# README - HR Analytics Project

Project Title: Predicting Employee Attrition using Data Analytics

Submitted By: AFRAH ASHRAF

Date: June 2025

## Project Summary

This project aims to analyze employee data to identify key factors contributing to attrition and build predictive models to estimate future attrition risk. The goal is to help organizations retain employees by understanding and acting on these insights.

## Tools & Technologies Used

- Python (Pandas, Seaborn, Scikit-learn, SHAP)

- Jupyter Notebook

- Power BI

- MS Word/PDF for documentation

## Folder Structure

HR_Attrition_Project/
├── Complete_Attrition_Project_Report.pdf # Final merged report with cover, EDA, models, visuals
├── HR_Attrition_Analysis.ipynb # Python analysis notebook
├── HR_Attrition_Dashboard.pbix # Power BI dashboard file
├── Attrition_Prevention_Strategy.pdf # Strategy recommendations based on model insights

├── HR_Employee_Attrition.csv # Dataset used
├── README.docx # Project instructions

## How to Run This Project

1. Open `HR_Attrition_Analysis.ipynb` in Jupyter Notebook or VS Code.

2. Ensure required libraries are installed (Pandas, Seaborn, Sklearn, SHAP).

3. Run all cells to view data analysis, model building, and SHAP explainability.

4. Open `HR_Attrition_Dashboard.pbix` in Power BI to view and explore the dashboard.

## 📌 Important Notes

- The dataset is clean and contains 1470 rows with 35 features.

- Two models are implemented: Logistic Regression and Decision Tree.

- SHAP values are used to interpret the Logistic Regression model.

- Prevention strategy recommendations are based on EDA and model output.