

HR Attrition Prediction – Insights & Analysis Report

Dataset Name: hr_employee_attrition.csv

Author: Vaishnavi H

Date: October 2025

1. Project Overview

The **HR Attrition Prediction Project** aims to analyze employee attrition patterns and identify the major factors contributing to employee turnover.

By combining descriptive analytics and predictive modeling, the project provides valuable insights for HR decision-making, helping organizations improve employee retention and satisfaction.

The project uses an HR dataset containing employee demographics, job-related features, satisfaction levels, and attrition status.

The analysis was performed using Python (for data cleaning, preprocessing, and EDA) and Tableau (for dashboard creation and visualization).

2. Objectives

The main objectives of this project were to:

- Analyze employee attrition trends by department, job role, gender, and age group.
- Identify high-risk groups with elevated attrition rates.
- Build a simple machine learning model to predict attrition likelihood.
- Provide actionable insights through interactive Tableau dashboards.

3. Data Summary

Attribute Type	Examples	Description
Employee Info	Employee ID, Department, Job Role	Organizational details

Demographics	Age, Gender, Education, Marital Status	Basic employee data
Job & Performance	JobSatisfaction, JobLevel, MonthlyIncome	Work-related variables
Attrition Info	Attrition (Yes/No)	Target variable for prediction

Dataset size: ~1,400 records

Target column: Attrition

Positive class: Employees who left the company (Attrition = "Yes")

4. Data Preparation & Cleaning

The following steps were performed during data cleaning:

1. Handled missing or inconsistent values.
2. Standardized categorical variables (e.g., Gender → "Male" / "Female").
3. Converted the Attrition column into numeric form (Attrition_Flag: Yes = 1, No = 0).
4. Created new features such as Age Groups for segmentation.
5. Exported clean datasets for Tableau dashboards:
 - o attrition_by_department.csv
 - o attrition_by_jobrole.csv
 - o attrition_by_gender.csv
 - o attrition_by_age.csv

5.Exploratory Data Analysis (Python)

The exploratory phase focused on understanding attrition distribution and correlations among features.

Key Python operations included:

- Distribution plots for Age, JobSatisfaction, and MonthlyIncome.

- Bar charts comparing attrition rates across departments and job roles.
- Correlation matrix to identify key influencing features.

6.Dashboard Insights (Tableau)

Two main Tableau dashboards were designed:

Dashboard 1: HR Attrition Overview

Visuals:

- Gender-wise Attrition (Pie Chart)
- Department-wise Attrition (Bar Chart)

Key Insights:

- The overall attrition rate stands at approximately X%.
- **Sales and HR departments** show the highest attrition rates.
- **Male employees** have a slightly higher attrition rate than females.

Dashboard 2: Attrition Drivers

Visuals:

- Job Role vs Attrition (Bar Chart)
- Age Group vs Attrition (Line Chart)
- Job Satisfaction vs Attrition (Stacked Bar)

Key Insights:

- **Younger employees (26-35 age group)** are more likely to leave the organization.
- **Sales Executives** and **Laboratory Technicians** exhibit higher turnover rates.

- Lower **Job Satisfaction** scores correlate strongly with higher attrition rates.

7.Predictive Modeling Summary

A Random Forest Classifier was trained to predict the probability of employee attrition.

Metric Value

Accuracy ~84%

Precision 0.80

Recall 0.78

F1-score 0.79

Top contributing features:

OverTime, Age, MonthlyIncome, JobLevel, JobSatisfaction

8.Business Recommendations

Based on the analysis:

1. Employee Engagement:

Improve engagement programs for Sales and R&D employees to reduce turnover.

2. Compensation Review:

Review pay scale for early-career and younger employees who show higher attrition.

3. Workload Management:

Manage overtime hours; frequent overtime strongly correlates with attrition.

4. Career Growth:

Enhance promotion and learning opportunities in departments with low satisfaction scores.

5. Data-driven Retention Strategy:

Implement predictive monitoring to identify and support at-risk employees early.

9. Tools & Technologies Used

Category	Tools
Data Storage	CSV, Pandas DataFrame
Data Cleaning & Analysis	Python, Pandas, NumPy
Visualization	Matplotlib, Seaborn, Tableau
Modeling	Scikit-learn (Random Forest, Logistic Regression)
Reporting	Tableau Dashboard, PDF Report

10. Key Deliverables

Deliverable Location

Cleaned Dataset HR_Attrition_Prediction/data/

Python Notebooks HR_Attrition_Prediction/notebooks/

Tableau Dashboards HR_Attrition_Prediction/dashboard/

Deliverable	Location
Final Report	HR_Attrition_Prediction/reports/HR_Attrition_Insights_Report.pdf

11.Conclusion

The HR Attrition Analysis project successfully identified key drivers influencing employee turnover.

Using a combination of data analytics and predictive modeling, it provided actionable insights that can help HR departments improve retention and satisfaction.

The dashboards and model results serve as practical tools for ongoing monitoring of workforce stability and employee engagement.