HR Analytics Project – End-to-End Analysis Report

1. Introduction -

In this project, I performed a complete HR Analytics study to understand employee behaviour, attrition patterns, and performance trends. The goal was to use data-driven insights to help HR teams improve workforce planning and retention.
I handled the entire pipeline—data cleaning, exploratory analysis, SQL analytics, Python visualizations, and Power BI dashboard creation.
This project helped me strengthen my analytical thinking, technical skills, and ability to convert raw data into business insights.

2. Dataset Description -

The dataset contains employee information across demographics, employment type, job role, performance, and attrition.
Key columns used:

- Employee_ID
- Gender
- Age
- Marital_Status
- City
- Department
- Job_Role
- Education
- Experience_Years
- Employment_Type
- Salary
- Performance_Rating
- Performance_Improvement_Plan
- Total_Trainings
- Leave_Count
- Manager_ID
- Attrition
- Age_Group
- Experience_Level
- Employment_Status

The dataset is realistic and covers most HR-related use cases such as Attrition analysis, performance evaluation, and workforce planning.

3. Objective of the Project -

The main objectives were:

1. Identify factors influencing employee attrition.
2. Study patterns across job roles, departments, and experience levels.
3. Analyze training impact, performance scores, and salary distribution.
4. Build an interactive dashboard to present insights to HR stakeholders.

4. Tools & Technologies Used -

- Python: Pandas, NumPy, Matplotlib, Seaborn
- SQL (MySQL): Data exploration and analytical queries
- Power BI: Dashboard, DAX measures
- Excel: VLOOKUP, INDEX-MATCH, data validation, Pivot Tables & Charts, Dashboard
- GitHub: Version control & project hosting

5. Data Cleaning & Pre-Processing -

I performed systematic data cleaning to prepare the dataset:

- Check for duplicates.
- Handled missing values in Date column.
- Rename the column names.
- Corrected data types (Salary, Experience, Joining Date, Last Working Date, etc.)
- Created two new derived features:
    - Age_Group (Young, Adult, Senior)
    - Experience_Level (Junior, Mid-Level, Senior, Expert)
- Created Attrition_Flag (0/1) for modeling and visualizations

This stage ensured that the data was consistent, structured, and ready for analysis.

6. Python Analysis (Pandas + Visualizations) -

- Exploratory Data Analysis (EDA)
- Attrition segmentation by department, job role, experience, age group, Marital Status
- Salary Distribution analysis
- Performance rating Vs training relationship
- Boxplots, bar charts, distribution plots
- Training impact and leave count patterns
- Attrition by Education

Key Python Insights:

- Support Lead role showed the highest attrition.
- Employees are higher in the age of 30-35.
- Employees who received fewer trainings tended to have lower performance scores.
- Fewer Employees are set to join with Performance Improvement Plan.

7. SQL Analysis -

I loaded the cleaned dataset into MySQL and wrote analytical SQL queries:

- Attrition Rate by Department, Role, and City
- Top 10 highest-paid employees
- New Calculated columns
- Correlation check of average salary by employees who left Vs Stayed
- Subqueries:
    - Attrition Rate Higher than overall Attrition Rate by Department
    - Employees with Salary Higher than Average Salary
    - Job Roles with Salary higher than overall average salary
    - Employee with Salary higher than their departments average salary

This SQL analysis demonstrated my ability to answer business HR questions using structured queries.

8. Power BI Dashboard Development -

I built an interactive HR Dashboard containing:

KPIs

- Total Employees
- Attrition Count
- Attrition Rate
- Average Age
- Active Employees

Visuals

- Attrition by Department & Education
- Attrition by Job Role
- Attrition by Age Group & Marital Status
- Salary Distribution
- Attrition by Performance Improvement Plan
- Attrition by City

Filters

- Department
- Age Group
- City
- Attrition

This dashboard converts raw data into easily digestible insights for HR managers.

9. Key Business Insights -

1. Support Lead role showed the highest attrition (22%).
2. Adults and employees with <3 years of experience left more frequently.
3. Salary imbalance exists between some job roles within the same department.
4. Training frequency has a strong relationship with performance improvement.
5. Certain managers had significantly higher attrition under them.

10. Project Outcome -

This project demonstrates my capability to perform complete end-to-end data analytics:

- Data Cleaning
- Python EDA
- SQL Analytics
- Power BI Dashboard
- Business Insights
- Excel Analytics
- Final Reporting

I gained strong experience in converting raw HR data into actionable findings that can support real organizational decision-making.

12. What I Added to GitHub -

- Jupyter Notebooks (Python Data Cleaning, EDA)
- SQL Queries File
- Power BI Analytics with Dashboard
- Excel Analysis with Dashboard
- Final Report
- Visuals images
- ReadMe with project explanation and dashboard screenshots

This project is perfect for showcasing both technical and analytical skills to recruiters.