HR Attrition Data Analysis Project

Introduction

The HR Attrition Data Analysis Project aims to analyze employee attrition trends, identify key factors contributing to employee turnover, and provide actionable insights for HR departments. The project utilizes a combination of data preprocessing, analysis, and visualization techniques to present a comprehensive view of workforce dynamics within the organization.

Abstract

This project focuses on understanding the patterns of employee attrition using the HR dataset. The raw dataset ('WA_Fn-UseC_-HR-Employee-Attrition.csv') was cleaned and processed in Python using Jupyter Notebook. The cleaned data ('Cleaned_HR_Attrition.xlsx') was further transformed in Power BI to create an interactive dashboard. Key metrics such as average monthly income, job satisfaction, performance rating, and attrition rate were analyzed to support HR decision-making and reduce turnover.

Tools Used

1. Python (Jupyter Notebook) – for data cleaning, transformation, and preprocessing.
2. Microsoft Excel – for storing the cleaned dataset.
3. Microsoft Power BI – for visualization and dashboard creation.
4. Pandas, NumPy – for data manipulation and analysis.
5. ReportLab – for generating the final report.

Steps Involved in Building the Project

1. **Data Collection:** The raw dataset was imported from the HR Attrition CSV file.
2. **Data Cleaning:** Handled missing values, standardized column names, and converted categorical data into numerical format using Python.
3. **Feature Engineering:** Created new columns such as Age Group, Salary Band, Tenure Band, Risk Score, and Risk Category.
4. **Data Transformation:** Unnecessary columns (e.g., OverTime_Yes, Department_Sales, JobRole dummies) were dropped in Power BI to optimize the dataset.
5. **Dashboard Creation:** Power BI measures were defined to calculate key KPIs like Total Employees, Average Monthly Income, Attrition Rate, and more.
6. **Visualization:** Developed charts and dashboards showing employee distribution by age, income, job level, and attrition risk.

Conclusion

The HR Attrition Data Analysis project successfully demonstrates how data analytics can assist HR professionals in understanding employee retention patterns. By visualizing and interpreting critical metrics, the organization can implement strategies to improve employee satisfaction, enhance performance, and reduce attrition rates. This project showcases the integration of data cleaning, transformation, and visualization techniques into a cohesive analytical workflow.