EMPLOYEE ATTRITION DATASET ANALYSIS

INTRODUCTION

Employee attrition, or turnover, is a critical concern for organizations as it directly impacts productivity, morale, and the financial health of a company. Understanding the factors contributing to employee attrition is essential for developing strategies to retain talent and reduce turnover rates. This analysis aims to explore an employee attrition dataset to identify key factors influencing employees' decisions to leave the organization. By leveraging statistical methods and machine learning techniques, we will examine variables such as job satisfaction, work-life balance, compensation, career progression, and organizational culture, among others.

AIM

The aim of analysing an employee attrition dataset is to identify the key factors that contribute to why employees leave an organization, such as job satisfaction, work environment, compensation, and career growth opportunities. By developing predictive models, the analysis seeks to pinpoint employees who are at a higher risk of leaving, enabling organizations to take proactive measures to retain valuable talent. Additionally, the analysis provides data-driven insights to inform and optimize human resource strategies aimed at reducing turnover and enhancing employee retention. Ultimately, this analysis supports decision-making processes that foster a positive work environment and improve employee satisfaction.

OBJECTIVES

1. Identify Key Drivers: Determine the primary factors that influence employee attrition within the organization.

2. Predict Attrition: Designed and executed SQL queries to extract relevant information from the database to forecast which employees are most likely to leave.

3. Profile High-Risk Groups: Identify specific employee segments (e.g., by department, role) that are more prone to attrition.

4. Inform HR Strategies: Provide actionable insights to help HR develop and refine strategies to improve employee retention.

5. Enhance Employee Satisfaction: Use findings to recommend changes in policies or practices that can increase job satisfaction and reduce turnover.

DATA OVERVIEW

Columns	Datatype
EmployeeID	int
Age	int
Attrition	text
BusinessTravel	text
DailyRate	int
Department	text
DistanceFromHome	int
Education	int
EducationField	text
EmployeeCount	int
EnvironmentSatisfaction	int
Gender	text
HourlyRate	int
JobInvolvement	int
JobLevel	int
JobRole	text
JobSatisfaction	int
MaritalStatus	text
MonthlyIncome	int
MonthlyRate	int
salary_category	varchar(50)
age_category	varchar(50)

DATA ANALYSIS

1. Descriptive Analysis: Summarize data to understand overall attrition trends, average tenure, and most affected departments.

2. Trend Analysis: Identify patterns over time, such as monthly or yearly attrition rates and changes in employee satisfaction.

3. Employee Segmentation: Group employees based on demographics or job roles to tailor retention strategies.

4. Performance Analysis: Analyse performance metrics, including average productivity and its correlation with attrition.

5. Attrition Reasons Analysis: Understand the reasons and frequency of employee attrition across different segments.

QUESTIONS

1.NUMBER OF EMPLOYEES WHO HAD ALREADY LEFT:

-->SELECT COUNT(*) FROM employee_attirtion_new

WHERE Attrition="Yes";

2.COUNT OF MALE AND FEMALE OUT OF THE ABOVE DATA:

-->select Gender,count(Gender) from employee_attirtion_new

where Attrition="Yes"

group by Gender;

3.EMPLOYEES OF AGE BETWEEN 18 AND 35,36 AND 60 AND ABOVE 60:

-->alter table employee_attirtion_new add column age_category varchar(20);

update employee_attirtion_new set age_category=case when Age between 18 and 35 then "Young adults"

when Age between 36 and 60 then "Middle Aged" else "Old aged" end;

4.EMPLOYEES WITH LOW,AVERAGE AND HIGH SALARY:

-->alter table employee_attirtion_new add column salary_category varchar(20);

update employee_attirtion_new set salary_category=case when monthlyincome<5000 then 'Low salary'

when monthlyincome>=5000 and monthlyincome<=10000 then 'Average salary'

else 'High salary' end;

5.Average age of Employees:

-->select avg(Age) from employee_attirtion_new;

6.Average age of Employees who have already left:

-->select avg(Age) from employee_attirtion_new where Attrition="Yes";

7.From which age category more employees have left the job?

-->select age_category,count(*) from employee_attirtion_new where Attrition="Yes" group by age_category;

8.From which salary category more employees have left the job?

-->select salary_category,count(*) from employee_attirtion_new where Attrition="Yes" group by salary_category;

9.Employees left the job according to Business Travel:

-->select BusinessTravel,count(*) from employee_attirtion_new where Attrition="Yes" group by BusinessTravels;

10.Employees in each department:

-->select Department,count(*) from employee_attirtion_new group by Department;

11.Employees left from each department:

-->select Department,count(*) from employee_attirtion_new where Attrition="Yes" group by Department;

12.Average distance of employee from Home:

-->select avg(DistanceFromHome) from employee_attirtion_new;

13.Maximum number of employees with how much year of education:

-->select Education,count(Education) from employee_attirtion_new group by Education;

14.Maximum number of employees in which EducationField:

-->select EducationField,count(EducationField) from employee_attirtion_new group by EducationField;

15.Average Environment Satisfaction Rating:

-->select avg(EnvironmentSatisfaction) from employee_attirtion_new;

16.Maximum number of the employee with which Job Role has left the job:

-->select max(JobRole) from employee_attirtion_new where Attrition="Yes";

17.Least number of employee with Job Role left the job:

-->select min(JobRole) from employee_attirtion_new where Attrition="Yes";

18.Marital status of Maximum employee who left the job:

```
select max(MaritalStatus) from employee_attirtion_new
where Attrition="Yes";
```

CONCLUSION

1.Total Employees Who Left:

199 out of 1676 employees have left.

2.Gender Distribution of Employees Who Left:

113 males and 86 females.

3.Age Distribution of Employees:

- Young adults (18-35): 833
- Middle-aged (36-60): 843
- No employees above 60 years.

4.Salary Distribution:

- Low Salary (<5000): 859
- Average Salary (5000-10000): 492
- High Salary (>10000): 325

5.Average Age:

- Overall: 36.87 years.
- Employees who left: 30.90 years.

6. Age Category with Most Attrition:

Young adults.

7.Salary Category with Most Attrition:

Low salary.

8.Attrition by Business Travel Frequency:

Most attrition from those who travel rarely.

9.Department Distribution:

- Cardiology: 531

- Maternity: 796
- Neurology: 349

10.Attrition by Department:

- Maternity: 98
- Cardiology: 74
- Neurology: 27

11.Average Distance from Home:

9.22 units.

12.Education Level with Maximum Employees:

Level 3 with 655 employees.

13.Education Field with Maximum Employees:

Life Sciences: 697 employees.

14.Average Environment Satisfaction:

2.71 rating.

15.Job Role with Maximum Attrition:

Therapist.

16.Job Role with Minimum Attrition:

Administrative.

17.Marital Status with Maximum Attrition:

Single.