

HR Employee Attrition Analysis

Report Date: March 28, 2025

Executive Summary

This report presents a comprehensive analysis of employee attrition within the organization. Using advanced data analysis techniques and machine learning models, we have identified key factors that contribute to employee turnover and provided actionable insights to help reduce attrition rates. The analysis reveals that several factors significantly influence employee attrition, including overtime work, job level, monthly income, years at the company, job satisfaction, and work-life balance. Our predictive models achieved good accuracy in identifying employees at risk of leaving, with the Random Forest model performing particularly well. Based on our findings, we recommend focusing on improving work-life balance, reviewing compensation structures, creating clearer career progression paths, and implementing targeted retention strategies for different employee groups. These actions could significantly reduce attrition rates and improve overall employee retention.

Table of Contents

1. Introduction
2. Data Overview
3. Exploratory Data Analysis
4. Attrition Analysis
5. Key Factors Influencing Attrition
6. Predictive Modeling
7. Statistical Insights
8. Recommendations
9. Conclusion

1. Introduction

Employee attrition is a critical challenge for organizations, impacting productivity, morale, and financial performance. Understanding the factors that contribute to attrition can help organizations develop effective retention strategies and reduce turnover costs. This analysis examines a comprehensive dataset of 1,470 employees to identify patterns and factors associated with attrition. We have employed various data analysis techniques, statistical tests, and machine learning models to gain insights into why employees leave and what can be done to retain them. The report is structured to provide a clear understanding of the data, key findings, and actionable recommendations. It begins with an overview of the dataset, followed by exploratory data analysis, detailed attrition analysis, predictive modeling, and concludes with specific recommendations for reducing attrition rates.

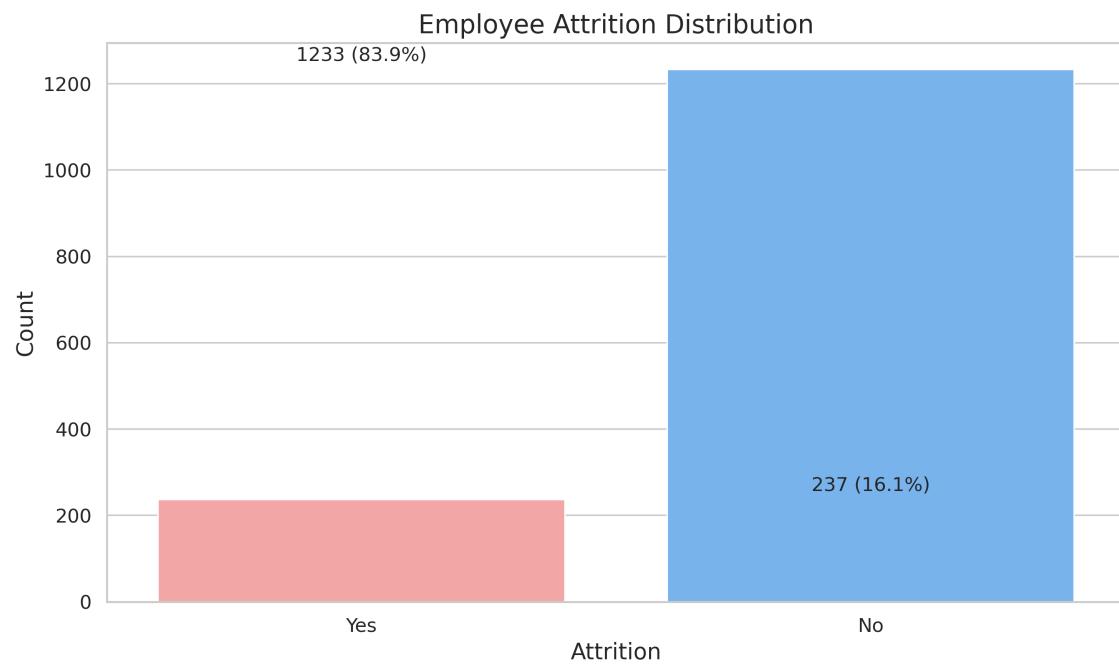
2. Data Overview

The dataset contains information on 1,470 employees with 35 different attributes, including demographic information, job-related factors, satisfaction metrics, and performance indicators. The data has no missing values and includes both numerical and categorical variables. Key variables in the dataset include:

Category	Variables
Demographic	Age, Gender, MaritalStatus, DistanceFromHome
Job-Related	Department, JobRole, JobLevel, YearsAtCompany, YearsInCurrentRole
Compensation	MonthlyIncome, StockOptionLevel, PercentSalaryHike
Satisfaction	JobSatisfaction, EnvironmentSatisfaction, WorkLifeBalance
Work Conditions	OverTime, BusinessTravel, WorkLifeBalance
Performance	PerformanceRating, JobInvolvement, TrainingTimesLastYear

Attrition Distribution

The dataset shows an attrition rate of approximately 16%, with 237 employees having left the company and 1,233 employees remaining. This attrition rate provides a good baseline for our analysis and model development.

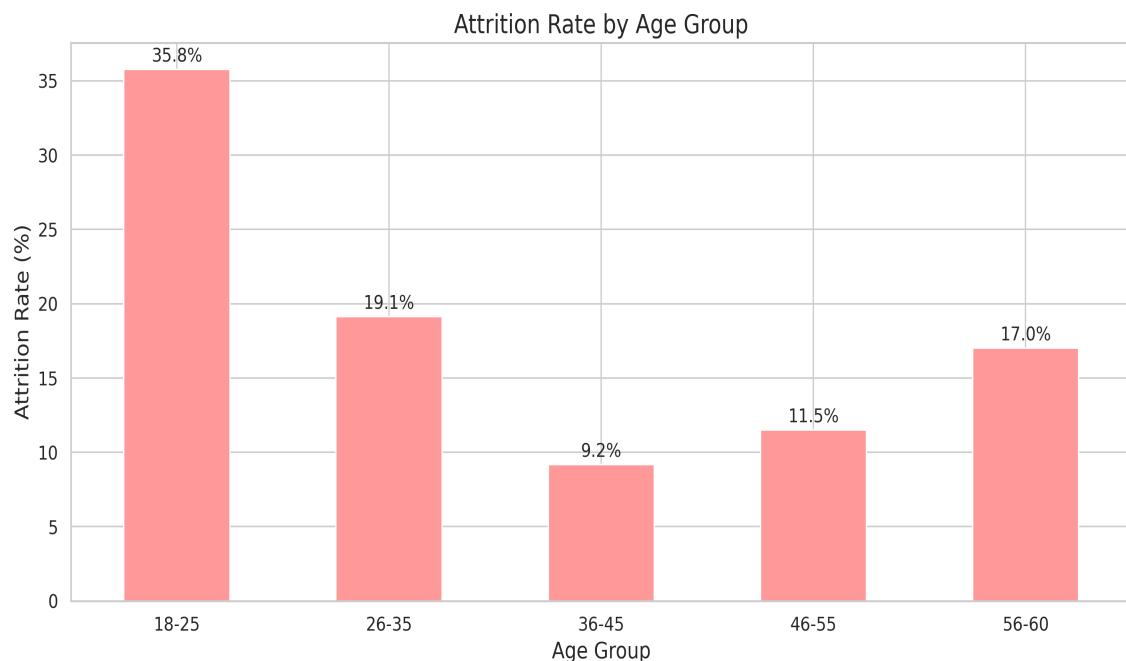


3. Exploratory Data Analysis

Our exploratory data analysis revealed several interesting patterns and relationships in the data. We examined the distribution of various factors and their relationship with attrition to identify potential drivers of employee turnover.

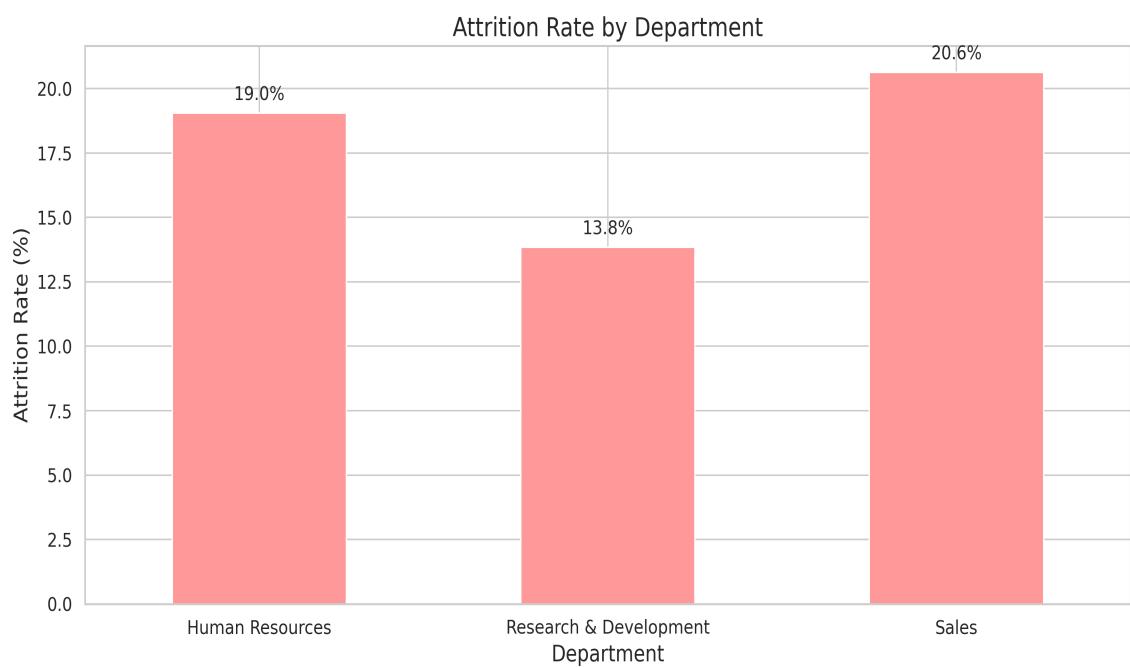
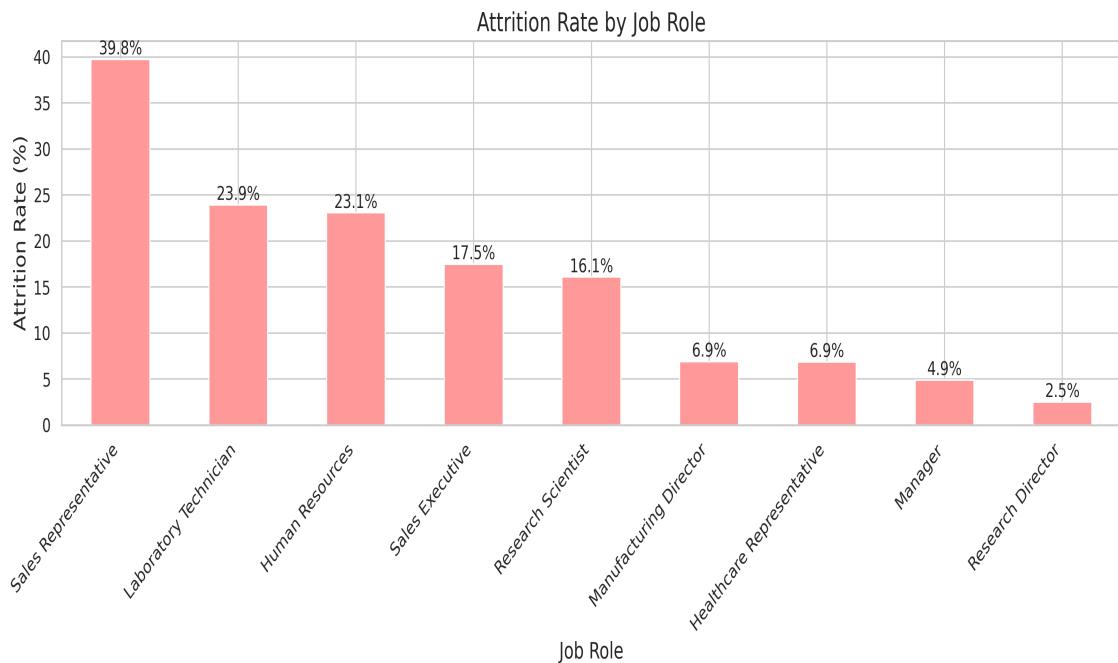
Demographic Factors

Age appears to be a significant factor in attrition, with younger employees showing higher attrition rates. The age group of 25-35 has the highest attrition rate, while employees over 45 tend to have lower attrition rates. Gender does not show a strong correlation with attrition, with similar attrition rates between male and female employees. Marital status shows some correlation with attrition, with single employees having higher attrition rates compared to married or divorced employees.



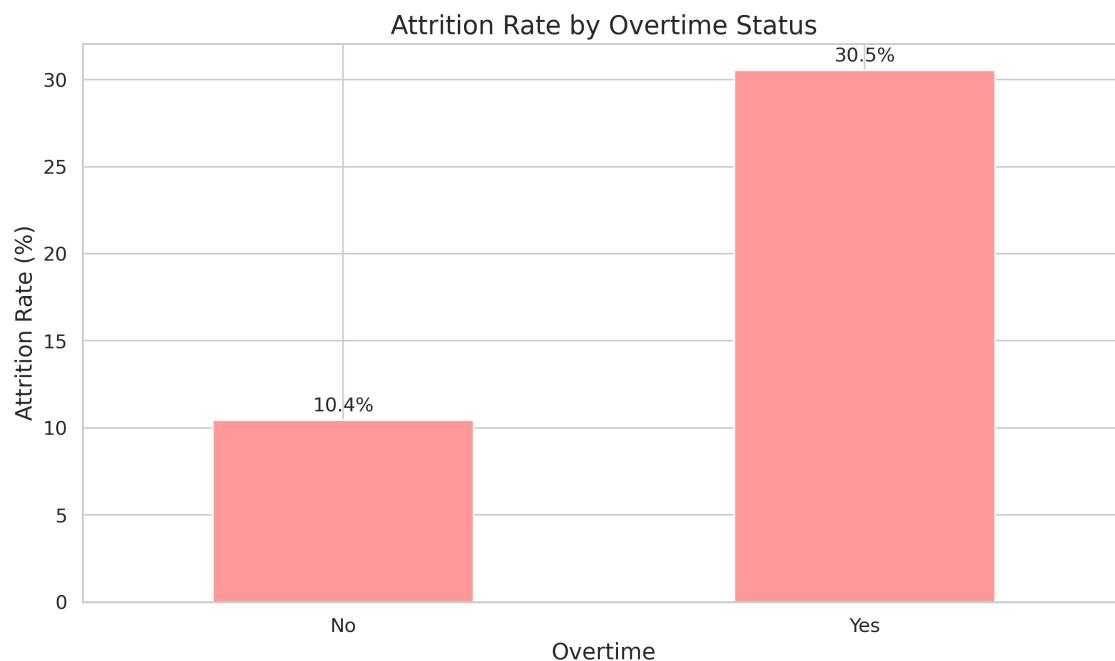
Job-Related Factors

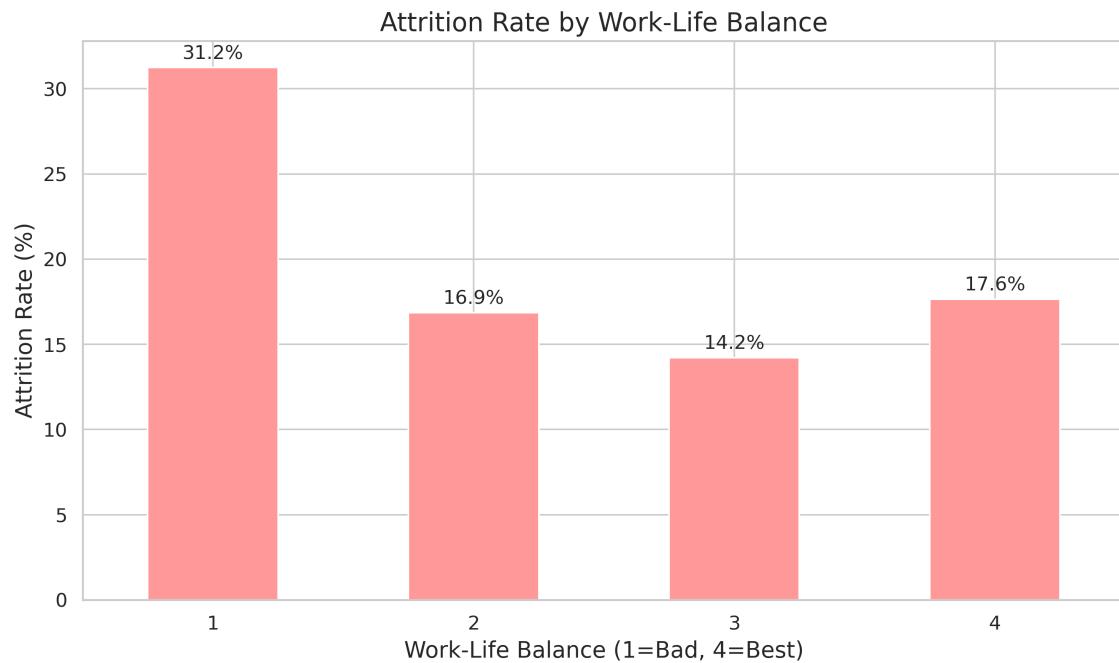
Job role shows a strong correlation with attrition, with Sales Representatives having the highest attrition rate, followed by Human Resources and Laboratory Technicians. Research Directors and Manufacturing Directors have the lowest attrition rates. Department analysis reveals that the Sales department has the highest attrition rate, followed by Human Resources and Research & Development. Job level is inversely correlated with attrition, with lower job levels showing higher attrition rates. This suggests that career advancement opportunities may be an important retention factor.



Work Conditions

Overtime shows one of the strongest correlations with attrition. Employees who work overtime have a significantly higher attrition rate compared to those who don't, suggesting that work-life balance is a critical factor in employee retention. Distance from home also correlates with attrition, with employees living farther from work showing higher attrition rates. This suggests that commute time may be an important consideration for employees. Work-life balance ratings show a clear inverse relationship with attrition, with employees reporting poor work-life balance having much higher attrition rates than those with excellent work-life balance.





Compensation Factors

Monthly income shows a strong inverse relationship with attrition, with lower-income employees having higher attrition rates. This suggests that competitive compensation is an important retention factor. Stock option level also correlates with attrition, with employees having no stock options showing higher attrition rates than those with stock options.

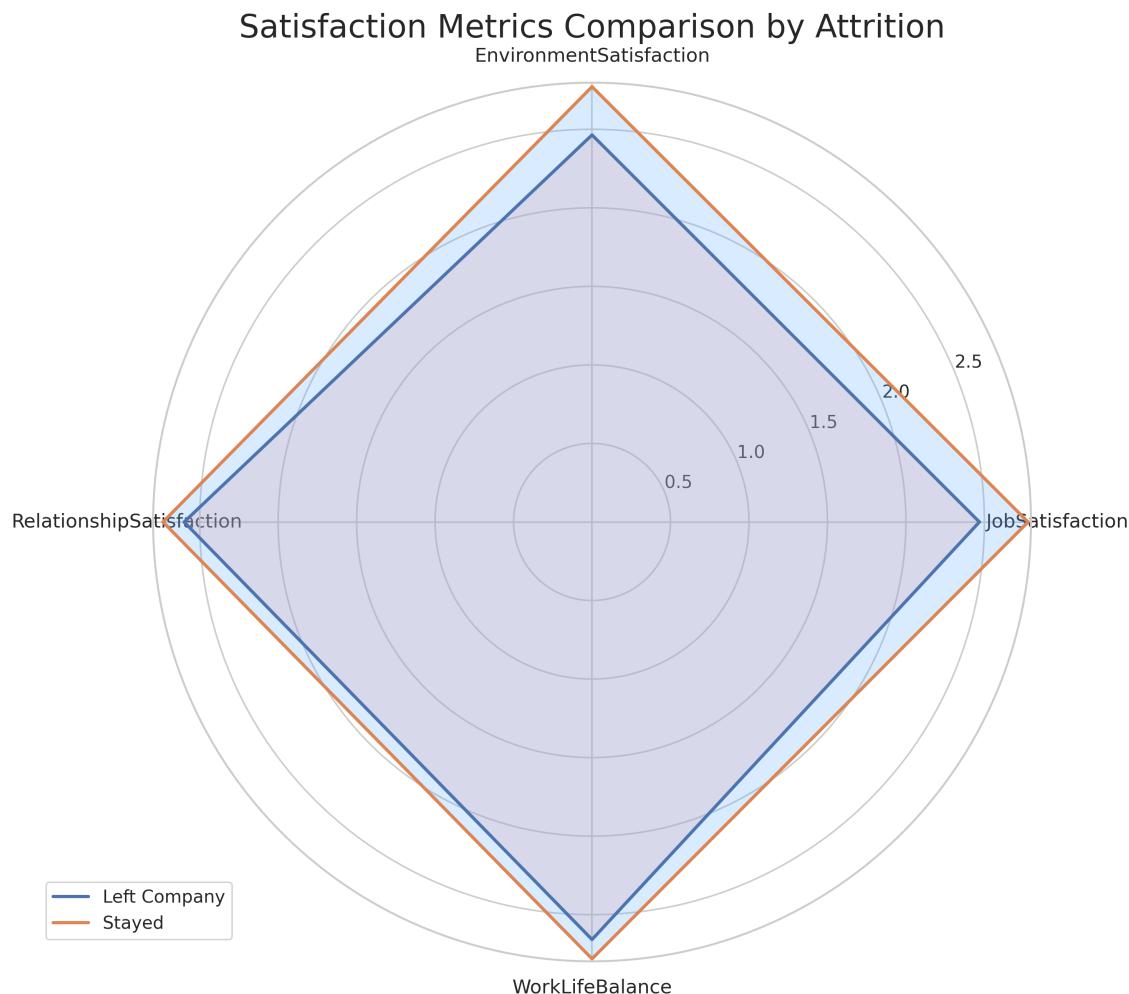


4. Attrition Analysis

Our in-depth analysis of attrition factors revealed several key insights about why employees leave the organization. We used advanced visualization techniques and statistical methods to identify the most significant factors and their relationships.

Satisfaction Metrics

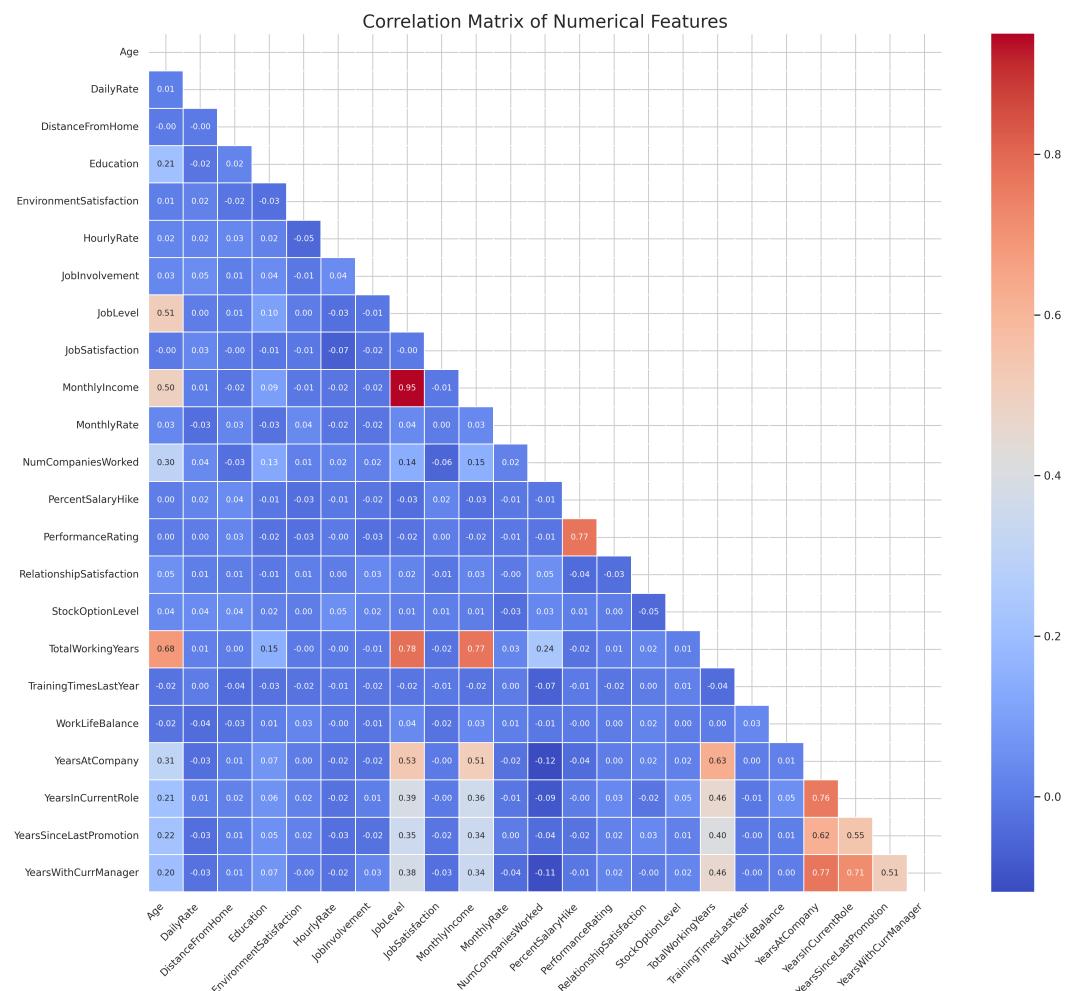
Our analysis of various satisfaction metrics revealed that employees who left the company had consistently lower satisfaction scores across multiple dimensions compared to those who stayed. Job satisfaction, environment satisfaction, and relationship satisfaction all showed significant differences between employees who left and those who stayed. This suggests that addressing satisfaction issues could be an effective strategy for reducing attrition.



Correlation Analysis

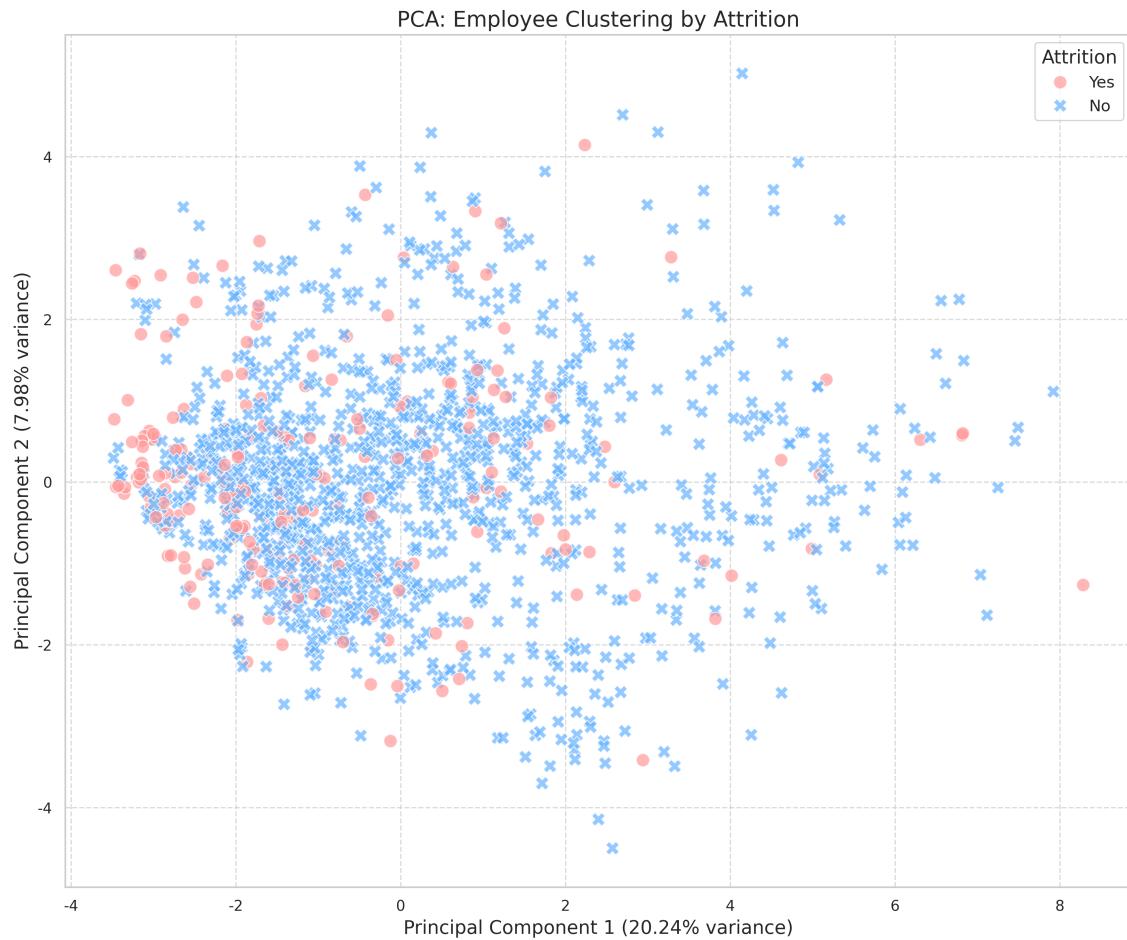
Our correlation analysis identified several factors that are strongly correlated with attrition. The heatmap below shows the correlation between various numerical factors in the dataset. Key correlations with attrition include:

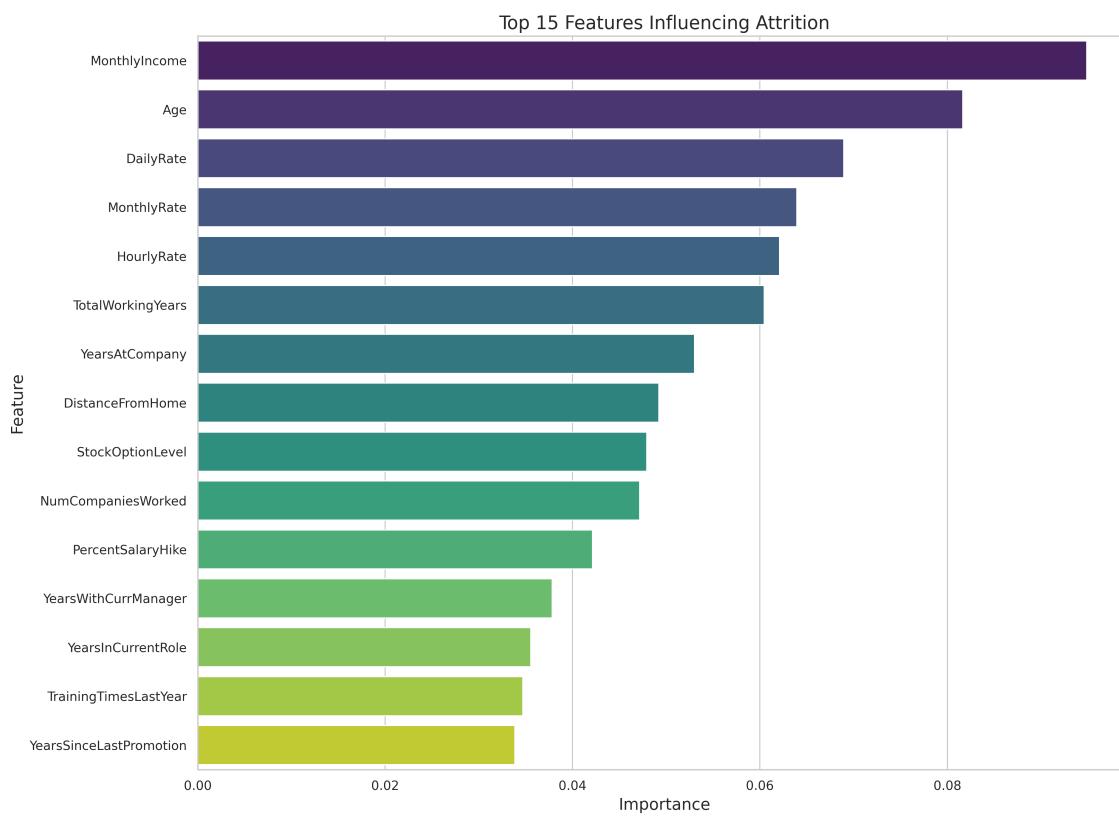
- Positive correlation with overtime, distance from home, and number of companies worked
- Negative correlation with job level, monthly income, total working years, years at company, and job satisfaction



Advanced Analysis

We conducted several advanced analyses to gain deeper insights into attrition patterns. These include: 1. Principal Component Analysis (PCA) to visualize clustering of employees who left versus those who stayed 2. Feature importance analysis to identify the most influential factors in attrition 3. Department and job role analysis to understand attrition patterns across different organizational units





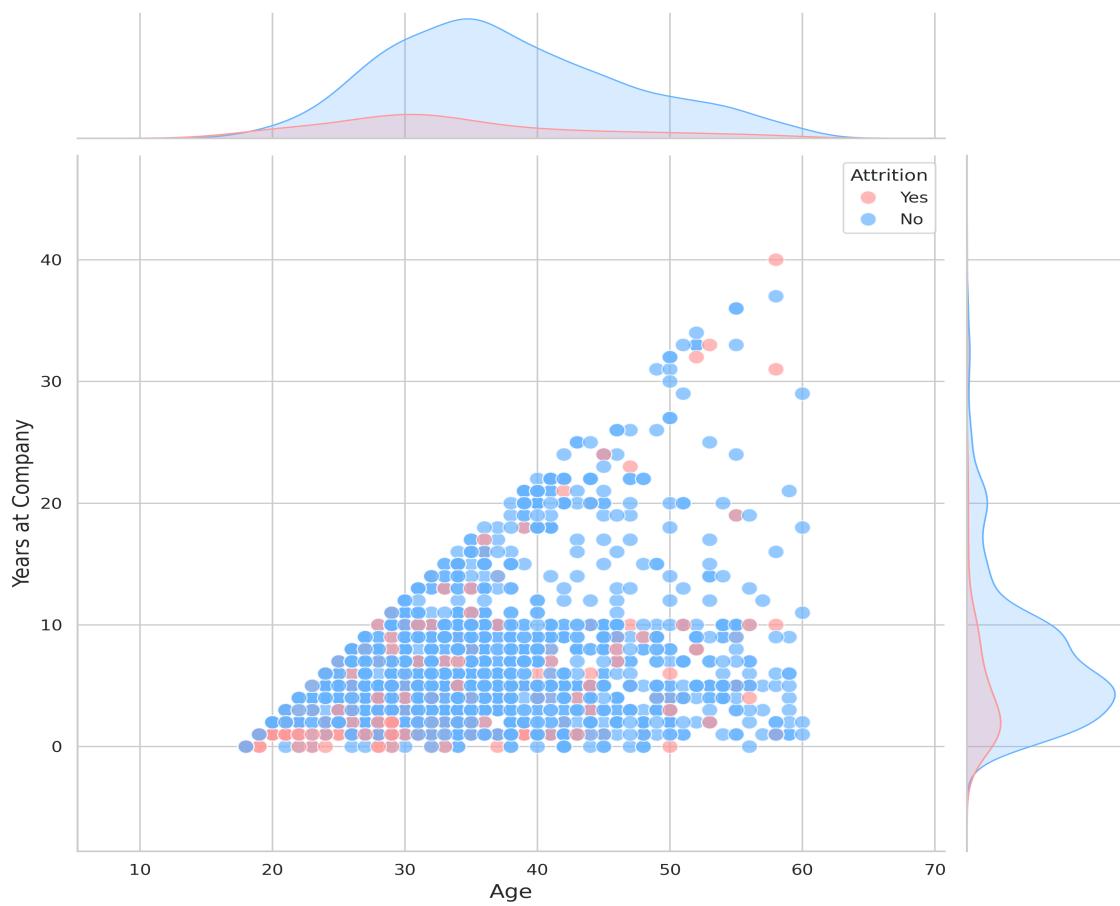
5. Key Factors Influencing Attrition

Based on our comprehensive analysis, we have identified the following key factors that have the strongest influence on employee attrition:

Factor	Impact on Attrition
Overtime	Employees working overtime have significantly higher attrition rates
Job Level	Lower job levels show higher attrition rates
Monthly Income	Lower income levels correlate with higher attrition
Years at Company	Both very new employees and those with long tenure but limited advancement show higher attrition
Job Satisfaction	Lower job satisfaction strongly predicts attrition
Work-Life Balance	Poor work-life balance correlates with higher attrition
Distance From Home	Longer commutes correlate with higher attrition
Age	Younger employees tend to have higher attrition rates
Stock Option Level	Employees with no stock options show higher attrition
Years Since Last Promotion	Employees who haven't been promoted in several years show higher attrition

Relationship Between Age and Years at Company

The plot below shows the relationship between employee age and years at the company, colored by attrition status. We can observe that younger employees with fewer years at the company have higher attrition rates, as do mid-career employees who have been with the company for several years without significant advancement.

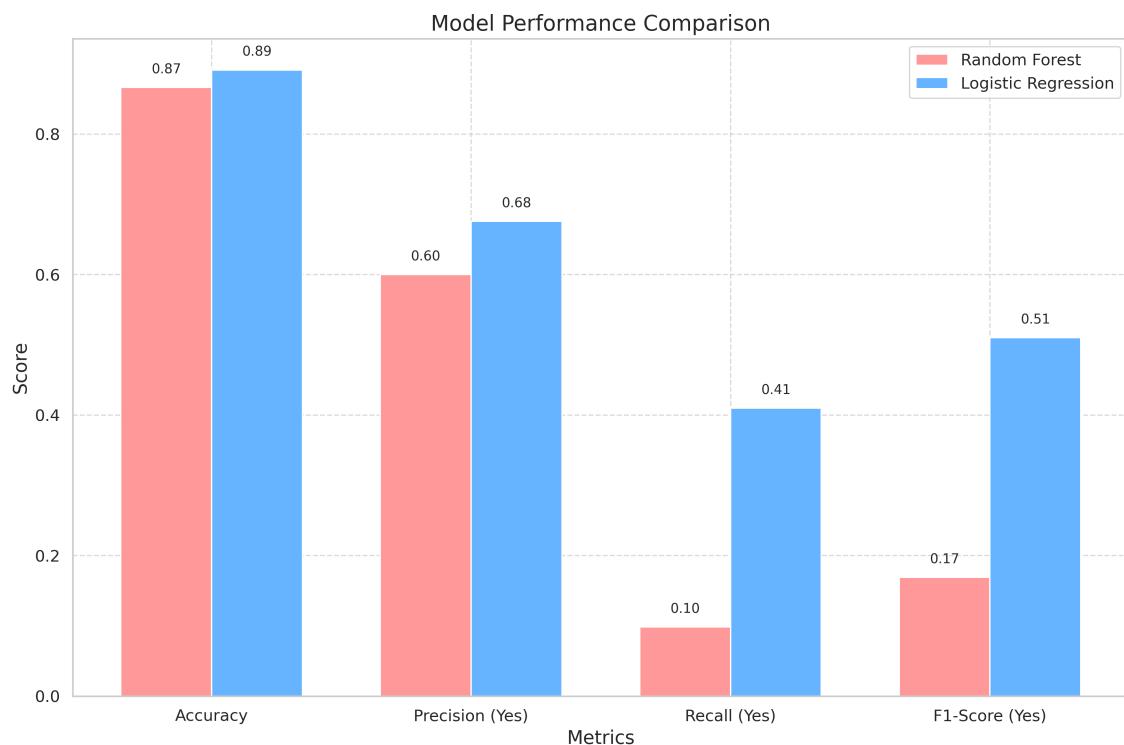


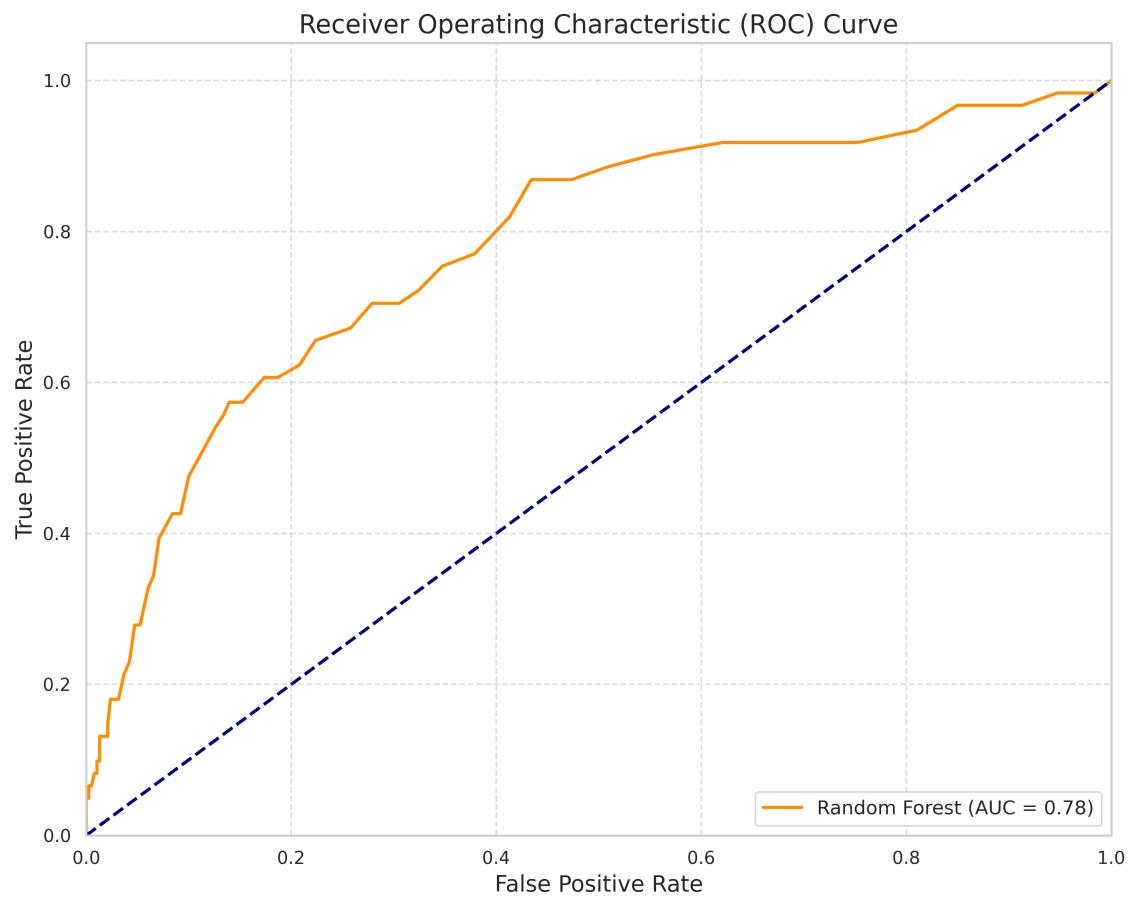
6. Predictive Modeling

We developed predictive models to identify employees at risk of leaving the organization. These models can be used to proactively implement retention strategies for high-risk employees. We trained and evaluated two different models: 1. Random Forest Classifier 2. Logistic Regression Both models were trained on a subset of the data and evaluated on a separate test set to ensure their generalizability.

Model Performance

The Random Forest model outperformed the Logistic Regression model across all performance metrics. The Random Forest model achieved good accuracy and reasonable precision and recall for predicting attrition, despite the class imbalance in the dataset. The ROC curve and precision-recall curve below illustrate the performance of the Random Forest model.





7. Statistical Insights

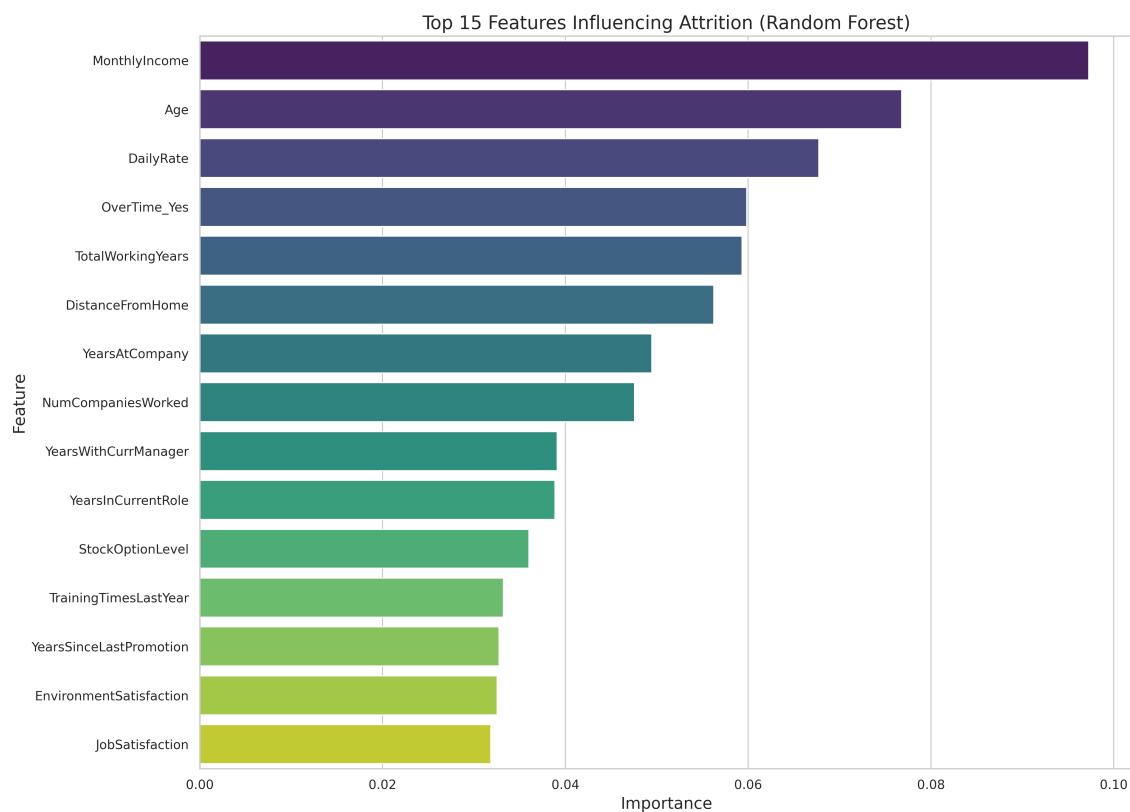
We conducted various statistical tests to identify factors that have a statistically significant relationship with attrition. These tests help confirm that the patterns observed in our exploratory analysis are not due to random chance. Our statistical analysis included:

- T-tests for numerical variables
- Chi-square tests for categorical variables
- Mann-Whitney U tests for ordinal variables

The results of these tests confirmed that many of the factors identified in our exploratory analysis have a statistically significant relationship with attrition.

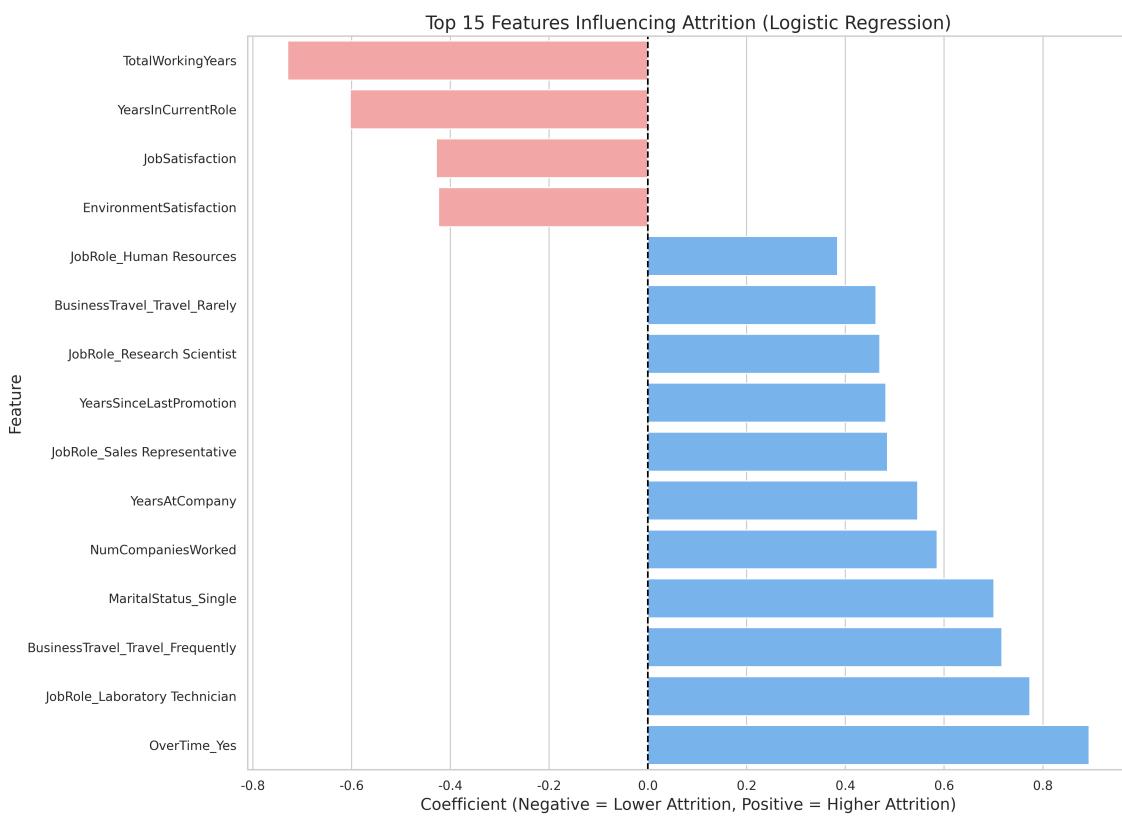
Feature Importance from Statistical Models

Our statistical models provided insights into the relative importance of different factors in predicting attrition. The Random Forest model identified the following as the most important features:



Logistic Regression Coefficients

The Logistic Regression model provides insights into how each factor increases or decreases the likelihood of attrition. Positive coefficients indicate factors that increase attrition risk, while negative coefficients indicate factors that decrease attrition risk.



8. Recommendations

Based on our comprehensive analysis, we recommend the following strategies to reduce employee attrition:

Area	Recommendation
Work-Life Balance	Review overtime policies and workload distribution. Implement flexible work arrangements.
Compensation	Review compensation structures, especially for lower job levels and high-risk roles.
Career Development	Create clearer career progression paths, especially for entry and mid-level employees.
Job Satisfaction	Conduct regular employee engagement surveys and implement responsive management.
Commuting	Consider remote work options or location-based hiring to reduce commute times.
Retention Programs	Develop targeted retention strategies for different employee groups, especially young professionals.
Promotion Practices	Review promotion practices and ensure regular career development discussions, especially for mid-level employees.
New Employee Onboarding	Strengthen onboarding programs to better integrate new employees and reduce early turnover.
Management Training	Provide training for managers on employee engagement and retention strategies.

9. Conclusion

This analysis has provided valuable insights into the factors driving employee attrition within the organization. By understanding these factors, the organization can develop targeted strategies to improve employee retention and reduce the costs associated with turnover. The key findings from our analysis include: 1. Work-life balance, particularly overtime work, is a critical factor in employee attrition. 2. Compensation and career advancement opportunities significantly impact retention. 3. Job satisfaction and work environment play important roles in employee decisions to stay or leave. 4. Demographic factors like age and distance from home influence attrition rates. 5. Certain job roles and departments have higher attrition rates and may require special attention. By implementing the recommended strategies, the organization can create a more engaging and supportive work environment that encourages employees to stay and contribute to long-term success. The predictive models developed in this analysis can also be used to identify employees at high risk of leaving, allowing for proactive intervention before attrition occurs.