# EMPLOYEE ATTRITION ANALYSIS

## NAME: Roy Mariam

## DATE: 19-03-2025

## TOOLS USED: Excel, Python, Tableau

**EXECUTIVE SUMMARY**

**● ANALYSIS OVERVIEW**

This analysis examines employee attrition (**turnover**) to identify key factors

influencing why employees leave an organization.

By understanding attrition trends, companies can develop **strategies to**

**improve retention, reduce costs, and enhance workforce stability.**

**● DATASET OVERVIEW**

This dataset contains 1,470 employees and 35 columns

● **KEY FINDINGS**

Key columns for attrition analysis

• Attrition (Yes / No) ————- Target Variable.

• Department, JobRole, Business Travel, OverTime ———— Categorical

Factors.

• Age, Distance From Home, Monthly Income, Job Satisfaction, Work Life

Balance, Years At Company ————- Potential influences on Attrition.

**KEY INSIGHTS**

• **Overall Attrition Rate**: About **16.2%** of employees leave the company.

• **Department Wise Attrition**: **Sales (20.6%)** has the highest turnover followed

by HR.

• **Work Life Balance & Overtime Impact:** Employees working **overtime** have

**30.5% attrition rate**, indicating overworking and burnout issues.

• **Salary & Job Satisfaction:** Lower **Monthly income** and **poor job**

**satisfaction** significantly contributes to attrition.

• **Age & Experience Factors:** Younger employees and those with **fewer years**

**at the company** are more likely to leave.

**CONCLUSION**

**● BUSINESS IMPACT**

• High turnover leads to **increased recruitment costs, loss of talent, and**

**lower productivity.**

• Identifying high-risk areas helps **HR teams develop targeted retention**

**strategies**.

• Data-driven decisions enable companies to **improve employee satisfaction**

**and reduce turnover rates**.

**● NEXT STEPS**

• Implement **work-life balance policies** to address overtime concerns.

• Offer **competitive salaries and career growth opportunities** to retain

employees.

• Use **Predictive analytics** to identify potential attrition risks early.

**INTRODUCTION**

**EMPLOYEE ATTRITION RATE ANALYSIS**

**● PROBLEM DEFINITION**

Employee attrition (turnover) is a **major challenge** for organizations, leading

to **financial losses, decreased productivity, and operational disruptions.**

High attrition rates can negatively impact company performance, morale, and

customer satisfaction.

● **BUSINESS PROBLEMS**

The organization wants to understand:

• **Why employees are leaving** (Key factors influencing attrition).

• **Which departments, age groups, or job roles are most affected.**

• **How to reduce attrition and improves employee retention strategies**.

● **KEY QUESTIONS TO ANSWER**

• What is the **Overall attrition rate?**

• Which **departments** and **job roles** have the highest turnover?

• How do factors like **overtime, salary, job satisfaction, and work-life**

**balance** influence attrition?

• Can we use **data-driven insights** to predict and reduce future attrition?

**SIGNIFICANCE OF EMPLOYEE ATTRITION ANALYSIS**

**● IDENTIFYING KEY ATTRITION TRENDS**

• Helps HR and management **understand why employees are leaving**.

• Reveals which **departments**, **age groups**, **or job roles** have the highest

attrition rates.

• Allows organizations to focus on **high-risk areas** for employee turnover.

● **IMPROVING EMPLOYEE RETENTION STRATEGIES**

• **Work-Life Balance & OverTime:** High attrition among employees working

overtime suggests a **need to reduce work load** or **improve compensation**.

• **Salary & Job Satisfaction**: Employees with **lower salaries or job**

**satisfaction** leave more often, highlighting **the importance of competitive**

**pay and career growth**.

• **Age & Experience**: Younger employees tend to leave for better opportunities,

suggests the **need for mentorship programs** and **career development**

**plans**.

● **FINANCIAL & OPERATIONAL IMPACT**

• **Hiring new employees is expensive —** costs include **recruitment, training,**

**and lost productivity**.

• Reducing attrition helps **save money** and **improve team stability**.

• High turnover can **affect customer satisfaction** if key employees leave

frequently.

● **DATA- DRIVEN DECISION MAKING**

• Using **data analytics (Power BI, Tableau, Machine Learning)** helps make

**informed HR decisions**.

• Helps predict future attrition risks and proactively address employee

concerns.

• Identifies **which retention strategies are most effective.**

**● ENHANCING WORKPLACE CULTURE**

• By addressing issues like **workload, compensation, and career growth**,

companies can create a **more supportive and engaging work**

**environment.**

• A strong retention strategy **boosts employee morale and productivity.**

**OBJECTIVE OF ANALYSIS**

• Identify **patterns and trends** in employee attrition.

• Provide **data-driven recommendations** to improve retention.

• Help HR team make **informed decisions** about employee satisfaction,

workload, and compensation.

**DATA COLLECTION AND PREPARATION**

**● DATA SOURCES FOR THE EMPLOYEE ATTRITION ANALYSIS**

**● HR INFORMATION SYSTEM (HRIS)**

• The primary source of employee demographic details (e.g. Age, Gender,

Department).

• Tracks **employee status** (active or terminated) and **attrition** data (Yes/No).

**● PAY ROLL RECORDS**

• Provides **Monthly Income, Hourly Rate,** and **Salary Hike** information.

• Used to understand how compensation relates to attrition.

**● EMPLOYEE SURVEYS**

• Captures **Job Satisfaction, Work-Life Balance, Environment Satisfaction,**

etc.

• Offers insights into **employee sentiment** and reason for leaving.

● **TIME & ATTENDANCE SYSTEMS**

• Includes **OverTime** hours and **Standard Hours** worked.

• Identifies potential **burnout** or **workload** issues affecting turnover.

● **INTERNAL PERFORMANCE RECORDS**

• Tracks **Performance Rating, Job Involvement, Years At Company,** etc.

• Helps assess **career progression** and **employee engagement levels**.

Consolidating these sources into a single data.

**HR-Employee-Attrition-All.csv sourced from GitHub.com.**

This analysis provides a comprehensive view of the factors influencing employee attrition.

**DATA CLEANING**

**● Column Consistency:** Verified that columns such as ‘Attrition’, ‘OverTime’,

and ‘Department’ had consistent naming.

● **Removing Irrelevant Columns**: Some columns (e.g., ’Employee Count’,

‘Standard Hours’, if they’re the same for all employees) might not provide

value and can be dropped or reduce noise.

● **Checking for Duplicates**: Ensured there were no exact duplicate rows

representing the same employee record.

**DATA TRANSFORMATION**

**● Converting Categorical Data:**

• Applied **Label Encoding** to convert columns like ‘Attrition’ (Yes/No) and

‘OverTime’ (Yes/No) into numeric (1/10).

• Standardized any inconsistent string entries (‘yes’, YES’) into a common

format.

● **Feature Creation**:

• Created an **AgeGroup** column to group employees into age ranges (e.g. 18-

25, 26-35).

**● Handling Missing Values**

• **Initial Inspection**: Checked for missing Values in each column using

methods like df. (). sum ()

• **Result**: The dataset had **no missing values** (all columns fully populated). If

missing values were found, common strategies include:

• Dropping Rows/ Columns if missing values are minimal.

• Imputing (filling) missing values with the mean/median for numeric columns

or the mode for categorical columns.

● **Standardization**: If any numeric columns had missing values after imputation,

they were re-checked to ensure no NaNs remained before modeling.

By performing these steps, the final dataset is consistent, numeric where

needed, and free of missing values, ensuring accurate analysis and robust

modeling for attrition insights.

**KEY ASSUMPTIONS IN THE EMPLOYEE ATTRITION ANALYSIS**

**● DATA ACCURACY AND COMPLETENESS**

• The dataset is assumed to be **accurate, up-to-date, and representative** of

the entire employee population.

• No major **data entry errors or missing values** significantly impacting

analysis.

● **VOLUNTARY VS. INVOLUNTARY ATTRITION**

• The analysis assumes that **all attrition is voluntary**, meaning employees

left due to dissatisfaction, workload, or better opportunities.

• If involuntary attrition (e.g., layoffs, termination) is included, it may distort the

findings.

● **CASUAL VS CORRELATION RELATIONSHIPS**

• The analysis identifies **correlation** between factors (e.g. low Job Satisfaction

and high attrition) but does not confirm **causation.**

• External Factors (e.g. economic downturn, Job market trends) could also

influence attrition.

● **UNIFORM WORK CONDITIONS**

• Assumes that employees within the same **department and job role** have

similar work conditions (e.g. Workload, pay structure).

• Differences in **management style or team culture** are not explicitly

considered.

● **PREDICTIVE MODEL ASSUMPTIONS**

• Machine learning models assume **linear relationships** between features like

**salary, Job satisfaction, and attrition** unless non-linear techniques are

applied.

• Model performance depends on the dataset and **may not generalize**

**perfectly** to future employees.

**METHODOLOGY**

**ANALYTICAL APPROACH FOR EMPLOYEE ATTRITION ANALYSIS**

In this analysis, I followed a structured approach to **understand, visualize,**

**and predict** employee attrition.

**● DATA EXPLORATION AND CLEANING**

**● Objective:** Ensured data Was clean, structured and ready for analysis.

● **Steps:**

• Loaded the dataset and inspected column names, data types, and summary

statistics.

• Handling **missing values, duplicate records, and inconsistencies.**

• Converted categorical variables (e.g. ‘Attrition’, ‘OverTime’) into numerical

values using **Label Encoding**.

• Standardized / normalized numeric columns in places needed for modeling.

● **DESCRIPTIVE ANALYSIS (UNDERSTAND ATTRITION TRENDS)**

**● Objective:** Identified key trends and patterns in employee attrition.

● **Techniques: Attrition rate calculation**

**Attrition rate = (Employees who left)**

**——————————— X 100**

**(Total Employees)**

**● Attrition by Department:** Checked for which departments have the highest

turnover.

● **Attrition by Age & Experience:** Identified whether younger employees or

those with fewer years at the company left more often.

● **Attrition by OverTime & Work-LifeBalance:** Examined if working overtime contributes to higher attrition.

● **Tools used: Python (Pandas, Matplotlib, Seaborn), Tableau.**

**● STATISTICAL ANALYSIS**

**● Objective:** Validating findings with statistical tests.

● **Techniques:**

• **Chi- Square Test:** Used to check if categorical variables (e.g. OverTime,

Attrition) are significantly related.

• **T-Test / ANOVA:** Used to compare attrition rates across different salary levels and job roles.

● **Tools used: Python (SciPy, Stats models).**

**● PREDICTIVE MODELING (MACHINE LEARNING)**

**● Objective:** Built a model to predict which employees are likely to leave.

● **Steps:**

• Defined **Independent variables (X) (e.g. Job Satisfaction, Work-Life**

**Balance, Salary).**

• Defined **dependent variables (y) (Attrition: Yes =1, No =0).**

• Split data into **training and test sets** (80% training, 20% testing).

• Trained a **Logistic Regression model.**

• Evaluated using **accuracy, precision, recall and F1-score.**

**● DATA VISUALIZATION & REPORTING**

**● Objective:** Presented findings in a clear, actionable format.

● **Deliverables:**

• **Tableau Dashboard:** Interactive visuals showing attrition trends.

• **PDF /Word Report:** Summarized key insights, statistical analysis and

recommendations.

**● Tools used: Tableau, Python (Matplotlib, Seaborn)**

**● BUSINESS RECOMMENDATIONS**

**● Objective:** Provided data-driven solutions to reduce attrition.

● **Key Focus Areas:**

• Improve **Work-Life Balance** policies to reduce overtime driven attrition.

• Offer **Competitive salaries** and career growth opportunities.

• Enhance **Job Satisfaction** through employee engagement programs.

**KEY PERFORMANCE INDICATORS (KPIs) AND**

**METRICS USED**

**● ATTRITION RATE (PRIMARY KPI)**

• **Definition:** The percentage of employees who leave the organization over a

given period.

• **Formula:**

**Attrition rate = (Number of employees who left)**

**————————————————— X 100**

**(Total Employees)**

• **Significance:** Helps HR understand the **Overall turnover** in the company

and whether it is increasing or decreasing.

● **DEPARTMENT- WISE ATTRITION RATE**

• **Definition:** The percentage of employees who leave with each department.

• **Formula:**

**Attrition Rate by department = (Employees who left each dept)**

**———————————————— X100**

**(Total employee)**

• **Significance:** Identified which **departments** are most affected by the attrition

allowing targeted retention efforts.

● **ATTRITION BY TENURE (YEARS AT COMPANY)**

• **Definition:** Tracked the percentage of employees leaving based on their

years of service.

• **Significance:** Assesses whether new hires leave quickly **(onboarding**

**issues)** or if long- term employees are dissatisfied.

● **ATTRITION BY WORK-LIFE BALANCE & OVERTIME**

• **Definition:** Measured the correlation between **workload and employee**

**turnover.**

• **Metrics used:**

• % of employees working overtime who left.

• Average work-life balance score for employees who stayed vs who left.

• **Significance:** Determined if **overtime and poor work-life balance** are

major factors in attrition.

● **JOB SATISFACTION & ATTRITION CORRELATION**

• **Definition:** Analyzed how Job satisfaction levels impact attrition rates.

• **Metrics used:**

• % of employees with **low job satisfaction** who left vs those who stayed.

• **Significance:** Helps HR improve **employee engagement** and **work place**

**policies.**

**● SALARY & ATTRITION RELATIONSHIP**

• **Definition:** Examined whether **low salary levels contribute to higher**

**turnover.**

• **Metric used:**

• Average salary of employees who left vs those who stayed.

• Salary distribution by attrition status.

• **Significance:** Determined if **compensation adjustments** are needed to

improve retention.

● **PREDICTIVE MODEL PERFORMANCE**

**Metrics used:**

• **Accuracy Score:** Measured how well the model predicts attrition.

• **Precision & Recall:** Evaluated how accurately the model identifies

employees likely to leave.

• **Feature importance:** Identified the **top factors** contributing to attrition.

• **Significance:** Enables HR to proactively **identify employees at risk of**

**leaving.**

**STATISTICAL TECHNIQUES & MODELS APPLIED IN**

**EMPLOYEE ATTRITION ANALYSIS**

Used to extract insights and predict employee attrition, various **statistical**

**techniques and machine learning models** were applied.

● **DESCRIPTIVE STATISTICS (INITIAL DATA ANALYSIS)**

• **Purpose:** Summarized key trends in the dataset before deeper analysis.

• **Techniques used:**

• **Mean, median and standard deviation:** Understanding salary distribution,

years at company and job satisfaction levels.

• **Attrition rate calculation:** Proportion of employees who left.

• **Frequency Distributions:** Checking how attrition varies by department,

tenure, salary and work- life balance.

• **Significance: Identifying patterns and high-risk groups** for attrition.

● **INFERENTIAL STATISTICS (HYPOTHESIS TESTING)**

• **Purpose:** Identify significant factors affecting attrition.

• **Techniques used:**

• **Chi - square test (for categorical variables)**

• **Used for:** Checking relationships between categorical features (e.g. Attrition,

Department, Overtime, Job Satisfaction).

• **Hypothesis:**

• Null (H0): No relationship between the variables.

• Alternative (H1): There is a significant relationship.

• **T-test / ANOVA (For Numeric Variables)**

• **Used for:** Comparing the means of numerical variables across attrition

groups.

• **Significance:** These tests **confirm whether specific factors (e.g. salary,**

**job satisfaction, overtime) truly influence attrition** rather than being

random patterns.

**●** **CORRELATION ANALYSIS**

• **Purpose:** Measure relationships between numerical variables and attrition.

• **Technique used:** Pearson correlation coefficient.

• **Significance:** Identifies whether high salary, job satisfaction or experience

**reduce** attrition risk.

● **PREDICTIVE MODELING (MACHINE LEARNING FOR ATTRITION**

**PREDICTION)**

• **Purpose:** Built a model to **predict which employees are at a risk of leaving.**

• **Steps taken:**

• **Defined features & Target variable:**

**Independent variable (X):** Factors like salary, Overtime, years at

company, job satisfaction.

**Dependent variable (y):** Attrition (1= Yes, 0 = No)

• **Data preprocessing:**

**Label Encoding** Categorical Values (e.g. Overtime: Yes = 1, No = 0)

**Standardization** for numerical values.

• **Splitting Data:**

80% Training, 20% Testing

(train\_test\_split).

● **MODELS APPLIED**

• **Logistic Regression (Baseline Model)**

Best for binary classification problems like attrition (Yes/No).

Estimates the probability of an employee leaving based on input factors.

Coefficients that show which factors **increase or decrease attrition**

**likelihood.**

• **Random Forest (Alternative Model for higher Accuracy)**

More robust for non-linear relationships between features.

Uses multiple decision trees to classify employees as likely to leave or stay.

Feature importance scores - shows which variables have the most impact on

attrition.

• **Model Evacuation metrics:**

**Accuracy score:** Measures how well the model predicts attrition.

**Precision & Recall:** Ensures the model correctly identifies employees at risk

of leaving.

**Confusing Matrix:** Breaks down corrects vs incorrect predictions.

• **Significance:** Enables HR team to **proactively identify high-risk**

**employees** and take action before they leave.

**DATA ANALYSIS AND FINDINGS**

**SUMMARY TABLE**

|  |  |
| --- | --- |
| **Metrics** | **Values** |
| **Total Employees** | 1,470 |
| **Attrition Rate** | 16.12% |
| **Highest Dept. Attrition** | Sales (20.63%) |
| **Lowest Dept. Attrition** | R&D (13.84%) |
| **Avg. Monthly income (Left)** | $4,787 |
| **Avg. Monthly Income (stayed)** | $6,833 |
| **High Overtime Attrition** | 30.53% (OverTime=Yes) |

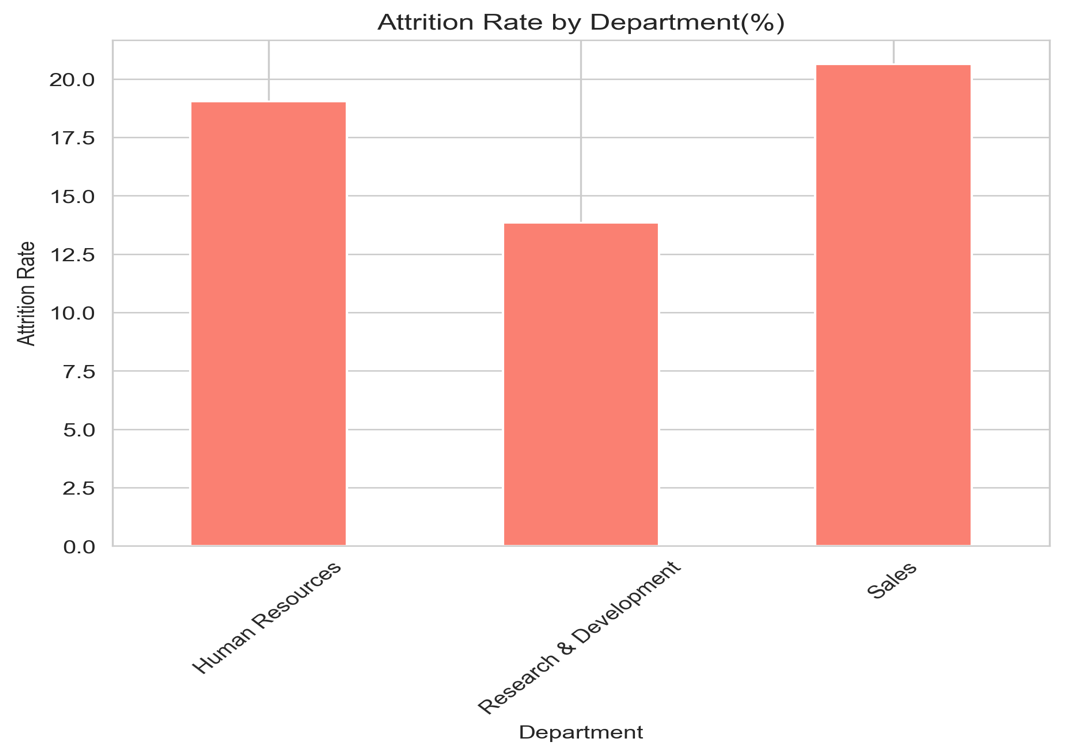
**KEY OBSERVATIONS:**

• Overall, **16.12%** of employees left.

• **Sales** has the highest attrition, **R&D** the lowest.

• Employees who left earn **less** on average and **often work overtime.**

**ATTRITION BY DEPARTMENT (BAR CHART)**

****

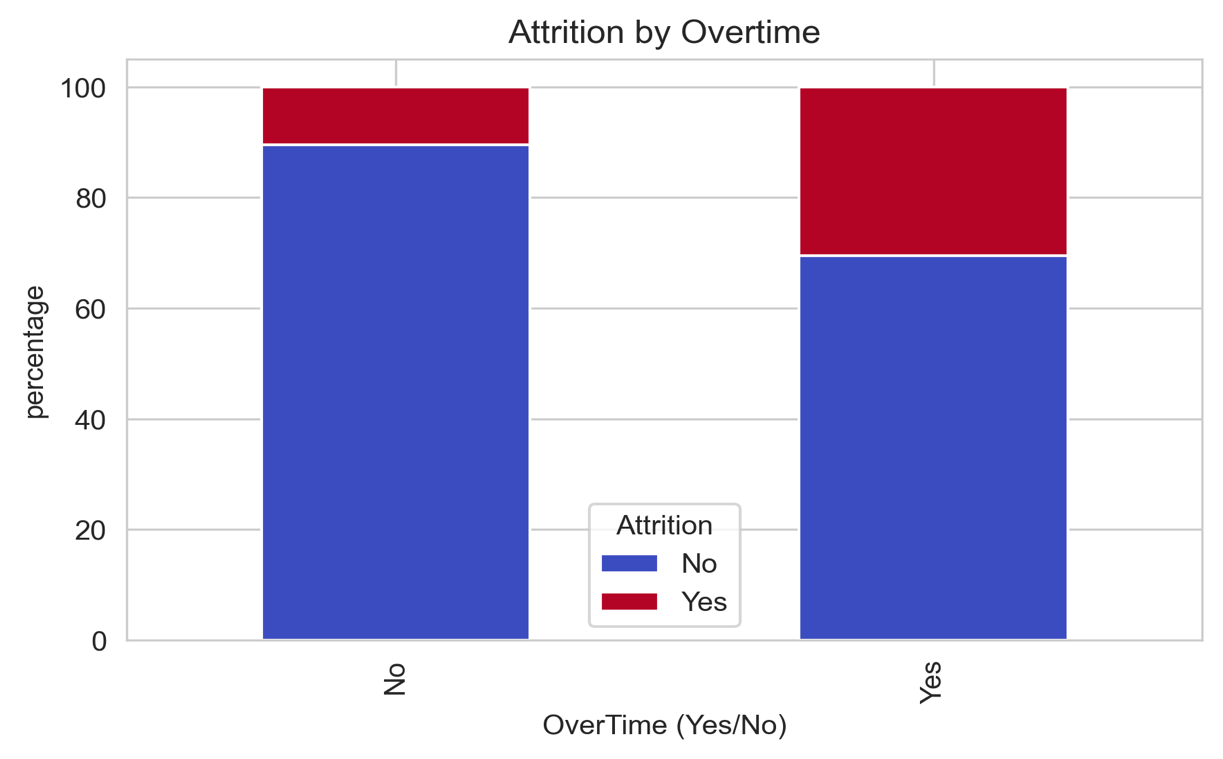
**INSIGHT:**

• **Sales** department leads with a **20.63%** attrition rate.

• **Human Resources** follows at **19.05%**.

• **R&D** has the lowest attrition rate **(13.84%)**.

**OVERTIME VS. ATTRITION (STACKED BAR CHART)**

****

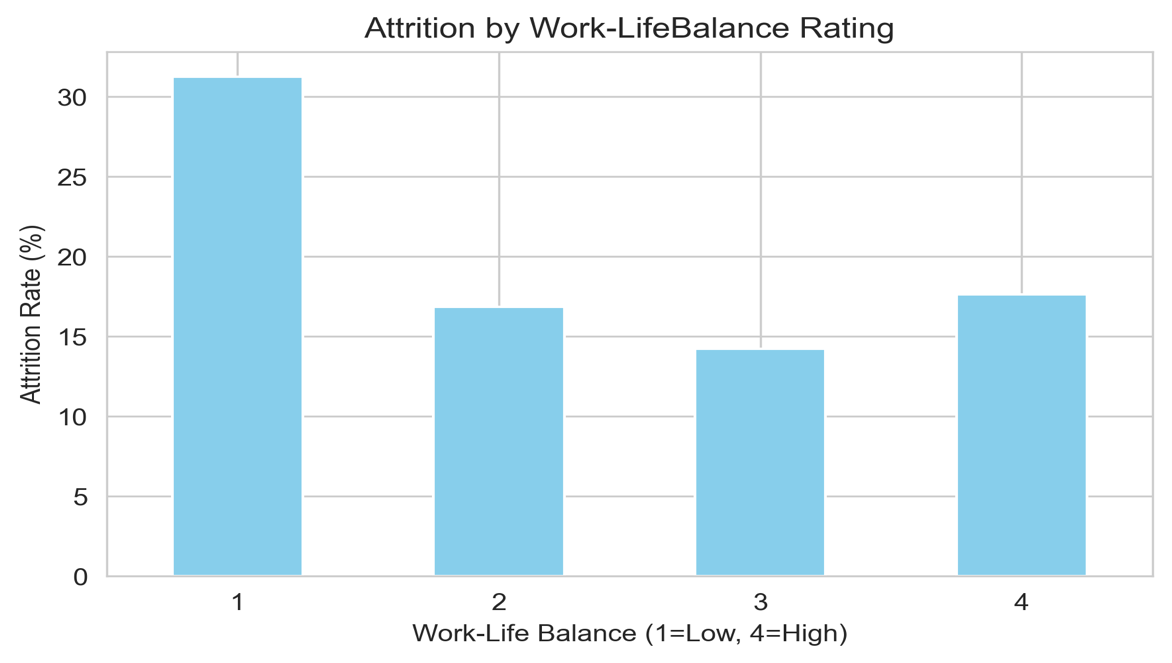
**INSIGHT:**

• **No overtime:** 10.44% attrition.

• **Yes overtime:** 30.53% attrition.

• Indicates **burnout** or **workload** issues contributing to turnover.

**WORK LIFE BALANCE & ATTRITION (BAR CHART)**

****

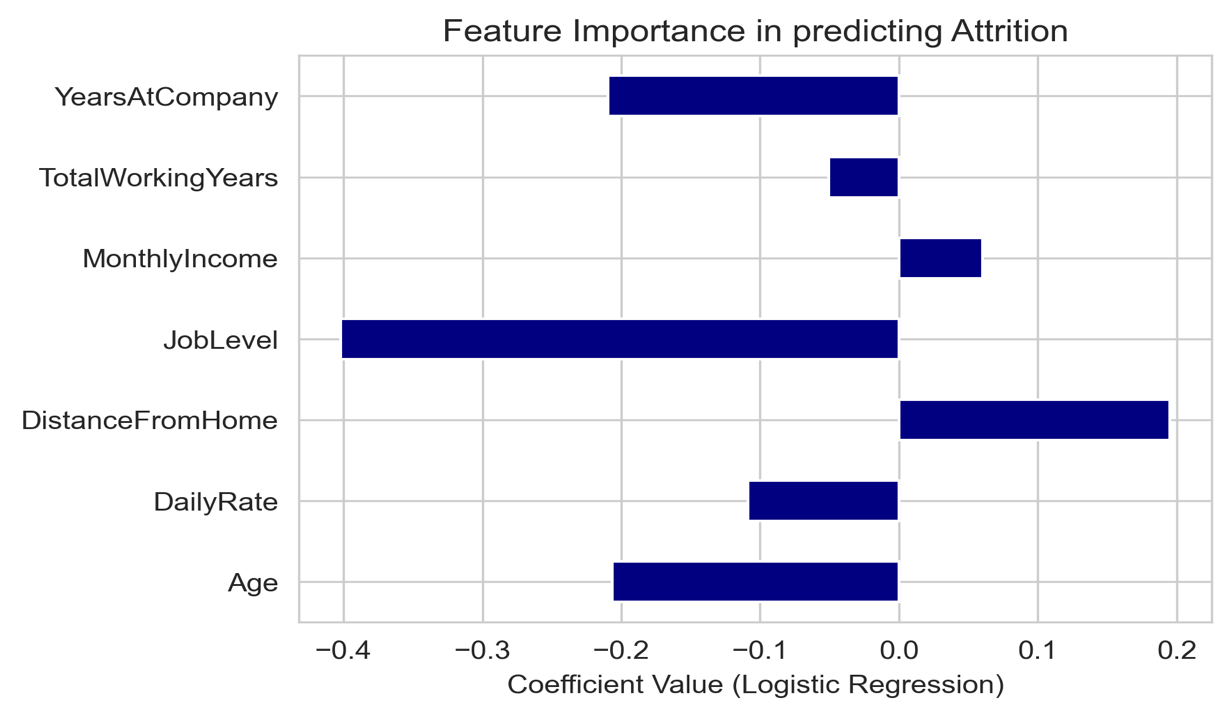
**INSIGHT:**

• **Rating = 1 (poor balance):** 31.25% attrition.

• **Rating = 4 (Good balance):** 17.65% attrition.

• Emphasizes the **need for flexible schedules and reduced overtime.**

**FEATURE IMPORTANCE (HORIZONTAL BAR CHART)**

****

**INSIGHT:**

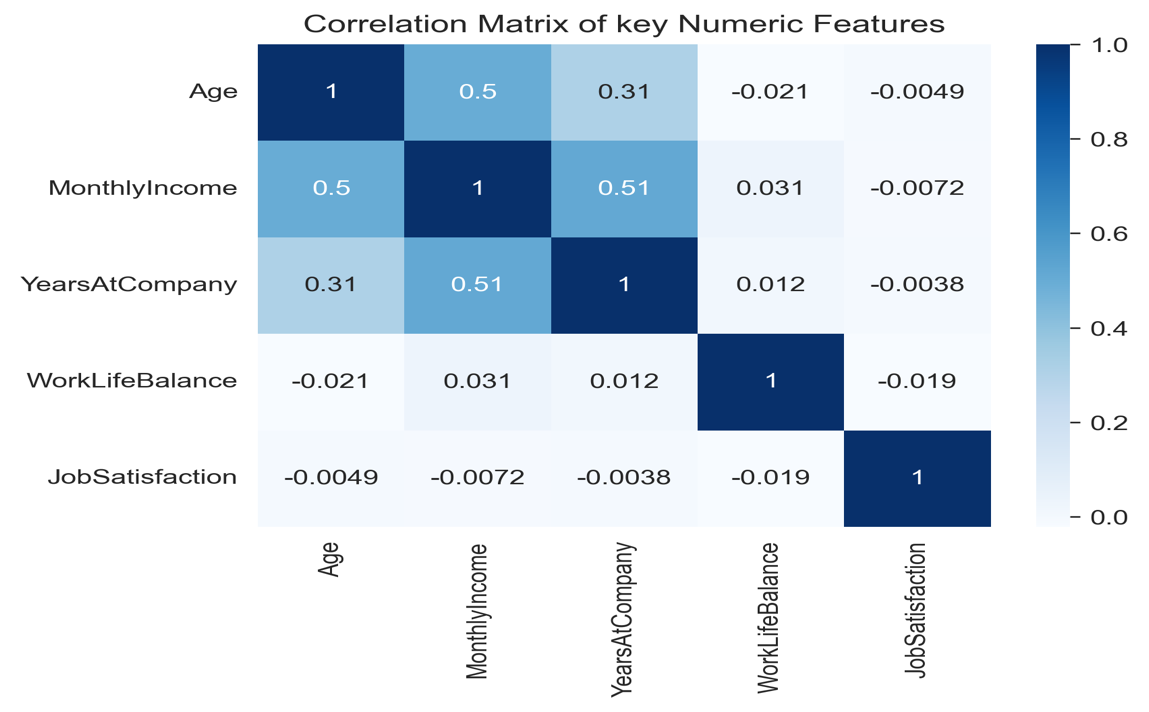
• **OverTime** is the strongest positive prediction of attrition.

• **MonthlyIncome, Job Satisfaction, Work-Life Balance** have **negative**

**coefficients** meaning **improvements** in these areas **reduce** attrition risk.

**CORRELATION HEAT MAP**

A **heat map** can quickly show how different numeric factors relate to each other (e.g, Age, MonthlyIncome, YearsAtCompany).



**INSIGHT:**

• Positive correlation between **Age** and **YearsAtCompany** (older employees

typically have longer tenure)

• Negative correlation between **Work-Life Balance** and **OverTime,** indicating

employees with high overtime often have poor work life balance.

**KEY TAKEAWAYS**

• **Sales & Overtime** are major drivers of attrition.

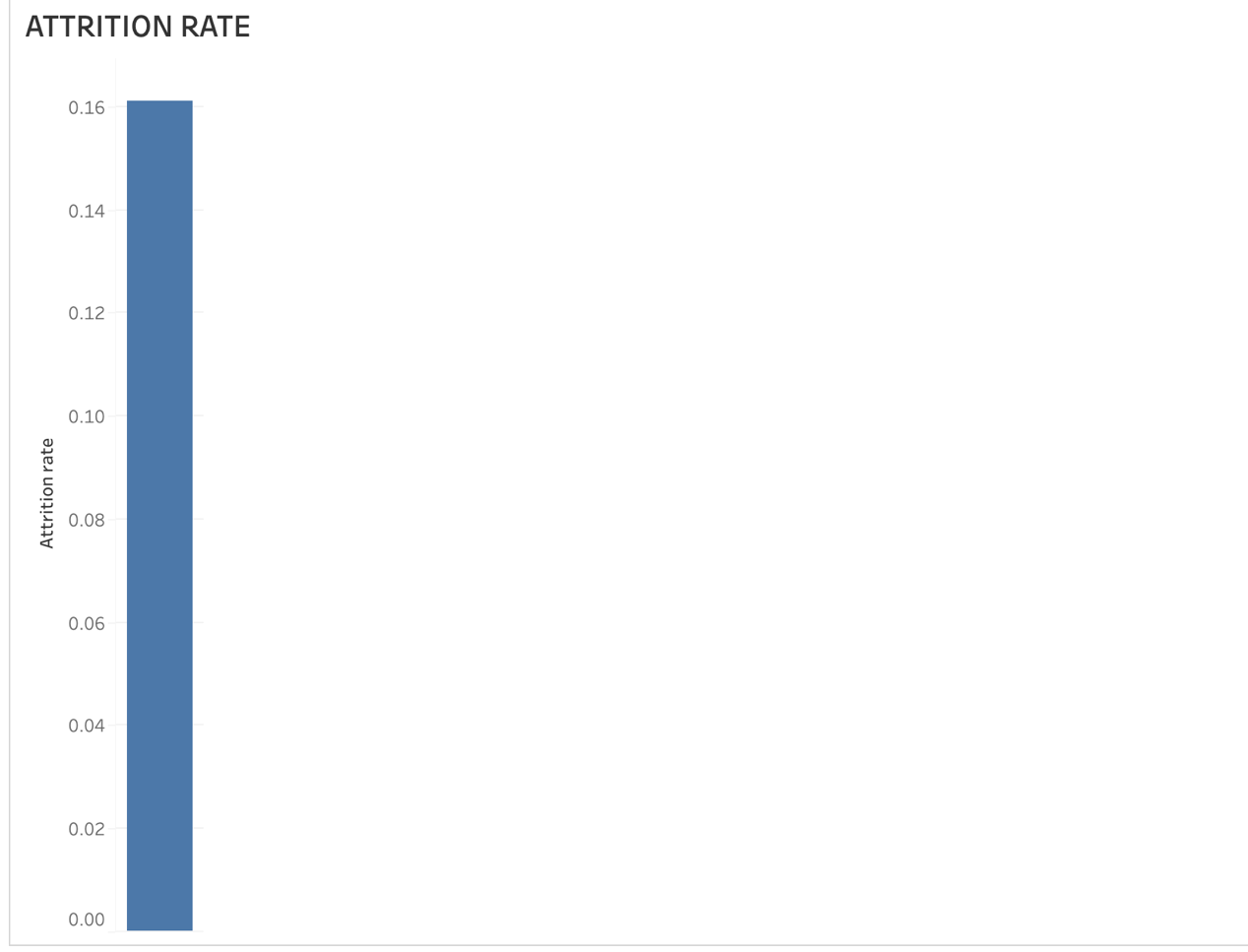
• **Lower salaries and job satisfaction** also increase turnover.

• **Youngest employees** (18-25) have the **highest** attrition rate.

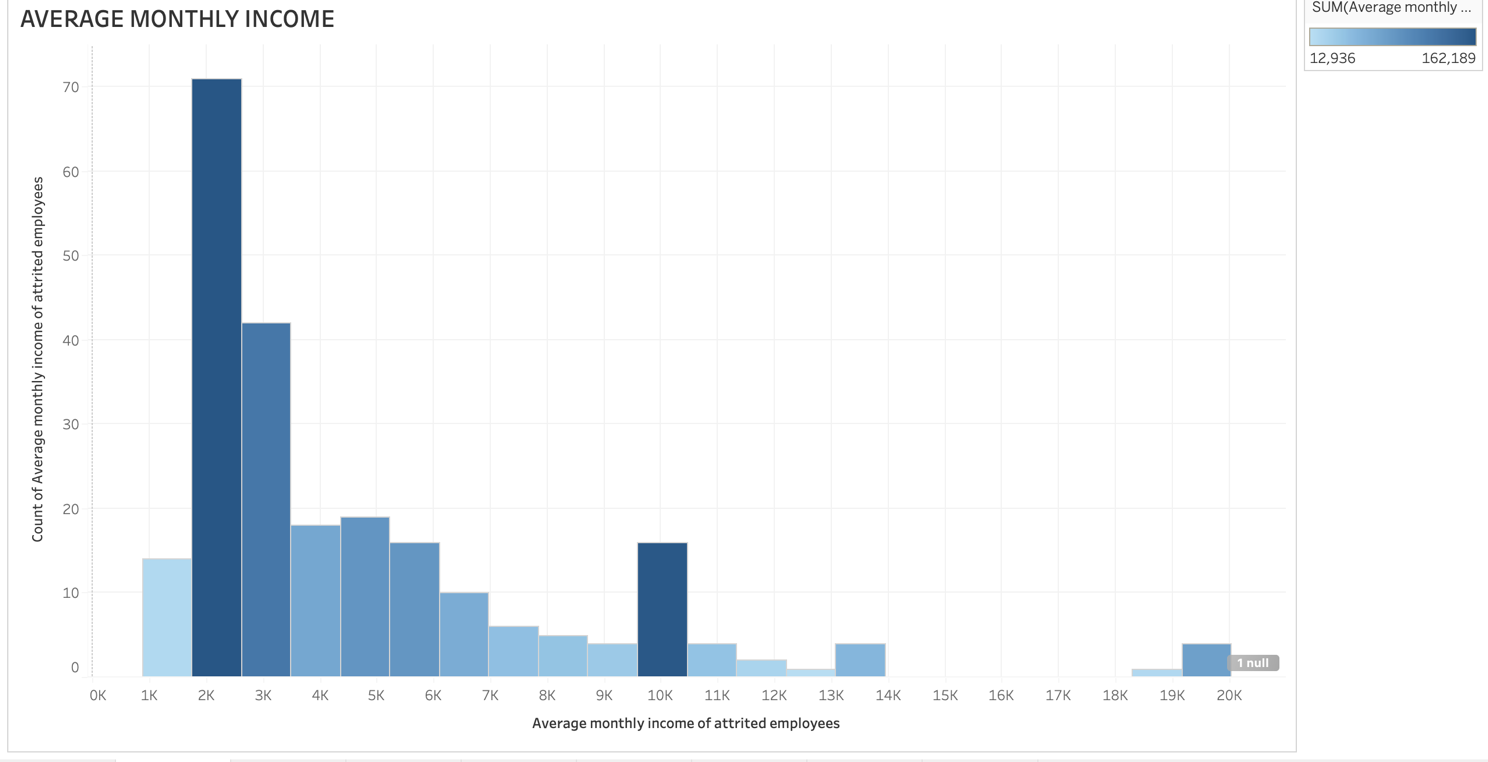
• **Improving work- life balance** (especially for overtime workers) could

**significantly reduce** attrition.

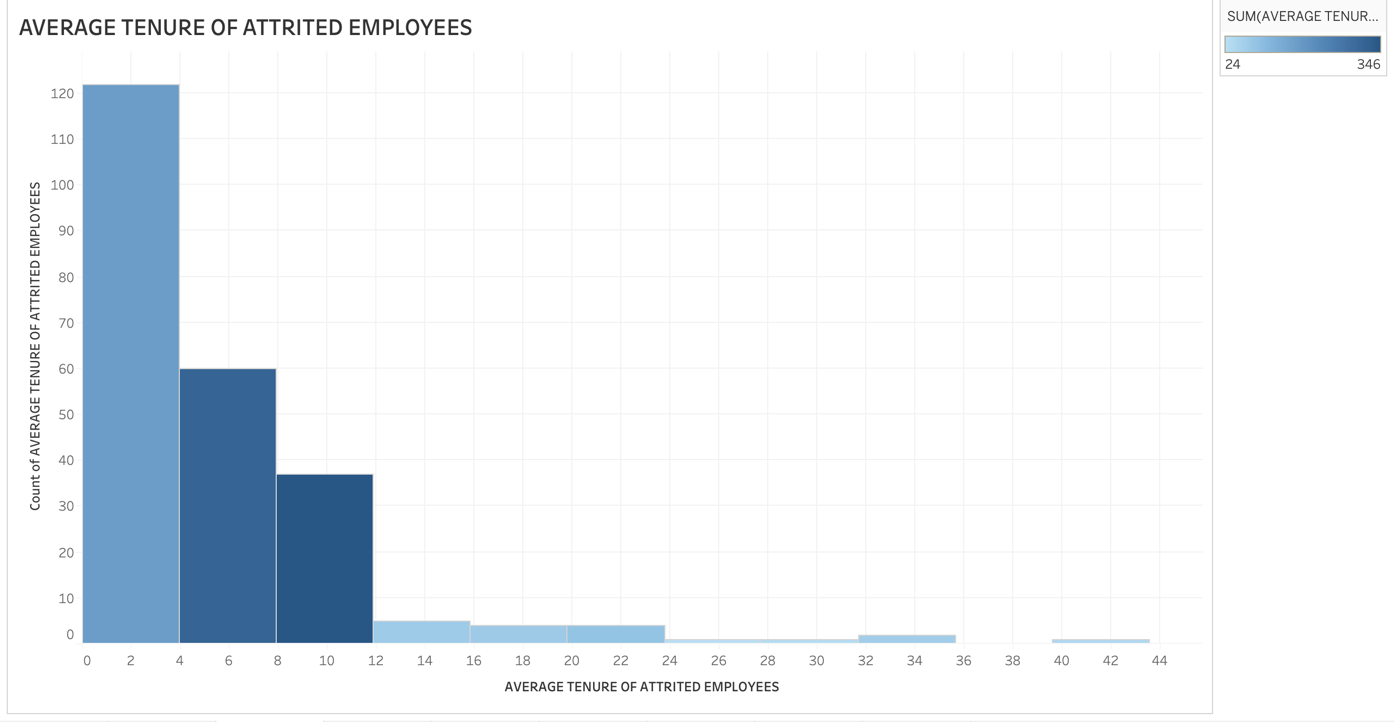
**DATA VISUALIZATION**

****

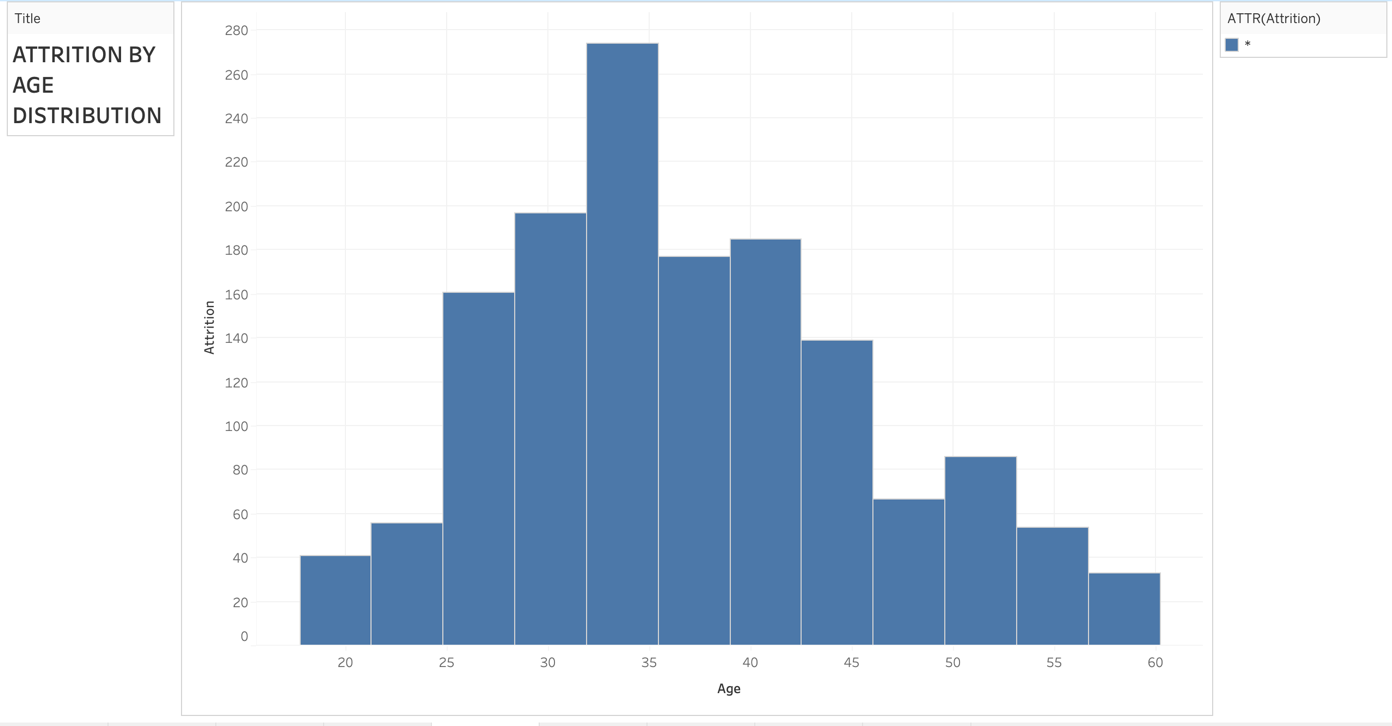
**ATTRITION RATE**

****

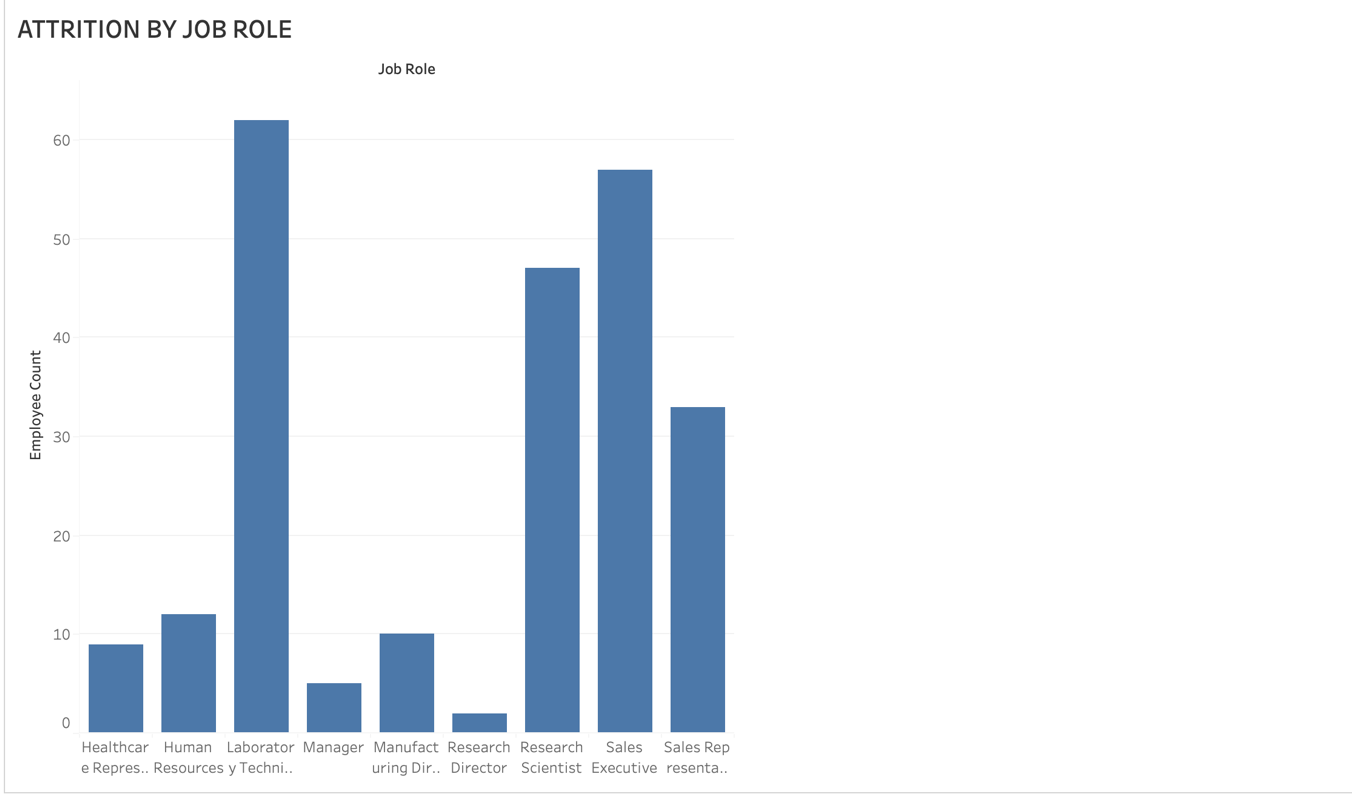
**AVERAGE MONTHLY INCOME**

****

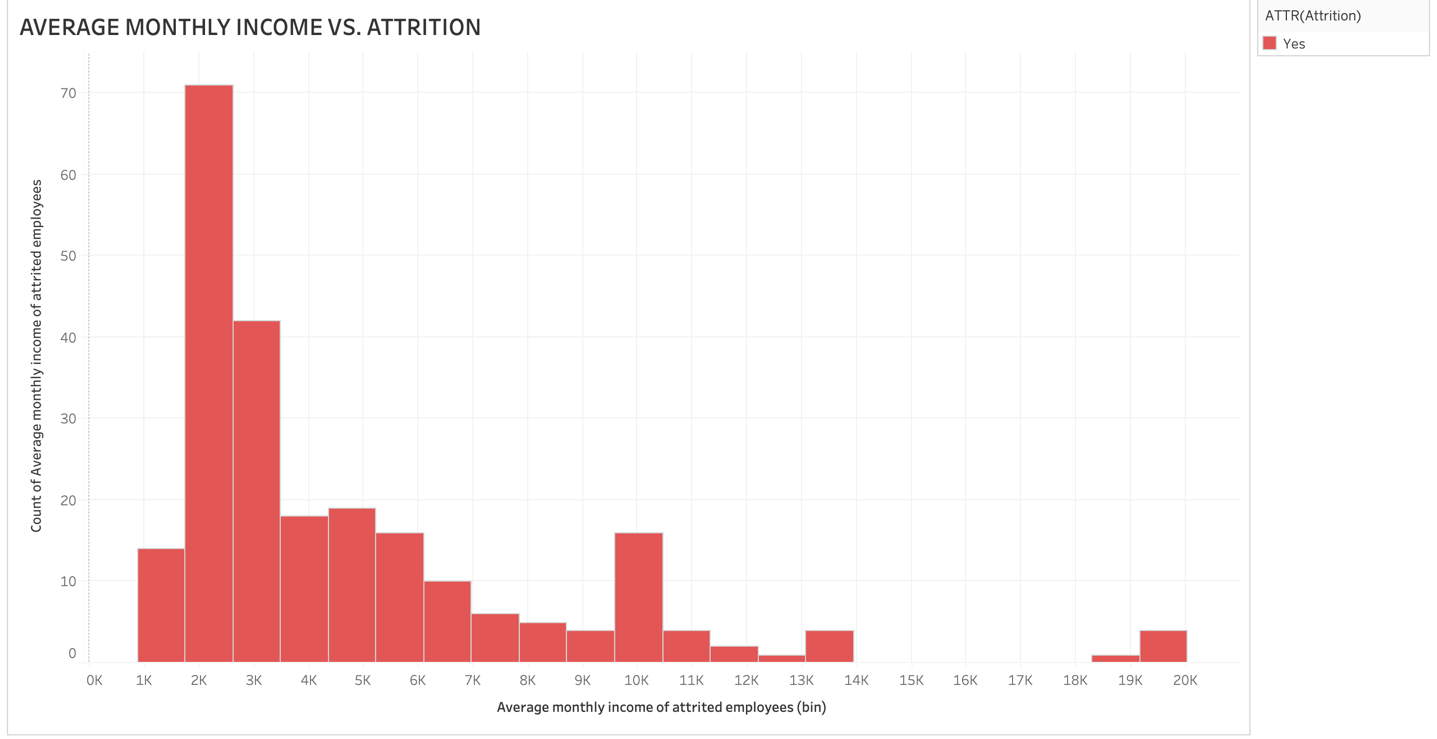
**AVERAGE TENURE OF ATTRITED EMPLOYEES**

****

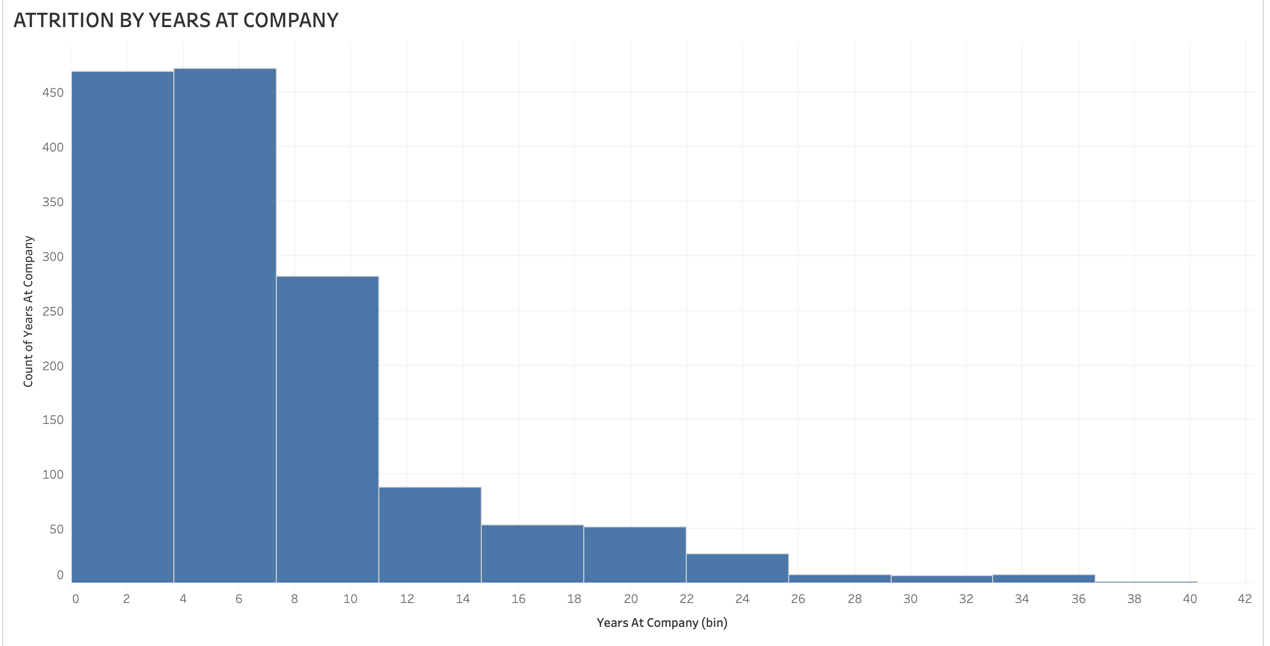
**ATTRITION BY AGE DISTRIBUTION**

****

**ATTRITION BY JOB ROLE**

****

**AVERAGE MONTHLY INCOME VS. ATTRITION**

****

**ATTRITION BYY YEARS AT COMPANY**

**DASHBOARD VISUALIZATION**

****

****

**DASHBOARD PRESENTATION**

**COMPARING EXPECTED VS. ACTUAL TRENDS**

**● OVERALL ATTRITION RATE**

• **Expected:** Many organizations experience a 10-15% annual turnover rate,

although this can vary widely by industry.

• **Actual (16.12%)**: Slightly above the typical range, suggesting the company

might have **marginally higher turnover** than average.

• **Interpretation**: The overall attrition rate is **close to, but somewhat higher**

**than,** common benchmarks -an area for possible improvement.

● **DEPARTMENT-WISE TURNOVER**

• **Expected:**

• **Sales** roles often have **higher attrition** due to target pressure and high

burnout.

• **R&D** roles typically show **lower attrition** because they require specialized

skills and often involve longer-term projects.

•  **Actual:**

• **Sales** indeed has the **highest attrition (20.63%)**, aligning with the

expectation that sales roles are more stressful.

• **Research & Development** has the **lowest attrition (13.84%)**, also matching

typical industry patterns.

• **Interpretation:** Department-wise results are **in line** with general HR

expectations- sales is the biggest flight risk, while R&D is more stable.

● **OVERTIME & WORK-LIFE BALANCE**

• **Expected:** Employees working excessively overtime or with **poor work life**

**balance** is more prone to **burnout** and thus more likely to leave.

• **Actual:**

• Employees who work **overtime** show a **30.53% attrition rate** vs **10.44%** for

non- overtime workers.

• **Poor work- life balance** (rating 1/4) has a **31.25%** attrition rate.

• **Interpretation:** The **high correlation between overtime and attrition**

confirms **industry research** that burnout is a major driver of turnover. This

**matches** the expected trend.

● **SALARY & JOB SATISFACTION**

• **Expected:** Lower pay and poor job satisfaction are **commonly associated**

with higher turnover.

• **Actual:**

• Employees who left earned about **$4,787/month** on the average, where as

those who stayed earned **$6,833/month.**

• **Low job satisfaction** groups has higher attrition rates.

• **Interpretation: Actual data** strongly supports the **expected.**

• **Relationship: Competitive compensation and positive work**

**environment** reduce attrition.

● **AGE & TENURE**

• **Expected**

• **Younger employees** (entry-level) are often more likely to job-hop for better

opportunities.

• **Longer-tenured employees** typically show high loyalty and lower attrition.

• **Actual**

• Age group **18-25** has the **highest attrition (34.78%).**

• **YearsAtCompany** correlates **negatively** with attrition, indicating **longer**

**tenure ——> lower turnover.**

• **Interpretation:** The **data aligns** with common patterns - **early - career**

**employees** live frequently, while **long-tenured employees** tend to stay.

**OVERALL CONCLUSION**

• The **slightly higher overall turnover** than the typical 10-15% range may

prompt **further investigation** and **enhanced retention strategies** (e.g.,

reducing overtime, improving compensation packages, and boosting job

satisfaction).

• By contrasting **expected vs. actual** findings, the organization can **validate**

**known risk areas** (e.g. sales, overtime) and **pinpoint where important** can

align with best practices.

**KEY OBSERVATIONS AND ANOMALIES**

**KEY OBSERVATIONS**

**● Overall Attrition Rate (16.12%)**

• Slightly higher than the typical 10-15% range often cited in industry

benchmarks.

• Suggests a **marginally higher turnover** than average.

● **High Attrition in sales (20.63%)**

• **Sales** department has the **highest attrition rate,** which aligns with **industry**

**norms** for high-stress, target-driven roles.

• **R&D** shows the **lowest attrition (13.84%),** consistent with specialized,

project-based positions.

● **Overtime correlates with burnout**

• **30.53%** attrition among employees who work **overtime** vs **10.44%** for

employees who do not.

• Reinforce the **strong link** between **excessive workload** and turnover.

● **Work-life balance influences turnover**

• **Poor work life balance** (rating= 1/4) has a **31.25%** attrition rate much higher

than those with better ratings.

• Emphasizes the **need for policies** that reduce burnout (e.g., Flexible hours,

remote work).

● **Salary & job satisfaction**

• Employees who left had a **lower average monthly salary** ($4,787) than

those who stayed ($6,833).

• **Low job satisfaction** correlates with higher turnover, suggesting

**compensation and engagement** are key retention levers.

● **Younger employees more likely to leave**

• Age group 18-25 experiences the **highest attrition (34.78%)**.

• Likely due to **early career mobility** and **search for growth opportunities**.

**ANOMALIES & ATYPICAL FINDINGS**

**● Departmental patterns largely match expectations**

• While **sales** are typically expected to have high attrition **Human Resources**

also shows a relatively high rate **(19.05%).**

• This may indicate **Internal process challenge** or **resources constraints**

within HR.

● **Relatively High overtime Attrition**

• A **3x difference** between overtime vs. non-overtime attrition is **larger** than

many industry benchmarks (often 1.5x - 2x).

• Suggests a **particular severe workload** or **lack of compensatory benefits**.

● **No Missing Values** in the Dataset

• It’s somewhat **unusual** to have a complete dataset with **zero** missing values.

• This might be a result of **strong day governance** or **automated HR**

**systems** but warrants a **double-check** for data entry errors.

● **Longer Tenure Employees still leaving**

• While attrition **does decrease** with more years at the company, a **notable**

**minority** of long-tenured employees still left.

• Could indicate **plateauing career paths** or **lack of advancement** for

experienced staffs.

**IMPLICATIONS OF THE OBSERVATIONS & ANOMALIES**

**● High Overtime Attrition:** The **severe gap** calls for **immediate review** of

workload distribution and overtime policies.

● **HR Department Turnover:** Potential **internal morale issues** or **resource**

**shortages** within the HR team.

● **Young Employee Flight:** Emphasizes **career development** and

**mentorship programs** to retain entry - level talent.

● **No Missing Data:** While convenient, it’s wise to **confirm** data completeness

and **accuracy**.

**BUSINESS CONTEXT OF THE FINDINGS**

**● SLIGHTLY HIGHER THAN AVERAGE ATTRITION (16.12%)**

**● Context:** A typical annual turnover rate for many industries hovers around

**10-15%**

**● Business Implication:**

• The company attrition rate is at the **higher end** of the typical range,

suggesting **increased recruitment costs, knowledge loss** and **potential**

**disruptions** in team dynamics.

• Retaining key talent becomes **more urgent** to maintain continuity and reduce

hiring expenditures.

**● HIGH TURNOVER IN SALES (20.63%)**

**● Context:** Sales is often a **revenue - generating** department with high-

performance pressure.

**● Business Implication:**

• Losing **sales staff** can directly affect **revenue** and **client relationships.**

• Repeated turnover in sales can **hamper pipeline management** and **client**

**acquisition.**

• Implementing **targeted retention strategies** (e.g., better incentives, realistic

targets, supportive management) can **stabilize revenue.**

**● OVERTIME AND BURNOUT ATTRITION (30.53%)**

**● context:** Employees working **overtime** show nearly **triple** the attrition rate

compared to those who do not.

● **Business Implication:**

• **Productivity** may suffer if **overworked employees** become **disengaged** or

**burned out.**

• **Health care cost** could rise due to stress-related issues.

• Companies risk a **toxic work culture** that can **damage employer branding**

and **deter top talent** from joining.

● **IMPACT OF WORK-LIFE BALANCE**

**● Context:** Poor work life balance (rating = 1/4) leads to a **31.25%** attrition rate.

● **Business Implication:**

• Excessive turnover due to **work-life imbalance** can **drive away high**

**performers** seeking healthier environments.

• Impacts **morale** and **team cohesion** if employees see colleagues frequently

leaving due to burnout.

• **Flexible work policies** and **wellness programs** can reduce turnover and

**improve employee reputation**.

● **SALARY & JOB SATISFACTION**

**● Context:** Employees who left earned on average **$4,787/month** vs.

**$6,833/month** for those who stayed, indicating **compensation** is a key factor.

● **Business Implication:**

• **Underpaid employees** often seek higher paying opportunities elsewhere,

raising **recruitment cost** for replacements.

• Low job satisfaction also correlates with poor **productivity** and **team e**

**engagement**.

• **Competitive compensation** and **recognition programs** can **boost**

**retention** and **performance.**

**● YOUNGER EMPLOYEES LEAVING MORE OFTEN (34.78% for 18-25)**

**● Context:** Early career- professionals are more prone to **job hopping** for

better pay or growth opportunities.

● **Business Implication:**

• High turnover among younger staff leads to **repetitive onboarding** and

**training expenses**.

• Loss of **future leadership pipeline** if the company fails to retain and develop

young talent.

• **Mentorship programs** and **clear career paths** can **reduce early attrition**

and foster loyalty.

**OVERALL BUSINESS IMPACT & RECOMMENDATIONS**

**● FINANCIAL COSTS:**

• **Recruitment, onboarding and training** expenses increase with higher

turnover.

• Frequent departures disrupt **project timelines** and **client’s relationships**

**(**especially in sales).

● **STRATEGIC CONSIDERATIONS:**

• **Retaining key talent** in revenue critical roles (sales) is vital for **meeting**

**growth targets**.

• **Improving work life balance** can **elevate the employer brand** and **reduce**

**burnout** related losses.

● **OPERATIONAL EFFICIENCY:**

• **Burn-out and overtime** can reduce **long- term productivity**.

• **Career development programs** for younger employees can **build internal**

**leadership** and **reduce churn**.

● **CULTURE & MORALE:**

• High attrition can create a **negative feedback loop** of **low morale** and

**further resignations**.

• Addressing **salary disparities** and **job satisfaction** fosters a **positive**

**culture** that **retains talent**.

**POTENTIAL CAUSES OF OBSERVED ATTRITION**

**PATTERNS**

**● HIGH TURNOVER IN SALES (20.63%)**

**● High-pressure targets:** Sales roles often come with **aggressive quotas** and

performance metrics, leading to **stress** and **burnout**.

● **Commission- based pay fluctuations:** if compensation heavily relies on

**commission**, inconsistent earnings can push employees to seek more stable

opportunities.

● **Competitive job market:** Skilled sales people are **in high demand**, making

it easier for them to **switch companies** for better pay or perks.

● **SIGNIFICANT OVERTIME & BURNOUT (30.53% ATTRITION)**

**● Resource Constraints:** If department are **understaffed**, existing employees

may shoulder heavier workloads, driving **excessive overtime**.

● **Poor Work-Life Balance Policies:** A lack of **flexible schedules, remote**

**work options,** or **time-off policies** can exacerbate burnout.

● **Organizational Culture:** A culture that **rewards ‘always-on’** behavior can

normalize overtime, leading to **chronic fatigue** and higher attrition.

● **YOUNGER EMPLOYEES LEAVING MORE FREQUENTLY (34.78% for Age**

**18-25)**

**● Career Growth & Exploration**: Early-career professionals often **job-hop** to

**explore different roles** or **find better growth paths**.

● **Lower Initial Salaries:** Entry-level employees might feel **underpaid**,

especially if pay raises or promotions are slow.

●**Lack of Engagement or Mentorship:** Younger employees may leave if they

**don’t see a clear career trajectory**, training opportunities or **strong**

**mentorship programs**.

● **SALARY DISCREPANCIES (LEAVING EMPLOYEES EARN $4,787 VS**

**$6,833 FOR STAYING)**

**● Below-Market Pay:** Employees discover **competitive offers** elsewhere,

leading to **voluntary resignation**.

● **Inequitable Pay Structures:** If compensation is not regularly reviewed or

adjusted for **market changes,** pay can **lag** and cause frustration.

● **Performance- Based vs. Fixed Pay:** High performers might feel

**undervalued** if **pay structures** don’t reward them proportionally.

● **LOWER JOB SATISFACTION CORRELATING WITH ATTRITION**

**● Limited Recognition & feedback:** Employees may feel **undervalued** if

managers provide **infrequent feedback** or **recognition.**

**● Monotonous Roles:** Lack of **variety, challenging tasks,** or **growth**

**opportunities** can reduce job satisfaction.

● **Workplace Culture & Management Style:** Unsupportive or **micromanaging**

leadership often leads to **dissatisfaction** and resignations.

● **HR DEPARTMENT TURNOVER (19.05%)**

**● Under- Resourced HR Team:** HR might be handling **excessive**

**responsibilities** (recruiting, payroll, compliance), leading to **stress** and

burnout.

● **High Expectations, Low Support:** HR is often expected to **solve**

**organizational issues** without adequate **executive support**.

● **Conflict of Interest:** HR professionals can feel **pressured** if they must

**enforce policies** they personally disagree with, causing **moral or job**

**dissatisfaction.**

**OVERALL IMPLICATIONS**

These potential causes highlight **operational gaps** and **cultural issues** that

contribute to turnover. By **addressing workload imbalances, improving**

**compensation practices,** and **fostering a supportive environment,**

organizations can **alleviate many of the factors driving attrition.**

**DATA - DRIVEN RECOMMENDATIONS TO REDUCE**

**ATTRITION**

**● ADDRESS OVERTIME & WORKLOAD IMBALANCES**

**● Finding:** Employees with overtime have a **30.53% attrition rate,** nearly 3x

higher than non-overtime workers.

● **Recommendation:**

• **Optimizing Staffing Levels:** Ensure sufficient headcount to distribute

workloads more evenly.

• **Implement Clear Overtime Policies:** Limit weekly overtime hours and offer

**compensatory time off.**

• **Introduce Flexible Work Options:** Remote work or staggered schedules

can **improve work-life balance.**

**● Data -Driven Rationale:** Reducing overtime should **lower burnout** and

**improve retention**, especially among teams with heavy workloads.

● **IMPROVE COMPENSATION COMPETITIVENESS**

**● Finding:** Employees who leave earn an average of **$4,787** vs. **$6,833** for

those who stay.

● **Recommendation:**

• **Conduct Salary Benchmarking:** Regularly compare salaries with **industry**

**standards** to stay competitive.

• **Performance-Based Raises:** Reward **high performers** more quickly to be

**prevent dissatisfaction**.

• **Transparent Pay Structures:** Provide **clear communication** on how pay is

determined and updated.

● **Data-Driven Rationale: Higher salaries** and **performance based**

**incentives** can **reduce flight risk,** especially in **Sales** and **technical roles**.

● **ENHANCE WORK-LIFE BALANCE INITIATIVES**

**● Finding:** Poor work-life balance (rating = 1/4) leads to a **31.25%** attrition rate.

● **Recommendation:**

• **Flexible Schedules:** Allow **flex-time** or **compressed workweeks** to

accommodate personal needs.

• **Promote Well-Being Programs:** Offer stress management workshops,

mental health support, or wellness stipends.

• **Encourage Time-Off:** Actively support vacation usage to **reduce burnout**.

● **Data-Driven Rationale:** Improved work-life balance policies **directly**

**address** the **high attrition** seen in employees with poor ratings.

● **TARGETED RETENTION IN SALES DEPARTMENT**

**● Finding: Sales** has the **highest attrition (20.63%),** impacting revenue

generation.

● **Recommendation:**

• **Revise Sales Target & Quotas:** Ensure targets are **achievable** to reduce

pressure.

• **Sales-Specific Incentives:** Offer **bonus structures** tied to **team**

**collaboration** as well as individual performance.

• **Ongoing Training & Support:** Provide **sales enablement tools**, coaching,

and **clear career paths.**

**● Data-Driven Rationale:** Focusing on **Sales** retention can **stabilize revenue**

and reduce **costly turnover** in key client-facing roles.

●**ENGAGE YOUNGER EMPLOYEES (18-25)**

**● Finding:** Attrition is **34.78%** in the **18-25** age group, the highest among all

age brackets.

● **Recommendation:**

• **Mentorship & Career Development:** Pair younger employees with

**experienced mentors** and define **growth pathways**.

• **Frequent Feedback & Recognition:** Conduct **regular check-ins** to ensure

they feel **valued and supported**.

• **Structured Onboarding Programs:** Strengthen **early engagement** to

reduce quick quits.

● **Data-Driven Rationale:** Investing in **career growth** and **support** for younger

employees can **significantly lower** early-career turnover.

● **FOSTER HIGHER JOB SATISFACTION**

**● Finding:** Lower **job satisfaction** correlates with **increased attrition**.

● **Recommendation:**

• **Employee Recognition Programs:** Celebrate milestones, achievements,

and **peer-to-peer recognition**.

• **Skill Development & Training:** Offer **continuous learning** and internal

mobility options to keep roles **challenging**.

• **Open Communication Channels:** Encourage **feedback loops** (e.g., town

halls, pulse surveys) to **quickly address issues**.

● **Data-Drive Rationale:** Higher engagement and satisfaction **directly reduce**

the likelihood of employees seeking other opportunities.

**OVERALL IMPACT**

Implementing these data-driven recommendations can:

• **Reduce turnover costs** (recruitment, onboarding, lost productivity).

• **Increase employee morale** and **long-term loyalty**.

• **Strengthen the employer brand**, attracting top talent who value balanced

workloads and competitive pay.

By **prioritizing the above interventions**, the organization can **address the root causes** of attrition uncovered in the analysis, ultimately **improving performance** and **retention** across all departments.

**ACTIONABLE STEPS FOR IMPROVEMENT**

● **Reduce Overtime & Improve Work-Life Balance**

**1. Set Overtime Limits:**

• Implement a **weekly cap** on overtime hours (e.g., max 5 hours/week).

• Require **manager approval** for overtime above the limit.

**2. Flexible Scheduling Options:**

• Introduce **remote work** or **flexible start/end times**.

• Pilot a **compressed workweek** (e.g., four 10-hour days).

**3. Workload Assessment:**

• Conduct quarterly **resource audits** to ensure departments are

adequately staffed.

• Automate or **streamline repetitive tasks** to reduce manual workload.

**● Review & Update Compensation Packages**

**1. Benchmark Salaries:**

• Compare current pay scales with **industry averages** to identify gaps.

• Adjust salaries **annually** or **biannually** to remain competitive.

**2. Performance-Based Raises:**

• Offer **merit increases** tied to clear performance metrics.

• Provide **spot bonuses** for outstanding achievements.

**3. Transparent Pay Policy:**

• Publish **salary bands** or ranges for each role.

• Communicate how promotions and **pay adjustments** are decided.

● **Boost Job Satisfaction & Engagement**

**1. Recognition Programs:**

• introduce **Employee of the Month awards** or **peer-to-peer**

**recognition platforms.**

• Publicly celebrate **team achievements** (e.g., hitting project

milestones).

**2. Professional Growth Opportunities:**

• Sponsor **training workshops**, online courses, or certifications.

• Encourage **internal promotions** to show clear career paths.

**3. Regular Feedback & Check-Ins:**

• Managers should hold **1:1 meetings** at least monthly.

• Use **pulse surveys** to gauge morale and address concerns quickly.

● **Focus on High-Risk Groups (Sales & Younger Employees)**

**1. Sales-Specific Retention Measures:**

• Refine **commission structures** to balance individual and team

incentives.

• Provide **sales coaching** and **pipeline management tools**.

**2. Early-Career Development:**

• Launch a **mentorship program** pairing entry-level employees with

experienced staff.

• Offer **clear career tracks** so younger employees see potential growth

within the company.

**3. Onboarding Enhancements:**

• Extend **onboarding** to 90 days with **scheduled check-ins** to ensure

new hires feel supported.

• Assign a **buddy or mentor** from Day 1 for quick knowledge transfer.

● **Implement & Monitor KPIs**

**1. Track Overtime Hours & Burnout Indicators:**

• Use **HR dashboards** to monitor employees exceeding weekly

overtime caps.

• Conduct **exit interviews** to see if overtime was a factor in

resignations.

**2. Attrition & Retention Metrics:**

• Measure **departmental attrition** monthly/quarterly.

• Set **retention targets** (e.g., reduce Sales turnover by 5% in 12

months).

**3. Employee Satisfaction Surveys:**

• Collect **Net Promoter Scores (NPS)** or **engagement scores**

quarterly.

• Compare **survey results** over time to gauge improvement.

**EXPECTED OUTCOMES**

**• Lower Attrition Rate:** By addressing **workload, compensation, and job**

**satisfaction** issues, expect a **5–10% reduction** in turnover within a year.

**• Higher Employee Engagement:** Clear career paths and recognition

programs lead to **improved morale** and **productivity**.

**• Stronger Employer Brand:** Positive work-life balance policies and fair pay

**attract top talent**, reducing hiring costs.

**POTENTIAL LIMITATIONS AMD FUTURE SCOPE**

● **Data Quality & Completeness**

• **Limitation:** While this dataset has no missing values, there is always a

possibility of **data entry errors, inaccurate records,** or **outdated**

**employee information**.

• **Future Scope:**

• **Regular Data Audits:** Implement systematic checks and validation

rules to ensure accuracy.

• **Integration of Additional Data Sources:** Incorporate performance

reviews, engagement surveys, or exit interviews to get a fuller picture

of attrition drivers.

● **Voluntary vs. Involuntary Turnover**

**• Limitation:** The analysis treats all attrition as though employees chose

to leave voluntarily. In reality, some departures may be **involuntary**

(e.g., layoffs, terminations).

• **Future Scope:**

**• Separate Turnover Categories:** Classify and analyze **voluntary vs.**

**involuntary** exits for clearer insights.

• **Tailored Interventions:** Address the **root causes of each type of**

**turnover differently.**

● **Generalizability to Other Departments or Industries**

**• Limitation:** The findings and recommendations are specific to the

**departments and roles** in the dataset, which may not represent all

industries or job functions.

• **Future Scope:**

• **Benchmarking Against Industry Peers:** Compare attrition rates and

practices with similar organizations to gauge competitiveness.

• **Scalable Framework:** Adapt the analysis approach to **other**

**departments**, **subsidiaries**, **or global branches** to see if trends hold.

● **Model Assumptions & Interpretations**

**• Limitation:** Predictive models like **Logistic Regression** assume

certain **linear relationships** and may not capture all complexities.

Additionally, **causation** cannot be conclusively established from

correlational data.

• **Future Scope:**

• **Use Advanced Models:** Experiment with **Random Forest, Gradient**

**Boosting**, or **Neural Networks** for potentially higher accuracy.

• **Causal Inference Techniques:** Employ methods like **propensity**

**score matching** or **instrumental variables** to get closer to causal

insights.

● **External Economic & Market Factors**

**• Limitation:** The analysis does not account for **macro-economic**

**conditions**, **job market trends**, or **competitor actions** that can

influence attrition rates.

• **Future Scope:**

• **External Data Integration:** Merge labor market indices,

unemployment rates, or competitor salary data to contextualize

turnover.

• **Scenario Planning:** Model how economic upturns/downturns might

affect attrition to prepare strategic responses.

● **Cultural and Managerial Differences**

**• Limitation:** The data does not capture **managerial styles**, **team**

**culture**, or **leadership practices**, which can vary significantly within a

company.

• **Future Scope:**

• **Leadership Effectiveness Metrics:** Include **manager ratings** or

**leadership assessment** results to see if certain managerial styles

correlate with lower attrition.

• **Qualitative Feedback:** Conduct focus groups or interviews to

understand **cultural nuances** behind turnover decisions.

**KEY TAKEAWAYS FROM EMPLOYEE ATTRITION**

**ANALYSIS**

**● OVERALL ATTRITION RATE:**

• The company’s attrition stands at **16.12%**, slightly above common

industry benchmarks (10–15%).

● **DEPARTMENT-WISE TRENDS:**

• **Sales** experiences the **highest turnover (20.63%)**, impacting

revenue-generating roles.

• **Research & Development** sees the **lowest turnover (13.84%)**,

reflecting the stability of specialized roles.

● **WORKLOAD & OVERTIME:**

• **30.53%** attrition among employees working overtime (vs. **10.44%** for

non-overtime), indicating **burnout** as a critical factor.

● **COMPENSATION & JOB SATISFACTION:**

• Employees who leave earn an average of **$4,787/month**, significantly

less than **$6,833/month** for those who stay.

• **Low job satisfaction** also correlates with higher turnover, highlighting

**pay and engagement** issues.

● **YOUNGER EMPLOYEES (18-25) AT HIGHER RISK:**

• Attrition for this age group reaches **34.78%**, suggesting **early-career**

**employees** often leave for better opportunities.

● **KEY PREDICTORS OF ATTRITION:**

• **Overtime** is the most significant positive predictor, while **salary**, **job**

**satisfaction, and work-life balance** are major factors reducing

turnover when improved.

● **BUSINESS IMPLICATIONS:**

• High turnover leads to **increased recruitment/training costs,**

**productivity loss**, and **potential revenue impact** (especially in

Sales).

• **Improving work-life balance**, **compensation structures, and**

**career development** can mitigate these issues.

● **RECOMMENDATIONS:**

• **Limit overtime** through clear policies and resource planning.

• **Enhance pay competitiveness** and **transparent compensation**.

• **Develop mentorship programs** for younger staff.

• **Offer growth paths** and **recognition initiatives** to boost job

satisfaction.

**REINFORCING THE SIGNIFICANCE OF EMPLOYEE**

**ATTRITION ANALYSIS**

Employee attrition directly impacts a company’s **financial performance,**

**operational continuity**, and **organizational culture**. By **quantifying**

**turnover rates** and **uncovering the root causes** (e.g., overtime,

compensation, work-life balance), this analysis equips decision-makers with

**actionable insights** to:

**1. Optimize Resource Allocation:** Lowering attrition means **fewer**

**recruitment costs, reduced training overhead,** and **higher productivity**

from a stable workforce.

**2. Retain High-Value Talent:** Targeted interventions (e.g., **improved pay**,

**flexible schedules**) help **preserve institutional knowledge** and

**maintain customer relationships**, especially in high-impact departments

like **Sales**.

**3. Enhance Employee Engagement & Satisfaction:** Addressing **burnout**

and **career development** needs fosters a **positive work environment**,

reducing the risk of losing top performers.

**4. Strengthen Employer Branding:** A company known for **work-life**

**balance** and **fair compensation** is more attractive to **top-tier candidates**,

ensuring a **robust talent pipeline**.

**AREAS FOR FURTHER RESEARCH**

● **VOLUNTARY VS. INVOLUNTARY TURNOVER**

• **Rationale:** Not all attrition stems from employees voluntarily leaving.

Some cases involve **terminations** or **layoffs**.

• **Next Step:** Segment the data into **voluntary** and **involuntary** exits to

**pinpoint distinct causes** and refine retention strategies.

● **MANAGERIAL INFLUENCE & LEADERSHIP STYLES**

• **Rationale:** The **quality of management** often impacts job satisfaction

and employee loyalty.

• **Next Step:** Collect **manager ratings**, 360-degree feedback, or

**leadership style assessments** to see if certain managers or

leadership approaches correlate with higher or lower attrition.

● **TEAM CULTURE & DYNAMICS**

• **Rationale:** Employee turnover can vary within departments based on

t**eam cohesion**, **communication**, and **collaboration**.

• **Next Step:** Survey employees on **team culture** and use **network**

**analysis** (e.g., social network mapping) to identify **pockets of high**

**turnover** within the same department.

● **LONGITUDINAL STUDY OF TENURE & CAREER PATHS**

• **Rationale:** Understanding **how tenure influences attrition** over time

can reveal **critical windows** where employees are most likely to

leave.

• **Next Step:** Track employees across **multiple years**, analyzing

**promotion history**, **role changes**, and **skill developmen**t to

determine the optimal **timing** for interventions (e.g., promotions,

raises).

● **EXTERNAL ECONOMIC FACTORS**

• **Rationale:** Economic conditions (e.g., **job market competitiveness,**

**regional unemployment rates**) can heavily influence turnover.

• **Next Step:** Combine internal HR data with **macro-economic**

**indicators** (e.g., local unemployment rate, competitor hiring trends) to

see if **attrition spikes** coincide with **economic upswings** or

**downturns**.

● **PSYCHOMETRIC & BEHAVIORAL INDICATORS**

• **Rationale:** Employee **personality traits**, **engagement scores**, and

**behavioral patterns** (e.g., absenteeism, lateness) can foreshadow

attrition.

• **Next Step:** Integrate **psychometric assessments** or **digital**

**footprints** (e.g., productivity logs) to create **early warning systems**

for potential flight risks.

● **DIVERSITY & INCLUSION METRICS**

• **Rationale:** Attrition rates can differ by **gender**, **ethnicity**, **or other**

**demographic factors**, indicating potential **equity** or **inclusion** issues.

• **Next Step:** Expand the dataset to include **diversity metrics**, then

analyze whether certain groups experience **disproportionate**

**turnover** and why.

● **ADVANCED PREDICTIVE MODELING & CASUAL ANALYSIS**

• **Rationale:** While **logistic regression** provides correlation-based

insights, more advanced methods might capture **non-linear**

**relationships** or **causal effects**.

• **Next Step:**

• **Machine Learning:** Try **Random Forest**, **Gradient Boosting**, or

**Neural Networks** for improved accuracy.

• **Causal Inference:** Use techniques like **propensity score matching**

to better isolate **cause-and-effect** relationships.

**DATA DICTIONARY**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Data Type** | **Possible Values/Formats** | **Description** |
| **Age** | Integer | 18 to 65+ (typical  range) | Employee’s age  (in years). |
| **Attrition** | String | “Yes”, “No” | Indicates whether  the employee left  the company  (Yes) or stayed  (No). |
| **Business Travel** | String | “Travel\_Rarely”,  “Travel\_Frequently”, “Non-Travel” | Frequency of  employee’s  business travel. |
| **Daily Rate** | Integer | 1 to 1500  (approx.) | Random daily  salary rate for  employees (for  demonstration  purposes). |
| **Department** | String | “Sales”, “Human  Resources”,  “Research &  Development” | The department  in which the  employee works. |
| **Distance From**  **Home** | Integer | 1 to 29 | Distance in miles  from the  employee’s home  to the workplace. |
| **Education** | Integer | 1 to 5 | Education level  (1=Below  College,  2=College,  3=Bachelor,  4=Master,  5=Doctor). |
| **Education Field** | String | “Life Sciences”,  “Medical”,  “Marketing”, etc. | Field of study for  the highest  education level  attained. |
| **Employee Count** | Integer | Typically, 1  (Constant in  sample data) | Used for  demonstration;  often all rows  have value 1 in  the sample  dataset. |
| **Employee Number** | Integer | Unique integer ID | A unique identifier  for each  employee. |
| **Environment Satisfaction** | Integer | 1 to 4 | Employee’s  satisfaction with  the work  environment  (1=Low, 4=Very  High). |
| **Gender** | String | “Male”, “Female” | Employee’s gender. |
| **Hourly Rate** | Integer | 1 to 100 (approx.) | Hourly wage rate  (For  demonstration). |
| **Job Involvement** | Integer | 1 to 4 | Employee’s level  of involvement in  their job (1=Low,  4=Very High). |
| **Job Level** | Integer | 1 to 5 | Job seniority level  (1=Entry-level,  5=Executive). |
| **Job Role** | String | “Sales  Executive”,  “Research  Scientist”, etc. | The specific role  or title of the  employee within  the company. |
| **Job Satisfaction** | Integer | 1 to 4 | Employee’s  satisfaction with  their job (1=Low,  4=Very High). |
| **Marital Status** | String | “Single”,  “Married”, “Divorced” | Employee’s  marital status. |
| **Monthly Income** | Integer | 1000 to 20,000+  (approx.) | Monthly salary in  USD (numeric  range depends  on dataset). |
| **Monthly Rate** | Integer | 1 to 30,000  (approx.) | Random monthly  rate assigned to  employees (for  demonstration). |
| **NumCompanies Worked** | Integer | 0 to 9 | Number of  companies the  employee worked  for prior to joining  the current  organization. |
| **Over 18** | String | “Y” | Indicates if  employee is over  18 (always “Y” in  sample data). |
| **OverTime** | String | “Yes”, “No” | Indicates whether  the employee  regularly worked  overtime. |
| **Percent Salary Hike** | Integer | 11 to 25 (approx.) | Percentage  salary increase  granted during  the last  performance  appraisal. |
| **Performance Rating** | Integer | 1 to 4 (often 3 or 4  in sample data) | Rating of the  employee’s job  performance  (1=Low,  4=Outstanding). |
| **Relationship Satisfaction** | Integer | 1 to 4 | Satisfaction with  co-worker or  supervisor  relationships  (1=Low, 4=Very  High). |
| **Standard Hours** | Integer | Typically, 80  (Constant in  sample data) | Standard hours of work (e.g., 80 in a  2-week period). |
| **StockOption Level** | Integer | 0 to 3 | Stock option  grant level (0=No  options, 3=High-level  options). |
| **TotalWorking Years** | Integer | 0 to 40+ | Total years of  professional experience (including current  and previous  roles). |
| **TrainingTimes Last Year** | Integer | 0 to 6 | Number of  training sessions  attended by the  employee in the  last year. |
| **WorkLife Balance** | Integer | 1 to 4 | Work-life balance  rating (1=Bad,  2=Good,  3=Better, 4=Best). |
| **YearsAt Company** | Integer | 0 to 40+ | Total number of  years the  employee has  been at the  current company. |
| **YearsIn Current Role** | Integer | 0 to 18+ | Number of years  the employee has  been in their  current role. |
| **Years Since Last Promotion** | Integer | 0 to 15+ | Number of years  since the  employee’s last  promotion. |
| **YearsWithCurr Manager** | Integer | 0 to 17+ | Number of years  the employee has  been under their  current manager. |

**NOTES ON DATA DICTIONARY**

• **Data Types:** The sample dataset uses a mix of **integer** columns for

numeric data and **string** columns for categorical data.

• **Value Ranges:** The **approximate ranges** come from the sample

dataset and may vary in real-world HR systems.

• **Column Purpose:** Some columns (e.g., EmployeeCount,

StandardHours) are **constant** or have minimal variance in the sample

data but can be more relevant in real HRIS systems.

• **Categorical Columns:** Many string columns (e.g., Attrition, OverTime,

MaritalStatus) can be **encoded** numerically for analysis (e.g., Label

Encoding).

• **Interdependence:** Certain columns (like Age and TotalWorkingYears)

may correlate, reflecting typical patterns (e.g., older employees often

have more years of experience).

**REFERENCES AND EXTERNAL SOURCES**

**1. Bureau of Labor Statistics (BLS), Job Openings and Labor Turnover**

**Survey (JOLTS).**

• Provides industry-wide turnover and hiring benchmarks.

• https://www.bls.gov/jlt/

2. **Society for Human Resource Management (SHRM).**

• Offers HR best practices, turnover metrics, and retention strategies.

• https://www.shrm.org/

**3. Data sourced from GitHub**

• A commonly used sample dataset for demonstration of employee

attrition analyses.

• https://github.com/pplonski/datasets-for-start/blob/master/

employee\_attrition/HR-Employee-Attrition-All.csv

4. **CIPD (Chartered Institute of Personnel and Development).**

• Publishes annual HR and workforce insights, including turnover rates

and engagement strategies.

• https://www.cipd.co.uk/

5. **Gartner HR Research & Insights.**

• Provides data-driven research on talent management, retention, and

organizational development.

• https://www.gartner.com/en/human-resources

6.**Academic Research on Turnover & Retention**

• Hom, P. W., & Griffeth, R. W. (1995). Employee Turnover. Cincinnati:

South-Western College Publishing.

• Mitchell, T. R., Holtom, B. C., & Lee, T. W. (2001). How to keep your

best employees: Developing an effective retention policy. Academy of

Management Executive, 15(4), 96–108.