DATA ANALYSIS INTERNSHIP

Task 1: HR DATA ANALYSIS ASSESSMENT

BY

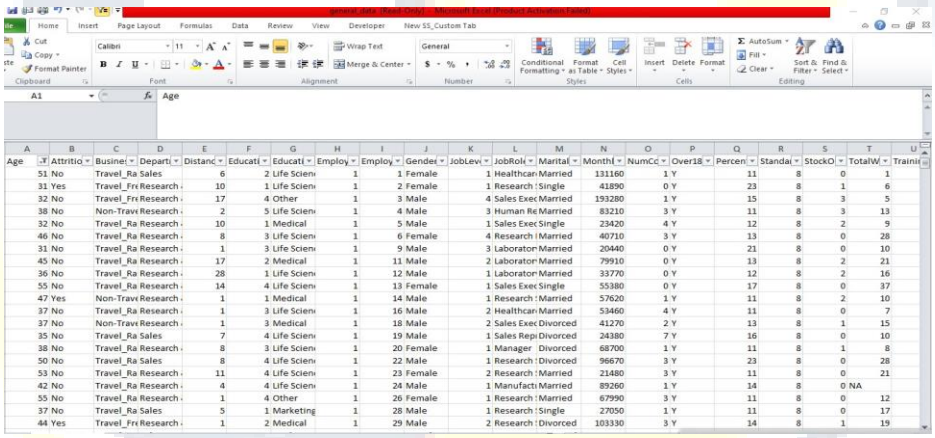
SHUBHANGI SHINDE

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

Get the general_data.csv into Excel.

Go to the Home tab > Sort & Filter > Filter. Before applying filter make sure to choose

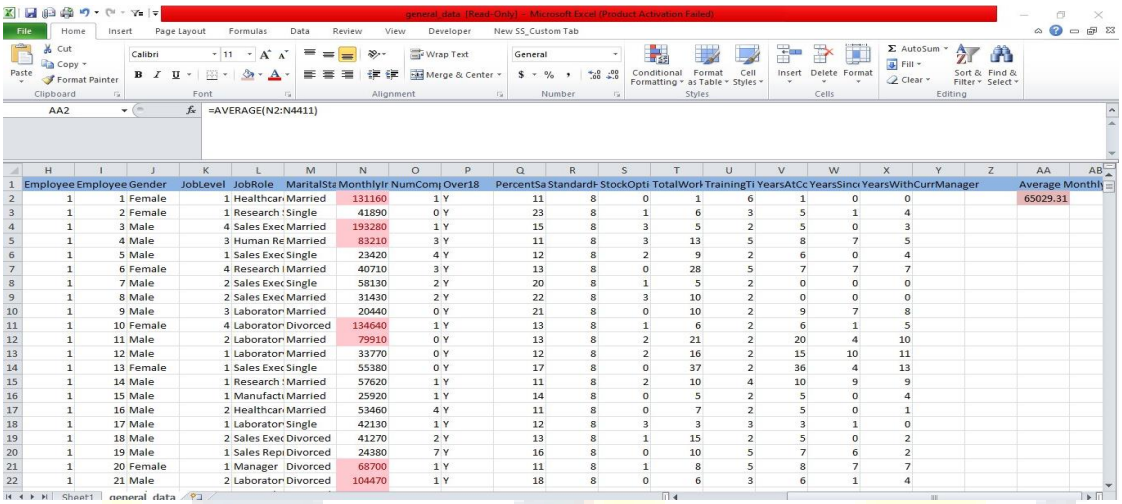
1. right column or CTRL+ SHIFT + L for applying filter
2. 2. Select Age column and click on drop down and select the Number filter > Greater than > Open Custom AutoFilter mention the age value click on OK.



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

Job Role	Average of 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.



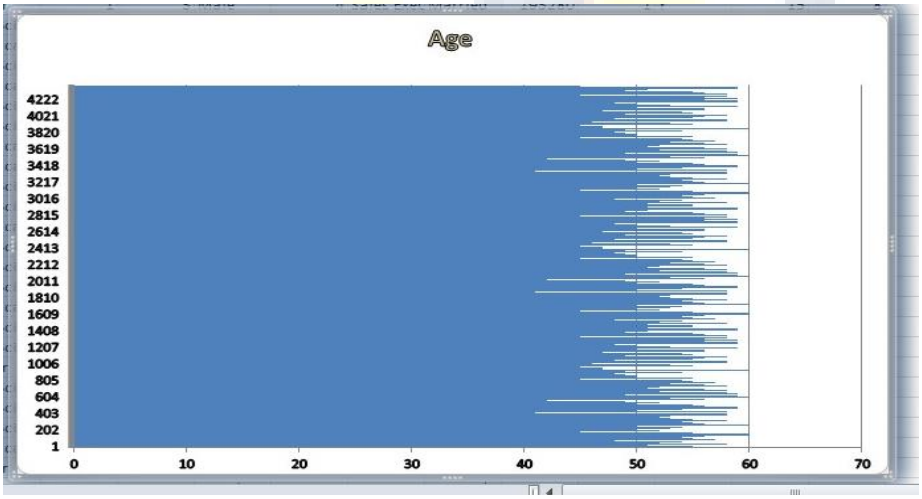
Select Monthly Income column and from Home tab > Conditional formatting > Top/Bottom Rules > Above Average.

For Calculating the Average Monthly income – =Average(range).

Observation:

We came to know that there are 1479 employees having monthly income greater than company’s average income is 65029.

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

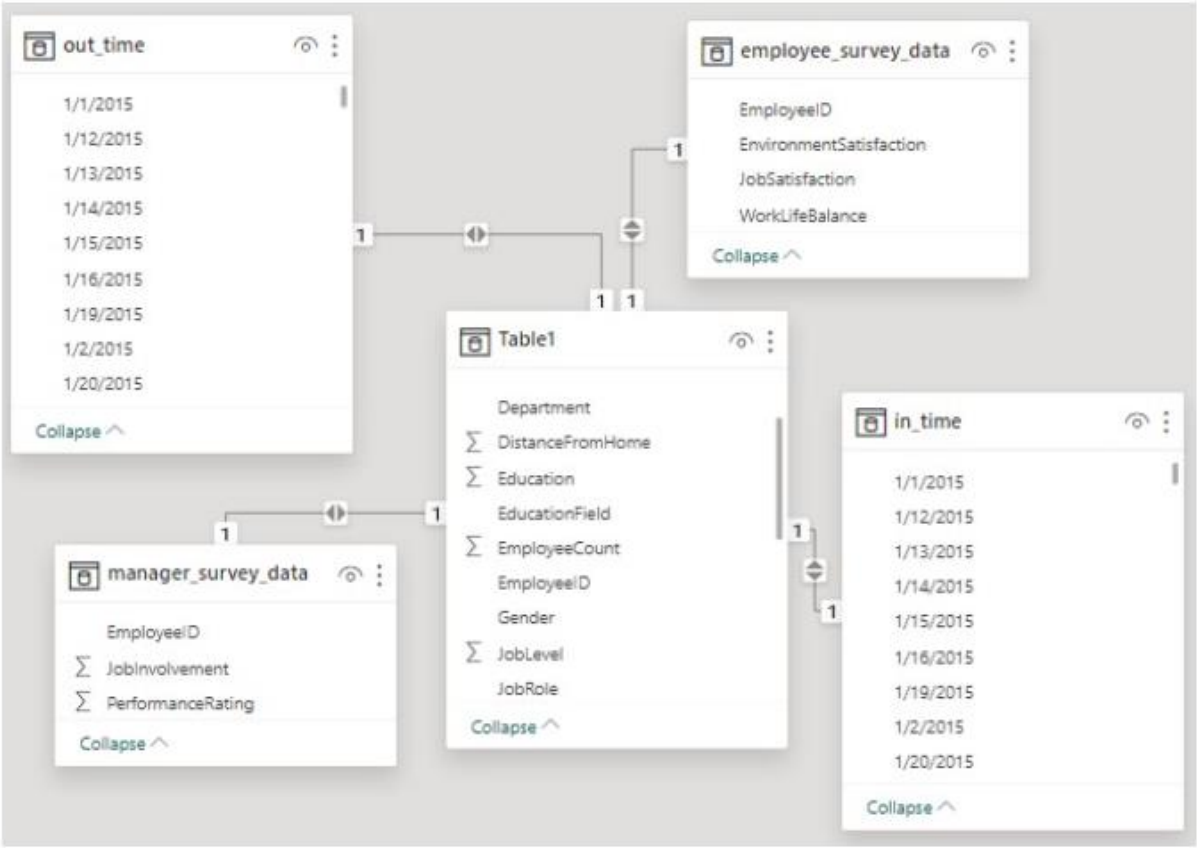


5. Identify and clean any missing or inconsistent data in the "Department" column.

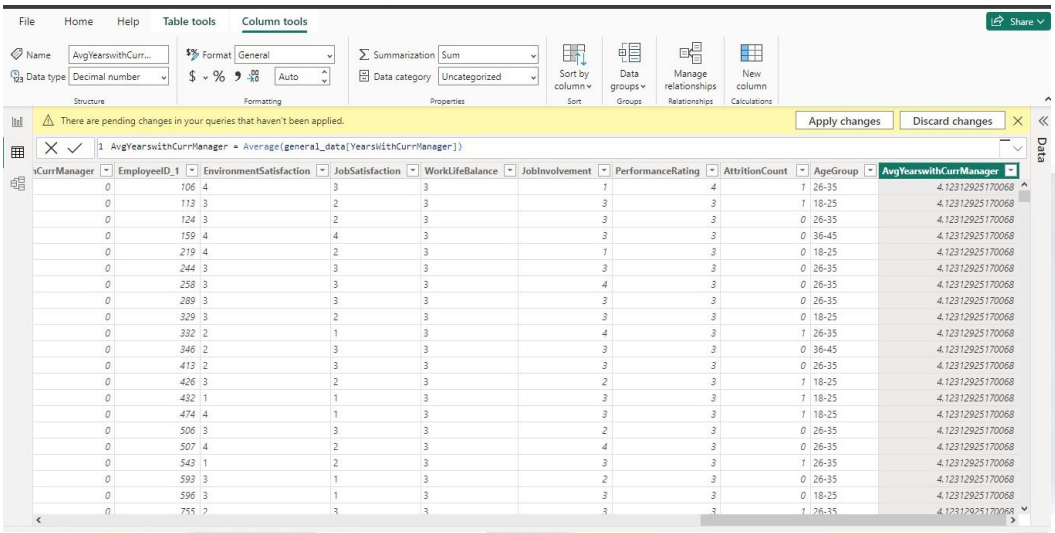
While analyzing the data, there are neither missing values not inconsistent data in the department column.

If we wanted to know about the missing values from an particular column. You can add filter and select the balnk values OR you can use COUNTBLANK funcation- =COUNTBLANK(given_range) OR from Home tab >conditional fromatting > New rules > format only cells that contain > from format only cells with : select cell value dropdown select - blank dropdown.

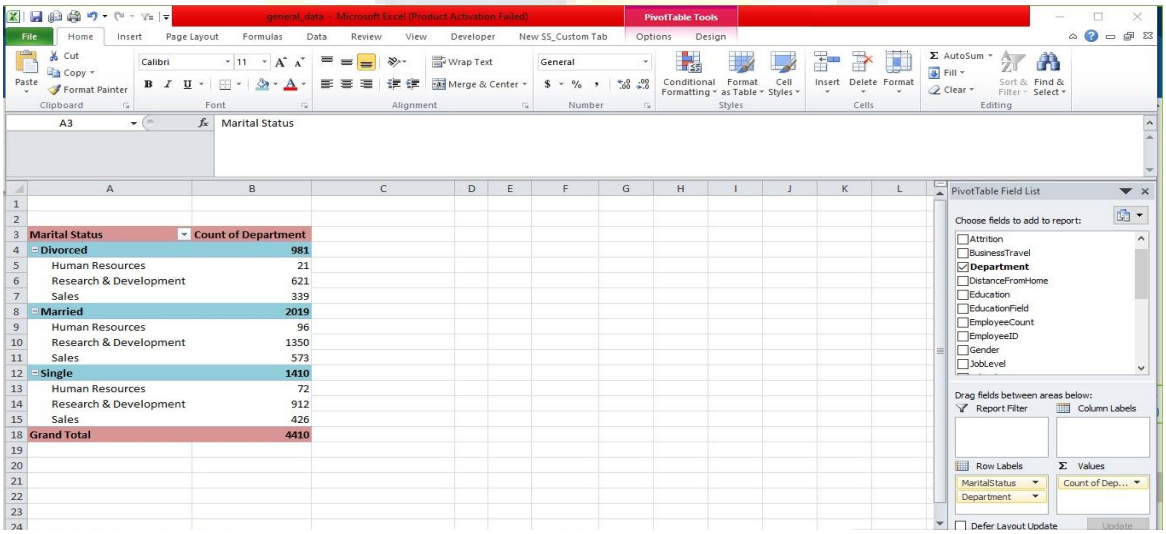
6. In Power BI, establish a relationship between the "EmployeeID" in the employee data and the "EmployeeID" in the time tracking data.



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



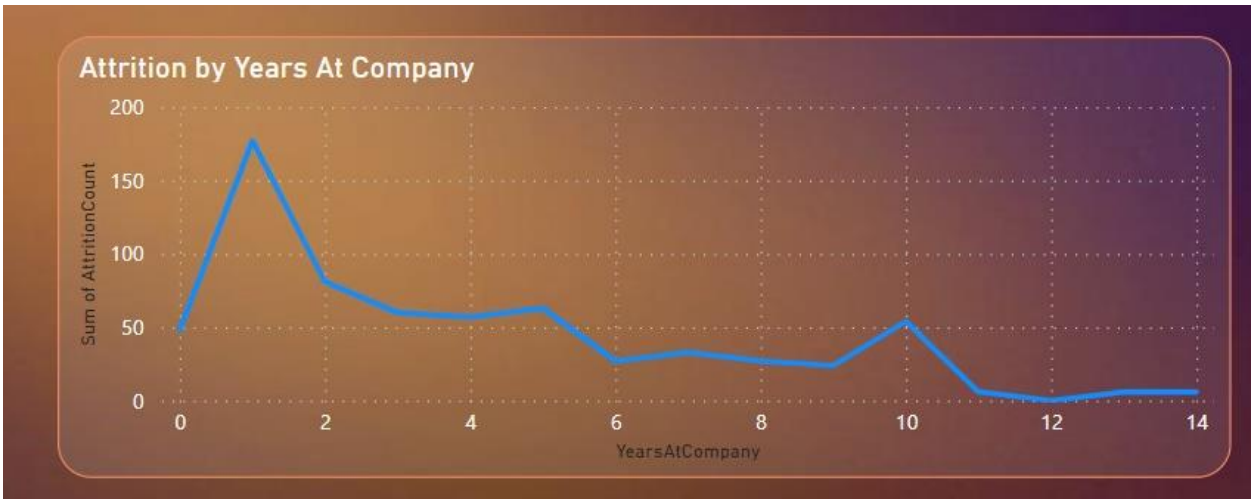
8. Using Excel, create a pivot table that displays the count of employees in each Marital Status category, segmented by Department.



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

I	N	Y	Z	AA
Employee	MonthlyIr	JobSatisfaction		
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.

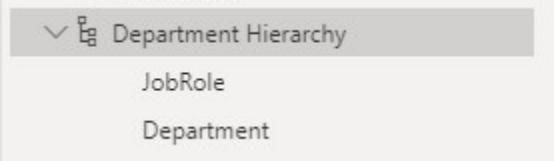


11. Describe how you would create a star schema for this dataset, explaining the benefits of doing so.

The Benefits of Star Schema

- ❖ Star schema are efficient for storing data and updating data by reducing the duplication.
- ❖ It is extremely simple to understand and build.
- ❖ No need for complex joins when querying data.

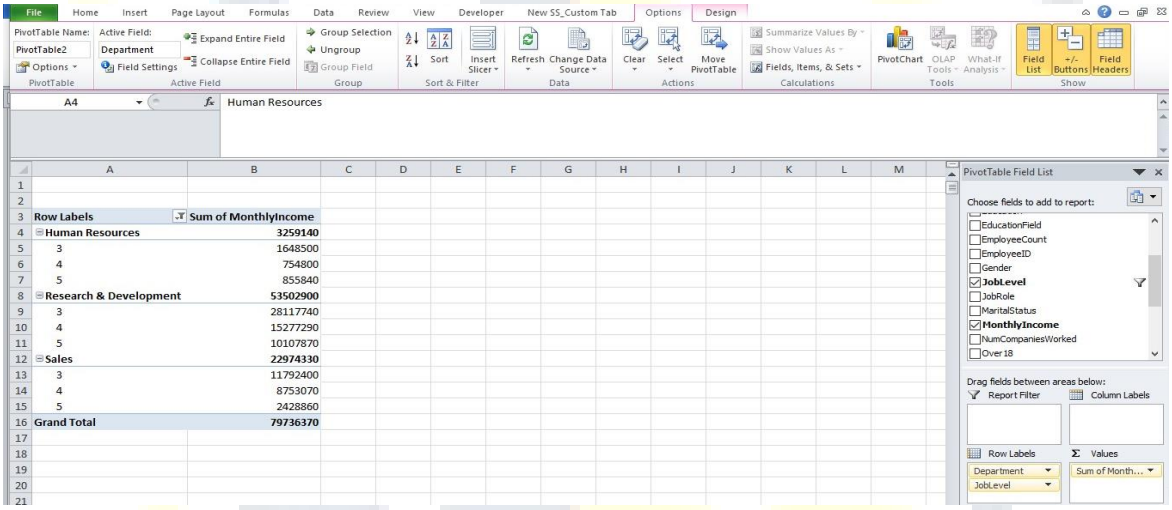
12. Create a hierarchy in Power BI that allows users to drill down from Department to Job Role to further narrow their analysis.



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

- ❖ Select Home > Transform data > Transform data to open the Power Query Editor.
- ❖ In the Power Query Editor, select New Parameters under Manage Parameters in the ribbon.
- ❖ In the Manage Parameters window, fill out the information about the parameter

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



Row Labels	Sum of MonthlyIncome
Human Resources	3259140
3	1648500
4	754800
5	855840
Research & Development	53502900
3	28117740
4	15277290
5	10107870
Sales	22974330
3	11792400
4	8753070
5	2428860
Grand Total	79736370

15. Explain how to perform a What-If analysis in Excel to understand the

- Go to the Data Tab

- In the "Data Tools" group, click "What-If Analysis" and choose "Data Table."

- In "Row input cell," enter the reference to the cell with Percent Salary Hike.

- Click OK

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

Steps to Verify Data Adherence to a Predefined Schema:

➢ Understand Schema: Define expected data types and constraints.

➢ Use Profiling Tools: Analyse data characteristics with profiling tools.

➢ Check Data Types: Verify data types match the schema.

➢ Validate Constraints: Ensure constraints are not violated.

➢ Identify Missing Values: Check for missing values in mandatory fields.

➢ Look for Outliers: Investigate unexpected values or outliers.

➢ Validate Relationships: Confirm valid relationships between tables.

Actions if Inconsistencies are Found:

➢ Data Cleansing: Correct inaccurate data.

➢ Communicate: Discuss inconsistencies with stakeholders.

➢ Document: Document issues and resolutions.

➢ Implement Validation Rules: Enforce data validation rules.

➢ Review ETL Processes: Ensure ETL processes align with schema.

➢ Implement Quality Checks: Integrate automated data quality checks.

➢ Consider Governance: Evaluate and improve data governance practices.

➢ Training and Documentation: Train personnel and update documentation

DATA INSIGHTS ➢

Among 4410 employees, there are 2650 males and 1760 females. Also, 3699 active employees with average age comes with 37.

- The average monthly income of the employees is 65030 and there are 711 people left the company.
- Around 453 people from the dept. Research and development left the company, 201 employees from the dept. sales and less amount of 57 from the dept. Human Resources.
- Attrition rate is max for Sales Executives dept
- Sales department employees are mostly satisfied in there work.
- Maximum peoples are working for 1-2 years


HR ANALYTICS DASHBOARD
Department
Human Resources
Research & Development
Sales
Attrition by Gender
Male
Female
2.65K
1.76K
Count Of Employee
4410
Attrition
711
Avg Age
37
Avg Salary
65K
Avg Years
7
Sum of AttritionCount by EducationField
303 (42.62%)
45 (6.3...)
75 (10.55%)
225 (31.65%)
Attrition by Age
2,000
1,500
1,000
500
0
26-35
36-45
46-55
18-25
55+
AgeGroup
Job Satisfaction
JobRole
Healthcare Representative 18 9 18 12 57
Human Resources 6 3 6 6 21
Laboratory Technician 36 36 36 18 126
Manager 11 3 15 12 41
Manufacturing Director 21 9 3 15 48
Research Director 18 3 27 9 57
Research Scientist 48 48 33 30 159
Sales Executive 27 24 72 42 165
Total 197 138 219 156 710
Attrition by Job Role
Sales Executive
Research Scientist
Laboratory Technician
Healthcare Representative
Research Director
Manufacturing Director
Manager
0
50
100
150
200
Attrition by Years At Company
200
150
100
50
0
0
2
4
6
8
10
12
14