HR Analytics Project :

Employee Data Insights Using MySQL

Presented by : Tejas Phalke

linkedin.com/in/tejas-phalke

Project Objectives

- Analyze employee demographics, job roles, and salary patterns.
- Evaluate job satisfaction and work-life balance from surveys.
- Identify attrition rates and key turnover risk factors.
- Assess employee performance and job involvement.
- Provide data-driven HR recommendations for retention and engagement.

Datasets & Database Structure

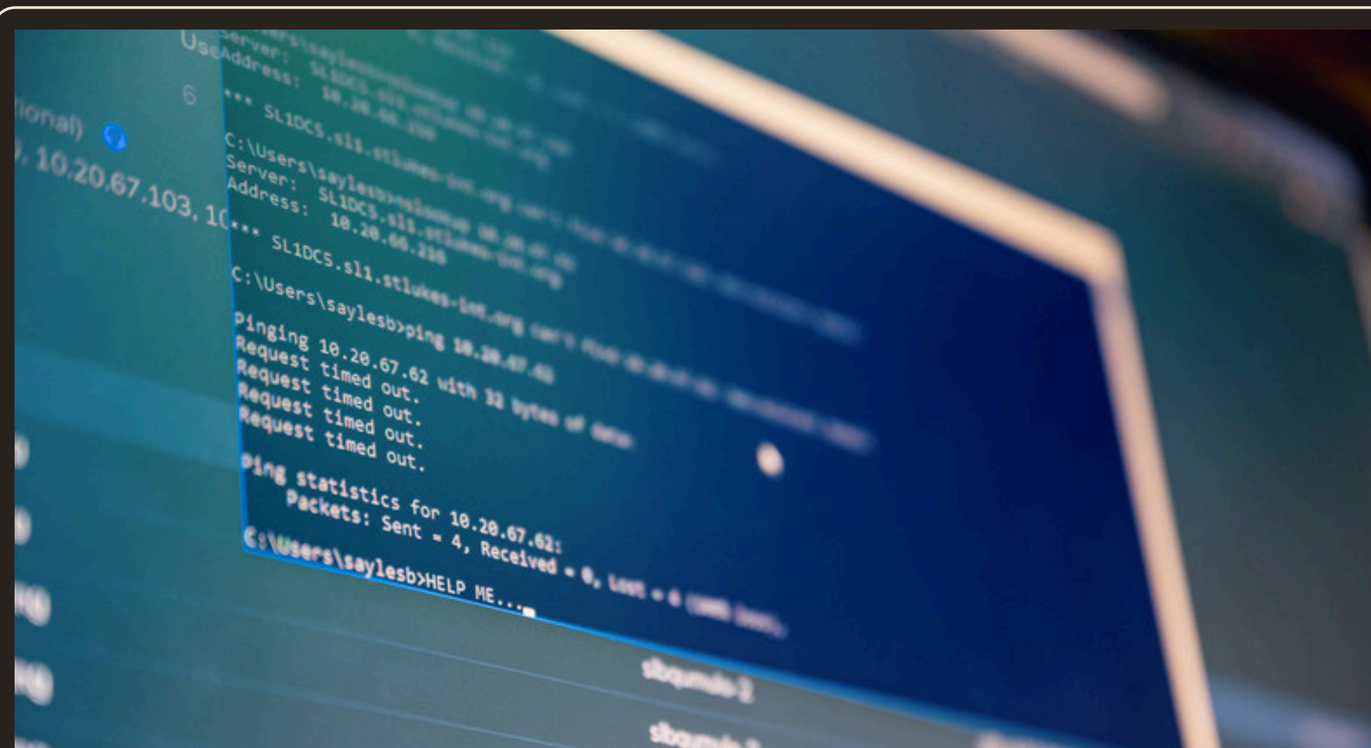
Database Name: employees_db

Process:

Created database → Created tables → Imported CSV files → Connected tables via EmployeeID

Tables:

- general_data
- employee_survey_data
- manager_survey_data



SQL Analytical Plan

Employee Demographics & Distribution

Employee Satisfaction & Work-Life Balance

Attrition Analysis

Salary & Compensation Insights

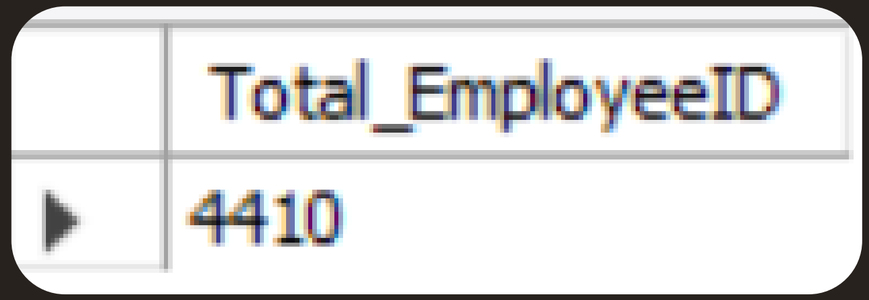
Cross-Analysis & Correlation Insights

Manager Feedback

Employee Demographics & Distribution

1) Determine the total number of employees.

```
SELECT
    COUNT(EmployeeID) AS Total_EmployeeID
FROM
    general_data;
```



Total_EmployeeID
4410

2) Count of employees by Department, Job Role, Gender, Education Field, etc.

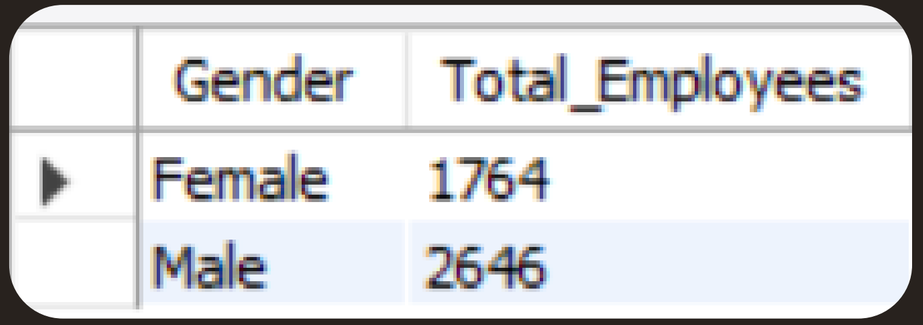
```
SELECT
   Department, COUNT(EmployeeID) AS Total_Employees
FROM
   general_data
GROUP BY Department
ORDER BY Total_Employees DESC;
```

Department	Total_Employees
Research & Development	2883
Sales	1338
Human Resources	189

```
SELECT
   JobRole, COUNT(EmployeeID) AS Total_Employees
FROM
   general_data
GROUP BY JobRole
ORDER BY Total_Employees DESC;
```

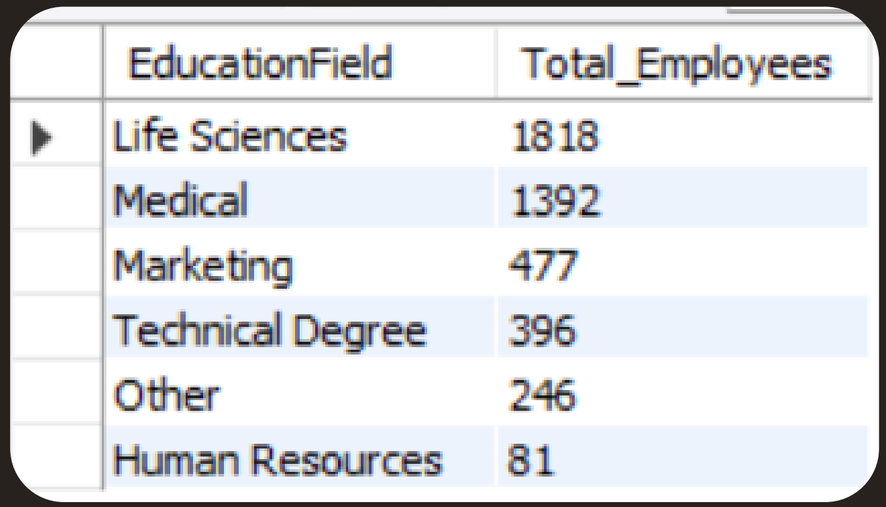
JobRole	Total_Employees
Sales Executive	978
Research Scientist	876
Laboratory Technician	777
Manufacturing Director	435
Healthcare Representative	393
Manager	306
Sales Representative	249
Research Director	240
Human Resources	156

```
SELECT
    Gender, COUNT(EmployeeID) AS Total_Employees
FROM
    general_data
GROUP BY Gender;
```



Gender	Total_Employees
Female	1764
Male	2646

```
SELECT
    EducationField, COUNT(EmployeeID) AS Total_Employees
FROM
    general_data
GROUP BY EducationField
ORDER BY Total_Employees DESC;
```



EducationField	Total_Employees
Life Sciences	1818
Medical	1392
Marketing	477
Technical Degree	396
Other	246
Human Resources	81

3) Distribution of employees by Age Group and Marital Status.

```
SELECT MaritalStatus, Age_Group, COUNT(EmployeeID) AS Total_Employees FROM (
   SELECT EmployeeID, MaritalStatus,
      CASE
         WHEN Age BETWEEN 18 AND 28 THEN '18-28'
         WHEN Age BETWEEN 29 AND 37 THEN '29-37'
         WHEN Age BETWEEN 38 AND 47 THEN '38-47'
         WHEN Age BETWEEN 48 AND 57 THEN '48-57'
         ELSE '58+'
      END AS Age_Group FROM general_data) AS grouped_data
GROUP BY MaritalStatus, Age_Group
ORDER BY Age_Group DESC;
```

MaritalStatus	Age_Group	Total_Employees
Divorced	58+	21
Married	58+	39
Single	58+	27
Married	48-57	306
Single	48-57	147
Divorced	48-57	108
Married	38-47	567
Divorced	38-47	312
Single	38-47	339
Single	29-37	567
Married	29-37	798
Divorced	29-37	405
Single	18-28	330
Divorced	18-28	135
Married	18-28	309

Employee Satisfaction & Work-Life Balance

1) Calculate the average Job Satisfaction, Environment Satisfaction, and Work-Life Balance scores.

```
SELECT
    AVG(JobSatisfaction) AS Avg_JobSatisfaction,
    AVG(EnvironmentSatisfaction) AS Avg_EnvironmentSatisfaction,
    AVG(WorkLifeBalance) AS Avg_WorkLifeBalance
FROM
    employee_survey_data;
```

Avg_JobSatisfaction	Avg_EnvironmentSatisfaction	Avg_WorkLifeBalance
2.7250	2.7245	2.7615

2) Identify departments or job roles with lower satisfaction levels.

```
-- departments
SELECT
    general_data.Department,
    AVG(employee_survey_data.JobSatisfaction) AS Avg_JobSatisfaction
FROM
    general_data
        JOIN
    employee_survey_data ON general_data.EmployeeID = employee_survey_data.EmployeeID
GROUP BY Department
ORDER BY Avg_JobSatisfaction;
```

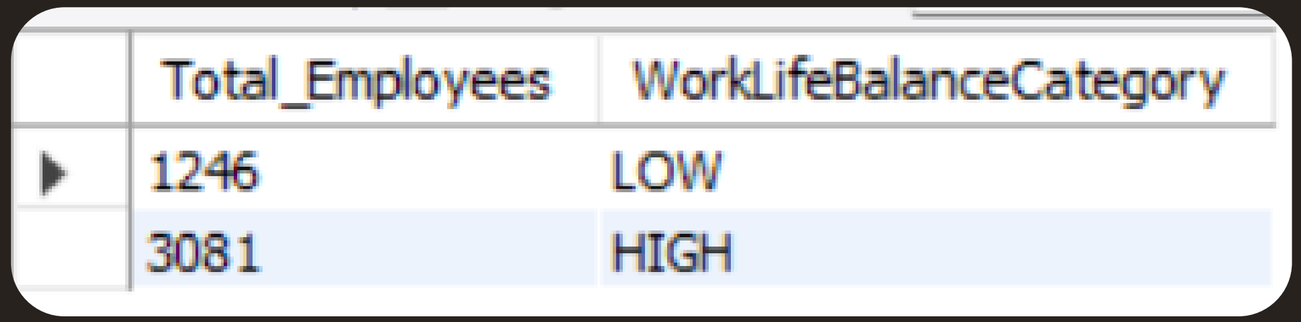
Department	Avg_JobSatisfaction
Human Resources	2.6809
Research & Development	2.7047
Sales	2.7749

```
-- job role
SELECT
    general_data.JobRole,
    ROUND(AVG(employee_survey_data.JobSatisfaction),2) AS Avg_JobSatisfaction
FROM
    general_data
        JOIN
    employee_survey_data ON general_data.EmployeeID = employee_survey_data.EmployeeID
GROUP BY JobRole ORDER BY Avg_JobSatisfaction DESC;
```

JobRole	Avg_JobSatisfaction
Human Resources	3.01
Research Director	2.85
Manufacturing Director	2.78
Healthcare Representative	2.72
Sales Executive	2.72
Manager	2.72
Laboratory Technician	2.69
Research Scientist	2.67
Sales Representative	2.64

3) Determine the number of employees with high/low Work-Life Balance ratings.

```
SELECT
   COUNT(EmployeeID) AS Total_Employees,
   WorkLifeBalance AS WorkLifeBalanceCategory
FROM
   (SELECT  employeeID,
         CASE
            WHEN WorkLifeBalance BETWEEN 1 AND 2 THEN 'LOW'
            WHEN WorkLifeBalance BETWEEN 3 AND 4 THEN 'HIGH'
            ELSE 'UNKNOW'
         END AS WorkLifeBalance
   FROM
      employee_survey_data) AS WorkLifeBalance
GROUP BY WorkLifeBalance;
```



Total_Employees	WorkLifeBalanceCategory
1246	LOW
3081	HIGH

Attrition Analysis

1) Calculate the overall attrition rate (percentage of employees who left the company).

```
SELECT
   ROUND(
      (SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) / COUNT(*)) * 100, 2
   ) AS Attrition_Percentage
FROM general_data;
```

Attrition_Percentage
16.12

2) Determine attrition rate by: Department, Job Role, Gender, Marital Status, Age Group.

Department	Total_Employees	Attrition_Count	Attrition_Percentage
Human Resources	189	57	30.16
Research & Development	2883	453	15.71
Sales	1338	201	15.02

```
-- Attrition Rate by Department
SELECT
    Department,
    COUNT(EmployeeID) AS Total_Employees,
    SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,
    ROUND(
        (SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) / COUNT(EmployeeID)) * 100, 2
    ) AS Attrition_Percentage
FROM general_data
GROUP BY Department
ORDER BY Attrition_Percentage DESC;
```

```sql
-- Attrition Rate by Job Role
SELECT
    JobRole,
    COUNT(EmployeeID) AS Total_Employees,
    SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,
    ROUND(
        (SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) / COUNT(EmployeeID)) * 100, 2
    ) AS Attrition_Percentage
FROM general_data
GROUP BY JobRole
ORDER BY Attrition_Percentage DESC;
```

JobRole	Total_Employees	Attrition_Count	Attrition_Percentage
Research Director	240	57	23.75
Research Scientist	876	159	18.15
Sales Executive	978	165	16.87
Laboratory Technician	777	126	16.22
Healthcare Representative	393	57	14.50
Sales Representative	249	36	14.46
Manager	306	42	13.73
Human Resources	156	21	13.46
Manufacturing Director	435	48	11.03

```sql
-- Attrition Rate by Gender
SELECT
    Gender,
    COUNT(EmployeeID) AS Total_Employees,
    SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,
    ROUND(
        (SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) / COUNT(EmployeeID)) * 100, 2
    ) AS Attrition_Percentage
FROM general_data
GROUP BY Gender
ORDER BY Attrition_Percentage DESC;
```

Gender	Total_Employees	Attrition_Count	Attrition_Percentage
Male	2646	441	16.67
Female	1764	270	15.31

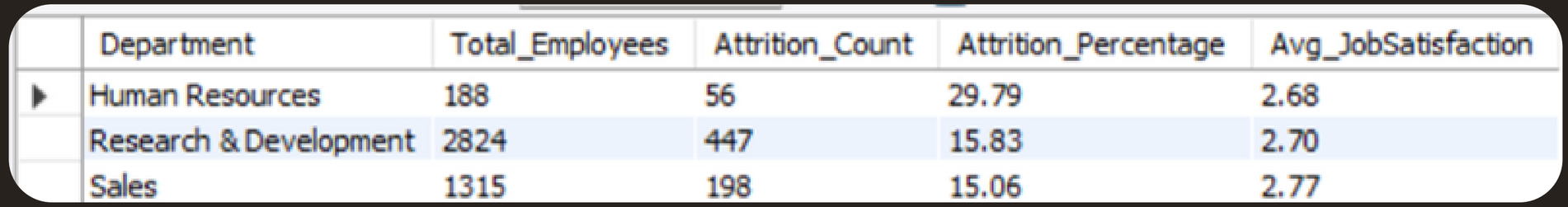
```sql
-- Attrition Rate by Marital Status
SELECT
    MaritalStatus,
    COUNT(EmployeeID) AS Total_Employees,
    SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,
    ROUND(
        (SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) / COUNT(EmployeeID)) * 100, 2
    ) AS Attrition_Percentage
FROM general_data
GROUP BY MaritalStatus
ORDER BY Attrition_Percentage DESC;
```

MaritalStatus	Total_Employees	Attrition_Count	Attrition_Percentage
Single	1410	360	25.53
Married	2019	252	12.48
Divorced	981	99	10.09

```sql
-- Attrition Rate by Age Group
SELECT
    CASE
        WHEN Age BETWEEN 18 AND 28 THEN '18-28'
        WHEN Age BETWEEN 29 AND 37 THEN '29-37'
        WHEN Age BETWEEN 38 AND 47 THEN '38-47'
        WHEN Age BETWEEN 48 AND 57 THEN '48-57'
        ELSE '58+'
    END AS Age_Group,
    COUNT(EmployeeID) AS Total_Employees,
    SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,
    ROUND(
        (SUM(CASE WHEN Attrition = 'Yes' THEN 1 ELSE 0 END) / COUNT(EmployeeID)) * 100, 2
    ) AS Attrition_Percentage
FROM general_data
GROUP BY Age_Group
ORDER BY Attrition_Percentage DESC;
```

Age_Group	Total_Employees	Attrition_Count	Attrition_Perce
18-28	774	219	28.29
58+	87	15	17.24
29-37	1770	297	16.78
48-57	561	66	11.76
38-47	1218	114	9.36

3) Identify key factors affecting attrition (e.g., high attrition in a department with low job satisfaction).



Department	Total_Employees	Attrition_Count	Attrition_Percentage	Avg_JobSatisfaction
Human Resources	188	56	29.79	2.68
Research & Development	2824	447	15.83	2.70
Sales	1315	198	15.06	2.77

```
SELECT
    g.Department,
    COUNT(g.EmployeeID) AS Total_Employees,
    SUM(CASE WHEN g.Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,
    ROUND(
        (SUM(CASE WHEN g.Attrition = 'Yes' THEN 1 ELSE 0 END) / COUNT(g.EmployeeID)) * 100, 2
    ) AS Attrition_Percentage,
    ROUND(AVG(e.JobSatisfaction), 2) AS Avg_JobSatisfaction
FROM general_data g
JOIN employee_survey_data e ON g.EmployeeID = e.EmployeeID
GROUP BY g.Department
ORDER BY Attrition_Percentage DESC;
```

Salary & Compensation Insights

1) Find the average, minimum, and maximum Monthly Income.

```
SELECT
   ROUND(AVG(MonthlyIncome), 2) AS Average_Income,
   MIN(MonthlyIncome) AS Minimum_Income,
   MAX(MonthlyIncome) AS Maximum_Income
FROM general_data;
```

Average_Income	Minimum_Income	Maximum_Income
65029.31	10090.00	199990.00

2) Compare income levels between employees who left vs. those who stayed.

```
SELECT
    Attrition,
    COUNT(EmployeeID) AS Total_Employees,
    ROUND(AVG(MonthlyIncome), 2) AS Average_Income,
    MIN(MonthlyIncome) AS Minimum_Income,
    MAX(MonthlyIncome) AS Maximum_Income
FROM general_data
GROUP BY Attrition;
```

Attrition	Total_Employees	Average_Income	Minimum_Income	Maximum_Income
No	3699	65672.60	10510.00	199990.00
Yes	711	61682.62	10090.00	198590.00

Manager Feedback

1) Find the average Job Involvement and Performance Rating scores of each Department.

Department	Avg_JobInvolvement	Avg_PerformanceRating
Human Resources	2.7460	3.1429
Sales	2.7399	3.1368
Research & Development	2.7242	3.1623

```
SELECT
    general_data.Department,
    AVG(manager_survey_data.JobInvolvement) AS Avg_JobInvolvement,
    AVG(manager_survey_data.PerformanceRating) AS Avg_PerformanceRating
FROM
    general_data
        JOIN
    manager_survey_data ON general_data.EmployeeID = manager_survey_data.EmployeeID
GROUP BY general_data.Department
ORDER BY Avg_JobInvolvement DESC;
```

2) Identify departments where performance is consistently below average

```
SELECT
    general_data.Department,
    AVG(manager_survey_data.PerformanceRating)
FROM
    general_data
        JOIN
    manager_survey_data ON general_data.EmployeeID = manager_survey_data.EmployeeID
GROUP BY Department
HAVING AVG(manager_survey_data.PerformanceRating) < (SELECT
        AVG(PerformanceRating)
    FROM
        manager_survey_data);
```

Department	AVG(manager_survey_data.PerformanceRating)
Sales	3.1368
Human Resources	3.1429

Cross-Analysis & Correlation Insights

1) Examine whether low environment satisfaction is correlated with higher attrition.

```
SELECT
   employee_survey_data.JobSatisfaction,
   COUNT(CASE

      WHEN general_data.Attrition = 'Yes' THEN 1
   END) AS Total_Attrition,
   COUNT(general_data.EmployeeID) AS Total_Employees,
   ROUND(COUNT(CASE WHEN general_data.Attrition = 'Yes' THEN 1
         END) / COUNT(general_data.EmployeeID) * 100,2) AS Attrition_Percentage
FROM general_data
      JOIN
   employee_survey_data ON general_data.EmployeeID = employee_survey_data.EmployeeID
GROUP BY employee_survey_data.JobSatisfaction ORDER BY Attrition_Percentage DESC;
```

JobSatisfaction	Total_Attrition	Total_Employees	Attrition_Percentage
1	195	851	22.91
3	216	1304	16.56
2	137	830	16.51
4	153	1342	11.40

2) Analyze if employees with poor work-life balance ratings leave more frequently.

```
SELECT
   employee_survey_data.WorkLifeBalance,
   COUNT(general_data.EmployeeID) AS Total_Employees,
   SUM(CASE WHEN general_data.Attrition = 'Yes' THEN 1 ELSE 0 END) AS Attrition_Count,
   ROUND(100.0 * SUM(CASE WHEN general_data.Attrition = 'Yes' THEN 1 ELSE 0 END) /
COUNT(general_data.EmployeeID),2) AS Attrition_Percentage
FROM general_data
      JOIN
   employee_survey_data  ON general_data.EmployeeID = employee_survey_data.EmployeeID
GROUP BY employee_survey_data.WorkLifeBalance
ORDER BY employee_survey_data.WorkLifeBalance;
```

WorkLifeBalance	Total_Employees	Attrition_Count	Attrition_Percentage
1	237	74	31.22
2	1009	167	16.55
3	2630	379	14.41
4	451	81	17.96

3) Investigate whether higher income reduces attrition risk.

```
SELECT Attrition,
    COUNT(EmployeeID) AS Total_Employees,
    ROUND(AVG(MonthlyIncome),2) AS Avg_MonthlyIncome
FROM general_data
GROUP BY Attrition;
```

Attrition	Total_Employees	Avg_MonthlyIncome
No	3699	65672.60
Yes	711	61682.62

4) Identify which departments have the most satisfied and best-performing employees.

```
SELECT
    general_data.Department,
    SUM(CASE WHEN employee_survey_data.JobSatisfaction = 4 THEN 1 ELSE 0 END) AS HighlySatisfiedCount,
    SUM(CASE WHEN manager_survey_data.PerformanceRating = 4 THEN 1 ELSE 0 END) AS
HighPerformerCount
FROM
    general_data
JOIN
    employee_survey_data ON general_data.EmployeeID = employee_survey_data.EmployeeID
JOIN
    manager_survey_data ON general_data.EmployeeID = manager_survey_data.EmployeeID
GROUP BY
    general_data.Department
HAVING
    HighlySatisfiedCount > 0 AND HighPerformerCount > 0
```

Department	HighlySatisfiedCount	HighPerformerCount
Sales	429	182
Research & Development	850	456
Human Resources	63	27

Some Important Key Insights

Majority of employees are aged between 28-37

R&D department records highest attrition.

Low work-life balance strongly correlates with higher attrition.

Higher job satisfaction and income reduce attrition risk.

High performers primarily concentrated in Sales and R&D departments.

Recommendations

1. Improve work-life balance initiatives, especially in R&D.
2. Adjust compensation structures in high-turnover, low-salary roles.

3. Focus retention programs on employees aged 28-37.

4. Provide recognition and career progression opportunities for high performers.

Conclusion

This HR analytics project successfully demonstrates the application of SQL in extracting meaningful insights from employee data to support strategic HR decisions. The recommendations provided can aid leadership in enhancing employee engagement, optimizing compensation, and reducing turnover.

In summary, This HR analytics project successfully demonstrates the application of SQL in extracting meaningful insights from employee data to support strategic HR decisions. The recommendations provided can aid leadership in enhancing employee engagement, optimizing compensation, and reducing turnover.

Thank You

Contact Information

Address: Pune, Maharashtra, India.
Phone: +91- 8956761150
Email: phalketejas89@gmail.com