

NAME: Nischitha S G

EMAIL: nischithasg@gmail.com

COLLEGE: B.M.S. College of Engineering

Title: IBM HR Analytics Employee Attrition & Performance

Objective

The primary objective of this project is to analyze employee attrition using IBM's HR Analytics dataset through data exploration, visualization, and machine learning techniques to uncover key patterns and drivers of employee turnover.

This analysis seeks to:

- Understand the overall attrition rate within the organization and determine how it varies across different demographic and job-related factors.
- Identify the most influential features contributing to employee attrition, such as job role, overtime, income level, and satisfaction ratings.
- Analyze and visualize patterns based on age, gender, department, and job satisfaction to detect high-risk attrition groups.
- Examine the relationship between work-life balance, performance rating, and tenure with respect to employee retention.
- Provide actionable insights through data-driven recommendations that can help organizations improve employee engagement and reduce attrition.

Tools and Technologies Used

Category	Tools/Technologies	Purpose/Usage
Programming Language	Python	Core language used for data analysis and machine learning
Data Analysis Libraries	Pandas, NumPy	Data loading, manipulation, and numerical analysis
Visualization Libraries	Matplotlib, Seaborn	Creating graphs and charts for data visualization
Machine Learning	Scikit-learn	Applying classification models and evaluating performance
IDE/Notebook	Jupyter Notebook	Writing, executing, and documenting Python code
Dataset Source	IBM HR Analytics Dataset	Realistic HR data used for attrition analysis
PDF/Reference Material	Guided Project Document (PDF)	Provides structure, use cases, and supporting theory

Dataset Information

The dataset used in this project is the IBM HR Analytics Employee Attrition & Performance Dataset, which contains detailed information on 1,470 employees. It is a structured, tabular dataset in CSV format (WA_Fn-UseC_-HR-Employee-Attrition.csv) and includes 35 features, covering both numerical and categorical data types.

This dataset does not have any missing or duplicate values, ensuring that data quality is suitable for analysis. The target variable is Attrition, which indicates whether an employee has left the organization (Yes) or stayed (No).

The dataset includes a wide range of attributes such as:

- Demographic details (Age, Gender, MaritalStatus, Education)
- Job-related information (JobRole, Department, BusinessTravel, JobLevel, JobInvolvement)
- Compensation details (MonthlyIncome, PercentSalaryHike, StockOptionLevel)
- Tenure and performance (PerformanceRating, YearsAtCompany, YearsInCurrentRole, YearsSinceLastPromotion)
- Satisfaction scores (JobSatisfaction, EnvironmentSatisfaction, RelationshipSatisfaction, WorkLifeBalance)
- Behavioral aspects (OverTime, NumCompaniesWorked, TrainingTimesLastYear)

The dataset provides a comprehensive view of employee profiles and is ideal for analyzing the factors that contribute to workforce attrition.

Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is a crucial step in understanding the underlying patterns, trends, and relationships within the dataset. The purpose of EDA in this project is to identify the key features influencing employee attrition and to visualize how various factors differ between employees who left the organization and those who stayed.

1. Attrition Distribution

The dataset shows that approximately 16.12% of employees have left the company, indicating a moderate attrition rate. This imbalance suggests that the dataset is somewhat skewed, which is common in real-world HR data.

2. Age Distribution

Most employees fall within the age range of 30 to 35, indicating a young and productive workforce. A KDE (Kernel Density Estimation) plot revealed that younger employees tend to leave more frequently than older ones.

3. Gender and Attrition

The gender distribution shows more male employees than female. When comparing attrition rates, it was observed that both genders experience attrition, though patterns may differ based on roles and departments.

4. Department Distribution

The majority of employees belong to the Research & Development department, followed by Sales and Human Resources. Understanding attrition within each department helps identify specific teams needing retention strategies.

5. Monthly Income

Employees with lower monthly income appear to leave more frequently than those earning higher salaries, highlighting compensation as a potential attrition driver.

6. Overtime

A strong correlation was observed between employees working **overtime** and higher attrition rates, suggesting that work-life imbalance significantly contributes to employee turnover.

7. Work-Life Balance and Job Satisfaction

Employees with poor work-life balance and low job satisfaction scores were more likely to leave. These insights emphasize the need for supportive HR policies.

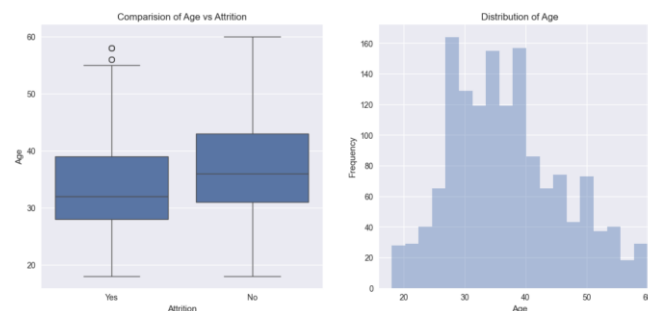
8. Tenure and Years at Company

The average tenure of employees at the company is approximately 7 years. Attrition is higher among employees with fewer years of service, particularly within the first 1–3 years.

Insights and Observations

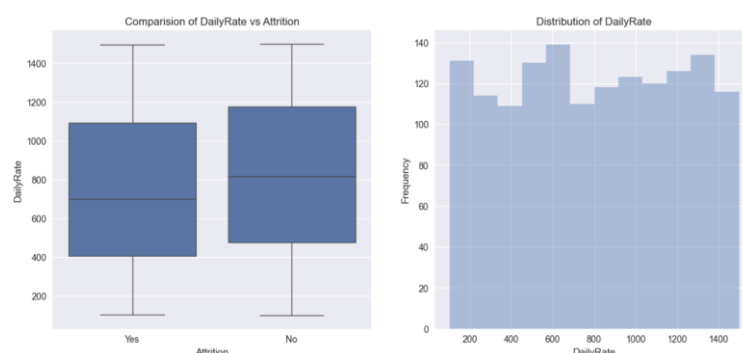
• Age vs Attrition:

Employees who left the company tend to be younger, with a median age around 30, whereas retained employees have a higher median age (~36). Younger employees are more prone to attrition.



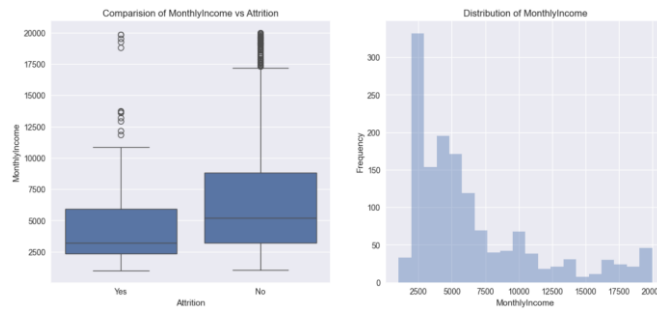
• DailyRate vs Attrition:

DailyRate shows a wide spread for both groups. However, there's no significant difference between attrition and non-attrition groups, indicating that DailyRate may not be a strong predictor of attrition.



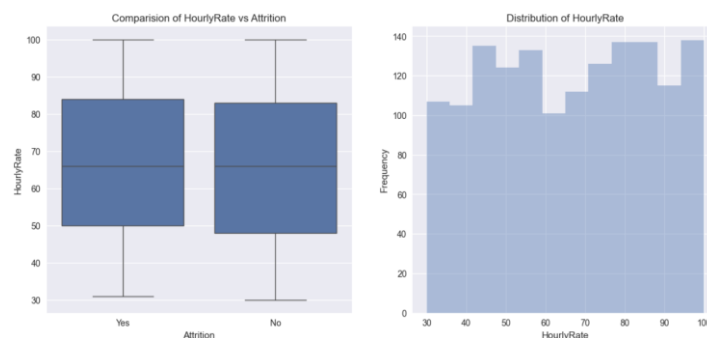
• MonthlyIncome vs Attrition:

Employees with lower monthly incomes are more likely to leave. Those who stayed generally earned higher salaries, suggesting that compensation is a key factor in employee retention.



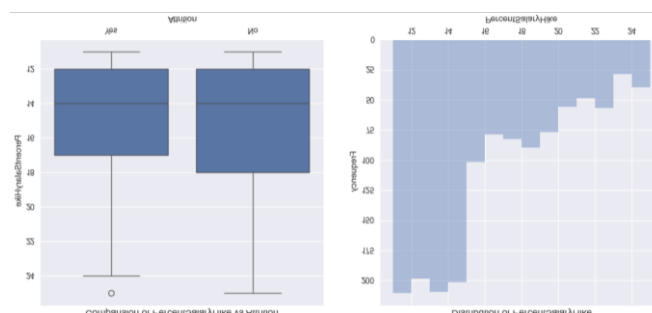
• HourlyRate vs Attrition:

The distribution of hourly rates is fairly uniform across both groups. Median values are almost identical, suggesting hourly rate has little to no impact on attrition.



• PercentSalaryHike vs Attrition:

Salary hikes are slightly lower for employees who left, but the difference is minimal. This implies PercentSalaryHike alone may not be a strong driver of attrition unless considered with other features.



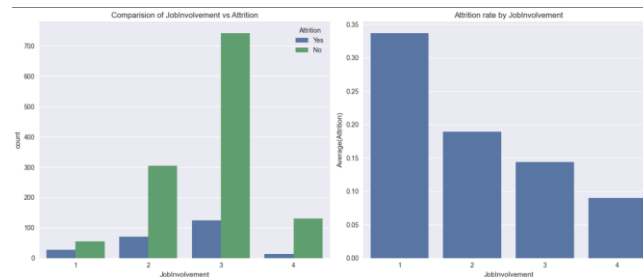
• TotalWorkingYears vs Attrition:

Employees with fewer total working years are more likely to leave. Those with longer career spans tend to stay, possibly due to better job stability or role fit.



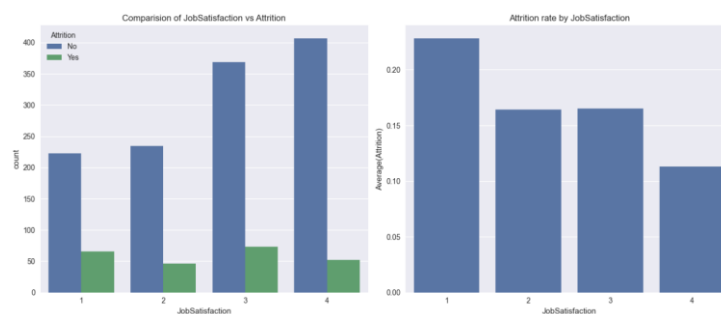
• Age, Monthly Income & Job Satisfaction vs Attrition

Younger employees with lower monthly income and poor job satisfaction are much more likely to leave the company. As employees age, their income and satisfaction levels tend to increase, which correlates with significantly lower attrition. This suggests that retention improves with both experience and adequate compensation, especially when paired with job satisfaction.



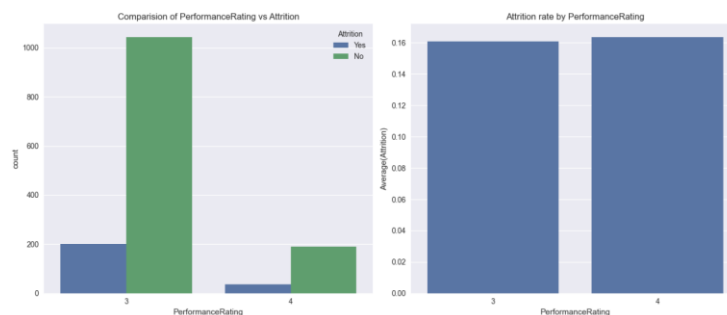
• Attrition by Job Role

Attrition is noticeably higher among operational roles such as Sales Executive and Laboratory Technician, which may be due to limited growth, repetitive tasks, or less rewarding job content. On the other hand, senior roles like Manager and Research Director exhibit very low attrition, indicating that employees in these positions are more satisfied and find greater value in staying with the organization.



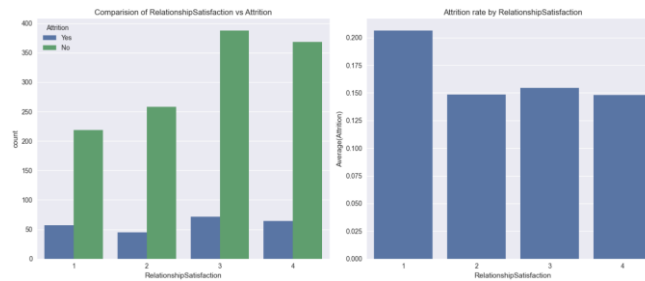
• Attrition by Job Level

Employees at lower job levels, particularly Job Level 1, show the highest attrition rates. As job level increases, attrition decreases sharply, suggesting that with career growth and better compensation, employees are more likely to stay. This pattern indicates a need to focus on development and retention efforts for early-career employees.



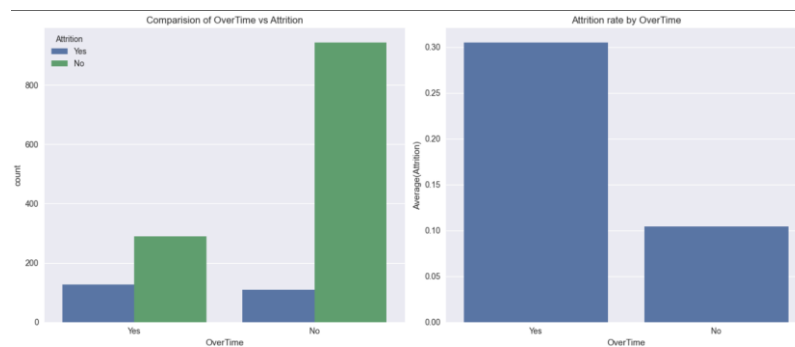
• Attrition by Marital Status

Single employees are significantly more likely to leave compared to their married or divorced counterparts. This may reflect greater mobility, fewer personal obligations, or a stronger drive to explore new opportunities among single individuals.



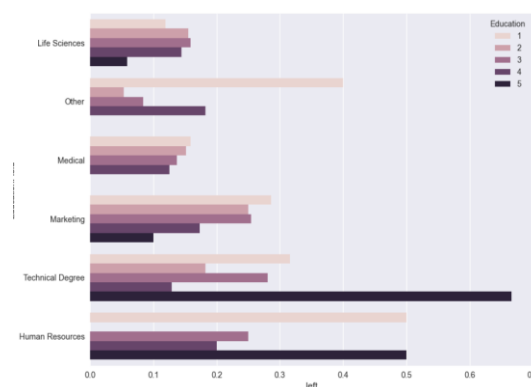
• Attrition by Years at Company

Most attrition occurs within the first 1–3 years of employment, highlighting a critical early period where employees are most vulnerable to leaving. Employees who remain with the company beyond five years are much more likely to stay longer, emphasizing the importance of onboarding, engagement, and growth opportunities in the initial years.



• Attrition by Education Field and Level

Employees with Technical Degrees and Education Level 5 show the highest attrition, indicating possible mismatch or better opportunities elsewhere. Human Resources also sees high attrition at higher education levels. Fields like Medical and Life Sciences show lower attrition, suggesting more job satisfaction or stability in those areas.



• Attrition by Monthly Income and Education

Attrition is more common among employees with lower monthly incomes, regardless of education level. As income increases, attrition decreases, indicating financial compensation as a strong factor in employee retention. Notably, even employees with higher education tend to leave if the income is not competitive, emphasizing the importance of aligning pay with qualifications.



Conclusion

This project successfully analyzed employee attrition using the IBM HR Analytics dataset by applying data preprocessing, exploratory data analysis (EDA), and visualization techniques. The analysis revealed several important insights into the factors influencing employee turnover.

Key conclusions drawn from the study include:

- **Attrition Rate:** Approximately 16.12% of employees have left the organization, indicating a moderate level of turnover that can impact productivity and stability.
- **High-Risk Groups:** Employees who are younger, have lower income, and shorter tenure are more likely to leave the company. Those frequently working overtime or reporting poor work-life balance also face higher attrition risk.
- **Satisfaction & Engagement:** Lower scores in job satisfaction, relationship satisfaction, and environment satisfaction are closely linked to higher attrition, emphasizing the need for continuous employee engagement.
- **Compensation & Tenure:** Employees with higher monthly income and more total working years tend to stay longer, suggesting that financial growth and career stability play crucial roles in retention.
- **Feature Significance:** Some features like HourlyRate and DailyRate showed minimal impact on attrition, indicating they may not be effective indicators for retention analysis.

The analysis also reveals that education level and monthly income significantly influence attrition. Highly educated employees, particularly in domains such as Human Resources and Technical Degrees, exhibit higher attrition, potentially due to role mismatch or unmet career expectations. Furthermore, lower income levels are consistently associated with higher attrition, reinforcing the importance of competitive compensation structures.

To reduce turnover, organizations should:

- Align job roles with employee qualifications, especially for highly educated individuals.
- Offer fair and performance-driven salaries to ensure financial satisfaction.
- Focus on enhancing job satisfaction and work-life balance through supportive HR policies and employee engagement programs.