PSYLIQ

HR Data Analysis

In today's dynamic corporate landscape, understanding and leveraging human resources data is imperative for informed decision-making and strategic planning within organizations. This HR data analysis project aims to harness the power of Excel and Power BI tools to delve into comprehensive employee data, gaining valuable insights to optimize workforce management by understanding the different causes of employee attrition, and drive organizational growth.

Presented By:
Vikram Satale

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

There are total of 4410 employees whose ages were ranging from 18 years to 60 years. Among all employees, there are 3252 employees who are 30 years old or older than 30 years.

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

Average Monthly Income by Job Role

Job Role	Average Monthly Income	Number of Employees
Manufacturing Director	₹ 69,183.72	435
Laboratory Technician	₹ 66,314.05	777
Research Director	₹ 65,473.13	240
Sales Representative	₹ 65,370.96	249
Sales Executive	₹ 65,186.69	978
Research Scientist	₹ 64,975.68	876
Manager	₹ 63,395.88	306
Healthcare Representative	₹ 60,983.74	393
Human Resources	₹ 58,528.08	156
Total	**₹ 65,029.31**	**4410**

Among these different job roles, **Manufacturing Director** is the high paying job role with ₹ 69,183.72 is the average salary while Human Resources is the lowest paying job role.

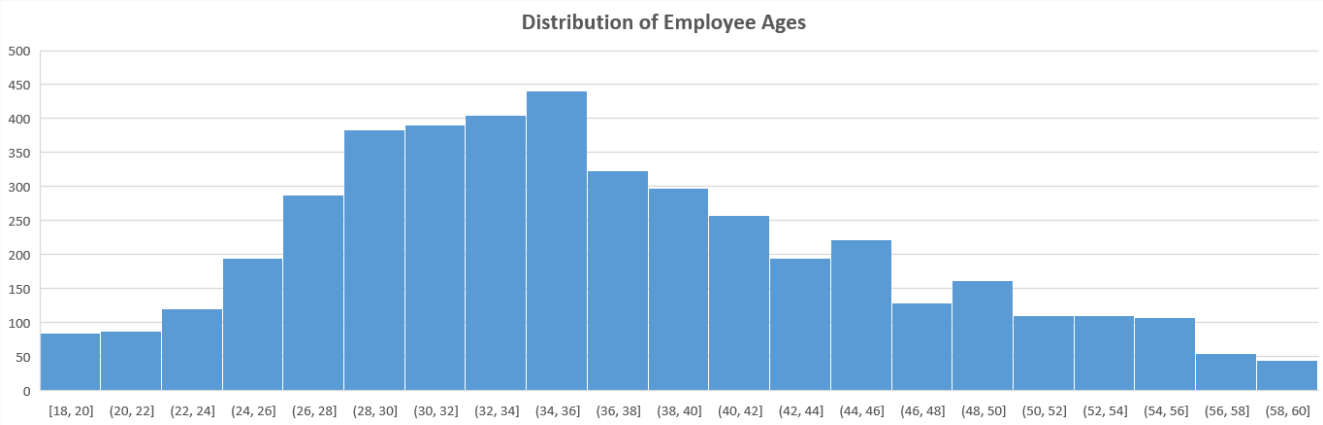
Que 3: Apply conditional formatting to highlight employees with Monthly Income above the company's average income.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Age	Attrition	BusinessTr	Departmei	DistanceFr	Education	EducationI	EmployeeC	EmployeeI	Gender	JobLevel	JobRole	MaritalSta	MonthlyIn	NumComp	Over18	PercentSal	StandardH	StockOptic	TotalWork	TrainingTir	YearsAtCo	YearsSince
2	51	No	Travel_Rai	Sales	6	2	Life Scienc	1	1	Female	1	Healthcare	Married	131160	1	Y	11	8	0	1	6	1	0
3	31	Yes	Travel_Fre	Research &	10	1	Life Scienc	1	2	Female	1	Research S	Single	41890	0	Y	23	8	1	6	3	5	1
4	32	No	Travel_Fre	Research &	17	4	Other	1	3	Male	4	Sales Exec	Married	193280	1	Y	15	8	3	5	2	5	0
5	38	No	Non-Trave	Research &	2	5	Life Scienc	1	4	Male	3	Human Re	Married	83210	3	Y	11	8	3	13	5	8	7
6	32	No	Travel_Rai	Research &	10	1	Medical	1	5	Male	1	Sales Exec	Single	23420	4	Y	12	8	2	9	2	6	0
7	46	No	Travel_Rai	Research &	8	3	Life Scienc	1	6	Female	4	Research I	Married	40710	3	Y	13	8	0	28	5	7	7
8	28	Yes	Travel_Rai	Research &	11	2	Medical	1	7	Male	2	Sales Exec	Single	58130	2	Y	20	8	1	5	2	0	0
9	29	No	Travel_Rai	Research &	18	3	Life Scienc	1	8	Male	2	Sales Exec	Married	31430	2	Y	22	8	3	10	2	0	0
10	31	No	Travel_Rai	Research &	1	3	Life Scienc	1	9	Male	3	Laboratori	Married	20440	0	Y	21	8	0	10	2	9	7
11	25	No	Non-Trave	Research &	7	4	Medical	1	10	Female	4	Laboratori	Divorced	134640	1	Y	13	8	1	6	2	6	1
12	45	No	Travel_Rai	Research &	17	2	Medical	1	11	Male	2	Laboratori	Married	79910	0	Y	13	8	2	21	2	20	4
13	36	No	Travel_Rai	Research &	28	1	Life Scienc	1	12	Male	1	Laboratori	Married	33770	0	Y	12	8	2	16	2	15	10
14	55	No	Travel_Rai	Research &	14	4	Life Scienc	1	13	Female	1	Sales Exec	Single	55380	0	Y	17	8	0	37	2	36	4
15	47	Yes	Non-Trave	Research &	1	1	Medical	1	14	Male	1	Research S	Married	57620	1	Y	11	8	2	10	4	10	9
16	28	No	Travel_Rai	Research &	1	3	Life Scienc	1	15	Male	1	Manufactu	Married	25920	1	Y	14	8	0	5	2	5	0
17	37	No	Travel_Rai	Research &	1	3	Life Scienc	1	16	Male	2	Healthcare	Married	53460	4	Y	11	8	0	7	2	5	0
18	21	No	Travel_Rai	Research &	3	2	Life Scienc	1	17	Male	1	Laboratori	Single	42130	1	Y	12	8	3	3	3	3	1
19	37	No	Non-Trave	Research &	1	3	Medical	1	18	Male	2	Sales Exec	Divorced	41270	2	Y	13	8	1	15	2	5	0
20	35	No	Travel_Rai	Sales	7	4	Life Scienc	1	19	Male	1	Sales Repr	Divorced	24380	7	Y	16	8	0	10	5	7	6
21	38	No	Travel_Rai	Research &	8	3	Life Scienc	1	20	Female	1	Manager	Divorced	68700	1	Y	11	8	1	8	5	8	7
22	26	No	Travel_Fre	Research &	1	4	Other	1	21	Male	2	Laboratori	Divorced	104470	1	Y	18	8	0	6	3	6	1
23	50	No	Travel_Rai	Sales	8	4	Life Scienc	1	22	Male	1	Research S	Divorced	96670	3	Y	23	8	0	28	2	10	1
24	53	No	Travel_Rai	Research &	11	4	Life Scienc	1	23	Female	2	Research S	Married	21480	3	Y	11	8	0	21	2	5	1
25	42	No	Travel_Rai	Research &	4	4	Life Scienc	1	24	Male	1	Manufactu	Married	89260	1	Y	14	8	0	NA	4	20	11

The average monthly income of all the employees in the company is ₹ 65029.31. Only 1479 employees are there whose salary is greater than average monthly salary of all the employees in the entire organisation.

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

The following bar graph represents the distribution of employee ages.

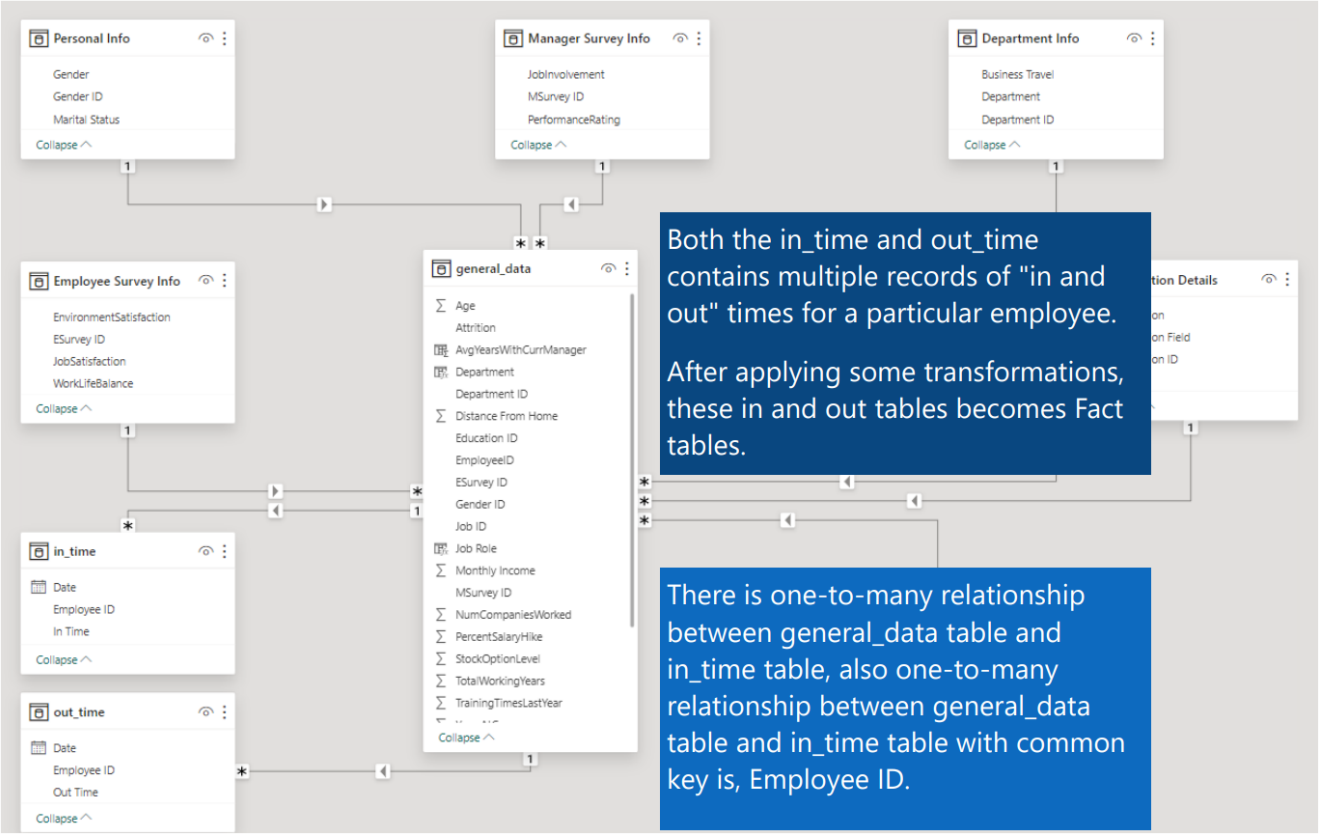


Most of the employees in the organisation are between the age groups 24 years to 46 years.

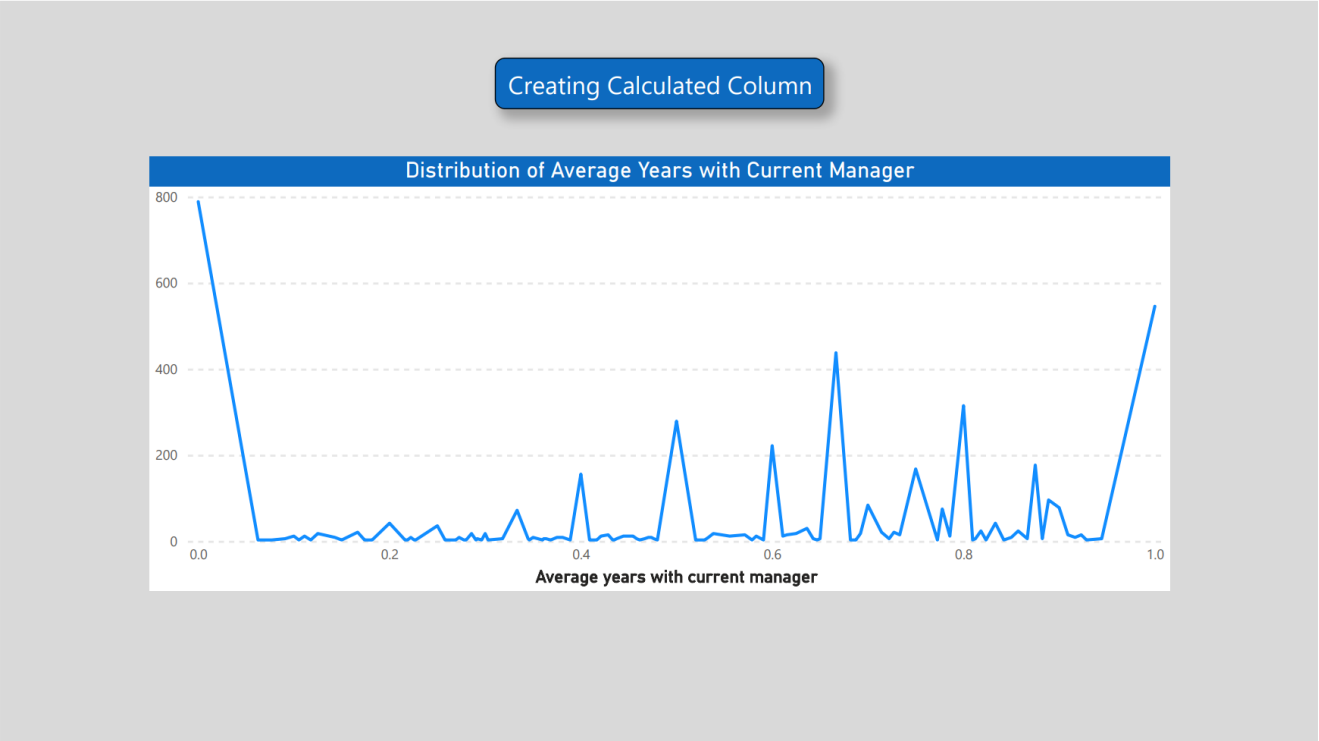
Que 5: Identify and clean any missing or inconsistent data in the "Department" column.

None of the records in the "Department" column was missing nor inconsistent.

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



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



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

Pivot Table for Marital Status Vs Department

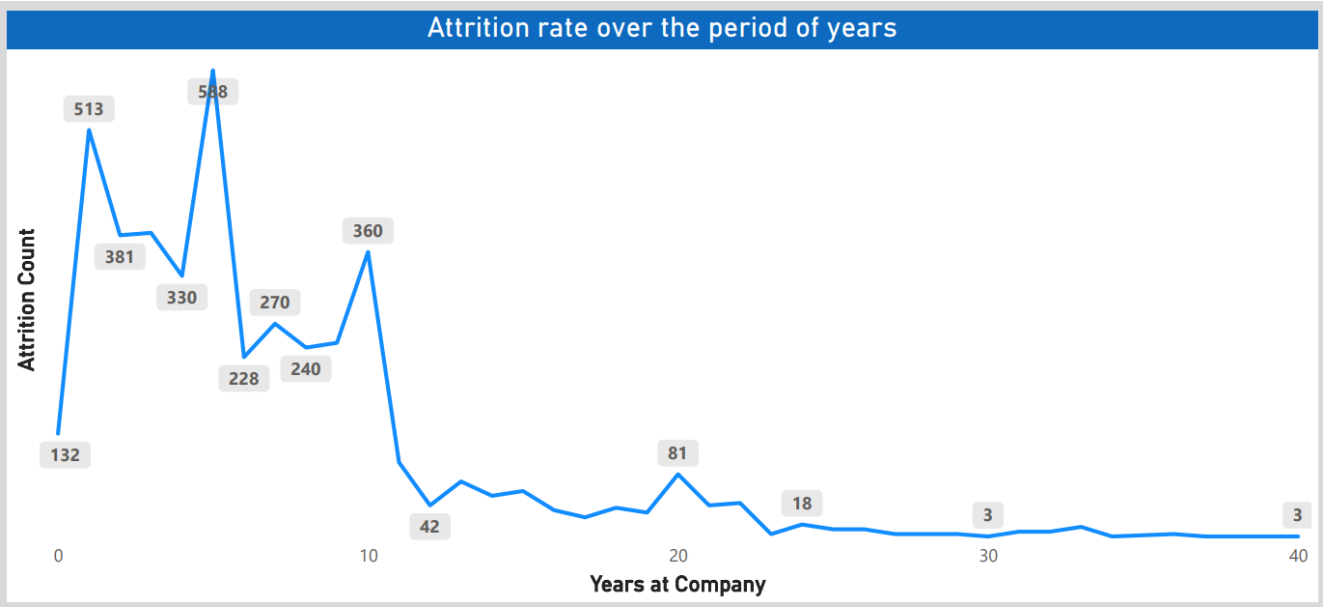
Department	Employee Count
Divorced	**981**
Human Resources	21
Research & Development	621
Sales	339
Married	**2019**
Human Resources	96
Research & Development	1350
Sales	573
Single	**1410**
Human Resources	72
Research & Development	912
Sales	426
Total	**4410**

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

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Age	Attrition	BusinessTr	Departmer	DistanceFr	Education	EducationI	EmployeeC	EmployeeI	Gender	JobLevel	JobRole	MaritalSta	MonthlyIn	Job Satisfa	NumComp	Over18	PercentSal	StandardH	StockOptic	TotalWork	TrainingTir	YearsAtCo
2	51	No	Travel_Ra	Sales	6	2	Life Scienc	1	1	Female	1	Healthcare	Married	131160	4	1	Y	11	8	0	1	6	1
3	31	Yes	Travel_Fre	Research &	10	1	Life Scienc	1	2	Female	1	Research S	Single	41890	2	0	Y	23	8	1	6	3	5
4	32	No	Travel_Fre	Research &	17	4	Other	1	3	Male	4	Sales Exec	Married	193280	2	1	Y	15	8	3	5	2	5
5	38	No	Non-Trave	Research &	2	5	Life Scienc	1	4	Male	3	Human Re	Married	83210	4	3	Y	11	8	3	13	5	8
6	32	No	Travel_Ra	Research &	10	1	Medical	1	5	Male	1	Sales Exec	Single	23420	1	4	Y	12	8	2	9	2	6
7	46	No	Travel_Ra	Research &	8	3	Life Scienc	1	6	Female	4	Research I	Married	40710	2	3	Y	13	8	0	28	5	7
8	28	Yes	Travel_Ra	Research &	11	2	Medical	1	7	Male	2	Sales Exec	Single	58130	3	2	Y	20	8	1	5	2	0
9	29	No	Travel_Ra	Research &	18	3	Life Scienc	1	8	Male	2	Sales Exec	Married	31430	2	2	Y	22	8	3	10	2	0
10	31	No	Travel_Ra	Research &	1	3	Life Scienc	1	9	Male	3	Laboratori	Married	20440	4	0	Y	21	8	0	10	2	9
11	25	No	Non-Trave	Research &	7	4	Medical	1	10	Female	4	Laboratori	Divorced	134640	1	1	Y	13	8	1	6	2	6
12	45	No	Travel_Ra	Research &	17	2	Medical	1	11	Male	2	Laboratori	Married	79910	4	0	Y	13	8	2	21	2	20
13	36	No	Travel_Ra	Research &	28	1	Life Scienc	1	12	Male	1	Laboratori	Married	33770	4	0	Y	12	8	2	16	2	15
14	55	No	Travel_Ra	Research &	14	4	Life Scienc	1	13	Female	1	Sales Exec	Single	55380	1	0	Y	17	8	0	37	2	36
15	47	Yes	Non-Trave	Research &	1	1	Medical	1	14	Male	1	Research S	Married	57620	2	1	Y	11	8	2	10	4	10
16	28	No	Travel_Ra	Research &	1	3	Life Scienc	1	15	Male	1	Manufactu	Married	25920	4	1	Y	14	8	0	5	2	5
17	37	No	Travel_Ra	Research &	1	3	Life Scienc	1	16	Male	2	Healthcare	Married	53460	4	4	Y	11	8	0	7	2	5
18	21	No	Travel_Ra	Research &	3	2	Life Scienc	1	17	Male	1	Laboratori	Single	42130	3	1	Y	12	8	3	3	3	3
19	37	No	Non-Trave	Research &	1	3	Medical	1	18	Male	2	Sales Exec	Divorced	41270	4	2	Y	13	8	1	15	2	5
20	35	No	Travel_Ra	Sales	7	4	Life Scienc	1	19	Male	1	Sales Repr	Divorced	24380	2	7	Y	16	8	0	10	5	7
21	38	No	Travel_Ra	Research &	8	3	Life Scienc	1	20	Female	1	Manager	Divorced	68700	1	1	Y	11	8	1	8	5	8
22	26	No	Travel_Fre	Research &	1	4	Other	1	21	Male	2	Laboratori	Divorced	104470	2	1	Y	18	8	0	6	3	6
23	50	No	Travel_Ra	Sales	8	4	Life Scienc	1	22	Male	1	Research S	Divorced	96670	2	3	Y	23	8	0	28	2	10
24	53	No	Travel_Ra	Research &	11	4	Life Scienc	1	23	Female	2	Research S	Married	21480	3	3	Y	11	8	0	21	2	5
25	42	No	Travel_Ra	Research &	4	4	Life Scienc	1	24	Male	1	Manufactu	Married	89260	3	1	Y	14	8	0	NA	4	20

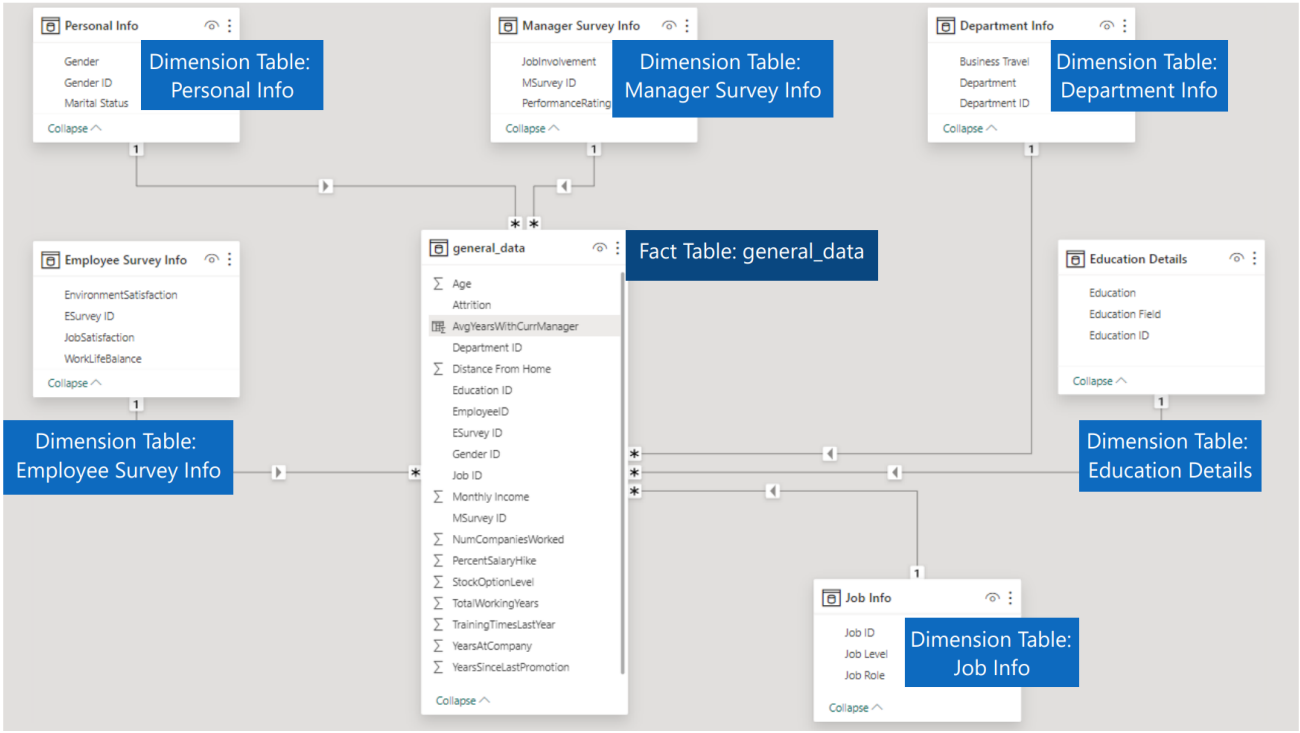
The average monthly income of all the employees in the company is ₹ 65029.31, while average job satisfaction ratings from all the employees in the company 2.728. Only 955 employees are there whose salary is greater than average monthly salary and whose job satisfaction rating is greater than average job satisfaction rating of the entire organisation.

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



The above graph shows that most of the employees are more likely to leave the company within first 5 years of their employment with the company.

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



A star schema is a data modeling technique used in data warehousing where a central "fact table" is connected to multiple "dimension tables" through primary-foreign key relationships.

Benefits of using a star schema:

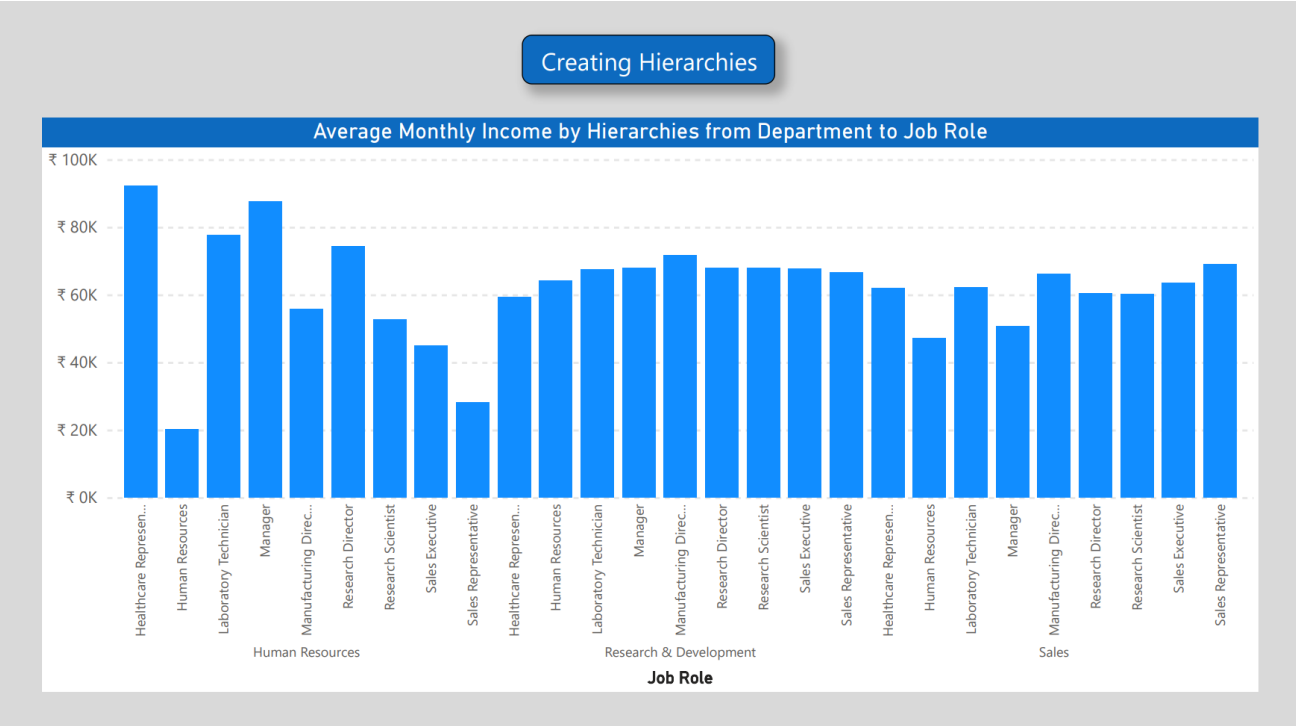
Simplified Queries, Improved Query Performance, Enhanced Understandability, Scalability and Flexibility, Optimized for Aggregation, etc. are some of the uses of a star schema.

A star schema facilitates efficient querying, analysis, and reporting, making it a preferred choice for data warehousing and analytical purposes, especially when dealing with complex datasets and analytical requirements.

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

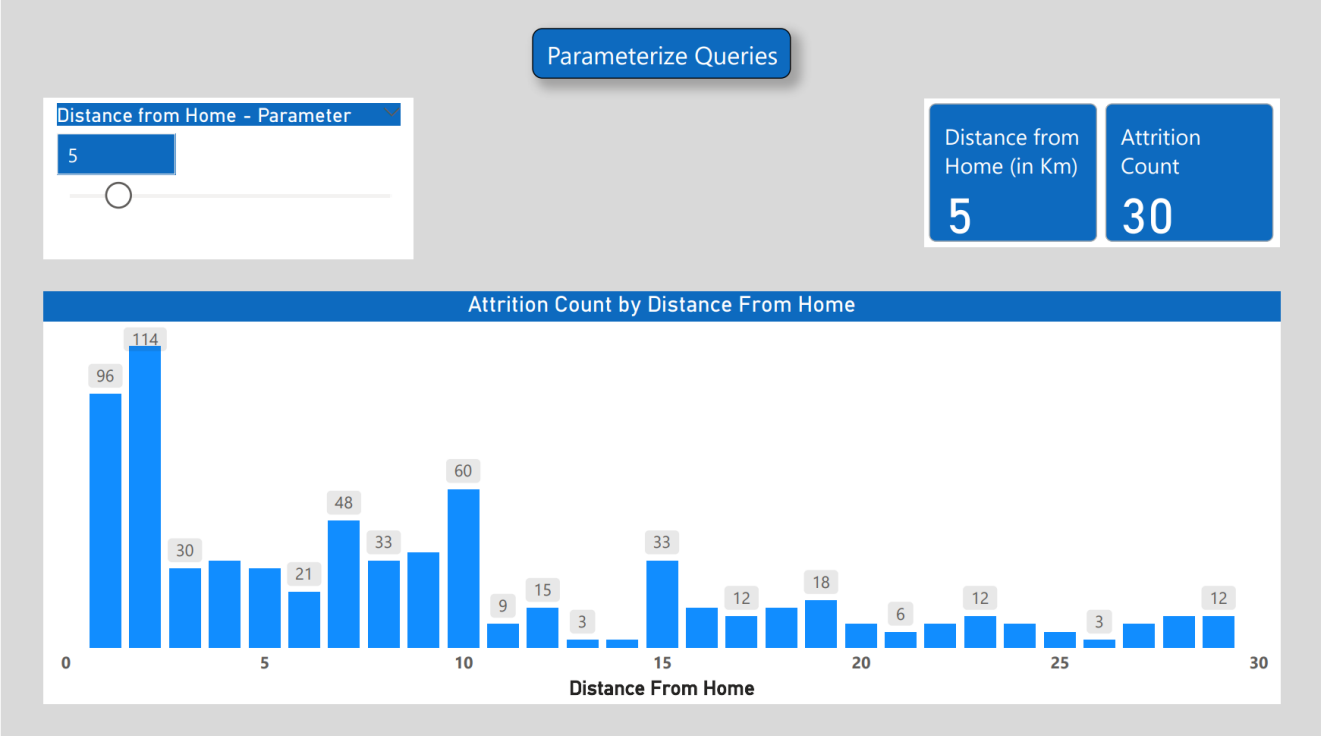
The rolling 3-month average calculates the average for last two months and the current month, for each month. There is not a date column in the given data, that means we do not have monthly salaries date-wise. So, we cannot calculate rolling 3-month average of Monthly Income.

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



After creating a hierarchy in Power BI, "Department" becomes the higher-level category and "Job Role" becomes the subcategory within the hierarchy. By creating this hierarchy, we can navigate from a broader level (Department) to a more specific level (Job Role) in their analysis, providing a more detailed view of the data.

Que 14: How can you set up parameterized queries in Power BI to allow users to filter data based on the Distance from Home column?



After setting the parameter for Distance from Home column, we can filter the attrition count corresponding to distance (in Km) from Home. For example, 30 employees left the organisation whose home were 5 Km away from the office.

Que 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 by Department

Department	Total Monthly Income
Human Resources	₹ 32,59,140
Research & Development	₹ 5,35,02,900
Sales	₹ 2,29,74,330
Total	**₹ 7,97,36,370**

JobLevel

1
2
3
4
5

From the pivot table, it can be clearly seen that the company is spending much on Research and Development Department.

Que 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.

Scenario Summary		Current Values:	5% increase	10% increase
Changing Cells:				
	Percent_Increase_in_Salary_Hike	0	0.05	0.1
Result Cells:				
	Final_SalaryE01	₹ 1,45,587.6	₹ 1,46,309.0	₹ 1,47,030.4
	Final_SalaryE02	₹ 51,524.7	₹ 52,006.4	₹ 52,488.2
	Final_SalaryE03	₹ 2,22,272.0	₹ 2,23,721.6	₹ 2,25,171.2
	Final_SalaryE04	₹ 92,363.1	₹ 92,820.8	₹ 93,278.4
	Final_SalaryE05	₹ 26,230.4	₹ 26,370.9	₹ 26,511.4
Notes: Current Values column represents values of changing cells at time Scenario Summary Report was created. Changing cells for each scenario are highlighted in gray.				

In the above table, the column “0” shows the final salaries of the randomly selected 5 employees with respect to their percent salary hikes when there is no increase in the percent salary hike. The column 0.05 shows the final salaries of those employees after 5% increase in percent salary hike, while the column 0.1 shows the final salaries of those employees after 10% increase in percent salary hike.

Using the above table one can compare the impact of increase in the percent salary hikes.

Que 17: Verify if the data adheres to a predefined schema. What actions would you take if you find inconsistencies?

The given data does not adhere to the star schema, so I split the entire data into different tables like Department Info, Education Details, Employee Survey Info, Job Info, Manager Survey Info, Personal Info, etc. All these tables are Dimension Tables since these table contains unique records, whereas general_data table is Fact table.

After gathering the information from different tables one can establish relationships between table to make star schema.

Thank you!