

# Task 1

HR Data Analysis

# By

# Akash Aman

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

**STEPS**:

➢Select the age column and use the filter option

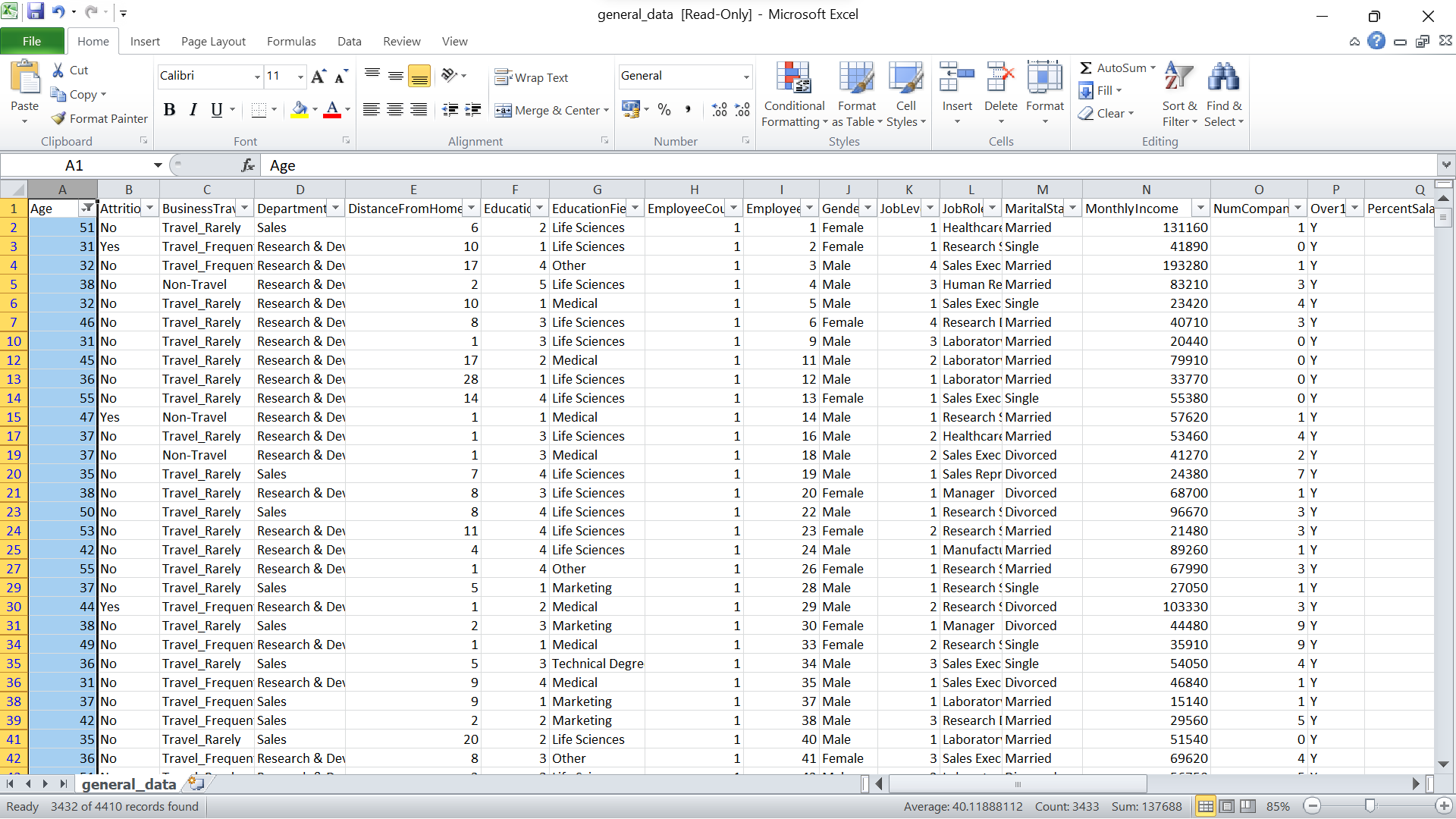
➢Select the Number filtering option, then select “Greater or Equal to

➢In the text box type 30

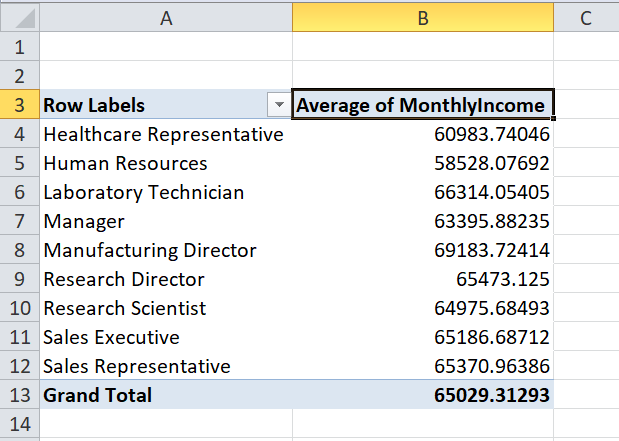
➢After clicking ok the rows will be filtered accordingly

**RESULT**:

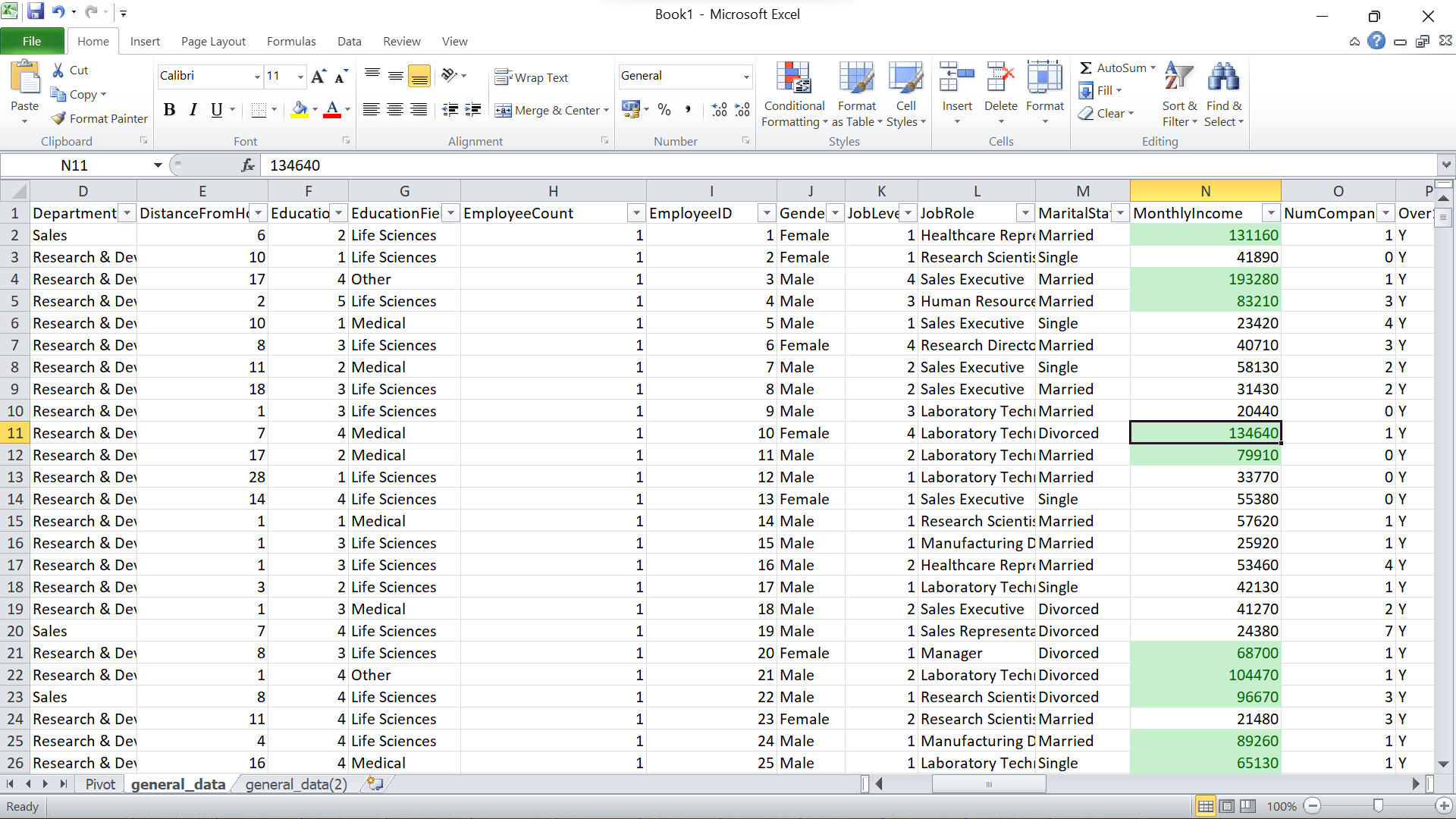
3432 employees are aged 30 and above.



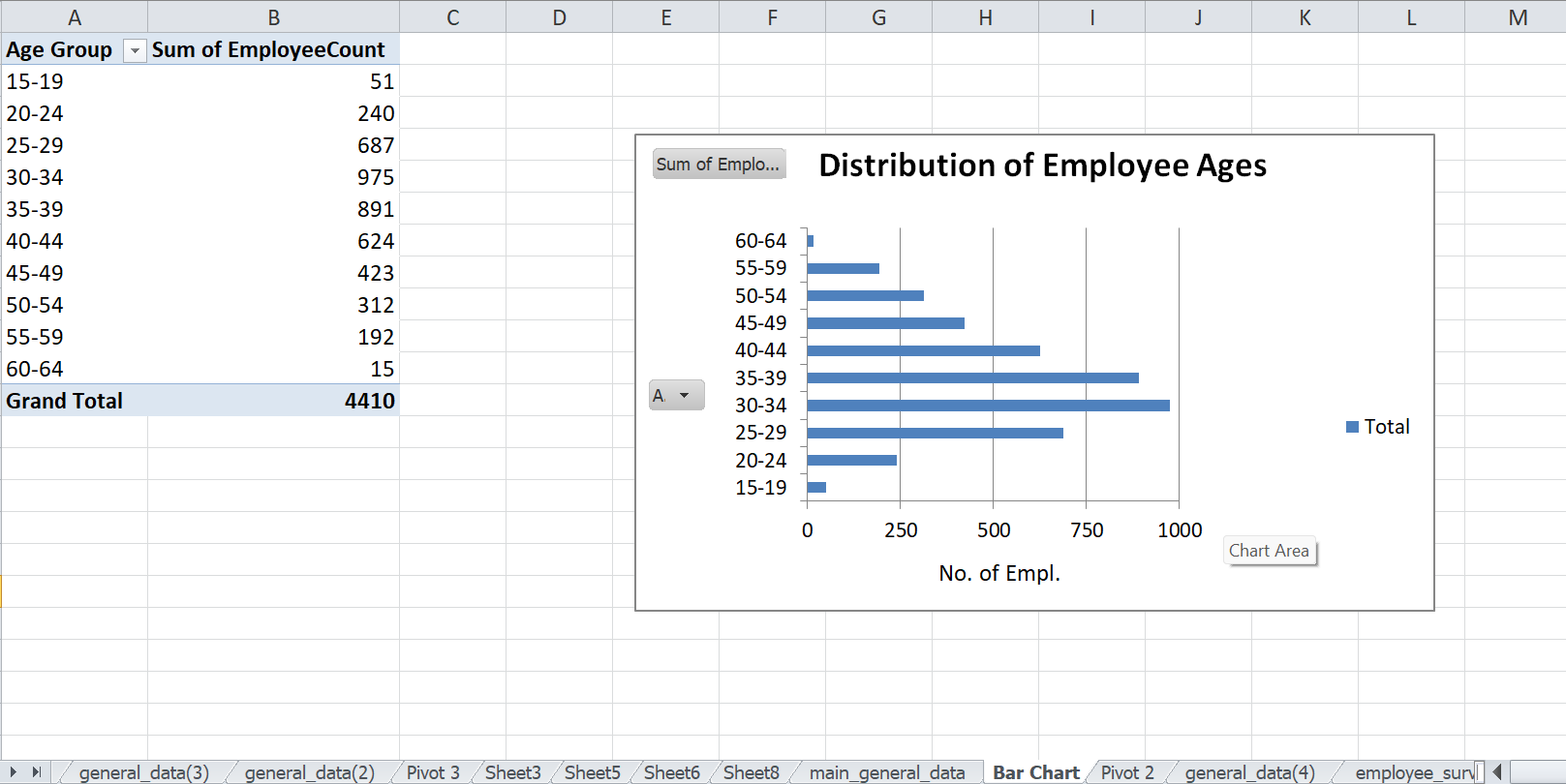
2. Create a pivot table to summarize the average Monthly Income by Job Role.



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



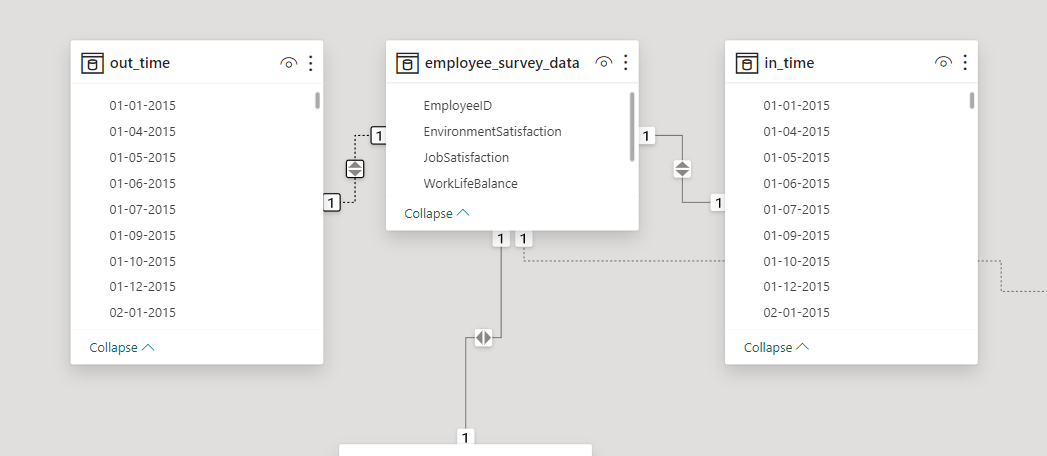
4. Create a bar chart in Excel to visualize the distribution of employee ages.

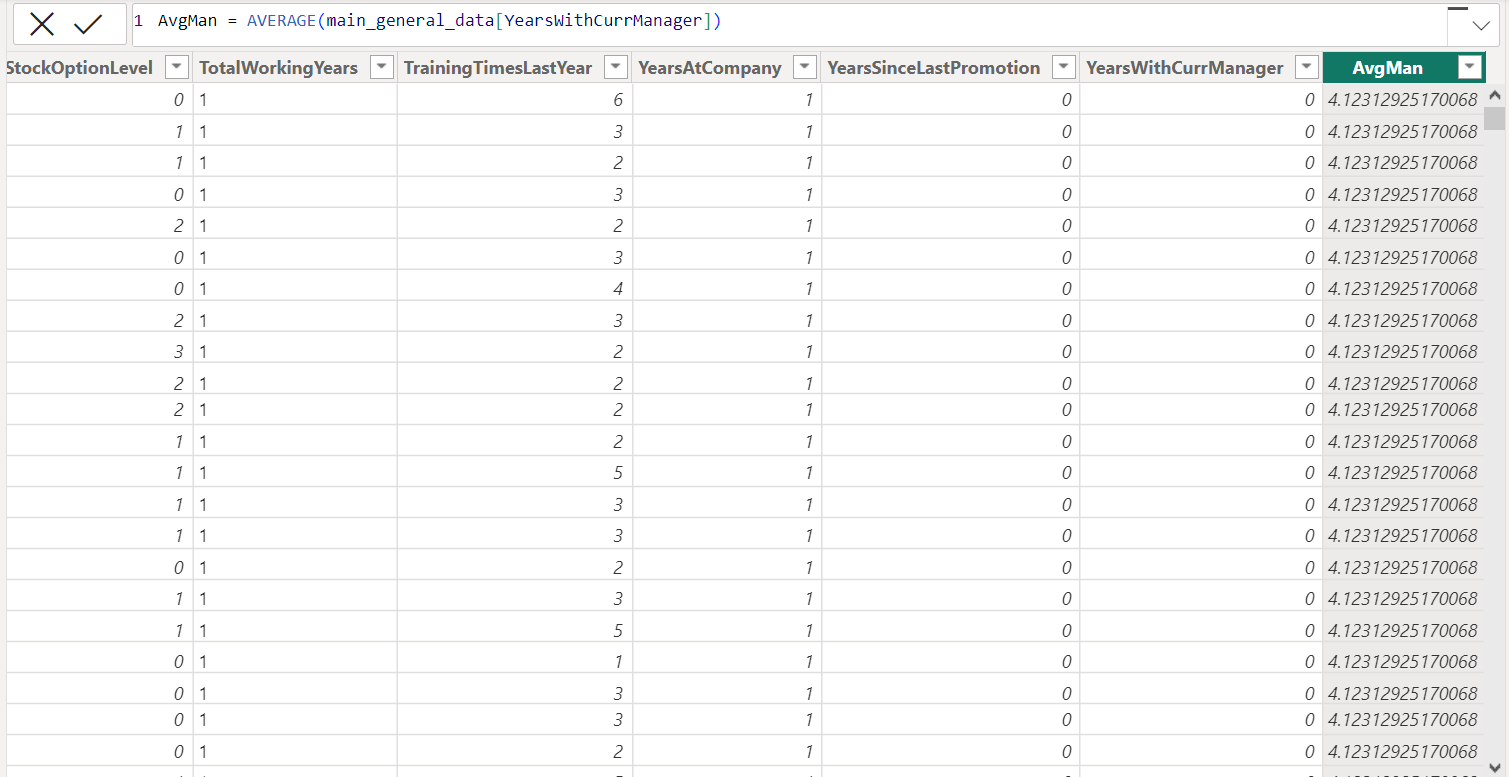


5. Identify and clean any missing or inconsistent data in the "Department" column.

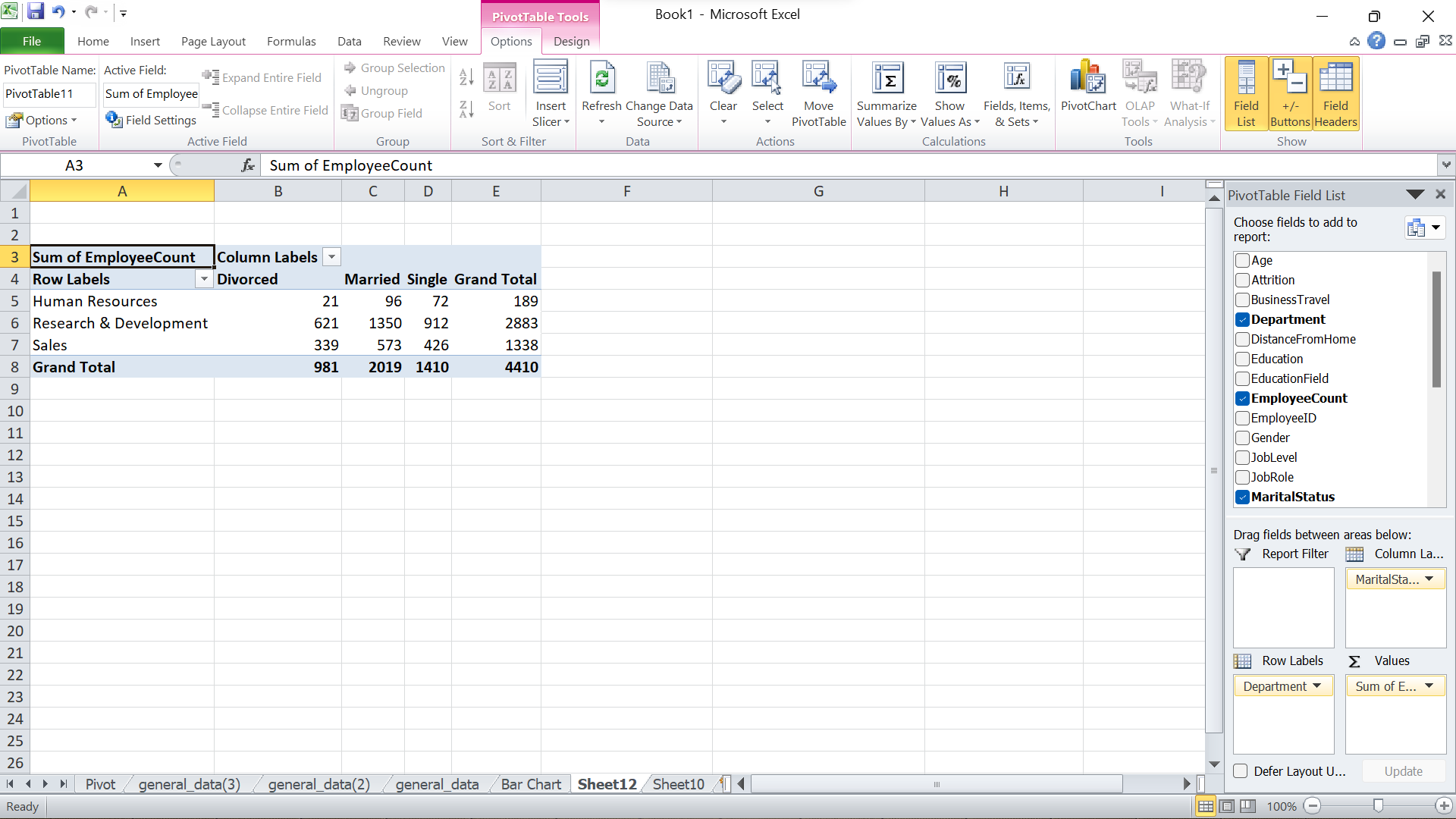
The Department column indicates that the data is consistent, clean and doesn’t contain any missing values.

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



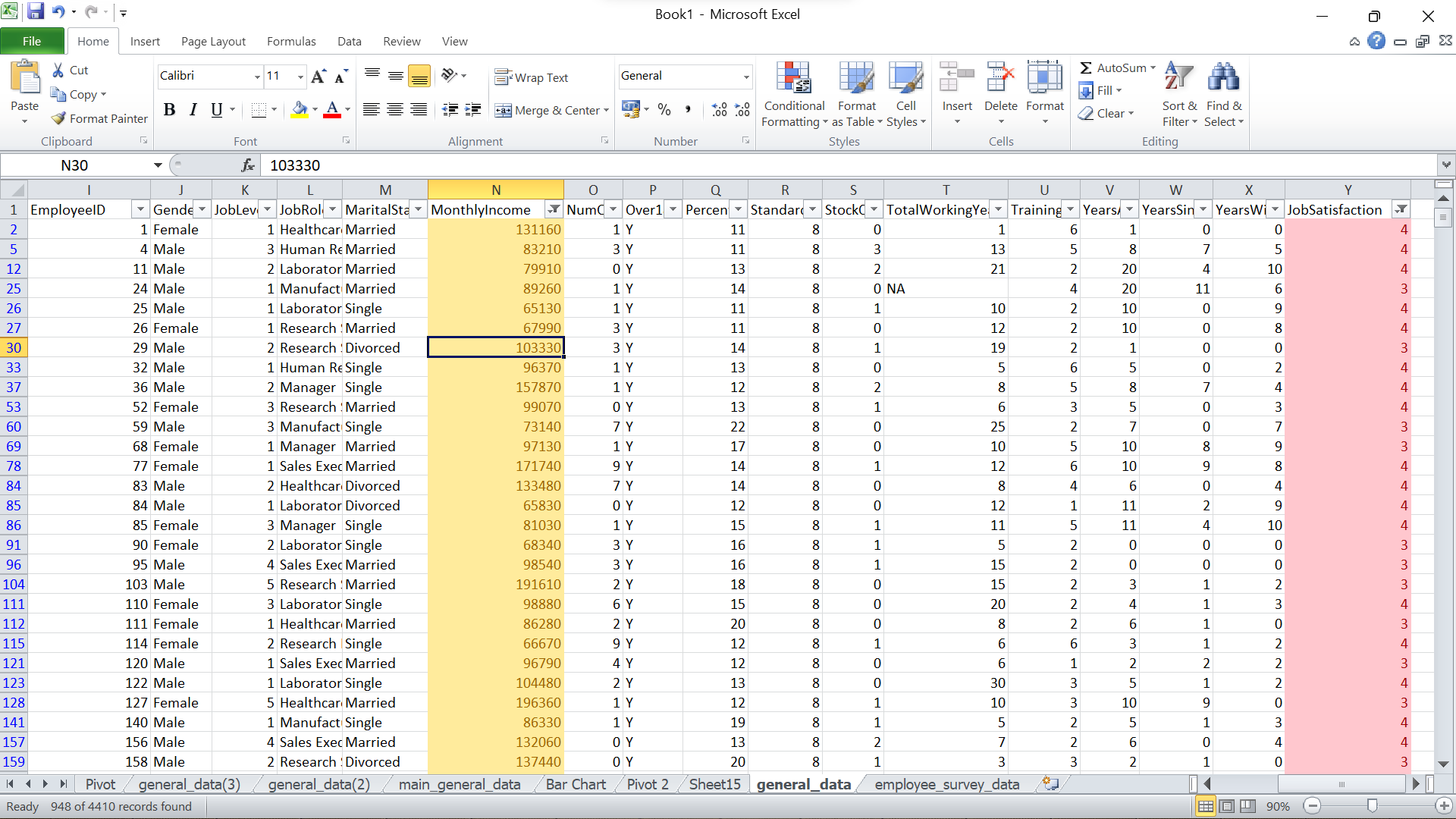
7. Using DAX, create a calculated column that calculates the average years an employee has  
spent with their current manager.

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

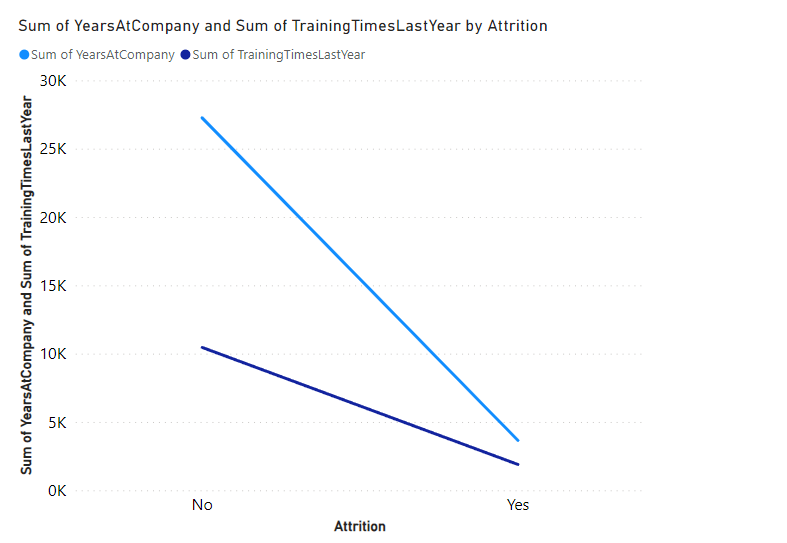


9. Apply conditional formatting to highlight employees with both above-average Monthly Income and above-average Job Satisfaction.

Total ‘948’ employees have both above-average Monthly Income and above-average Job Satisfaction, for this result, I used “VLOOKUP” to get the job satisfaction column from employee\_survey\_data.



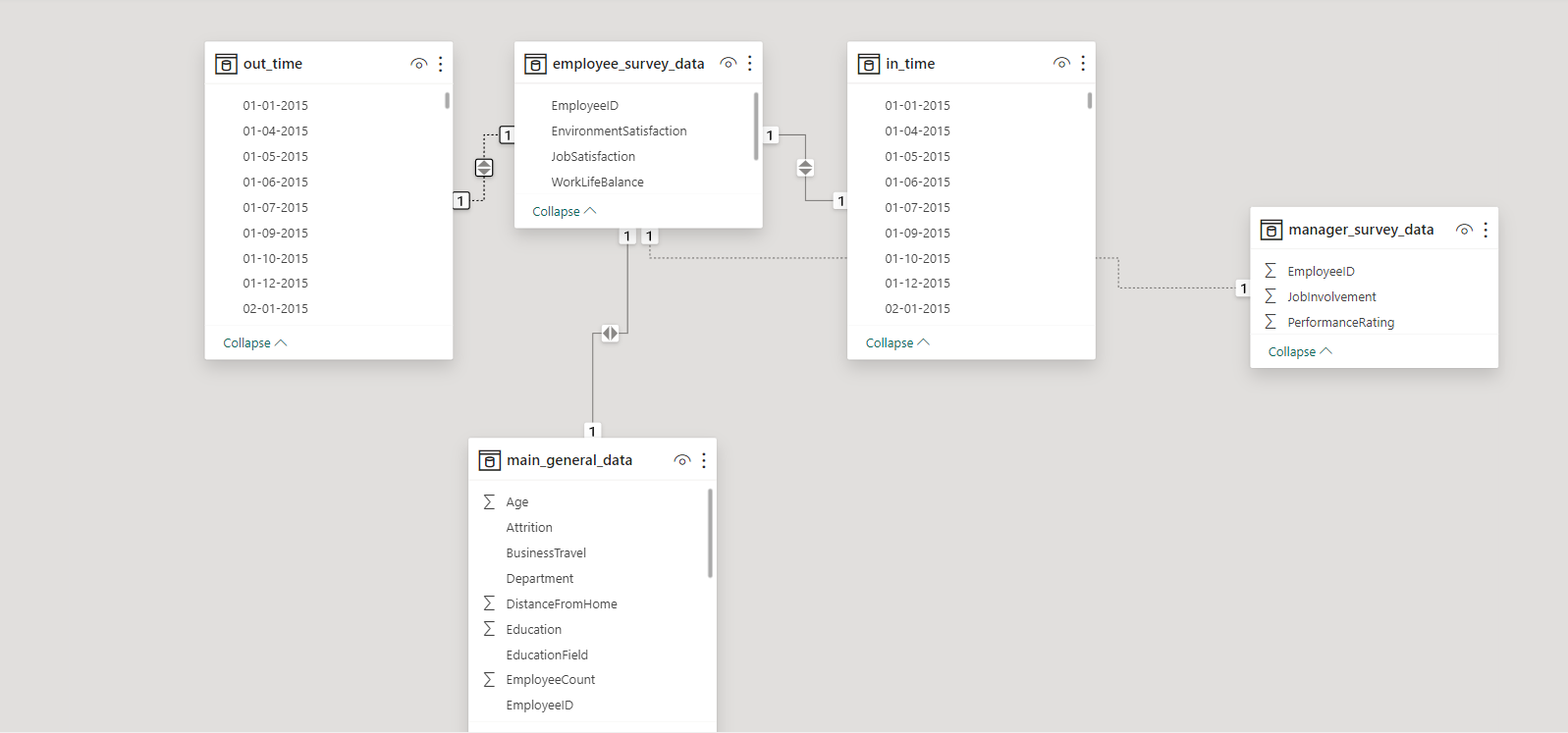
10. In Power BI, create a line chart that visualizes the trend of Employee Attrition over the  
years.



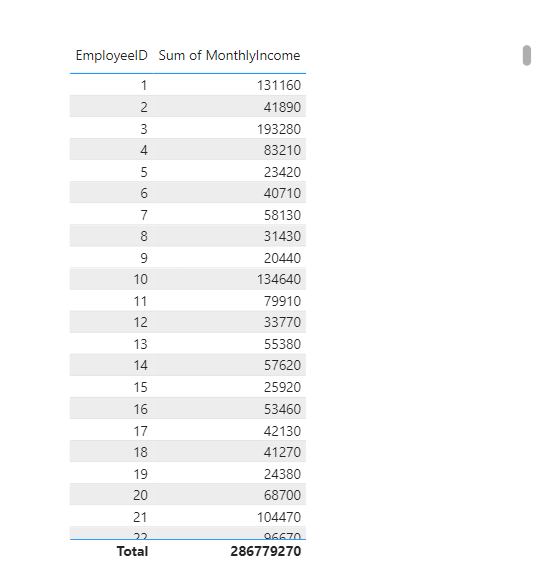
11. Describe how you would create a star schema for this dataset, explaining the benefits of  
doing so.

Primary key of fact table :-“Employee ID”  
Dimension Tables:

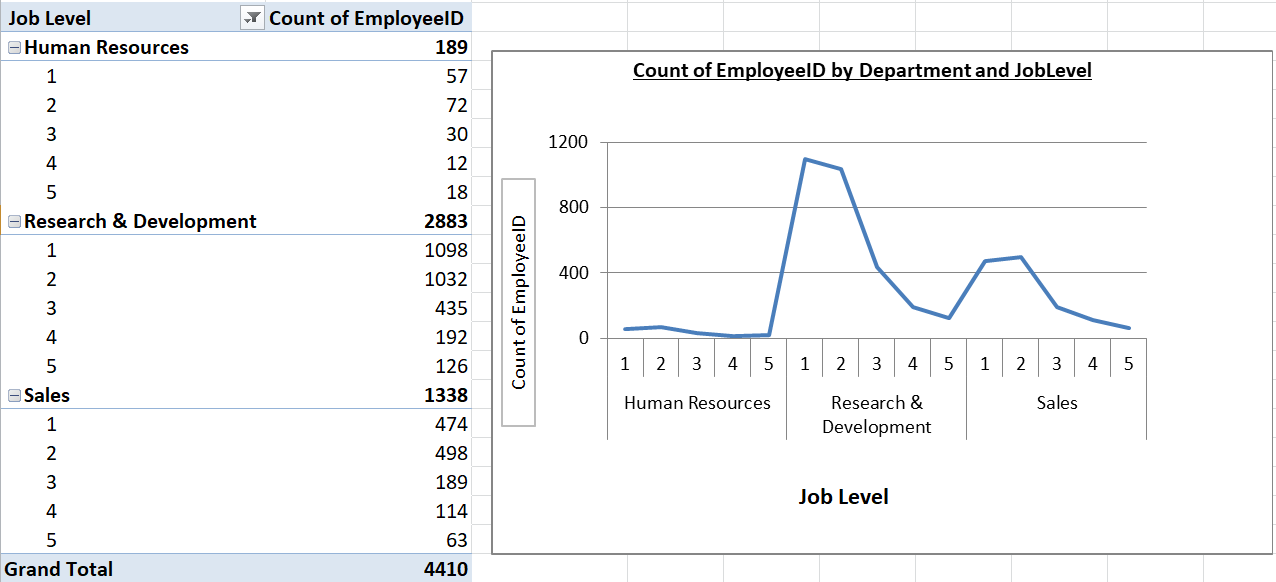
• Manager\_Details  
• Employee\_Data  
• Time\_Data  
The benefits of Star Schema:  
**Simpler Queries** – Join logic of star schema is quite cinch in comparison to other join logic which are needed to fetch data from a transactional schema that is highly normalized.  
**Simplified Business Reporting Logic** – In comparison to a transactional schema that is highly normalized, the star schema makes simpler common business reporting logic, such as of reporting and period-over-period.  
**Feeding Cubes** – Star schema is widely used by all OLAP systems to design OLAP cubes efficiently. In fact, major OLAP systems deliver a ROLAP mode of operation which can use a star schema as a source without designing a cube structure.  
**Easier Maintenance**- The separation of dimensions and facts simplifies maintenance tasks. Updating or adding new data elements can be done independently within each dimension, making the schema more modular and easier to maintain.  
**Scalability**- Star schemas are scalable and flexible, allowing for the addition of new dimensions or facts without disrupting the existing structure. This scalability is crucial for accommodating evolving business requirements and data growth.



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



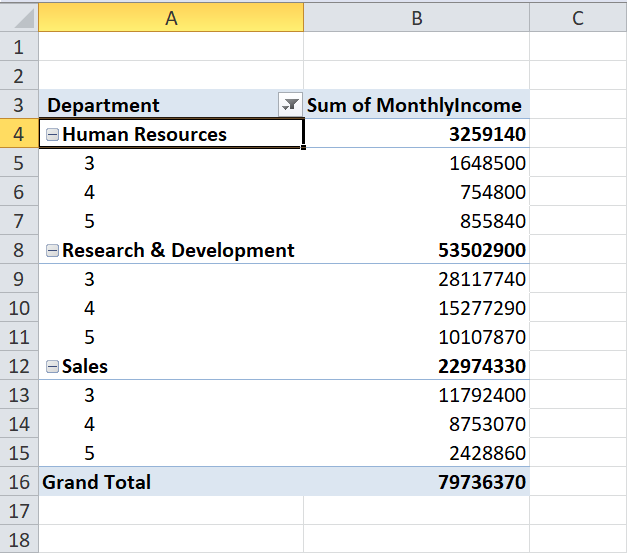
13. Create a hierarchy in Power BI that allows users to drill down from Department to Job  
Role to further narrow their analysis.



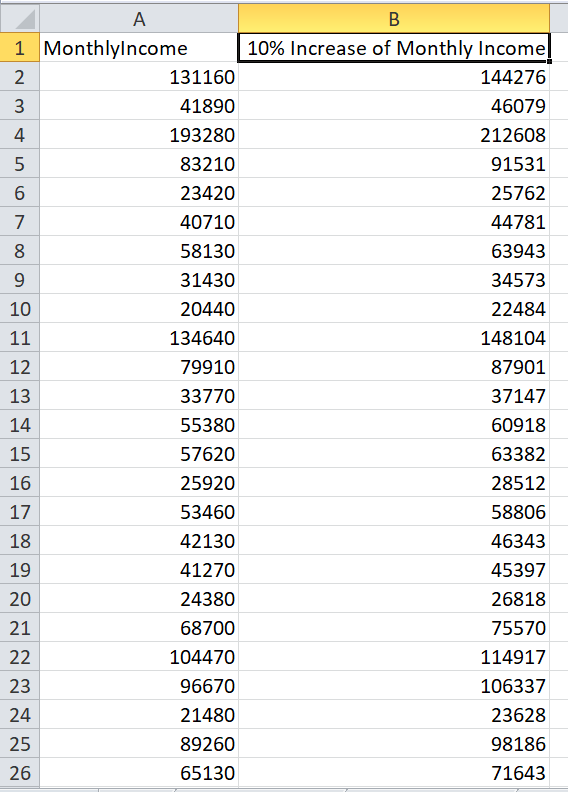
14. How can you set up parameterized queries in Power BI to allow users to filter data based on the Distance from Home column?

* Create a parameter in the Manage Parameters window.
* In Power Query, apply filter to the “Distance from Home” column using the created parameter.
* Load and apply changes to the data.
* Create a visual in PowerBI report.
* Add a slicer for the parameter in the slicer visualization.
* Users can use the slicer to set the distance filter, dynamically updating the data in their reports.

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?

Verifying data adherence to a predefined schema is critical for ensuring its reliability and consistency. If inconsistencies are detected, a systematic approach is necessary. First, pinpoint the specific discrepancies, noting fields or records that deviate from the expected schema. Investigate the root causes, which could range from data entry errors to changes in source systems. Once identified, take corrective actions by cleansing or transforming the data to align with the predefined schema. Implement  
validation rules to prevent future inconsistencies and communicate findings to relevant stakeholders. Ongoing monitoring processes should be established to promptly catch and rectify any emerging discrepancies, maintaining data quality over time. Continuous improvement efforts, including documentation updates and user training, are essential to enhance data quality measures systematically. The goal is to create a robust system that consistently produces reliable and accurate data for informed decision-making.

# 