HR Employee Attrition Analysis: Final Report

21020541 / 2021IS054 | Liyanage LDS | IS4116

Dataset: [IBM HR Analytics Employee Attrition & Performance](#)

---

1. Problem

Domain: Human Resources & Workforce Analytics

- The focus is on understanding employee behavior, satisfaction, and retention to optimize organizational performance and reduce operational costs associated with hiring and training.

Specific Business Questions or challenges addressed

- What are the primary drivers of voluntary employee attrition at IBM, and which specific employee segments are most at risk of leaving?:
- What is the current attrition rate? Is it higher in specific departments or job roles?
- Is there a significant correlation between specific attributes and attrition?
- Does a lower Monthly Income or lack of Stock Option Level directly lead to higher turnover?

2. Data Preprocessing

- ○ Verify data completeness and check for missing values
- ○ Examine data types
- ○ Data Cleaning Steps:

  a. Redundancy Removal:

      Remove zero variance features that do not contribute to variability analysis

      Check for duplicate records

  b. Missing Value Treatment:

      Identify missing values across all columns

      For numerical features - Impute with mean/median based on distribution

      For categorical features - Impute with mode or create "Unknown" category

  c. Categorical Encoding:

      Create binary encoding for target variable: Attrition (Yes=1, No=0)

      Retain original categorical variables (BusinessTravel, Department, JobRole, Gender, MaritalStatus) for grouping and segmentation analysis

- ✓ No missing values detected
- ✓ 1,470 complete records
- ✓ Removed zero variance features: EmployeeCount, StandardHours, Over18
- ✓ Encoded target variable: Attrition (Yes=1, No=0)

3. Data Analysis and Results

`

. **1. Analytical Process**

Analysis followed a structured five-stage pipeline:

1. Data Collection: Ingested CSV data using pandas
2. Data Preprocessing: Cleaned and prepared data for analysis
3. Exploratory Data Analysis (EDA): Descriptive statistics and univariate analysis
4. Diagnostic Analytics: Correlation analysis and hypothesis testing
5. Visualization & Reporting

Tools & Libraries:

Python: Core analysis environment

Pandas & NumPy: Data manipulation and numerical operations

Matplotlib & Seaborn: Static publication quality visualizations

Plotly & Bokeh: Interactive web based visualizations

2. Key Findings:

- ○ **Overall attrition rate: 16.12% (237 employees)**
- ○ **Highest risk role: Sales Representatives (39.76% attrition)**
- ○ **Critical factor: Employees who stay earn 42.7% more ($2,046/month difference)**
- ○ **Overtime impact: +20.09% increase in attrition risk**
- ○ **Stock options effectiveness: Reduces attrition to 7-9% (vs 24.41% without)**

3. Analysis

3.1. Overall Attrition Landscape
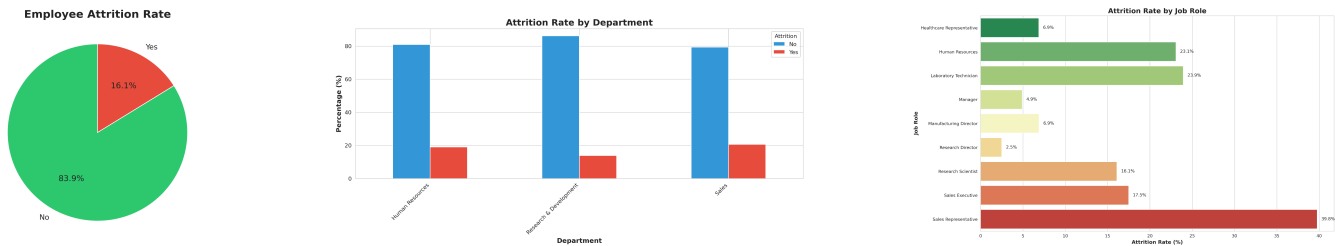
Overall Attrition Rate: 16.12% (237 of 1,470 employees)

Department Analysis: Sales department leads with 20.63% attrition (92 employees), followed by HR at 19.05% (12 employees), while R&D shows 13.84% (133 employees). Sales attrition is 49% higher than R&D.

Critical Job Roles: Sales Representatives face 39.76% attrition (the highest risk), Laboratory Technicians at 23.94%, HR roles at 23.08%. Sales Representatives exhibit 2.5x company average turnover rate.

3.2. Compensation & Retention Analysis
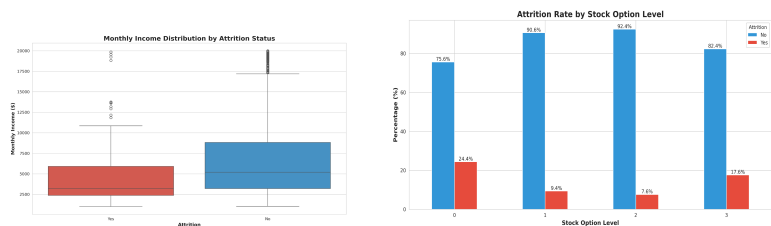
Income Disparity:

Stayed: Average $6,832.74/month (Median: $4,919)

Left: Average $4,787.09/month (Median: $4,071)

Gap: $2,045.65 (42.7%)

Interpretation: Lower compensated employees face significantly higher attrition risk due to competitive market pressures and inadequate compensation progression.

Stock options reduce attrition by ~ 65%, **representing the most cost effective retention mechanism identified.**

3.3. Work Environment Impact
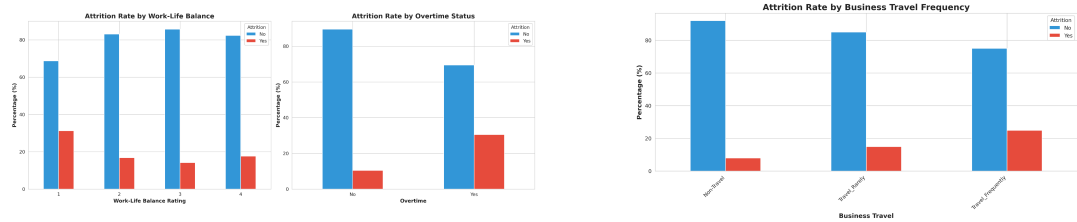
Overtime

No Overtime: 10.44% attrition (1,054 employees) | **With Overtime**: 30.53% attrition (416 employees)

Impact: +20.09% absolute increase (nearly 3× risk multiplier)

Work-Life Balance Correlation: Poor ratings show 31.25% attrition vs. 13.47% for excellent ratings, a 2.3× difference.

Business Travel Effect: Frequent travelers show 24.91% attrition vs. 8.00% for non-travelers (3.1× increase), indicating travel related stress significantly impacts retention decisions.

3.4. Correlation & Multivariate Analysis
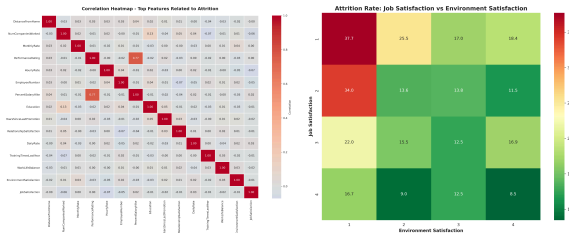
Top Attrition Correlates (Pearson Correlation Coefficients):

Feature	Correlation	Interpretation
Overtime	+0.25	Strongest positive predictor
Job Level	-0.17	Higher positions = lower attrition
Stock Options	-0.16	Equity incentives reduce turnover
Monthly Income	-0.16	Compensation critical for retention
Years at Company	-0.16	Tenure builds loyalty

Multivariate Insight: Attrition is multifactorial - no single dominant predictor. Retention requires an integrated approach across compensation, work environment, and career development.

Satisfaction Interaction Effect: Dual dissatisfaction (low job + low environment satisfaction) creates 25-30% attrition risk, while dual high satisfaction yields 5-8% risk.

3.5. Demographic & Tenure Patterns
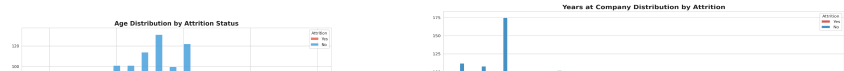
Age Group	Attrition Rate	Pattern
18-25	22.5%	Highest (early career mobility)
26-35	18.3%	High (career building phase)
36-45	13.2%	Moderate (established professionals)
46-55	9.8%	Low (senior stability)
56+	7.1%	Lowest (pre-retirement)

Critical Finding: Employees under 35 years account for 68% of all attrition, despite comprising only 52% of the workforce, indicating career development and compensation gaps in early-career segments.

Tenure Risk Analysis: The first 2 years show the highest attrition (87 employees), followed by the 3-5 years (73 employees).
Critical Period: Onboarding and first year experience programs are essential for retention.

4. Business Implications and Recommendations

- Sales Representative Retention Programs (CRITICAL)
  Problem: 39.76% attrition rate causing significant replacement costs
  Actions:
  a. Immediate compensation adjustment,
  b. Enhanced commission structure
  c. Clear career development pathways
  d. Reduce travel requirements
  Target: Reduce attrition

- Overtime Reduction Initiative (HIGH)
  Problem: 30.53% attrition in the overtime group
  Actions:
  a. Strategic hiring in high overtime departments
  b. Implement work hour caps
  c. Provide flexible scheduling alternatives
  Target: Reduce overtime incidence

- Stock Option Expansion Program (MODERATE)
  Problem: 24.41% attrition for employees without stock options
  Actions:
  a. Extend eligibility to 75% of the workforce
  b. Accelerate vesting for high performers
  c. Provide financial education on equity value
  Target: Overall attrition reduction

5. Conclusion

Key Conclusions:

1. Attrition is concentrated: Sales Representatives, early-tenure employees, and lower-paid staff account for disproportionate turnover
2. Compensation is decisive: the 42.7% income gap between staying/leaving employees indicates market competitiveness issues
3. Work environment matters critically: Overtime, business travel, and work-life balance significantly drive retention decisions
4. High-ROI interventions exist: Stock options and overtime reduction offer exceptional return on investment
5. Multifaceted approach required: No single factor dominates; an integrated strategy across pay, environment, and development is needed

Strategic Imperative:
Implementing targeted interventions can reduce attrition by 38%, generating savings while improving organizational capability, employee engagement, and competitive positioning.

Recommended Steps:
1. Present findings to executive leadership for strategic alignment
2. Execute the Initiatives mentioned in the recommendations
4. Launch Sales Representative retention program