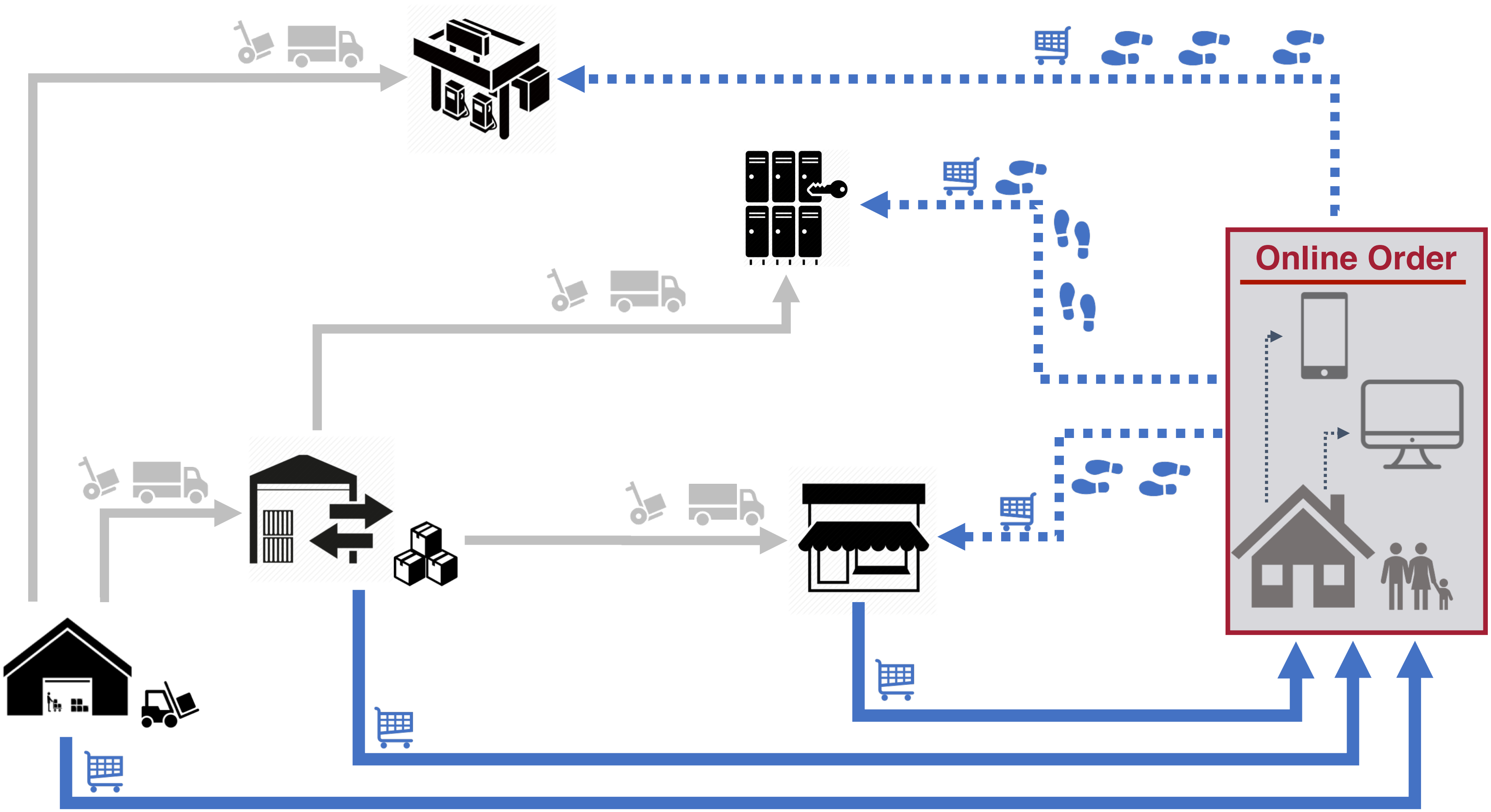
Online Grocery & Omnichannel Strategy: Predicting Home Delivery Adoption

Ryan Alexander Alberts, M.A.Sci. Candidate

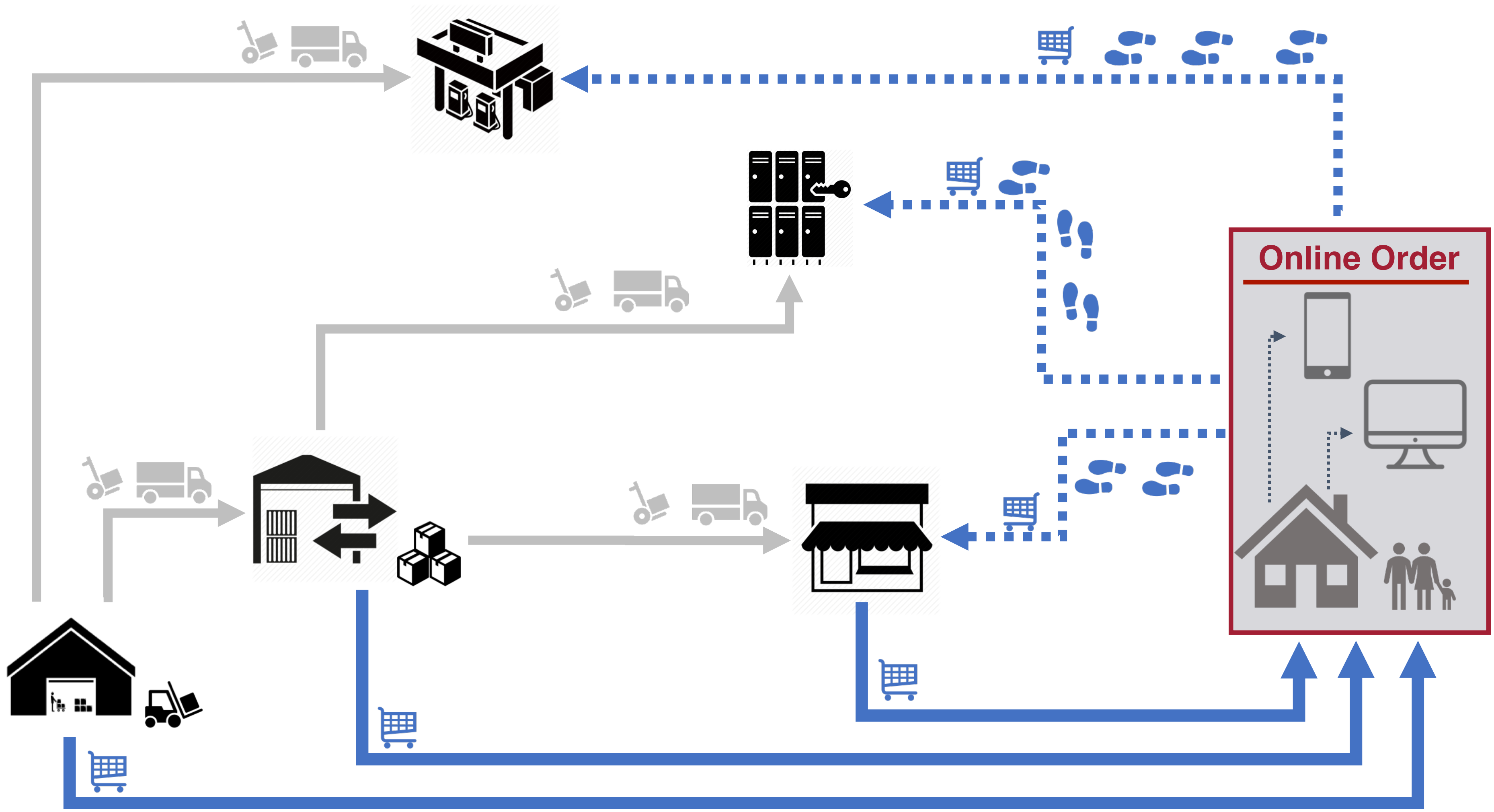
Antoine Lahad Abinader, M.A.Sci. Candidate

Dr. Eva Ponce, Ph.D., Advisor

Online Grocery: The Omnichannel Revolution



Online Grocery: The Omnichannel Revolution



Online Grocery: The Omnichannel Revolution

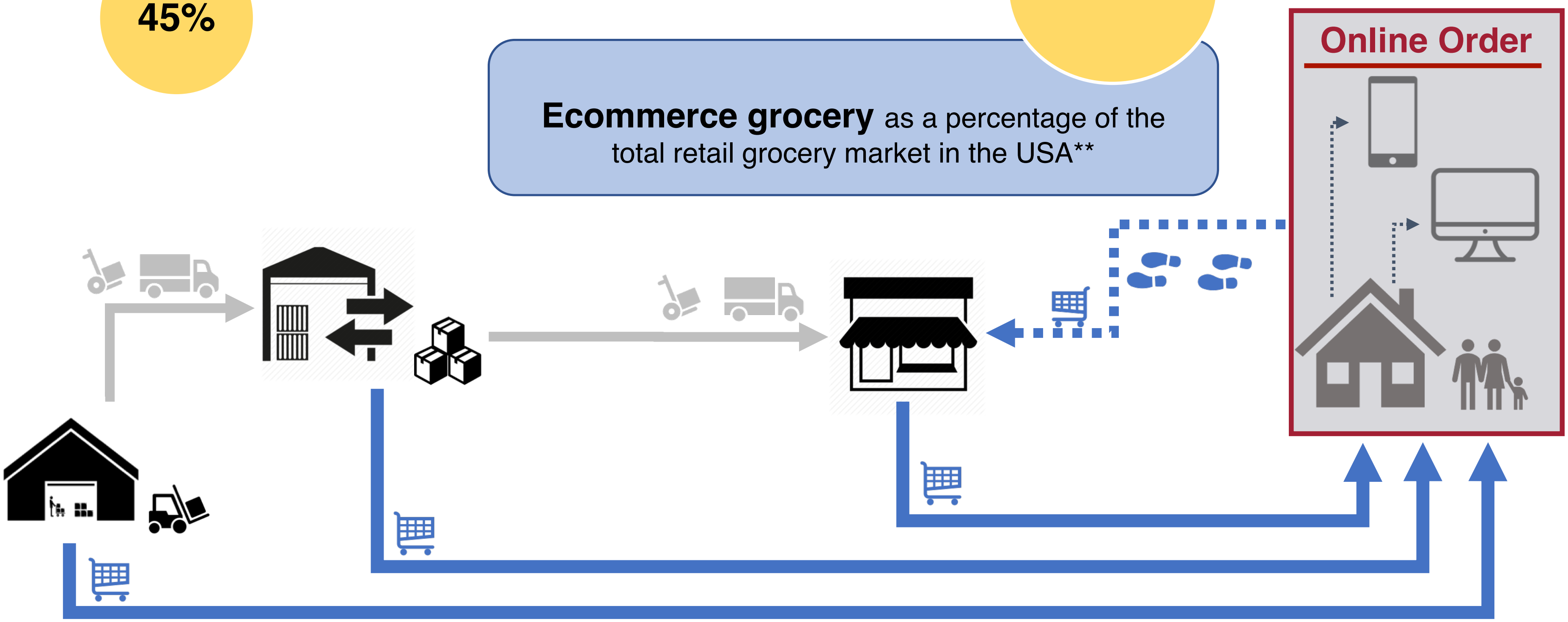
"Customers that shop with us in multiple channels **spend three to four times more** than a customer that shops in one channel."*

Online orders picked up in-store in 2016 for Home Depot, Internet Retailer's omnichannel merchant of the year***

45%

3.07%

Ecommerce grocery as a percentage of the total retail grocery market in the USA**



*Michael Koppel, CFO Nordstrom
**The Nielsen Company and Food Marketing Institute: "The Digitally Engaged Food Shopper"
***Business Insider, "Home Depot and Lowe's succeed at omnichannel"

Methodology Overview

What are the critical US markets for Home Delivery?

What drives customer channel choice?

Historical Data

Shopping Behavior

Geo-Location

→

Predictive Tool

- Heat-Map for deploying Home Delivery capabilities

Survey

Demographics

Competitive Landscape

Channel Preference

→

Channel Preference Model

- Quantified effects of channel features on channel choice

Exploratory Modeling

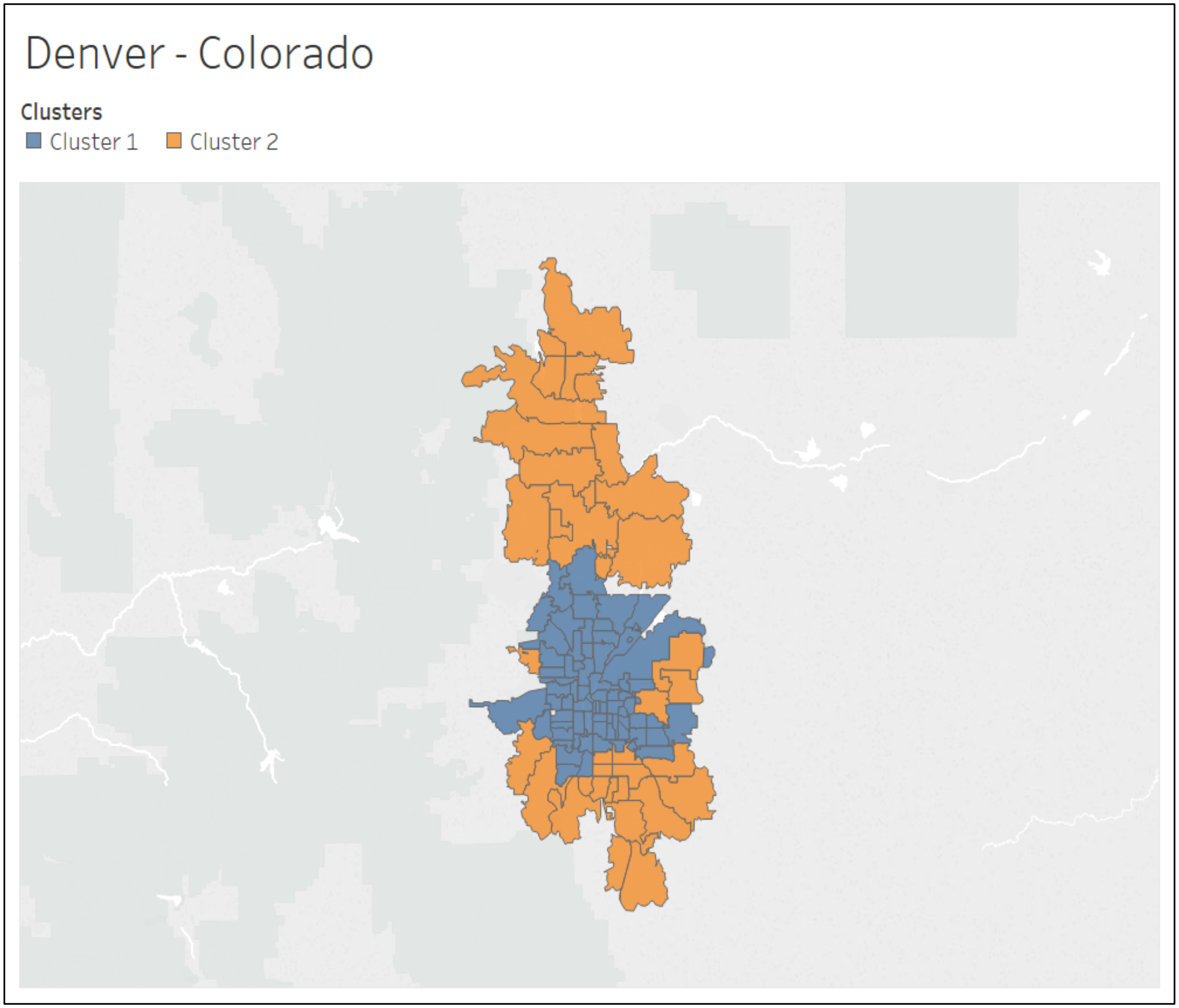
Cluster 1 (Blue)

- Urban
- Dense population
- Many competitors
- Higher mean sales

Cluster 2 (Orange)

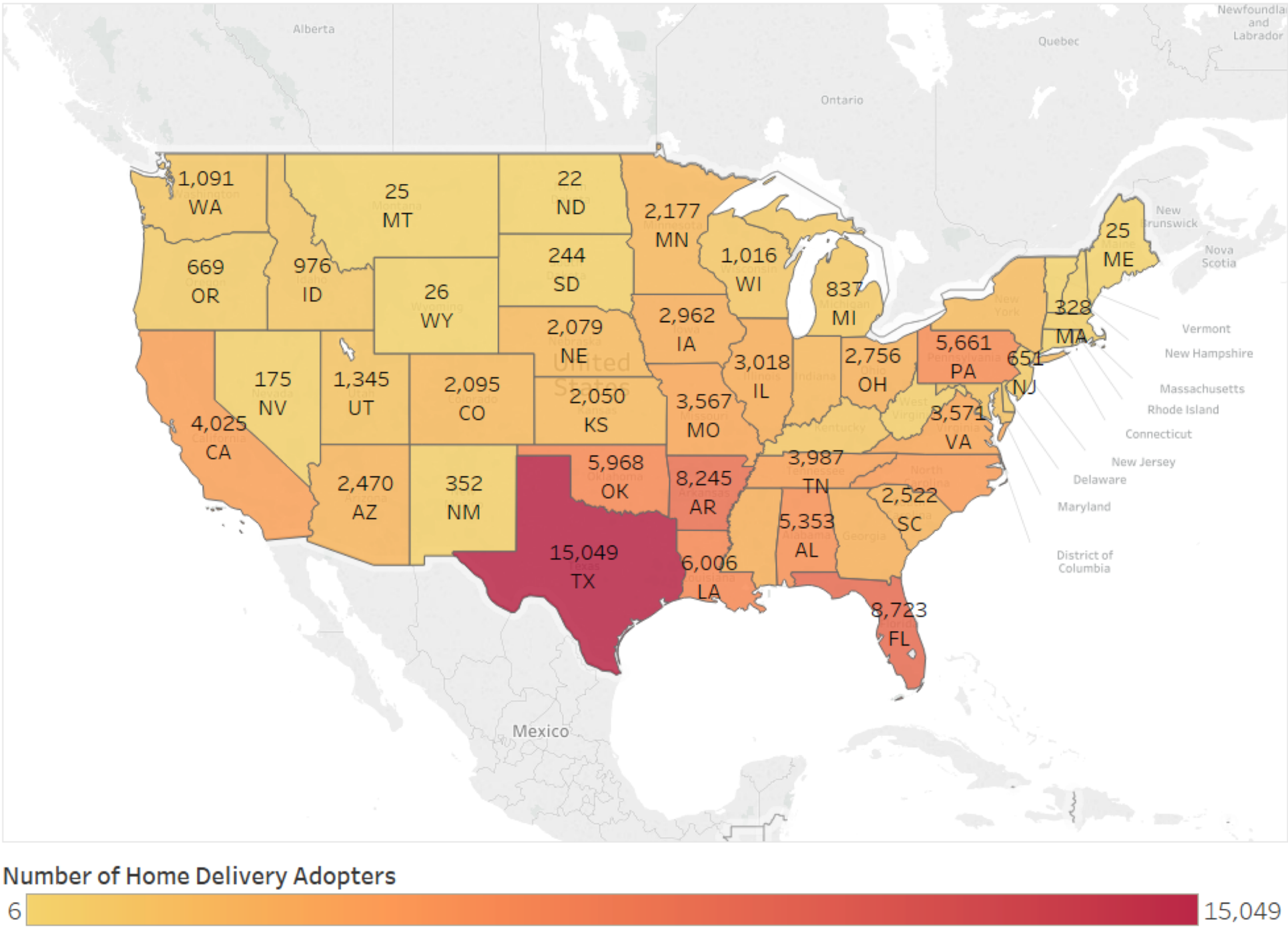
- Suburban / Rural
- Less dense population
- Few competitors
- Lower mean sales

	Centers		
	Avg. Pop Density	Avg. Mean Sales	Avg. Total Comp
Cluster 1	4422	123	2.99
Cluster 2	1250	138	1.13



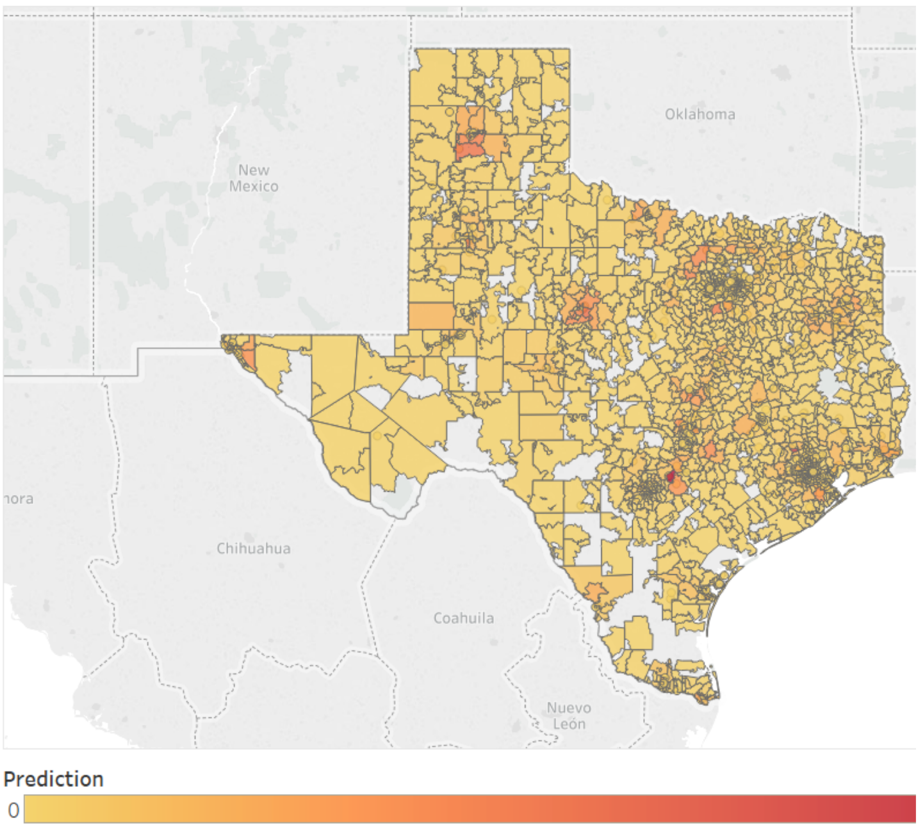
Findings – Heat-Map of Predictions

Home Delivery Adopters Volume by State

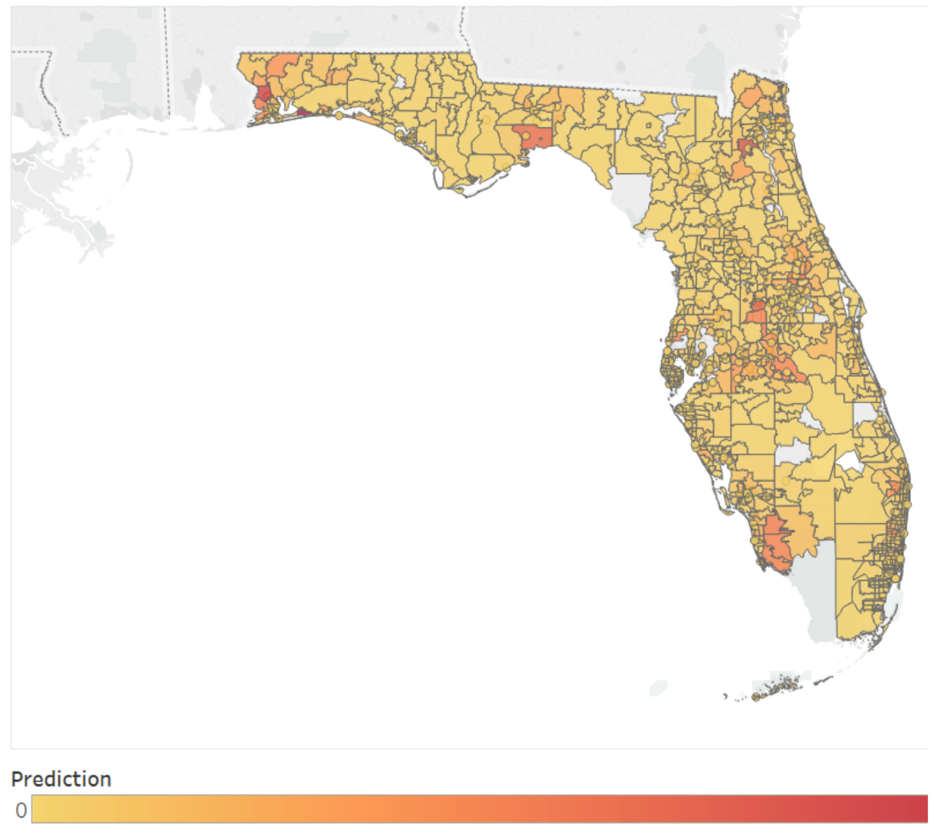


Findings – Heat-Map of Predictions

Texas - All Adopters

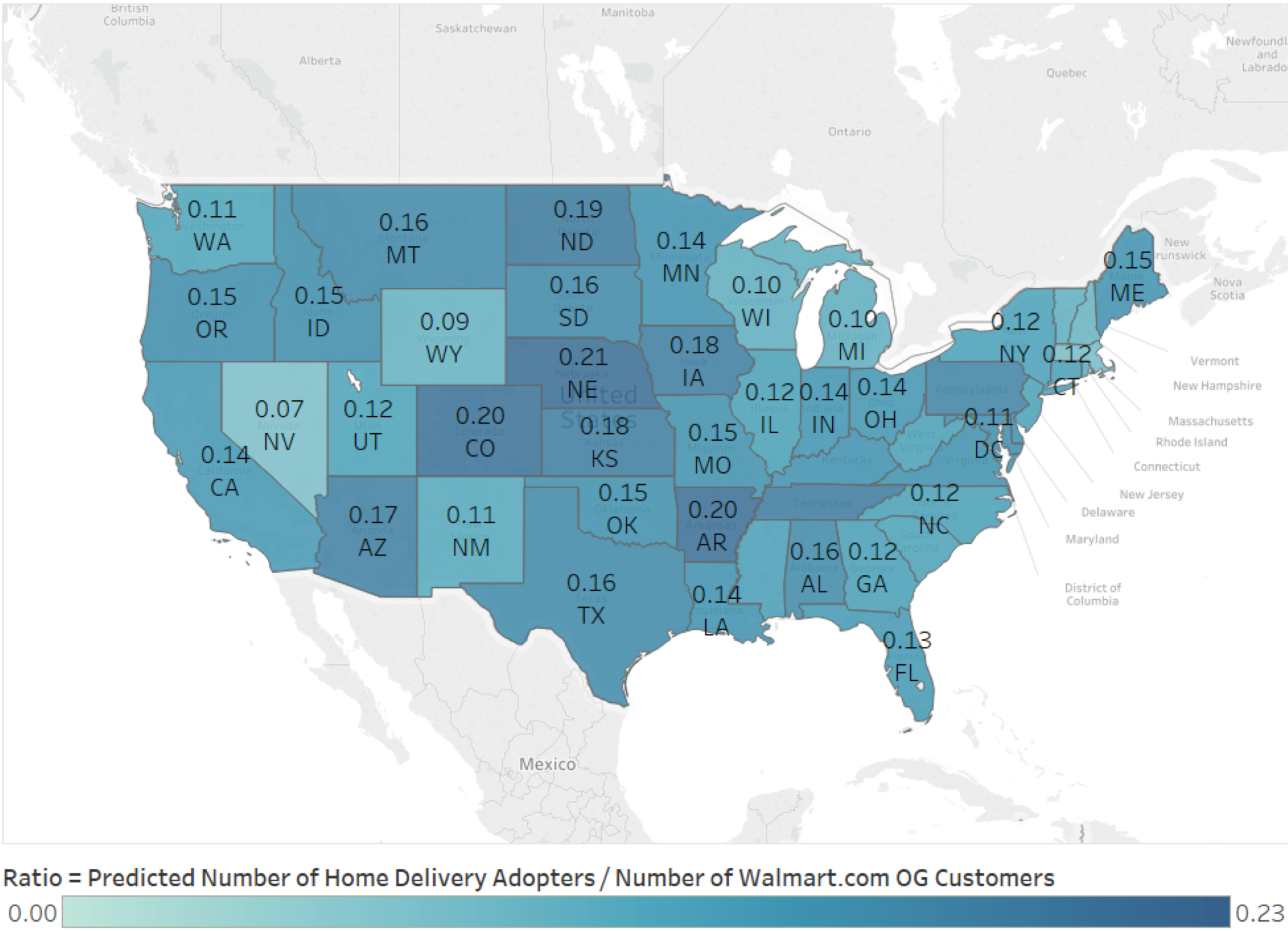


Florida - All Adopters



Findings – Heat-Map of Predictions

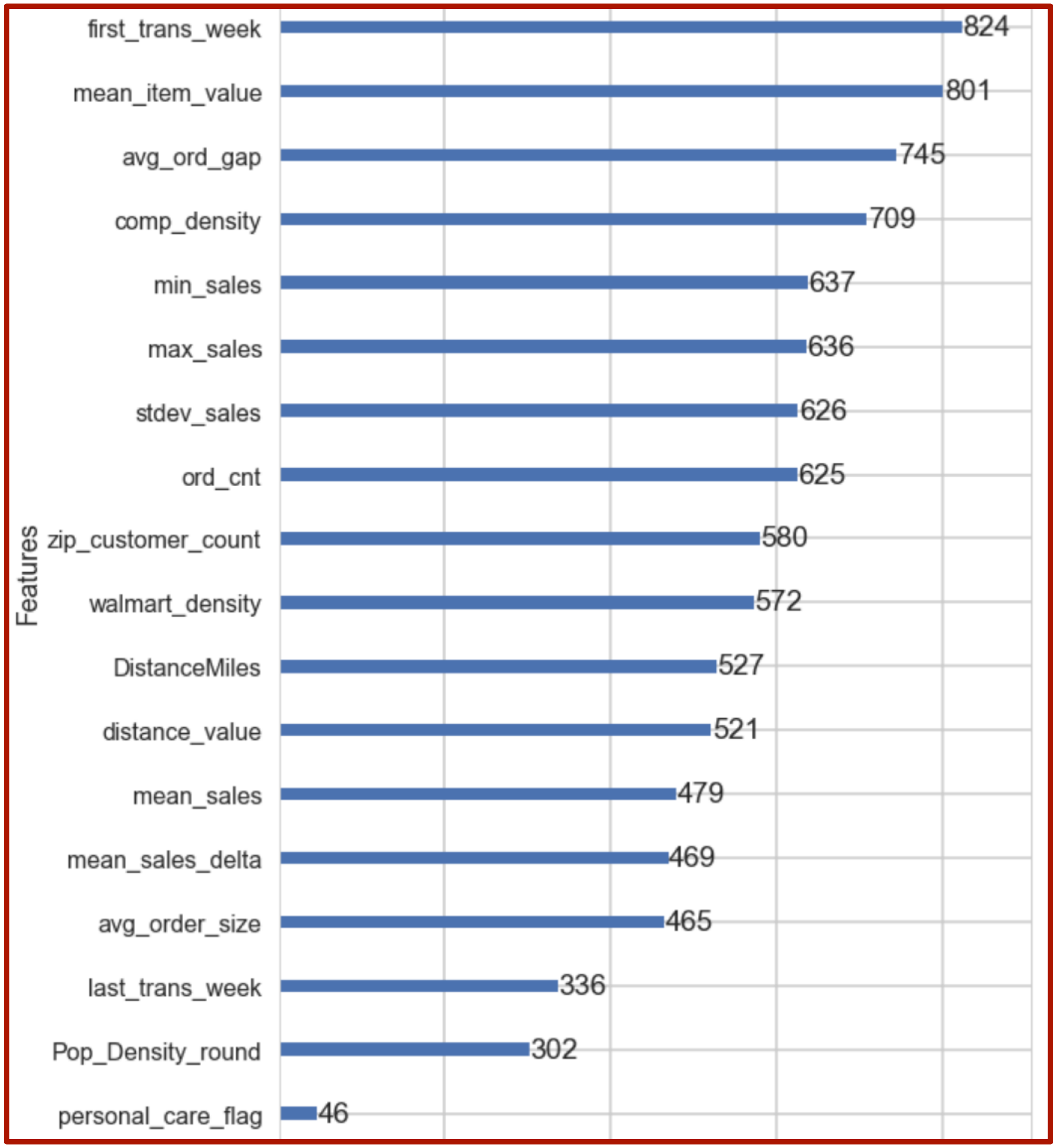
Home Delivery Adopters Density By State



Findings – Feature Ranking

Feature Rank:

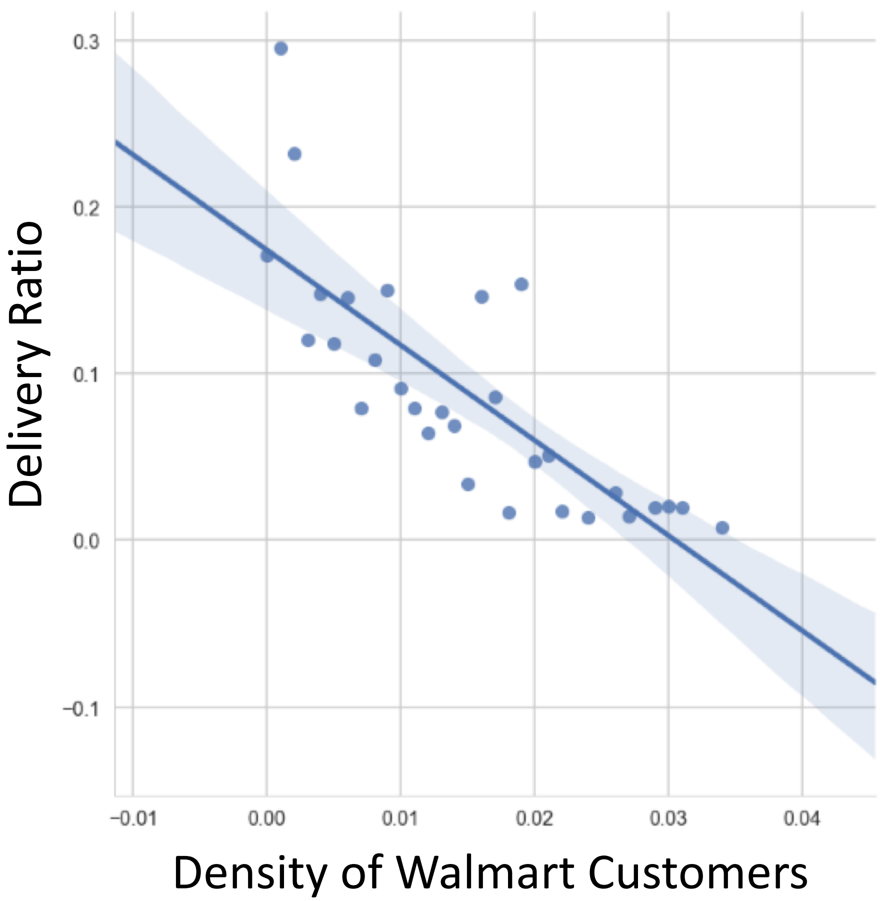
- Relative weight of each feature in model, i.e. predictive capacity



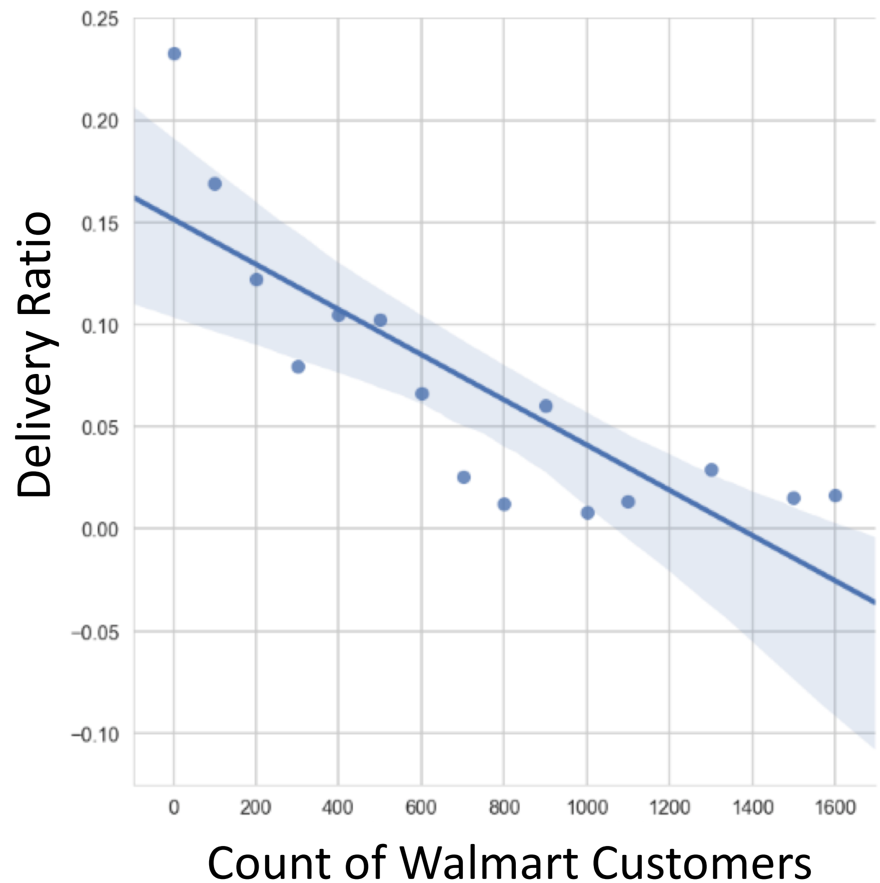
Findings – Feature Importance

1. Walmart Location Matters

Walmart Customer Density



Walmart Customer Count



Channel Choice Pipeline

Data Collection
- Survey released via Harris Poll
- 801 respondents from USA general population

Scenario Design
- Selection of attributes to define each choice
- Combination of levels to build choice sets

Statistical Analysis
- Random Effect Logit Model
- Findings

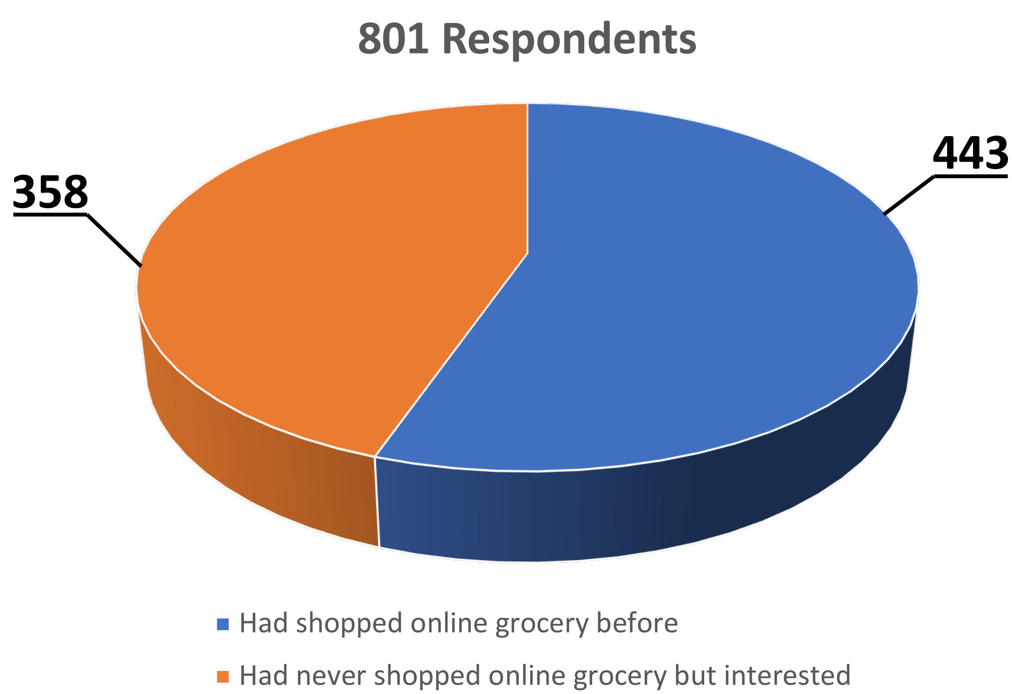
Data Collection

Survey released via Harris Poll to the USA general population

Current Shopping Behavior Questions

Demographics Questions

Choice Sets



Scenario Design – Attributes & Levels

- **Delivery Window**
- **Delivery Fee**
- Delivery Agent
- Distance from Store
- Store on commute

1.	Order placed anytime today and home delivered the next day	$6.99
2.	Order placed by 1pm and home delivered as soon as 4 hours	$9.99
3.	Order placed by 1pm and home delivered as soon as 1 hours	$14.99

Scenario Design – Example of Choice Set

Scenario & Grocery Definitions	Home Delivery	Pick-up from Store
Delivery Window	Order placed by 1pm and delivered as soon as one hour	Order placed by 1pm and ready for pick-up as soon as one hour
Delivery Fee	Delivery fee is $14.99	Free
Delivery Agent	Order delivered by Walmart Associate	Curbside Pick-up
Distance from Store	Store is between 10 and 15 miles from home	Store is between 10 and 15 miles from home
Store on commute	Store is on your daily commute	Store is on your daily commute
	You must be present during delivery window	
Which option would you choose?		
	❑	❑

Scenario Design – Attributes & Levels

- Delivery Window
- Delivery Fee
- Delivery Agent
- Distance from Store
- Store on commute

1.	Order placed anytime today and picked up from store the next day	Free
2.	Order placed by 1pm and picked up from store as soon as 4 hours	Free
3.	Order placed by 1pm and picked up from store as soon as 1 hours	Free

Scenario Design – Attributes & Levels

- Delivery Window
- Delivery Fee
- **Delivery Agent**
- Distance from Store
- Store on commute

1.	Walmart Associate
2.	3rd Party (e.g. Uber/Deliv)

Scenario Design – Attributes & Levels

- Delivery Window
- Delivery Agent
- Delivery Fee
- **Distance from Store**
- Store on commute

1.	Less than 10 miles
2.	10 to 15 miles
3.	More than 15 miles

Scenario Design – Attributes & Levels

Delivery Window

Delivery Agent

Delivery Fee

Distance from Store

Store on commute

1.	Yes
2.	No

Findings

1) Sensitivity to Delivery Window + Cost

- Home Delivery:

Window	Cost
Order placed anytime today and delivered the next day	6.99
Order placed by 1pm and delivered as soon as 4 hours	9.99
Order placed by 1pm and delivered as soon as 1 hours	14.99

20.7% less likely to choose home delivery

- Pick up from Store:

Window	Cost
Order placed anytime today and delivered the next day	Free
Order placed by 1pm and delivered as soon as 4 hours	Free
Order placed by 1pm and delivered as soon as 1 hours	Free

Pick-up window does NOT seem to be a factor in channel choice

Findings

2) Sensitivity to Store Distance

A customer is 2.77 times more likely to choose home delivery when a store is 15+ miles away as compared to a customer with a store that is less than 10 miles away

3) Sensitivity to Delivery Agent

Agent
Walmart Associate
3rd party (e.g. Uber/Deliv)

Customers are NOT sensitive to the delivery agent. It does not affect their channel choice

Findings

4) Sensitivity to Car Availability

Having a car reduces the likelihood of home delivery by 74.88%

5) Sensitivity to Age

Seniors (65+ years old) are 63.64% less likely to choose home delivery than Generation Z (18 – 24 years old)

Limitations & Future Research

Overfitting, disentangling variables, hypothetical bias

Conclusion

Methodology for implementing omnichannel distribution strategy for online grocery

Online Grocery & Omnichannel Strategy: Predicting Home Delivery Adoption

Thank you!

Ryan Alexander Alberts, M.A.Sci. Candidate

Antoine Lahad Abinader, M.A.Sci. Candidate

Dr. Eva Ponce, Ph.D., Advisor

Model Selection

Gradient Boosting Machines

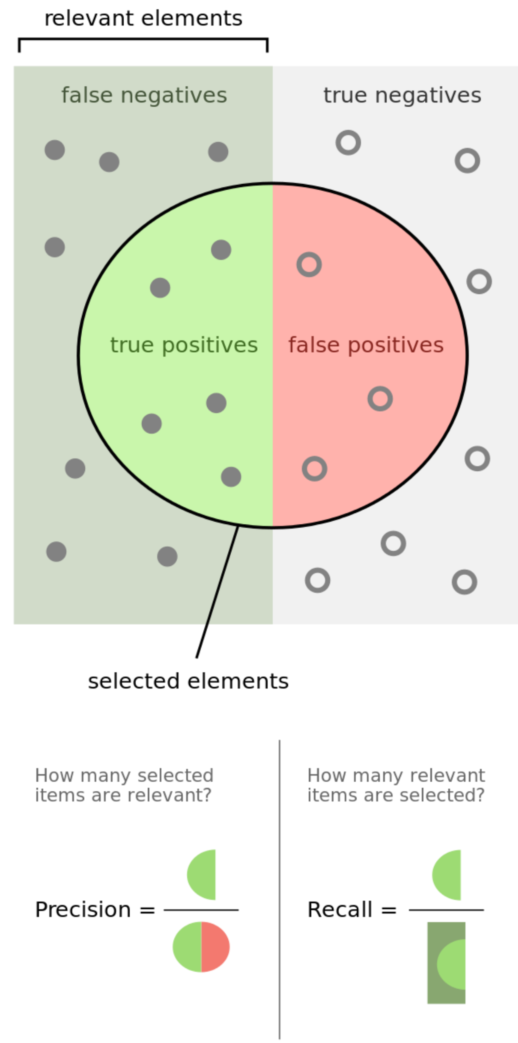
		Predictions		
		Precision	**Recall**	**Support**
Actuals	**0**	0.92 *(8,428)*	0.94 *(568)*	*(8,996)*
	1	0.63 *(689)*	0.59 *(978)*	*(1,667)*
		f1 score = 0.88		*(10,663)*

k-Nearest Neighbors

		Predictions		
		Precision	**Recall**	**Support**
Actuals	**0**	0.90 *(8,580)*	0.95 *(416)*	*(8,996)*
	1	0.64 *(915)*	0.45 *(752)*	*(1,667)*
		f1 score = 0.87		*(10,663)*

Naïve Bayes

		Predictions		
		Precision	**Recall**	**Support**
Actuals	**0**	0.90 *(8,638)*	0.96 *(358)*	*(8,996)*
	1	0.67 *(927)*	0.44 *(740)*	*(1,667)*
		f1 score = 0.87		*(10,663)*



Contact Information:

Ryan Alberts | ralberts@mit.edu
Antoine Abinader | antoine4@mit.edu
Eva Ponce | eponce@mit.edu

Ryan Alexander Alberts, M.A.Sci. Candidate

Antoine Lahad Abinader, M.A.Sci. Candidate

Dr. Eva Ponce, Ph.D., Advisor

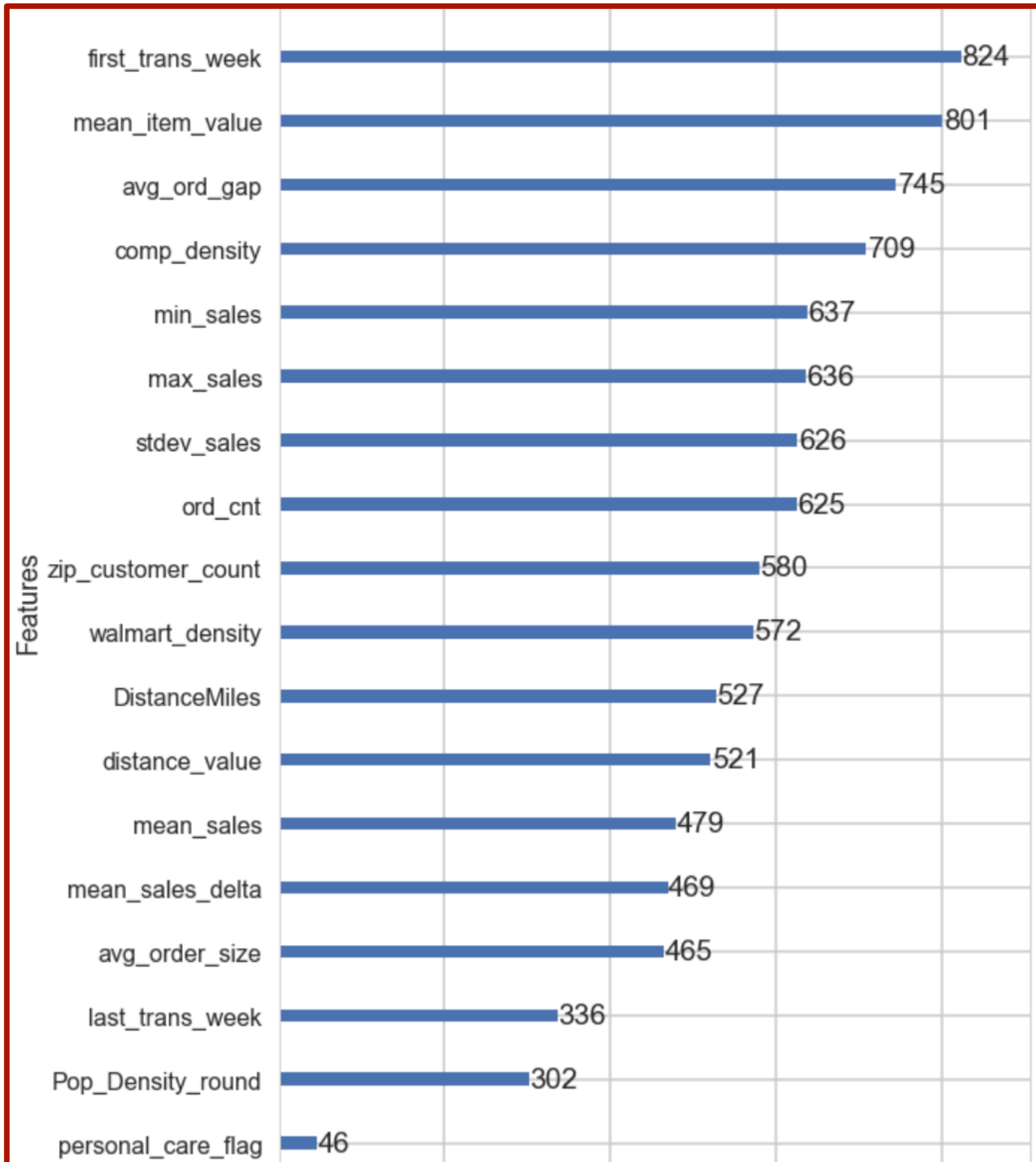
Findings

Feature Rank:

- Relative weight of each feature in model, i.e. predictive capacity

Feature Engineering:

- Mean Item Value
- Walmart Customer Density
- Zip-Code Customer Count



Classification Models: kNN, Naïve Bayes, GBM

Gradient Boosting Machines

Ensemble tree method of iterative gradient descent focusing on misclassified observations, using stepwise prediction function at each leaf:

$$f(x) = \sum_{m}^{M} c_m I(x, R_m)$$

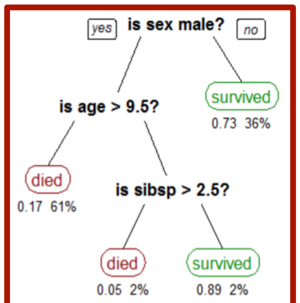
M = Number of Leaves
R_m = Region in feature space leaf m
c_m = feature constant at leaf m
I = Indicator function where:

$$I = \begin{cases} 1, & x \in R_m \\ 0, & x \notin R_m \end{cases}$$

Logistic Regression: binary classifier
Evaluative metric: log loss, AUC

Loss Function
+
Regularization

$$-\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij} \log p_{ij}$$



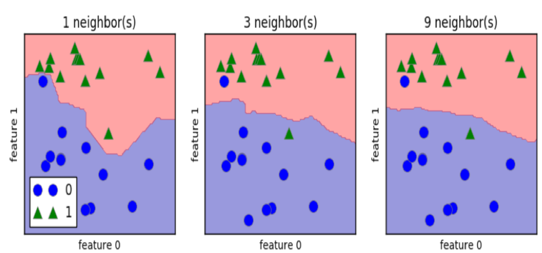
k-Nearest Neighbors

Non-parametric majority vote of closest observations according to Euclidean distance (p=2):

$$d(x, y) = \sum_{i=0}^{N-1} |x_i - y_i|^{1/p}$$

N = Number of Observations
x_i = Observation x at point i
y_i = Observation y at point i

No feature distribution assumptions
No explicit training phase



Naïve Bayes

Probabilistic classifier that applies Bayes Rule:

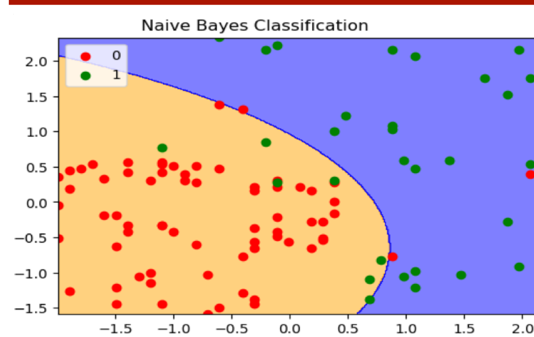
$$P(y \mid x_1, \ldots, x_n) \propto P(y) \prod_{i=1}^{n} P(x_i \mid y)$$

$$\Downarrow$$

$$\hat{y} = \arg\max_{y} P(y) \prod_{i=1}^{n} P(x_i \mid y),$$

y = Class Variable
x_i = Feature vector x at point i
$\hat{y}$ = Predicted y at point i

Feature Likelihood is Gaussian
Naïve independence assumptions



Online Grocery Customer Analysis