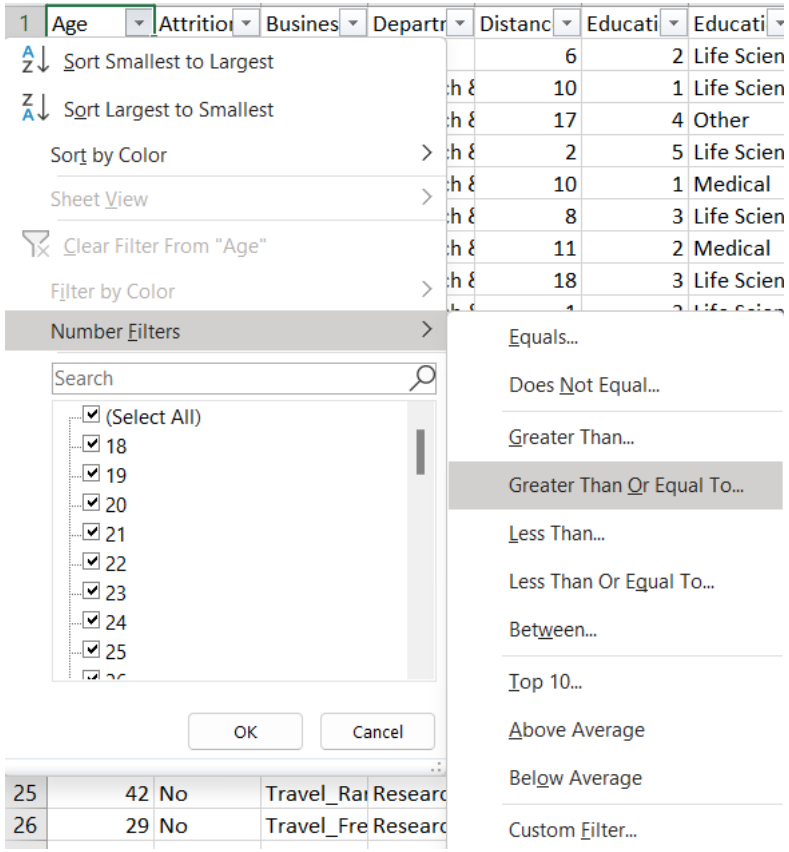
HR DATA ANALYSIS

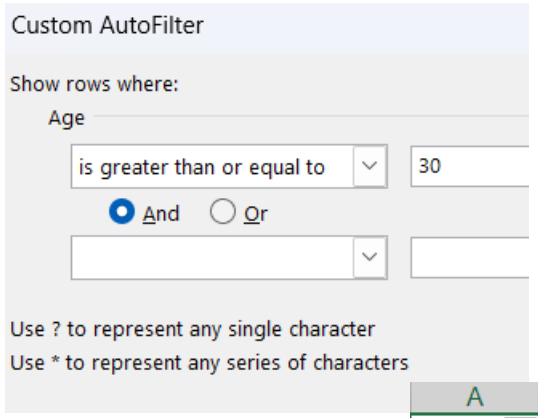
PSYLIQ DATA ANALYST INTERNSHIP

LILIANA MORONES ALBA

TASK 1

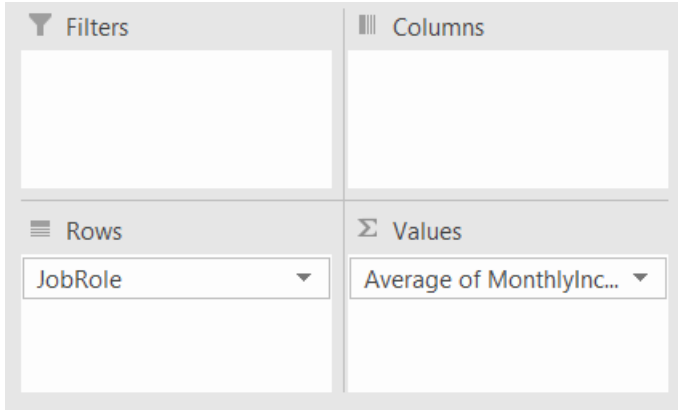
1. Using Excel, how would you filter the dataset to only show employees aged 30 and above?





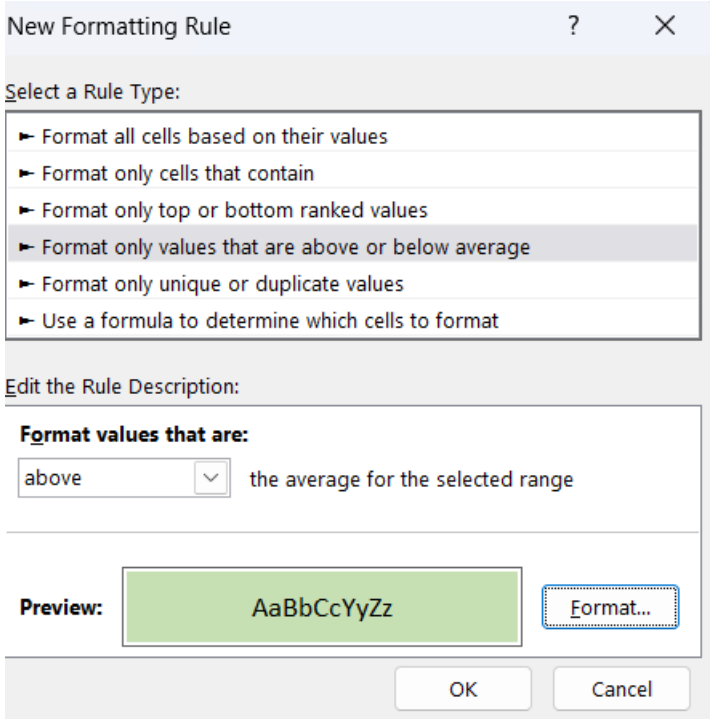
A	B	C	D	E	F	G
Age	Attritio	Busines	Departr	Distanc	Educati	Educa
51	No	Travel_Ra	Sales	6	2	Life S
31	Yes	Travel_Fre	Research &	10	1	Life S
32	No	Travel_Fre	Research &	17	4	Other
38	No	Non-Trave	Research &	2	5	Life S
32	No	Travel_Ra	Research &	10	1	Medic
46	No	Travel_Ra	Research &	8	3	Life S
31	No	Travel_Ra	Research &	1	3	Life S
45	No	Travel_Ra	Research &	17	2	Medic
36	No	Travel_Ra	Research &	28	1	Life S
55	No	Travel_Ra	Research &	14	4	Life S

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



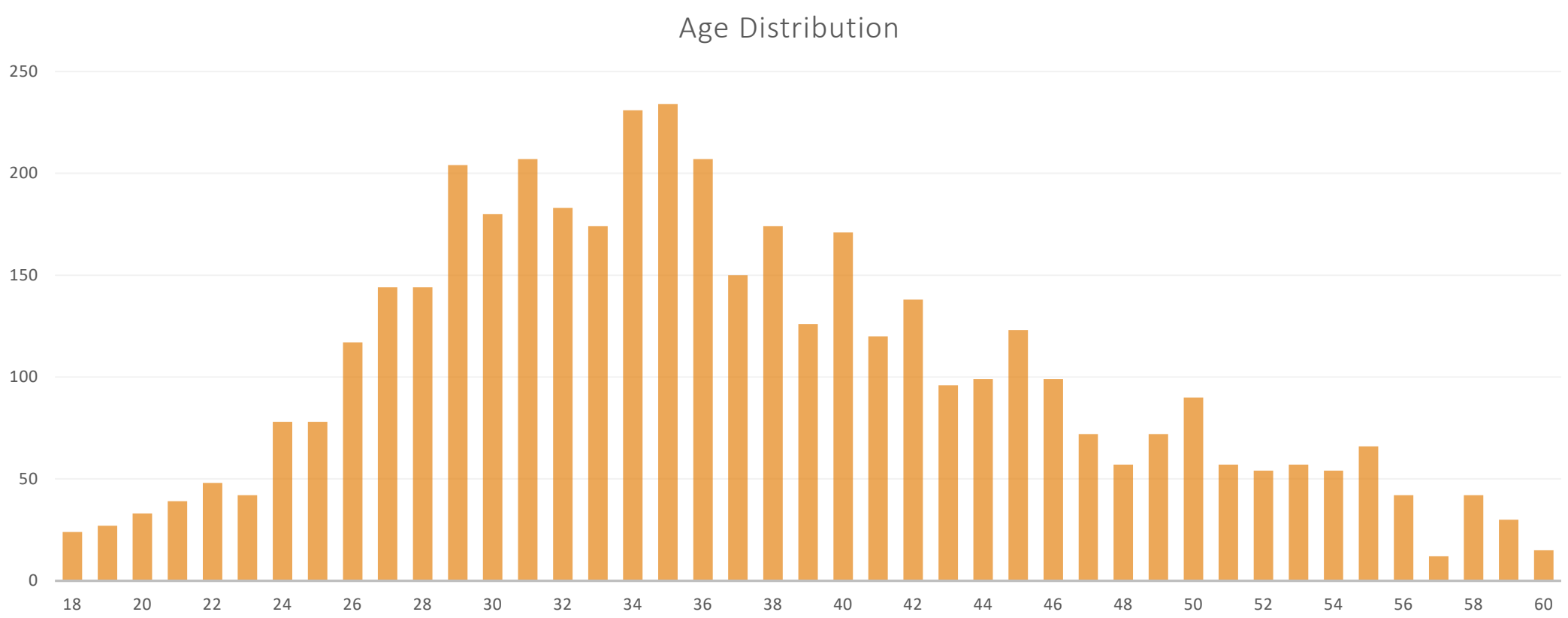
Job Role	Average of MonthlyIncome
Healthcare Representative	$ 60,983.74
Human Resources	$ 58,528.08
Laboratory Technician	$ 66,314.05
Manager	$ 63,395.88
Manufacturing Director	$ 69,183.72
Research Director	$ 65,473.13
Research Scientist	$ 64,975.68
Sales Executive	$ 65,186.69
Sales Representative	$ 65,370.96
Grand Total	**$ 65,029.31**

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



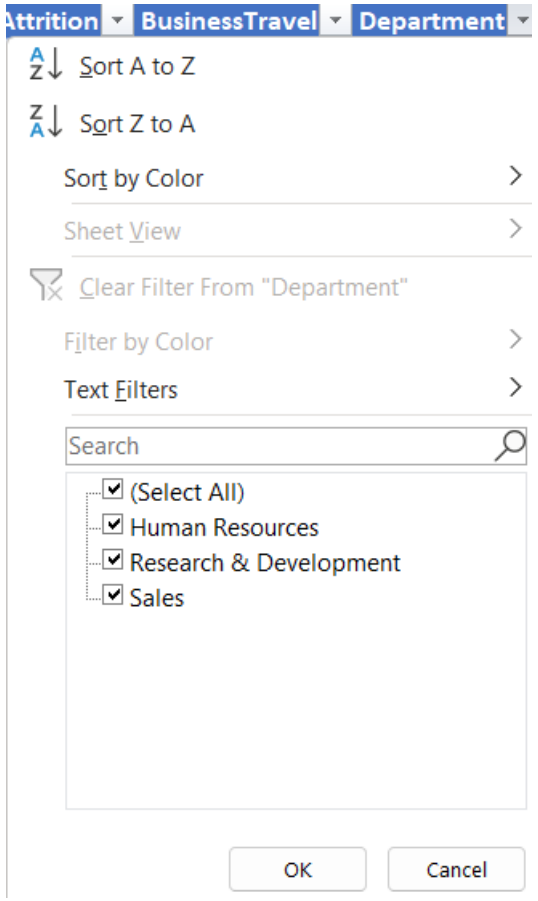
Employ	Employ	Gender	JobLeve	JobRole	Marital	MonthlyIncome
1	1	Female	1	Healthcare	Married	$ 131,160
1	2	Female	1	Research S	Single	$ 41,890
1	3	Male	4	Sales Exec	Married	$ 193,280
1	4	Male	3	Human Re	Married	$ 83,210
1	5	Male	1	Sales Exec	Single	$ 23,420
1	6	Female	4	Research I	Married	$ 40,710
1	7	Male	2	Sales Exec	Single	$ 58,130
1	8	Male	2	Sales Exec	Married	$ 31,430
1	9	Male	3	Laboratory	Married	$ 20,440
1	10	Female	4	Laboratory	Divorced	$ 134,640

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

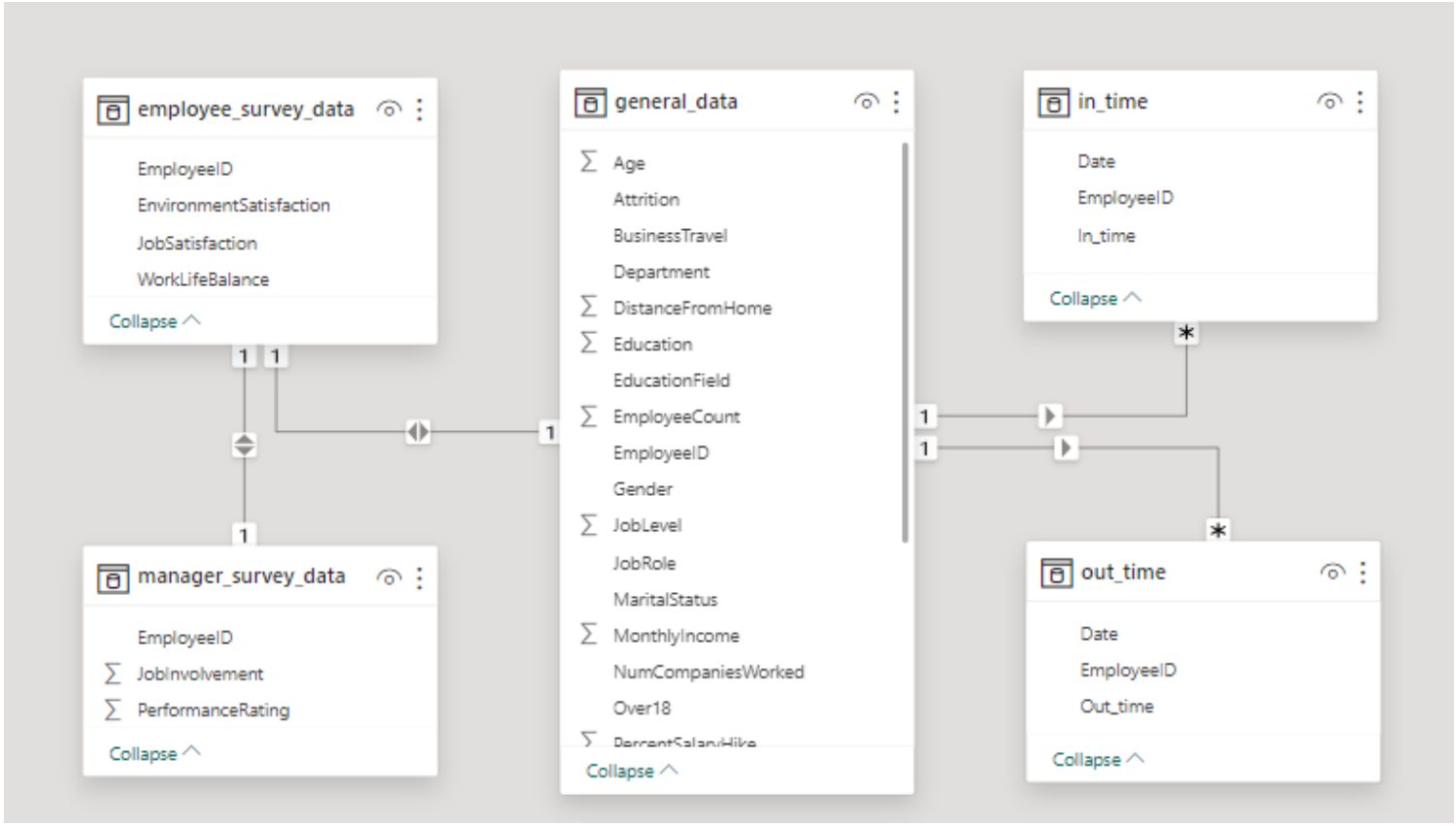


5. Identify and clean any missing or inconsistent data in the "Department" column.

There is no missing or inconsistent data.



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.

```
AVG_YearswManager = AVERAGE(general_data[YearsWithCurrManager])
```

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

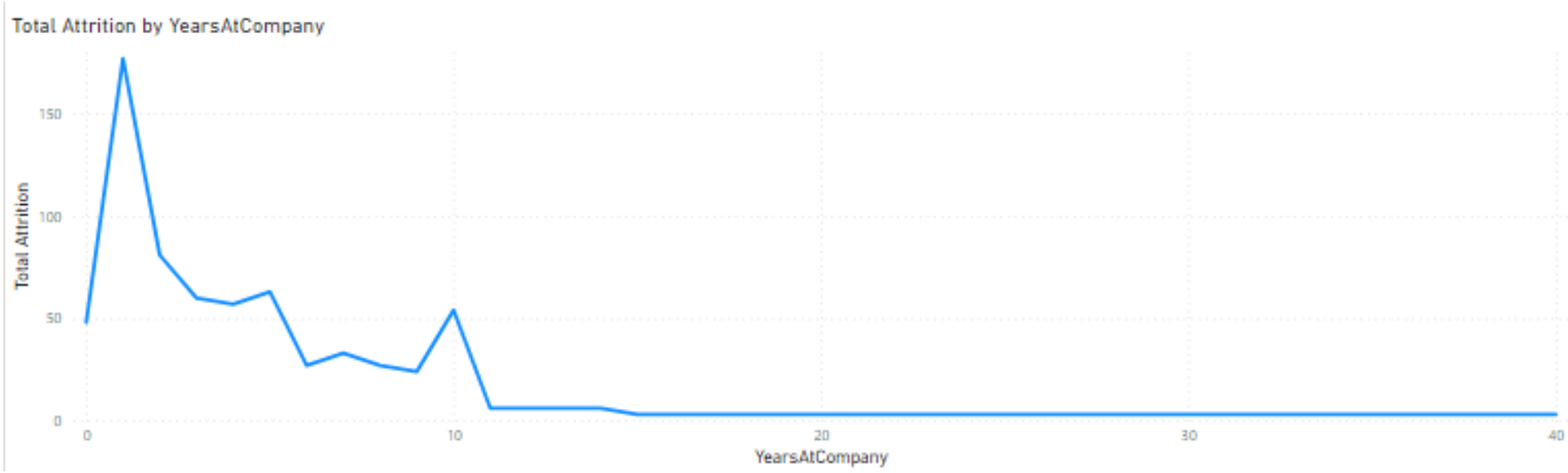
Row Labels	Employees
Divorced	**981**
Human Resources	21
Research & Development	621
Sales	339
Married	**2019**
Human Resources	96
Research & Development	1350
Sales	573
Single	**1410**
Human Resources	72
Research & Development	912
Sales	426
Grand Total	**4410**

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

H	I	J	K	L	M	N	O	P
EmployeeC	EmployeeI	Gender	JobLevel	JobRole	MaritalSta	MonthlyIncome	Job Satisfaction	NumComp
1	1	Female	1	Healthcare	Married	$131,160	0	1
1	2	Female	1	Research S	Single	$41,890	2	0
1	3	Male	4	Sales Exec	Married	$193,280	2	1
1	4	Male	3	Human Re	Married	$83,210	4	3
1	5	Male	1	Sales Exec	Single	$23,420	1	4
1	6	Female	4	Research D	Married	$40,710	2	3
1	7	Male	2	Sales Exec	Single	$58,130	3	2
1	8	Male	2	Sales Exec	Married	$31,430	2	2
1	9	Male	3	Laboratory	Married	$20,440	4	0
1	10	Female	4	Laboratory	Divorced	$134,640	1	1
1	11	Male	2	Laboratory	Married	$79,910	4	0
1	12	Male	1	Laboratory	Married	$33,770	4	0
1	13	Female	1	Sales Exec	Single	$55,380	1	0
1	14	Male	1	Research S	Married	$57,620	2	1
1	15	Male	1	Manufactu	Married	$25,920	4	1
1	16	Male	2	Healthcare	Married	$53,460	4	4
1	17	Male	1	Laboratory	Single	$42,130	3	1
1	18	Male	2	Sales Exec	Divorced	$41,270	4	2
1	19	Male	1	Sales Repr	Divorced	$24,380	2	7

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

```
Total Attrition = COUNTROWS(FILTER('general_data', 'general_data'[Attrition] = "Yes"))
```



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

Fact Table

- Metrics or measures of interest
- Primary Key: EmployeeID

Dimension Tables

- Descriptive information
- Employee, Manager and Time Tables

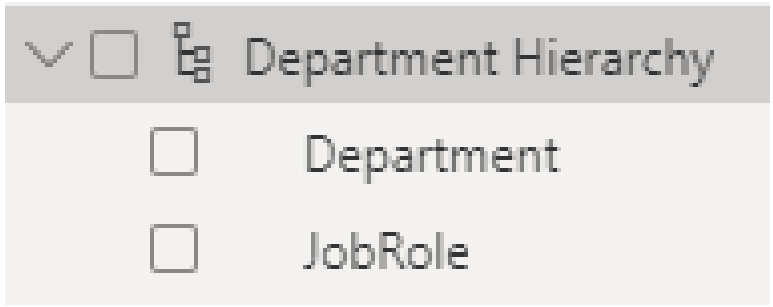
- Simple and easy-to-understand structure.
- Better performance for analytical queries.
- Fewer joins to access data.

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

```
3MonthIncome = 3*general_data[MonthlyIncome]
```

```
AVG 3MonthIncome = AVERAGE(general_data[3MonthIncome])
```

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



Department	Total Attrition
Human Resources	**57**
Healthcare Representative	3
Laboratory Technician	18
Manager	3
Manufacturing Director	6
Research Scientist	3
Sales Executive	24
Research & Development	**453**
Sales	**201**
Total	**711**

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.

```
=SUMIFS(Gral[ [ MonthlyIncome ] ], Gral[Department], [@Department], Gral[JobLevel], ">=" & 3)
```

Department	Monthly Income
Human Resources	$ 3,259,140
Research & Development	$ 53,502,900
Sales	$ 22,974,330

15. 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.

Calculate the average monthly income.

Multiply the average monthly income by 1 + the 10% of salary hike.

% Salary hike	Avg Monthly Inco	Income after salary hike
10%	$ 65,029	$ 71,532

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

To verify if the data adheres to a predefined schema we need to first understand the ‘business rules’ to make sure that our relations, tables, columns and data types are the correct ones.

In the case I find inconsistencies, I would correct some of them manually, and others using a software process to clean my data, communicate my findings with the interested parts and actualize the scheme if necessary.