driving strategic decisions with data

WRITTEN BY VICTOR CHANG | MBA 610 – FOUNDATIONS OF ANALYTICS

Employee Retention

**Executive Summary**

Employee retention remains one of the most pressing challenges in today’s competitive labor market. High turnover rates not only increase recruitment and training costs but also disrupt organizational continuity and morale. This paper explores employee retention by examining its scope, causes, and business impact. It uses a relevant dataset to uncover patterns and root causes through descriptive and diagnostic analytics, with recommendations aimed at helping organizations improve employee engagement and reduce turnover.

**Introduction**

**Employee Retention**

Employee retention refers to an organization’s ability to keep its employees over time. It is a critical aspect of workforce management, directly influencing operational efficiency, company culture, and financial performance. Retention issues often arise from dissatisfaction with leadership, limited growth opportunities, inadequate compensation, or poor work-life balance (Society for Human Resource Management [SHRM], 2022).

**Scope of the Problem and Business Impact**

High employee turnover costs businesses significantly. According to Gallup (2023), replacing an employee can cost one-half to two times the employee’s annual salary. For a mid-sized organization, this could mean hundreds of thousands to millions of dollars annually. Furthermore, turnover negatively affects productivity and employee morale. For instance, U.S. companies lose about $1 trillion annually due to voluntary turnover (Gallup, 2023).

**Data Exploration**

**Data Source**

A relevant dataset for analyzing employee retention is the “HR Analytics: Employee Attrition & Performance” dataset, available on Kaggle: https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset. This dataset includes information on employees from a fictional IBM dataset, covering personal demographics, job roles, performance metrics, and attrition status. It is commonly used in HR analytics studies due to its well-structured format and completeness.

**Describe the Data**

The dataset contains 1,470 rows and 35 columns. Each row represents an employee, while each column includes features such as:

* Age, Gender, Education, Job Role
* Monthly Income, YearsAtCompany, Work Life Balance
* Attrition (Yes/No – target variable)

The diversity and richness of this dataset allow for exploring various factors contributing to employee attrition.

**Analytical Techniques**

**Descriptive Analytics**

*Trends and Patterns*

Expected analysis: Use frequency counts, distributions, and cross-tabulations. You’ll likely observe that:

* Younger employees and those with fewer years at the company have higher attrition.
* Roles with high travel requirements or lower job satisfaction correlate with higher turnover.
* Departments like Sales or Human Resources might show higher attrition rates.

**Diagnostic Analytics**

*Causes of Issues*

Expected analysis: Multivariate analysis and correlation. Common findings include:

* Low job satisfaction and poor work-life balance are strong predictors of attrition.
* Employees with lower income or fewer opportunities for promotion are more likely to leave.

What to look for in Python:

* Heatmaps of correlations
* Logistic regression to quantify feature importance
* Decision trees or SHAP values for explainability

**Exploratory Visualization**

Suggestions:

* Bar chart of attrition rate by job role
* Box plots comparing monthly income between those who stayed and those who left
* Heatmap showing correlation of variables with attrition

Use seaborn and matplotlib libraries to create these visuals.

**Insights and Recommendations**

**Actionable Insights**

Based on expected trends:

* Employees with less tenure, lower salaries, and poor job satisfaction are most likely to leave.
* Work-life balance and career development opportunities are consistent drivers of retention.

**Solutions and Strategies**

Recommendations include:

* Implement targeted retention strategies such as mentorship for new employees.
* Increase transparency around career progression and provide continuous learning.
* Use predictive analytics models to identify at-risk employees early and intervene (SHRM, 2022).

**Communication**

**Findings**

Create dashboards or infographics that highlight:

* Top 5 features most associated with attrition
* Summary of employee demographics vs. attrition likelihood
* “What-if” scenarios using model predictions

Ensure visuals are intuitive for a non-technical audience by minimizing jargon and emphasizing takeaways.

**Conclusion**

Employee retention is both a human and business imperative. With the right data and analytics, organizations can understand why employees leave and implement strategies to increase engagement and loyalty. By proactively addressing root causes, companies can create a more stable and productive workforce.

**References**

Gallup. (2023). *The True Cost of Employee Turnover*. https://www.gallup.com/workplace

Society for Human Resource Management (SHRM). (2022). *Managing for Employee Retention*. https://www.shrm.org/resourcesandtools/hr-topics/employee-relations/pages/managing-employee-retention.aspx

Kaggle. (n.d.). *HR Analytics: Employee Attrition & Performance*. https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset