

August 2019
Clark Consulting

Frito Lay Employee Analysis



Outline

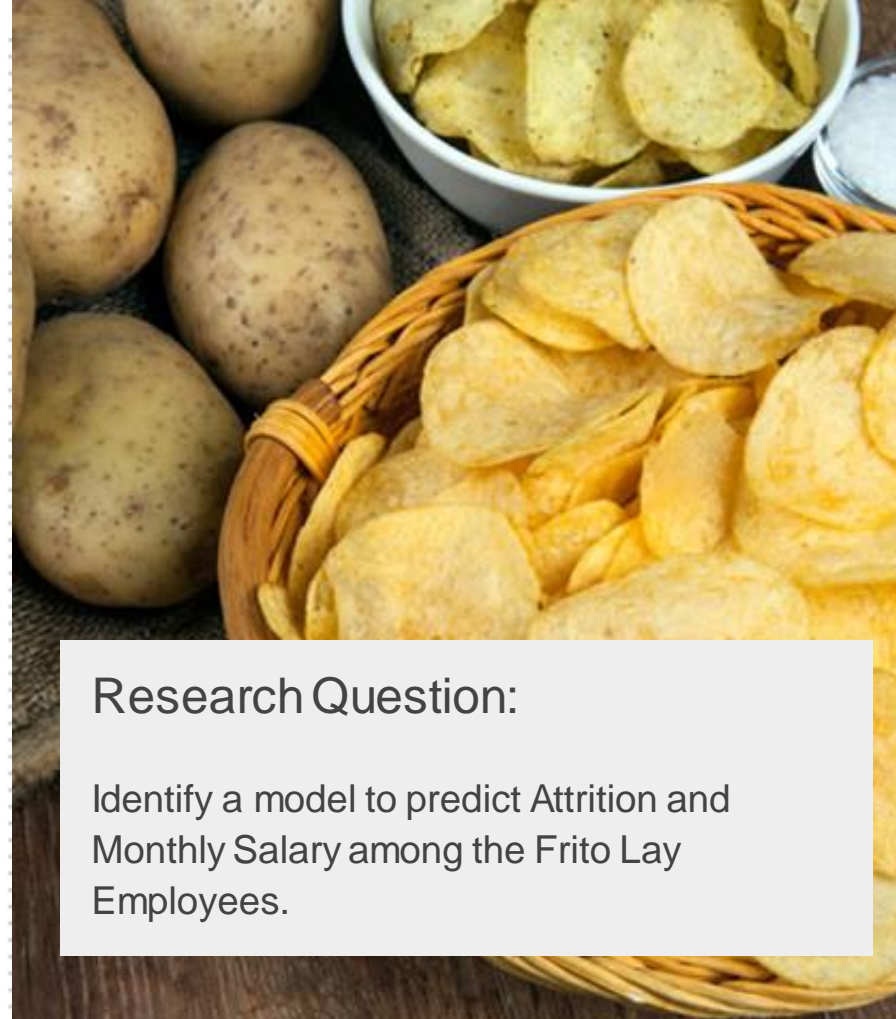
- **Business Question and Situation Analysis**
- **Data Overview and Cleaning**
- **Salary Prediction**
- **Salary Variable Importance**
- **Attrition Prediction**
- **Attrition Variable Importance**
- **Conclusion**



Potato Chips At a Glance

In 2016, the U.S. Potato Chip industry grew to \$7.74 billion and is expected to grow consistently over the next 10 years.

With new competitors emerging in the industry, it's important to ensure Frito Lay maintains top talent to maintain an edge



Research Question:

Identify a model to predict Attrition and Monthly Salary among the Frito Lay Employees.

Data Overview

Data Available:

Traindata.csv - dataset with 870 different employees at Frito Lay with information on salary, performance and company sentiment.

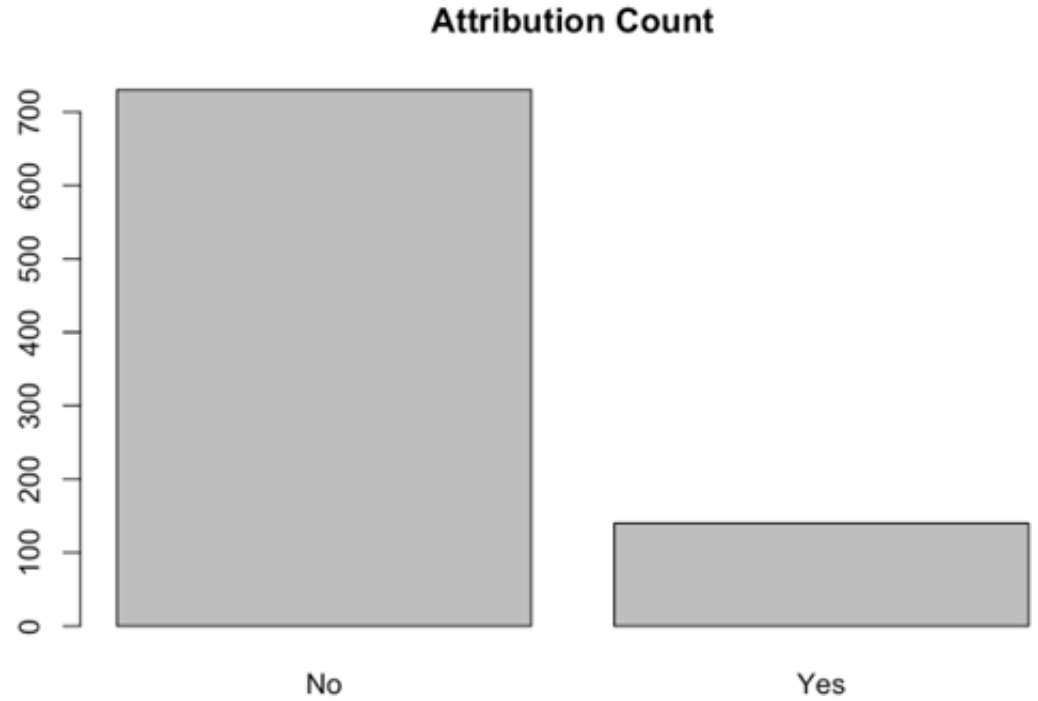
TestSalary.csv - dataset of 300 employees with the same information as Traindata, but with Salary information removed for testing

TestAttrition - dataset of 300 employees with the same information as Traindata, but with Attrition information removed



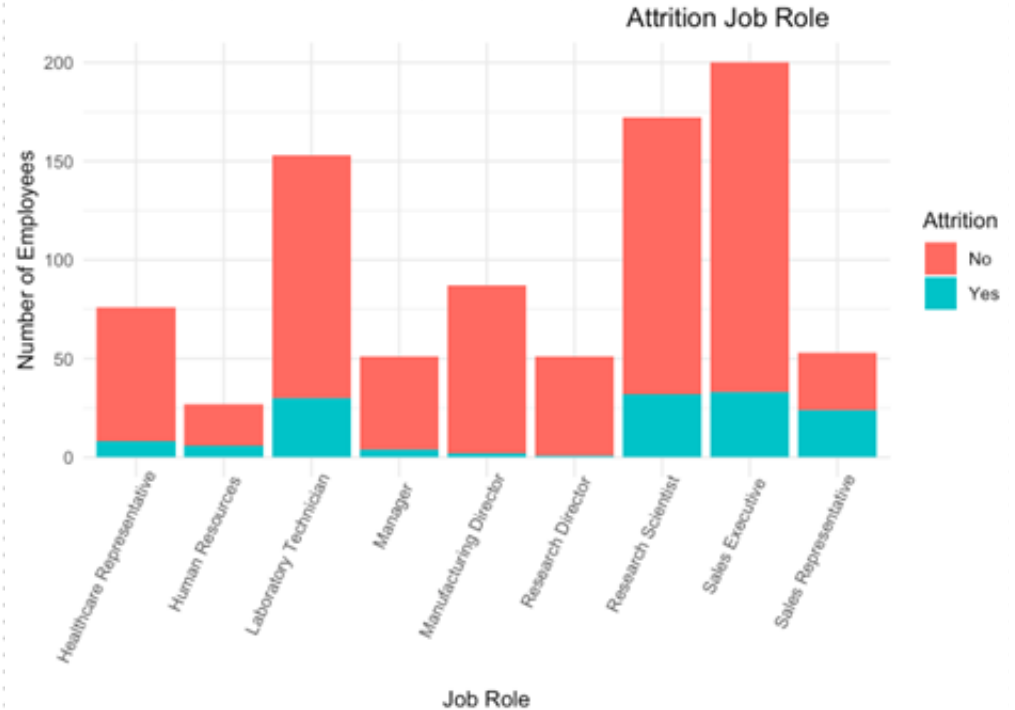
Employee Data Analysis

- In the year of the study, employee attribution was roughly 16%

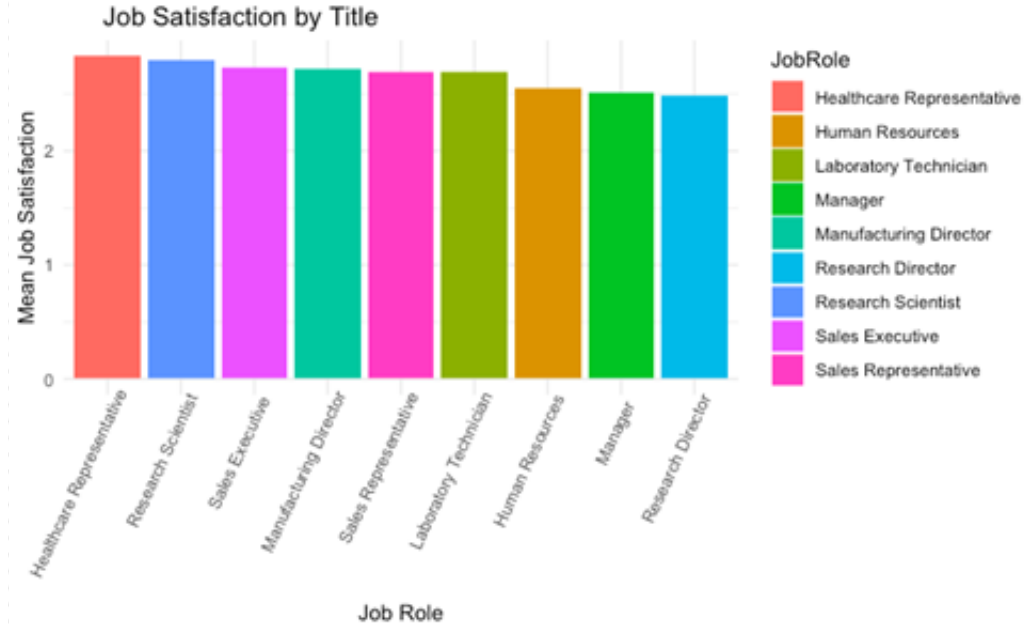


Employee Data Analysis

- Attrition is the highest among Sales Executives and Laboratory Technicians.



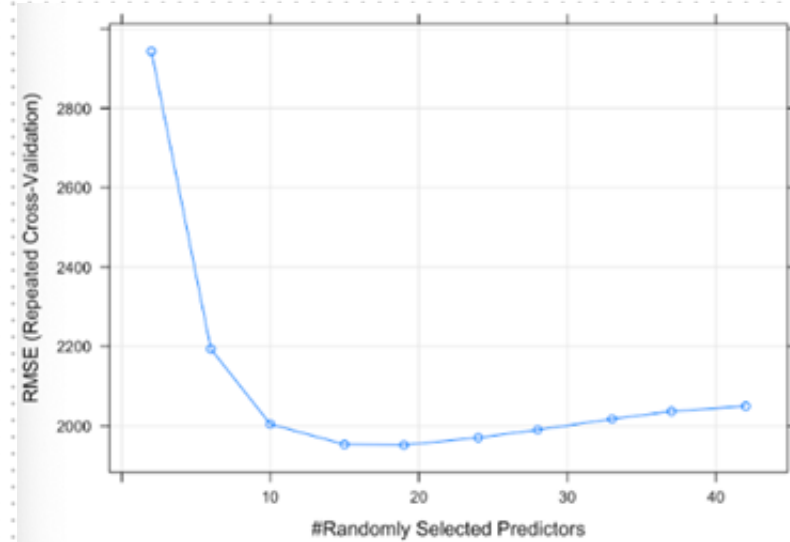
Employee Data Analysis



- Healthcare Representative and Research Scientists have the highest Job satisfaction while managers have the lowest.

Salary Prediction

	K Nearest Neighbor	Random Forest	Support Vector Method
RMSE	\$2,299.51	\$1,951.66	\$2,035.27



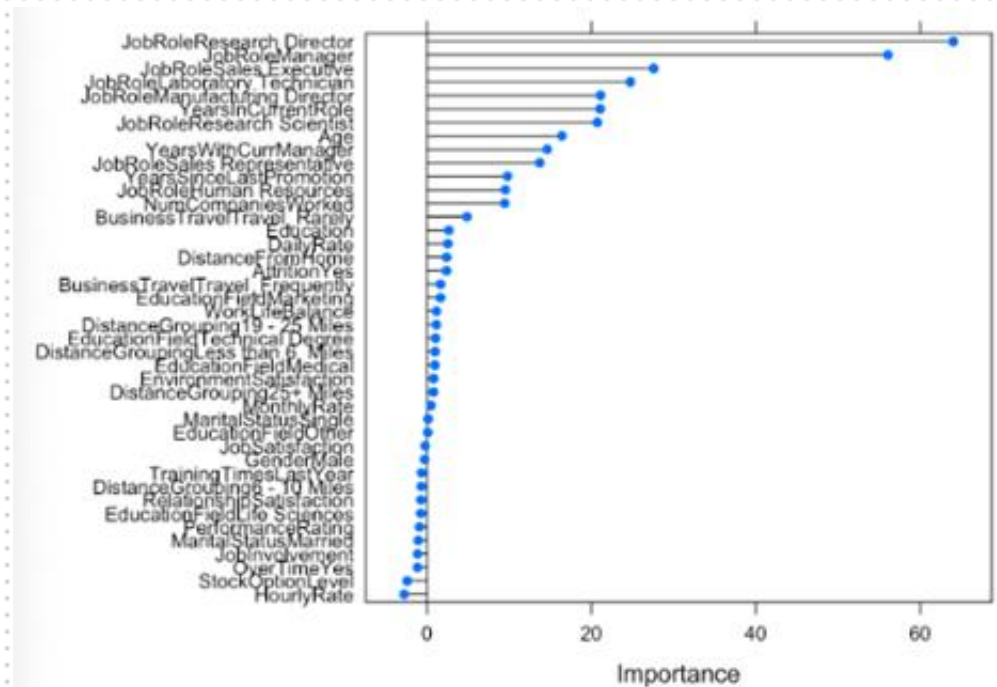
Salary Prediction Feature Importance

- Top Indicators

- Job Role

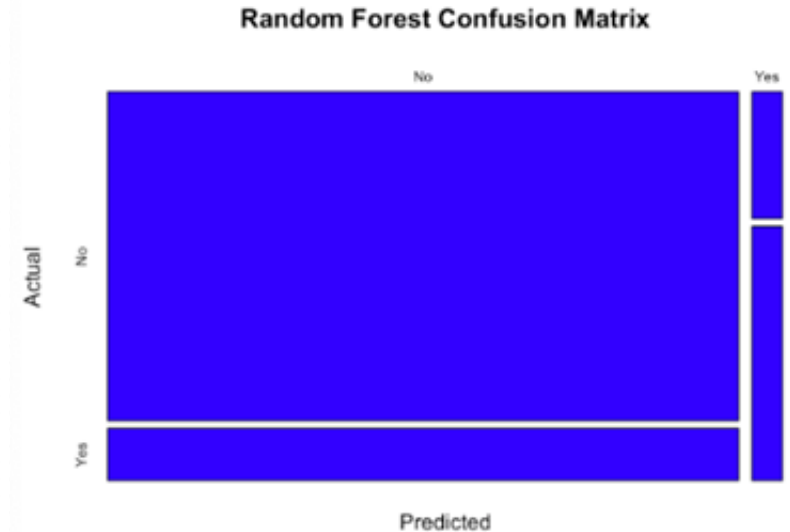
- Years in Current Role

- Age



Attrition Prediction

	Naive Bayes	Random Forest	Support Vector Method
Sensitivity / Specificity	0.88 / 0.45	0.86 / 0.67	0.85 / 0.57



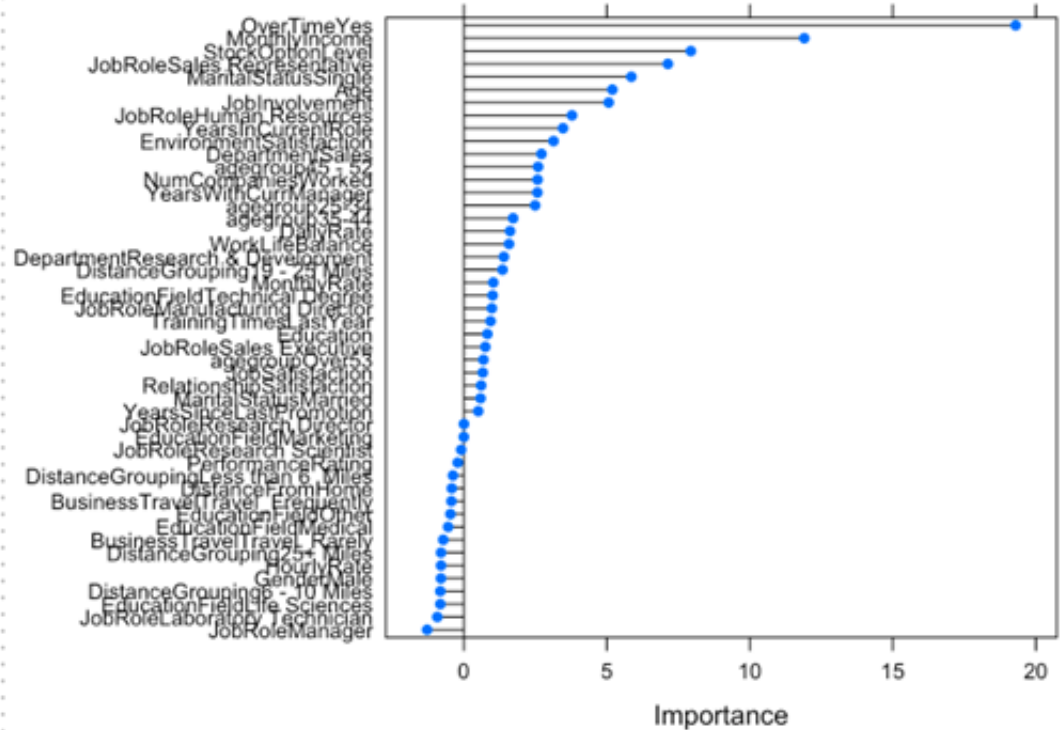
Attrition Feature Importance

- Top Indicators

- Overtime Workers

- Monthly Income

- Stock Option Level



Conclusion

- **Employ the RF model to predict when new employees are meeting the indicators for high attrition and provide a counter offer to entice them to stay.**

- **Employ the RF model to show employees how their salaries are slated to grow at a company to help them better control their futures.**



Appendix



Salary Cleaning

Looking at the interaction rate of all the variables, we can see that the following variables show to have high collinearity per our model, which we can remove. For the variables where we are seeing high collinearity among categorical variables, we will leave in. This is all validated through our significant p value of <0.0001 .

Below are the variables with high collinearity

- Age/Age Group
- Distance Grouping
- Department



Presentation Videos



-Daniel Clark: <https://youtu.be/YTCdQHtOVpw>