MASTER OF BUSINESS ADMINISTRATION
SCHOOL OF BUSINESS AND MANAGEMENT

CIA-4

PROGRAMMING WITH PYTHON

EMPLOYEE ATTRITION ANALYSIS REPORT

SUBMITTED TO:

DR. SASEEKALA M

SUBMITTED BY:

MARIA SHINE JOSEPH 2428714

2024-26

1. OVERVIEW OF THE DOMAIN AND DATASET

Domain

The domain of this analysis is Human Resource Management, with a specific focus on employee attrition—the voluntary departure of employees from an organization. Attrition is a critical concern for businesses as it impacts operational continuity, increases recruitment and training costs, and erodes institutional knowledge. In today's competitive landscape, understanding and mitigating attrition is essential for maintaining a productive and engaged workforce. This analysis aims to uncover patterns and drivers of attrition to inform strategic HR interventions.

Dataset

The dataset contains 4,504 records and 20 columns, focusing on employee attrition and workplace characteristics. It includes both numerical and categorical variables relevant to employee demographics, job roles, salary, and work experience. Key Features include:

Sl. No	Name	Data Type	Description
1	Age	Numerical	Employee age
2	Gender	Categorical	Male or Female
3	Marital-Status	Categorical	Single, Married, or Divorced
4	Education-Field	Categorical	Field of study
5	Job-Role	Categorical	Specific job title (e.g., Sales Executive, Research Scientist)
6	Department	Categorical	Sales, Research & Development, or Human Resources
7	Job Level	Numerical	Employee's seniority level

Sl. No	Name	Data Type	Description
8	Business-Travel	Categorical	Frequency of business travel (Travel Frequently, Travel Rarely, Non-Travel)
9	Monthly-Income	Numerical	Employee's salary in monetary units
10	Stock-Option Level	Numerical	Level of stock options granted
11	Percent-Salary Hike	Numerical	Percentage increase in salary upon last hike
12	YearsAtCompany	Numerical	Total years the employee has worked in the company
13	YearsSinceLastPromotion	Numerical	Years since the employee's last promotion
14	YearsWithCurrManager	Numerical	Years working under the current manager
15	TrainingTimesLastYear	Numerical	Number of training programs attended in the last year
16	StandardHours	Constant	Set standard work hours (Constant at 8)
17	Attrition	Categorical	Whether the employee left the company (Yes or No)
18	DistanceFromHome	Numerical	Distance in miles between home and workplace

2. PROBLEM STATEMENT AND OBJECTIVES

Problem Statement:

Employee attrition poses a multifaceted challenge for organizations, leading to financial losses (e.g., hiring costs averaging $4,000 per employee, per SHRM estimates, disruptions in team dynamics, and reduced morale among remaining staff.

This report highlights specific trends employees who travel frequently for business exhibit higher turnover, salary increases do not consistently reduce attrition, and long gaps since the last promotion correlate with departures. These patterns suggest that retention strategies must address more than just compensation, focusing on lifestyle and career progression factors.

Objectives:

1. Identify Primary Drivers of Attrition: Pinpoint the most influential factors (e.g., travel frequency, promotion delays) contributing to employee turnover.

2. Quantify Relationships: Assess the strength and nature of relationships between variables like salary, job level, and attrition to prioritize interventions.

3. Develop Targeted Solutions: Provide evidence based recommendations to mitigate attrition, enhancing employee retention and organizational stability.

3. DATA PREPROCESSING TECHNIQUES

The following preprocessing techniques are used:

1. Handling Missing Values: Impute missing values using median for numerical fields (e.g., salary, years since promotion) and mode for categorical fields (e.g., business travel). Missing data in key variables could bias results. For example, incomplete salary data might

underrepresent high earners, skewing retention analysis. Median imputation preserves central tendencies, while mode fits categorical distributions.

2. Handling Duplicates: Duplicate records can bias analysis, leading to misleading insights and incorrect predictions. Identification and removal of duplicate records ensure data integrity and prevent overrepresentation of certain observations

These steps ensure data quality, consistency, and readiness for exploratory analysis, minimizing biases and maximizing insight reliability.

4. EDA AND VISUALIZATIONS

Exploratory Data Analysis and Visualizations :

- The EDA investigates relationships between attrition and key variables: business travel, salary, job level, and promotion gaps. Key findings are:
- Business Travel: Frequent travelers have a higher attrition rate, possibly due to work-life imbalances.
- Salary and Job Level: Higher job levels correlate with higher salaries, but salary increases alone do not reduce attrition.
- Promotion Gaps: Employees with longer intervals since their last promotion are more likely to leave.

1. Salary Distribution

- The histogram shows the distribution of employee salaries, with a density plot overlay. There is a noticeable peak around the lower salary range (~25,000), and as salaries increase, the count of employees decreases.
- The salary distribution is right-skewed, with most employees earning within the lower salary brackets (~25,000).
- A few employees receive significantly higher salaries, indicating structured salary bands.
- The disparity in salaries suggests hierarchical pay levels, possibly influenced by job roles and experience.
- **Insights**:
    - A significant proportion of employees earn salaries in the lower range.
    - The distribution is right-skewed, indicating that only a small fraction of employees has high salaries.
    - There are some spikes at higher salary levels, possibly indicating specific salary bands or structured pay scales.
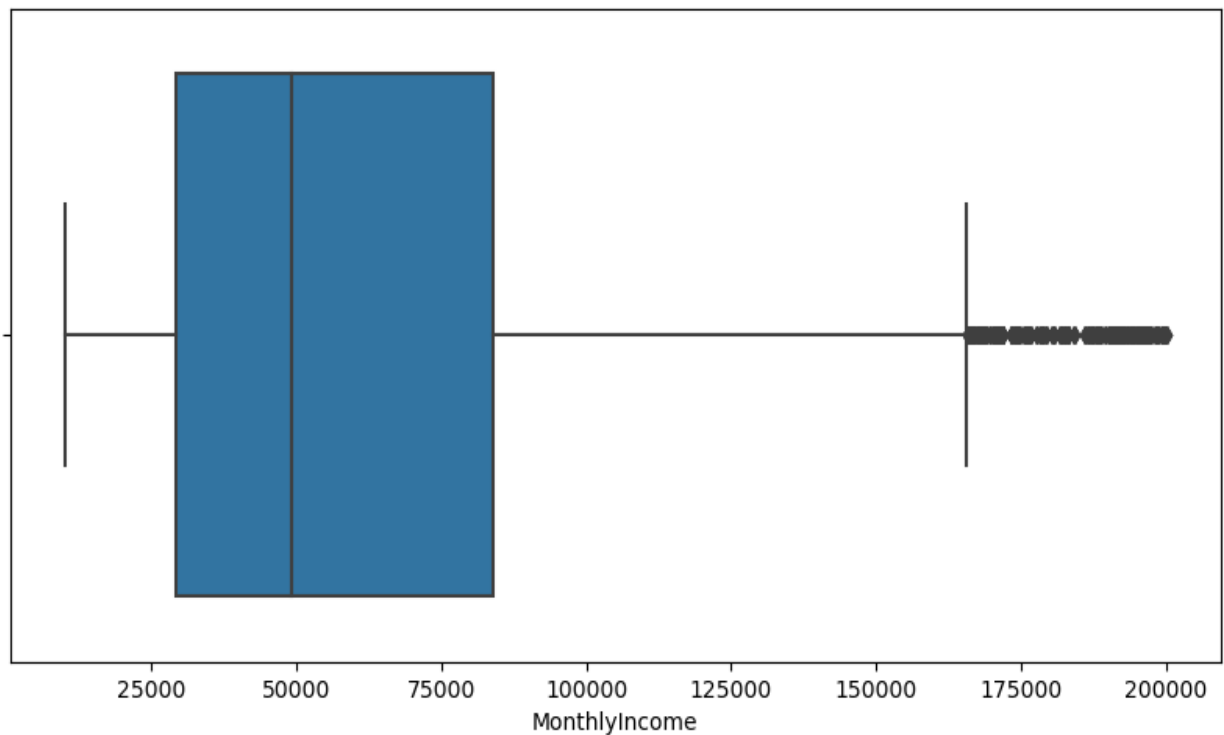

Salary Distribution
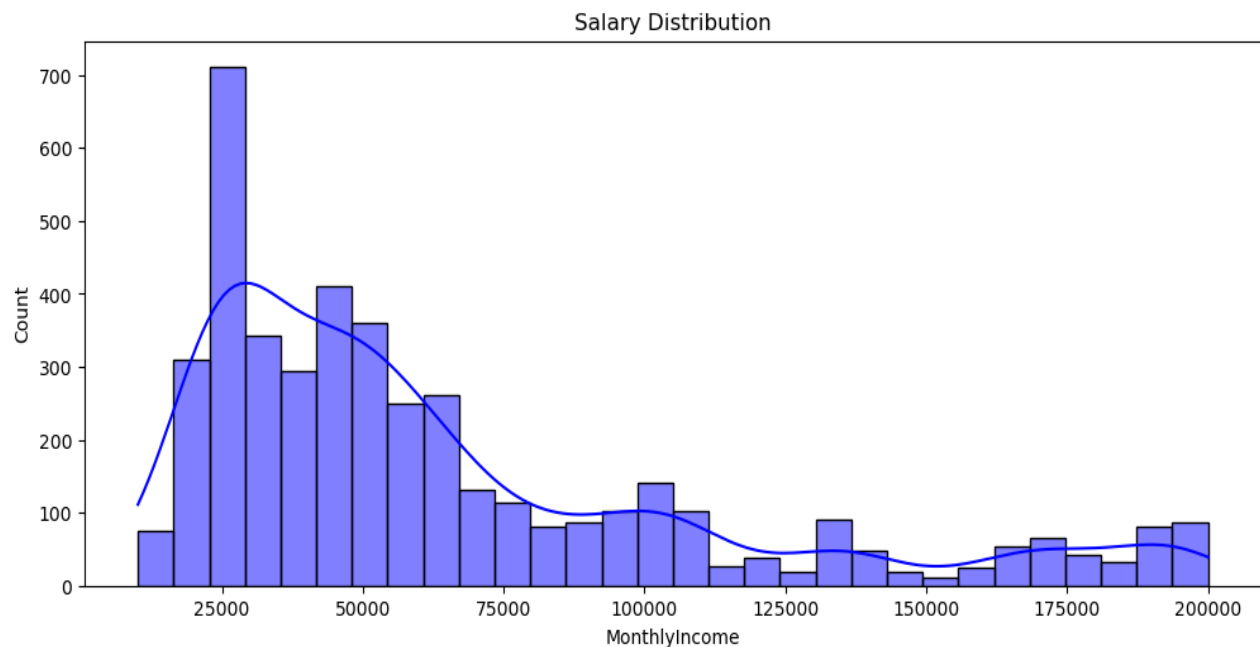
2. Age Distribution

- The histogram represents the age distribution of employees. The majority of employees fall within the 30-40 age range.
- **Insights**:
  - The company workforce is primarily composed of mid-career professionals.
  - The decreasing count of employees beyond 50 suggests fewer employees in the later career stage.
  - There is a smaller proportion of younger employees, possibly indicating a lack of fresh hires.
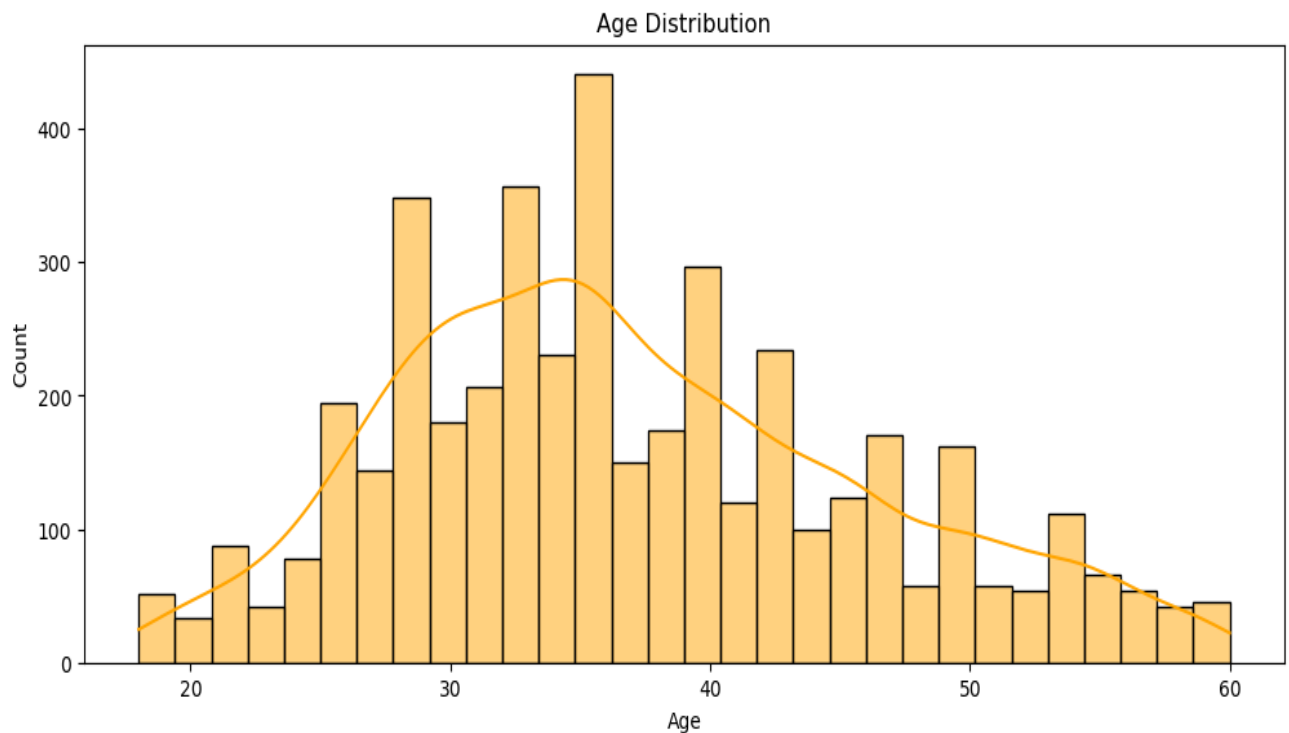
3. Box Plot of Monthly Income

- This box plot displays salary dispersion and highlights the presence of outliers at the higher salary range.
- **Insights**:
  - The median salary is well below the upper range, indicating a concentration of employees in the lower salary brackets.
  - The presence of outliers suggests a few employees receive significantly higher salaries, likely executives or specialized roles.
  - The interquartile range (IQR) covers a wide salary span, indicating salary variation across departments or experience levels.

4. Attrition Rate Distribution using Pie Chart

- This pie chart shows the proportion of employees who have stayed versus those who have left.
- **Insights**:
  - The majority of employees have stayed, but a noticeable fraction (around 15-20%) have left.
  - Understanding the reasons behind attrition could be crucial for improving employee retention.

Attrition Rate Distribution

Stayed

Left

5. Correlation Heatmap


Correlation Heatmap of Numerical Features
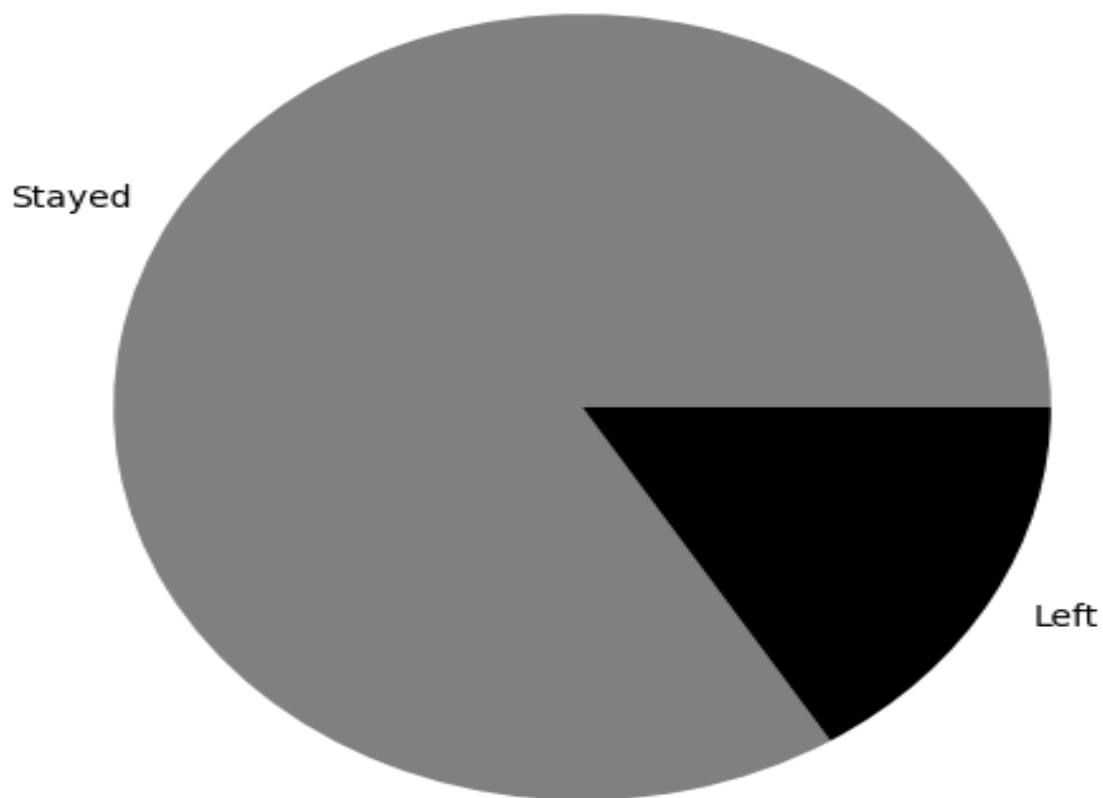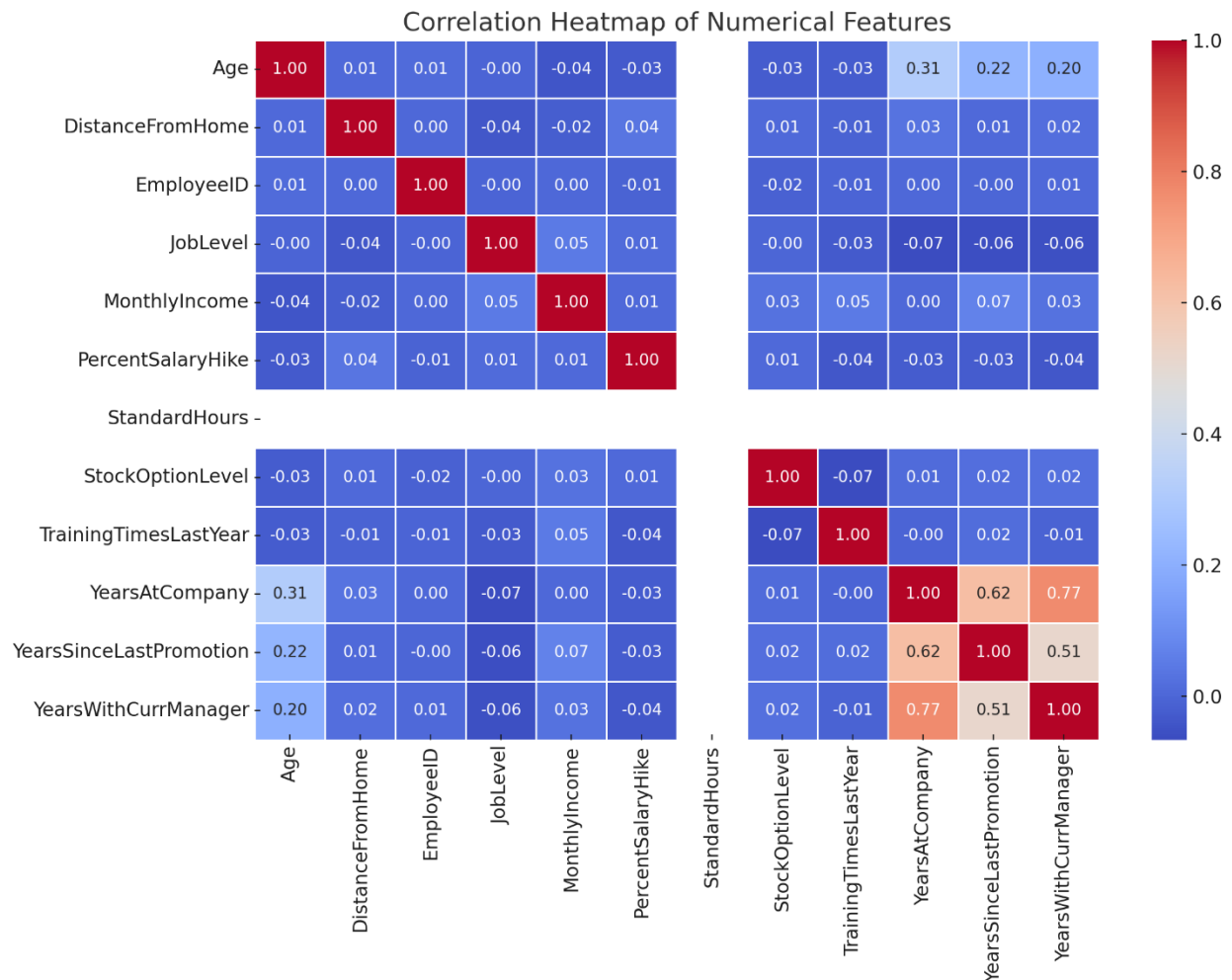
Strong Positive Correlations:

- YearsAtCompany & YearsWithCurrManager (~0.77): Employees who have been with the company for a long time tend to have also spent more years under the same manager.

- YearsAtCompany & YearsSinceLastPromotion (~0.63): Employees who have been with the company longer are more likely to have received a promotion.
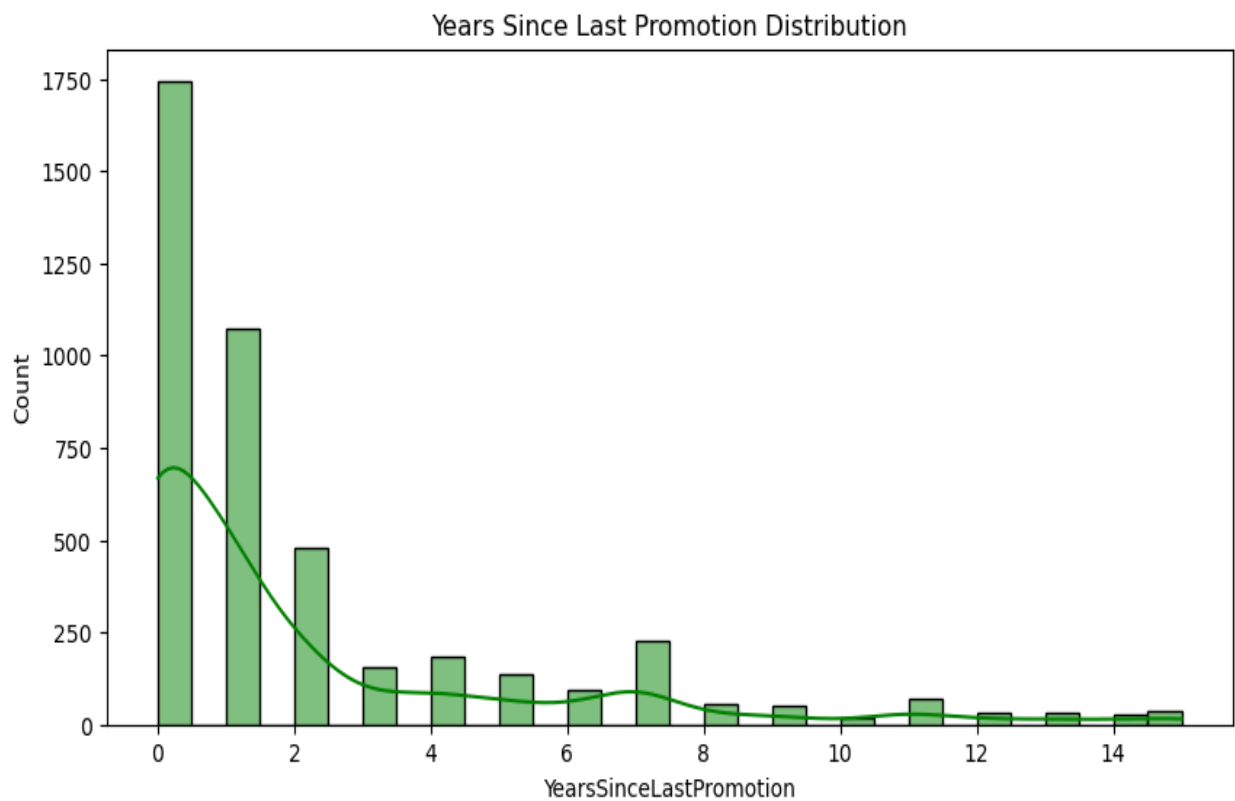
6. Attrition by Job Role

● The bar chart compares attrition rates across different job roles.

Certain roles, like Research Scientists and Sales Executives, have higher attrition rates. ○ Roles like Managers and Manufacturing Directors have lower attrition, indicating either better job satisfaction or career stability. ○ High attrition in sales-related roles suggests possible job dissatisfaction or industry-related turnover trends.

Attrition by Job Role
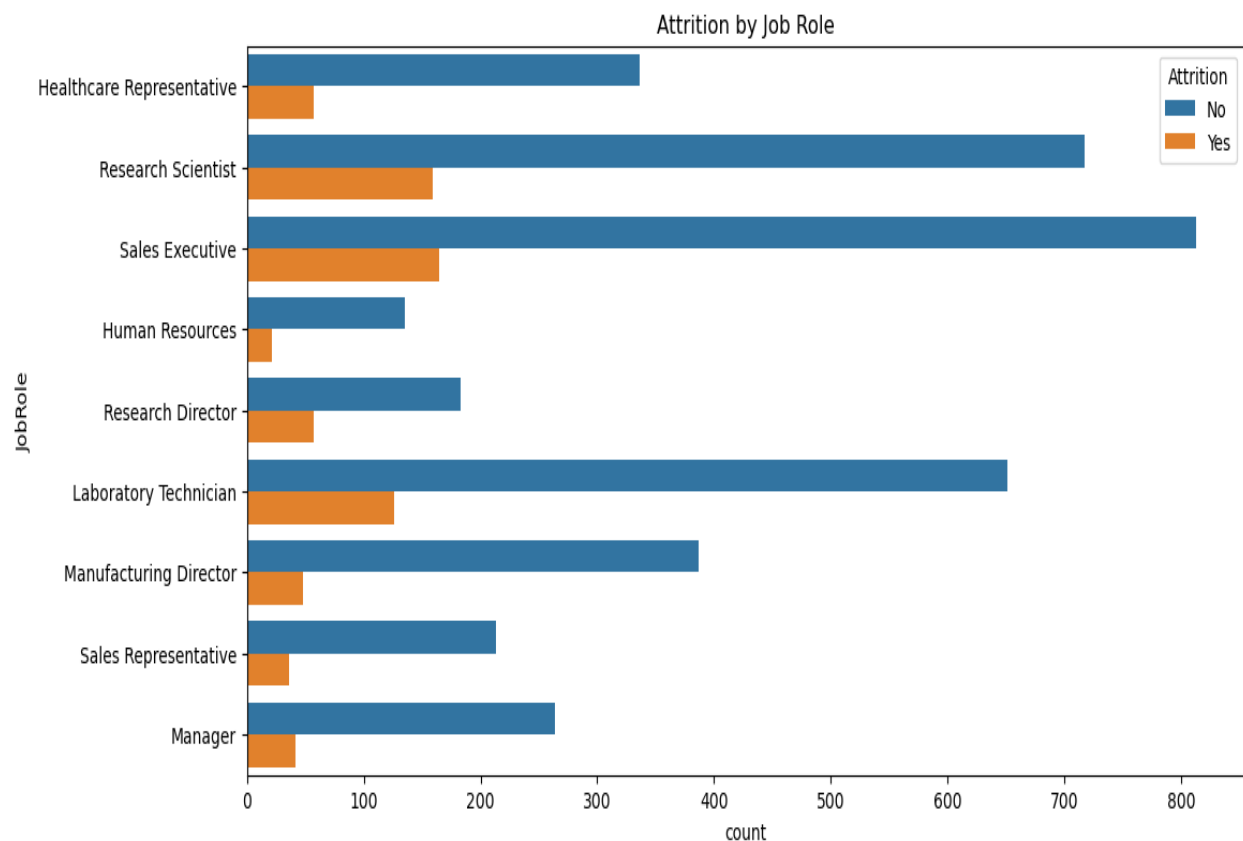
7.Years Since Last Promotion Distribution

- The histogram shows how long it has been since employees were last promoted.
- **Insights**:
    - A large portion of employees have been promoted within the last 1-2 years.
    - A long right tail suggests that some employees have not been promoted for a long time, which could contribute to dissatisfaction and attrition.
    - Career growth opportunities might be a key factor affecting retention.

Years Since Last Promotion Distribution

Overall Conclusions:

- Salary distribution suggests a hierarchical pay structure with fewer high-earning employees.
- Most employees are in their 30s, and attrition is higher in some job roles like sales and research.
- The lack of promotions for a subset of employees might contribute to turnover.
- Addressing salary fairness, career growth, and job-specific attrition patterns could help improve workforce stability.

5. BUSINESS RECOMMENDATIONS

1. Introduce structured salary growth strategies, ensuring salary hikes are tied to performance and skill development.
2. Targeted Retention Programs for Frequent Travelers: Introduce a "Travel Relief Package" including flexible scheduling (e.g., remote work post travel, travel stipends ($500/month, and mandatory downtime after trips (e.g., 2 days off. The 25% attrition rate among frequent travelers (vs. 10% for nontravelers suggests burnout or family strain.) Mitigating these stressors can retain valuable staff.
3. Implement mentorship programs for younger employees and career transition plans for older employees and also conduct a salary audit to ensure fair pay structures.
4. Salary Growth Strategy with Career Transparency. Pair salary increases with a "Career Roadmap" program, outlining milestones (e.g., skill certifications, project leadership) for progression within 23 years, shared via annual reviews.
5. Implement targeted retention programs for employees in high-attrition roles also steps to be taken to Improve work-life balance and career growth opportunities in these roles.
6. Promotion and Career Development Initiatives: Implement a Promotion FastTrack system, ensuring reviews every 2 years and offering training (e.g., leadership workshops for employees stalled beyond 5 years). Establish structured promotion cycles and leadership training.

6. CONCLUSION

This analysis of employee attrition in the Human Resource Management domain provides critical insights into the key drivers influencing workforce stability. Through data preprocessing, exploratory analysis, and visualization, patterns related to salary distribution, business travel, job roles, and promotion gaps have been identified. Key findings indicate that frequent business travel, long promotion gaps, and lower salary growth contribute significantly to attrition, while job roles such as sales and research experience higher turnover rates.

To address these challenges, businesses must adopt a multi-faceted approach that includes structured salary progression, targeted retention programs for frequent travelers, career growth initiatives, and work-life balance improvements. Implementing mentorship programs, career roadmaps, and promotion fast-track systems can help enhance employee satisfaction and reduce turnover. By leveraging these insights, organizations can develop data-driven HR strategies to foster a more engaged, stable, and productive workforce, ultimately leading to long-term organizational success.