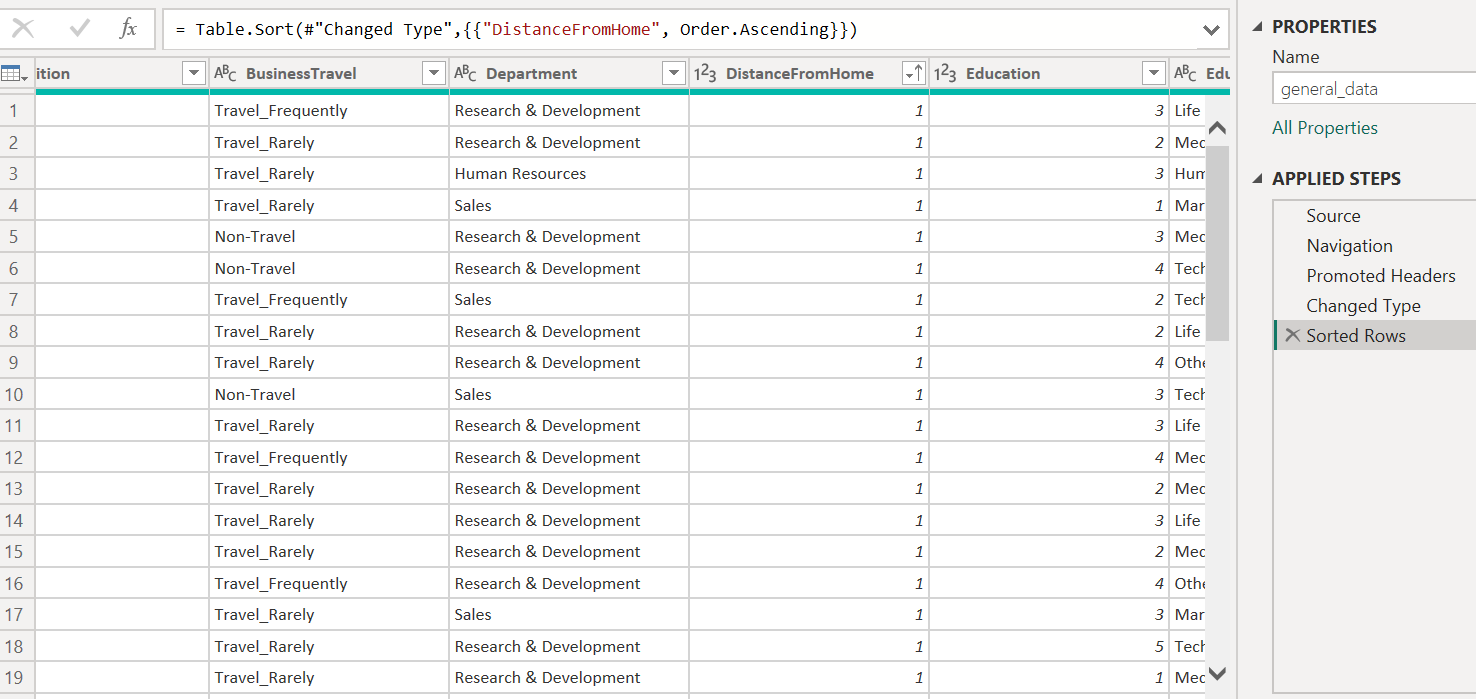


13. Users can drill down from Department to Job Role to further narrow their analysis as shown below. This indicates that “Research and Development” department and “Sales Executive” job role receive the highest monthly income.



14. To filter the data based on the Distance from Home column can be achieved by opening the data in the “Transform Page”.



15. Considering the Job Level greater than or equal to 3, the Research & Development Department has the highest sum of Monthly Income while Human Resources has the least.

16. What-If analysis in Excel to understand the impact of a 10% increase in Percent Salary Hike on Monthly Income can be done using data table under the What-If analysis tab in the Data tab..

17. The data adheres to a predefined schema except for the absence of “EmployeeID” title in the “in-time” and “out-time” datasets. In powerBi model view, the EmployeeID column in the general data table was dragged to the time table for linking purpose.

**Conclusion**

It is found that employees in “Research and Development” department and “Sales Executive” job role receive the highest monthly payment.

**Dashboard**

