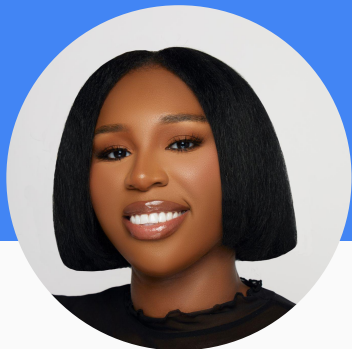


Frito Lay

Employee Attrition Analysis for Frito Lay

The team



Brittany Blackmon

Data Scientist



Tom Cruise

Data Scientist



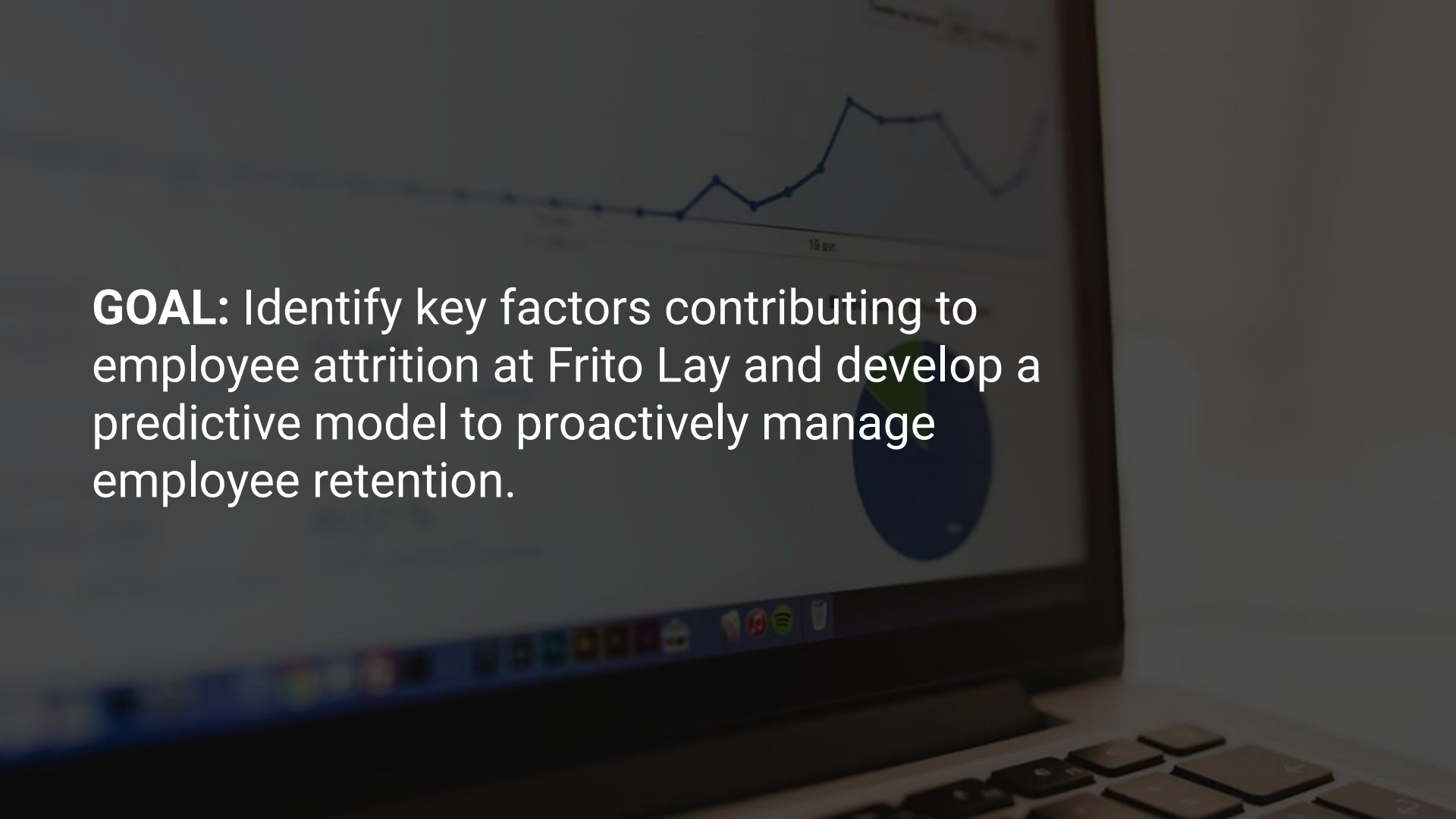
Dolly Parton

Data Scientist



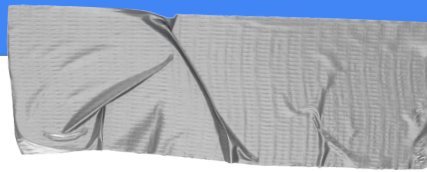
Will Smith

Data Scientist

A laptop screen is shown with a dark overlay. On the screen, there is a line graph with a blue line showing fluctuations over time, and a pie chart with green and blue segments. The text '19 av.' is visible near the graph. A large text overlay in white reads:

GOAL: Identify key factors contributing to employee attrition at Frito Lay and develop a predictive model to proactively manage employee retention.

19 av.



Identifying Key Predictors of Attrition: Problem Statement and Hypothesis

→ **Motivation:**

What is the importance of identifying employees likely to leave?

→ **Problem**

The key question we're addressing is whether employee attrition at Frito Lay can be predicted using factors such as job satisfaction, monthly income, and department.

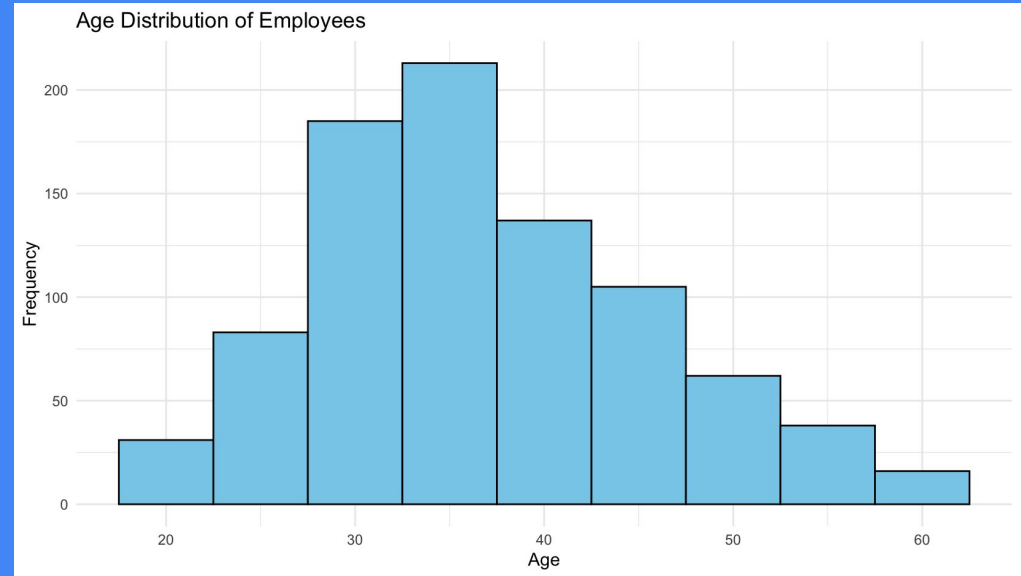
→ **Hypothesis**

We hypothesize that factors such as lower job satisfaction and lower income levels are significantly associated with higher employee attrition.

Our analysis is based on Frito Lay's employee dataset, containing 870 records with 36 variables. Key metrics include age, job satisfaction, income, job role, work-life balance, and whether the employee left (Attrition status).

Key Variables and Summary:

- **Age** ranges from 18 to 60, with an average of 36.8 years.
- **Daily Rate** spans from 103 to 1499, with a median of 818.
- **Monthly Income** ranges between \$1,081 and \$19,999, with an average of \$6,390.
- **Job Satisfaction** scores vary from 1 (low) to 4 (high), with a mean of 2.7.
- **Environment Satisfaction** is also rated from 1 to 4, with an average of 2.7.

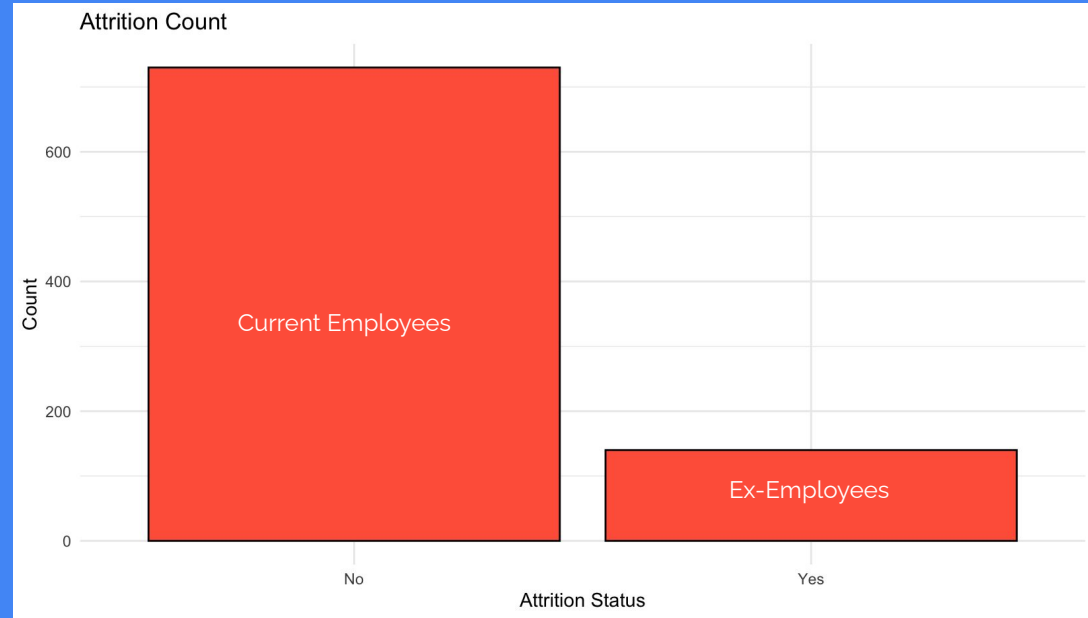


****Attrition** is a binary target variable, with 730 employees labeled as 'No' (they stayed) and 140 labeled as 'Yes' (they left).*

Our analysis is based on Frito Lay's employee dataset, containing 870 records with 36 variables. Key metrics include age, job satisfaction, income, job role, work-life balance, and whether the employee left (Attrition status).

Key Variables and Summary:

- **Age** ranges from 18 to 60, with an average of 36.8 years.
- **Daily Rate** spans from 103 to 1499, with a median of 818.
- **Monthly Income** ranges between \$1,081 and \$19,999, with an average of \$6,390.
- **Job Satisfaction** scores vary from 1 (low) to 4 (high), with a mean of 2.7.
- **Environment Satisfaction** is also rated from 1 to 4, with an average of 2.7.



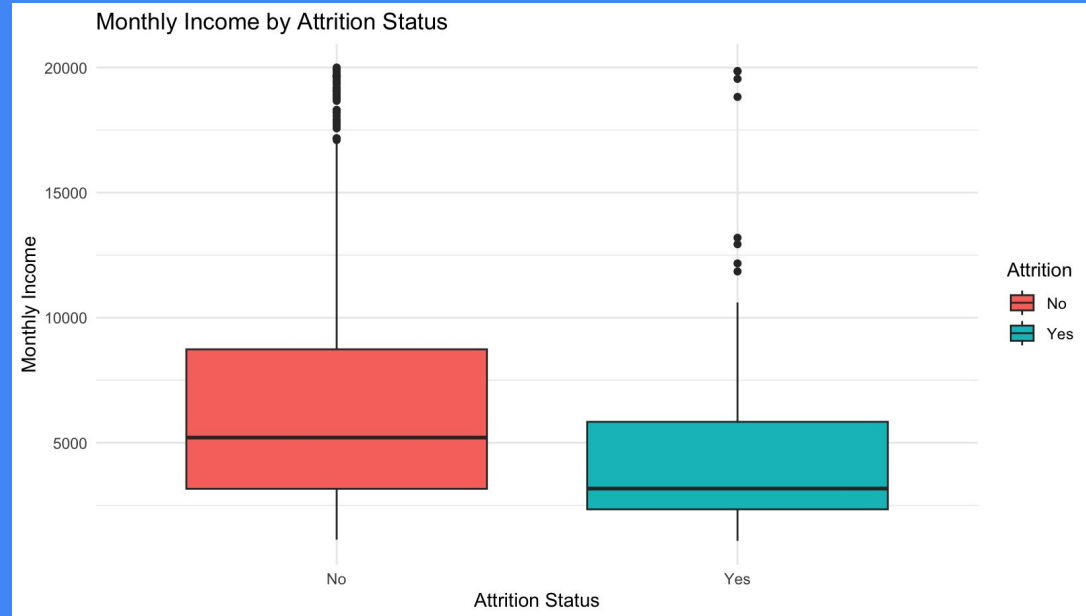
The Distribution of Incomes is Greater Among Employees Who Stayed vs. Those Who Didn't

**Attrition is a binary target variable, with 730 employees labeled as 'No' (they stayed) and 140 labeled as 'Yes' (they left).*

Our analysis is based on Frito Lay's employee dataset, containing 870 records with 36 variables. Key metrics include age, job satisfaction, income, job role, work-life balance, and whether the employee left (Attrition status).

Key Variables and Summary:

- **Age** ranges from 18 to 60, with an average of 36.8 years.
- **Daily Rate** spans from 103 to 1499, with a median of 818.
- **Monthly Income** ranges between \$1,081 and \$19,999, with an average of \$6,390.
- **Job Satisfaction** scores vary from 1 (low) to 4 (high), with a mean of 2.7.
- **Environment Satisfaction** is also rated from 1 to 4, with an average of 2.7.



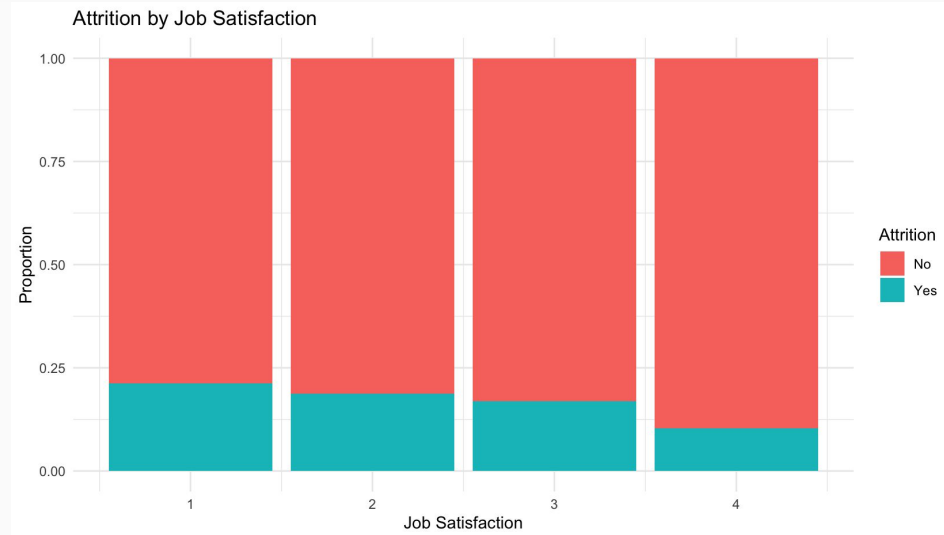
****Attrition** is a binary target variable, with 730 employees labeled as 'No' (they stayed) and 140 labeled as 'Yes' (they left).*

Exploratory Data Analysis: Attrition by Key Variables

Overview: Attrition by Job Satisfaction

Observation: Employees with lower job satisfaction scores (1 or 2) tend to leave more frequently than those with higher satisfaction scores (3 or 4).

Key Insight: Job satisfaction appears to have a clear relationship with attrition. The proportion of employees who leave decreases as job satisfaction increases, indicating that improving job satisfaction could help reduce attrition.



Both income and job satisfaction are key predictors of attrition. Employees with lower incomes and lower job satisfaction are more likely to leave the company. These findings support the hypothesis that financial and emotional incentives play a critical role in employee retention.

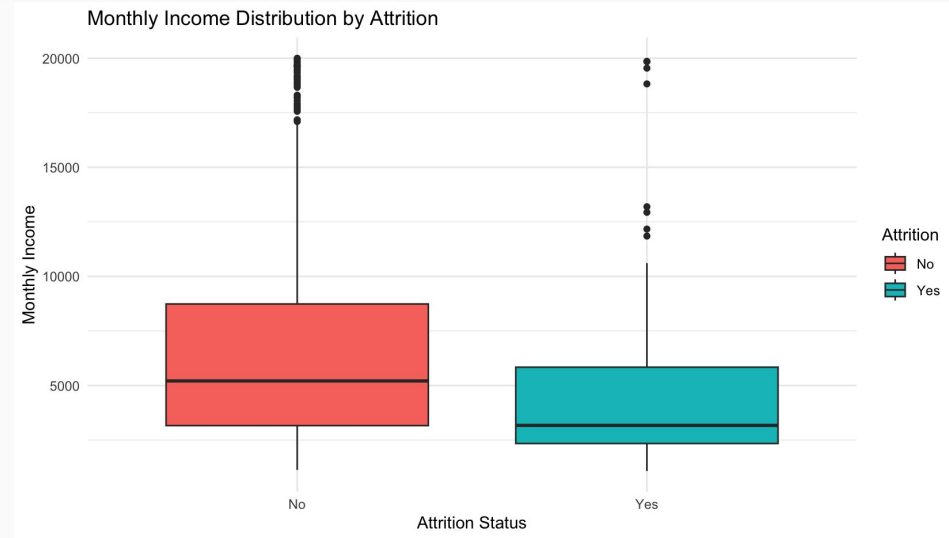
Exploratory Data Analysis: Attrition by Key Variables

Overview: Monthly Income Distribution by Attrition

Observation: Employees who stayed (Attrition = "No") generally have higher median monthly incomes compared to those who left (Attrition = "Yes").

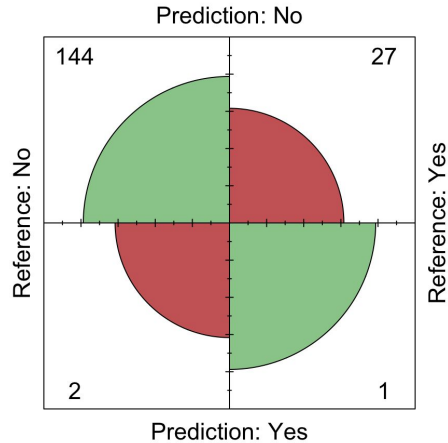
Key Insight: The employees with lower monthly income levels tend to leave more frequently. There is a visible difference in the median income, showing that financial compensation may be a significant factor influencing attrition.

Outliers: Both groups have outliers, but those who stayed have a wider range of high incomes.



Both income and job satisfaction are key predictors of attrition. Employees with lower incomes and lower job satisfaction are more likely to leave the company. These findings support the hypothesis that financial and emotional incentives play a critical role in employee retention.

Confusion Matrix for KNN Model

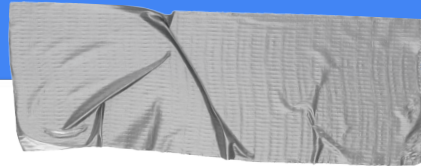


Results & Cost Analysis of the KNN Model

- **Objective:** Use KNN to classify employee attrition based on key factors like **JobSatisfaction** and **MonthlyIncome**.
- **Data Preparation:**
 - Split data into 80% training and 20% testing sets.
 - Ensured consistent factor levels for **Attrition** in both datasets.
- **Model Configuration:**
 - Set $k = 11$ for KNN.
- **Initial Accuracy:** Achieved 83.3% overall accuracy.
- **Limitations:**
 - The model's sensitivity is low (3.57%), meaning it struggles to accurately detect employees likely to leave (attrition cases).

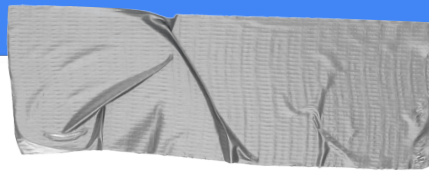
GBM is the preferred model due to its ability to detect attrition cases effectively (sensitivity of 60.71% vs. 3.57% in KNN)

METRIC	KNN	GBM
Overall Accuracy	83.3%	Not specified in final, typically high
Sensitivity	3.57%	60.71%
Strengths	Simplicity, Interpretability	High sensitivity, can capture complex relationships
Limitations	Poor sensitivity to attrition cases, struggles with imbalance	Longer training time, more complex to interpret



Comparison of the Models

- **Objective:** Classify employee attrition with emphasis on accurately identifying employees likely to leave.
- **KNN:**
 - **Selected Features:** Key factors like **JobSatisfaction** and **MonthlyIncome**.
 - **Performance Metrics:**
 - **Accuracy:** Achieved an overall accuracy of **83.3%**.
 - **Sensitivity:** The model's sensitivity was **3.57%**—extremely low. This indicates that while the KNN model performed well in overall accuracy, it struggled significantly with detecting employees likely to leave (true positive rate for attrition).
- **GBM:**
 - **Selected Features:** Expanded to include **JobSatisfaction**, **MonthlyIncome**, **YearsAtCompany**, **OverTime**, and **EnvironmentSatisfaction**.
 - **Performance Metrics:**
 - **Achieved Sensitivity:** 60.71%, which is a substantial improvement over KNN.



Results Interpretation


Based on the comparison between the KNN and Gradient Boosting Machine (GBM) models, here is an updated analysis and conclusion for the attrition prediction task:

KNN

- **True Positives:** 1 instance of correctly predicted attrition.
- **False Positives:** 2 instances where non-attrition was predicted as attrition.
- **True Negatives:** 144 instances of correctly predicted non-attrition.
- **False Negatives:** 27 instances where attrition was incorrectly classified as non-attrition.

GBM

- **True Positives:** 17 instances of correctly predicted attrition.
- **False Positives:** 14 instances where non-attrition was predicted as attrition.
- **True Negatives:** 132 instances of correctly predicted non-attrition.
- **False Negatives:** 11 instances where attrition was incorrectly classified as non-attrition.



Key Insights and Actionable Recommendations

→ Conclusion

- The gradient boosting model is a more effective tool for predicting employee attrition due to its superior sensitivity. Implementing this model could enable Frito Lay to proactively address factors contributing to employee turnover, fostering a more stable workforce and reducing replacement costs

→ Recommendations

- **Enhance Work-Life Balance and Career Development:** Provide **flexible scheduling** and **remote work options** to help employees balance work and personal responsibilities, especially in demanding roles.
- **Develop Data-Driven Retention Strategies:** Focus on **competitive compensation reviews** and **performance-based incentives** for Sales and low-income employees to address specific retention needs.
- **Implement Targeted Support for High-Risk Departments:** Introduce **skill-building workshops** to empower employees in high-stress departments, fostering both confidence and career advancement.

Thank You