Understanding Customer Behaviors in Car Insurance

Yaman Shadid

09/13/2023

Project Overview

- **For who?**
    - My stakeholder is a car insurance company looking to gain insights into customer behaviors.
- **What problem am I solving for them?**
    - I aim to identify patterns in the data to help the company make:
        - Informed decisions
        - Reduce risk
            - The insurance company aims to improve their risk assessment processes.
        - Optimize their services

Data Introduction

- **Brief Intro:**
    - I am working with an annual car insurance dataset containing features, logs, and an outcome column indicating whether a customer claimed their insurance (1) or not (0)
- **Source and acknowledgments:**
    - The data is a combination of real-world data and some generated for analysis

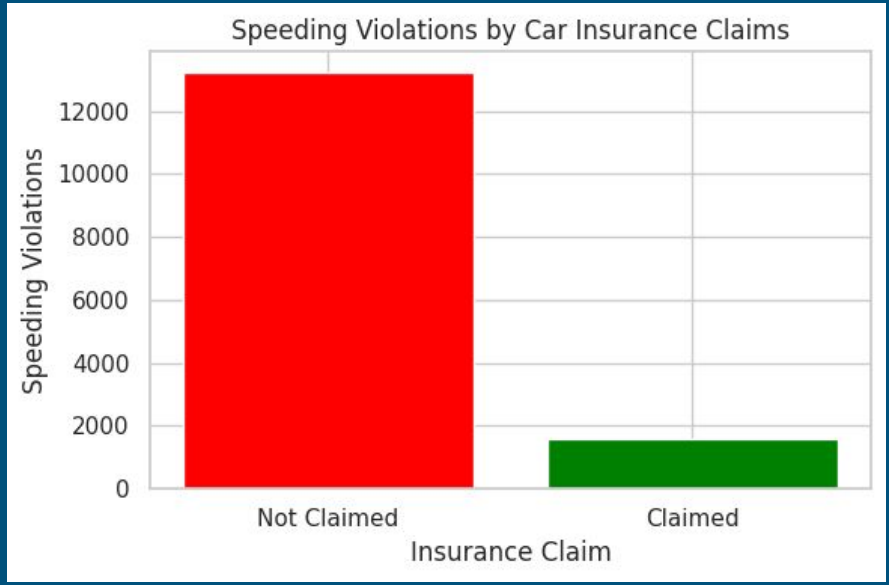
Key Findings - Visual 1

- **Visual 1:**

  Bar Plot (X-axis: Speeding Violations, Y-axis: Proportion of Customers Not Claiming Insurance)

- **Finding 1:** Speeding Violations vs. Insurance Claims

  - I observed that as the number of speeding violations increases, the proportion of customers not claiming their insurance also increases.
  - This suggests a strong correlation between speeding violations and a reduced likelihood of insurance claims.
  - Customers with more speeding violations appear to be more self-reliant when it comes to covering expenses, impacting their claim behavior.



Key Findings - Visual 2

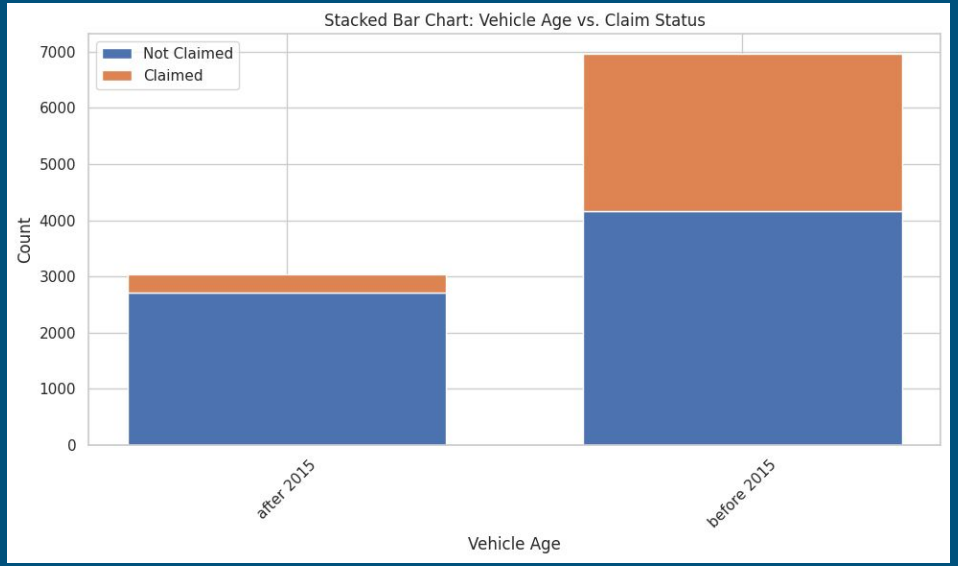
- **Visual 2:**

  Stacked Bar Chart (X-axis: Vehicle Age, Y-axis: Count, Stacked by Claim Status)

- **Finding 2:** Vehicle Age and Claims

  - Customers with newer vehicles are less likely to make claims compared to those with older vehicles.

  - This suggests a potential correlation between vehicle age and claim probability.



Strengths

- **Strengths of My Model**

    - My model can provide valuable insights into customer behaviors, helping the company optimize pricing, marketing, and risk assessment.

    - It handles imbalance data effectively, crucial for insurance datasets with fewer claimants.

Limitations

- **Limitations of My Model**
    - While my model is effective, it may not capture nuanced customer behaviors that could be essential for risk assessment.
    - The model’s performance might be influenced by the quality of the data, including the generated part.

Classification Consideration

- **Classification Problem: False Positives and False Negatives**
    - **False Positives:** Predicting a claim when there isn't one may lead to increased operational costs.
    - **False Negatives:** Failing to predict a claim when there is one may result in dissatisfied customers

Class Balance

- **Class Balance in Evaluation**
  - I considered the imbalance nature of the dataset when evaluating my classification model
  - I used precision, recall, and F-1 score to assess model performance, ensuring a balanced approach.

Final Recommendations

- Final Recommendations
  - Utilize the insights gained from the model to refine pricing strategies, tailor marketing efforts, and optimize risk assessment.
  - Continuously monitor and improve data quality to enhance the model’s accuracy.
  - Consider additional features or external data sources to further refine the model’s predictions.

Thank you

Understanding Customer Behaviors
in Car Insurance

By:

Yaman Shadid