



HR Data Analysis Assessment

By Arfa Afsheen

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

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	F
1	Age	Attrition	BusinessTravel	Department	DistanceFromHome	Education	EducationField	EmployeeCount	Employee	Gender	JobLevel	JobRole	MaritalStatus	MonthlyIncome	NumCompaniesWorked	OverseasAssignments
2	51	No	Travel_Rarely	Sales	6	2	Life Sciences	1	1	Female	1	Healthcare Representative	Married	131160		1 Y
3	31	Yes	Travel_Frequently	Research & Development	10	1	Life Sciences	1	2	Female	1	Research Scientist	Single	41890		0 Y
4	32	No	Travel_Frequently	Research & Development	17	4	Other	1	3	Male	4	Sales Executive	Married	193280		1 Y
5	38	No	Non-Travel	Research & Development	2	5	Life Sciences	1	4	Male	3	Human Resources	Married	83210		3 Y
6	32	No	Travel_Rarely	Research & Development	10	1	Medical	1	5	Male	1	Sales Executive	Single	23420		4 Y
7	46	No	Travel_Rarely	Research & Development	8	3	Life Sciences	1	6	Female	4	Research Director	Married	40710		3 Y
10	31	No	Travel_Rarely	Research & Development	1	3	Life Sciences	1	9	Male	3	Laboratory Technician	Married	20440		0 Y
12	45	No	Travel_Rarely	Research & Development	17	2	Medical	1	11	Male	2	Laboratory Technician	Married	79910		0 Y
13	36	No	Travel_Rarely	Research & Development	28	1	Life Sciences	1	12	Male	1	Laboratory Technician	Married	33770		0 Y
14	55	No	Travel_Rarely	Research & Development	14	4	Life Sciences	1	13	Female	1	Sales Executive	Single	55380		0 Y
15	47	Yes	Non-Travel	Research & Development	1	1	Medical	1	14	Male	1	Research Scientist	Married	57620		1 Y
17	37	No	Travel_Rarely	Research & Development	1	3	Life Sciences	1	16	Male	2	Healthcare Representative	Married	53460		4 Y
19	37	No	Non-Travel	Research & Development	1	3	Medical	1	18	Male	2	Sales Executive	Divorced	41270		2 Y
20	35	No	Travel_Rarely	Sales	7	4	Life Sciences	1	19	Male	1	Sales Representative	Divorced	24380		7 Y
21	38	No	Travel_Rarely	Research & Development	8	3	Life Sciences	1	20	Female	1	Manager	Divorced	68700		1 Y
23	50	No	Travel_Rarely	Sales	8	4	Life Sciences	1	22	Male	1	Research Scientist	Divorced	96670		3 Y
24	53	No	Travel_Rarely	Research & Development	11	4	Life Sciences	1	23	Female	2	Research Scientist	Married	21480		3 Y
25	42	No	Travel_Rarely	Research & Development	4	4	Life Sciences	1	24	Male	1	Manufacturing Director	Married	89260		1 Y
27	55	No	Travel_Rarely	Research & Development	1	4	Other	1	26	Female	1	Research Scientist	Married	67990		3 Y
29	37	No	Travel_Rarely	Sales	5	1	Marketing	1	28	Male	1	Research Scientist	Single	27050		1 Y
30	44	Yes	Travel_Frequently	Research & Development	1	2	Medical	1	29	Male	2	Research Scientist	Divorced	103330		3 Y
31	38	No	Travel_Rarely	Sales	2	3	Marketing	1	30	Female	1	Manager	Divorced	44480		9 Y
34	49	No	Travel_Frequently	Research & Development	1	1	Medical	1	33	Female	2	Research Scientist	Single	35910		9 Y
35	36	No	Travel_Rarely	Sales	5	3	Technical Degree	1	34	Male	3	Sales Executive	Single	54050		4 Y
36	31	No	Travel_Frequently	Research & Development	8	4	Medical	1	35	Male	1	Sales Executive	Divorced	46940		1 Y

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

The screenshot displays an Excel spreadsheet with a pivot table and the PivotTable Fields task pane. The pivot table summarizes the average monthly income by job role. The task pane shows the source data fields and the current pivot table configuration.

Row Labels	Average of MonthlyIncome
Healthcare Representative	60983.74046
Human Resources	58528.07692
Laboratory Technician	66314.05405
Manager	63395.88235
Manufacturing Director	69183.72414
Research Director	65473.125
Research Scientist	64975.68493
Sales Executive	65186.68712
Sales Representative	65370.96386
Grand Total	65029.31293

PivotTable Fields

Choose fields to add to report:

Search

- ☐ MaritalStatus
- ☒ **MonthlyIncome**
- ☐ NumCompaniesWorked
- ☐ Over18
- ☐ PercentSalaryHike
- ☐ StandardHours
- ☐ StockOptionLevel

Drag fields between areas below:

Filters	Columns

Rows	Values
JobRole	Average of MonthlyInco...

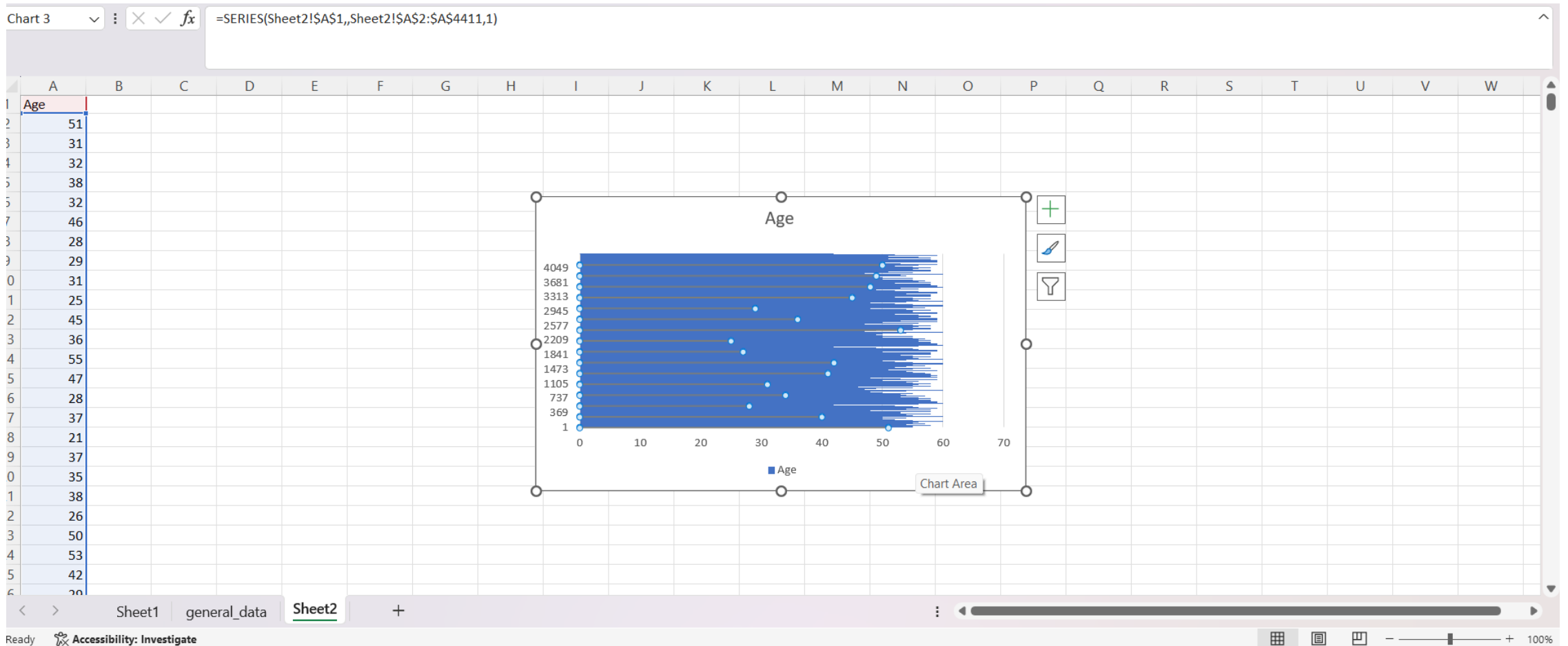
☐ Defer Layout Update Update

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

	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
1	EducationField	EmployeeCount	EmployeeID	Gender	JobLevel	JobRole	MaritalStatus	MonthlyIncome	NumCompaniesWorked	Over18	PercentSalaryHike	StandardHours	StockOptions	TotalWorking	TrainingTimesLastYear	
2	Life Sciences	1	1	Female	1	Healthcare Representative	Married	131160	1	Y	11	8	0	1	6	
3	Life Sciences	1	2	Female	1	Research Scientist	Single	41890	0	Y	23	8	1	6	3	
4	Other	1	3	Male	4	Sales Executive	Married	193280	1	Y	15	8	3	5	2	
5	Life Sciences	1	4	Male	3	Human Resources	Married	83210	3	Y	11	8	3	13	5	
6	Medical	1	5	Male	1	Sales Executive	Single	23420	4	Y	12	8	2	9	2	
7	Life Sciences	1	6	Female	4	Research Director	Married	40710	3	Y	13	8	0	28	5	
8	Medical	1	7	Male	2	Sales Executive	Single	58130	2	Y	20	8	1	5	2	
9	Life Sciences	1	8	Male	2	Sales Executive	Married	31430	2	Y	22	8	3	10	2	
10	Life Sciences	1	9	Male	3	Laboratory Technician	Married	20440	0	Y	21	8	0	10	2	
11	Medical	1	10	Female	4	Laboratory Technician	Divorced	134640	1	Y	13	8	1	6	2	
12	Medical	1	11	Male	2	Laboratory Technician	Married	79910	0	Y	13	8	2	21	2	
13	Life Sciences	1	12	Male	1	Laboratory Technician	Married	33770	0	Y	12	8	2	16	2	
14	Life Sciences	1	13	Female	1	Sales Executive	Single	55380	0	Y	17	8	0	37	2	
15	Medical	1	14	Male	1	Research Scientist	Married	57620	1	Y	11	8	2	10	4	
16	Life Sciences	1	15	Male	1	Manufacturing Director	Married	25920	1	Y	14	8	0	5	2	
17	Life Sciences	1	16	Male	2	Healthcare Representative	Married	53460	4	Y	11	8	0	7	2	
18	Life Sciences	1	17	Male	1	Laboratory Technician	Single	42130	1	Y	12	8	3	3	3	
19	Medical	1	18	Male	2	Sales Executive	Divorced	41270	2	Y	13	8	1	15	2	
20	Life Sciences	1	19	Male	1	Sales Representative	Divorced	24380	7	Y	16	8	0	10	5	
21	Life Sciences	1	20	Female	1	Manager	Divorced	68700	1	Y	11	8	1	8	5	
22	Other	1	21	Male	2	Laboratory Technician	Divorced	104470	1	Y	18	8	0	6	3	
23	Life Sciences	1	22	Male	1	Research Scientist	Divorced	96670	3	Y	23	8	0	28	2	
24	Life Sciences	1	23	Female	2	Research Scientist	Married	21480	3	Y	11	8	0	21	2	
25	Life Sciences	1	24	Male	1	Manufacturing Director	Married	89260	1	Y	14	8	0	NA	4	
26	Medical	1	25	Male	1	Laboratory Technician	Single	65130	1	Y	11	8	1	10	2	

Sheet1 general_data

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



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

There are neither missing nor inconsistent values in the 'Department' column to be cleaned.

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

The screenshot displays the Power BI Desktop interface. The main workspace shows a relationship diagram with two tables: 'general_data' and 'employee_survey_data'. A line connects the 'EmployeeID' field in 'general_data' to the 'EmployeeID' field in 'employee_survey_data', indicating a one-to-one relationship. The 'general_data' table has fields: Department, DistanceFromHome, Education, EducationField, EmployeeCount, EmployeeID, Gender, JobLevel, and Inheritance. The 'employee_survey_data' table has fields: EmployeeID, EnvironmentSatisfaction, JobSatisfaction, and WorkLifeBalance. The 'Properties' pane on the right shows the 'general_data' table selected, with fields like Name, Description, Synonyms, Row label, Key column, Is hidden, and Is featured table. The 'Data' pane on the right shows a list of tables: 'employee_survey_data' and 'general_data'. The bottom status bar shows 'All tables' and a '+' button.

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

HR-Analytics • Last saved: Today at 5:39 PM

Search

Sign in

File Home Help Table tools Column tools

Name: AverageYearWithC... Format: General Summarization: Sum Data type: Decimal number Data category: Uncategorized

Structure Formatting Properties Sort Groups Relationships Calculations

	StandardHours	StockOptionLevel	TotalWorkingYears	TrainingTimesLastYear	YearsAtCompany	YearsSinceLastPromotion	YearsWithCurrManager	AverageYearWithCurrentManager
11	8	0	1	6	1	0	0	4.12312925170068
21	8	1	1	3	1	0	0	4.12312925170068
12	8	1	1	2	1	0	0	4.12312925170068
13	8	0	1	3	1	0	0	4.12312925170068
15	8	2	1	2	1	0	0	4.12312925170068
22	8	0	1	3	1	0	0	4.12312925170068
12	8	0	1	4	1	0	0	4.12312925170068
13	8	2	1	3	1	0	0	4.12312925170068
13	8	3	1	2	1	0	0	4.12312925170068
13	8	2	1	2	1	0	0	4.12312925170068
12	8	2	1	2	1	0	0	4.12312925170068
13	8	1	1	2	1	0	0	4.12312925170068
14	8	1	1	5	1	0	0	4.12312925170068
15	8	1	1	3	1	0	0	4.12312925170068
15	8	1	1	3	1	0	0	4.12312925170068
11	8	0	1	2	1	0	0	4.12312925170068
17	8	1	1	3	1	0	0	4.12312925170068
13	8	1	1	5	1	0	0	4.12312925170068
15	8	0	1	1	1	0	0	4.12312925170068
15	8	0	1	3	1	0	0	4.12312925170068
13	8	0	1	3	1	0	0	4.12312925170068
12	8	0	1	2	1	0	0	4.12312925170068
19	8	1	1	5	1	0	0	4.12312925170068
14	8	1	1	2	1	0	0	4.12312925170068
14	8	1	1	4	1	0	0	4.12312925170068
16	8	0	1	3	1	0	0	4.12312925170068
13	8	0	1	2	1	0	0	4.12312925170068

Table: general_data (4,410 rows) Column: AverageYearWithCurrentManager (1 distinct values)

Data

Search

- employee_survey_data
- general_data
 - Age
 - Attrition
 - AverageYearWithCurrentManager
 - BusinessTravel
 - Department
 - DistanceFromHome
 - Education
 - EducationField
 - EmployeeCount
 - EmployeeID
 - Gender
 - JobLevel
 - JobRole
 - MaritalStatus
 - MonthlyIncome
 - NumCompaniesWorked
 - Over18
 - PercentSalaryHike
 - StandardHours
 - StockOptionLevel
 - TotalWorkingYears

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

The screenshot shows an Excel spreadsheet with a pivot table and the PivotTable Fields task pane. The pivot table is located in the range A1:O10 and displays the count of employees by marital status, segmented by department. The PivotTable Fields task pane is open on the right side of the screen, showing the fields available for the report.

PivotTable Fields

Choose fields to add to report:

Search

☐ BusinessTravel
☒ **Department**
☐ DistanceFromHome
☐ Education
☐ EducationField
☐ EmployeeCount
☒ **EmployeeID**

Drag fields between areas below:

Filters

Columns
Department

Rows
MaritalStatus

Values
Count of EmployeeID

☐ Defer Layout Update Update

Count of EmployeeID	Column Labels	Human Resources	Research & Development	Sales	(blank)	Grand Total
Divorced		21	621	339		981
Married		96	1350	573		2019
Single		72	912	426		1410
(blank)						
Grand Total		189	2883	1338		4410

Apply conditional formatting to highlight employees with above average Job Satisfaction

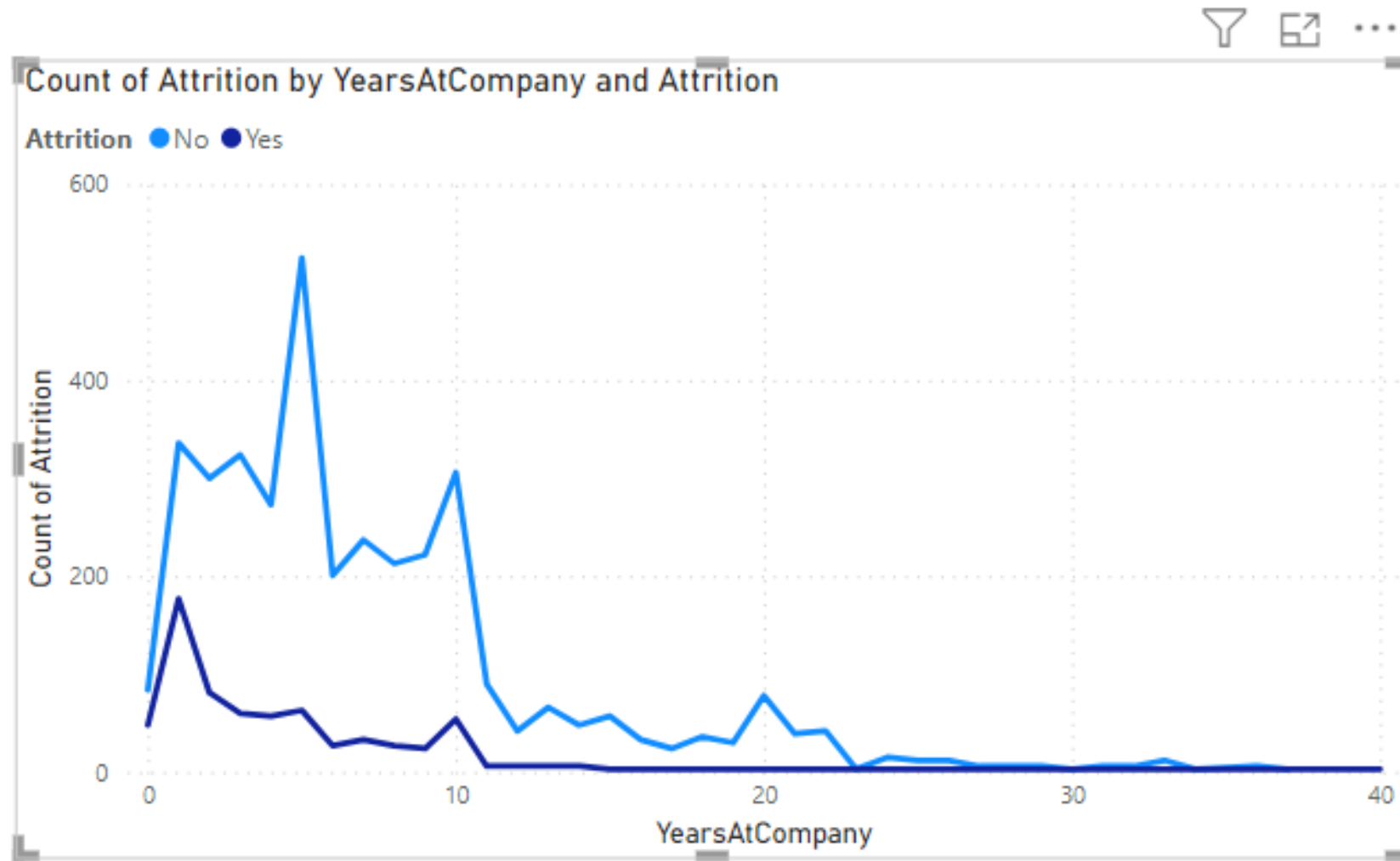
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	EmployeeID	EnvironmentSatisfaction	JobSatisfaction	WorkLifeBalance														
2	1	3	4	2														
3	2	3	2	4														
4	3	2	2	1														
5	4	4	4	3														
6	5	4	1	3														
7	6	3	2	2														
8	7	1	3	1														
9	8	1	2	3														
10	9	2	4	3														
11	10	2	1	3														
12	11	3	4	3														
13	12	NA	4	3														
14	13	4	1	3														
15	14	1	2	2														
16	15	4	4	2														
17	16	3	4	4														
18	17	4	3	4														
19	18	1	4	3														
20	19	2	2	2														
21	20	1	1	3														
22	21	3	2	1														
23	22	1	2	2														
24	23	3	3	2														
25	24	2	3	3														
26	25	2	4	2														
										Average of JobSatisfaction								
										2.728246014								

employee_survey_data

Ready Accessibility: Good to go

100%

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.

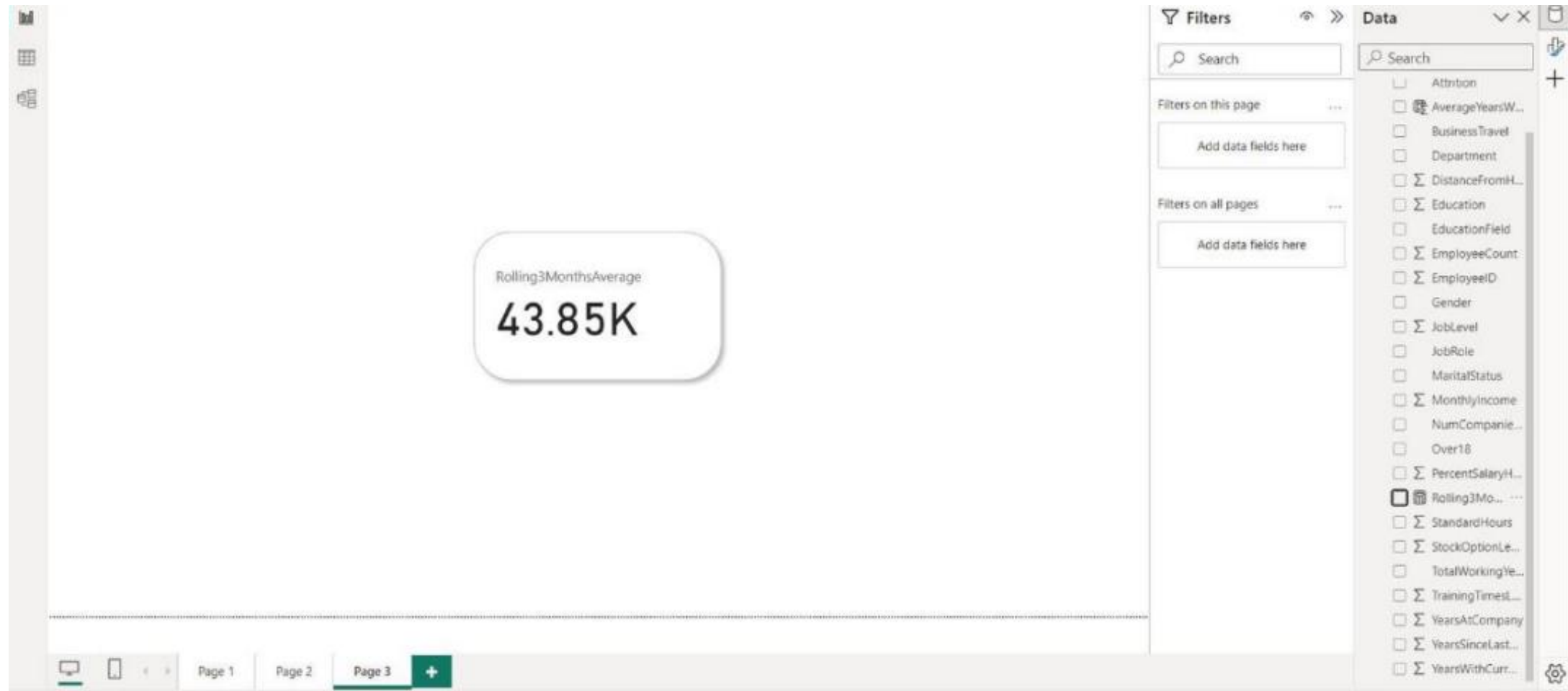
To create a star schema out of this data set we would need 'fact' and 'Dimension' table separated with appropriate columns and keys; the general step should be

- 1) Normalize the table and divide it into more tables i.e we may create 'Employee Attrition fact table' , 'Satisfaction fact table' , 'job performance fact table' which would be the center of star schema.
- 2) Furthermore we need dimension table as well consisting of 'Employee Dimension table' , 'satisfaction dimension table' , 'job performance dimension table' , 'in-out dimension table'
- 3) All these dimension tables would contain a primary key 'EmployeeID' which is connected to the foreign key 'EmpID' of the fact table
- 4) This ensures that the data is correctly related and power BI can perform accurate analysis and reporting

The Benefits of star schema would be:

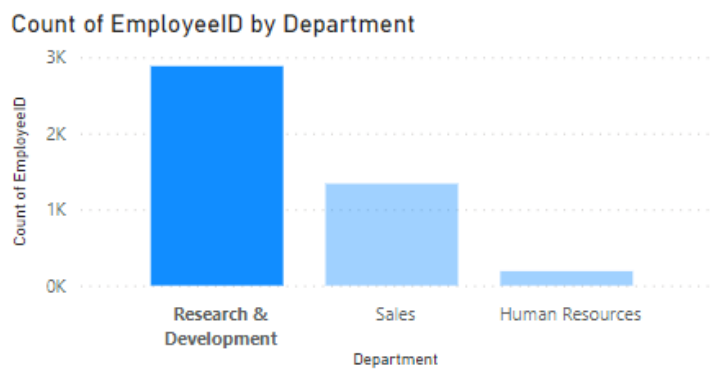
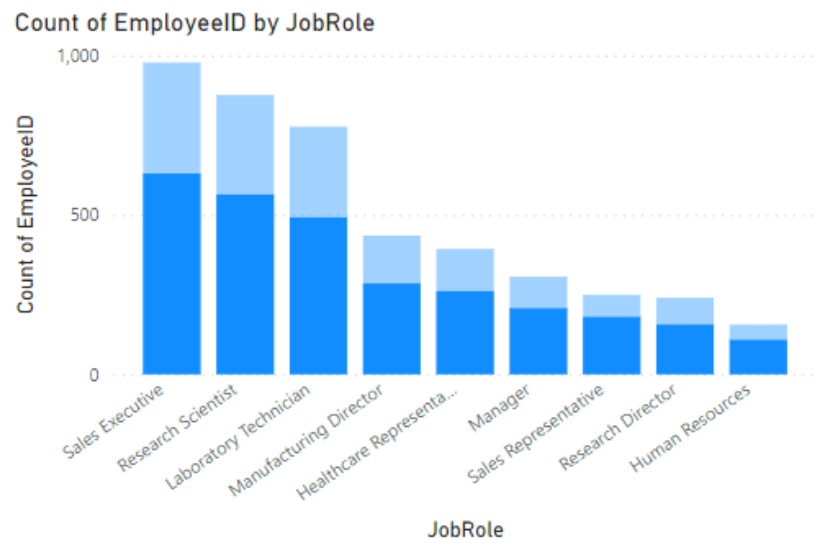
- A) Simplicity and understandability for technical as well as non-technical user
- B) The separation of dimension and facts allows for efficient querying
- C) The star schema is scalable and can handle large datasets and they are flexible and can adapt to changing reporting requirements
- D) Data redundancy is minimized in star schema

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



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

Department	JobRole
Research & Development	Healthcare Representative
Research & Development	Human Resources
Research & Development	Laboratory Technician
Research & Development	Manager
Research & Development	Manufacturing Director
Research & Development	Research Director
Research & Development	Research Scientist
Research & Development	Sales Executive
Research & Development	Sales Representative



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

1)Parameterized queries can be set up using Power Query, to allow users to filter data based on the Distance from Home column using parameters.

2)Load your data, create a new parameter in the Power Query Editor for Distance Parameter, and set its datatype to decimal/whole number.

3)In the Power Query Editor, locate the query that loads your data and add a filter step to filter the data based on the parameter.

```
= Table.SelectRows(YourPreviousStep, each [DistanceFromHome] <= DistanceParameter)
```

4)Click "Close & Apply" in the Home tab to apply the changes. 5)In your Power BI report, create a slicer visual or any other method for users to input the parameter value and manually update the filter condition in your visual to reference the parameter.

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

The screenshot displays an Excel spreadsheet with a PivotTable and the PivotTable Fields task pane. The PivotTable is located in the range B3:E5 and shows the sum of monthly income for each department, filtered by job level greater than or equal to 3. The PivotTable Fields task pane is open on the right side of the screen, showing the available fields and their placement in the PivotTable layout.

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

PivotTable Fields

Choose fields to add to report:

Search

- ☐ Education
- ☐ EducationField
- ☐ EmployeeCount
- ☐ EmployeeID
- ☐ Gender
- ☒ **JobLevel**
- ☐ JobRole

Drag fields between areas below:

Filters

Columns

Department

Rows

JobLevel

Values

Sum of MonthlyIncome

☐ Defer Layout Update Update

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

- 1) In a cell (let's say A1), enter the initial Percent Salary Hike value (e.g., 5%), or else, highlight the cells where the percent hike is stored.
- 2) another cell, calculate the Monthly Income based on the formula that includes the Percent Salary Hike.
- 3) Now that you have both the columns 'Monthly Income' and 'Percent Salary Hike', drag down the cell with the initial Percent Salary Hike to, say, A10. Excel will fill in the remaining values.
- 4) We can see how Monthly Income changes with different Percent Salary Hike values.
- 5) Likewise, you can also go to Data Tab in Excel and click on 'What-If analysis' and select 'Data Table'. Choose the cell with the formula for Monthly Income as the "Column input cell" and play with different Percent Salary Hike values, and Excel will show you the corresponding Monthly Income.

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

1) Verifying if data adheres to a predefined schema involves checking whether the actual data in a dataset aligns with the expected structure and rules outlined in the predefined schema.

2) As per current status of the data, there is a need to reorder EmployeeID column, changing data type of 'TotalWorkingHours' column, filling NA values and blank values from general_data, employee_survey_data, manager_survey_data.

3) There is 'EmployeeID' label missing in the 'in-time' and 'out-time' data.

4) It is very important to address these inconsistencies in the data files and check if data profiling is done carefully and the data is validated correctly.

5) The dataset would then conform to the predefined schema following the resolution of all identified inconsistencies and the implementation of necessary data quality measures.

DASHBOARD

