A report submitted by

Kevin Cherian George

Topic - IBM HR Analytics Employee Attrition & Performance

Submitted on –

Index

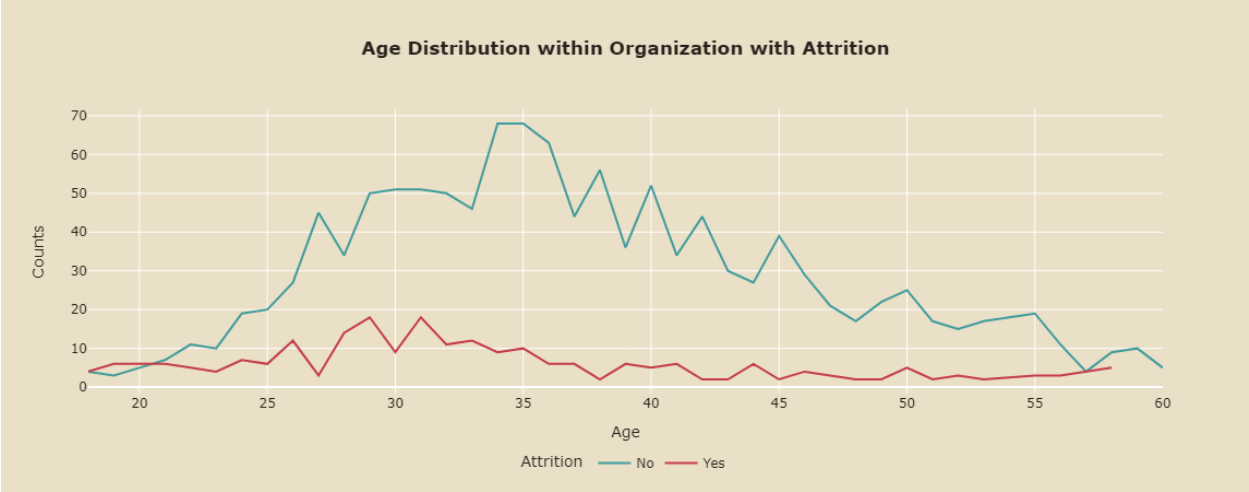
1. Data Loading and Exploration
2. Data Visualization
3. Data Preprocessing
4. Model Development
5. Findings
6. Recommendations

Data Loading and Exploration

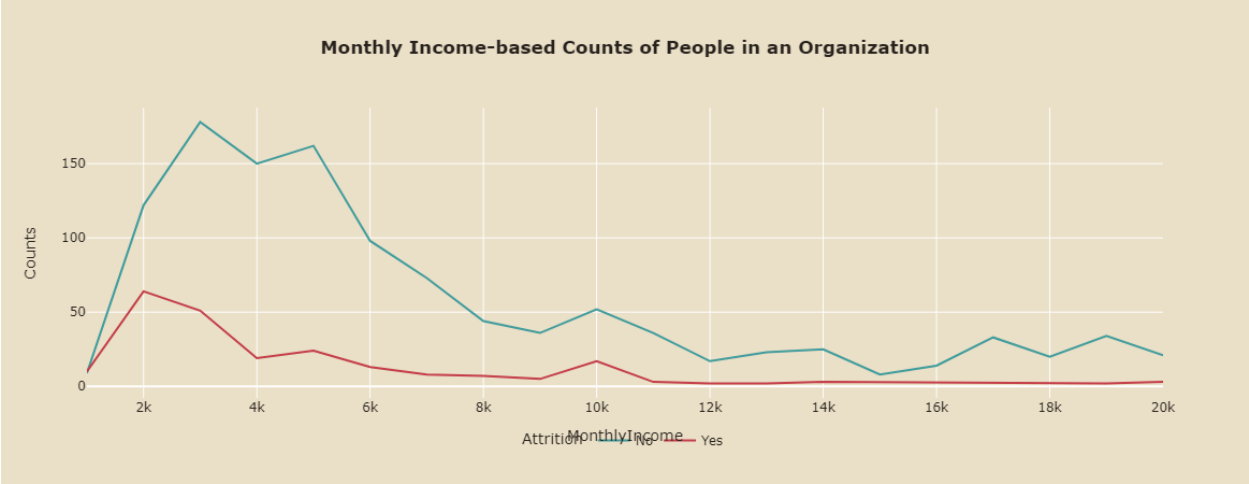
- The code begins by importing the necessary libraries and loading the dataset into a pandas DataFrame.
- It performs initial data exploration by checking the data types, null values, and unique values in categorical and numerical columns.
- Irrelevant columns like 'EmployeeCount', 'StandardHours', 'EmployeeNumber', and 'Over18' are dropped from the dataset.

Data Visualization

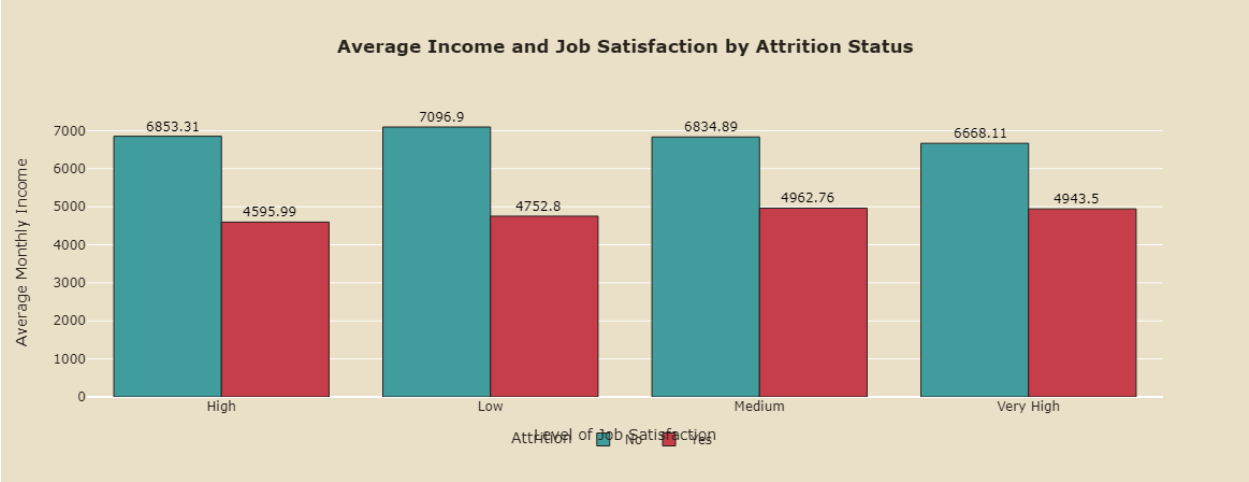
- The code employs several visualization techniques using Plotly and Seaborn to explore the relationships between different variables and employee attrition.
- It examines the age distribution within the organization and its relation to attrition.



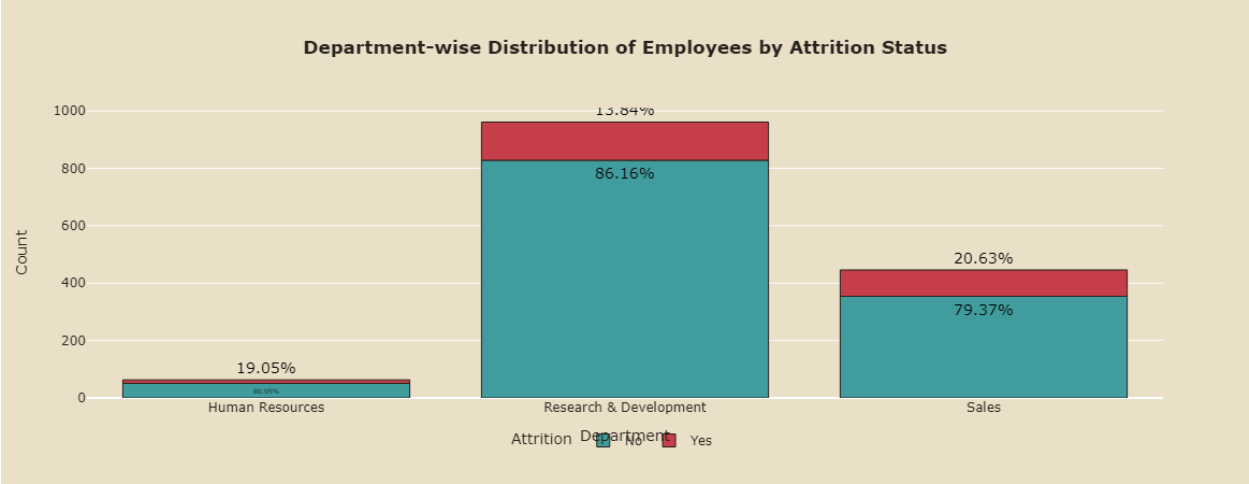
- The influence of monthly income on employee attrition rates is explored.



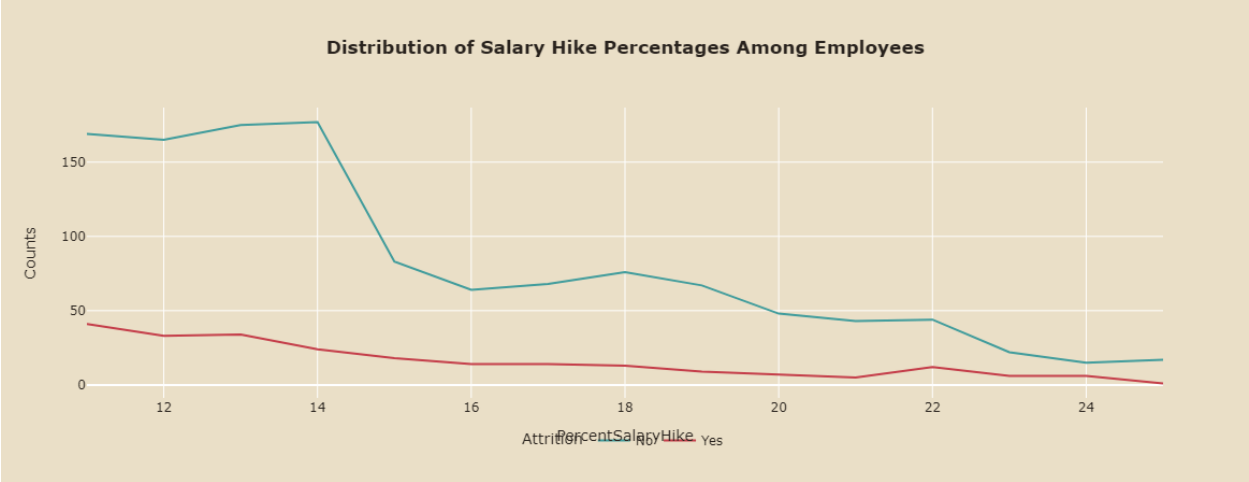
- The relationship between job satisfaction and attrition is visualized, considering the average monthly income.



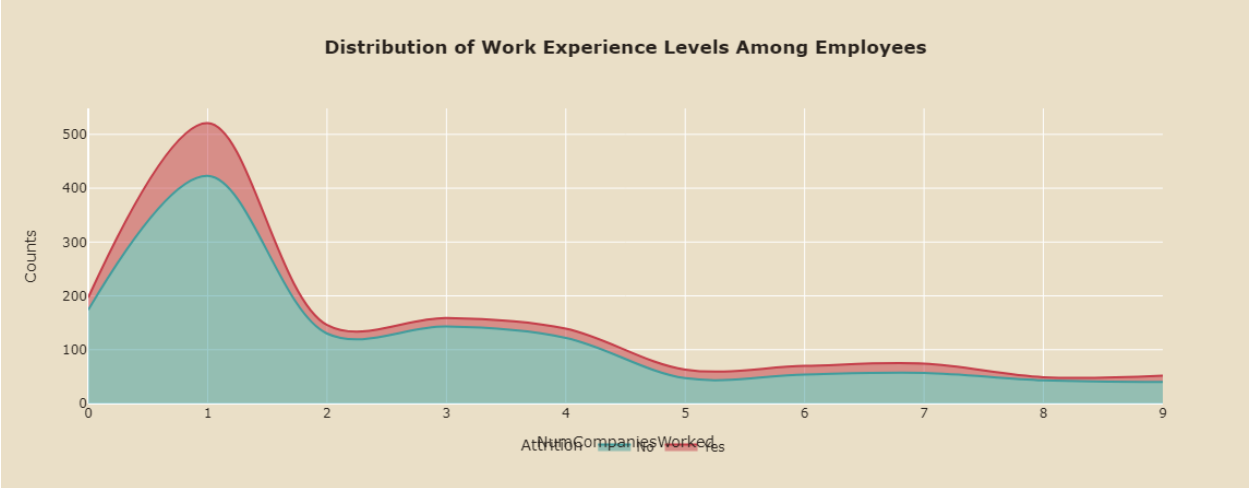
- Department-wise and job role-wise distributions of employees by attrition status are analyzed using bar plots.



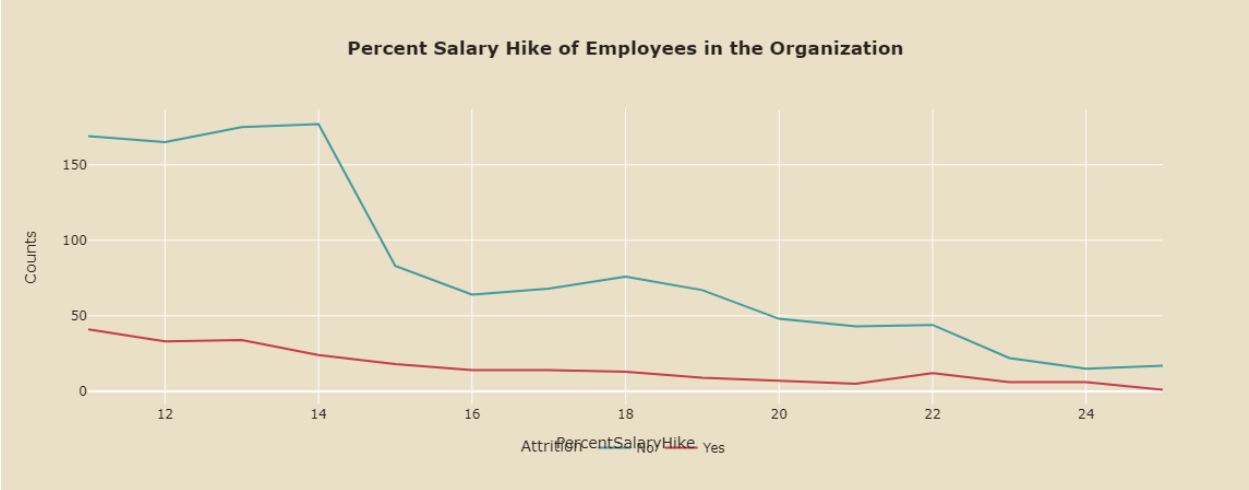
- The impact of salary hikes on employee retention is investigated.



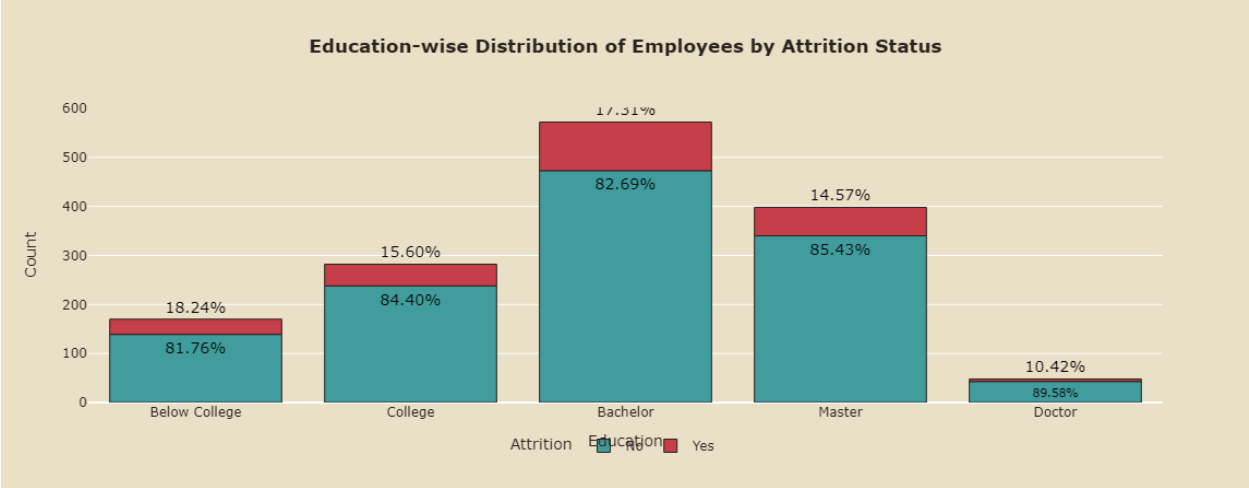
- Work experience diversity among employees is examined using an area plot.



- The effect of salary hikes on employee retention and motivation is visualized using a line plot.



- Attrition rates across different education levels are explored using a bar plot.



Data Preprocessing

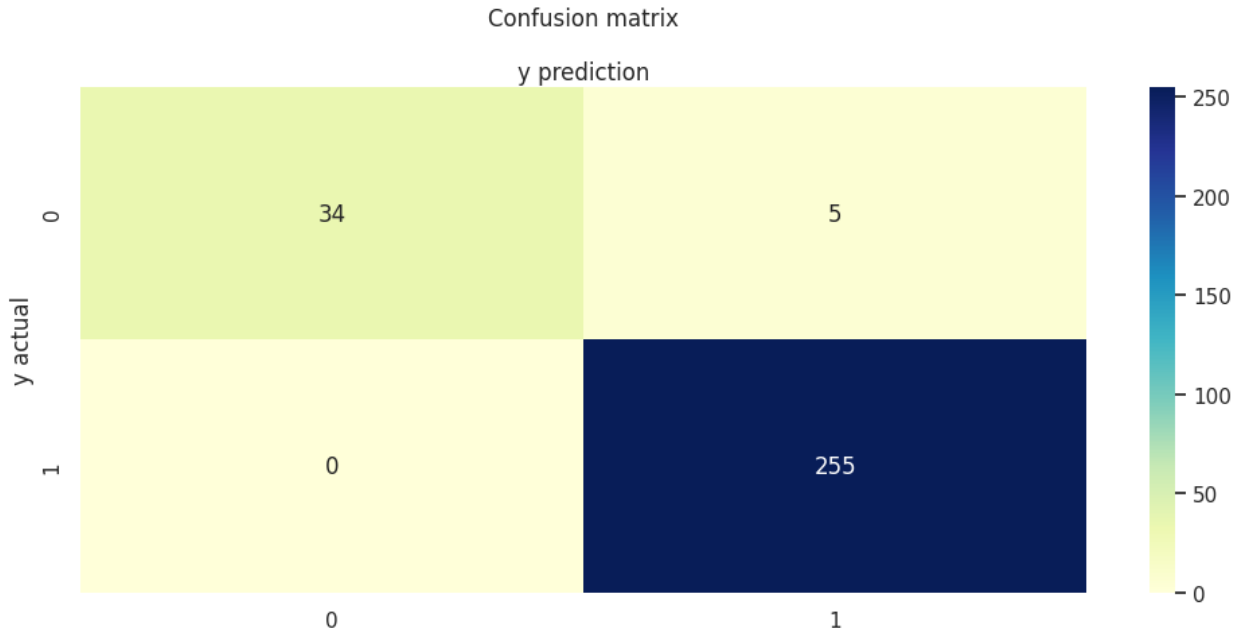
- The code handles categorical variables by converting those with only two distinct values to numeric using label encoding and those with more than two distinct values to numeric using one-hot encoding.
- Correlation analysis is performed to identify highly correlated features.
- Ordinal encoding is applied to categorical variables.
- Numerical features like 'MonthlyIncome', 'YearsInCurrentRole', 'YearsAtCompany', and 'YearsWithCurrManager' are dropped.
- Normalization is performed using MinMaxScaler.
- The data is split into training and testing sets.
- SMOTE (Synthetic Minority Over-sampling Technique) is applied to handle class imbalance in the target variable (Attrition).

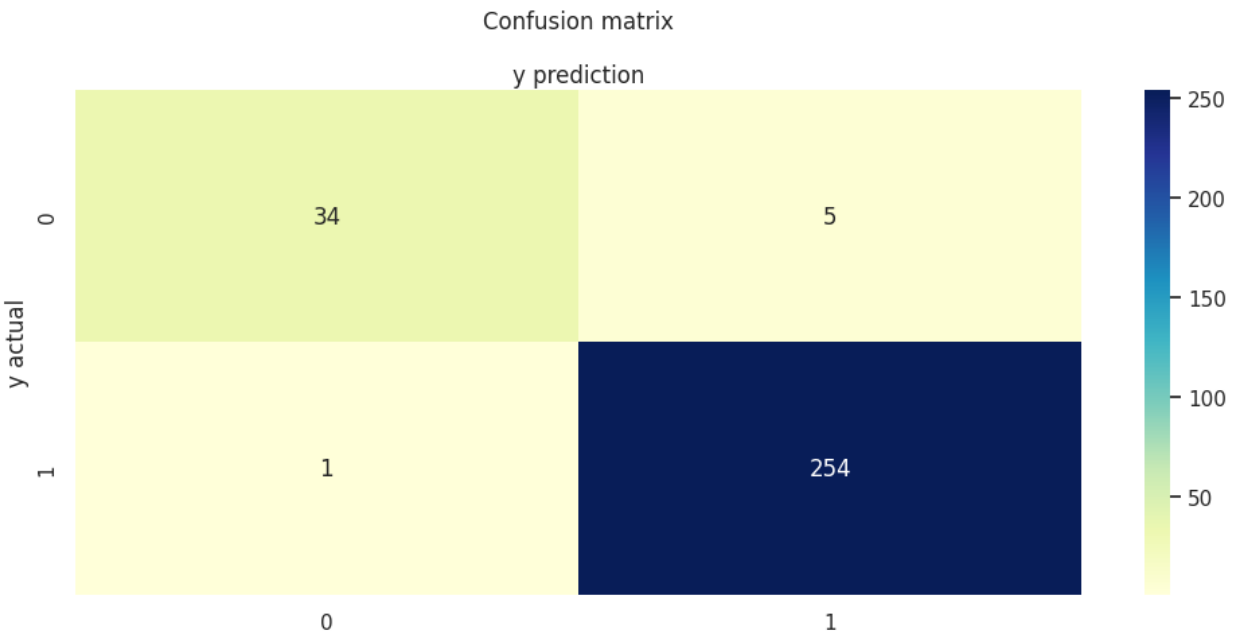
Model Development

- The code trains and evaluates several machine learning models for employee attrition prediction:
  - Logistic Regression
  - Random Forest Classifier
  - Gradient Boosting Classifier
  - AdaBoost Classifier
  - Gaussian Process Classifier

- For each model, the code computes the accuracy score, plots the confusion matrix, and displays the classification report.

- The models' accuracies are compared, and the best-performing models are identified as Random Forest Classifier and Gradient Boosting Classifier with an accuracy of **98%**




Confusion matrix
y prediction
y actual
0
1
34
5
1
254
0
1
250
200
150
100
50

Findings

- Gender Disparity: Males exhibit a higher attrition rate compared to females, hinting at potential disparities in job satisfaction, career opportunities, and workplace environment.

- Age Dynamics: Attrition rates vary across different age groups, with individuals between 28-32 experiencing the highest attrition. This trend declines with advancing age, indicating a shift towards job stability and long-term commitments as individuals progress in their careers.

- Income Levels: Attrition rates are influenced by income levels, with significant spikes observed at very low income levels and a gradual decrease as income rises. This underscores the importance of financial stability in employee retention.

- Job Satisfaction: Lower levels of job satisfaction correlate with higher attrition rates, particularly among employees with average monthly salaries of 4596. Conversely, higher satisfaction levels, especially among those earning 6853, contribute to employee retention.

- Departmental Differences: The Sales department exhibits the highest attrition rate, followed by Human Resources, while Research and Development demonstrate lower rates. This suggests variations in work culture, opportunities, and satisfaction levels across departments.

- Job Role Impact: Higher-level job roles show lower attrition rates compared to lower-level roles, indicating the importance of career advancement opportunities and job stability in retaining talent.

- Salary Increment Influence: Enhanced salary increments serve as a significant incentive for retention, motivating employees to perform better and remain committed to the organization.

- Educational Background: Individuals with higher education levels, such as master's and doctorate degrees, demonstrate lower attrition rates, highlighting the value of specialized skills and advanced qualifications in job satisfaction and retention.

- Salary and Stock Options: Salary and stock options serve as significant motivators for employees, leading to higher loyalty and reduced attrition rates. Employees who receive higher pay and more stock options are more likely to remain committed to their organization, highlighting the importance of competitive compensation packages in retaining talent.

- Work-Life Balance: Work-life balance emerges as a crucial factor influencing employee motivation and retention. While a good work-life balance is often considered a motivation factor, it can also lead employees to seek better opportunities and a higher standard of living elsewhere. Balancing work demands with personal life priorities is essential for maintaining employee satisfaction and reducing turnover.

- Single employees demonstrate a higher rate of departure compared to their married and divorced counterparts.

- Approximately 10% of employees leave upon reaching their 2-year anniversary with the company.

- Employees who are loyal, hold higher salaries, and assume more responsibilities exhibit a lower likelihood of leaving compared to their peers.

- Individuals residing farther away from their workplace exhibit a higher likelihood of leaving compared to those who live closer.

- Employees who frequently travel for work display a higher propensity to leave compared to their counterparts.

- Those required to work overtime demonstrate a higher likelihood of leaving compared to those who do not.

- Sales representatives comprise a significant proportion of leavers within the dataset.

- Employees with a history of working at multiple companies in the past exhibit a higher likelihood of leaving compared to their counterparts.

Recommendations

1. Ensure equal opportunities for all genders, addressing potential disparities in job satisfaction and career advancement.

2. Provide support and development opportunities for employees in their late twenties to early thirties to improve retention during this pivotal career phase.

3. Adjust salary structures to offer competitive compensation, particularly for lower income levels, enhancing financial stability and reducing turnover.

4. Prioritize initiatives to enhance job satisfaction, such as recognition programs and skill development opportunities.

5. Assess and address factors contributing to higher attrition in departments like Sales and Human Resources, improving workload management and work culture.

6. Provide clear career paths and skill development opportunities for lower-level roles to increase job stability.

7. Review and optimize salary increment policies and offer competitive compensation packages, including stock options, to motivate loyalty.

8. Implement policies promoting work-life balance, such as flexible work arrangements and wellness programs, to improve satisfaction and reduce turnover.

9. Address specific factors such as support for single employees, workload management for anniversary dates, and challenges faced by employees who live far or travel frequently.

10. Implementing these strategies will help mitigate attrition rates and retain valuable talent, contributing to long-term organizational success.