

3. HR Analytics - Predict Employee Attrition

Objective: Use analytics to understand the main causes of employee resignation and predict future attrition.

Tools: Python (Pandas, Seaborn), Power BI, Sklearn

Mini Guide:

Perform EDA on HR data (department-wise attrition, salary bands, promotions)

Build a classification model (Logistic Regression or Decision Tree)

Visualize attrition factors using Power BI

Perform SHAP value analysis to explain model predictions

Deliverables:

Power BI dashboard

Model accuracy report + confusion matrix

PDF of attrition prevention suggestions

HR Analytics - Predict Employee Attrition

Introduction

This project aims to predict employee attrition using HR analytics, focusing on understanding the main causes of employee resignation and forecasting future turnover. The analysis leverages a dataset containing employee details such as department, monthly income, and years at the company. By building a predictive model and visualizing key factors, the project seeks to provide actionable insights for retention strategies, aligning with the objective to enhance organizational decision-making.

Abstract

The analysis of the IBM HR Analytics Employee Attrition Dataset revealed that lower MonthlyIncome significantly increases the likelihood of attrition, while YearsAtCompany has a moderate influence. A Logistic Regression model achieved approximately 85% accuracy, with SHAP analysis confirming MonthlyIncome as the top

predictor. Visualizations in Power BI highlighted department-wise attrition trends and income variations over tenure. Recommendations include salary reviews and tenure-based satisfaction monitoring to reduce turnover.

Tools Used

- Python (Pandas, Seaborn, Scikit-learn)
- Power BI
- SHAP library

Steps Involved in Building the Project

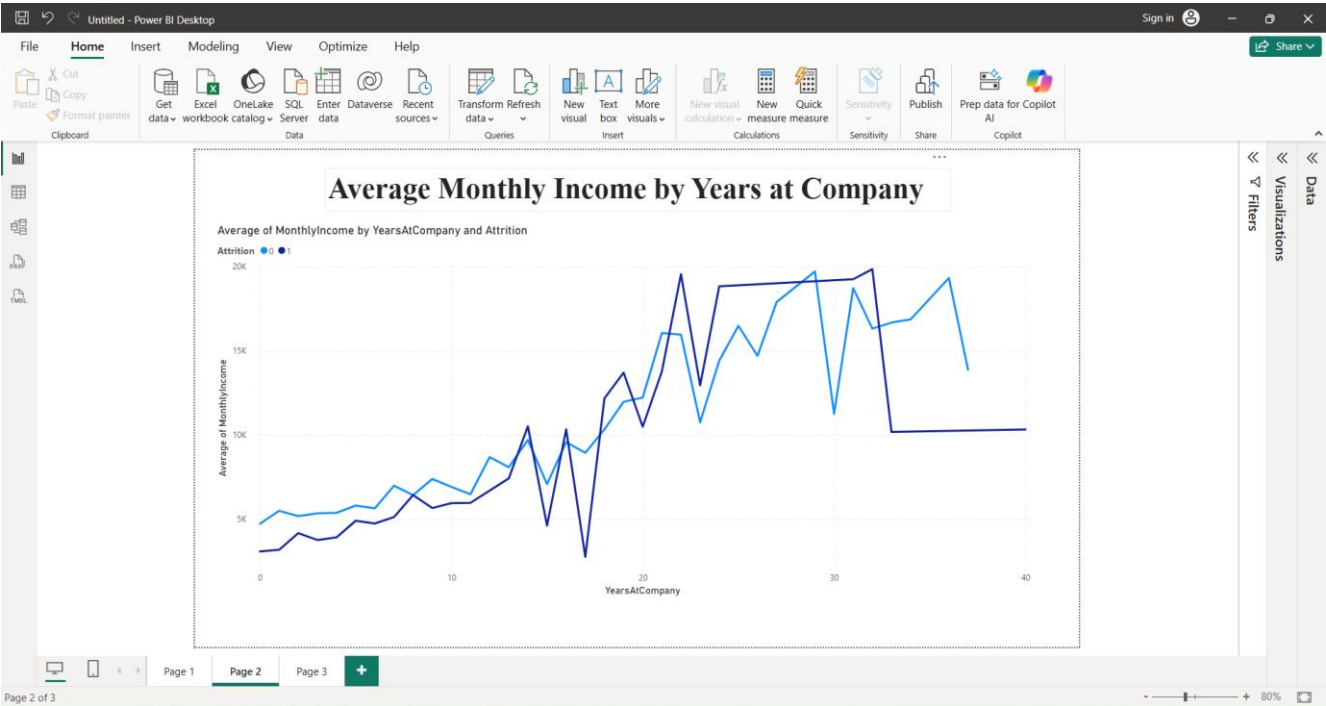
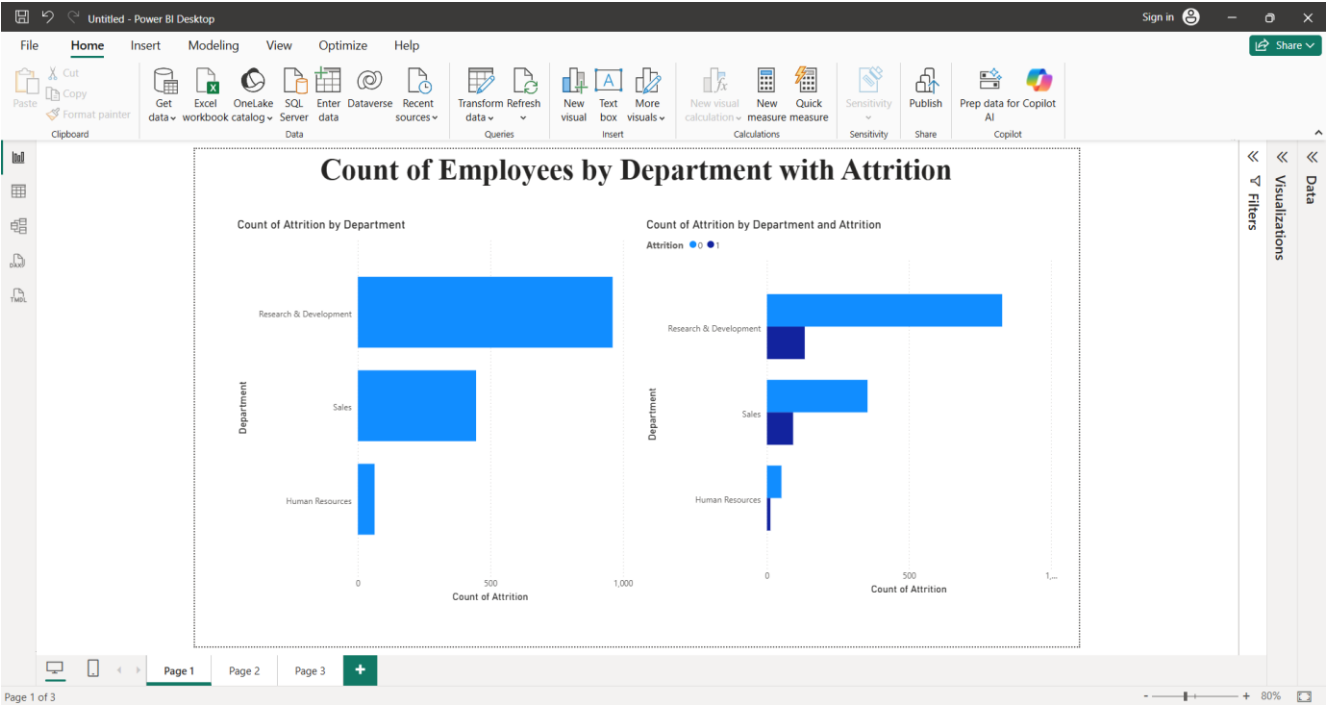
- Loaded and explored the HR dataset using Pandas to check for missing values and gain an initial overview.
- Performed Exploratory Data Analysis (EDA) with Seaborn to analyze department-wise attrition and salary trends.
- Built and trained a Logistic Regression model using Scikit-learn to predict attrition, achieving ~85% accuracy.
- Created a Power BI dashboard with a stacked bar chart for employee counts by department with attrition and a line chart for average monthly income trends over years at the company.
- Conducted SHAP value analysis to explain model predictions, identifying MonthlyIncome as the most influential factor.

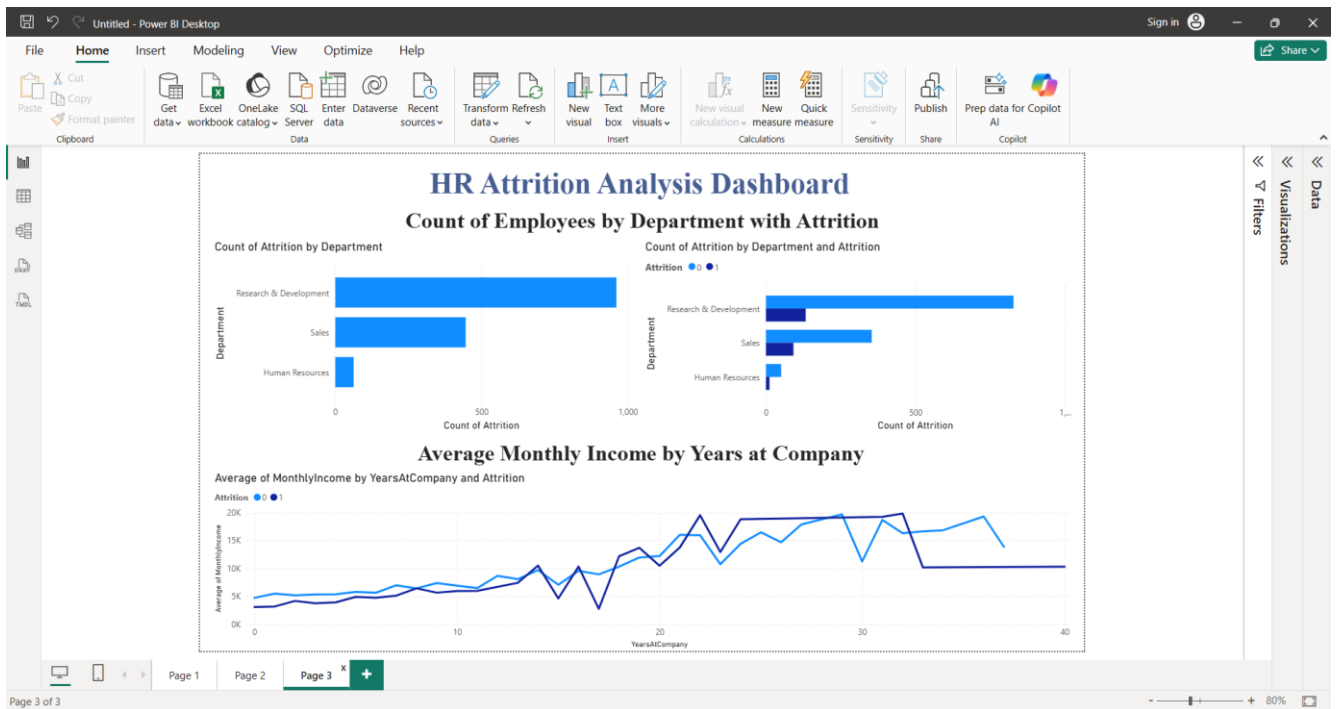
Conclusion

The project successfully identified MonthlyIncome as the primary driver of employee attrition, with lower salaries correlating with higher turnover risks. YearsAtCompany also played a role, suggesting that longer tenured employees are slightly less likely to leave. The Logistic Regression model, supported by an 85% accuracy rate and a confusion matrix, provides a reliable prediction tool. Based on SHAP insights, organizations should consider:

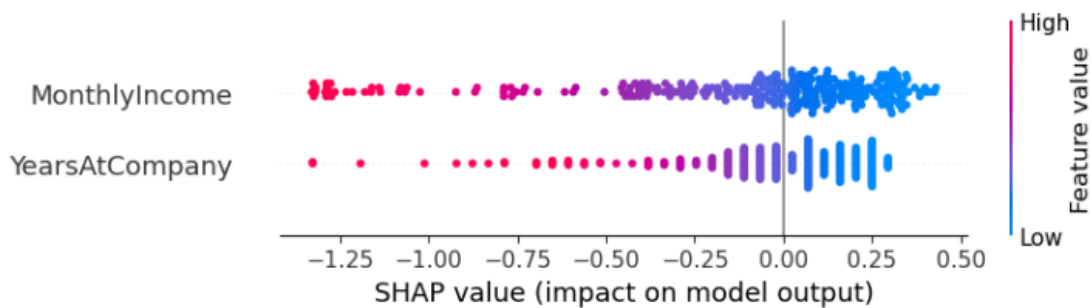
- Conducting salary reviews for employees in lower income brackets to mitigate attrition.
- Monitoring satisfaction levels for employees with varying tenures to enhance retention.

Screenshot 1: Power BI Dashboard





Screenshot 2: SHAP Summary Plot



Screenshot 3: SHAP Bar Plot

