Employee Performance and Attrition

Project Overview: The Employee Performance & Attrition Analysis project aims to leverage Data Warehouse, Python, SQL, and Power BI to analyze workforce trends, identify key drivers of employee attrition, and assess performance metrics. The insights derived from this analysis will help organizations optimize employee retention strategies and improve productivity.

Objectives

- Identify factors influencing employee attrition, including salary, work-life balance, and tenure.
- Helping Management to bring down the attrition rate by finding the root causes.
- Create interactive reports in Power BI for insights.
- Improve HR decision-making using data-driven insights.

Problem Statement: Organizations face significant challenges in managing employee performance and retention. High employee attrition not only leads to increased hiring costs but also results in the loss of institutional knowledge, reduced productivity, and potential disruptions in business operations. Many organizations struggle to identify the root causes of attrition and performance issues due to the lack of comprehensive data analysis.

Key concerns include:

- Inability to accurately assess employee performance trends over time.
- Difficulty in identifying critical factors that contribute to voluntary and involuntary attrition.
- Inefficient HR decision-making due to the absence of deeper insights.
- Challenges in correlating compensation, job satisfaction, and work-life balance with employee retention.

Technology Stack

- **Data Warehouse:** Centralized storage of structured data from various sources.
- **SQL:** Data extraction, transformation, and loading (ETL) processes.
- **Python:** Data preprocessing, statistical analysis.
- **Power BI:** Visualization and reporting.

Data Sources:

The project utilizes three key datasets:

Dataset Name	**Description**
Employee Data	Contains demographic and job-related details about employees.
Attrition Data	Tracks employee attrition and exit interview scores.
Employee Performance Data	Includes performance ratings, promotions, and job satisfaction.

Employee Data

Column Name	**Data Type**	**Description**
Employee_ID	int	Unique identifier for each employee.
Age	int	Age of the employee.
first_name	str	First name of the employee.
last_name	str	Last name of the employee.
gender	str	Stated gender of the employee.
Department	str	Department where the employee works.
Job_Role	str	Specific role/title of the employee.
Education_Level	str	Highest education qualification.
Marital_Status	str	Marital status of the employee.
Job_Tenure	int	Number of years the employee has worked in the organization.
Distance_From_Home	int	Distance (in km) from employee's residence to the workplace.

Attrition Data

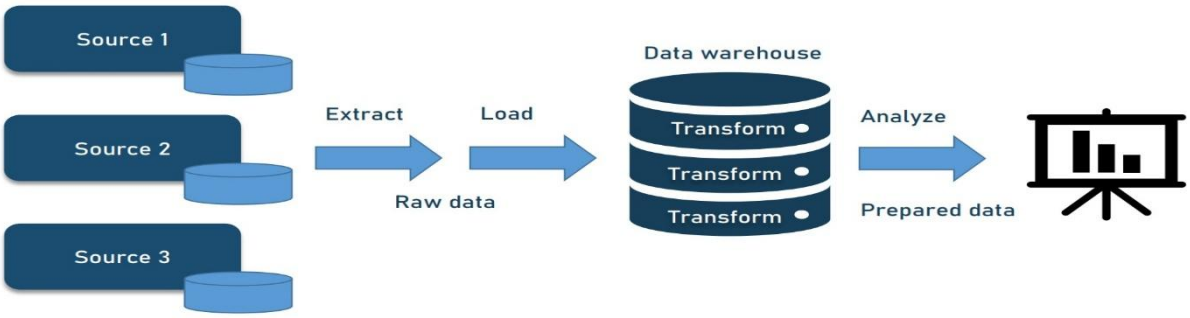
Column Name	Data Type	Description
employee_ID	int	Employee identifier.

attrition	bool	Whether the employee has left the company (True = Yes, False = No).
Exit_Interview_Score	int	Score given during the exit interview.

Employee Performance Data

Column Name	Data Type	Description
Employee_ID	int	Employee identifier.
Performance_Rating	int	Rating of employee performance.
Last_Promotion_Year	int	Year when the employee was last promoted.
Training_Hours	int	Number of hours spent in training programs.
Work_Life_Balance	int	Work-life balance rating.
Job_Satisfaction	int	Employee's satisfaction with their job.

1. Data Warehousing:



- **Design a Data Model:**
  - **Star Schema:** Centralize your data around a fact table capturing employee performance metrics, linked to dimension tables like Time, Department, Role, and Personal Attributes.
  - **Data Sources:** Aggregate data from HR databases, performance reviews, and even anonymous employee surveys.
- **ETL Process with SQL & Python:**
  - **Extract:** Use SQL to pull data from relational databases.

- **Transform:** Employ Python (Pandas, NumPy) to clean, normalize, and transform data—handling missing values, encoding categorical variables, and scaling numerical features.
- **Load:** Populate your data warehouse, ensuring optimized query performance for analytics.

2. SQL:

SQL Scope in the Project SQL plays a crucial role in the Employee Performance & Attrition Analysis project by enabling efficient data management, transformation, and analysis. Its key responsibilities include:

1. **Data Extraction:**
   - Querying HR databases to retrieve employee performance, payroll, and attrition data.
   - Fetching survey results and attendance logs for further analysis.
2. **Data Transformation & Cleaning:**
   - Handling missing values, duplicates, and inconsistent data formats.
   - Standardizing salary structures, job roles, and performance ratings.
3. **ETL (Extract, Transform, Load) Processes:**
   - Loading processed data into the Data Warehouse for centralized storage.
   - Aggregating data for trend analysis and reporting.
4. **Data Analysis & Querying:**
   - Generating insights on attrition trends, department-wise performance, and salary distributions.
   - Identifying correlations between job satisfaction scores and attrition rates.

Key Performance Indicators (KPIs):

- **Employee Satisfaction Score**: Aggregate of employee feedback on work conditions, leadership, and job satisfaction.
- **Average Tenure**: The average length of time employees remain with the company.
- **Attrition Rate**: Percentage of employees leaving over a given period.
- **Performance Rating Distribution**: Categorization of employees based on performance scores.
- **Department-wise Attrition Trends**: Comparison of attrition across different departments.

- **Exit Interview Sentiment Analysis**: Text analysis of exit interviews to identify common reasons for attrition.

3. **Python:**

Python Scope in the Project Python is a crucial component of the Employee Performance & Attrition Analysis project, enabling advanced data processing, statistical analysis. Its key responsibilities include:

1. **Data Preprocessing & Cleaning:**
    - Handling missing data, outlier detection, and normalization.
    - Standardizing categorical variables and encoding them for analysis.
2. **Exploratory Data Analysis (EDA):**
    - Identifying patterns in employee performance, satisfaction, and attrition.
    - Generating visualizations and summary statistics to detect trends.
3. **Feature Engineering:**
    - Creating new attributes from existing employee data to enhance modeling.
    - Calculating tenure, satisfaction scores, and performance growth metrics.
4. **Integration with SQL & Data Warehouse:**
    - Extracting structured datasets from SQL databases for modeling.
    - Storing processed and scored data back into the Data Warehouse.
5. **Automated Reporting & Monitoring:**
    - Generating automated reports with Python libraries (Pandas, Matplotlib, Seaborn).
    - Scheduling periodic updates and retraining models as new data becomes available.

4. Power BI Visualization: Telling the Story

Power BI Scope in the Project Power BI serves as the primary visualization tool for this project, enabling real-time monitoring and interactive analysis of employee performance and attrition trends. Its key responsibilities include:

1. **Data Integration & Processing:**
    - Connect directly to the Data Warehouse for real-time updates.
    - Utilizing SQL queries to fetch, filter, and aggregate data for dashboards.

2. **Dashboard & Report Development:**
    - Designing interactive Reports to visualize employee performance trends.
    - Creating reports to analyze attrition at department and individual levels.
3. **KPI Tracking & Monitoring:**
    - Displaying key metrics like attrition rate, satisfaction scores, and performance trends.
    - **Key Performance Indicators (KPIs)**
        1. Attrition Rate
        2. Retention Rate
        3. Average Tenure
        4. Department-wise Employee Score
        5. Average Performance Rating
        6. Average Exit Interview Satisfaction Score
        7. Department wise Attrition Rate
    - Implementing dynamic filters and slicers for detailed analysis.

Expected Outcomes

- A clear understanding of factors affecting employee performance and attrition.
- Actionable insights for HR teams to enhance employee retention.
- A user-friendly dashboard for continuous monitoring and decision-making.