Employee Attrition Data Analysis

1. Introduction

1.1 Background

The Company stated in the Employee Attrition Data project, established a few years ago, has been experiencing an employee attrition rate of approximately 16% over the past couple of years. This high attrition rate is negatively impacting various aspects of the organization, including productivity, morale, and overall performance. To address this issue and reduce employee turnover, the company has engaged an HR analytics consultancy to analyze its employee data. The goal is to identify the underlying causes of employee attrition and develop strategies to improve retention.

1.2 Problem Statement

Company is facing a significant challenge with employee attrition, which is affecting organizational efficiency and productivity. The company seeks to understand the key factors leading to employee departures. By analyzing employee data, the organization aims to uncover patterns and reasons for attrition and implement data-driven decisions to reduce turnover rates.

1.3 Project Objectives

Identify Key Factors: Analyze employee demographic, performance, and engagement data to determine the primary factors contributing to employee attrition.
Create a Dashboard: Develop a user-friendly dashboard that presents actionable insights and visualizations to help the HR team make informed decisions.
Provide Recommendations: Based on the analysis, suggest strategic initiatives to mitigate attrition and improve employee retention.

2. Methodology

2.1 Data Exploration
I) Data Set

The dataset provided for the Employee Attrition Analysis project includes various attributes that capture employee demographics, job-related factors, performance metrics, and other relevant data. These details are crucial for analyzing the reasons behind employee turnover. Although the actual dataset file was not provided directly in the document, here are the common features typically found in such datasets:

- **Age:** The age of the employee.
- **Employee ID:** Depicts the unique employee identification number
- **Gender:** Male or female.
- **Attrition:** A binary value (Yes/No) indicating whether the employee has left the company or is still working there.
- **Employee Count:** Count of Employee(s)

- **Education Level:** Categories like high school, bachelor's degree, master's degree, etc.
- **Education Field:** Fields of education employees have gained education in (Medical, Life Sciences, Others, etc)
- **Department:** Department the employee works in (e.g., Sales, HR, R&D).
- **Total Working Years:** Total number of years the employee has worked with the company
- **Performance Rating:** Employee's performance rating in numbers
- **Job Involvement:** Employee's involvement in the specific job in numbers
- **TotalWorkingYears:** Employee's total number of working years.
- **Job Satisfaction:** Employee's job satisfaction on a scale.
- **Environment Satisfaction:** Satisfaction level with the working environment.
- **WorkLifeBalance:** Rating indicating the balance between work and personal life.
- **DistanceFromHome:** Distance between the employee's residence and the workplace.
- **YearsAtCompany:** Number of years at the Company
- **YearsSinceLastPromotion:** Number of years passed since Employee's last promotion
- **YearsWithCurrManager:** Employee's numbers of years under the current manager in numbers
- **TrainingTimesLastYear:** Number of times employee has been assigned training in last year
- **BusinessTravel:** Degree of business travel

The dataset includes a variety of quantitative and qualitative variables that can be used to:
- Identify trends in attrition based on demographic or job-related factors.
- Understand correlations between job satisfaction, compensation, and turnover.
- Develop targeted strategies to reduce turnover and improve employee retention.

These features are vital for performing the analysis needed to support the project's objectives and develop a dashboard that provides actionable insights to the HR team at the Company.

2.2 Data Preparation

I) Data Preprocessing: The processes that are included in data preparation include cleaning up data, verifying missing and null values, removing redundant data, and filling in null values.

We have made a deep copy of the dataset and processed every pre-processing step using this method because we are unable to convert some attributes, like Business travel and Education Fields, into numerical values. As a result, we have made dummies and processed them while verifying the null values; we have null values in some of the attributes for example.

1) Number of Companies Worked 2) Satisfaction with Environment 3) Workplace Contentment 4) Balance Work-Life and use a median to fill them in.

2.3 Exploratory Data Analysis

I) Data Overview: The dataset used for the overall analysis has a total of 29 attributes with 4410 records. As a part of this endeavor, a deep copy of an original data file has been created to glean better trends, and transform attributes in readable format that will aid in the insightful visualizations.

The new file consists of dummies that contribute to 48 attributes and 4410 rows. These dummies have improvised certain attributes like - Business Travel, Education Field, Department, Marital Status, Job Role, etc. The improvisation gives more granular controls to pick approximate data for enhanced and decision-driven insights.

II) Outlier Detection: Outlier detection for an attrition dataset using a box plot revealed many outliers. To eliminate these outliers, we employed a standardization process (z-score) and the results showed that the outliers were minimized from some of the attributes, including -
1) MonthlyIncome
2) Number of Companies Worked.
3) Years Worked Overall
4) Last Year's Training Schedule
5) Years at Company
6) Years Since Last Promotion
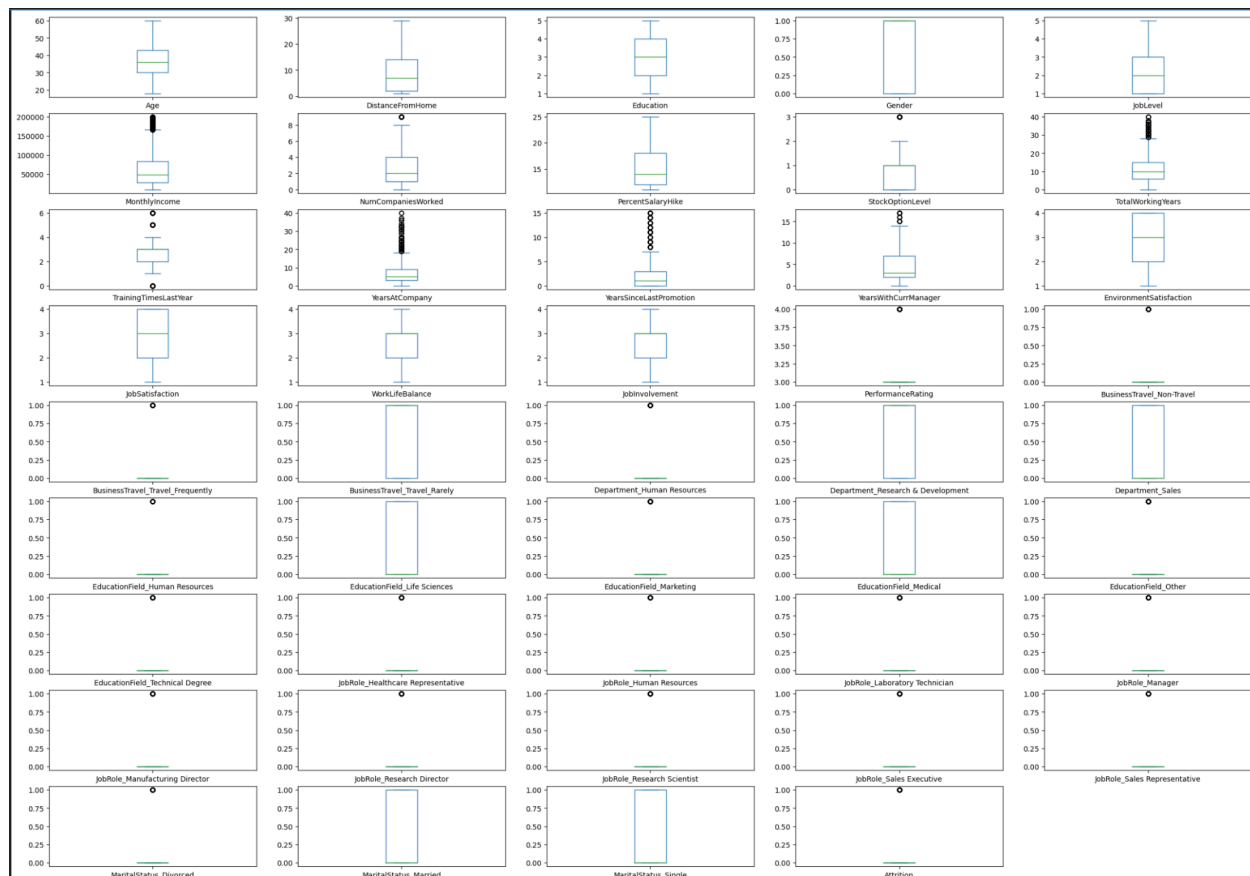7) Years Working with CurrManager

Figure: Shows outliers in the dataset

III) Correlation:

For the correlation aspect the heatmap and manual correlation have been developed that explicitly demonstrates the positive, the negative, and statistically significant correlations between the given data attributes, indicating that the attributes are influenced by each other.

# 3. Data Representation & Visualizations

Attrition Analysis

The dataset indicates that 16% of the entire employee population experiences attrition. This attrition rate provides a clear picture of workforce turnover and can serve as a key indicator of organizational health, signaling potential areas for improvement in employee retention strategies. The 16% attrition rate highlights a significant portion of the workforce affected by turnover, offering insight into the company's ability to retain talent.
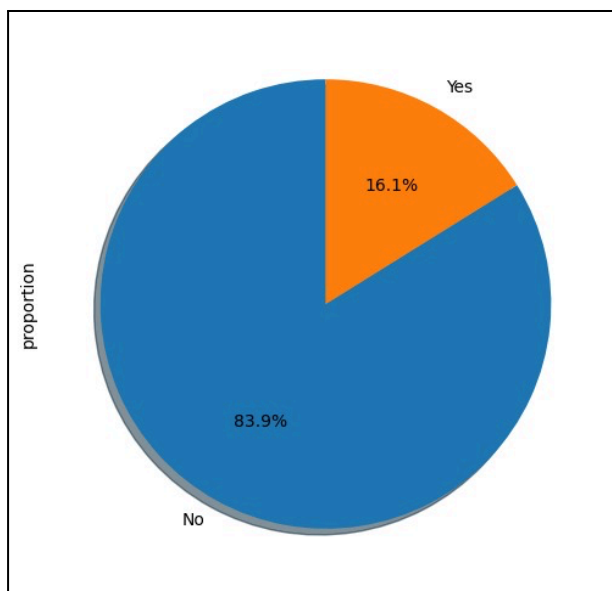
- **Effect of Gender Analysis**

The analysis reveals that gender has a notable yet variable influence on employee attrition. As illustrated in the figure, the attrition rate is significantly higher among males compared to females. Males account for 8% of the total workforce, and they also represent a disproportionately larger share of overall attrition cases. In contrast, the number of female employees is comparatively lower, and their contribution to attrition is relatively minimal. This disparity suggests that gender plays a role in shaping turnover rates, with males being

more   prone   to   leaving   the   organization   than   females.
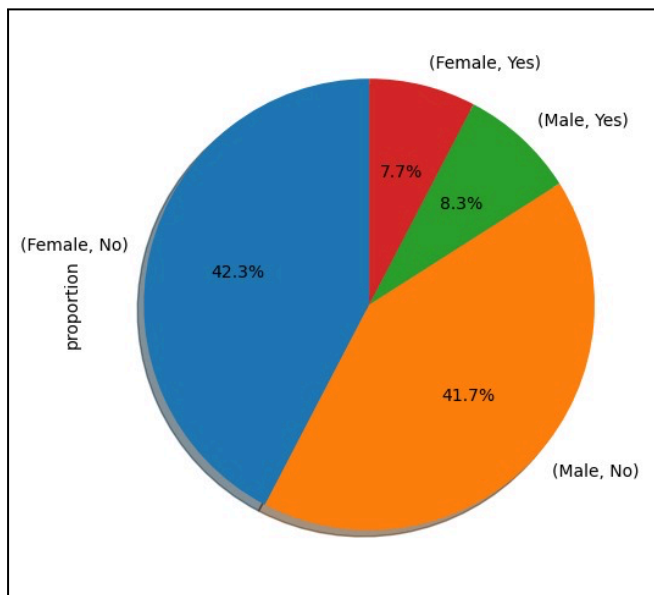
- **Business Travel**

The figure below reveals a distinct pattern regarding the impact of business travel on employee attrition. Employees who travel infrequently are the most affected, showing the highest attrition rates compared to their peers. Following closely are those who travel frequently, experiencing notable but slightly lower attrition. In stark contrast, employees who do not travel at all exhibit a significantly lower attrition rate, standing at just 0.8% of the total employee population. This suggests that travel demands may be a contributing factor to employee turnover, with non-travelers demonstrating greater stability and retention within the organization.
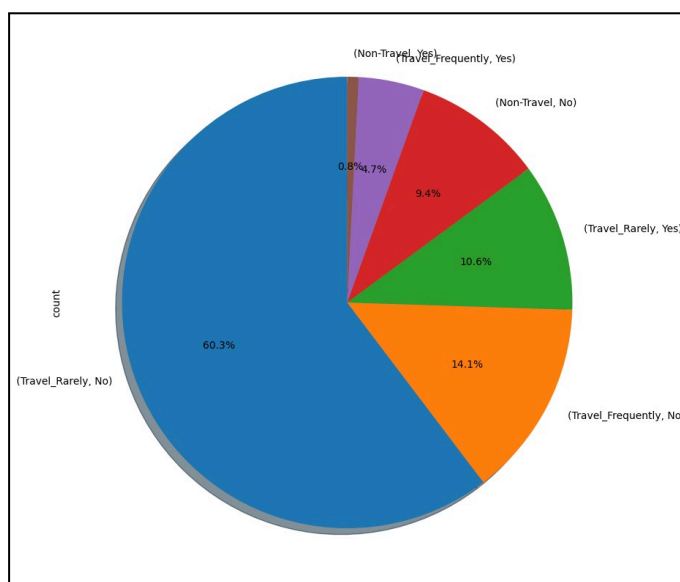
- Distance From Home

The figure below clearly demonstrates that employee attrition is highest for those living within a 0-10 kilometer radius from the workplace. This proximity seems to correlate with increased turnover rates, suggesting that employees who live closer to the office may be more likely to leave the organization. Whether due to shorter commutes providing easier access to other opportunities or other factors, this geographic range experiences a noticeably higher level of attrition compared to employees living farther away.
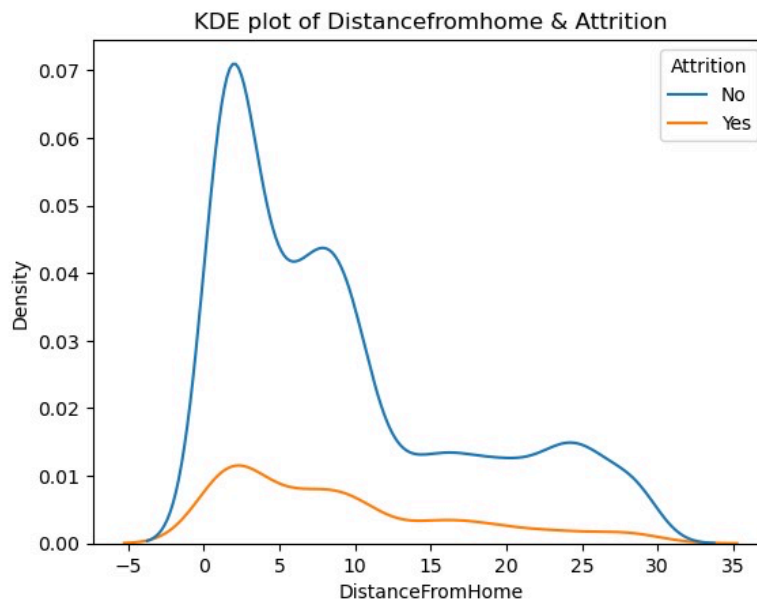

- Department

The analysis indicates that the "Research & Development" department experiences the highest attrition rate, with 10.3% of the total employee population leaving the organization. This suggests that R&D faces unique challenges in retaining its workforce, possibly due to the nature of the work or specific department-related factors. Following R&D, the "Sales" department also shows notable attrition, standing at 4.6% of the total workforce. While significantly lower than R&D, the Sales department's turnover still represents a meaningful portion of employee departures, highlighting potential areas for further investigation into retention strategies for these key departments.
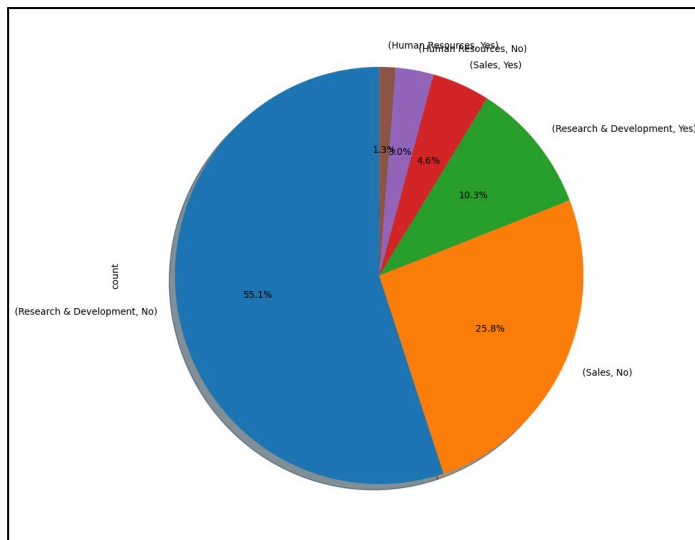
- **Job Position**

The analysis of the Job Role attribute reveals that the positions most impacted by attrition are Sales Executives and Research Scientists. Sales Executives show the highest attrition rates, making them the most affected role in the organization. Following closely behind are Research Scientists, who also experience significant turnover. These two roles, topping the list for attrition, suggest that employees in sales and research positions may face particular challenges, such as job demands, stress, or market conditions, contributing to higher turnover compared to other roles within the company.
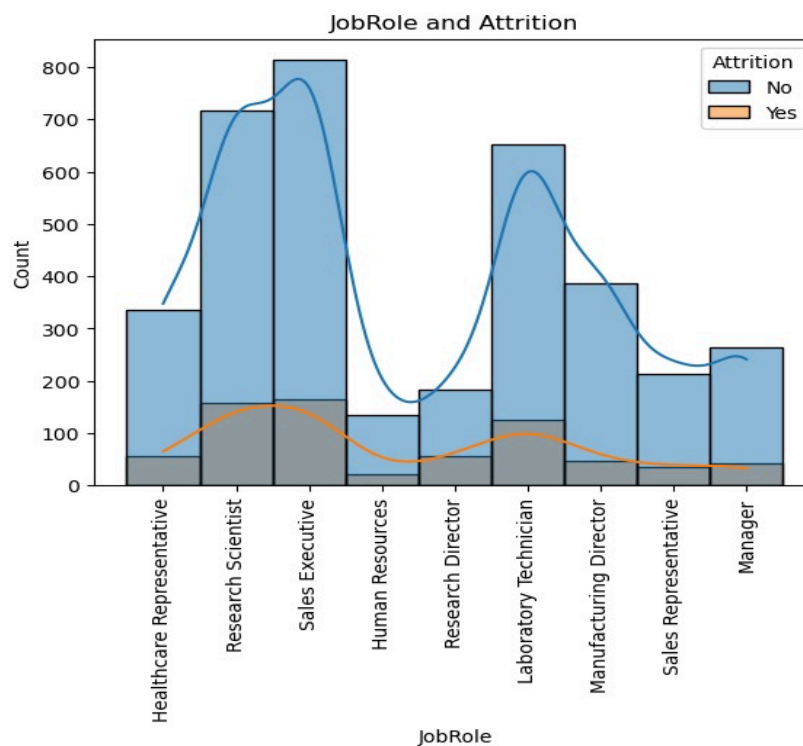
- **Education & Education Field**
  The analysis of the Education attribute shows that employees with a Level 3 education are the most impacted by attrition, experiencing the highest turnover rates. This group is followed by those with a Level 4 education, who also see a notable level of attrition, though not as high as Level 3. Interestingly, employees with a Level 1 education are the least affected by attrition, demonstrating the lowest turnover rates. This pattern suggests that employees with mid-level education qualifications may be more likely to leave the organization, while those with lower education levels exhibit greater stability in their roles.

  The analysis of the Education attribute shows that employees with a Level 3 education are the most impacted by attrition, experiencing the highest turnover rates. This group is followed by those with a Level 4 education, who also see a notable level of attrition, though not as high as Level 3. Interestingly, employees with a Level 1 education are the least affected by attrition, demonstrating the lowest turnover rates. This pattern suggests that employees with mid-level education qualifications may be more likely to leave the organization, while those with lower education levels exhibit greater stability in their roles.
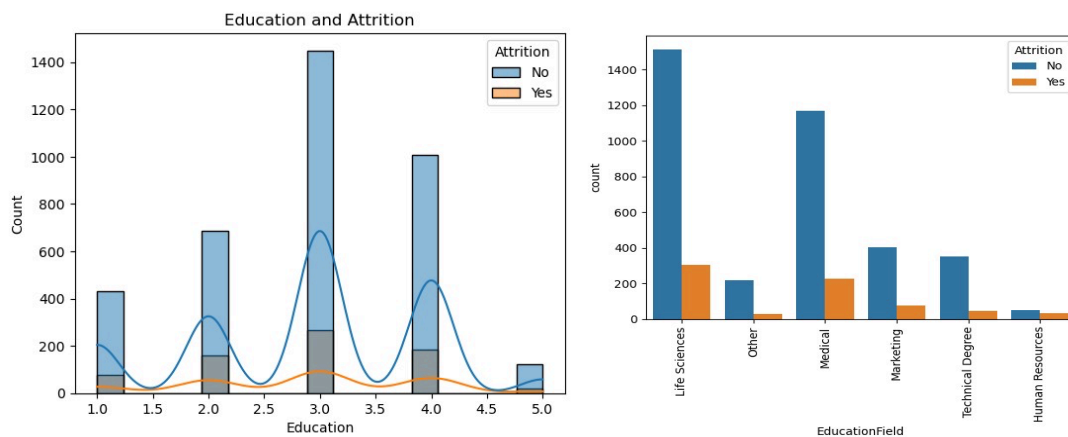

- **Effect of Age**
  The analysis reveals that employees in the age group close to 30 are significantly more impacted by attrition compared to other age groups. This suggests that individuals in this stage of their careers may be more likely to seek new opportunities or face challenges that contribute to higher turnover. The increased attrition in this demographic could be influenced by various factors such as career growth expectations, job satisfaction, or personal life transitions. As a result, the 30-year age group represents a critical area for understanding and addressing employee retention strategies.
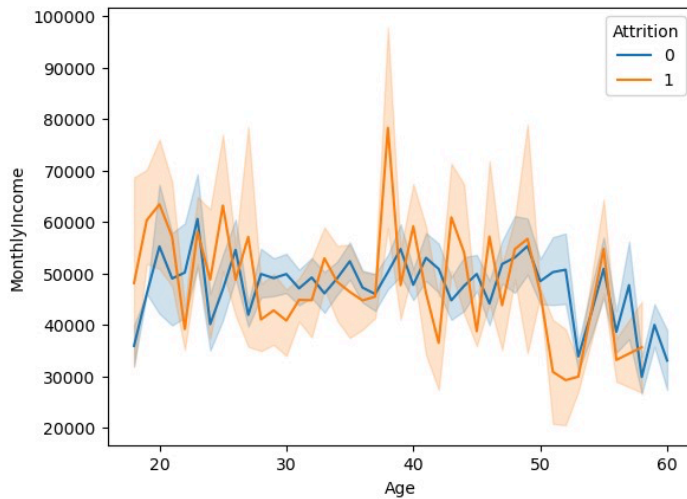
- **Effect of Salary**

   The analysis indicates that employees with salaries in the range of $30,000 to $75,000 are more likely to experience attrition compared to other salary brackets. This mid-level salary range shows a higher turnover rate, suggesting that employees earning within this range may be more inclined to leave the organization. Possible reasons for this could include dissatisfaction with compensation, the pursuit of better-paying opportunities, or a lack of growth potential. Understanding the factors driving attrition within this salary group can be crucial for developing targeted retention strategies aimed at improving employee satisfaction and reducing turnover.
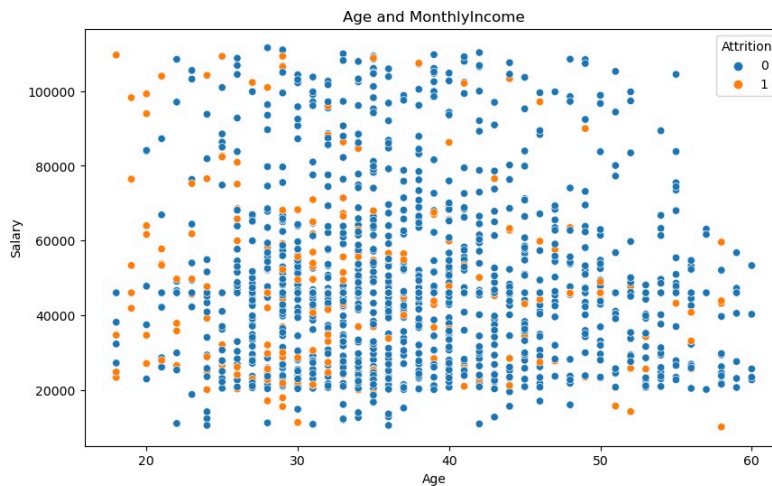
- **Work Life Balance**
  The data shows that employees with a Level 3 have the highest rate of attrition compared to others. Whether it's due to limited career advancement opportunities, job dissatisfaction, or other external factors, Level 3 employees are leaving the organization at a higher rate.
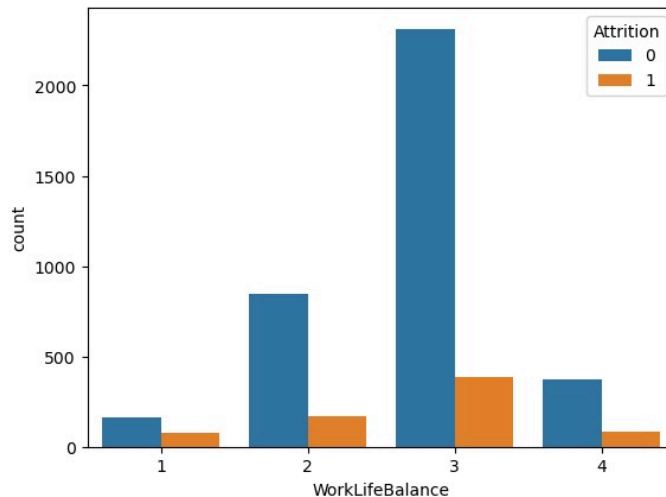

- **Effect of Current Manager**
  The analysis of the Current Manager attribute reveals that attrition is most pronounced during the initial years of employment, particularly within the first year. Following this, a notable level of turnover is also observed during the second year under the same manager. This pattern indicates that new hires may face challenges or adjustments during their early tenure that contribute to higher attrition rates. Understanding the dynamics of employee-manager relationships during these critical early years is essential for developing effective retention strategies and fostering a supportive work environment that encourages long-term employee engagement.
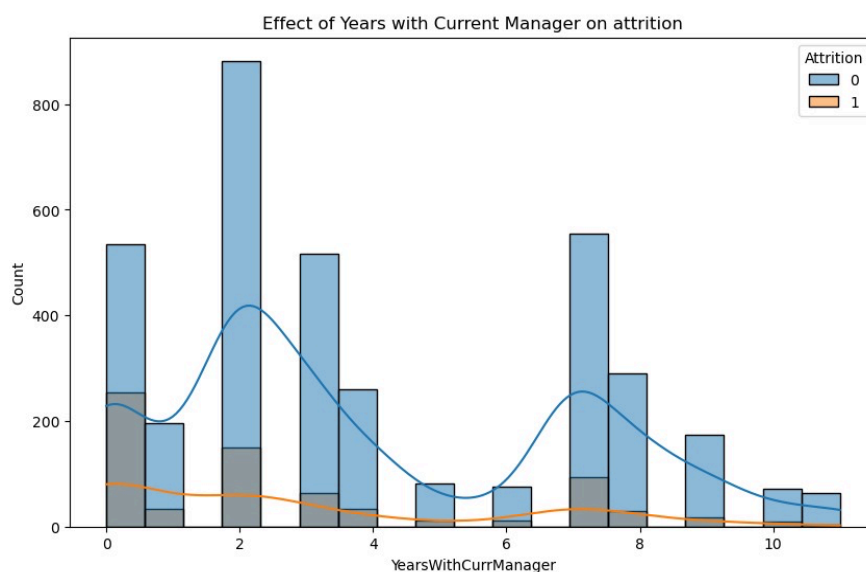
- **Effect of Last Promotion**

The data indicates that 330 employees have experienced attrition within their first year of association with the organization. This high turnover rate during such an early stage of employment suggests that these individuals may have encountered significant challenges that influenced their decision to leave. One notable factor appears to be the limited opportunities for promotion, which were perceived as bleak or nearly negligible. This lack of advancement potential could contribute to feelings of dissatisfaction and disillusionment among new hires, underscoring the importance of clearly defined career paths and growth opportunities in enhancing employee retention from the outset.
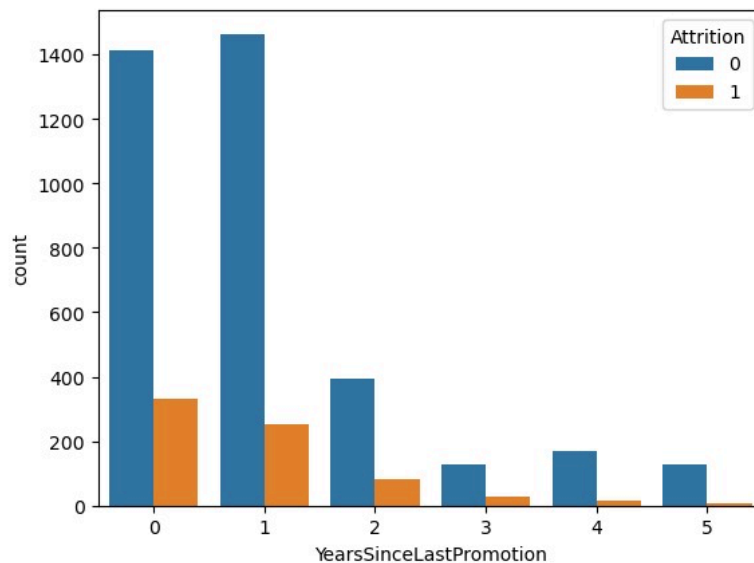

- **Effect of Environment Satisfaction**

The analysis indicates that employees with a poor level of job satisfaction are significantly more affected by attrition, demonstrating a clear link between dissatisfaction and turnover. Interestingly, employees who report a Level 3 satisfaction, which is generally considered good, experience the second highest rate of attrition. This suggests that even those who are moderately satisfied may still be at risk of leaving the organization, potentially due to unmet expectations or a desire for improvement. Understanding the nuances of employee satisfaction and its relationship to attrition is crucial for identifying areas where retention strategies can be enhanced, particularly for those in the "good" satisfaction range who may still be vulnerable to turnover.
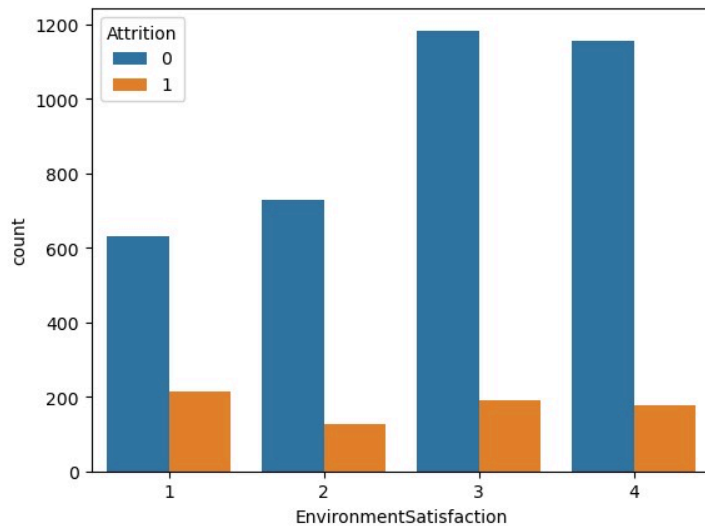
- **Job Satisfaction**

  The analysis reveals an intriguing trend: employees at Level 3, who are more satisfied with their jobs, have been the most affected by attrition. This suggests that even individuals who report relatively high job satisfaction are leaving the organization at significant rates. Surprisingly, those at Level 1, who express dissatisfaction with their jobs, also experience attrition but at a lower rate compared to Level 3. This indicates that while dissatisfaction contributes to turnover, there are other underlying factors influencing the departure of employees who are generally more content with their roles. It emphasizes the need to explore beyond job satisfaction, focusing on elements like career growth, recognition, and work-life balance to better understand and mitigate attrition.
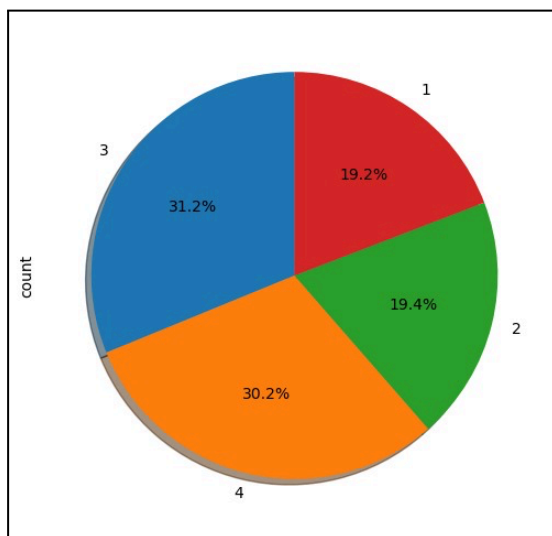
- **Job Involvement**

  The data shows that employees with a higher level of involvement, particularly those categorized at Level 3, have been significantly impacted by attrition, with approximately 399 people leaving the organization. This indicates that even those who are actively engaged and involved in their roles are not immune to turnover.

Despite their commitment, factors such as limited growth opportunities, dissatisfaction with certain aspects of their job, or external career prospects may drive these employees to leave. This highlights the importance of addressing the needs of engaged employees and ensuring that their involvement is met with appropriate rewards and opportunities for advancement to reduce the risk of attrition.
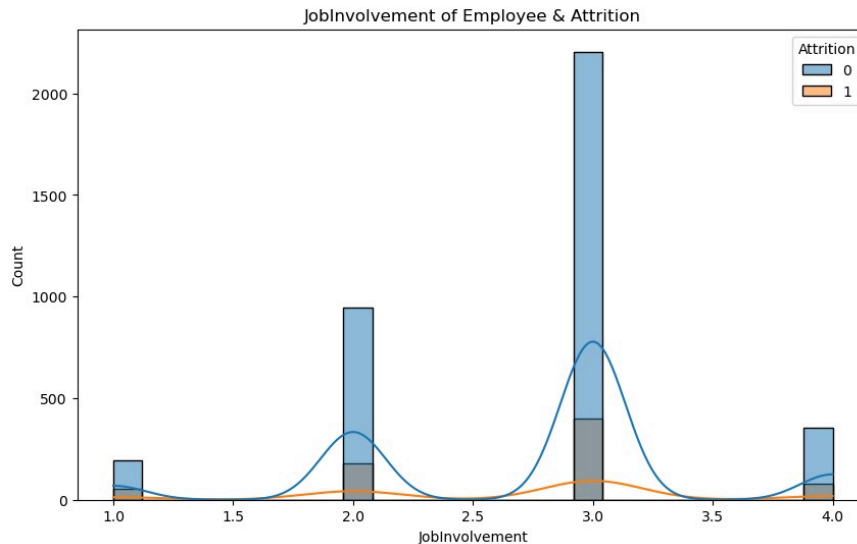
5. Conclusion & Recommendations

Conclusion:

- High Attrition Rate: The company is facing a significant employee attrition rate of 16%, which is impacting overall productivity, morale, and efficiency.

- Most Affected Groups: Employees with mid-level education (Level 3), those in the age group near 30, and those with salaries between $30,000 to $75,000 are the most affected by attrition.

- Business Travel Impact: Employees who travel rarely or frequently for work are more prone to leave, while non-travelers show much lower attrition rates.

- Departmental Challenges: The Research & Development department experiences the highest attrition rate (10.3%), followed by the Sales department (4.6%).

- Job Satisfaction Levels: Interestingly, employees with moderate job satisfaction (Level 3) are leaving at a higher rate than those with lower satisfaction levels, signaling deeper issues beyond immediate dissatisfaction.

Recommendations:

1. Develop Career Growth Opportunities: Improve promotion pathways and growth prospects, particularly for those within their first few years and Level 3 satisfaction employees, to enhance long-term retention.

2. Targeted Retention Programs: Focus retention strategies on the most affected groups (R&D and Sales, age group near 30, mid-level salaries) by addressing specific concerns like career development and workload.

3. Reduce Turnover Through Engagement: Increase job involvement and recognition programs, ensuring that highly involved employees feel valued and have clear pathways to progress.

4. Address Business Travel Fatigue: Consider reevaluating business travel policies to reduce strain on frequently traveling employees, improving work-life balance.

5. Improve Manager-Employee Relations: Provide leadership training to managers to better support employees during their early tenure, particularly in their first two years, which shows high turnover rates.