



# Project Report

## Data Analysis



**Project Title : HR Analytics - Predict Employee Attrition**

### Introduction :

In today's data-driven world, organizations rely heavily on data analytics to gain insights and make informed decisions. Employee-related data plays a vital role in understanding workforce trends, performance, diversity, and overall organizational health. This project leverages an employee dataset to perform data cleaning, analysis, visualization, and dashboard building. By using Python for preprocessing and visualization and Power BI for creating interactive dashboards, the project aims to provide clear insights into employee demographics, performance, and other critical HR metrics.

### Abstract:

The project focuses on analyzing employee data to uncover meaningful patterns and trends. Using **Python**, the dataset undergoes thorough cleaning, transformation, and exploratory data analysis (EDA). Key insights are visualized through Python libraries such as Matplotlib and Seaborn. To enhance interactivity and decision-making, a **Power BI dashboard** is developed, allowing users to filter, drill down, and explore employee-related metrics dynamically. The combined use of Python and Power BI provides both in-depth analysis and a user-friendly interface for business stakeholders.

### Tools Used :

#### 1. Python

- Data Cleaning (handling missing values, correcting data types, removing duplicates).
- Exploratory Data Analysis (EDA) and Visualization using Pandas, Matplotlib, Seaborn.

## **2. Power BI**

- Creation of an **interactive dashboard**.
- Visual representation of KPIs (employee count, attrition, department distribution, etc.).
- Filters and slicers for easy navigation and data-driven decision-making.

## **Steps Involved in Building the Project :**

### **1. Data Collection**

- Import the Employee dataset (Employee Sample Data.xlsx).

### **2. Data Cleaning (Python)**

- Remove duplicates.
- Handle missing values.
- Standardize categorical fields (e.g., department names, job roles).
- Correct data types (dates, numbers, categories).

### **3. Exploratory Data Analysis (Python)**

- Generate summary statistics.
- Visualize distributions (e.g., age, salary, years at company).
- Correlation analysis (e.g., factors affecting attrition).
- Plot graphs for trends (department size, gender ratio, job satisfaction, etc.).

### **4. Data Preparation for Dashboard**

- Export cleaned dataset to .csv or .xlsx format for Power BI integration.

### **5. Dashboard Development (Power BI)**

- Import cleaned dataset into Power BI.
- Design interactive visuals (bar charts, pie charts, KPI cards, slicers).
- Create dashboards highlighting key employee insights such as:
  - Attrition trends

- Average salary by department
- Employee distribution by gender, age, and role
- Years of service vs. performance

## **6. Testing and Refinement**

- Validate accuracy of data and visuals.
- Ensure interactivity (slicers, filters) works correctly.

## **Conclusion:**

This project demonstrates how Python and Power BI can be integrated to provide end-to-end data analysis and visualization. Python ensures the dataset is properly cleaned, analyzed, and explored, while Power BI translates these insights into interactive dashboards accessible to decision-makers. The approach improves workforce analysis, aids HR departments in identifying patterns such as attrition risk and departmental imbalances, and supports strategic decision-making with data-backed insights.

**Submitted by**

**B.Sarumathi**