HR Analytics Project Using MySQL

Introduction :

This project focuses on performing HR analytics by analyzing employee-related data to uncover actionable insights for improving employee engagement, retention, satisfaction, and overall organizational performance. Structured Query Language (SQL) is utilized to extract, join, and analyze data from multiple employee datasets including demographics, compensation, surveys, and performance feedback.

Project Objectives :

The primary objectives of this HR analytics project are:

1. To analyze employee demographics, job roles, and salary patterns for understanding workforce distribution and identifying trends.
2. To evaluate employee satisfaction and work-life balance using employee survey responses.
3. To conduct attrition analysis and identify risk factors contributing to employee turnover.
4. To assess employee performance and job involvement based on manager evaluations.
5. To provide data-driven recommendations to HR and leadership for workforce management improvements.

Datasets & Database Structure :

In this project, a custom relational database named employee_data_analysis has been designed and used for performing employee-related data analysis. The database consists of three structured tables, each representing different aspects of employee information within the organization.

Table Name	Description
general_data	Contains employee demographic details, job information, salary, attrition status, and career history.
employee_survey_data	Includes employee feedback on work environment, job satisfaction, and work-life balance.
manager_survey_data	Contains manager evaluations of employee job involvement and performance ratings.

The database was built by importing the following CSV files:

- general_data.csv
- employee_survey_data.csv
- manager_survey_data.csv

❖ Dataset Download Link - Employee Details Dataset

Data Import Process :

- A new database named employees_db was created using MySQL Workbench.
- Individual tables (general_data, employee_survey_data, manager_survey_data) were created within the database with appropriate data types for each column.
- CSV files were imported into these tables using Table Data Import Wizard.
- Data types were validated to ensure correct formats for numeric and text fields.
- Null and missing values were reviewed and handled where necessary.
- Primary key relationships were established using EmployeeID to maintain relational integrity across tables.

SQL-Based Analytical Plan :

To derive important insights from the given employee datasets, the following SQL queries were formulated and executed:

- Employee Demographics & Distribution

(From general_data.csv)

- Determine the total number of employees.
- Count of employees by Department, Job Role, Gender, Education Field, etc.
- Distribution of employees by Age Group and Marital Status.

- Employee Satisfaction & Work-Life Balance

(From employee_survey_data.csv and manager_survey_data.csv)

- Calculate the average Job Satisfaction, Environment Satisfaction, and Work-Life Balance scores.
- Identify departments or job roles with lower satisfaction levels.
- Determine the number of employees with high/low Work-Life Balance ratings.

- Attrition Analysis

(From general_data.csv)

- Calculate the overall attrition rate (percentage of employees who left the company).
- Determine attrition rate by:
    - Department
    - Job Role
    - Gender
    - Marital Status
    - Age Group
- Identify key factors affecting attrition (e.g., high attrition in a department with low job satisfaction).

- Salary & Compensation Insights

(From general_data.csv)

- Find the average, minimum, and maximum Monthly Income.
- Analyze income distribution by Job Role, Department, and Gender.
- Compare income levels between employees who left vs. those who stayed.

- Manager Feedback

(From manager_survey_data.csv)

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

- Identify managers or departments where performance is consistently below average.

- Cross-Analysis & Correlation Insights

(Combining data from all three datasets)

- Examine whether low environment satisfaction is correlated with higher attrition.
- Analyze if employees with poor work-life balance ratings leave more frequently.
- Investigate whether higher income reduces attrition risk.
- Identify which departments have the most satisfied and best-performing employees.

- Summary

Objectives, like:

- 📊 Demographics & Distribution
- 📈 Employee Satisfaction
- 🔍 Attrition Analysis
- 💰 Salary & Compensation
- 🌟 Performance & Career Progression
- 📝 Manager Feedback
- 🔄 Cross-Analysis

This structured SQL-based approach ensures a comprehensive analysis of employee demographics, satisfaction metrics, attrition trends, performance indicators, and managerial feedback. The insights derived will enable management to make informed, data-driven decisions for improving employee engagement, retention strategies, and workforce planning.

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 :

- 📌 Improve work-life balance initiatives, especially in R&D.
- 📌 Adjust compensation structures in high-turnover, low-salary roles.
- 📌 Focus retention programs on employees aged 28-37.
- 📌 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.