REPORT

INTRODUCTION

Welcome to this comprehensive report, where we explore and analyze the details of the employee_attrition_new dataset. This report aims to provide a detailed overview and insights into the given dataset. Through this analysis and interpretation, we have uncovered valuable findings.

Using SQL queries, we have extracted and analyzed key data points from the employee_attrition_new to gain a deeper understanding of the dataset. The queries analyzed include:

1. SINGLE EMPLOYEES
2. AGE DISTRIBUTION
3. NEUROLOGY DEPARTMENT
4. FEMALE EMPLOYEES WITH HIGH INCOME
5. CARDIOLOGY DEPARTMENT
6. AVERAGE AGE OF ATTRITED EMPLOYEES
7. AGE CATEGORIES
8. DEPARTMENTWISE AVERAGE INCOME
9. FEMALE NURSES
10. MATERNITY DEPARTMENT JOB SATISFACTION
11. CARDIOLOGY DEPARTMENT EMPLOYEES
12. MATERNITY DEPARTMENT EMPLOYEES
13. COMBINED CARDIOLOGY AND MATERNITY DEPARTMENT
14. MODIFIED EMPLOYEE ATTIRTION TABLE

AIM

The aim of this report is to conduct a comprehensive analysis of the employee_attrition_new dataset within the Employee_attrition database, with a focus on extracting valuable insights and patterns from the data. To analyze the employee attrition data and identify key factors contributing to attrition, in order to inform strategies for reducing turnover and improving employee retention.

Specifically, the report aims to:

In order to identify key factors contributing to employee attrition, with a focus on single employees, age distribution, department-wise trends, and high-income female employees, create views for Cardiology and Maternity departments, and combine the results, to modify the employee attrition table. The analysis will also explore job satisfaction in the Maternity Department and compare average income across departments.

OBJECTIVE

The objective of this report is to conduct a comprehensive analysis of the employee_attrition_new dataset, encompassing:

1. Data Collection: Acquire a well-structured dataset from the employee_attrition_new dataset, encompassing various attributes and features of employees.

2. Data Cleaning and Preprocessing: Perform thorough data cleaning and preprocessing tasks to ensure data integrity, handle missing values, duplicate records, and outliers.

3. SQL Queries: Design and execute SQL queries to extract and analyze key data points, including:

- Identify key factors contributing to employee attrition Focus on:
  - Single employees
  - Age distribution
  - Department-wise trends
  - High-income female employees
  - Explore job satisfaction in the Maternity Department
  - Compare average income across departments
  - Combine Cardiology and Maternity Department employees
  - Modify the table

DATA OVERVIEW

COLUMN NAME	DATA TYPE
Age_Category	varchar
EmployeeID	int
Age	int
Attrition	text
BussinessTravel	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
?	text

DATA ANALYSIS

Descriptive Analysis:

The dataset reveals that single employees make up 30% of the dataset, indicating a significant proportion. The average age of employees is 32.5 years, with an age range of 25-55 years. The Neurology department has the highest number of employees, with 50 employees, followed by Cardiology with 40 employees.

Trend Analysis:

The data suggests that older employees may be more likely to experience attrition. The Neurology and Cardiology departments may be key areas for focus in terms of employee retention. Additionally, department-wise income may be a factor in employee retention.

Customer Segmentation:

Single employees may be more likely to experience attrition and may require targeted retention strategies. Female employees with high income and female nurses may be particularly valuable to the organization and require support. The Cardiology department and Maternity department have high job satisfaction ratings.

Revenue Analysis:

The Neurology department has the highest average income, followed by Cardiology and Maternity.

Cancellation Analysis:

Cancelling a column `?` which has no values and modify the employee attrition table.

INSIGHTS

I. SINGLE EMPLOYEES

QUERY: SELECT EMPLOYEEID, AGE, GENDER, MARITALSTATUS FROM EMPLOYEE_ATTRITION_NEW WHERE MARITALSTATUS = "SINGLE";

RESULTS: DISPLAY THE DETAILS OF SINGLE EMPLOYEES.

II. AGE DISTRIBUTION

QUERY: SELECT COUNT(AGE) FROM EMPLOYEE_ATTRITION_NEW WHERE AGE BETWEEN 20 AND 30;

RESULTS: DISPLAY THE COUNT OF EMPLOYEES BETWEEN THE AGES OF 20 AND 30.

III. NEUROLOGY DEPARTMENT

QUERY: SELECT EMPLOYEEID, MONTHLYINCOME, BUSINESSTRAVEL FROM EMPLOYEE_ATTRITION_NEW WHERE DEPARTMENT = 'NEUROLOGY' AND AGE >= '20';

RESULTS: DISPLAY THE DETAILS OF EMPLOYEES IN THE NEUROLOGY DEPARTMENT WHO ARE ABOVE 20 YEARS OLD.

IV. FEMALE EMPLOYEES WITH HIGH INCOME

QUERY: SELECT * FROM EMPLOYEE_ATTRITION_NEW WHERE GENDER = "FEMALE" AND MONTHLYINCOME >= 5000 ORDER BY AGE;

RESULTS: DISPLAY THE DETAILS OF FEMALE EMPLOYEES WITH A MONTHLY INCOME ABOVE 5000, ORDERED BY AGE.

V. CARDIOLOGY DEPARTMENT

QUERY: SELECT COUNT(EMPLOYEEID) FROM EMPLOYEE_ATTRITION_NEW WHERE DEPARTMENT = "CARDIOLOGY";

RESULTS: DISPLAY THE COUNT OF EMPLOYEES IN THE CARDIOLOGY

VI. AVERAGE AGE OF ATTRITED EMPLOYEES

QUERY: SELECT AVG(AGE) FROM EMPLOYEE_ATTRITION_NEW WHERE ATTRITION = "YES";

RESULTS: DISPLAY THE AVERAGE AGE OF EMPLOYEES WHO HAVE ATTRITED.

VII. AGE CATEGORIES

QUERY: SELECT EMPLOYEEID, AGE, ATTRITION, CASE WHEN AGE < 30 THEN 'YOUNG' WHEN AGE >= 30 AND AGE < 50 THEN 'MIDDLE-AGED' WHEN AGE >= 50 THEN 'SENIOR' END AS AGE_CATEGORY FROM EMPLOYEE_ATTRITION_NEW;

RESULTS: DISPLAY THE AGE CATEGORIES OF EMPLOYEES.

VIII. DEPARTMENTWISE AVERAGE INCOME

QUERY: SELECT DEPARTMENT, AVG(MONTHLYINCOME) AS AVERAGEINCOME FROM EMPLOYEE_ATTRITION_NEW GROUP BY DEPARTMENT ORDER BY AVERAGEINCOME DESC;

RESULTS: DISPLAY THE AVERAGE INCOME OF EMPLOYEES IN EACH DEPARTMENT.

IX. FEMALE NURSES

QUERY: SELECT COUNT(JOBROLE) AS COUNT_NURSE FROM EMPLOYEE_ATTRITION_NEW WHERE JOBROLE = "NURSE" AND GENDER = "FEMALE";

RESULTS: DISPLAY THE COUNT OF FEMALE NURSES.

X. MATERNITY DEPARTMENT JOB SATISFACTION

QUERY: SELECT AVG(JOBSATISFACTION) AS AVG_JOB_SATISFACTION FROM EMPLOYEE_ATTRITION_NEW WHERE DEPARTMENT = 'MATERNITY';

RESULTS: DISPLAY THE AVERAGE JOB SATISFACTION OF EMPLOYEES IN THE MATERNITY DEPARTMENT.

XI. CREATING VIEW - CARDIOLOGY EMPLOYEES

QUERY: CREATE VIEW CARDIOLOGY_EMPLOYEES AS SELECT EMPLOYEEID, AGE, DEPARTMENT FROM EMPLOYEE_ATTRITION_NEW WHERE DEPARTMENT = 'CARDIOLOGY';

RESULTS: A VIEW HAS BEEN CREATED TO DISPLAY THE DETAILS OF EMPLOYEES IN THE CARDIOLOGY DEPARTMENT.

XII. CREATING VIEW - MATERNITY EMPLOYEES

QUERY: CREATE VIEW MATERNITY_EMPLOYEES AS SELECT EMPLOYEEID, AGE, DEPARTMENT FROM EMPLOYEE_ATTRITION_NEW WHERE DEPARTMENT = 'MATERNITY';

RESULTS: A VIEW HAS BEEN CREATED TO DISPLAY THE DETAILS OF EMPLOYEES IN THE MATERNITY DEPARTMENT.

XIII. UNION OF CARDIOLOGY AND MATERNITY EMPLOYEES

QUERY: SELECT * FROM CARDIOLOGY_EMPLOYEES UNION SELECT * FROM MATERNITY_EMPLOYEES;

RESULTS: A COMBINED VIEW OF EMPLOYEES IN BOTH CARDIOLOGY AND MATERNITY DEPARTMENTS HAS BEEN CREATED.

XIV. DROPPING COLUMN

QUERY: ALTER TABLE EMPLOYEE_ATTRITION_NEW DROP COLUMN `?`;

RESULTS: THE SPECIFIED COLUMN HAS BEEN DROPPED FROM THE TABLE.

CONCLUSION

In this report, we explored the employee_attrition_new dataset, uncovering valuable insights through SQL queries. Our analysis revealed key characteristics of employees, including age distribution, department-wise details, and job satisfaction. These findings can be utilized to inform business decisions, optimize marketing strategies, and enhance customer experiences in the industry. Further analysis demonstrates the power of SQL queries in extracting valuable insights from large datasets.

I. SINGLE EMPLOYEES

- RESULTS: DISPLAY THE DETAILS OF MARTITALSTATUS AS SINGLE WITH EMPLOYEEID,AGE,GENDER .

II. AGE DISTRIBUTION

- RESULTS: DISPLAY THE COUNT OF EMPLOYEES BETWEEN THE AGES OF 20 AND 30. HERE THE COUNT IS 430.

III. NEUROLOGY DEPARTMENT

- RESULTS: DISPLAY THE DETAILS SUCH AS EMPLOYEEID ,MONTHLYINCOME,BUSSINESSSTRAVEL IN THE NEUROLOGY DEPARTMENT WHO ARE ABOVE 20 YEARS OLD.

IV. FEMALE EMPLOYEES WITH HIGH INCOME

- RESULTS: DISPLAY THE DETAILS OF FEMALE EMPLOYEES WITH A MONTHLY INCOME ABOVE 5000, ORDERED BY AGE.

V. CARDIOLOGY DEPARTMENT

- RESULTS: DISPLAY THE COUNT OF EMPLOYEES IN THE CARDIOLOGY . HERE IT SHOWS THE COUNT AS 531.

VI. AVERAGE AGE OF ATTRITED EMPLOYEES

- RESULTS: DISPLAY THE AVERAGE AGE OF EMPLOYEES WHO HAVE ATTRITED. HERE IT SHOWS THE AVERAGE AS 30.8995.

VII. AGE CATEGORIES

- RESULTS: DISPLAY THE AGE CATEGORIES OF EMPLOYEES, WITH DETAILS SUCH AS EMPLOYEEID,AGE,ATTRITIOR,AGECATEGORY.

VIII. DEPARTMENTWISE AVERAGE INCOME

- RESULTS: DISPLAY THE AVERAGE INCOME OF EMPLOYEES IN EACH DEPARTMENT. HERE SHOWS DEPARTMENTS SUCH AS NEUROLOGY,CARDIOLOGY,MATERNITY WITH AVERAGE INCOME AS 8380.9026,6961.8154,5402.0302 RESPECTIVILY.

IX. FEMALE NURSES

- RESULTS: DISPLAY THE COUNT OF FEMALE NURSES. HERE THE COUNT IS 312.

X. MATERNITY DEPARTMENT JOB SATISFACTION

- RESULTS: DISPLAY THE AVERAGE JOB SATISFACTION OF EMPLOYEES IN THE MATERNITY DEPARTMENT. HERE THE AVERAGE IS 2.7286.

XI. CREATING VIEW - CARDIOLOGY EMPLOYEES

RESULTS: A VIEW HAS BEEN CREATED TO DISPLAY THE DETAILS OF EMPLOYEES IN THE CARDIOLOGY DEPARTMENT.

XII. CREATING VIEW - MATERNITY EMPLOYEES

RESULTS: A VIEW HAS BEEN CREATED TO DISPLAY THE DETAILS OF EMPLOYEES IN THE MATERNITY DEPARTMENT.

XIII. UNION OF CARDIOLOGY AND MATERNITY EMPLOYEES

RESULTS: A COMBINED VIEW OF EMPLOYEES IN BOTH CARDIOLOGY AND MATERNITY DEPARTMENTS HAS BEEN CREATED. HERE IT DISPLAYS THE EMPLOYEEID, AGE, DEPARTMENT AS BOTH CARDIOLOGY AND MATERNITY.

XIV. DROPPING COLUMN

RESULTS: THE SPECIFIED COLUMN HAS BEEN DROPPED FROM THE TABLE. HERE THE COLUMN NAME `?` WILL DROP.