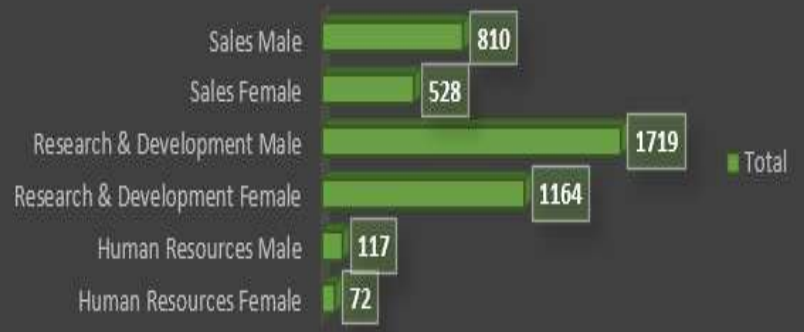
HR DATA ANALYSIS

Pysliq Task II

Departments count

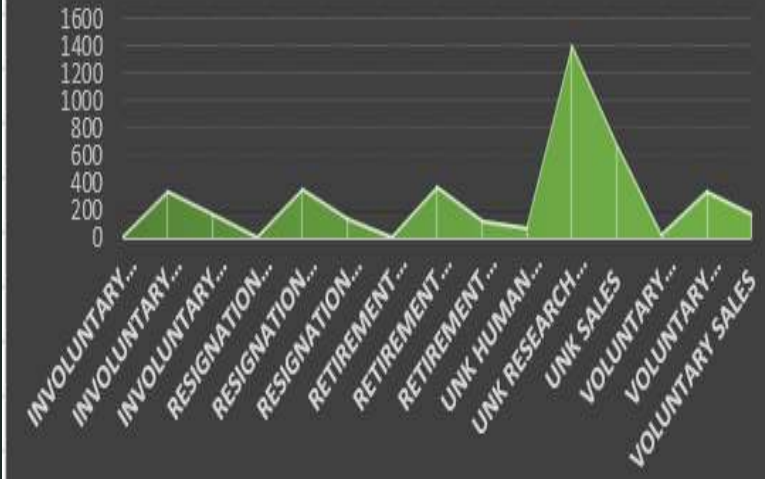
	Human Resources	Research & Development	Sales
Total	189	2883	1338

GENDER COUNT IN EACH DEPARTMENT

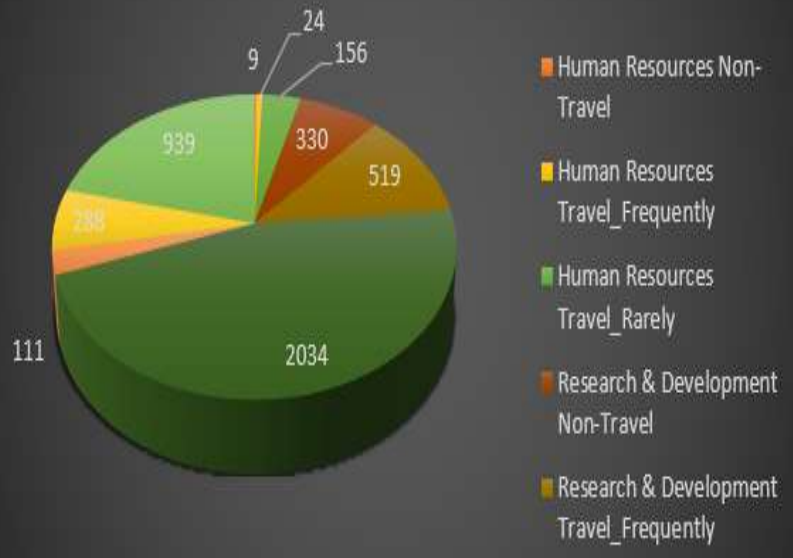


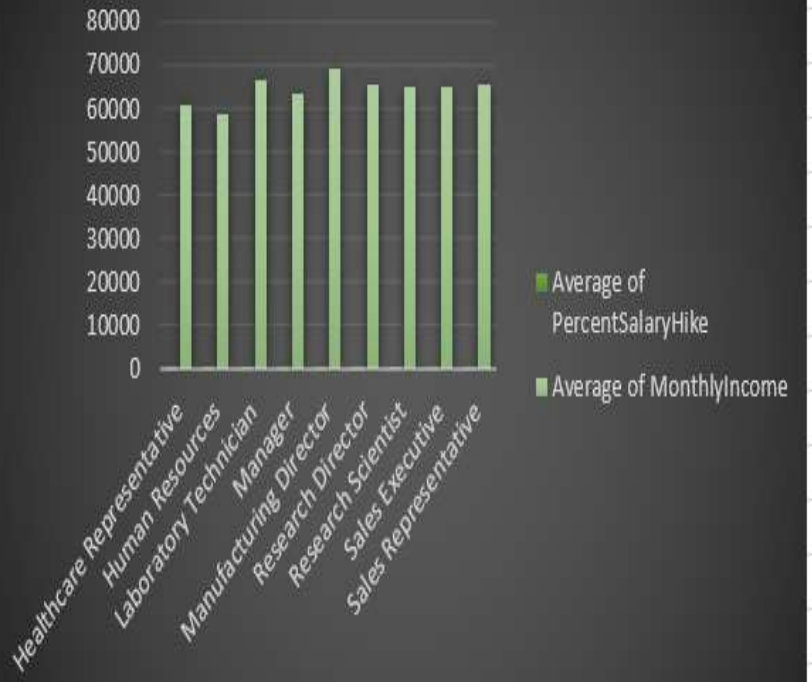
Count of MonthlyIncome with Martial Status

Termination type in each department

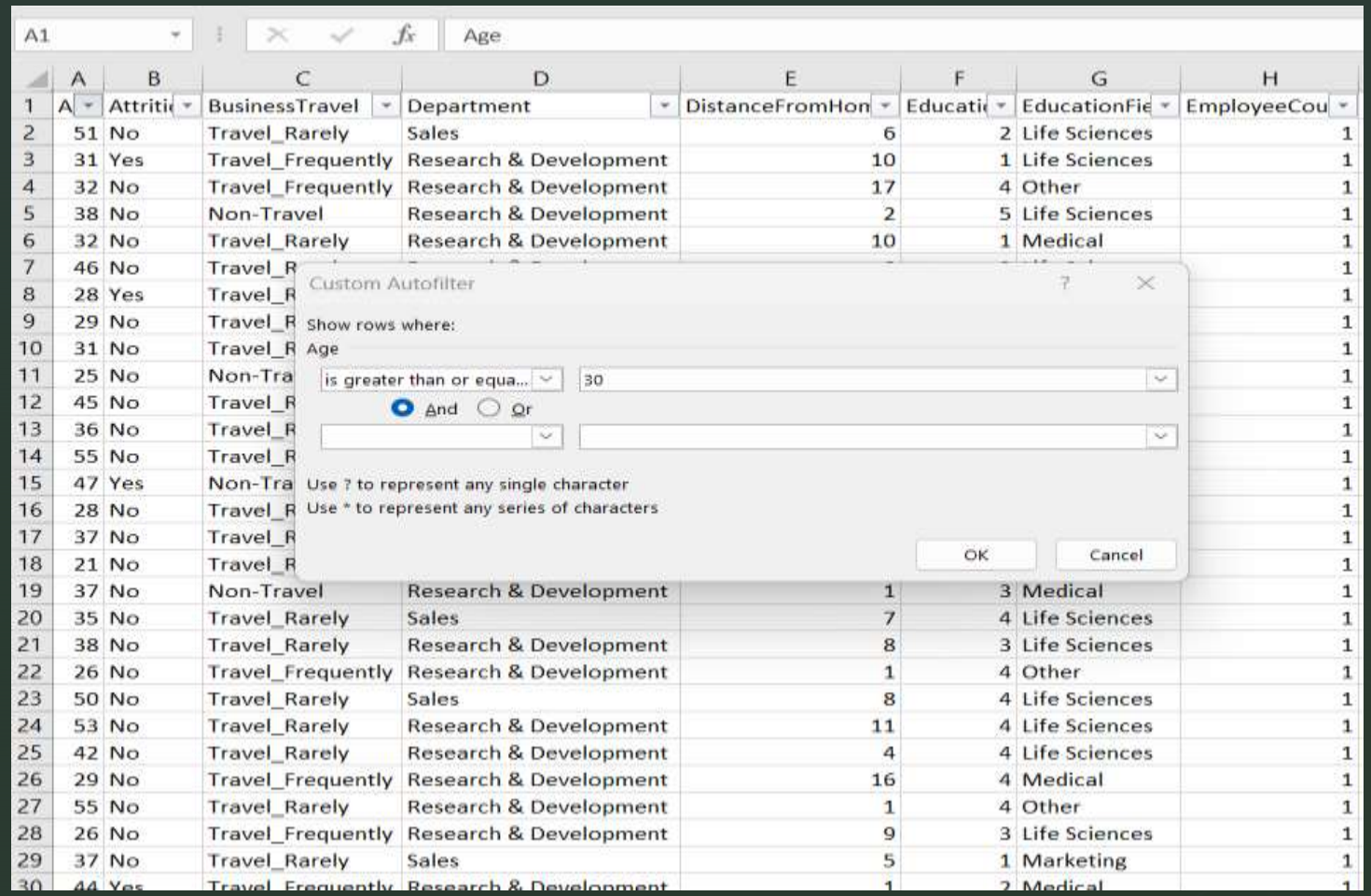


Business Travel in Each Department





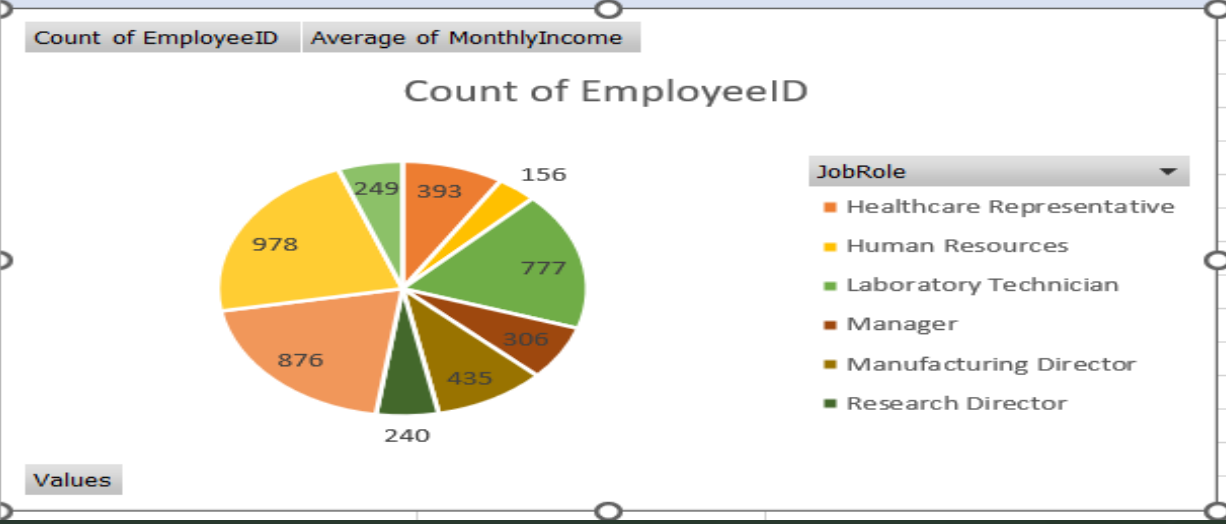
- Using Excel, how would you filter the dataset to only show employees aged 30 and above



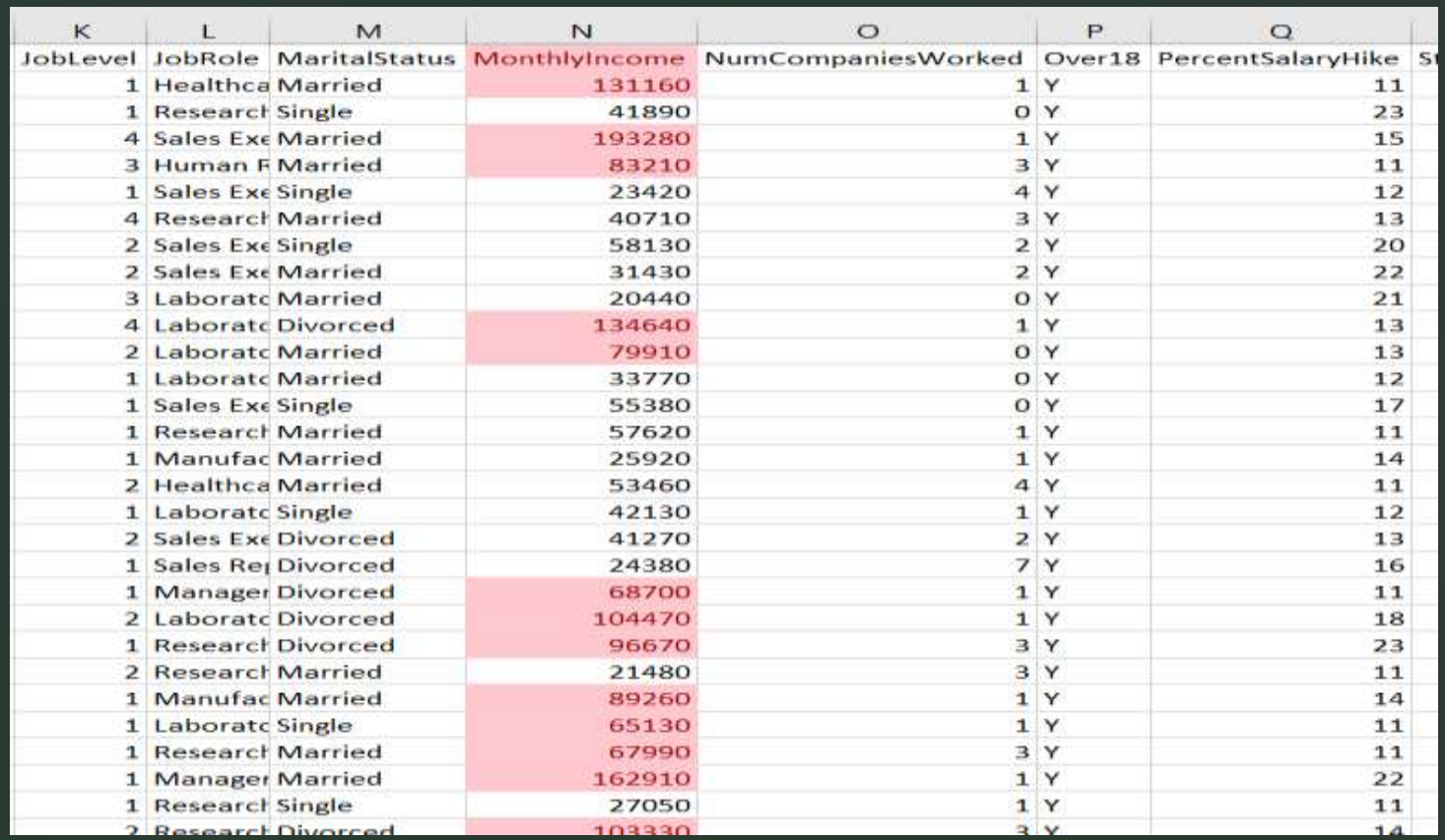
	A	B	C	D	E	F	G	H
1	A	Attriti	BusinessTravel	Department	DistanceFromHon	Educati	EducationFie	EmployeeCou
2	51	No	Travel_Rarely	Sales	6	2	Life Sciences	1
3	31	Yes	Travel_Frequently	Research & Development	10	1	Life Sciences	1
4	32	No	Travel_Frequently	Research & Development	17	4	Other	1
5	38	No	Non-Travel	Research & Development	2	5	Life Sciences	1
6	32	No	Travel_Rarely	Research & Development	10	1	Medical	1
7	46	No	Travel_R					1
8	28	Yes	Travel_R					1
9	29	No	Travel_R					1
10	31	No	Travel_R					1
11	25	No	Non-Tra					1
12	45	No	Travel_R					1
13	36	No	Travel_R					1
14	55	No	Travel_R					1
15	47	Yes	Non-Tra					1
16	28	No	Travel_R					1
17	37	No	Travel_R					1
18	21	No	Travel_R					1
19	37	No	Non-Travel	Research & Development	1	3	Medical	1
20	35	No	Travel_Rarely	Sales	7	4	Life Sciences	1
21	38	No	Travel_Rarely	Research & Development	8	3	Life Sciences	1
22	26	No	Travel_Frequently	Research & Development	1	4	Other	1
23	50	No	Travel_Rarely	Sales	8	4	Life Sciences	1
24	53	No	Travel_Rarely	Research & Development	11	4	Life Sciences	1
25	42	No	Travel_Rarely	Research & Development	4	4	Life Sciences	1
26	29	No	Travel_Frequently	Research & Development	16	4	Medical	1
27	55	No	Travel_Rarely	Research & Development	1	4	Other	1
28	26	No	Travel_Frequently	Research & Development	9	3	Life Sciences	1
29	37	No	Travel_Rarely	Sales	5	1	Marketing	1
30	44	Yes	Travel_Frequently	Research & Development	1	2	Medical	1

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

Row Labels	Count of EmployeeID	Average of MonthlyIncome
Healthcare Representative	393	60983.74046
Human Resources	156	58528.07692
Laboratory Technician	777	66314.05405
Manager	306	63395.88235
Manufacturing Director	435	69183.72414
Research Director	240	65473.125
Research Scientist	876	64975.68493
Sales Executive	978	65186.68712
Sales Representative	249	65370.96386
Grand Total	**4410**	**65029.31293**

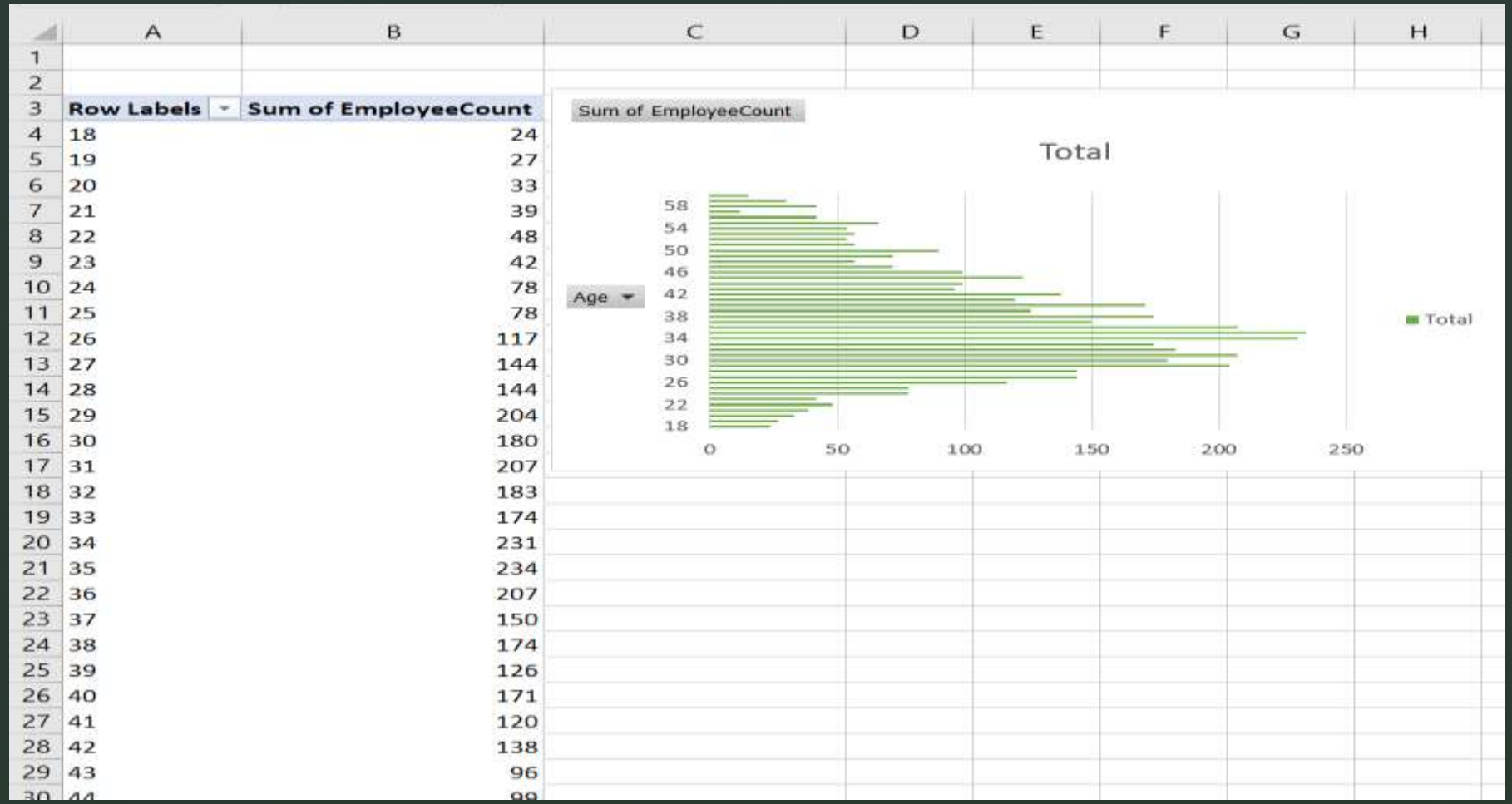


- Apply conditional formatting to highlight employees with Monthly Income above the company's average income.



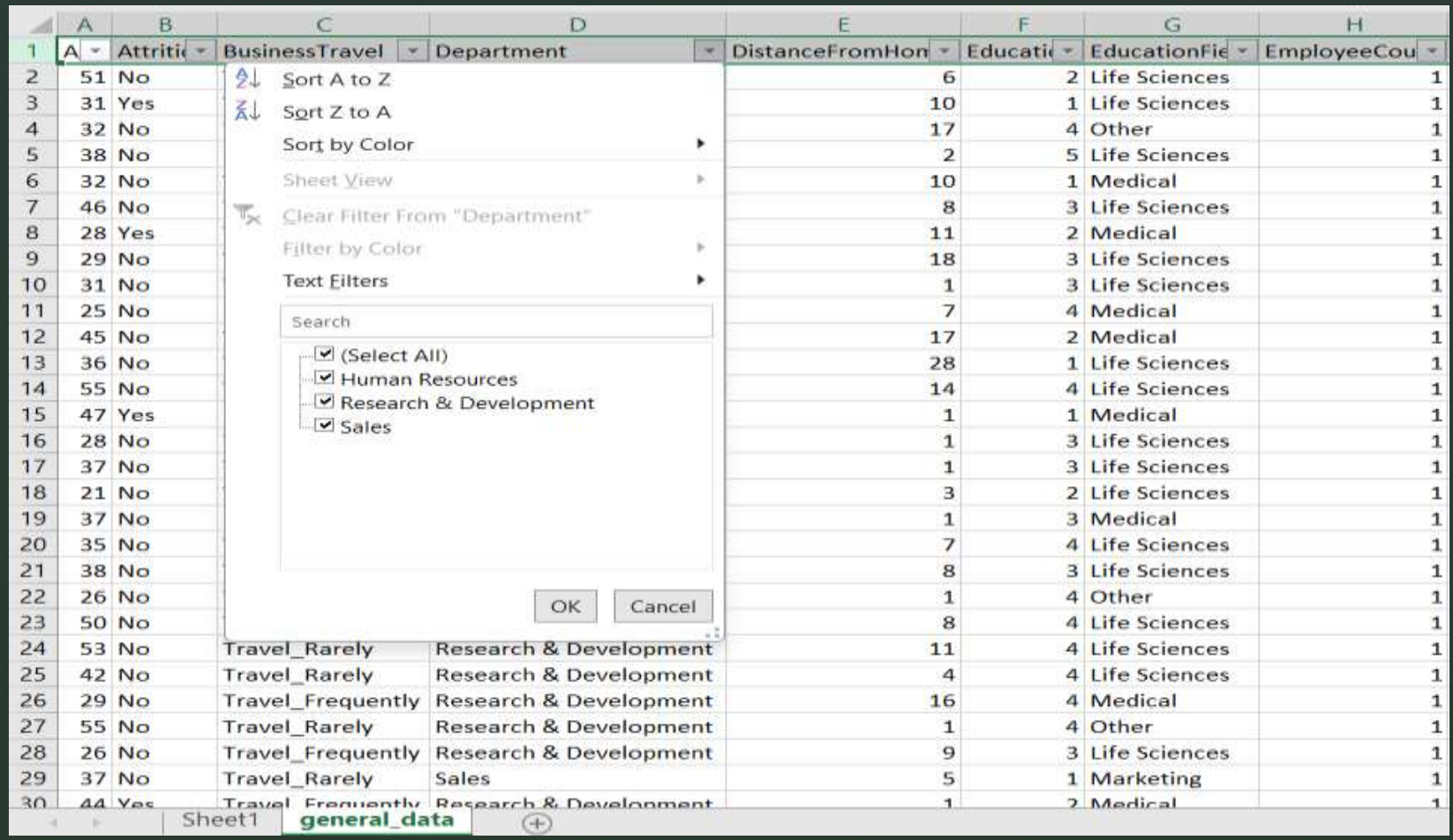
K	L	M	N	O	P	Q	
JobLevel	JobRole	MaritalStatus	MonthlyIncome	NumCompaniesWorked	Over18	PercentSalaryHike	St
1	Healthca	Married	131160	1	Y	11	
1	Researcl	Single	41890	0	Y	23	
4	Sales Exe	Married	193280	1	Y	15	
3	Human F	Married	83210	3	Y	11	
1	Sales Exe	Single	23420	4	Y	12	
4	Researcl	Married	40710	3	Y	13	
2	Sales Exe	Single	58130	2	Y	20	
2	Sales Exe	Married	31430	2	Y	22	
3	Laborato	Married	20440	0	Y	21	
4	Laborato	Divorced	134640	1	Y	13	
2	Laborato	Married	79910	0	Y	13	
1	Laborato	Married	33770	0	Y	12	
1	Sales Exe	Single	55380	0	Y	17	
1	Researcl	Married	57620	1	Y	11	
1	Manufac	Married	25920	1	Y	14	
2	Healthca	Married	53460	4	Y	11	
1	Laborato	Single	42130	1	Y	12	
2	Sales Exe	Divorced	41270	2	Y	13	
1	Sales Rep	Divorced	24380	7	Y	16	
1	Manager	Divorced	68700	1	Y	11	
2	Laborato	Divorced	104470	1	Y	18	
1	Researcl	Divorced	96670	3	Y	23	
2	Researcl	Married	21480	3	Y	11	
1	Manufac	Married	89260	1	Y	14	
1	Laborato	Single	65130	1	Y	11	
1	Researcl	Married	67990	3	Y	11	
1	Manager	Married	162910	1	Y	22	
1	Researcl	Single	27050	1	Y	11	
2	Researcl	Divorced	103330	3	Y	14	

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



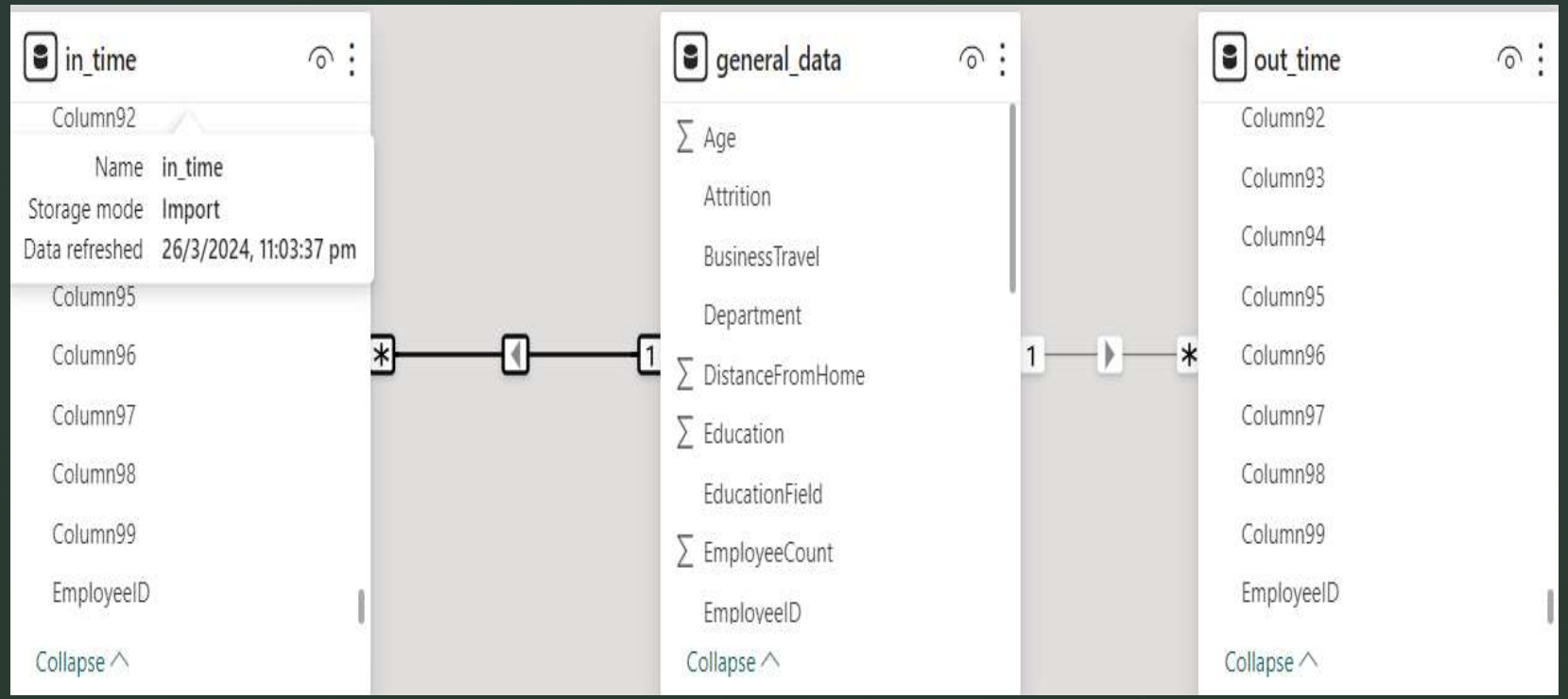
Row Labels	Sum of EmployeeCount
18	24
19	27
20	33
21	39
22	48
23	42
24	78
25	78
26	117
27	144
28	144
29	204
30	180
31	207
32	183
33	174
34	231
35	234
36	207
37	150
38	174
39	126
40	171
41	120
42	138
43	96
44	99

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

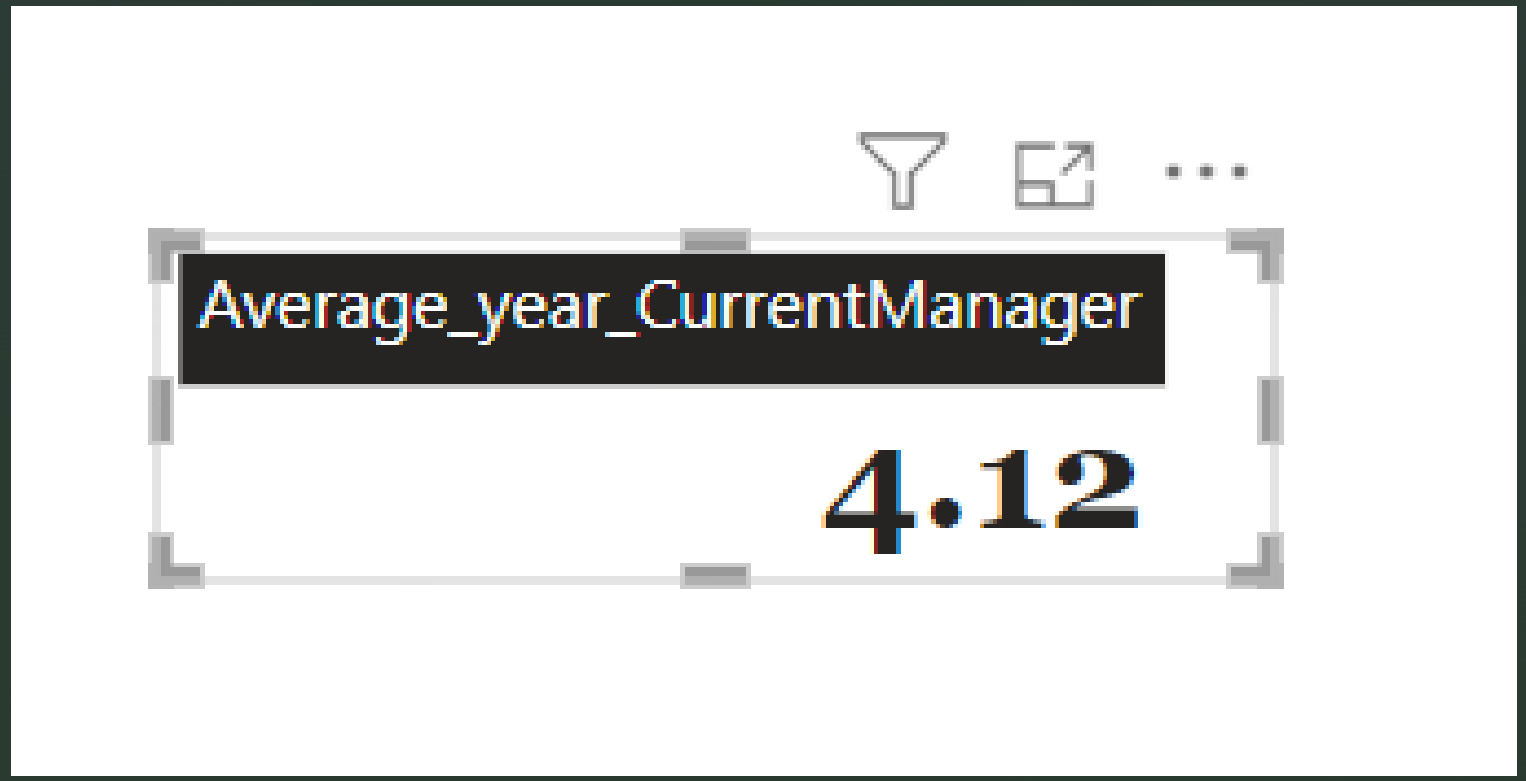


	A	B	C	D	E	F	G	H
1	A	Attriti	BusinessTravel	Department	DistanceFromHon	Educati	EducationFie	EmployeeCou
2	51	No			6	2	Life Sciences	1
3	31	Yes			10	1	Life Sciences	1
4	32	No			17	4	Other	1
5	38	No			2	5	Life Sciences	1
6	32	No			10	1	Medical	1
7	46	No			8	3	Life Sciences	1
8	28	Yes			11	2	Medical	1
9	29	No			18	3	Life Sciences	1
10	31	No			1	3	Life Sciences	1
11	25	No			7	4	Medical	1
12	45	No			17	2	Medical	1
13	36	No			28	1	Life Sciences	1
14	55	No			14	4	Life Sciences	1
15	47	Yes			1	1	Medical	1
16	28	No			1	3	Life Sciences	1
17	37	No			1	3	Life Sciences	1
18	21	No			3	2	Life Sciences	1
19	37	No			1	3	Medical	1
20	35	No			7	4	Life Sciences	1
21	38	No			8	3	Life Sciences	1
22	26	No			1	4	Other	1
23	50	No			8	4	Life Sciences	1
24	53	No	Travel_Rarely	Research & Development	11	4	Life Sciences	1
25	42	No	Travel_Rarely	Research & Development	4	4	Life Sciences	1
26	29	No	Travel_Frequently	Research & Development	16	4	Medical	1
27	55	No	Travel_Rarely	Research & Development	1	4	Other	1
28	26	No	Travel_Frequently	Research & Development	9	3	Life Sciences	1
29	37	No	Travel_Rarely	Sales	5	1	Marketing	1
30	44	Yes	Travel_Frequently	Research & Development	1	2	Medical	1

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



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



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

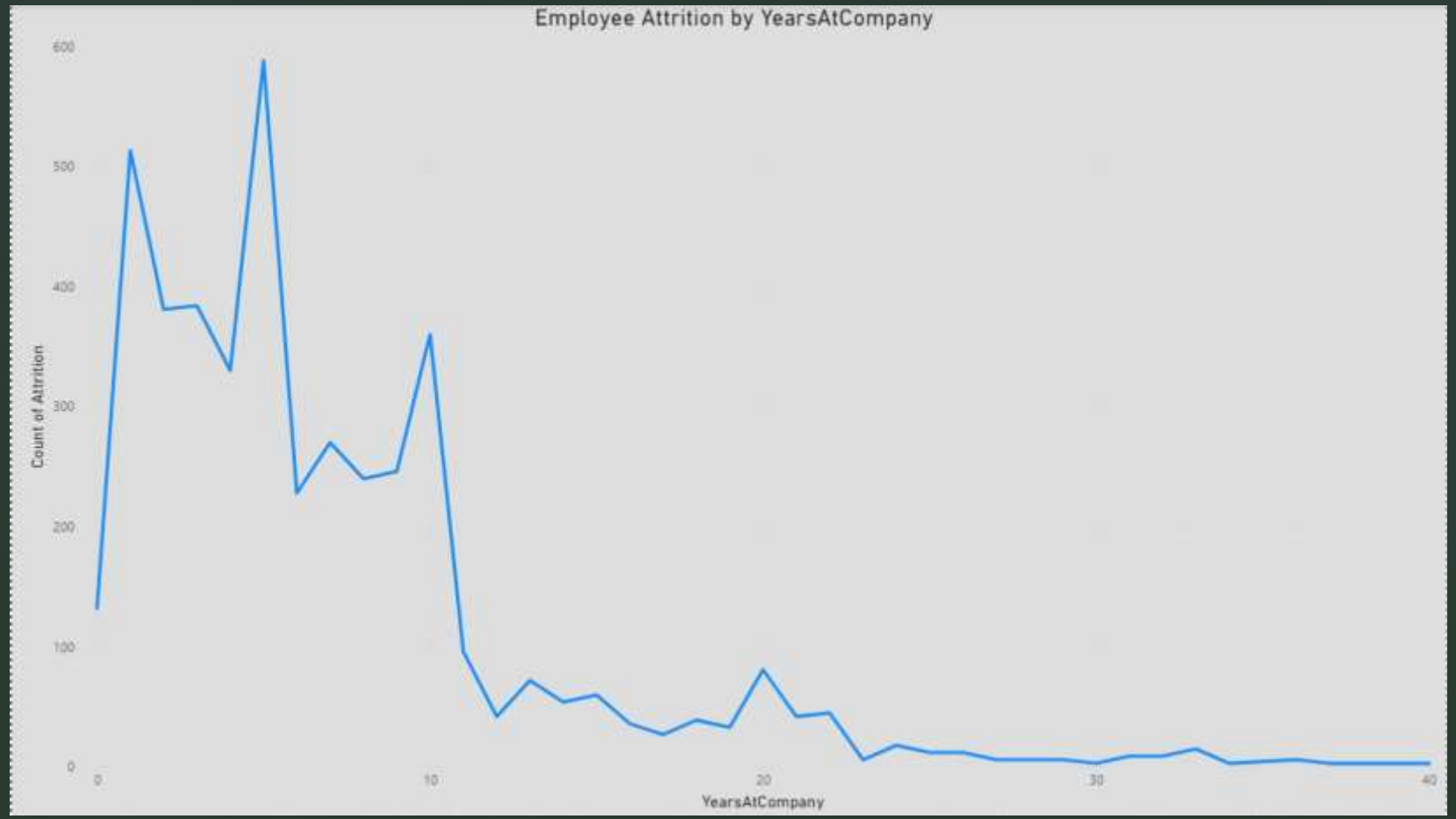
Count of EmployeeID	Column Label			
Marital Status	Human Resources	Research & Development	Sales	Grand Total
Divorced	21	621	339	981
Married	96	1350	573	2019
Single	72	912	426	1410
Grand Total	**189**	**2883**	**1338**	**4410**

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

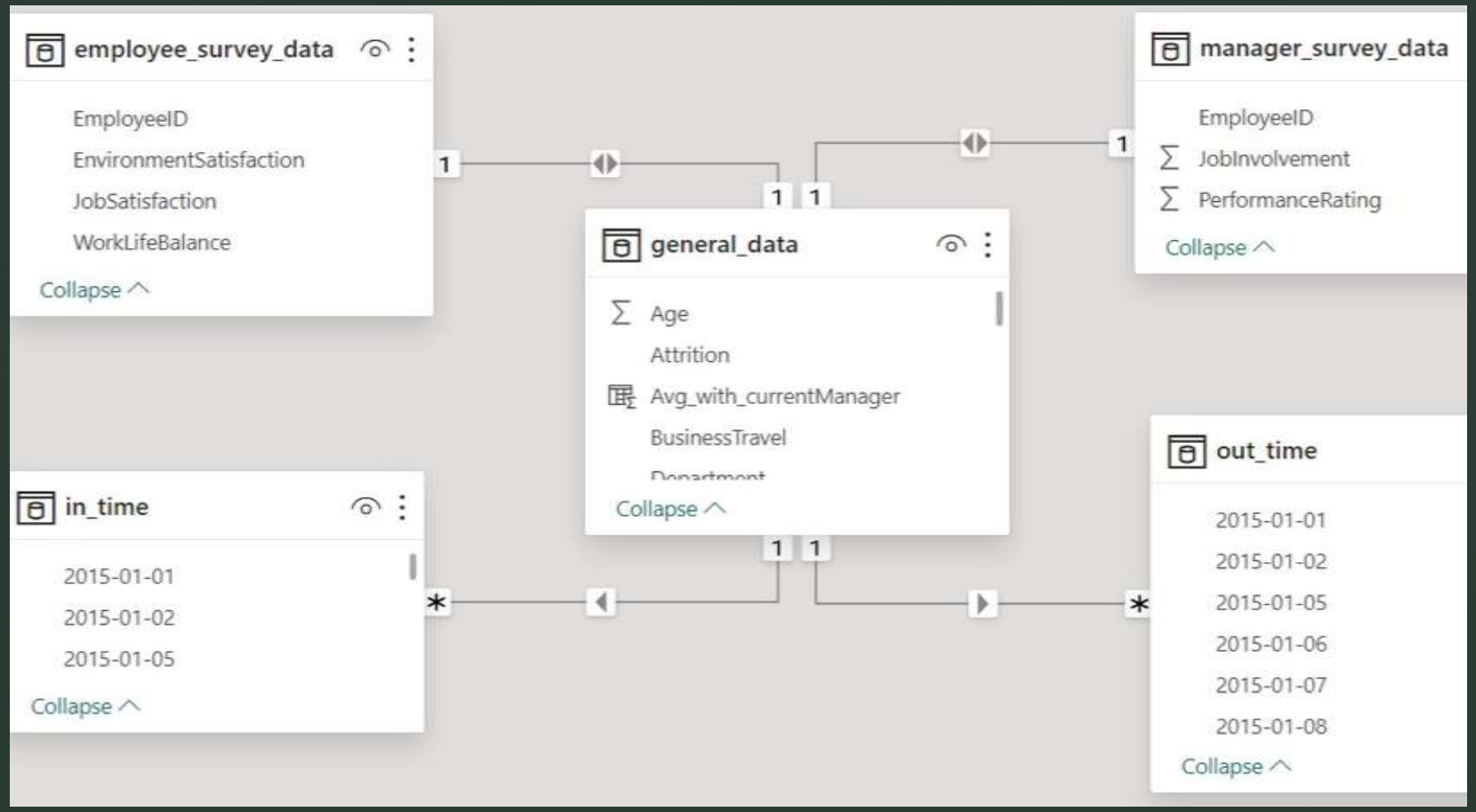
EnvironmentSatisfactio	JobSatisfactior	WorkLifeBalanc	MonthlyIncome
3	4	2	131160
3	2	4	41890
2	2	1	193280
4	4	3	83210
4	1	3	23420
3	2	2	40710
1	3	1	58130
1	2	3	31430
2	4	3	20440
2	1	3	134640
3	4	3	79910
NA	4	3	33770
4	1	3	55380
1	2	2	57620
4	4	2	25920
3	4	4	53460
4	3	4	42130

Avg. MonthlyIncome	65029.31293
Avg.JobSatisfaction	2.728246014

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



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



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

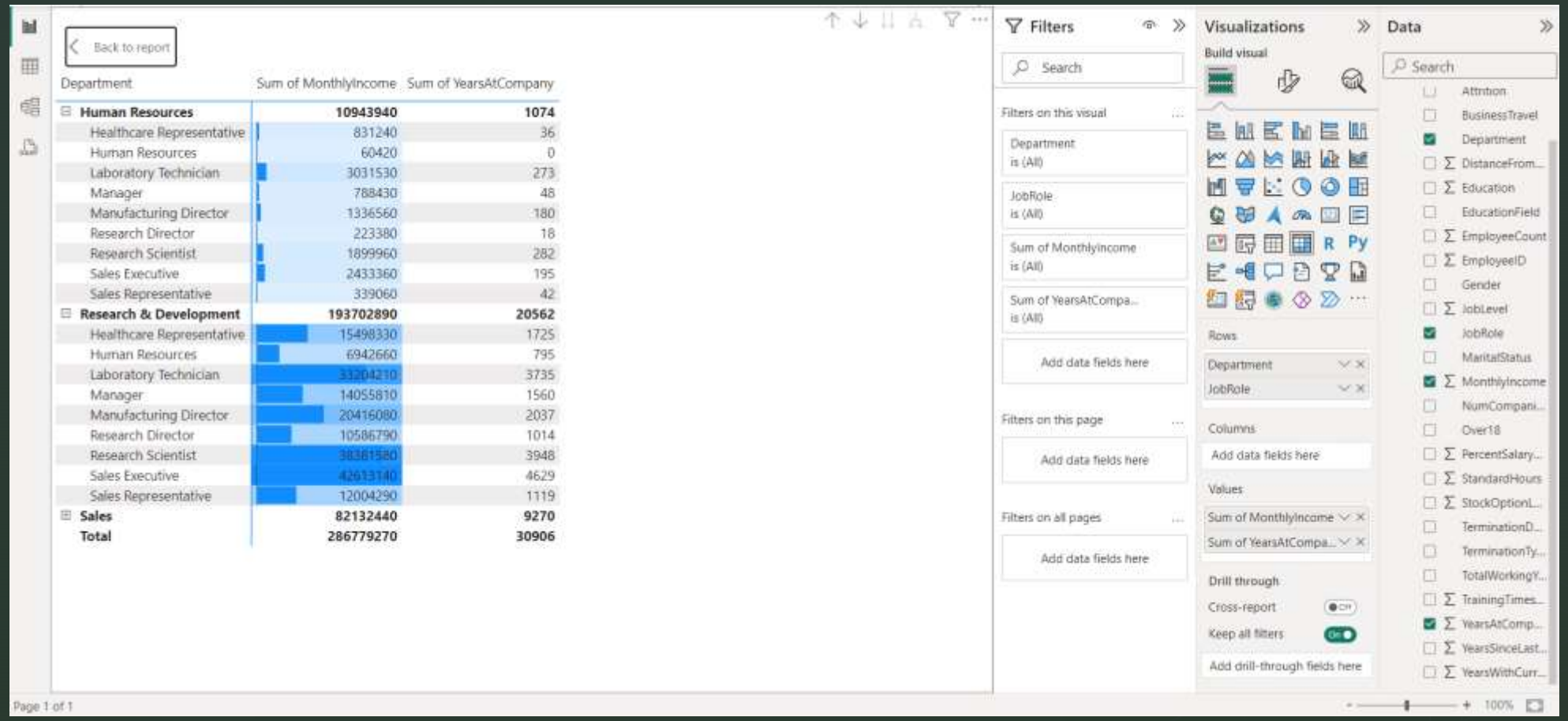
1. Identify the fact table: It contains the main metrics or measurements that you want to analyse. This table typically contains numerical values and is surrounded by dimension tables.
2. Identify the dimension tables: They contain descriptive attributes related to the data in the fact table. These attributes provide context to measurements in the fact table.
3. Create relationships between the fact Table and dimension tables: The fact table is connected to dimension tables through keys. Each dimension table will have a primary key that is referenced as a foreign key in the fact table.
4. Normalize the dimensions table: Dimension tables should be normalized to reduce redundancy and improve data consistency.

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

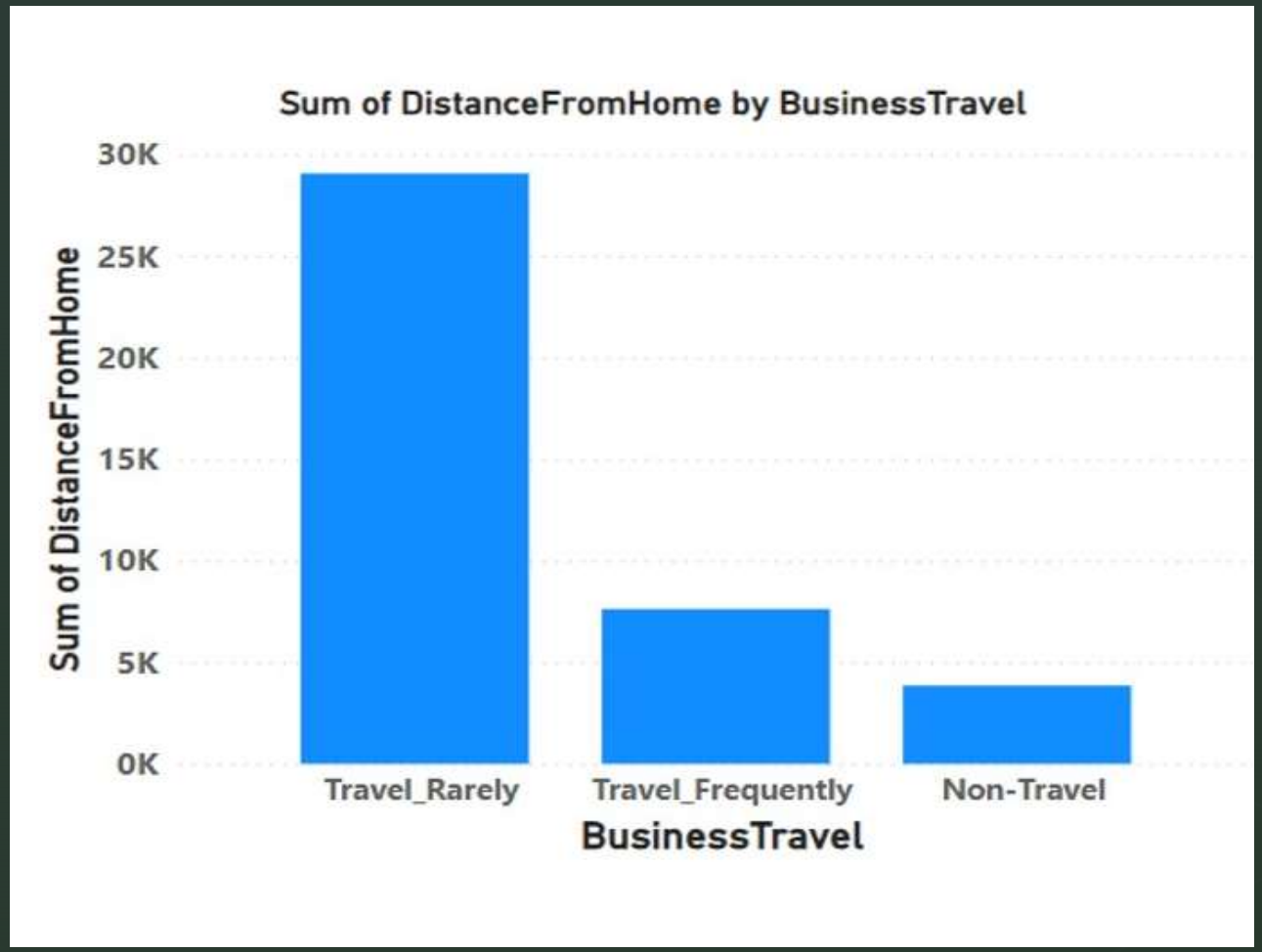
  - **Rolling #-month Average =**

```
VAR LastDate = LASTDATE(Employee'[Date])
RETURN
CALCULATE(
AVERAGEX(
VALUES('Employee'[Employee ID]),
CALCULATE(SUM('Employee'[Monthly Income]))
),
FILTER(
ALL('Employee),
'Employee'[Date]<= LastDate &&
'Employee'[Date] > DATEADD9LastDate, -3, MONTH0
)
)
```

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



- 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



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

- 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.

	Regular	10% Hike
Average Monthly Income	65029.31293	
Average Salary Hike	15.20952381	25.20952381
Monthly Income After Hike	74919.96176	81422.89305

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

1) Verify Data Adherence:

- Review data against predefined schema.
- Identify inconsistencies.

2) Actions for Inconsistencies:

- Communicate with stakeholders.
- Cleanse and transform data using Power Query Editor.
- Update data sources or processes.
- Adjust data model/schema as needed.

3) Documentation and Monitoring:

- Document changes made.
- Re-verify data integrity.
- Establish ongoing data quality monitoring.

Takeaway: Ensuring data consistency is critical for accurate analysis in Power BI. Implementing systematic processes for verification, correction, and monitoring maintains data integrity over time.