

MY FIRST SQL PROJECT!

Exploring Reasons underlying
Employee Attrition using SQL



SHREYAJAIN8593@GMAIL.COM

The project involves an exploratory analysis of a dataset of 1470 employees from a fictitious organization.

The objective is to identify the reasons/factors underlying employee attrition.

I've used MySQL Workbench to complete this project.

Each of the following slides has 3 parts: The research question, an SQL query to find an answer, and its result and interpretation.

Let's go!



Retrieve a list of unique job roles within the company.

```
SELECT DISTINCT JobRole  
FROM hr_attrition.dataset;
```

	JobRole
▶	Research Scientist
	Sales Representative
	Laboratory Technician
	Human Resources
	Healthcare Representative
	Sales Executive
	Manufacturing Director
	Research Director
	Manager

Compare the average monthly income of employees in different departments.

```
SELECT Department,  
ROUND(AVG(MonthlyIncome),2)  
AS "Average Monthly Income"  
FROM hr_attrition.dataset  
GROUP BY Department  
ORDER BY ROUND(AVG(MonthlyIncome),2);
```

	Department	Average Monthly Income
▶	Research & Development	6281.25
	Human Resources	6654.51
	Sales	6959.17

For each job role, calculate the average age and average monthly income of employees.

```
SELECT DISTINCT JobRole AS "Job Role",
ROUND(AVG(MonthlyIncome),0) AS 'Monthly Income',
ROUND(AVG(Age),0) AS "Average Age"
FROM hr_attrition.dataset
GROUP BY JobRole
ORDER BY `Monthly Income` DESC;
```

	Job Role	Monthly Income	Average Age
►	Manager	17182	47
	Research Director	16034	44
	Healthcare Representative	7529	40
	Manufacturing Director	7295	38
	Sales Executive	6924	37
	Human Resources	4236	36
	Research Scientist	3240	34
	Laboratory Technician	3237	34
	Sales Representative	2626	30

Count the number of employees who have left the company and are still active.

Step 1:

```
ALTER TABLE hr_attrition.dataset
ADD COLUMN EmployeeStatus VARCHAR(255)
AS
(CASE
    WHEN ATTRITION = 'YES'
        THEN 'Departed Employees'
    ELSE 'Active Employees' END);
```

Step 2:

```
SELECT EmployeeStatus AS 'Employee Status',
COUNT(EmployeeStatus) AS 'Headcount'
FROM hr_attrition.dataset
GROUP BY EmployeeStatus;
```



	Employee Status	Headcount
▶	Departed Employees	237
	Active Employees	1233

Calculate the attrition rate for each department.

```
SELECT Department,  
ROUND  
  (SUM(CASE WHEN ATTRITION='Yes' THEN 1 ELSE 0 END)  
  /COUNT(*) * 100,2)  
AS 'Attrition Rate'  
FROM hr_attrition.dataset  
GROUP BY DEPARTMENT  
ORDER BY `ATTRITION RATE` ASC;
```

	Department	Attrition Rate
▶	Research & Development	13.84
	Human Resources	19.05
	Sales	20.63



Among the three departments in this organization, the sales department has the highest rate of employee turnover.

Are there gender differences in employee attrition?

```
SELECT GENDER AS 'Gender',
ROUND(AVG(age),2) AS 'Avg. Age',
COUNT(*) AS 'Headcount',
→ (SUM
  [ (CASE WHEN Attrition='Yes' THEN 1 ELSE 0 END)
    /COUNT(*)) * 100
  AS 'Attrition Rate',
ROUND(AVG(MonthlyIncome),0) AS 'Avg Monthly Income'
FROM hr_attrition.dataset
GROUP BY Gender;
```



Although there are more male employees than female employees, women tend to leave the company less often than men. On average, both male and female employees receive nearly equal compensation, with women earning just 4% more than their male counterparts.

	Gender	Avg. Age	Headcount	Attrition Rate	Avg Monthly Income
▶	Male	36.65	882	17.0068	6381
	Female	37.33	588	14.7959	6687

Identify the top three job roles with the highest attrition rates.

```
SELECT JobRole AS 'Job Role',
ROUND((SUM(CASE WHEN ATTRITION='Yes' THEN 1 ELSE 0 END)/COUNT(*)) * 100,2)
AS 'Attrition Rate'
FROM hr_attrition.dataset
GROUP BY JobRole
ORDER BY `ATTRITION RATE` DESC
LIMIT 3;
```

	Job Role	Attrition Rate
▶	Sales Representative	39.76
	Laboratory Technician	23.94
	Human Resources	23.08



Supporting the earlier finding of increased turnover in the sales department, it appears that sales representatives are the most prone to resign from this organization.

Do more experienced employees tend to leave the company?

```
SELECT EmployeeStatus AS 'Employee Status',  
ROUND(AVG(Age),2) AS 'Average Age',  
ROUND(AVG(TotalWorkingYears),2) AS 'Total working years',  
ROUND(AVG(YearsAtCompany),2) AS 'Years at this company',  
ROUND(AVG(YearsInCurrentRole),2) AS 'Years in present role'  
FROM hr_attrition.dataset  
GROUP BY EmployeeStatus;
```



The overall findings suggest that younger employees with less work experience are more prone to resign. This could be due to the higher likelihood of younger individuals engaging in job hopping, as those with less experience may be more eager to advance their careers quickly.

	Employee Status	Average Age	Total working years	Years at this company	Years in present role
▶	Departed Employees	33.61	8.24	5.13	2.90
	Active Employees	37.56	11.86	7.37	4.48

Does compensation differ for employees who left and those who stayed?

```
SELECT EmployeeStatus AS 'Employee Status',  
MIN(MonthlyIncome) AS 'Min Monthly Income',  
ROUND(AVG(MonthlyIncome),0) AS 'Avg Monthly Income',  
MAX(MonthlyIncome) AS 'Max Monthly Income',  
ROUND(AVG(PercentSalaryHike),2) AS 'Avg Percent salary hike',  
AVG(YearsSinceLastPromotion) AS 'Years Since Last Promotion'  
FROM hr_attrition.dataset  
GROUP BY EmployeeStatus
```



Compensation and benefits could be driving resignations at this organization, as departed employees have 43% lower average monthly income.

	Employee Status	Min Monthly Income	Avg Monthly Income	Max Monthly Income	Avg Percent salary hike	Years Since Last Promotion
▶	Departed Employees	1009	4787	19859	15.10	1.9451
	Active Employees	1051	6833	19999	15.23	2.2344

Calculate the attrition rate for employees who worked overtime versus those who didn't.

```
SELECT OVERTIME AS 'Working Overtime (Yes/No)',  
(SUM(CASE WHEN ATTRITION='YES'  
THEN 1 ELSE 0 END)/COUNT(*))*100  
AS 'Attrition Rate',  
(SUM(CASE WHEN WorkLifeBalanceKey='Bad' THEN 1 ELSE 0 END)/COUNT(*)) * 100  
AS '% Employees with Bad Work Life Balance',  
(SUM(CASE WHEN WorkLifeBalanceKey='Best' THEN 1 ELSE 0 END)/COUNT(*)) * 100  
AS '% Employees with Best Work Life Balance'  
FROM hr_attrition.dataset  
GROUP BY OVERTIME;
```



A threefold higher attrition rate is observed among employees reporting overtime, suggesting that overtime conditions might be a key factor driving attrition.

	Working Overtime (Yes/No)	Attrition Rate	% Employees with Bad Work Life Balance	% Employees with Best Work Life Balance
▶	Yes	30.5288	5.2885	8.6538
	No	10.4364	5.5028	11.1006

How do departed and active employees differ on their psychological satisfaction levels?

```
SELECT EmployeeStatus AS 'Employee Status',
AVG(PerformanceRating) AS 'Avg Performance Rating',
(SUM(CASE WHEN Relationshipsatkey='Low' THEN 1 ELSE 0 END)/COUNT(*)) * 100
AS '% Employees with Low Relationship Satisfaction',
(SUM(CASE WHEN JobSatisfactionKey='Low' THEN 1 ELSE 0 END)/COUNT(*)) * 100
AS '% Employees with Low Job Satisfaction',
(SUM(CASE WHEN EnvironmentSatisfaction='1' THEN 1 ELSE 0 END)/COUNT(*)) * 100
AS '% Employees with Low Environment Satisfaction'
FROM hr_attrition.dataset
GROUP BY `Employee Status`;
```



A higher percentage of departed employees report lower satisfaction levels across all dimensions – relationships, job satisfaction, and the work environment, compared to their active counterparts. Dissatisfaction might be a contributing factor to employee departures.

Employee Status	Avg Performance Rating	% Employees with Low Relationship Satisfaction	% Employees with Low Job Satisfaction	% Employees with Low Environment Satisfaction
Departed Employees	3.1561	24.0506	27.8481	30.3797
Active Employees	3.1533	17.7616	18.0860	17.1938

THANK YOU!



SHREYAJAIN8593@GMAIL.COM