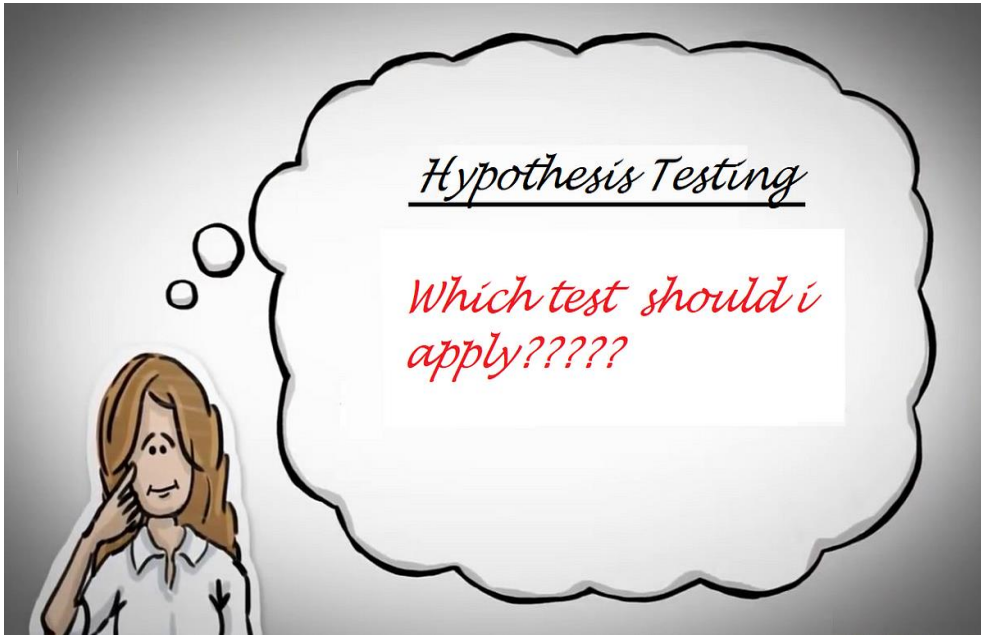# Restaurant Data Analysis

Yash Namdeo Nikhare

Hypothesis Testing

Identifying relationships between variables

This information can be used to improve the model's performance.

Kruskal-Wallis test doesn't rely on data conforming to a specific distribution, making it suitable for non-normally distributed datasets

Inspection type do not affect the Next Inspection Grade C or Below

Restaurant Categories affect the Next Inspection Grade C or Below

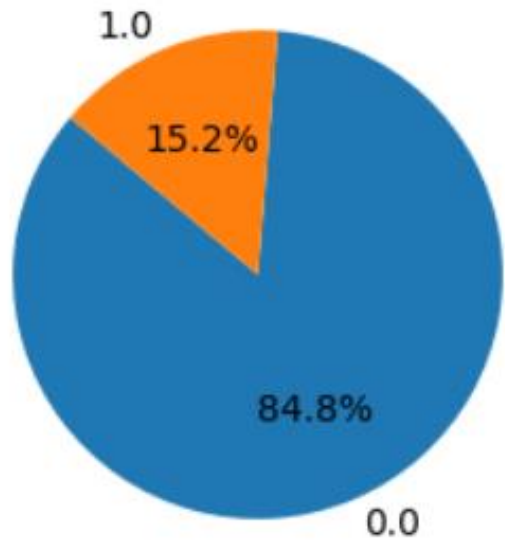**Median Employee Age** have an affect on Current Grade

Current Grades have significant affect on Next Inspection Grade C

Current Demerit affects the Next Inspection Grade C or Below

Data Imbalance

SMOTE Tomek

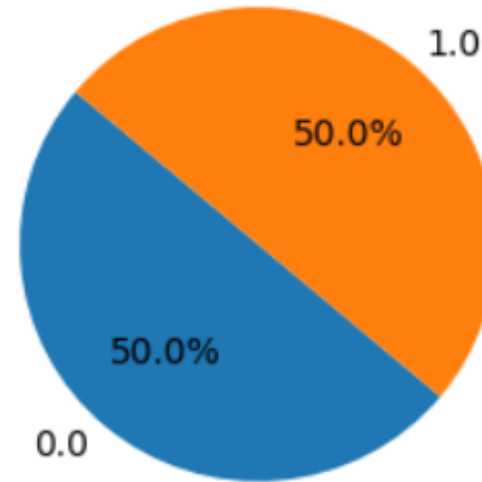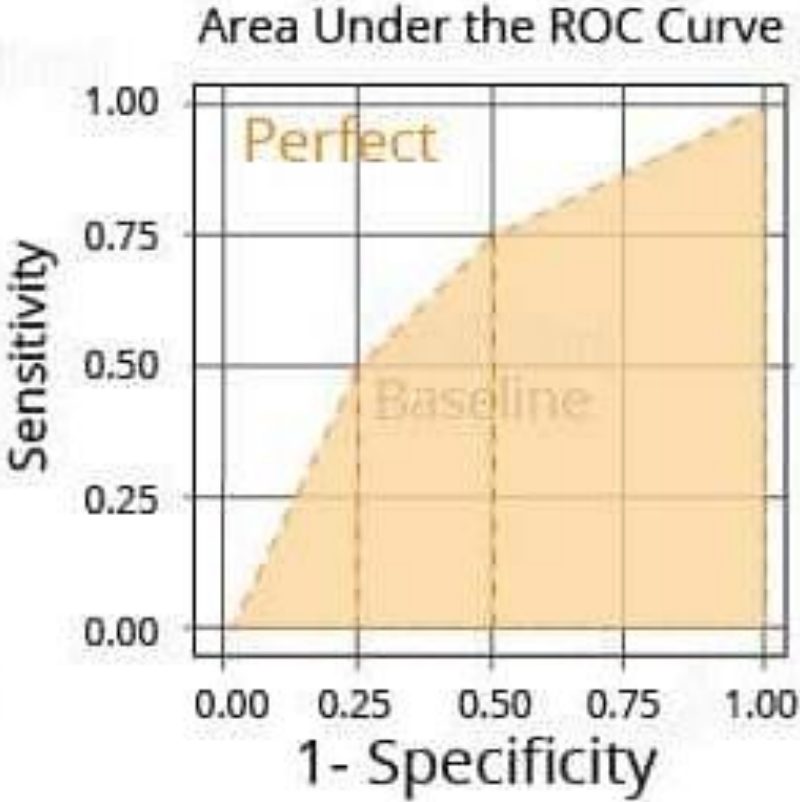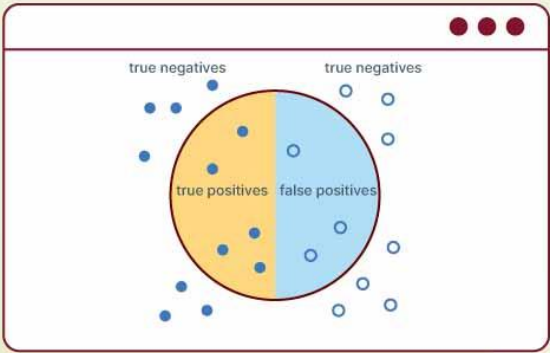Problem with Classification Model

Evaluation Metrics

Range of F1 score=[0,1]

Range of ROC-AUC curve=[0,1]

Algorithm	F1 score on training data	F1 score on test data	ROC-AUC score on training data	ROC-AUC score on test data
Catboost	0. 8872	0.9034	0.9487	0.9108
XGBoost	0.8783	0.8978	0.9427	0.8993

Top 4 Factors Affecting the Model for Prediction

Current Demerit

This key insight emphasizes the crucial role of maintaining a low demerit count for restaurants aspiring to achieve good inspection grades.

Current Grade

The current grade assigned to a restaurant serves as a pivotal indicator of its overall compliance and adherence to health and safety standards.

Beverage and Bar Service

The higher likelihood of next inspection grade C may be from challenges associated with alcohol service. Inspections focus alcohol practices, and overall compliance with health and safety standards in these establishments.
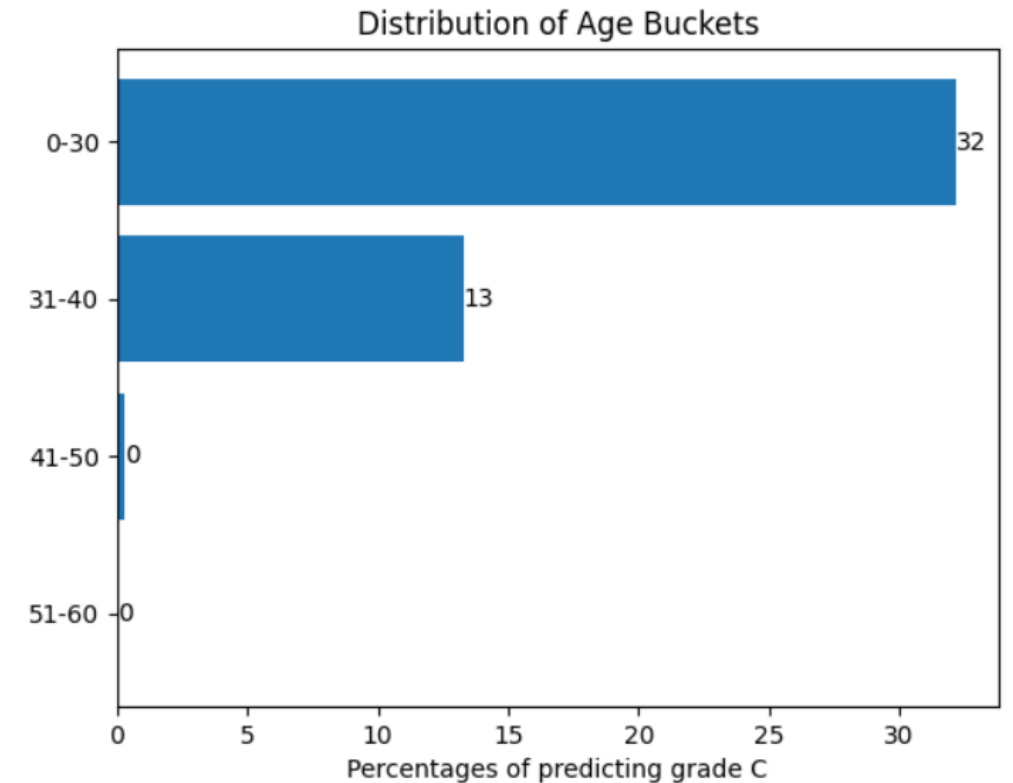
2nd Violation

It has the largest number Major violation categories .

Performance of Cat Boost Model on Unseen Data

Distribution of Restaurant Categories

Total number of Retail Food Outlets =5
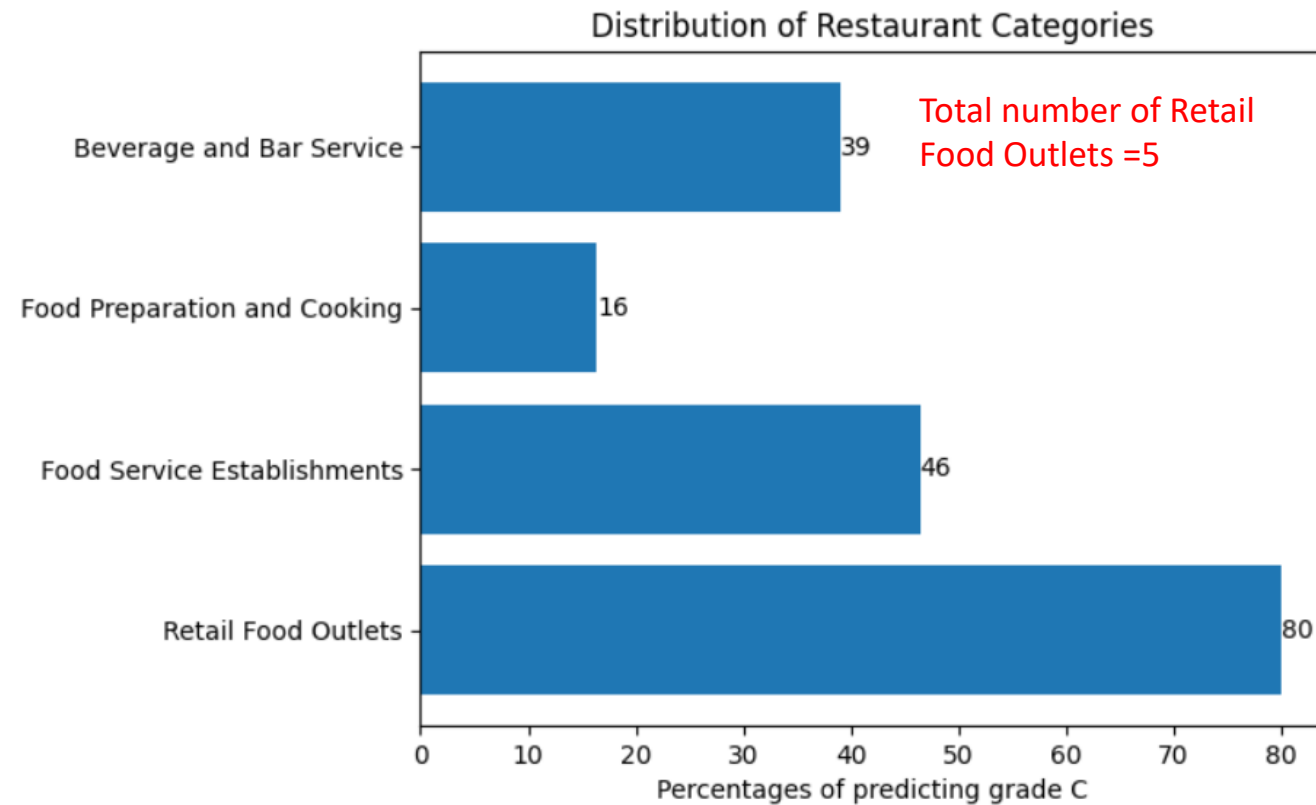
Bar chart offering valuable insights for targeted interventions and concerning in specific restaurant categories.

Distribution of Age Buckets

Bar chart illustrating the percentages of grade C across different age-wise buckets in the dataset.

Proposed Solutions to the Problems

Potential Steps to Avoid Receiving Current Demerits and Attaining a Lower Current Grade

Commercial General Liability (CGL): protects restaurants from liability claims arising from bodily injury or property damage caused by their products, operations, or employees.

Commercial Property Insurance: protects restaurants from financial losses caused by damage to their physical property, such as their building, equipment, and inventory.

Preventive Measures for Bear Beverage Restaurant Category to Avoid Receiving Grade C

Liquor liability coverage: Many restaurants serve alcohol, and they need to be protected from the risk of liquor liability lawsuits. Nationwide could offer liquor liability coverage to help restaurants cover the costs of damages and settlements if they are sued as a result of alcohol-related incidents.

Spoilage coverage: Spoilage is another common risk for restaurants, and it can be expensive to replace spoiled food. Nationwide could offer spoilage coverage to help restaurants cover the cost of replacing spoiled food.

Thank You