



TEXAS McCombs

The University of Texas at Austin
McCombs School of Business

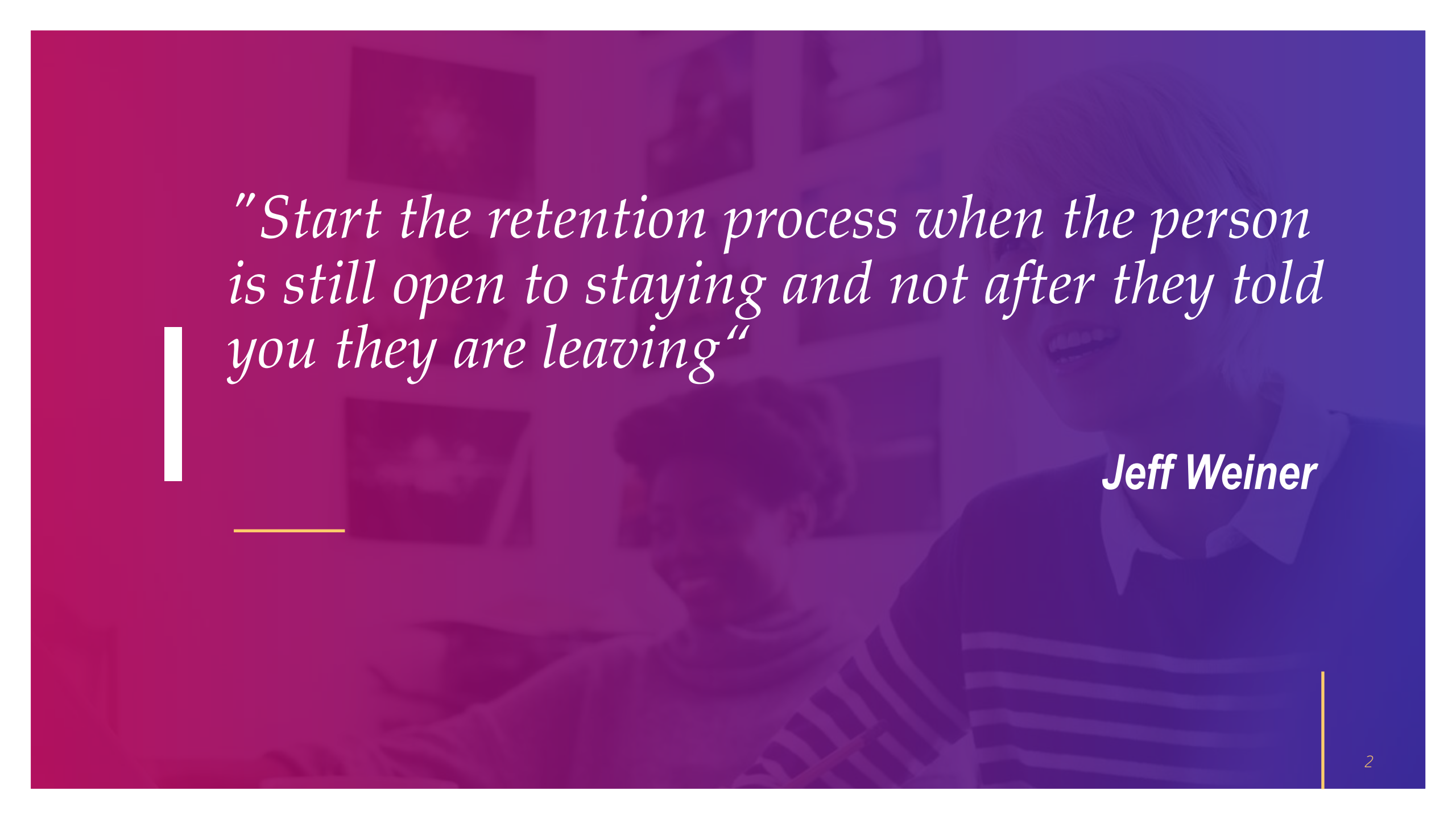
Employee Attrition Analysis and Prediction

Final project elaborated by:

Jessica Barker

Akshaya Sathiya

Younes Smyej



"Start the retention process when the person is still open to staying and not after they told you they are leaving"

Jeff Weiner



Estimated costs of losing an employee:

- 1.5 - 2x the employee's salary
- For **hourly** workers, an average of \$1,500 per employee
- For **technical** positions, 100-150% of salary

Employee Attrition Analysis and Prediction

- Recruitment - finding, talking, choosing, hiring
- Starting off - training costs, time spent
- Less work getting done - it might take 1-2 years for a new person to work as well
- Business goes down - new workers might not be as good at fixing problems and dealing with things
- Culture changes - when someone leaves, others wonder why and feel more tired

| We want to answer...



Why do employees leave?



What factors characterize employee attrition?



What can companies do to prevent losing employees



Employee Attrition

Understanding why employees leave...

- *Unsatisfying compensation and benefits*
- *Lack of development opportunity*
- *Lack of work-life balance*
- *Lack of recognition*
- *Poor management*
- *Poor work conditions*

Our Data

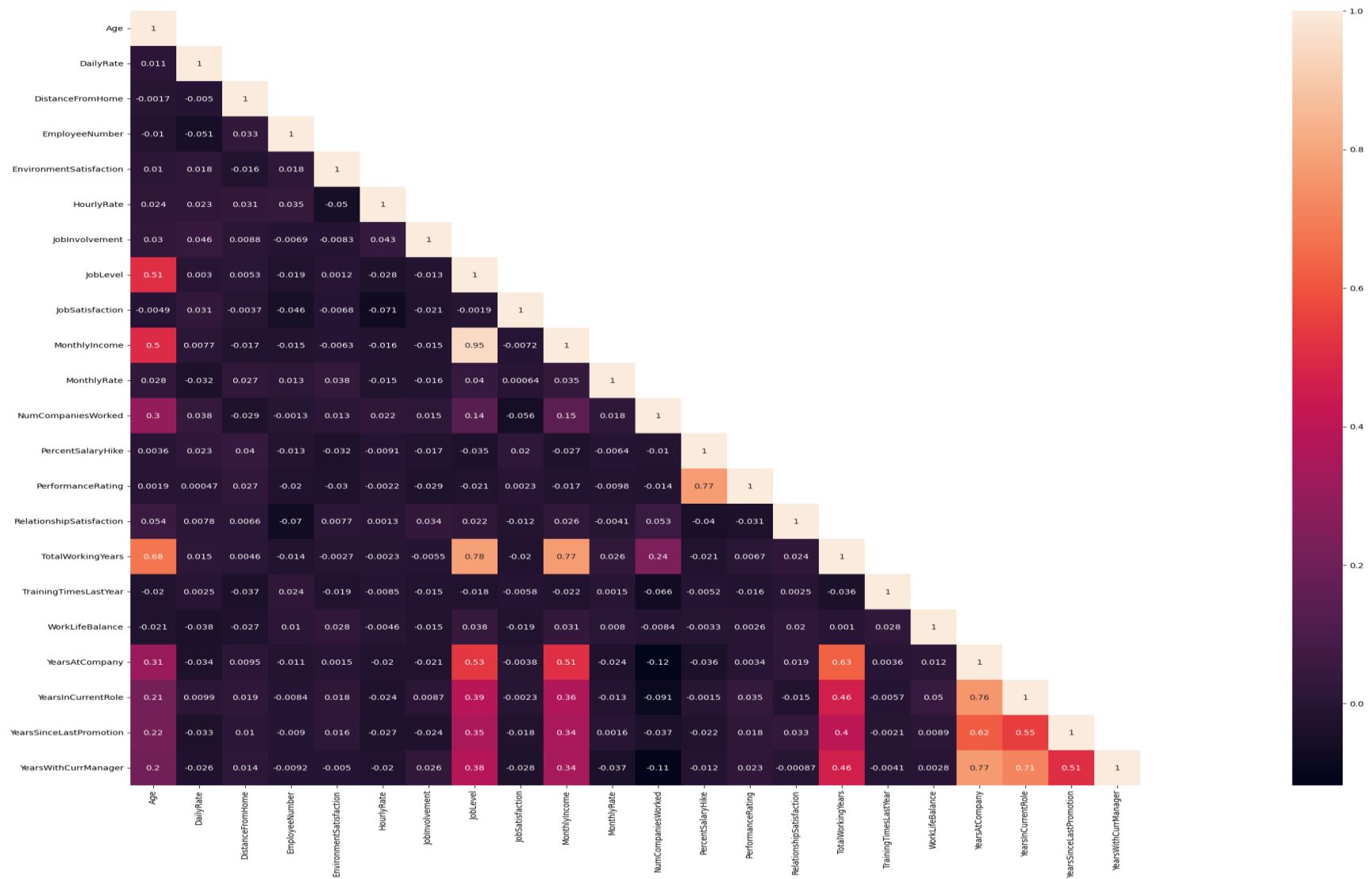


IBM data experts created information about HR employees, including:

- 1,470 instances (entries) and 35 features (details)
- Personal employee info like age, marital status, distance from home, gender, education, etc.
- Answers from surveys about things like job satisfaction and work-life balance
- Work-related data such as monthly income, years at the company, job role, and travel.

Exploratory Data Analysis (EDA)

- Count plot of 'Attrition' distribution shows higher 'No' than 'Yes'.
- Imbalanced classes suggest potential model performance challenges.
- Addressing class imbalance is essential for accurate prediction.
- EDA helps us understand dataset characteristics and informs subsequent analysis.

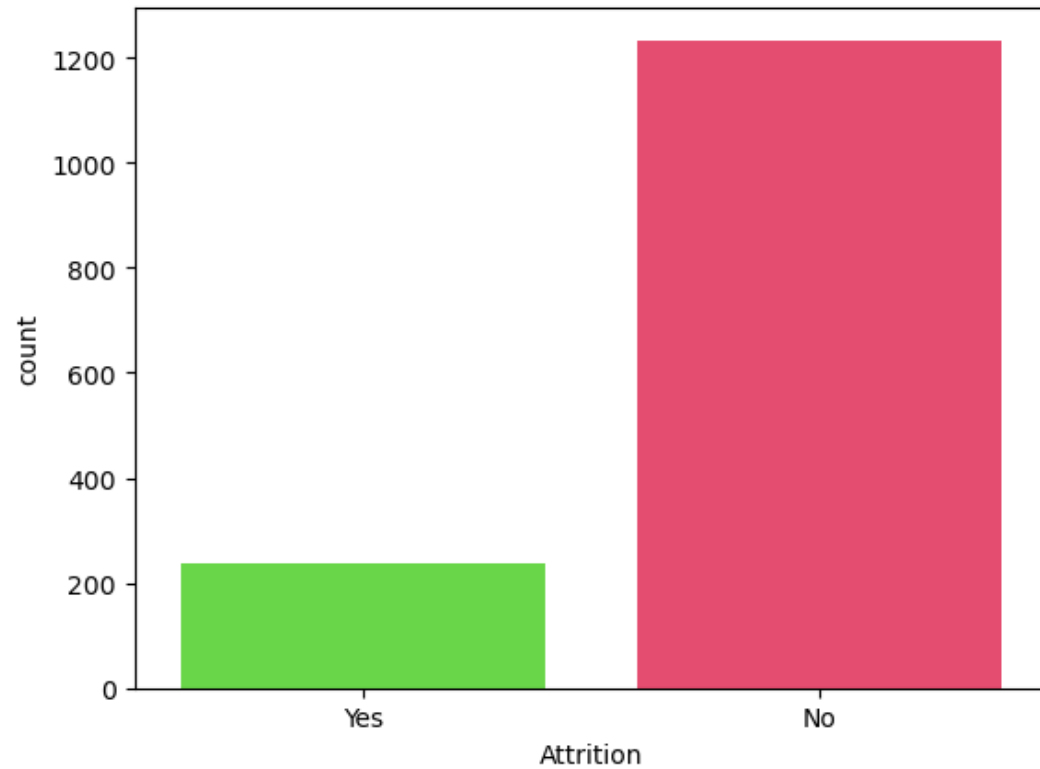


Correlation Matrix

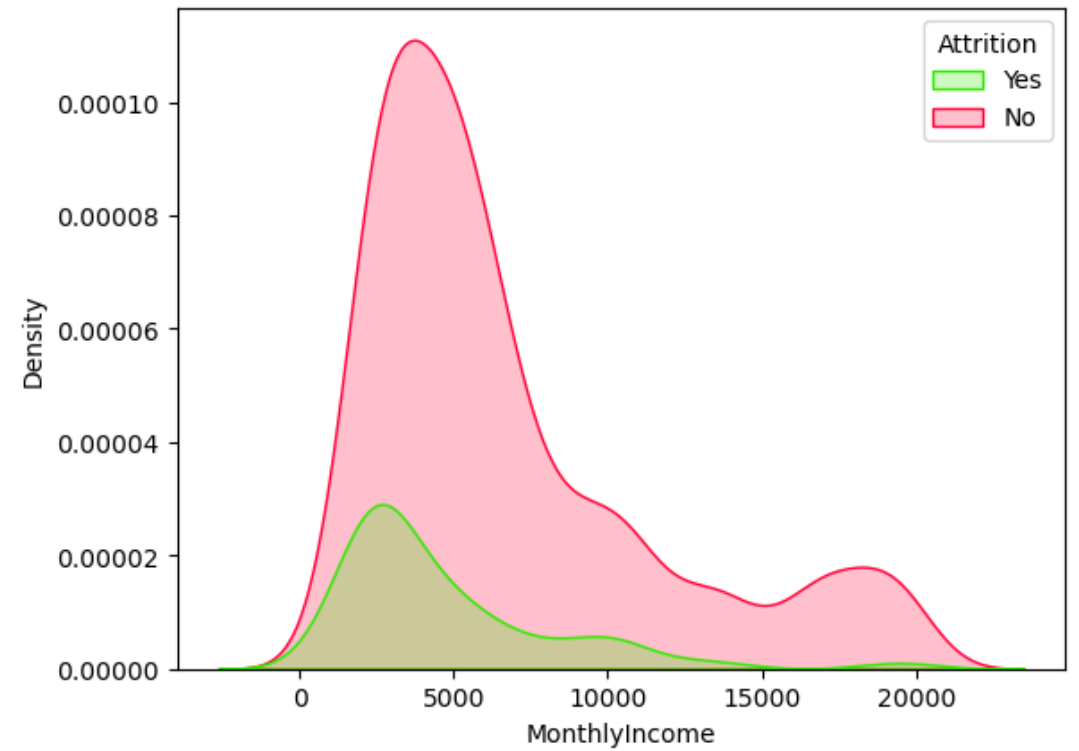
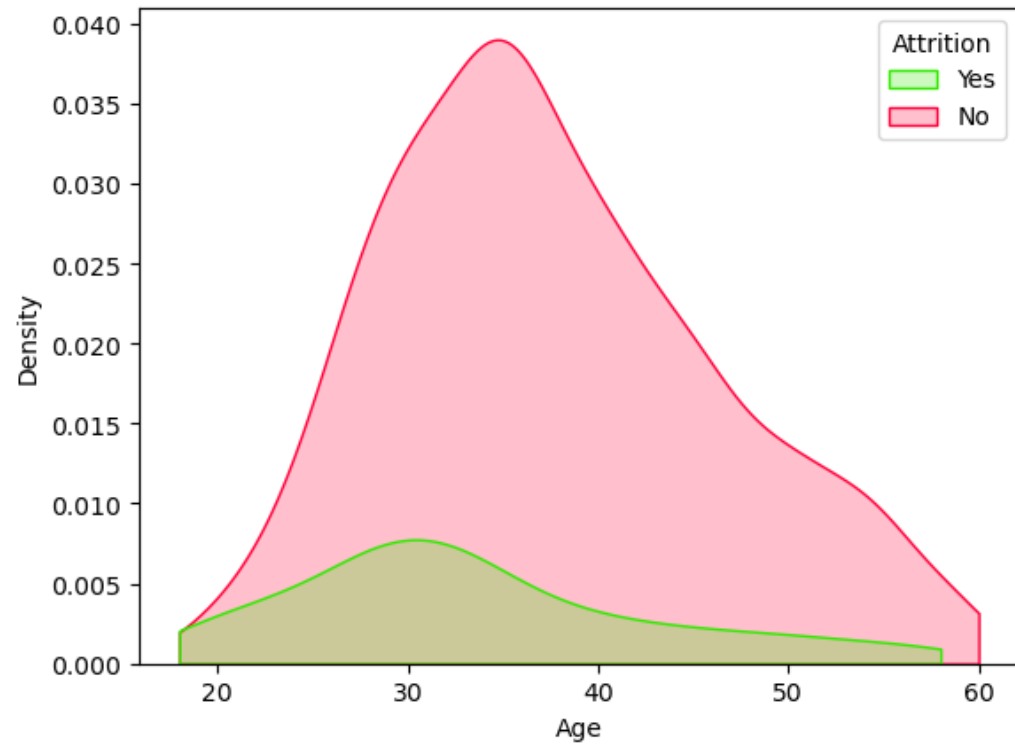
Attrition Analysis by Categories

Analyzing attrition by categories uncovers initial patterns, suggesting potential drivers of employee turnover. These insights guide subsequent analysis and strategy development:

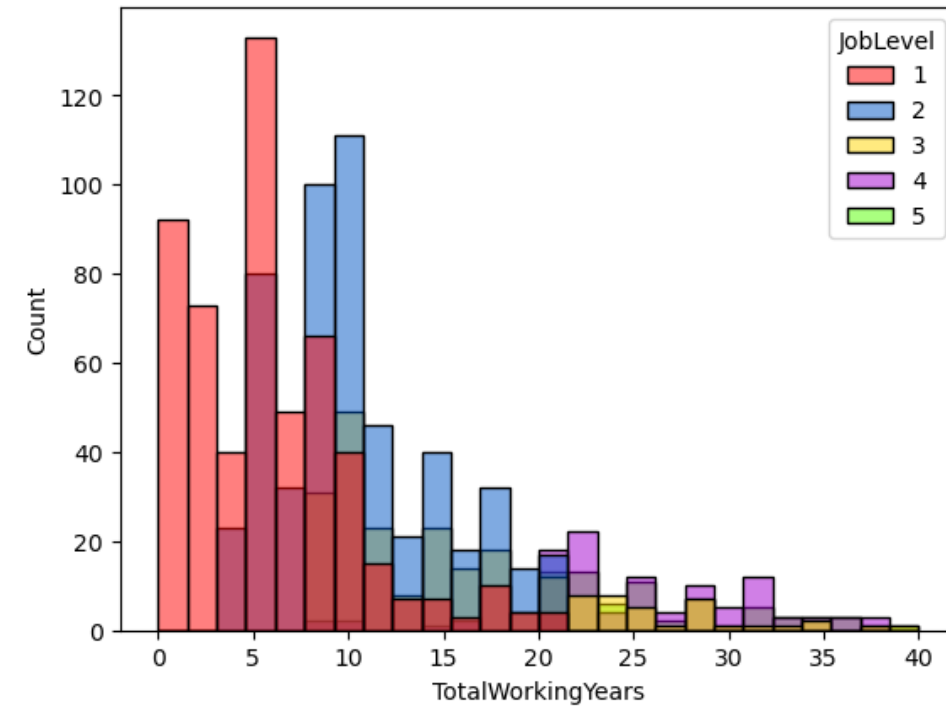
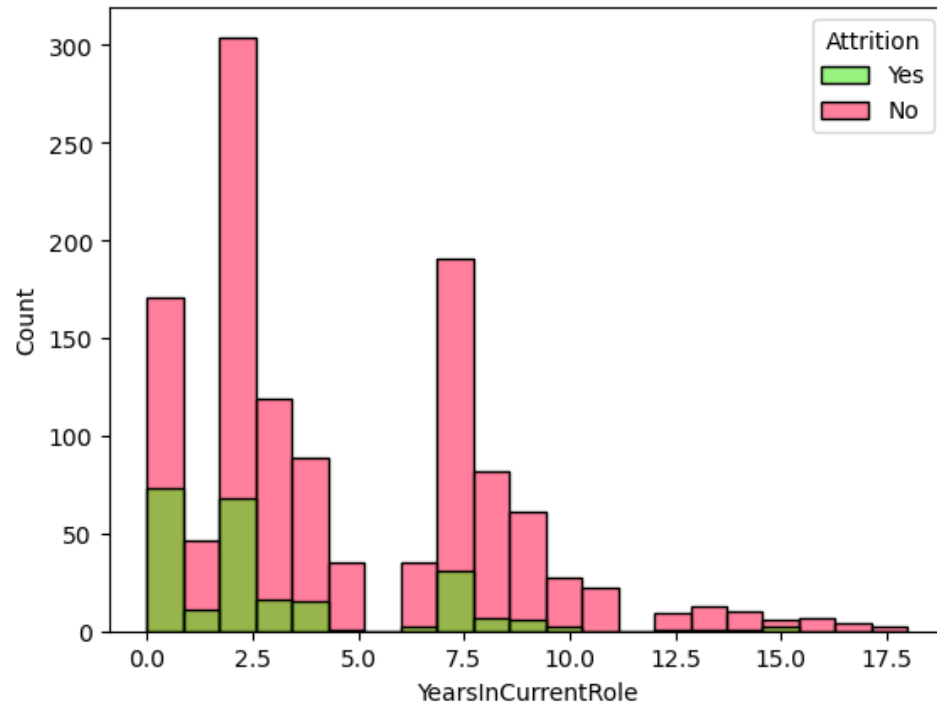
- Gender
- Age
- Years in Current Role
- Total Working Years
- Years since Last Promotion
- Monthly income
- Top 3 Departments with Attrition



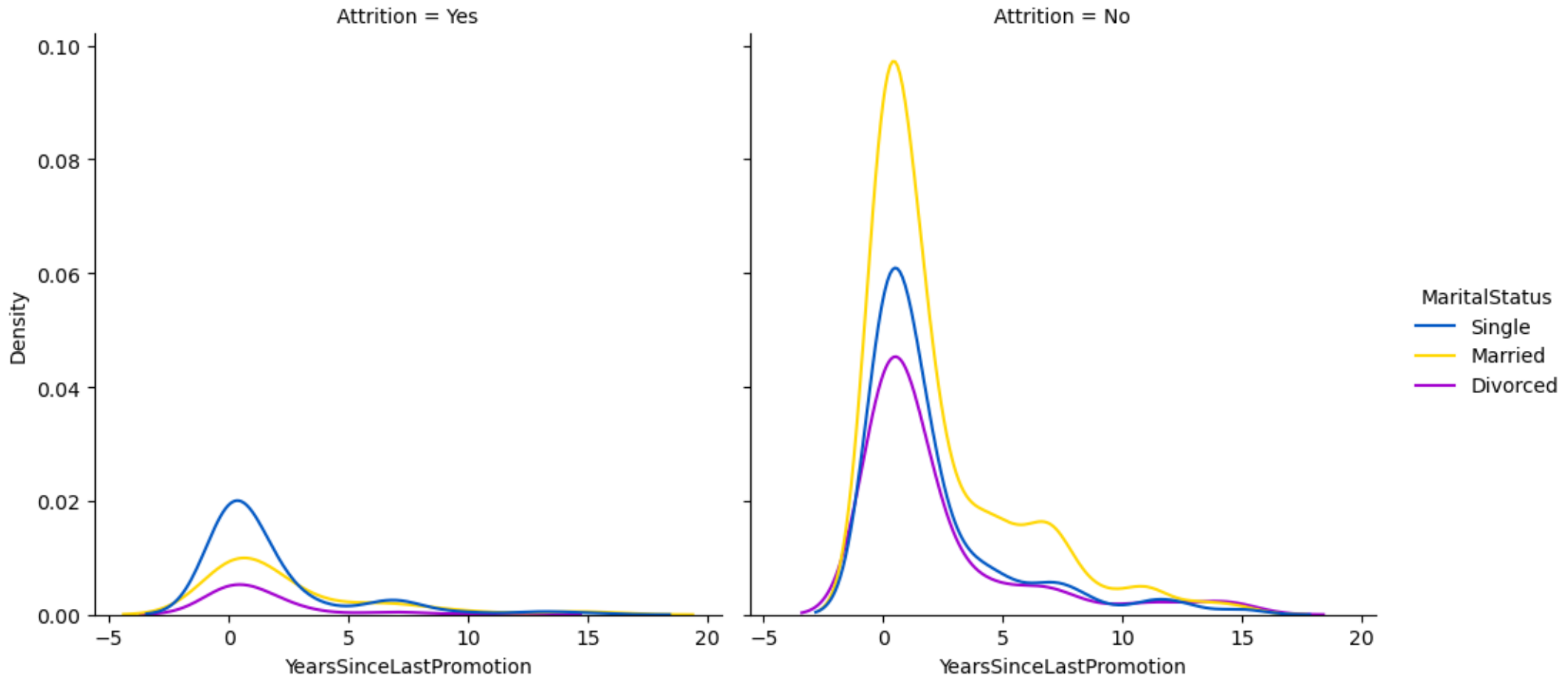
Overall Attrition and Attrition by Gender



Attrition Analysis by Age & Monthly Income

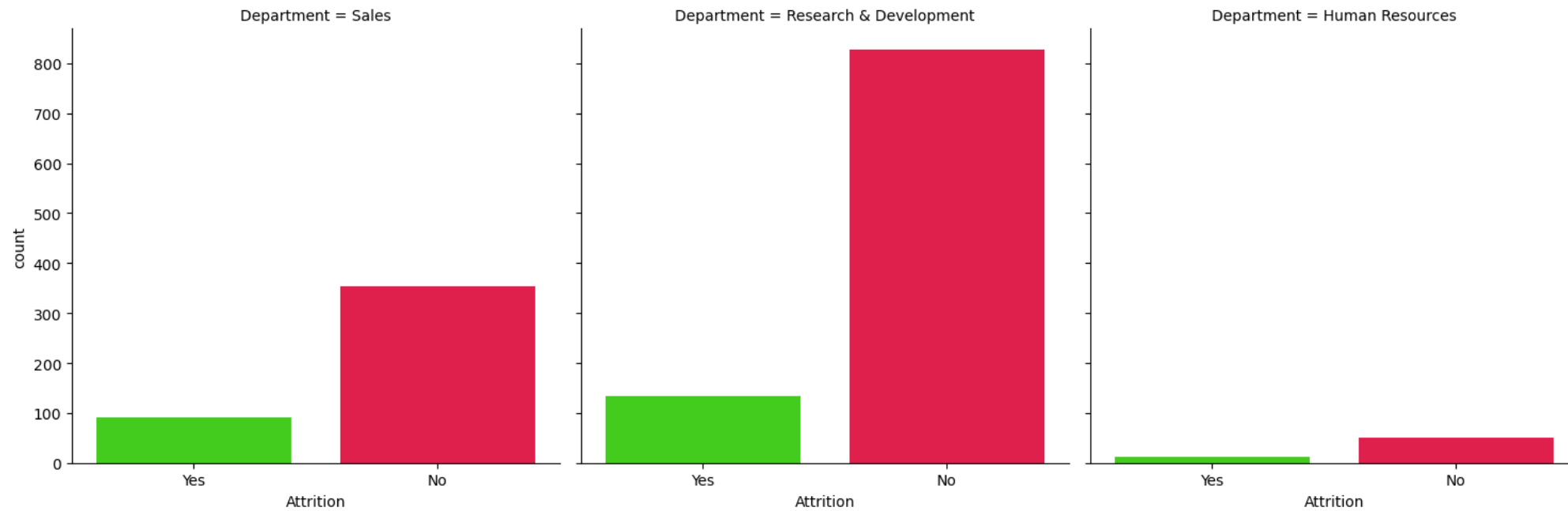


Attrition Analysis by Years in Current Role & Total Working Years



Attrition Analysis by Years since Promotion

Single, Married, and Divorce



Attrition Analysis on the top three Departments

Random Forest

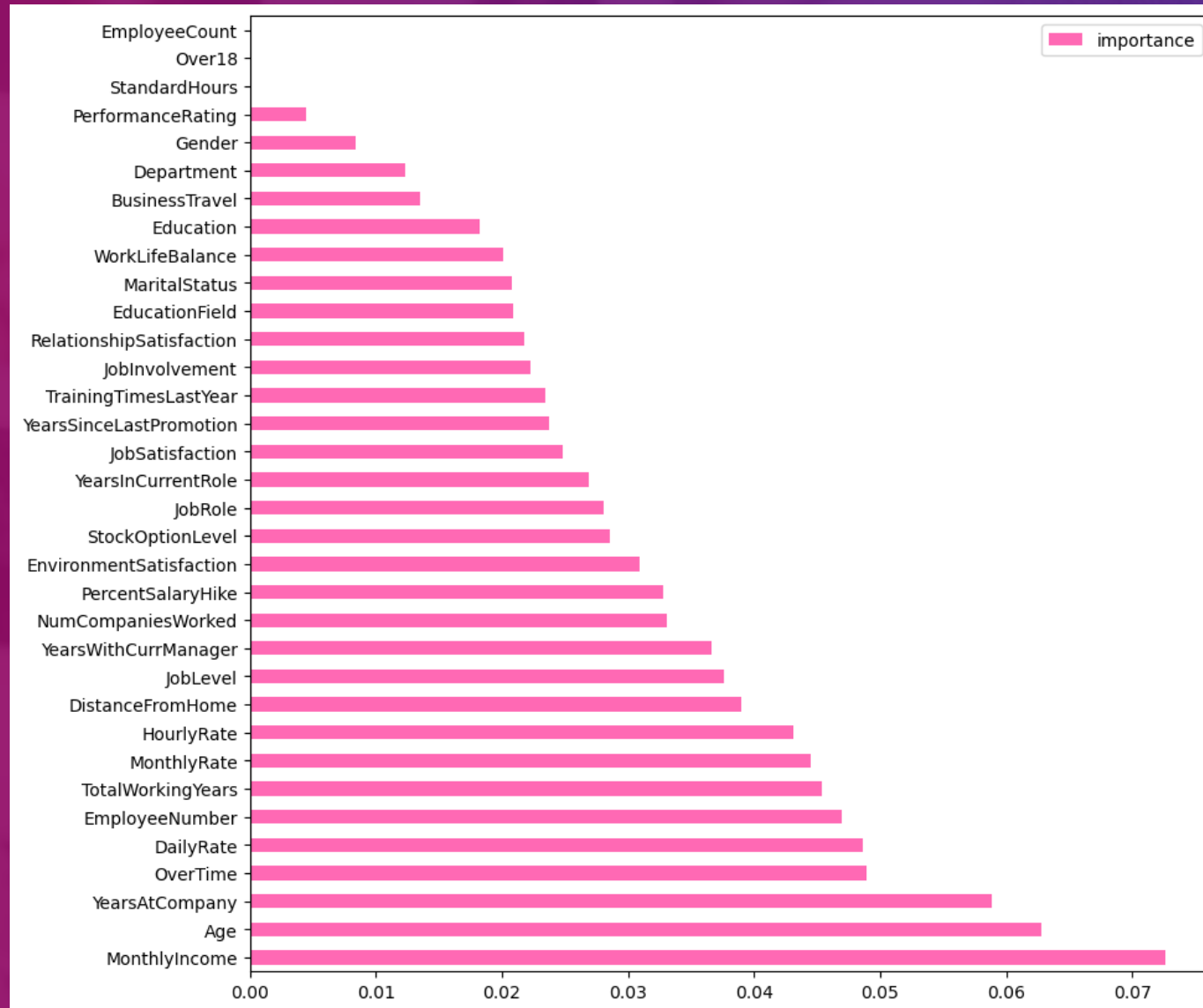
We used this model because it is good for binary classification, and for its ability to determine feature importance, which we planned to use to guide our exploratory analysis. It ended up with an 87% accuracy, but due to the low recall, we did not end up using this model. The classification report, confusion matrix, and feature importance can be seen below:

Accuracy Score : 0.8777173913043478

Classification Report

	precision	recall	f1-score	support
0	0.89	0.98	0.93	321
1	0.58	0.15	0.24	47
accuracy			0.88	368
macro avg	0.74	0.57	0.59	368
weighted avg	0.85	0.88	0.84	368

Random Forest Model Prediction Bar Graph



XGBoost

We used this as a second model because it is also good for binary classification, and ability to handle complex relationships within the data. XGBoost has built-in regularization techniques, such as L1 and L2 regularization, that can help prevent overfitting. We chose to use this mode because it had a higher accuracy rate of 89% and a higher recall rate of 32% .

Accuracy Score : 0.8858695652173914

Classification Report

	precision	recall	f1-score	support
0	0.91	0.97	0.94	321
1	0.60	0.32	0.42	47
accuracy			0.89	368
macro avg	0.75	0.64	0.68	368
weighted avg	0.87	0.89	0.87	368

A woman with dark curly hair and glasses is smiling and looking to the right. The background is a blurred office setting with a window and some framed pictures. A vertical red line is on the right side of the image.

THANK YOU!

Any questions?