Who's Leaving and Why: A Data-Driven Look at Employee Attrition

Part 1 of the HR Analytics Series

Problem

- 16% of employees were leaving, creating costs and operational risk.
- High-risk segments identified: fresh hires, Sales Representatives, and employees working overtime.
- HR lacked clear insights into why employees leave and which interventions work.

Process

Descriptive Analysis (Excel Dashboard)

- Segmented employees by department, role, tenure, gender, income, satisfaction, and overtime.
- Built an interactive dashboard to identify trends and high-risk groups instantly.
- Key metrics highlighted:
    - Fresh Hires (0-1 yr): **34.88% attrition**
    - Sales Representatives: **39.76% attrition**
    - Employees with Overtime: **30.53% attrition**
    - Lower-income employees: **28.61% attrition**

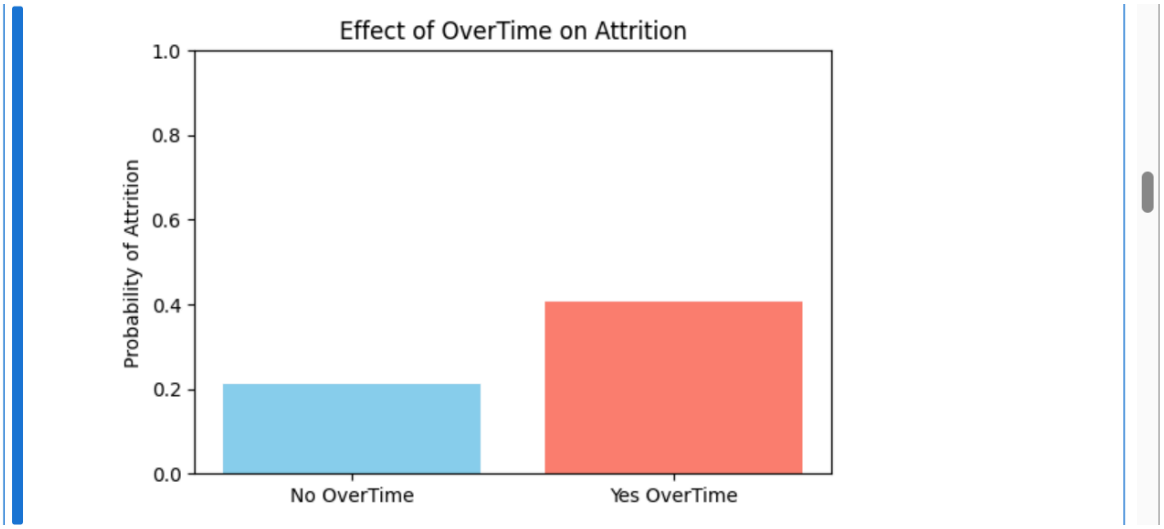
Predictive Analysis (Python - Logistic Regression)

- Built a **Logistic Regression model** to estimate attrition probability.
- **Top drivers of leaving:**
    - **OverTime**
    - **Low Job Satisfaction**
    - **Single employees**
    - Specific **job roles** (e.g., Laboratory Technicians)
- **Protective factors:** High Job Satisfaction, longer tenure (>5 yrs)
- Created **visualizations for recruiter-friendly storytelling:**
    - OverTime vs attrition probability bar chart

- Job Satisfaction Probability Curves
- Feature importance chart

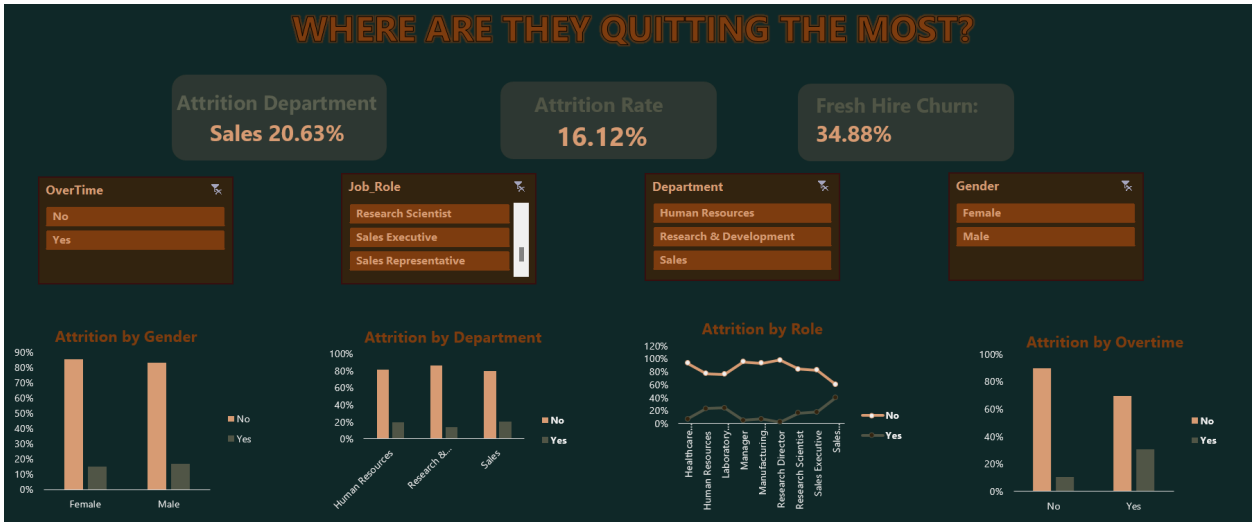
Visualizations:

OverTime Probability Bar Chart



Dashboard Preview

A snapshot of the Excel dashboard showing department-level churn, role-specific risks, and satisfaction breakdowns.



Model Evaluation:

Metric	Class 0 - Stay	Class 1 - Leave
Precision	0.88	0.69
Recall	0.99	0.15
F1-score	0.93	0.24
Accuracy	**0.87**	-

This model predicts employees are likely to **stay** very well, while predicting leaving employees is harder due to class imbalance. Still, it provides actionable insights into **which factors most influence turnover.**

Results/ Insights

- **High-risk segments:** Fresh Hires, Sales Representatives, Overtime employees
- **Predictive drivers:** OverTime, low Job Satisfaction, single employees, and certain job roles.
- **Protective factors:** High Job Satisfaction, Longer Tenure
- **Actionable outcomes:**
  - Monitor **overtime** and implement workload guardrails
  - Enhance **onboarding and early support** for fresh hires
  - Increase **job satisfaction and engagement initiatives** for at-risk employees

Combining **descriptive + predictive analysis,** HR can move from **reactive exits to proactive retention strategies.**

What I'd Explore Next

- Integrate **exit interview data** for qualitative insights
- Explore **seasonal attrition patterns**
- Layer **employee engagement survey data** for richer predictive signals
- Analyze **promotion history and remote vs in-office trends**

Tools Used

- **Microsoft Excel:** Dashboard creation, pivot tables
- **Python:** Pandas, scikit-learn (Logistic Regression), matplotlib (plots)

Why This Project Matters

- Demonstrates ability to combine descriptive and predictive HR analytics
- Highlights actionable insights for decision-making
- Showcases technical skills (Excel + Python) and business understanding

Explore The Project

- GitHub
- Portfolio