

Employee Turnover: Exploratory Data Analysis Report

To: HR Department, Portobello Tech

From: ML Development Team

Date: July 8, 2025

Subject: Analysis of Key Factors Driving Employee Attrition

1. Executive Summary

This report details the findings from an Exploratory Data Analysis (EDA) conducted on the company's historical employee data. The analysis reveals that employee turnover is not random but is driven by specific, identifiable factors.

The single most significant predictor of turnover is **employee satisfaction**. Our analysis has identified three distinct high-risk employee profiles:

1. **The Under-utilized:** Employees working on only **2 projects**.
2. **The Over-worked:** Employees working on **6 or 7 projects**, often correlated with logging significantly more hours than average.
3. **The Dissatisfied:** A large and distinct group of employees reporting very low satisfaction levels (below 0.2).

We recommend immediate, targeted interventions focused on workload management and employee engagement to mitigate turnover risk in these specific segments.

2. Key Factor Correlation Analysis

A correlation matrix was generated to understand the linear relationships between numerical features and employee turnover (left).

Key Findings:

- **Satisfaction Level:** This has the strongest relationship with turnover, with a significant negative correlation of **-0.39**. In business terms, this means as employee satisfaction declines, the probability of an employee leaving increases sharply.
- **Other Factors:** `time_spend_company` (0.14), `average_monthly_hours` (0.07), and `number_project` (0.02) have a weak positive correlation with turnover. While they are contributing factors, they are far less impactful than an employee's direct satisfaction.

3. Employee Work-Life Profile Analysis

Distribution plots of key metrics reveal distinct patterns within the workforce.

Key Findings:

- **Employee Satisfaction:** The workforce's satisfaction is highly polarized. There are large groups of employees who are either highly satisfied (0.6-1.0) or highly dissatisfied (0.1-0.4), with a noticeable trough in the middle. This indicates a lack of moderately satisfied employees and points to systemic issues creating a "love it or leave it" environment.
- **Average Monthly Hours:** The bimodal distribution clearly shows two dominant work patterns: a group working standard hours (approx. 140-160 hours/month) and a second, large group working significant overtime (approx. 240-260 hours/month). This overtime group likely contributes to burnout.
- **Employee Evaluation:** The evaluation scores are also bimodal, suggesting employees are generally perceived as either "average performers" (scores around 0.5-0.6) or "top performers" (scores >0.8).

4. Workload and Attrition Analysis

Analysis of the number of projects an employee works on provides one of the clearest indicators of turnover risk.

Key Findings:

- **High-Risk Zones:** The highest rates of turnover occur at the extremes of the project workload:
 - **2 Projects:** A very large number of employees who left were working on only two projects, suggesting potential under-utilization, boredom, or a poor fit for their role.
 - **6 & 7 Projects:** Employees in this bracket have an extremely high turnover rate, strongly indicating burnout.
- **The Retention "Sweet Spot":** Employees working on **3, 4, or 5 projects** are the most likely to stay with the company. This appears to be the optimal workload for engagement and retention.

5. Synthesis and Recommendations

The data tells a clear story: employee turnover at Portobello Tech is primarily a problem of **engagement and workload management**.

- **The Burnout Path:** Employees working long hours on 6-7 projects become dissatisfied and leave.
- **The Disengagement Path:** Employees who are not given enough work (2 projects) become disengaged, dissatisfied, and leave.

Actionable Recommendations:

1. **Conduct Targeted "Stay Interviews":** Immediately engage with employees currently working on **2, 6, or 7 projects**. The goal is to understand their challenges and satisfaction levels before they decide to leave.
2. **Review Workload Distribution:** Managers should review how projects are allocated to ensure employees are not consistently in the high-risk zones. The data strongly supports aiming for a 3-5 project workload per employee.
3. **Investigate the "Under-utilized" Group:** For employees with only 2 projects, determine if this is a temporary situation or if they are in roles that do not fully leverage their skills. This could be an indicator of poor job-fit.
4. **Monitor Satisfaction Proactively:** Given its strong correlation with turnover, regular pulse surveys on employee satisfaction should be considered a primary leading indicator for HR.