

---

# Health Insurance Cross-Sell

Predicting vehicle insurance cross-sells to current  
health insurance customers.

# TABLE OF CONTENTS

---

01

Business and Data  
Understanding

02

Predictive  
Modeling

03

Results

04

Recommendations  
and Next Steps

---

01

# Business and Data Understanding

# Business Understanding

---

With predictive modeling, we can:

- **Improve efficiency** Boost the percentage of leads that turn into conversions, without additional sales team output.
  - **Improve marketing** Which kinds of customers are we failing to attract? What are they concerned about?
-

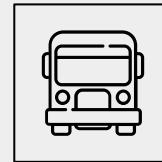
# Data Understanding

---



## Demographics

Gender  
Age  
Driver's License



## Vehicle

Previously Insured  
Vehicle Age  
Vehicle Damage



## Health Insurance Policy

Annual Premium  
Policy Sales Channel  
Vintage

**~ 380,000 customers**

---

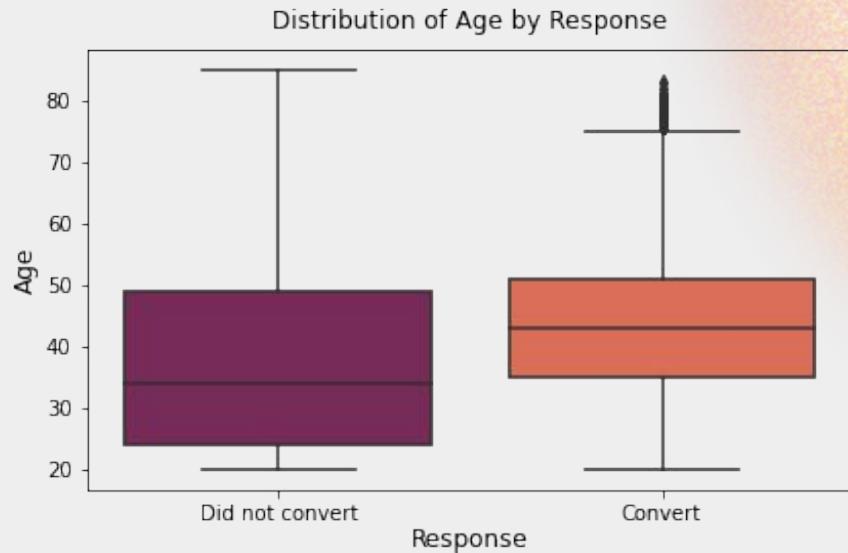
# Data Understanding

---

## Customer Age

**50%** of customers who don't convert  
are *younger* than 35.

**75%** of customers who do convert are  
*older* than 35.

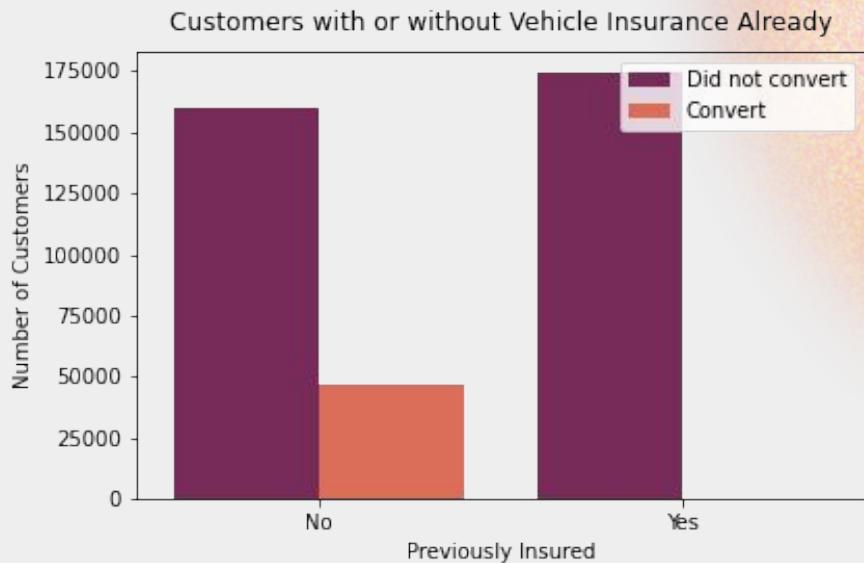


# Data Understanding

---

## Previously Insured

Only 150 of nearly 175,000 customers who already had vehicle insurance responded positively to the cross-sell.

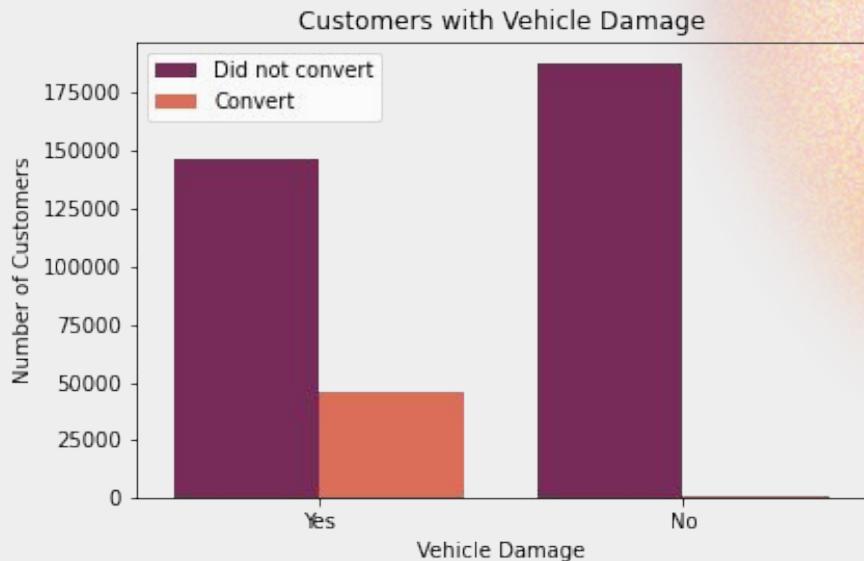


# Data Understanding

---

## Vehicle Damage

24% of customers with vehicle damage converted, whereas *less than 1%* of customers without vehicle damage converted.



02

# Predictive Modeling

# Baseline Model

---

Most features included.

Logistic regression; predicts the probability of a class by assuming log-odds are linear.

- Initial model relied on vehicle damage and previously insured.
  - Weighted the classes to encourage minority class detection.
  - Increased threshold to increase conversion rate.
-

# Other Models

---

K Neighbors Classifier - find the customers in the training data who are most similar, and predict that same class.

AdaBoost Classifier - ensemble method, builds lots of models and encourages models to learn from the mistakes of previous models.

---



# 03 Results

# Final Model Metrics

---

29%

PRECISION

Sales team's conversion  
rate

85%

RECALL

Percent of interested  
customers converted

43%

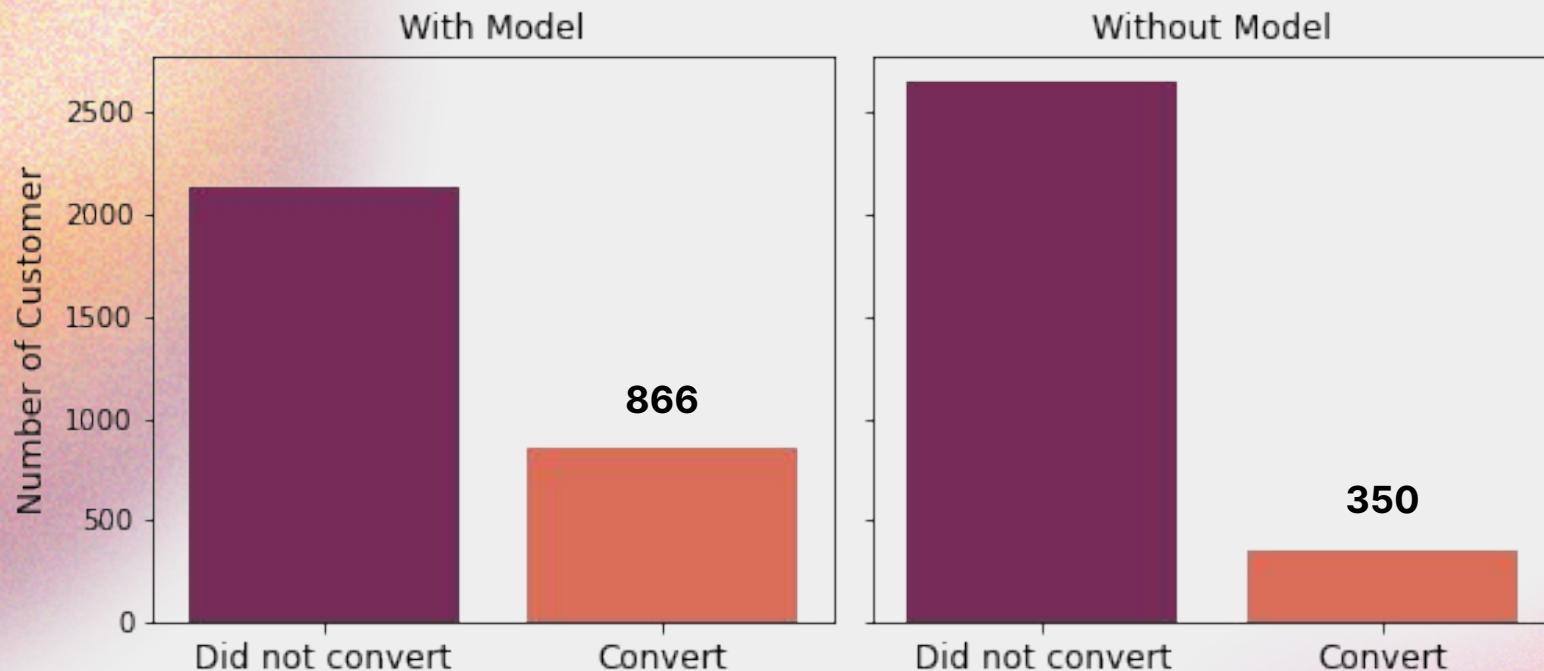
F1 SCORE

Balance between  
precision and recall



\*Assuming 10 person- sales team contacts 10 customers / day each.

## Contacts vs. Conversions over One Month Period



---



04

# Recommendations and Next Steps

---

# Recommendations

---

## Implement

- Deploy logistic regression model
- Use the top 2 sales channels (over half of all cross-sells come from these two.)

## Next Steps

Do some research for marketing & product development on:

- Customer segments
- Regions

Build a dashboard on conversion rates & demographics by region, by customer segment

---



# THANKS!

## ANY QUESTIONS?

[rebecca.clark@gmail.com](mailto:rebecca.clark@gmail.com)

[github.com/ran0027](https://github.com/ran0027)

[linkedin.com/rebecca-neel-ds](https://www.linkedin.com/rebecca-neel-ds)

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**