



Vehicle Insurance Customer Targeting

MSBA 7027 Group 17

Agenda

01

Introduction

02

Data Description

03

ML Methods

04

Conclusion



01

Introduction



Morning

Go through **emails** from potential and existing **customers**

After lunch

Make **cold call** to potential **customers**

Evening

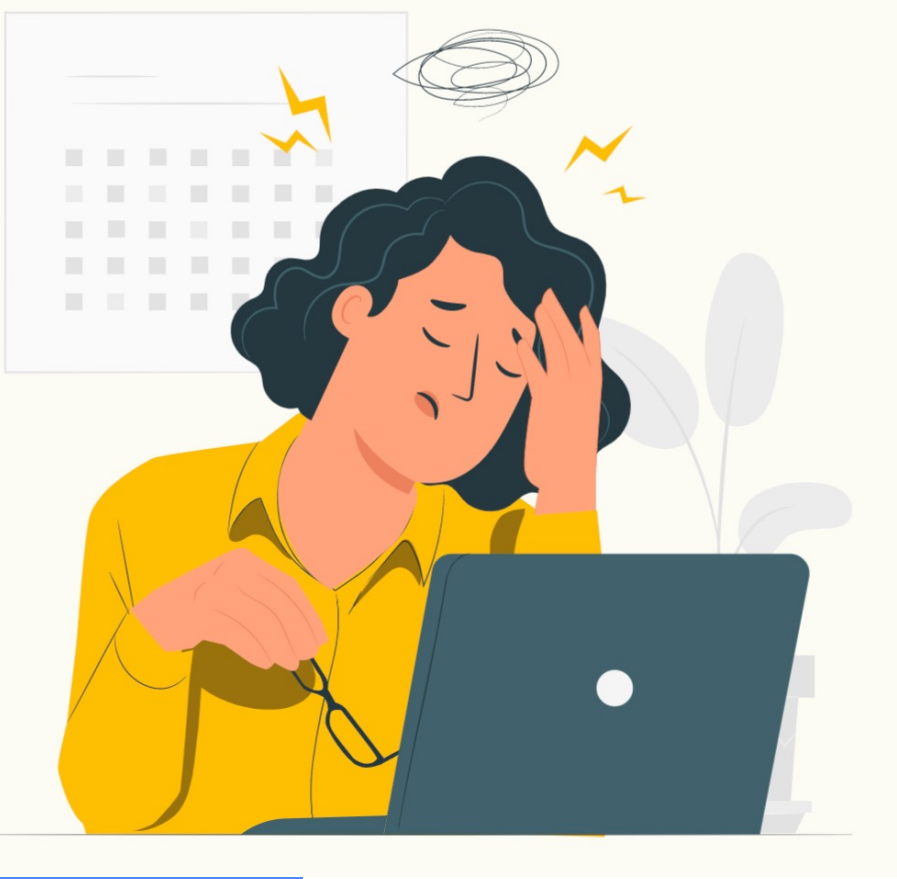
Meet with **customers** face-to-face to evaluate thier needs

Mid-Morning



Meet with **customers** to determine coverage

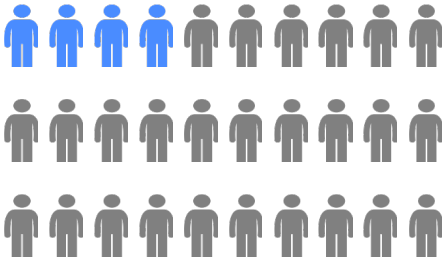
Mid-Afternoon

Work on various marketing options to secure **new leads**



Background

-  Insurance sales agents spend **more than 45 hours** a week on meeting clients.
-  However, only **14.3%** customers will buy insurance.



Objective



 Build a model to **predict whether** the clients is interested in extending their vehicle insurance **based on thier behaviour**.

 Accuracy: 99%




Benefit

WHO

-  Reach out those **potential clients** effectively
-  Increase the **conversion rate** of marketing campaigns.

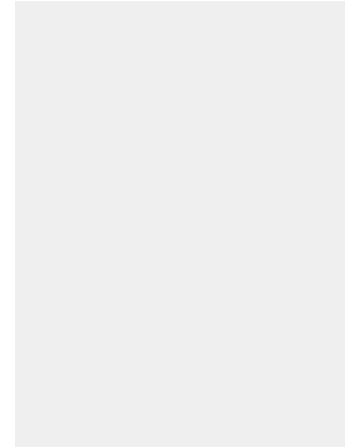
WHICH

-  Predict **which product** is going to be bought by each client to make more **efficient** campaigns.

02

Data

Data Description
Preliminary Data Analysis



Data Description

Response Variable

Response: “Yes” - the customers would like to renew their insurance
“No” - the customers would discontinue their insurance

Explanatory Variable

Demographic

1. Customer – Customer unique ID
2. Marital Status
3. Education
4. State - where customers live in
5. Gender
6. Location Code

Financial

7. Income
8. Employment Status
9. Vehicle Class
10. Vehicle Size
11. Monthly Premium Auto - Premium auto that customers need to pay

Consumption Records

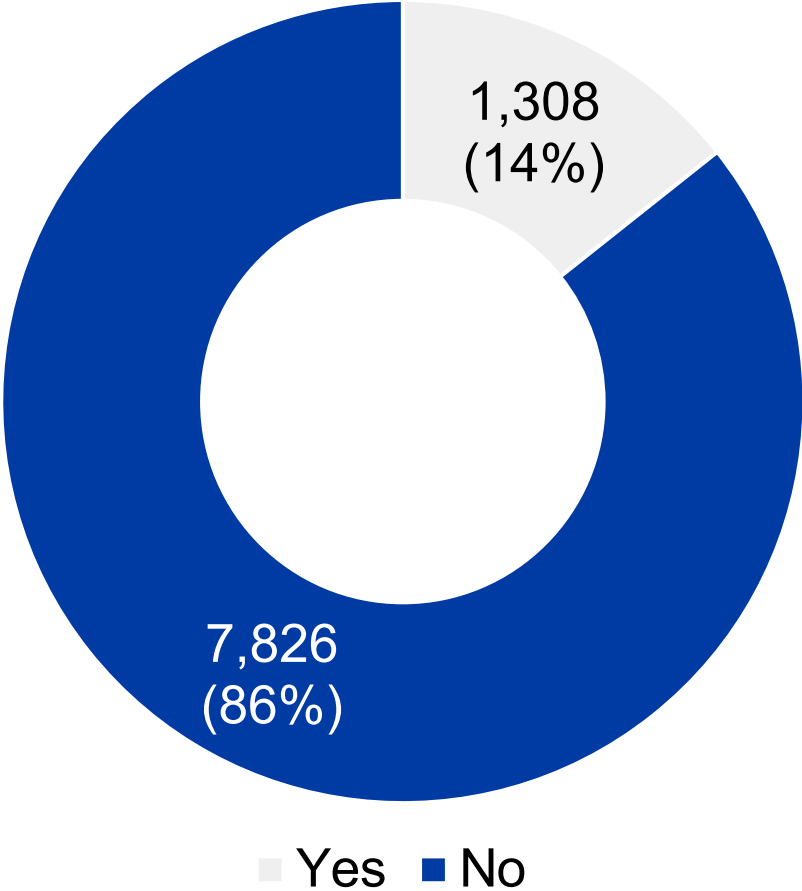
12. Effective To Month - The first month when customer would like to activate their car insurance
13. Months Since Last Claim
14. Months Since Policy Inception
15. Number of Open Complaints
16. Total Claim Amount

Renew Sales

17. Customer Lifetime Value – Equals to Customer Value * Customer lifespan
18. Number of Policies
19. Policy Type
20. Policy
21. Renew Offer Type
22. Sales Channel
23. Coverage

Preliminary Data Analysis

Y: Number of Response



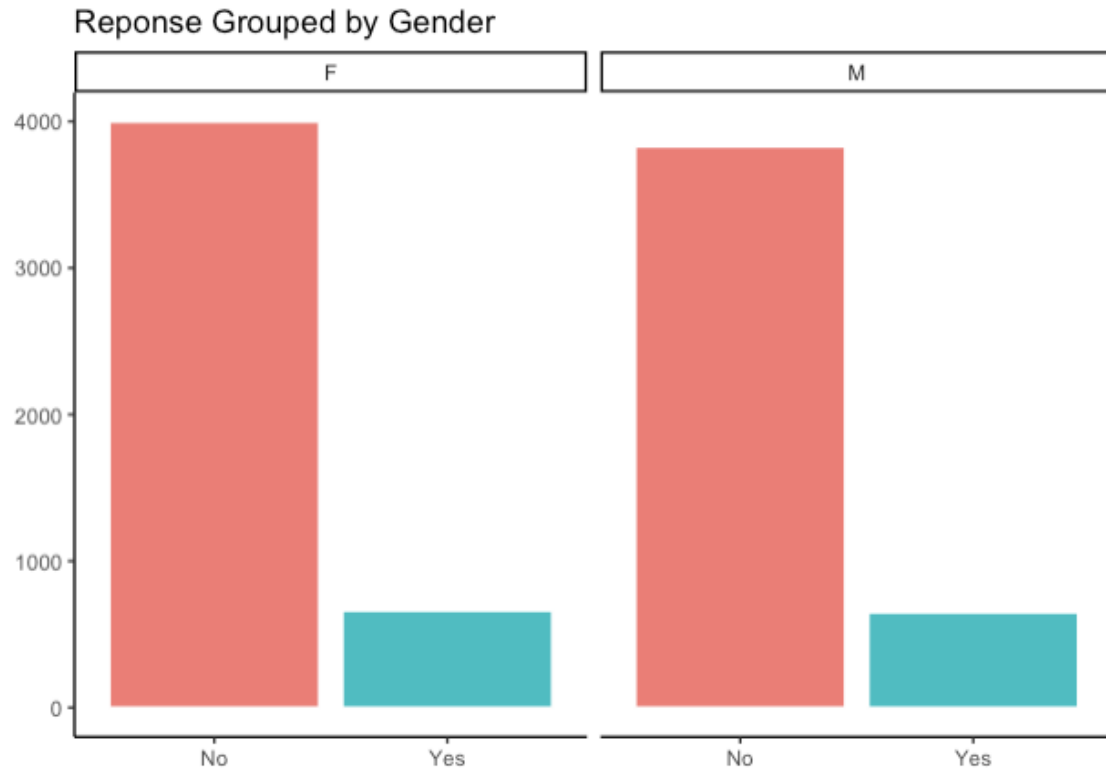
Much more customers
answering "No"

X: Numeric Features

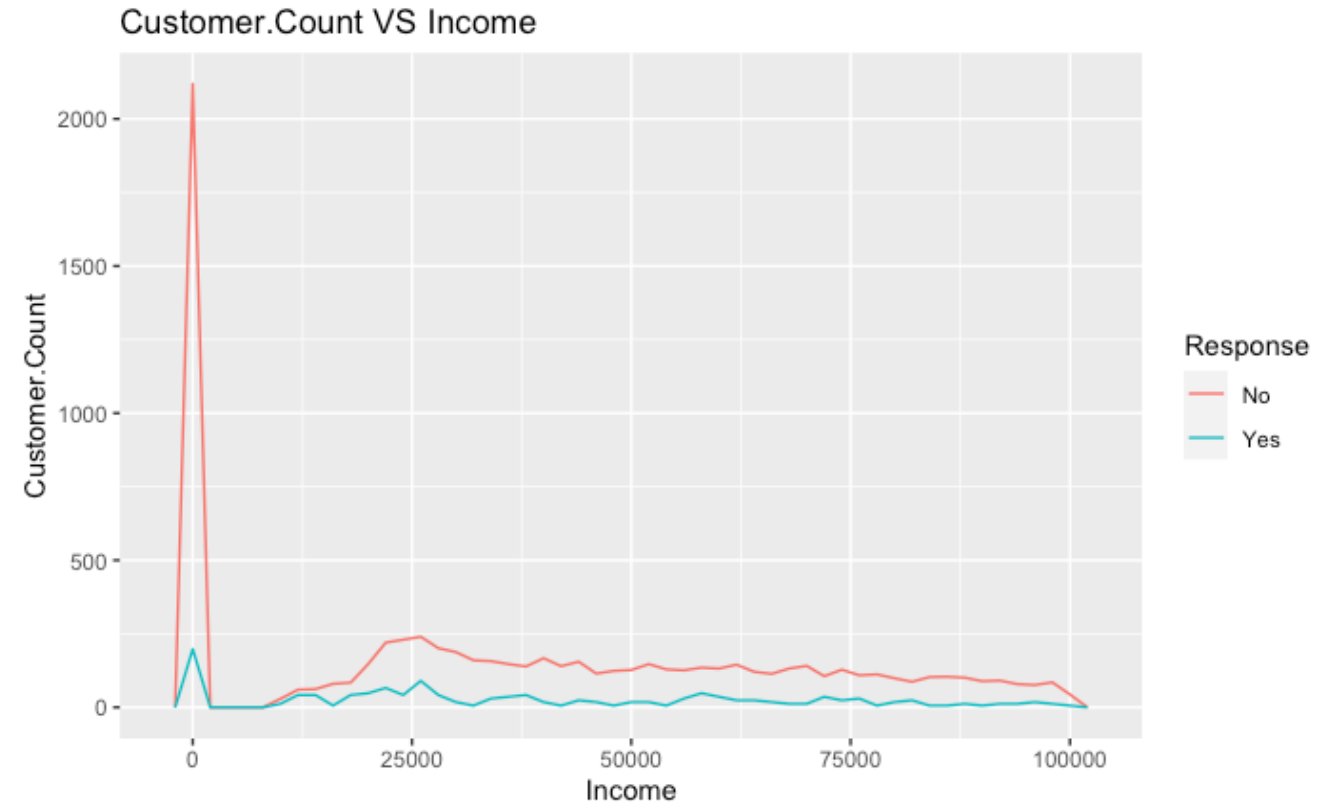
	mean	sd	max	min
Customer.Lifetime.Value	8004.94	6870.97	3658.9	1898.01
Income	37657.4	30379.9	42522.5	0
Monthly.Premium.Auto	93.22	34.41	26.69	61
Months.Since.Last.Claim	15.1	10.07	11.86	0
Months.Since.Policy.Inception	48.06	27.91	35.58	0
Number.of.Open.Complaints	0.38	0.91	0	0
Number.of.Policies	2.97	2.39	1.48	1
Total.Claim.Amount	434.09	290.5	213.58	0.1

"Customer.Lifetime.Value" and
"Income" have large ranges

Preliminary Data Analysis



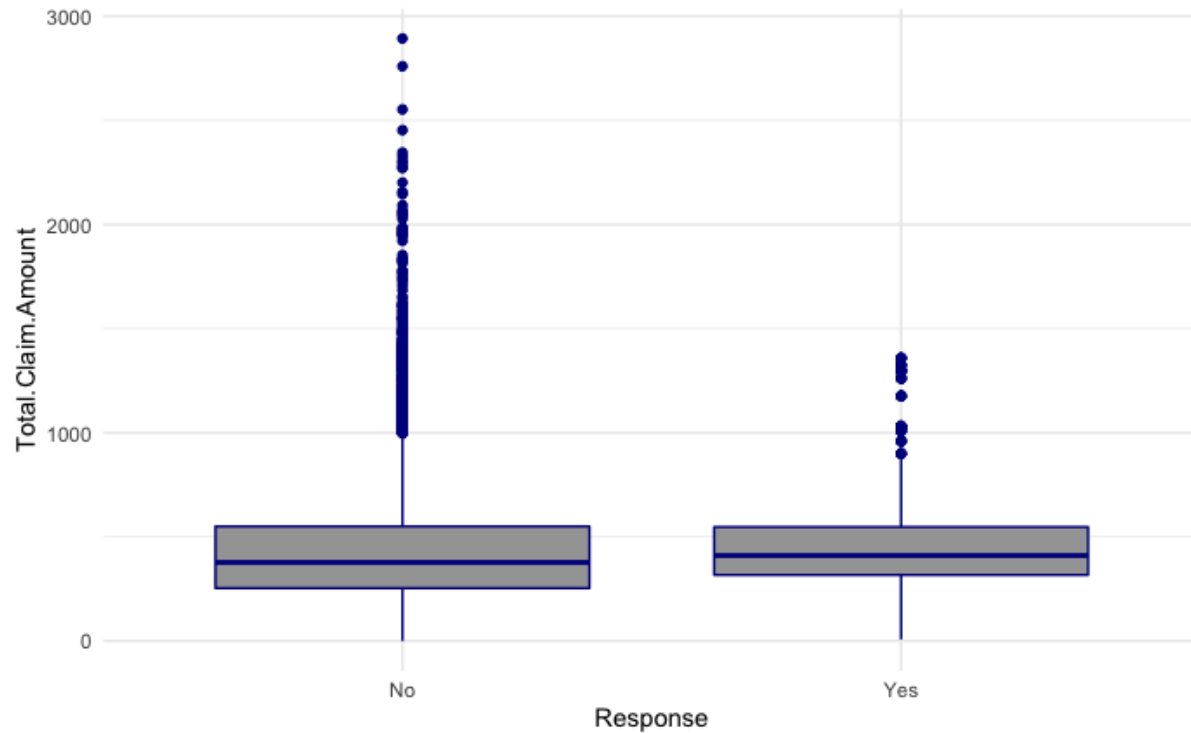
No obvious difference
between genders



For "No Response" customers, there is a
higher proportion of zero income customers

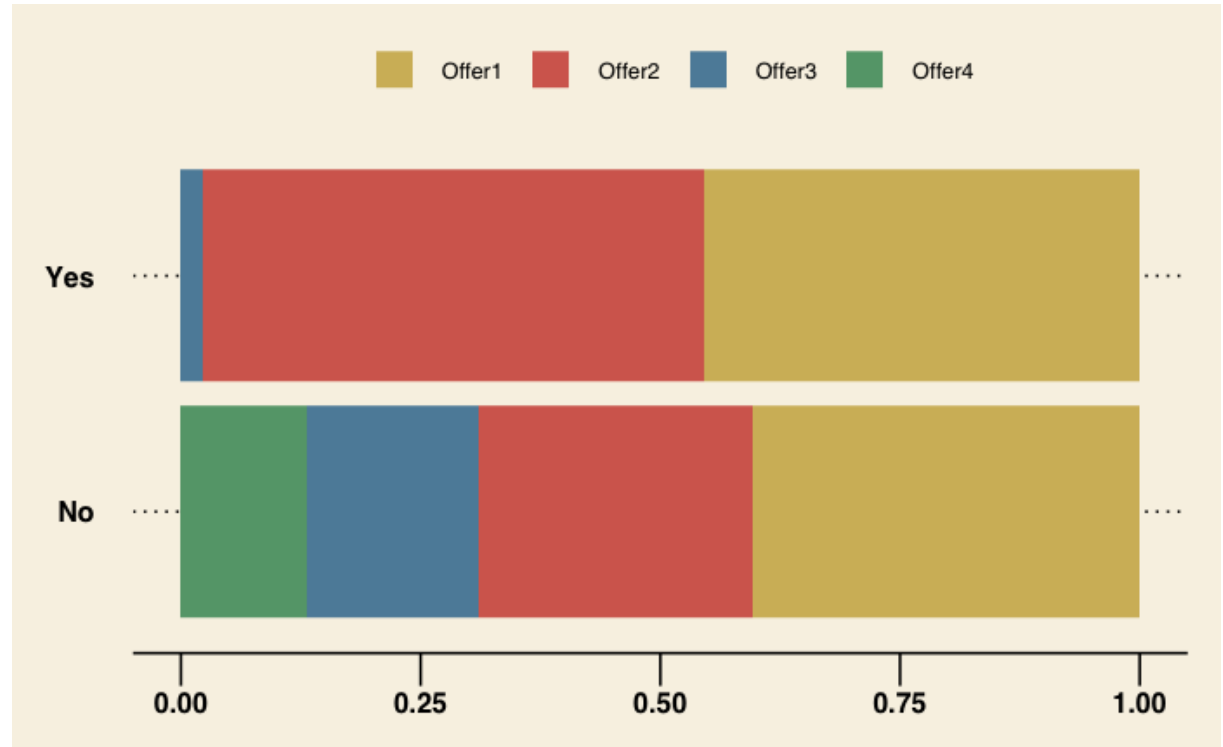
Preliminary Data Analysis

Total.Claim.Amount VS Response



Customers who have high Total Claim Amount tended to not to renew their vehicle insurance

Response Grouped by Offer Type



Offer2 is more attractive to existing customers, while Offer1 is somehow undesirable

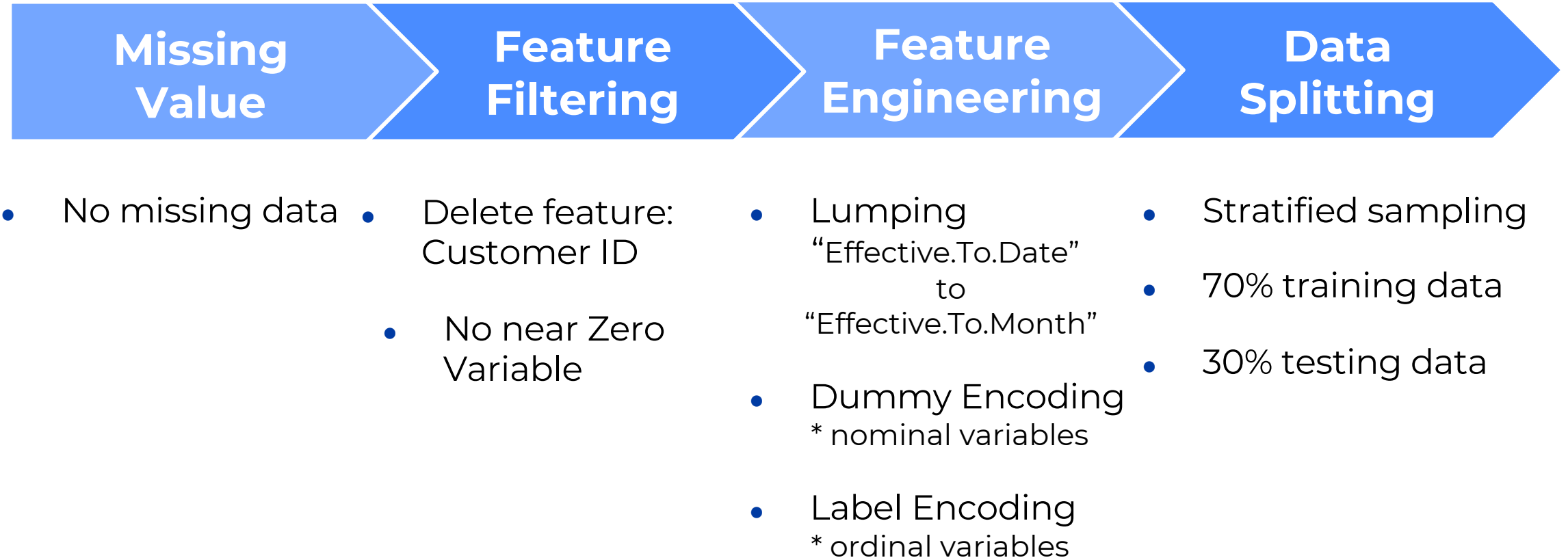


03

ML methods

Data Preprocessing
Machine Learning Algorithms

Data Preprocessing



Data Preprocessing



- Stratified sampling
 - 70% training data
 - 30% testing data
- Numerical variables
- `Step_center()`
 - `Step_scale()`
 - `Step_normalize()`

ML methods & comparison

	Random Forest	Basic GBM	SVM - radial	KNN	SVM - linear
accuracy	0.9949	0.9898	0.9376	0.9143	0.8723
run time	short	long	long	short	medium
	LDA	Decision Tree	Logistic Regression	SVM - polynomial	
accuracy	0.8713	0.8709	0.8709	0.8639	
run time	short	short	short	medium	

ML methods & comparison

	4 most important feature			
Random Forest	Income	Renew.Offer.Type	Total.Claim. Amount	Customer. Lifetime.Value
Basic GBM	Customer. Lifetime.Value	Income	Effective.To. Month	Employment Status
SVM - radial	Effective.To. Month	Employment Status	Renew.Offer.Type	Education

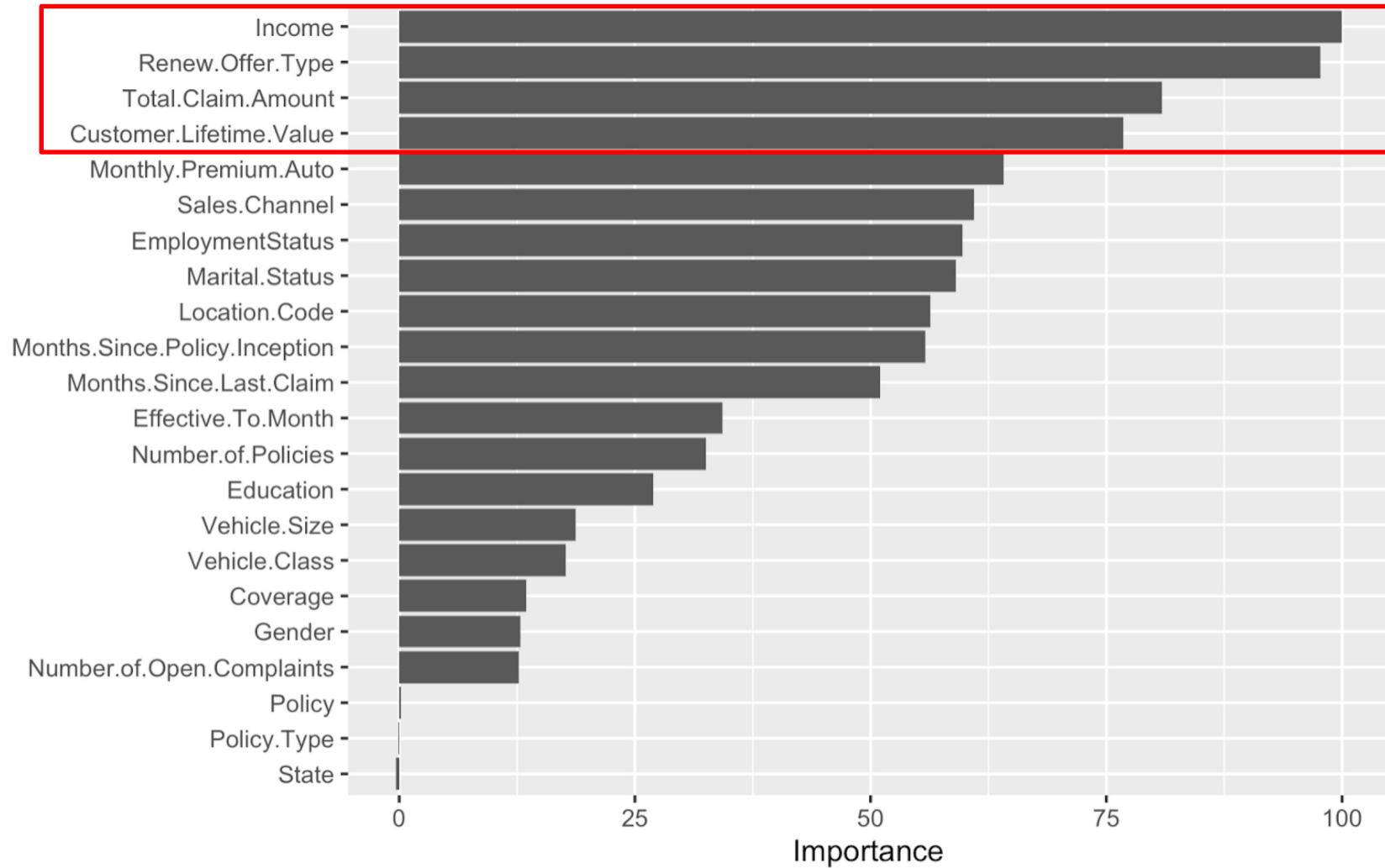


04

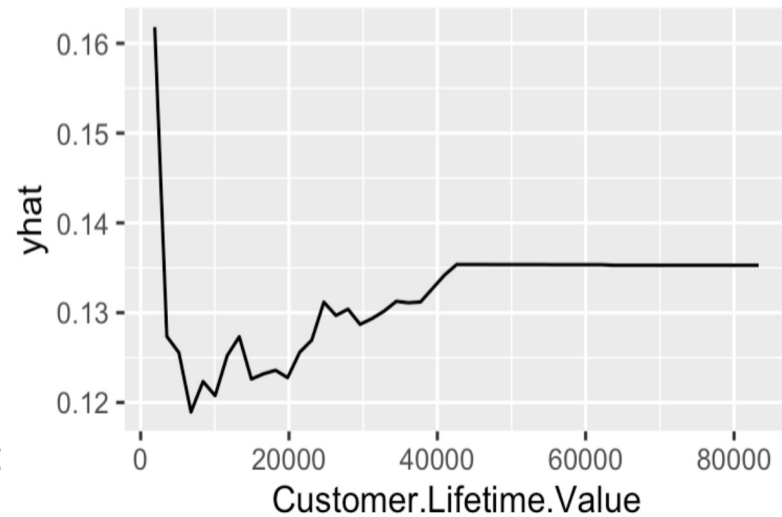
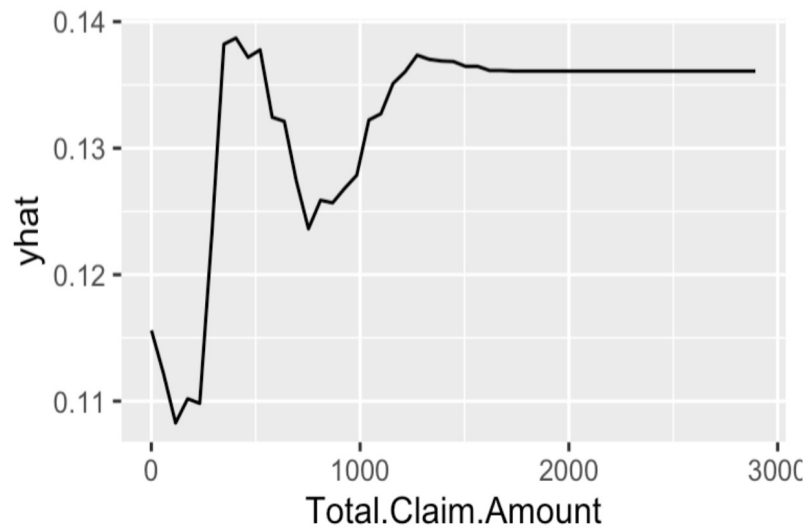
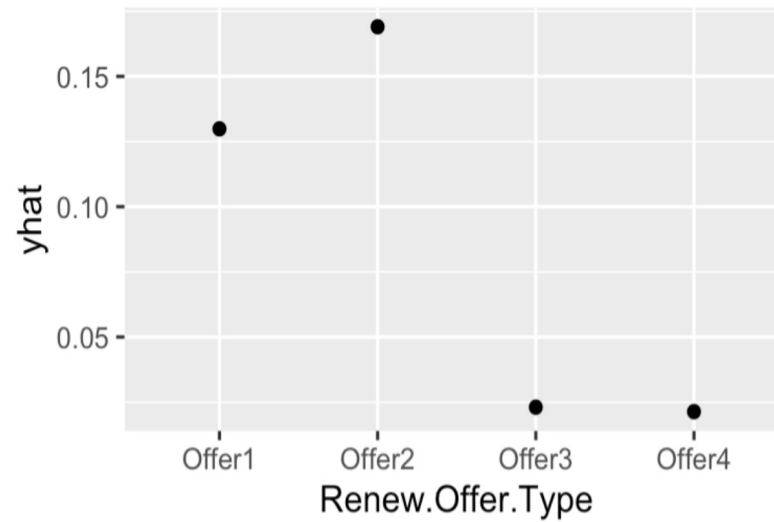
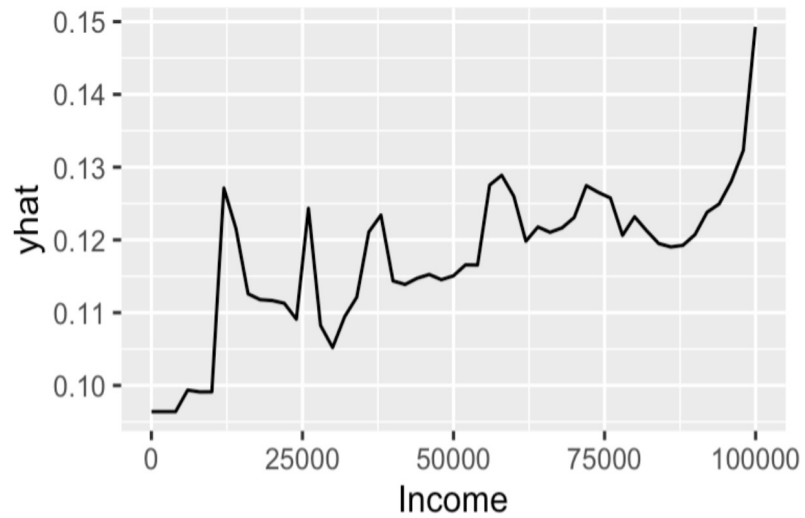
Conclusion

Most important features
Partial dependence plots

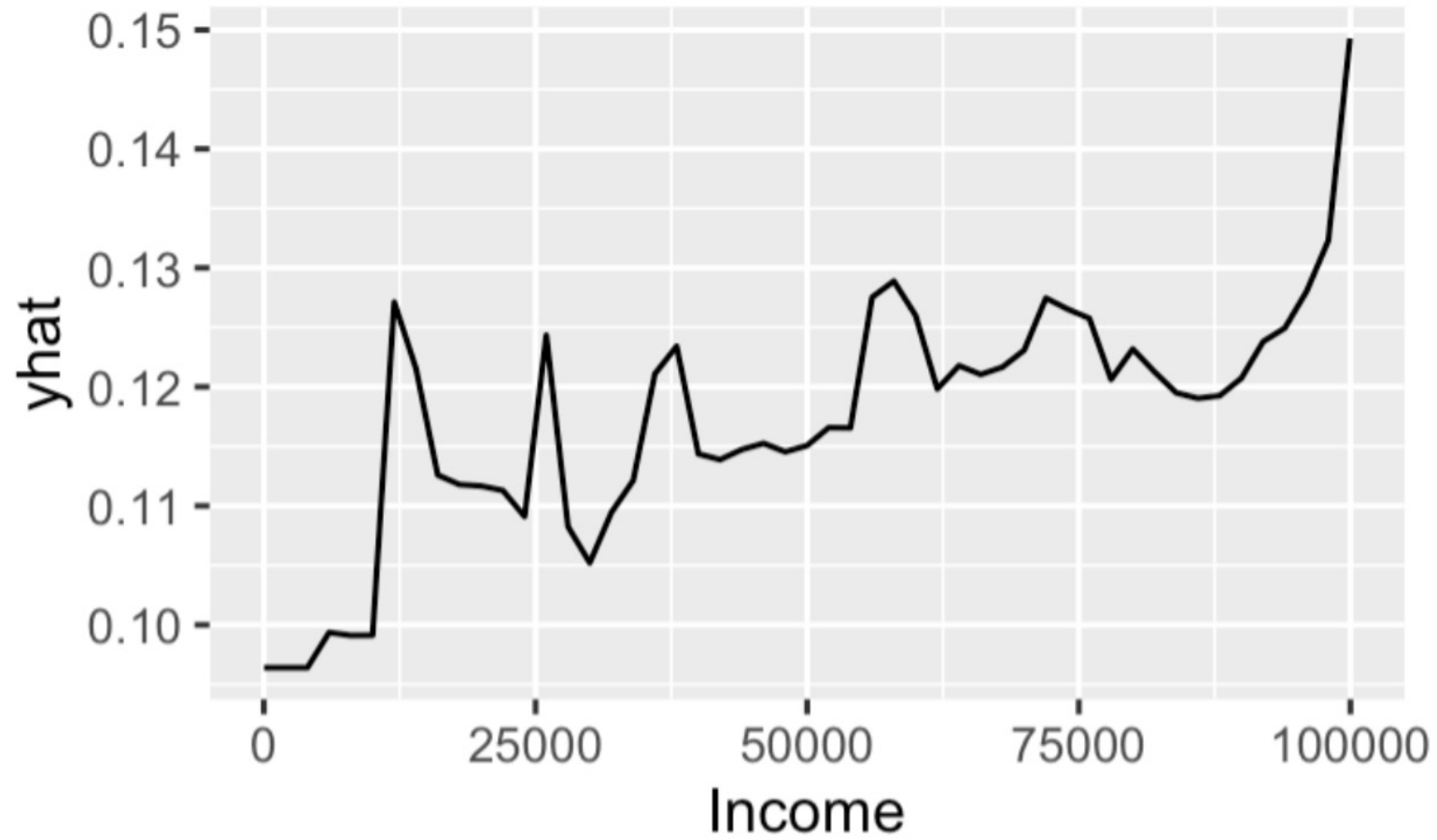
Most important features



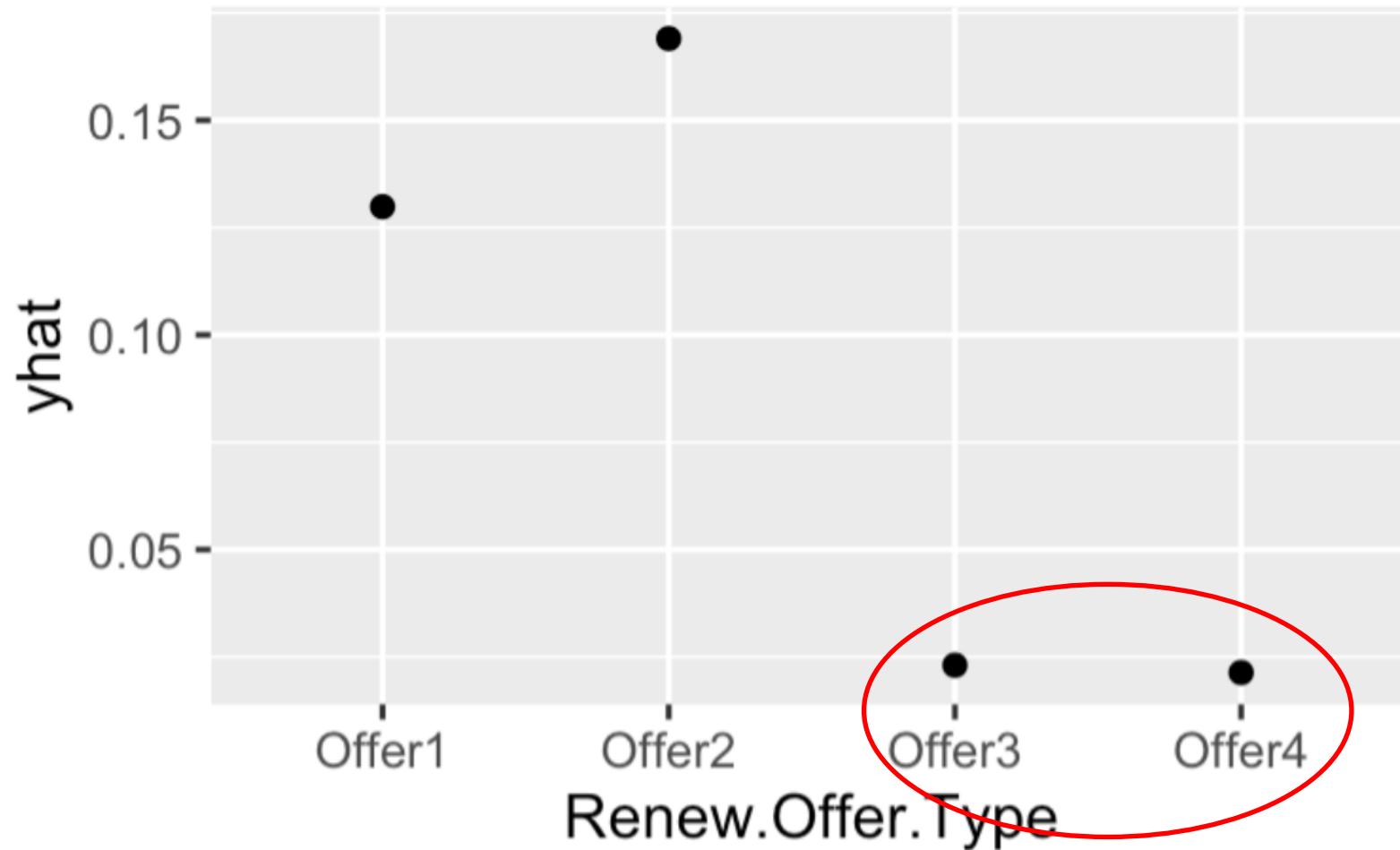
Partial dependence plots



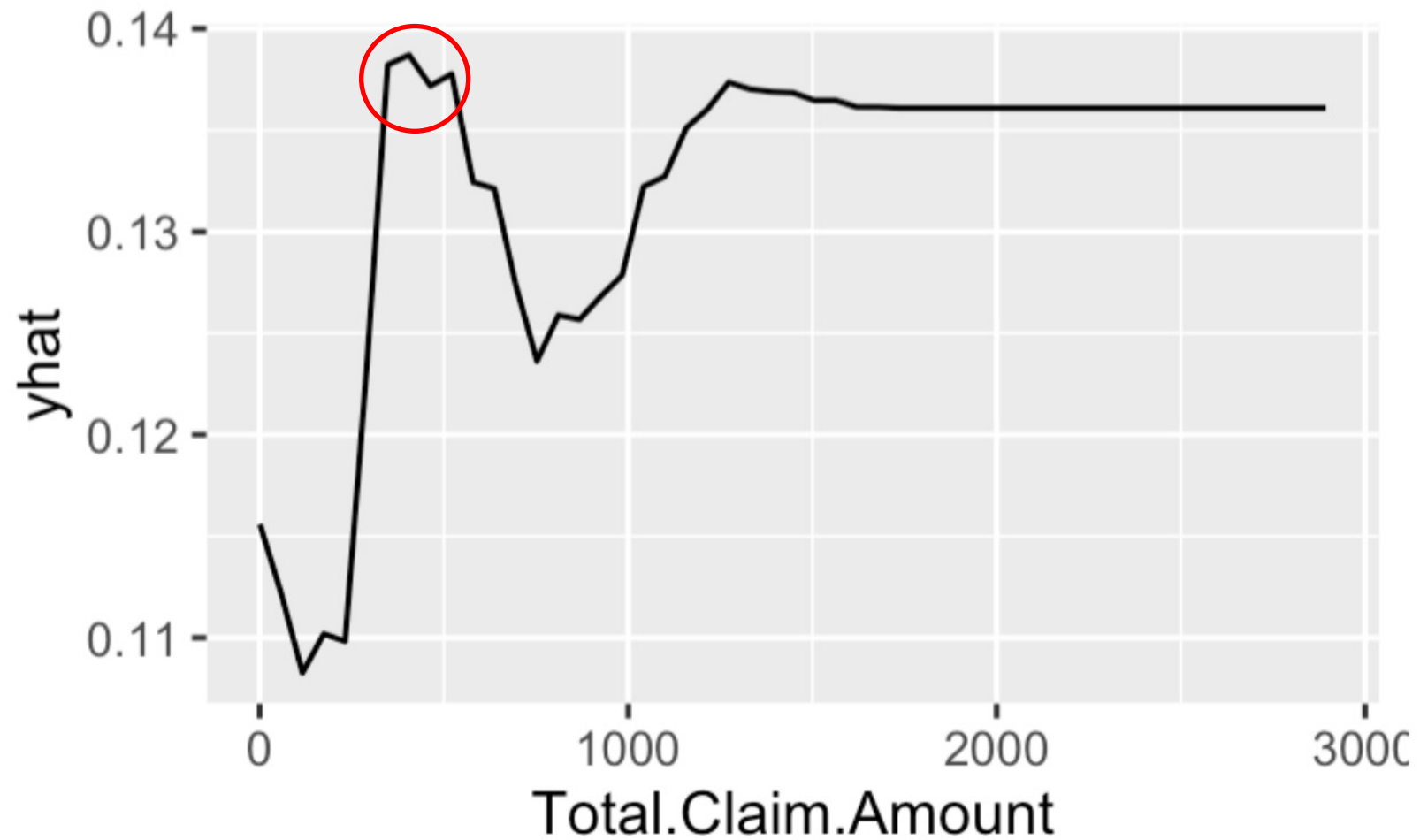
Partial dependence plots



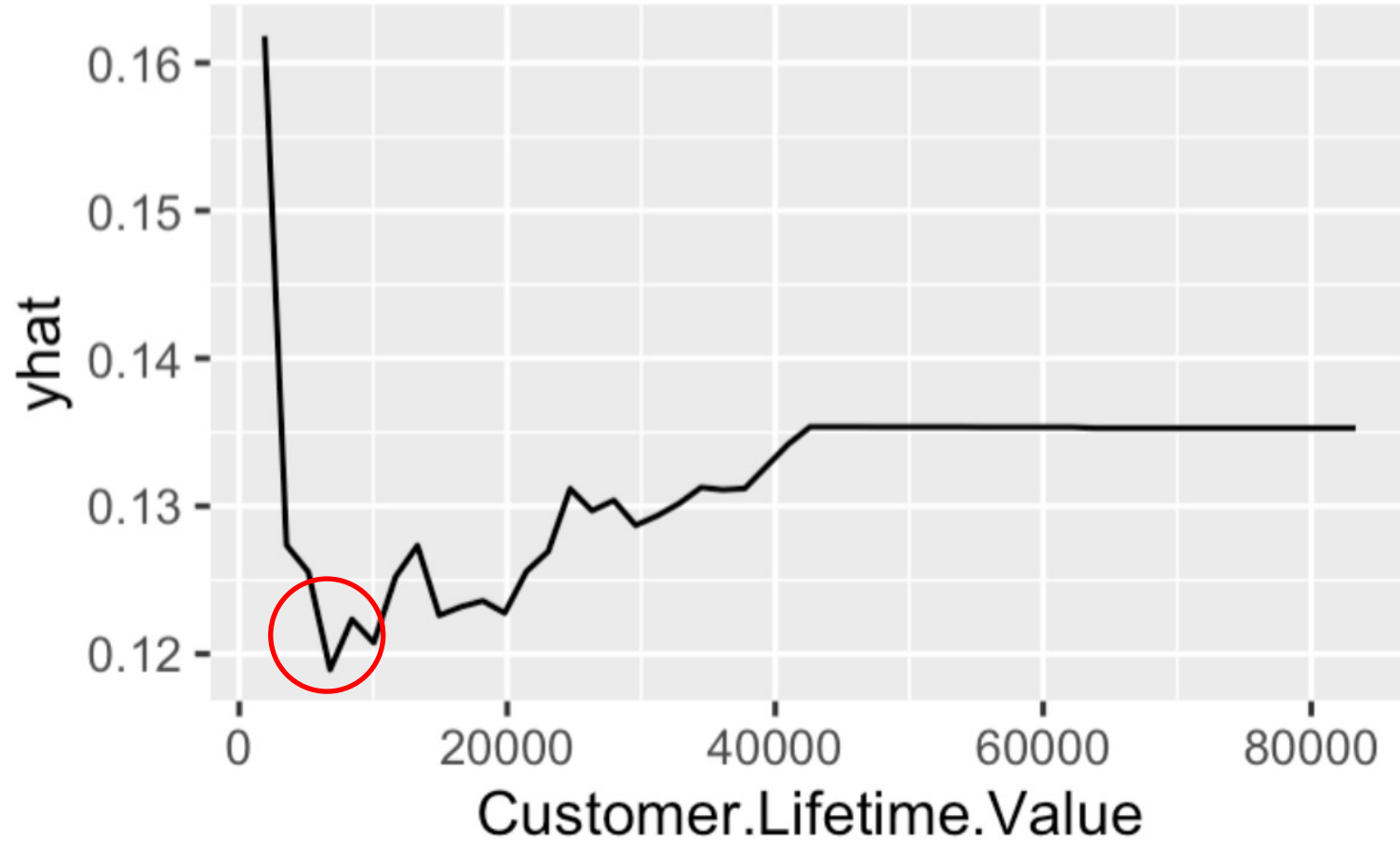
Partial dependence plots



Partial dependence plots



Partial dependence plots



Thanks

Q & A