Customer Marketing

Kate Larsen

February 10, 2017

Agenda

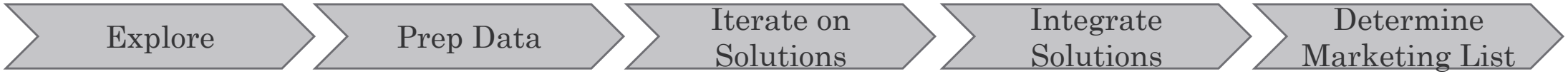
- Problem Overview
- Methodology
- Solution
- Results
- Next Steps

Problem Statement

- Marketing customers is an expensive sunk cost ($30).
- Maximum profit is achieved by targeting customers that will respond, purchase a policy, and generate profit for the company.
- Profit = Profit from Customers – 30 * (# customers marketed)
- Key Questions
  - Which customers are likely to respond?
  - How much profit is a customer likely to generate?
  - What are key features of high profitable customers?
  - How many customers should be targeted?

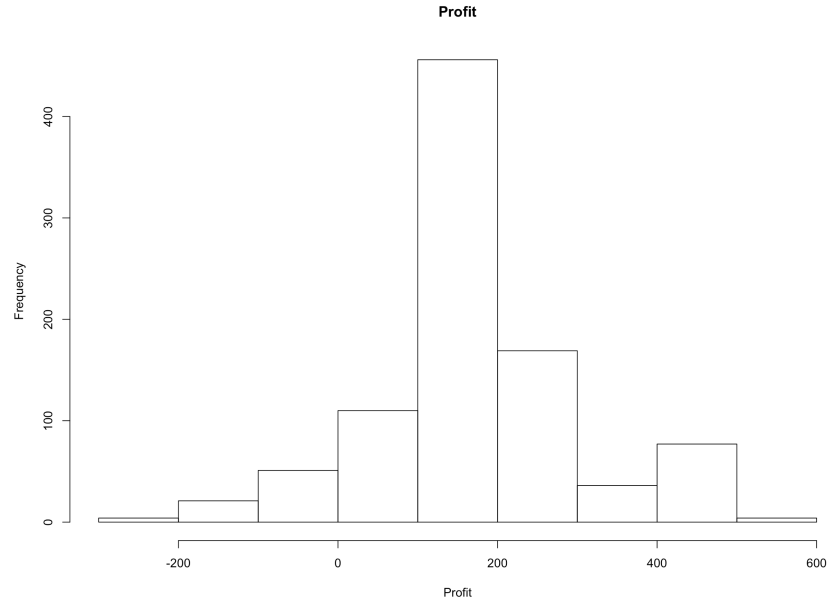
Methodology

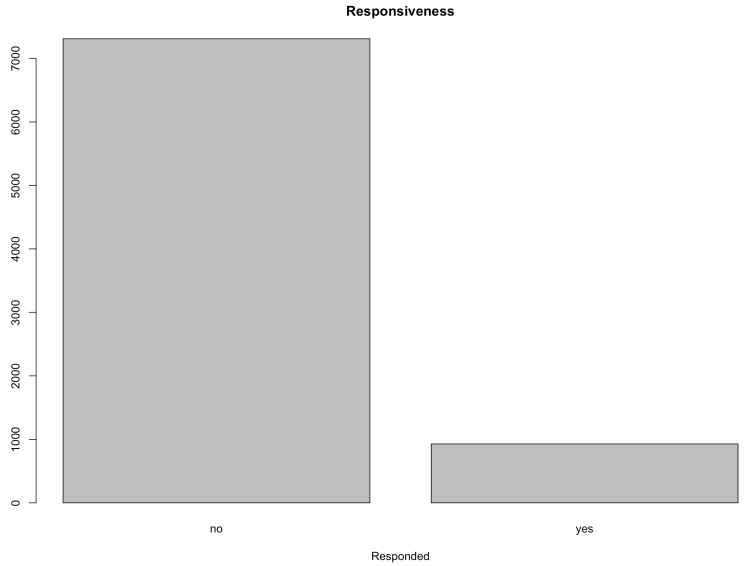
Phase	Purpose	Key Insight
1) Exploration	• Understand business problem • Understand data attributes • Explore relationships within data	*Profitability from a policy must be greater than $30, thus predicting responsiveness is important.*
2) Data Preparation	• Resolve missing data • Combine like attributes	*Multiple features had missing values or skewed data.*
3) Iterate Models	• Test solution accuracy	*Simple models often yielded better results.*
4) Develop Final Solution	• Implement best found approaches	*Factors such as profession and market indices showed significance.*
5) Deliver Marketing List	• Integrate solutions to determine contact list	*Find number predicted to deliver positive total profit.*



Initial Analysis

Data	Solution
Missing Values	Replace with common values in historical data set
Relationships between characteristics	Combine characteristics together
Ranges in data	Perform transformations to alleviate skew





Explore → Prep Data → Iterate on Solutions → Integrate Solutions → Determine Marketing List

Solution Overview

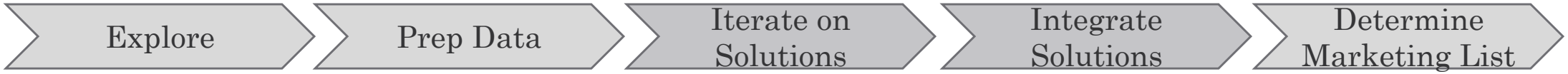
A two part solution combined the likelihood of a customer responding with the predicted profitability of that customer. These were combined to yield the total profitability to measure all customers.

	A) Responsiveness	B) Profitability
Measures	Will a customer respond?	How much will that customer's policy be profitable?
Outcome	yes / no	$ amount predicted
Implementation	Binary Logistic Model	Multiple Regression Model
Example:	70% chance of responding	$170 dollars

Solution Analysis

The following measures were significant in identifying responsiveness and profitability.

	A) Responsiveness	B) Profitability
Measures	Will a customer respond?	How much will that customer's policy be profitable?
Outcome	Yes or No	Profitability ($)
Key Factors	- Number of employees - Previous outcome - Last month contacted - Consumer Price Index - Consumer Confidence Index - Profession	- Loan - Profession



Marketing Comparison

Historical

Profitability	Customers
No Profit (NA)	7,310
Negative Profit (<$0)	75
Negative Profit ($0-$30)	41
Positive Profit (> $30)	810
	8,238

Future

Profitability	Customers
No Profit (NA)	30,888
Negative Profit (<$0)	13
Negative Profit ($0-$30)	102
Positive Profit (> $30)	1,947
	32,950

Proposed Next Steps

- Market the customers identified as high potential
- Incorporate additional data points to enable further insight
  - Number of policies purchased
  - Type of policies purchased
  - Number of children

Appendix

Attribute Deep Dive

Data Columns

Column	Details	Proposed Solution
custAge	Many just under 40, missing values	Set as average
profession	Only a few unknown	Set as mode
marital	Only a few unknown	Set as mode
schooling	2400 unknown	Set as mode
default	1600 unknown Only 1 as 'yes'	Set as mode
loan	Only 184 unknown	Set as mode (no)
contact	Good	Keep as is
month	Varied	Bucket for season
day_of_week	787 unknown	Set as mode
campaign	Heavy skew	Log
pday	Of those actually contacted (316), heavy right skew	Bucket or categorize with previous columns

Data Columns

Column	Details	Proposed solution
previous	Heavy right skew, 1178 with actual values	Bucket
poutcome	Varying histogram	Combine with previous columns
Emp.var.rate	Varying histogram	Watch for correlation
Cons.price.idx	Varying histogram	Watch for correlation
Cons.conf.idx	Varying histogram	Watch for correlation
euribor3m	Heavy on both ends of histogram	Bucket low / high
Nr.employed	Watch for correlations, left skew	log
pmonths	Right skew for those 316 records	Bucket
pastEmail	1100 emails sent, right skew on those sent	Make binary of yes/no