HR Employee Attrition Analysis

Data Analysis Project

February 23, 2025

1 Introduction

Welcome to this Python project, where you will be working with an HR Employee Attrition dataset to analyze factors that contribute to employee turnover. Employee attrition is a critical challenge for organizations, impacting productivity, team morale, and financial stability. Understanding the key drivers behind attrition can help businesses develop effective retention strategies and create a better work environment.

In this project, you will explore a dataset that includes various attributes related to employee satisfaction, performance, demographics, and job characteristics. You will use Python to clean, manipulate, and analyze the data, aiming to uncover patterns and insights that can help HR managers make informed decisions.

Your report should provide a detailed analysis of the data, including visualizations, charts, and key performance indicators (KPIs) that highlight important trends. You will investigate questions such as:

- How does distance from home impact employee attrition?
- What is the relationship between monthly income and job satisfaction?
- Do employees with higher work-life balance ratings tend to stay longer in the company?
- How do different job roles experience attrition rates?
- Show a breakdown of distance from home by job role and attrition.
- Compare average monthly income by education level and attrition.

By the end of this project, you will have gained valuable data analysis skills, enhanced your ability to work with real-world HR data, and developed insights that could be applied to improving employee retention strategies in organizations. This project also provides an opportunity to practice working with Python libraries such as `pandas`, `numpy`, `matplotlib`, and `seaborn`, as well as creating meaningful data-driven reports and visualizations.

Your findings will be presented in a well-structured report and dashboard, making it accessible even to those without a data analysis background.

2 Project Tasks

2.1 Data Retrieval and Merging

- Connect Python to the MySQL database.
- Write and execute SQL queries to extract relevant information.
- Merge multiple tables if necessary to form a structured dataset for analysis.

2.2 Exploratory Data Analysis (EDA)

- Identify missing values and handle them appropriately.
- Perform summary statistics and visualize data distributions.
- Identify potential correlations between features that may contribute to attrition.

2.3 KPI Calculation and Analysis

- Compute key performance indicators such as attrition rate, tenure analysis, and satisfaction index.
- Provide interpretations for each KPI and explain its relevance to HR decision-making.

2.4 Data-Driven Recommendations

- Based on the KPIs, derive actionable recommendations to reduce attrition.
- Justify your recommendations with supporting data.

2.5 Dashboard and Reporting

- Create a dashboard in Python (using Plotly Dash or Streamlit) to visualize key insights.
- Summarize findings in a structured report with clear explanations and visual representations.

3 Feature Explanation

The dataset consists of multiple features that provide insights into employee demographics, job roles, compensation, and satisfaction levels. Below is a structured breakdown of all features and their meanings.

3.1 1. Employee Demographics

- **Age:** The age of the employee in years.
- **Gender:** The gender of the employee (Male/Female).
- **Marital Status:** The employee's marital status (Single/Married/Divorced).
- **Education:** Highest level of education achieved (1 = Below College, 2 = College, 3 = Bachelor, 4 = Master, 5 = Doctor).
- **Education Field:** The primary field of education of the employee (e.g., Life Sciences, Medical, Marketing, Technical Degree).

3.2 2. Job and Employment Details

- **Job Role:** The specific job title of the employee (e.g., Sales Executive, Research Scientist, Manager).
- **Department:** The department where the employee works (e.g., Sales, RD, HR).
- **Job Level:** The level of seniority in the company (1 = Entry Level, 5 = Senior Management).
- **Years at Company:** The total number of years the employee has worked in the company.
- **Years in Current Role:** The number of years the employee has held their current position.
- **Years Since Last Promotion:** The number of years since the employee's last promotion.
- **Years with Current Manager:** The number of years the employee has worked under their current manager.
- **Training Times Last Year:** The number of training programs the employee attended in the last year.
- **Overtime:** Whether the employee works overtime (Yes/No).

3.3 3. Compensation and Benefits

- **Monthly Income:** The monthly salary of the employee.
- **Hourly Rate:** The employee's hourly wage.
- **Daily Rate:** The employee's daily earnings.
- **Monthly Rate:** The employee's total monthly earnings before deductions.

- **Stock Option Level:** The stock option benefits granted to the employee (0 = None, 3 = High).
- **Percent Salary Hike:** The percentage increase in salary after the last performance evaluation.

3.4 4. Work-Life Balance and Satisfaction Metrics

- **Job Satisfaction:** A rating of how satisfied the employee is with their job (1 = Low, 4 = Very High).
- **Environment Satisfaction:** A measure of satisfaction with the work environment (1 = Low, 4 = Very High).
- **Relationship Satisfaction:** A measure of the employee's satisfaction with workplace relationships (1 = Low, 4 = Very High).
- **Work-Life Balance:** A measure of work-life balance satisfaction (1 = Bad, 4 = Best).

3.5 5. Performance and Attrition

- **Performance Rating:** The employee's performance rating in the last review (1 = Low, 4 = Outstanding).
- **Attrition:** Indicates whether the employee left the company (Yes/No).
- **Distance from Home:** The distance between the employee's home and the workplace in kilometers.
- **Num Companies Worked:** The number of companies the employee has previously worked for.

This section provides a comprehensive understanding of each feature in the dataset, helping to guide analysis and ensure meaningful interpretations of the data.

4 Analysis Objectives and KPI Calculations

In this project, you will analyze the HR Employee Attrition dataset to uncover key trends and insights related to employee turnover. The following aspects should be examined during the analysis, supported by key performance indicators (KPIs) to quantify findings.

4.1 1. Attrition Trends and Patterns

- Identify overall attrition rates across the company.
- Examine attrition trends by department, job role, and tenure.
- Determine whether attrition is concentrated in specific groups (e.g., new hires, experienced employees, high performers).

KPI: Attrition Rate

$$\text{Attrition Rate} = \left(\frac{\text{Total Employees Who Left}}{\text{Total Employees}} \right) \times 100 \tag{1}$$

Example Calculation: If 50 employees leave out of 1000 employees:

$$\text{Attrition Rate} = \left(\frac{50}{1000} \right) \times 100 = 5 \tag{2}$$

Insights: A high attrition rate may indicate poor job satisfaction, limited career growth, or external job market opportunities.

4.2 2. Compensation and Benefits Analysis

- Compare average monthly income between employees who left and those who stayed.
- Analyze the relationship between salary levels and attrition rates.
- Assess the impact of stock options and salary hikes on employee retention.

KPI: Average Monthly Income by Attrition Status

$$\text{Average Monthly Income} = \frac{\sum \text{Monthly Income of Employees in a Group}}{\text{Total Employees in That Group}} \tag{3}$$

Example Calculation: If employees who left had total salaries of $100,000 across 10 employees$:
Avg Monthly Income = $\frac{100000}{10}$ =10,000 (4)

Insights: If employees with lower salaries leave more frequently, HR should consider adjusting salary structures to retain talent.

4.3 3. Work-Life Balance and Satisfaction

- Evaluate job satisfaction and work-life balance ratings for employees who left vs. those who stayed.
- Determine if employees with lower work-life balance scores are more likely to leave.

- Assess environmental and relationship satisfaction to understand their impact on retention.

KPI: Attrition Rate by Work-Life Balance

$$\text{Attrition Rate by Work-Life Balance} = \left(\frac{\text{Employees Who Left at a Given Work-Life Balance Score}}{\text{Total Employees at That Score}} \right) \times 100 \tag{5}$$

Insights: If employees with poor work-life balance ratings leave more often, HR should consider flexible working arrangements.

4.4 4. Career Growth and Job Stability

- Analyze the impact of promotions and years since last promotion on attrition.
- Investigate the correlation between job level and employee turnover.
- Assess how tenure and years in current role influence attrition likelihood.

KPI: Average Years Since Last Promotion by Attrition Status

$$\text{Avg Years Since Last Promotion} = \frac{\sum \text{Years Since Last Promotion of Employees in a Group}}{\text{Total Employees in That Group}} \tag{6}$$

Insights: If employees who leave have longer gaps between promotions, HR should review promotion policies.

4.5 5. Performance and Attrition

- Determine whether high performers are more or less likely to leave.
- Analyze performance ratings across different attrition statuses.
- Investigate the relationship between training participation and retention rates.

KPI: Average Performance Rating by Attrition Status

$$\text{Avg Performance Rating} = \frac{\sum \text{Performance Rating of Employees in a Group}}{\text{Total Employees in That Group}} \tag{7}$$

Insights: If high-performing employees are leaving, HR should investigate career advancement opportunities.

4.6 6. Geographic and Commuting Factors

- Assess whether employees with longer commute distances are more likely to leave.
- Examine the impact of different office locations on attrition rates.

KPI: Average Commute Distance by Attrition Status

$$\text{Average Distance from Home} = \frac{\sum \text{Distance from Home of Employees in a Group}}{\text{Total Employees in That Group}} \tag{8}$$

Insights: If long commutes correlate with higher attrition, HR should explore remote work policies.

4.7 7. Data-Driven HR Recommendations

- Based on the insights from the analysis, suggest actionable strategies to reduce attrition.
- Identify potential policy changes that could improve employee retention.
- Provide recommendations on compensation, work-life balance initiatives, and career growth opportunities.

This structured approach will ensure a comprehensive evaluation of the factors influencing employee attrition and enable the development of targeted HR strategies.

5 Final Deliverable: Structured Python Notebook and AI-Powered Dashboard

To complete this project successfully, students must submit a well-structured Jupyter Notebook that includes all aspects of the analysis, along with a Python-based AI-powered dashboard that visualizes key insights.