PSYLIQ

HR DATA ANALYSIS

PSYLIQ Internship Project

Created by: Divya Pardeshi

INTRODUCTION

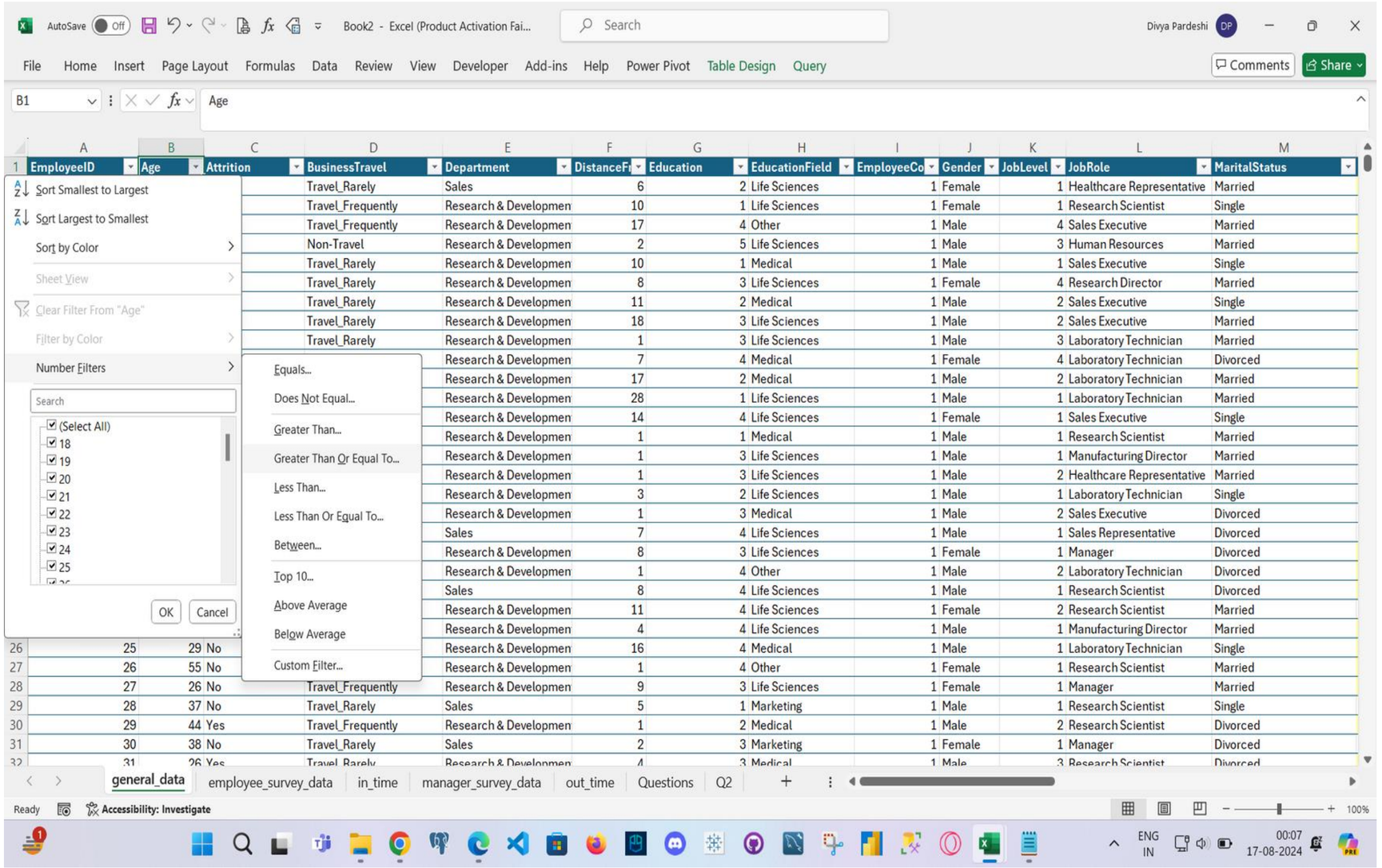
In this presentation, I've explored key HR questions within a dataset of over 4,400 employees, addressing a range of topics. I will demonstrate my approach to solving the HR Data Analysis Assessment tasks, which include:

- Filtering datasets to display specific employee groups
- Creating pivot tables to summarize data
- Applying conditional formatting to highlight key metrics
- Visualizing trends through charts and graphs

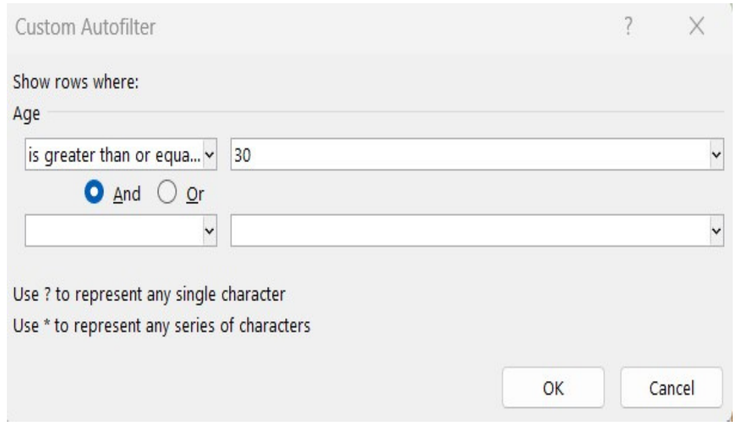
Additionally, I will cover establishing relationships between datasets in Power BI. Each task is designed to showcase practical skills in both Excel and Power BI, with the aim of deriving actionable insights from HR data.

PROJECT QUESTIONS
& Solutions

1. USING EXCEL, HOW WOULD YOU FILTER THE DATASET TO ONLY SHOW EMPLOYEES AGED 30 AND ABOVE?



EmployeeID	Age	Attrition	BusinessTravel	Department	DistanceF	Education	EducationField	EmployeeCo	Gender	JobLevel	JobRole	MaritalStatus
			Travel_Rarely	Sales	6	2	Life Sciences	1	Female	1	Healthcare Representative	Married
			Travel_Frequently	Research & Developmen	10	1	Life Sciences	1	Female	1	Research Scientist	Single
			Travel_Frequently	Research & Developmen	17	4	Other	1	Male	4	Sales Executive	Married
			Non-Travel	Research & Developmen	2	5	Life Sciences	1	Male	3	Human Resources	Married
			Travel_Rarely	Research & Developmen	10	1	Medical	1	Male	1	Sales Executive	Single
			Travel_Rarely	Research & Developmen	8	3	Life Sciences	1	Female	4	Research Director	Married
			Travel_Rarely	Research & Developmen	11	2	Medical	1	Male	2	Sales Executive	Single
			Travel_Rarely	Research & Developmen	18	3	Life Sciences	1	Male	2	Sales Executive	Married
			Travel_Rarely	Research & Developmen	1	3	Life Sciences	1	Male	3	Laboratory Technician	Married
				Research & Developmen	7	4	Medical	1	Female	4	Laboratory Technician	Divorced
				Research & Developmen	17	2	Medical	1	Male	2	Laboratory Technician	Married
				Research & Developmen	28	1	Life Sciences	1	Male	1	Laboratory Technician	Married
				Research & Developmen	14	4	Life Sciences	1	Female	1	Sales Executive	Single
				Research & Developmen	1	1	Medical	1	Male	1	Research Scientist	Married
				Research & Developmen	1	3	Life Sciences	1	Male	1	Manufacturing Director	Married
				Research & Developmen	1	3	Life Sciences	1	Male	2	Healthcare Representative	Married
				Research & Developmen	3	2	Life Sciences	1	Male	1	Laboratory Technician	Single
				Research & Developmen	1	3	Medical	1	Male	2	Sales Executive	Divorced
				Sales	7	4	Life Sciences	1	Male	1	Sales Representative	Divorced
				Research & Developmen	8	3	Life Sciences	1	Female	1	Manager	Divorced
				Research & Developmen	1	4	Other	1	Male	2	Laboratory Technician	Divorced
				Sales	8	4	Life Sciences	1	Male	1	Research Scientist	Divorced
				Research & Developmen	11	4	Life Sciences	1	Female	2	Research Scientist	Married
				Research & Developmen	4	4	Life Sciences	1	Male	1	Manufacturing Director	Married
25	29	No		Research & Developmen	16	4	Medical	1	Male	1	Laboratory Technician	Single
26	55	No		Research & Developmen	1	4	Other	1	Female	1	Research Scientist	Married
27	26	No	Travel_Frequently	Research & Developmen	9	3	Life Sciences	1	Female	1	Manager	Married
28	37	No	Travel_Rarely	Sales	5	1	Marketing	1	Male	1	Research Scientist	Single
29	44	Yes	Travel_Frequently	Research & Developmen	1	2	Medical	1	Male	2	Research Scientist	Divorced
30	38	No	Travel_Rarely	Sales	2	3	Marketing	1	Female	1	Manager	Divorced



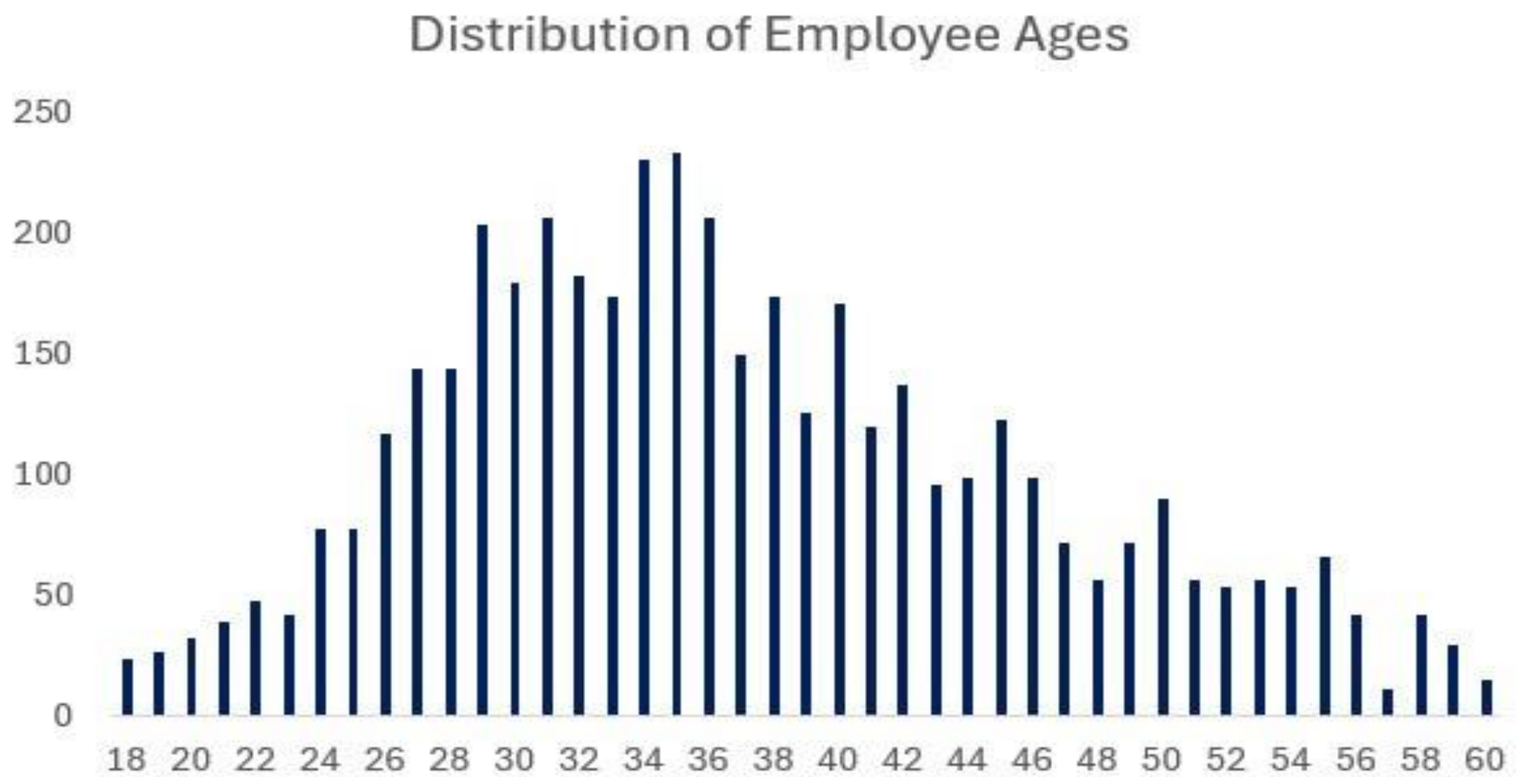
2. CREATE A PIVOT TABLE TO SUMMARIZE THE AVERAGE MONTHLY INCOME BY JOB ROLE.

Job Role	Avg. of Monthly Income
Healthcare Representative	$60.98 K
Human Resources	$58.53 K
Laboratory Technician	$66.31 K
Manager	$63.40 K
Manufacturing Director	$69.18 K
Research Director	$65.47 K
Research Scientist	$64.98 K
Sales Executive	$65.19 K
Sales Representative	$65.37 K

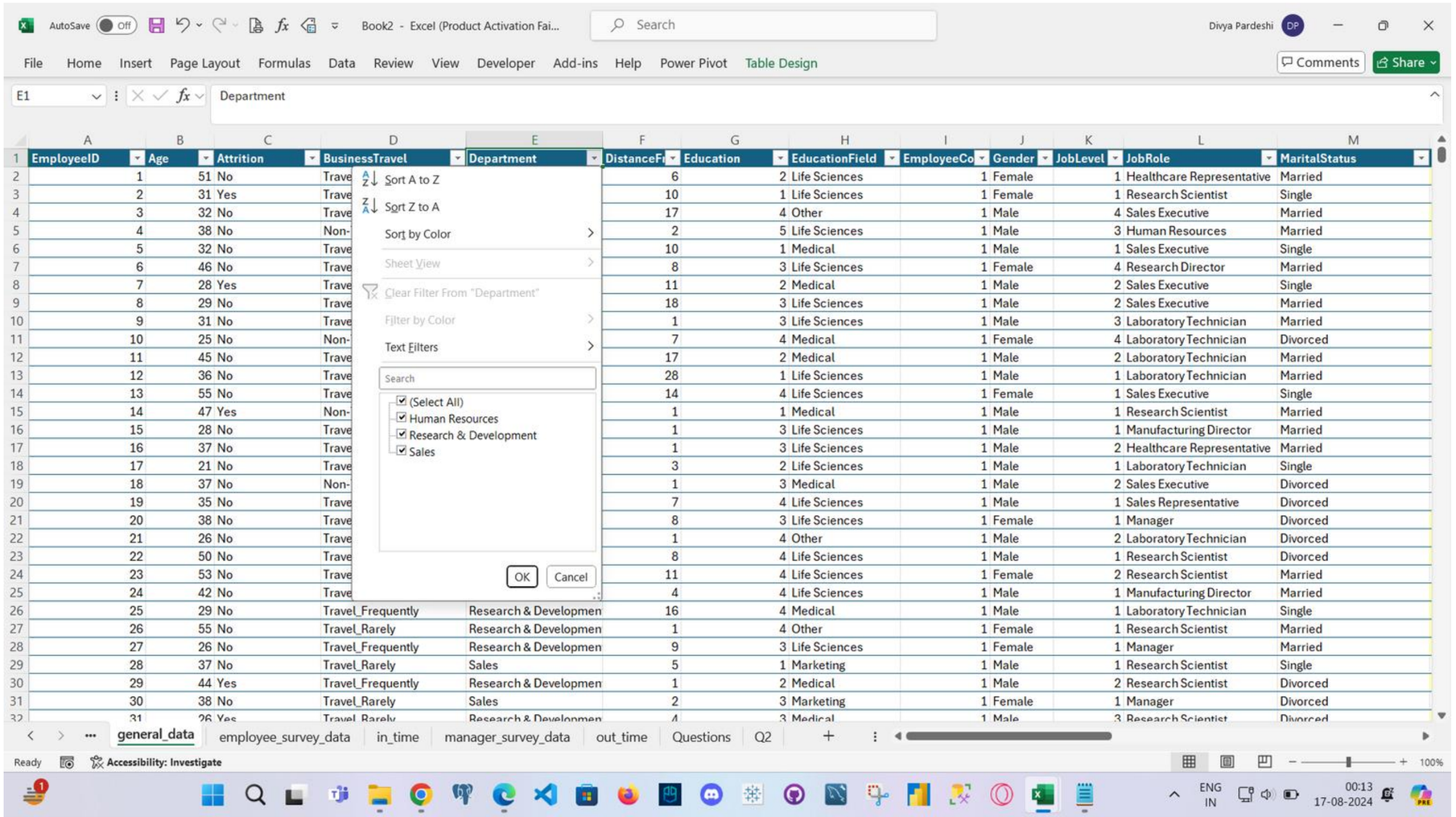
3. APPLY CONDITIONAL FORMATTING TO HIGHLIGHT EMPLOYEES WITH MONTHLY INCOME ABOVE THE COMPANY'S AVERAGE INCOME

Age	Attrition	BusinessTravel	Department	DistanceFr	Education	EducationField	EmployeeCo	Gender	JobLevel	JobRole	MaritalStatus	MonthlyIncome
51	No	Travel_Rarely	Sales	6	2	Life Sciences	1	Female	1	Healthcare Representative	Married	131160
31	Yes	Travel_Frequently	Research & Developmen	10	1	Life Sciences	1	Female	1	Research Scientist	Single	41890
32	No	Travel_Frequently	Research & Developmen	17	4	Other	1	Male	4	Sales Executive	Married	193280
38	No	Non-Travel	Research & Developmen	2	5	Life Sciences	1	Male	3	Human Resources	Married	83210
32	No	Travel_Rarely	Research & Developmen	10	1	Medical	1	Male	1	Sales Executive	Single	23420
46	No	Travel_Rarely	Research & Developmen	8	3	Life Sciences	1	Female	4	Research Director	Married	40710
28	Yes	Travel_Rarely	Research & Developmen	11	2	Medical	1	Male	2	Sales Executive	Single	58130
29	No	Travel_Rarely	Research & Developmen	18	3	Life Sciences	1	Male	2	Sales Executive	Married	31430
31	No	Travel_Rarely	Research & Developmen	1	3	Life Sciences	1	Male	3	Laboratory Technician	Married	20440
25	No	Non-Travel	Research & Developmen	7	4	Medical	1	Female	4	Laboratory Technician	Divorced	134640
45	No	Travel_Rarely	Research & Developmen	17	2	Medical	1	Male	2	Laboratory Technician	Married	79910
36	No	Travel_Rarely	Research & Developmen	28	1	Life Sciences	1	Male	1	Laboratory Technician	Married	33770
55	No	Travel_Rarely	Research & Developmen	14	4	Life Sciences	1	Female	1	Sales Executive	Single	55380
47	Yes	Non-Travel	Research & Developmen	1	1	Medical	1	Male	1	Research Scientist	Married	57620
28	No	Travel_Rarely	Research & Developmen	1	3	Life Sciences	1	Male	1	Manufacturing Director	Married	25920
37	No	Travel_Rarely	Research & Developmen	1	3	Life Sciences	1	Male	2	Healthcare Representative	Married	53460
21	No	Travel_Rarely	Research & Developmen	3	2	Life Sciences	1	Male	1	Laboratory Technician	Single	42130
37	No	Non-Travel	Research & Developmen	1	3	Medical	1	Male	2	Sales Executive	Divorced	41270
35	No	Travel_Rarely	Sales	7	4	Life Sciences	1	Male	1	Sales Representative	Divorced	24380
38	No	Travel_Rarely	Research & Developmen	8	3	Life Sciences	1	Female	1	Manager	Divorced	68700
26	No	Travel_Frequently	Research & Developmen	1	4	Other	1	Male	2	Laboratory Technician	Divorced	104470
50	No	Travel_Rarely	Sales	8	4	Life Sciences	1	Male	1	Research Scientist	Divorced	96670
53	No	Travel_Rarely	Research & Developmen	11	4	Life Sciences	1	Female	2	Research Scientist	Married	21480
42	No	Travel_Rarely	Research & Developmen	4	4	Life Sciences	1	Male	1	Manufacturing Director	Married	89260
29	No	Travel_Frequently	Research & Developmen	16	4	Medical	1	Male	1	Laboratory Technician	Single	65130
55	No	Travel_Rarely	Research & Developmen	1	4	Other	1	Female	1	Research Scientist	Married	67990
26	No	Travel_Frequently	Research & Developmen	9	3	Life Sciences	1	Female	1	Manager	Married	162910
37	No	Travel_Rarely	Sales	5	1	Marketing	1	Male	1	Research Scientist	Single	27050
44	Yes	Travel_Frequently	Research & Developmen	1	2	Medical	1	Male	2	Research Scientist	Divorced	103330
38	No	Travel_Rarely	Sales	2	3	Marketing	1	Female	1	Manager	Divorced	44480
26	Yes	Travel_Rarely	Research & Developmen	4	3	Medical	1	Male	3	Research Scientist	Divorced	68540

4. CREATE A BAR CHART IN EXCEL TO VISUALIZE THE DISTRIBUTION OF EMPLOYEE AGES.



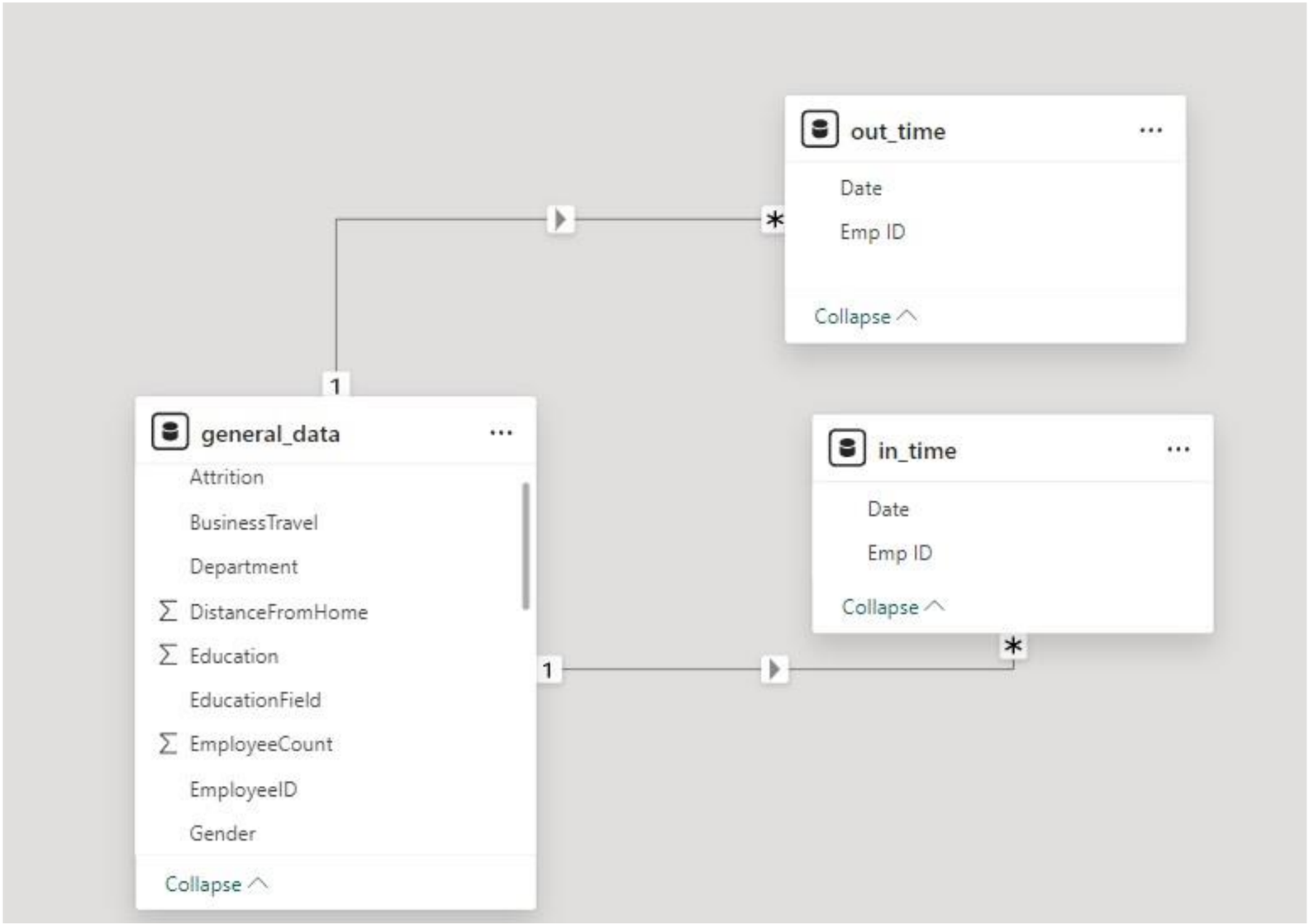
5. IDENTIFY AND CLEAN ANY MISSING OR INCONSISTENT DATA IN THE "DEPARTMENT" COLUMN.



EmployeeID	Age	Attrition	BusinessTravel	Department	DistanceFr	Education	EducationField	EmployeeCo	Gender	JobLevel	JobRole	MaritalStatus
1	51	No	Trave		6	2	Life Sciences	1	Female	1	Healthcare Representative	Married
2	31	Yes	Trave		10	1	Life Sciences	1	Female	1	Research Scientist	Single
3	32	No	Trave		17	4	Other	1	Male	4	Sales Executive	Married
4	38	No	Non-		2	5	Life Sciences	1	Male	3	Human Resources	Married
5	32	No	Trave		10	1	Medical	1	Male	1	Sales Executive	Single
6	46	No	Trave		8	3	Life Sciences	1	Female	4	Research Director	Married
7	28	Yes	Trave		11	2	Medical	1	Male	2	Sales Executive	Single
8	29	No	Trave		18	3	Life Sciences	1	Male	2	Sales Executive	Married
9	31	No	Trave		1	3	Life Sciences	1	Male	3	Laboratory Technician	Married
10	25	No	Non-		7	4	Medical	1	Female	4	Laboratory Technician	Divorced
11	45	No	Trave		17	2	Medical	1	Male	2	Laboratory Technician	Married
12	36	No	Trave		28	1	Life Sciences	1	Male	1	Laboratory Technician	Married
13	55	No	Trave		14	4	Life Sciences	1	Female	1	Sales Executive	Single
14	47	Yes	Non-		1	1	Medical	1	Male	1	Research Scientist	Married
15	28	No	Trave		1	3	Life Sciences	1	Male	1	Manufacturing Director	Married
16	37	No	Trave		1	3	Life Sciences	1	Male	2	Healthcare Representative	Married
17	21	No	Trave		3	2	Life Sciences	1	Male	1	Laboratory Technician	Single
18	37	No	Non-		1	3	Medical	1	Male	2	Sales Executive	Divorced
19	35	No	Trave		7	4	Life Sciences	1	Male	1	Sales Representative	Divorced
20	38	No	Trave		8	3	Life Sciences	1	Female	1	Manager	Divorced
21	26	No	Trave		1	4	Other	1	Male	2	Laboratory Technician	Divorced
22	50	No	Trave		8	4	Life Sciences	1	Male	1	Research Scientist	Divorced
23	53	No	Trave		11	4	Life Sciences	1	Female	2	Research Scientist	Married
24	42	No	Trave		4	4	Life Sciences	1	Male	1	Manufacturing Director	Married
25	29	No	Travel_Frequently	Research & Developmen	16	4	Medical	1	Male	1	Laboratory Technician	Single
26	55	No	Travel_Rarely	Research & Developmen	1	4	Other	1	Female	1	Research Scientist	Married
27	26	No	Travel_Frequently	Research & Developmen	9	3	Life Sciences	1	Female	1	Manager	Married
28	37	No	Travel_Rarely	Sales	5	1	Marketing	1	Male	1	Research Scientist	Single
29	44	Yes	Travel_Frequently	Research & Developmen	1	2	Medical	1	Male	2	Research Scientist	Divorced
30	38	No	Travel_Rarely	Sales	2	3	Marketing	1	Female	1	Manager	Divorced

NO 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.

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

4.12
Average_Years_With_Current_Manager

8. USING EXCEL, CREATE A PIVOT TABLE THAT DISPLAYS THE COUNT OF EMPLOYEES IN EACH MARITAL STATUS CATEGORY, SEGMENTED BY DEPARTMENT.

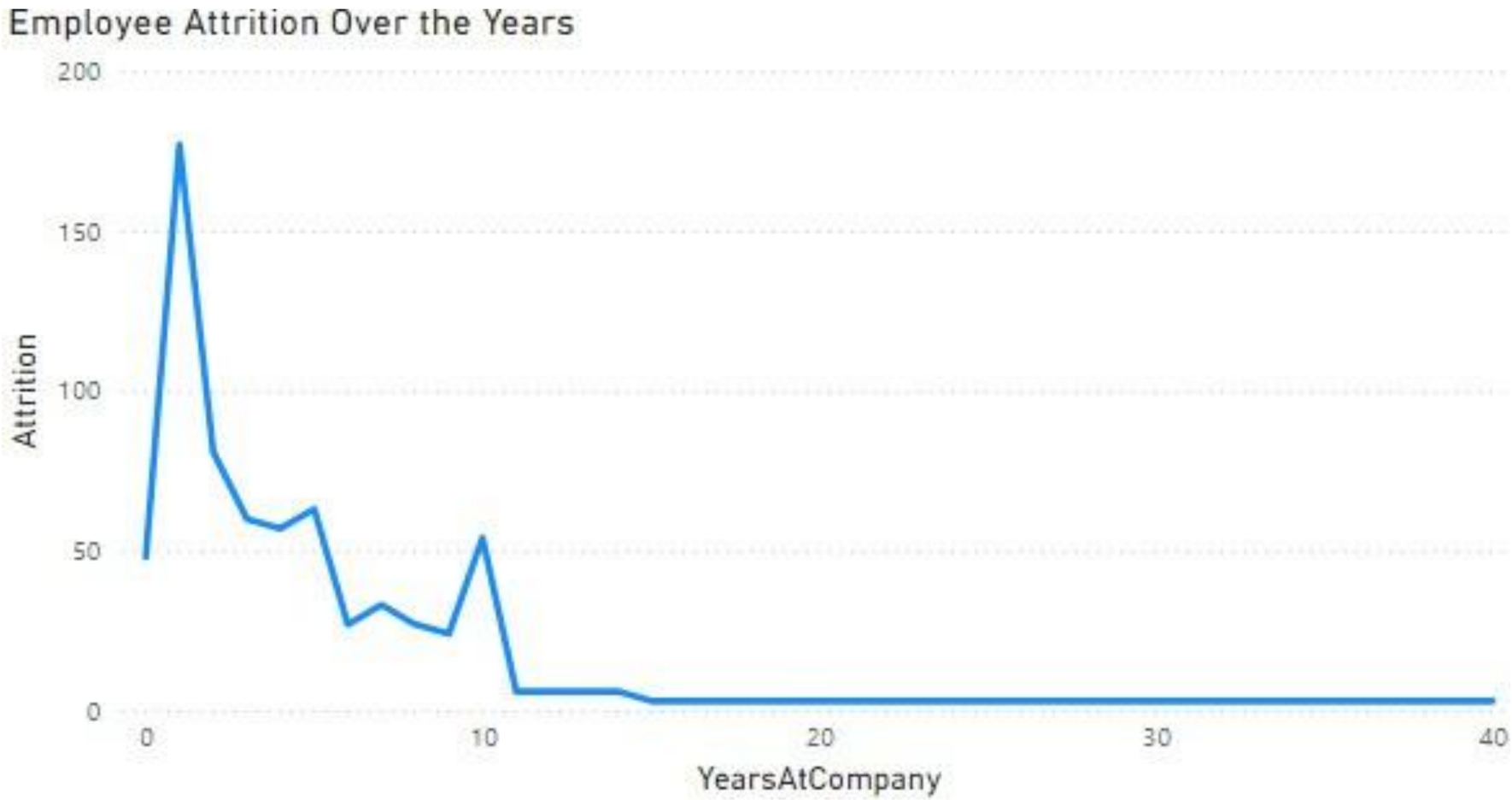
Department	Employee Count
Human Resources	**189**
Divorced	21
Married	96
Single	72
Research & Development	**2883**
Divorced	621
Married	1350
Single	912
Sales	**1338**
Divorced	339
Married	573
Single	426
Grand Total	**4410**

9. APPLY CONDITIONAL FORMATTING TO HIGHLIGHT EMPLOYEES WITH BOTH ABOVE-AVERAGE MONTHLY INCOME AND ABOVE-AVERAGE JOB SATISFACTION.

MonthlyIncome
131160
41890
193280
83210
23420
40710
58130
31430
20440
134640
79910
33770
55380
57620
25920
53460
42130
41270
24380
68700
104470
96670
21480
89260
65130
67990
162910
27050
103330
44480
68540

JobSatisfaction
4
2
2
4
1
2
3
2
4
1
4
4
1
2
4
4
3
4
2
1
2
2
3
3
4
4
1
4
3
4
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.

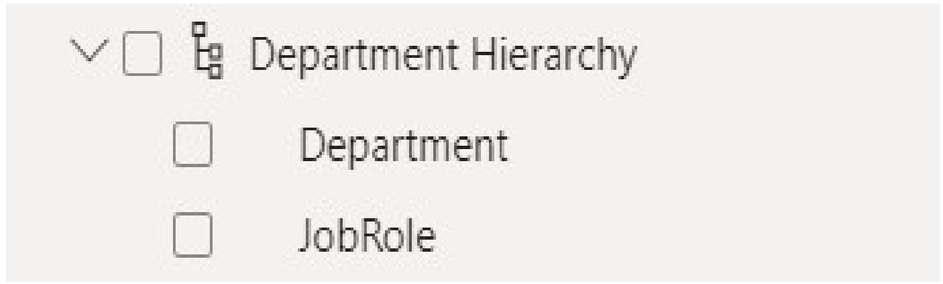
- **Optimized Query Performance:** Power BI's VertiPaq engine is highly optimized for star schemas, leading to faster data processing and report generation.
- **Simplified Data Model:** A clear separation between fact and dimension tables makes the data model easier to understand, maintain, and visualize within Power BI.
- **Efficient DAX Calculations:** DAX formulas work more efficiently with star schemas, especially for measures that involve complex aggregations or filtering across dimensions.
- **Improved Data Refresh:** Star schemas often result in smaller, more focused tables, which can lead to quicker data refresh times and more efficient memory usage in Power BI.
- **Better Visualization Performance:** The structured data model allows Power BI to generate visuals more quickly, improving the overall responsiveness of dashboards and reports.

12. USING DAX, CALCULATE THE ROLLING 3-MONTH AVERAGE OF MONTHLY INCOME FOR EACH EMPLOYEE.

```
Rolling Avg = CALCULATE(Average
(‘generaldata’[MonthlyIncome]),DATESINPERIOD(‘dimdate’[Date],
LASTDATE(‘dimdate’[Date]),-3,Month))
```

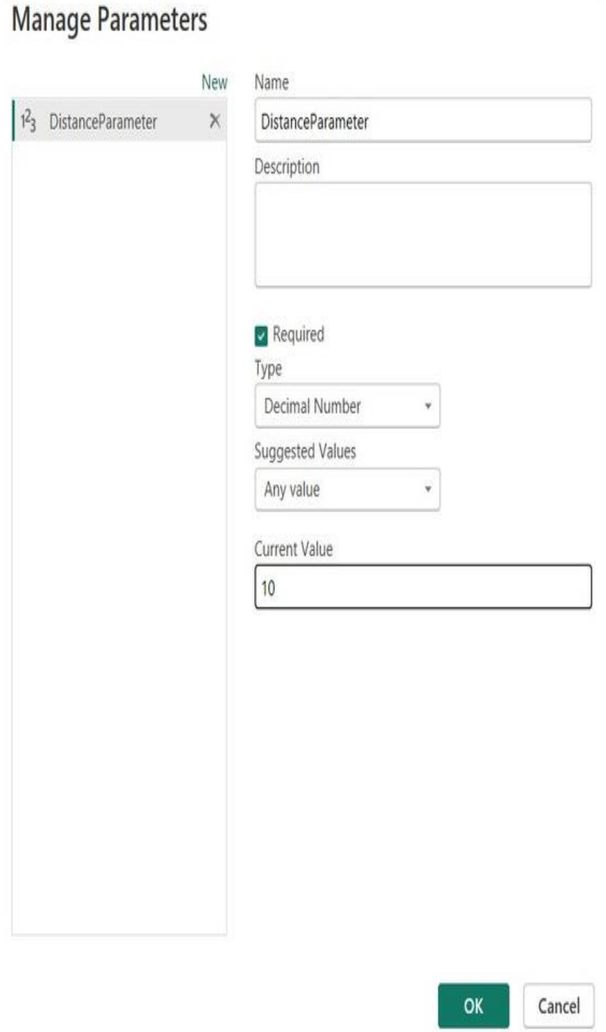
13. CREATE A HIERARCHY IN POWER BI THAT ALLOWS USERS TO DRILL DOWN FROM DEPARTMENT TO JOB ROLE TO FURTHER NARROW THEIR ANALYSIS.

Department	Total Employees
Human Resources	**189**
Healthcare Representative	9
Human Resources	3
Laboratory Technician	39
Manager	9
Manufacturing Director	24
Research Director	3
Research Scientist	36
Sales Executive	54
Sales Representative	12
Research & Development	**2883**
Sales	**1338**
Total	**4410**



14. HOW CAN YOU SET UP PARAMETERIZED QUERIES IN POWER BI TO ALLOW USERS TO FILTER DATA BASED ON THE DISTANCE FROM HOME COLUMN?

To set up parameterized queries in Power BI for filtering data based on the "Distance from Home" column:
1. Click on "Transform data" to open Power Query Editor.
2. In Power Query Editor, click on "Manage Parameters" in the "Home" tab.
3. Create a new parameter, for example, "DistanceParameter," and set its data type (like Decimal or Whole Number).
4. Close the "Manage Parameters" window and go back to the data view.
5. In the filter for the "Distance from Home" column, replace a constant value with the created parameter like "is less than or equal to DistanceParameter."
 6. Go back to the report view, and you'll see a new parameter in the right pane.
7. Users can now adjust the parameter to filter data based on different distances from home.



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.

Total Monthly Income	Departments		
Job Level	Human Resources	Research & Development	Sales
3	$1.65 M	$28.12 M	$11.79 M
4	$754.80 K	$15.28 M	$8.75 M
5	$855.84 K	$10.11 M	$2.43 M

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.

	A	B	C	D	E
1	EmployeeID	Monthly Income	PercentSalaryHike	Increased Percent Salary Hike	New Monthly Income
2	1	131160	11	12.1	1587036
3	2	41890	23	25.3	1059817
4	3	193280	15	16.5	3189120
5	4	83210	11	12.1	1006841
6	5	23420	12	13.2	309144
7	6	40710	13	14.3	582153
8	7	58130	20	22	1278860
9	8	31430	22	24.2	760606
10	9	20440	21	23.1	472164
11	10	134640	13	14.3	1925352
12	11	79910	13	14.3	1142713
13	12	33770	12	13.2	445764
14	13	55380	17	18.7	1035606
15	14	57620	11	12.1	697202
16	15	25920	14	15.4	399168
17	16	53460	11	12.1	646866
18	17	42130	12	13.2	556116
19	18	41270	13	14.3	590161
20	19	24380	16	17.6	429088
21	20	68700	11	12.1	831270
22	21	104470	18	19.8	2068506
23	22	96670	23	25.3	2445751
24	23	21480	11	12.1	259908
25	24	89260	14	15.4	1374604
26	25	65130	11	12.1	788073
27	26	67990	11	12.1	822679
28	27	162910	22	24.2	3942422
29	28	27050	11	12.1	327305
30	29	103330	14	15.4	1591282

17. 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 and handle inconsistencies, follow these steps:

1. Define the Schema:

•Clearly outline the expected structure of data, including data types, allowed categorical values, and any constraints.

2. Load the Data:

•Import the data into analysis environment, such as Excel, Power BI, or a database management system, ensuring it's loaded according to the predefined schema.

3. Check Data Types and Missing Values:

•**Data Types:** Ensure each column's data type matches the predefined schema (e.g., numeric, text, date).

•**Missing Values:** Identify and evaluate any missing values to determine if they are acceptable or need to be addressed.

4. Validate Categorical Data:

•Cross-check categorical columns against the list of predefined valid categories. Ensure no unexpected or invalid values are present.

17. VERIFY IF THE DATA ADHERES TO A PREDEFINED SCHEMA. WHAT ACTIONS WOULD YOU TAKE IF YOU FIND INCONSISTENCIES.

Handling Inconsistencies:

1.Incorrect Data Types:

Action: Convert the data to the correct type, or flag and correct any records that cannot be converted.

2.Missing Values:

Action: Handle missing data by either imputing values, removing the affected records, or flagging them for further investigation, depending on the context.

3. Invalid Categorical Values:

Action: Correct or remove any records with invalid categories, or map them to valid categories if possible.

THANK YOU