

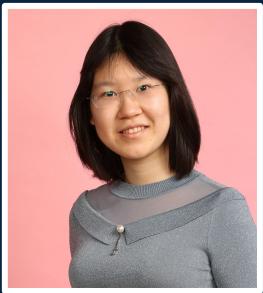


Notre Dame
Data Club

Fall 2023

34.51°

Team Introduction



Qingdi (Cindy) Du
Business Analytics + ACMS
Beijing, China



Brendan Ye
Finance/ACMS
The Woodlands, TX



Henrique Raposo
Business Analytics + Econ
Rio de Janeiro, Brazil



Martin Alvarez
Finance + ACMS
Guatemala City, Guatemala



Joseph Park
Computer Science
Austin, TX



Nathan Fanous
ACMS + Data Science
Davis, CA

Team Introduction Cont.



Alex Clark
Business Analytics
Atlanta, GA

Grant Flesher
Computer Science + Econ
Chesterton, IN

Patrick Oglesby
Econ
Huntersville, NC



Sean Kim
Business Analytics + ACMS
Orlando, FL

Elena Zhang
Statistics
Beijing, China

Executive Summary



01	Introduction	<ul style="list-style-type: none">• Coupon Design Overview• Purpose
02	Demographics VS Coupon Usage	<ul style="list-style-type: none">• Logistic regression• Descriptive Analysis• Campaign Type Effectiveness• Time Series Analysis• Customer/Product Based Coupon Design
03	Trend & Seasonality VS Coupon Usage	<ul style="list-style-type: none">• Correlation Analysis• Logistic Regression
04	Customer Behavior	<ul style="list-style-type: none">• Loyalty Program• Source of Coupon
05	Recommendation & Other Considerations	

Through our research, data manipulation, and coding, we conducted descriptive and predictive analysis to provide recommendations for more efficient coupon design.

Prompt & Dataset



- **Prompt:** Kroger is looking to **streamline advertising** by targeting specific households that are most likely to respond positively to certain coupons. We would like you to identify these households using any modeling method (Regression, Clustering, etc.) based on **previous household responses** to similar campaigns using household demographics, coupon usage, transactions, etc.
- **Dataset:** completejourney_py or completejourney package in R

Questions of Interest



- What targeted demographics led to higher coupon redemption rates?
- How did coupon usage change over time?
- How did customer behavior change upon receiving a coupon?



Three Types of Coupon Designs



- Weekly Digital Deals

This Week's Hottest Coupons

Simply clip your coupons and use up to 5 times through 11/28 delivery. Be sure to check back 11/29 for even more great deals.

The screenshot shows a mobile-style interface for 'This Week's Hottest Coupons'. It features a central card for a 'Monster Energy' coupon worth '\$1.49' with an expiration date of 'Exp. Nov. 28 - Today!'. Below the card are buttons for 'Shop All Items' and 'Sign In To Clip'. To the right, there are two more coupon cards: one for '\$4.99 Old Spice or Secret S...' with an expiration date of 'Exp. Nov. 28 - Today!', and another partially visible for 'JELL-O'. Each card also has 'Shop All Items' and 'Sign In To Clip' buttons.

- Pickup/Delivery Deals

SAVE on Pickup or Delivery

The screenshot shows a mobile-style interface for 'SAVE on Pickup or Delivery'. It features a central card for a '\$20 OFF' coupon with an expiration date of 'Exp. Dec. 18'. Below the card are buttons for 'Shop All Items' and 'Sign In To Clip'. To the right, there are two more coupon cards: one for 'Save 30% off Ryl Tea PICKL' with an expiration date of 'Exp. Nov. 28 - Today!', and another partially visible for a different product. Each card also has 'Shop All Items' and 'Sign In To Clip' buttons.

- Themed Deals

Explore Holiday Categories

The screenshot shows a mobile-style interface for 'Explore Holiday Categories'. It features four circular icons on a red background: 'Holiday Meal Planning' (a ham), 'Gift Guide' (a wrapped gift), 'Savings' (a piggy bank), and 'Holiday Baking' (a tray of cookies). Each icon has a small descriptive text label below it.

What is our Purpose?



- Low Redemption Rate of 1.81%
- **Goal:** boost redemption rate by analyzing demographics and time series by various models

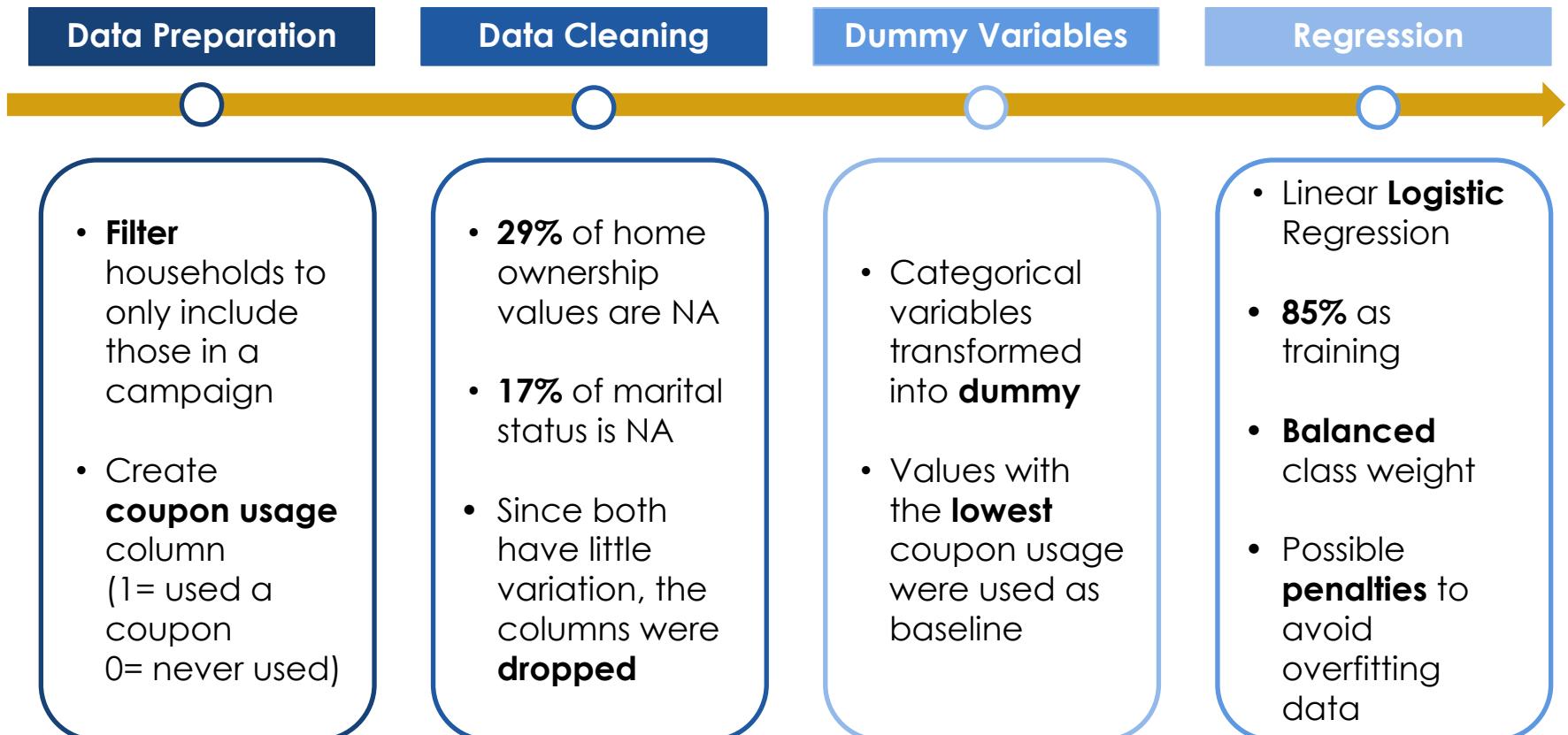


Demographics

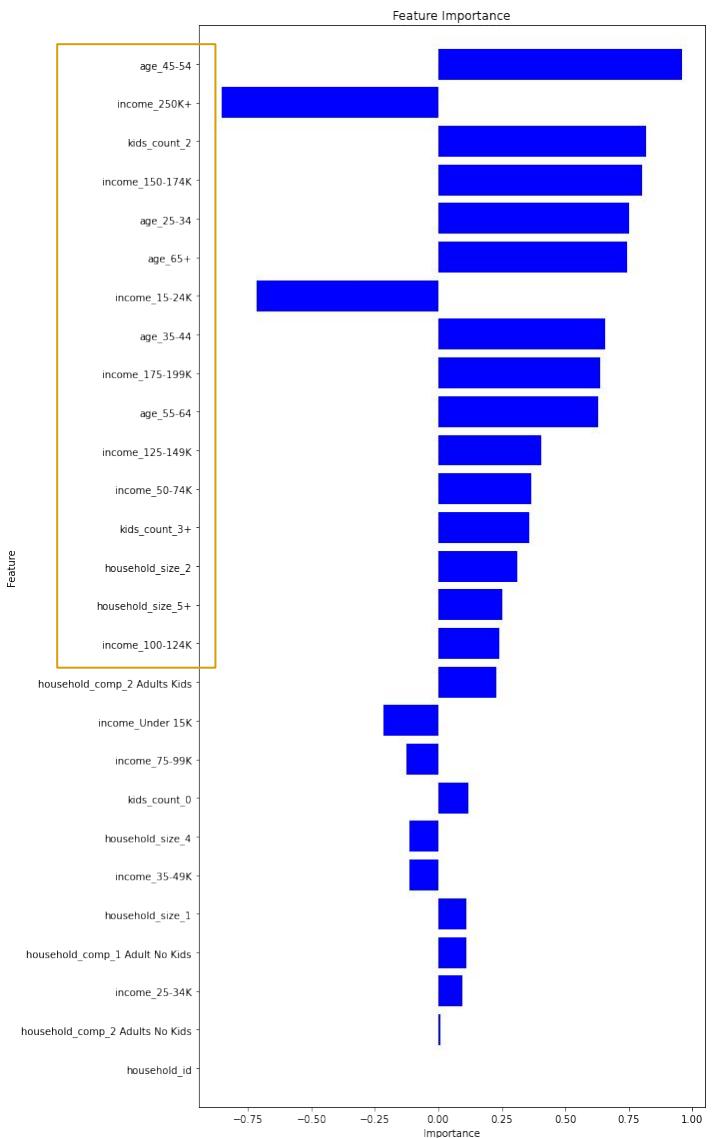
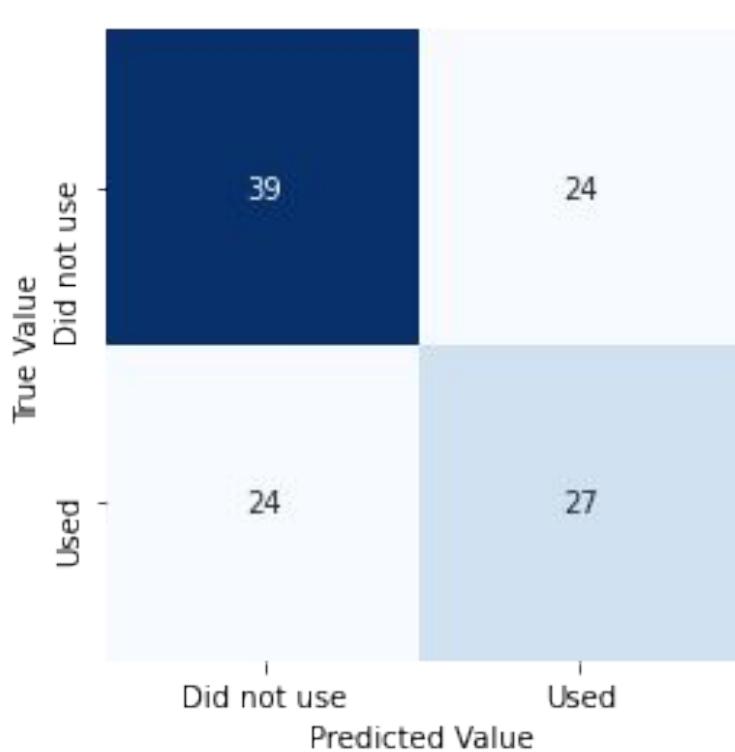
Demographics Logistic Regression: how to identify most important characteristics?



Goal: identify most important demographic characteristics in predicting coupon redemption rate

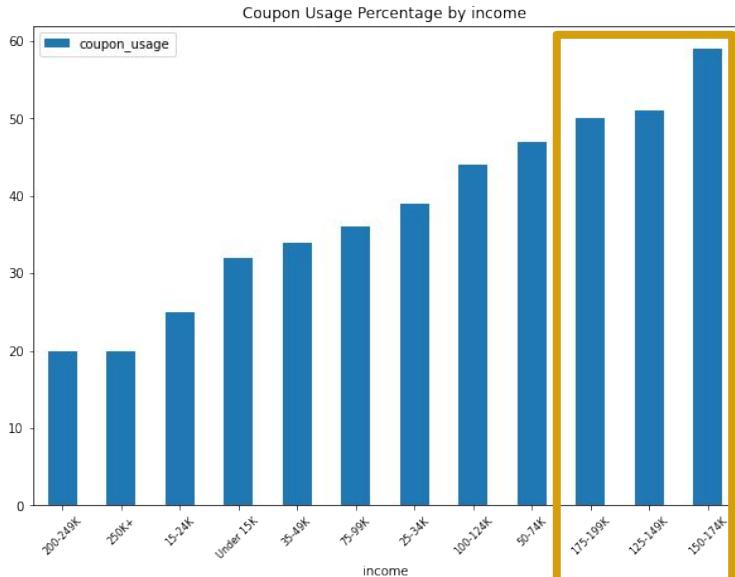


The focus should be on income, age, kids count and household size



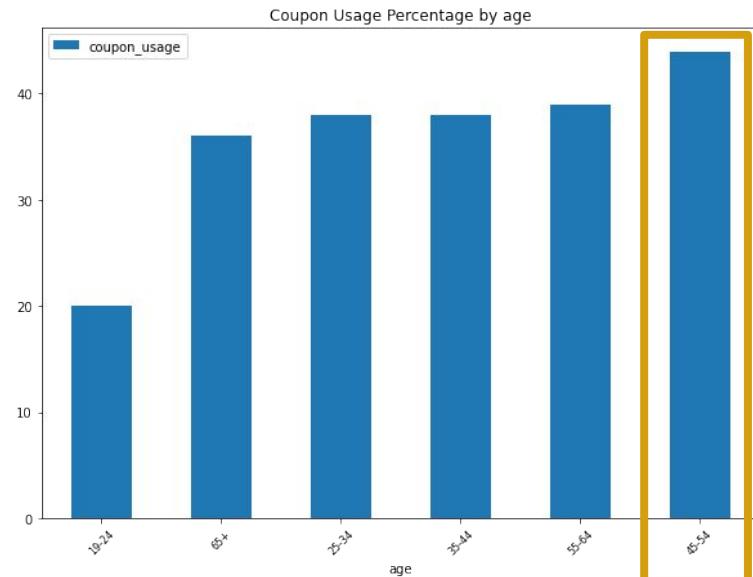
- 57.89% accuracy
- The most important variables in the model are income, age, kids count and household size

\$125-199K households and middle-aged customers used coupons more frequently



INCOME

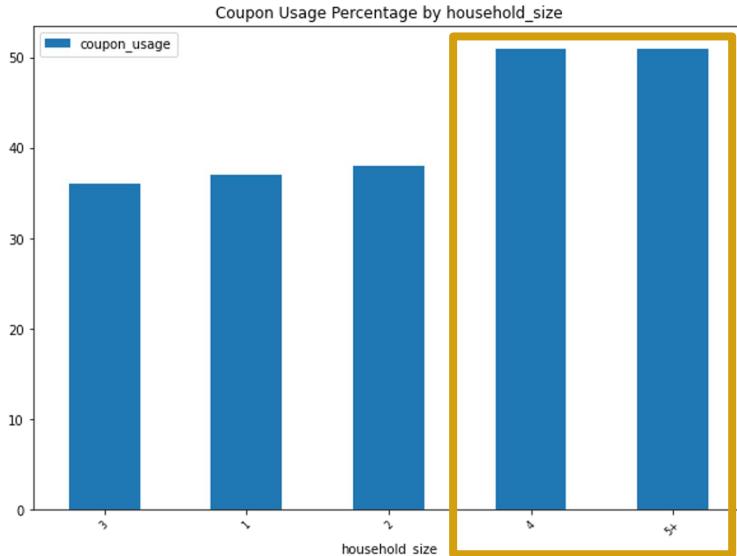
- \$125K-199K consistently used coupons more
- > \$200K used coupons the least



AGE

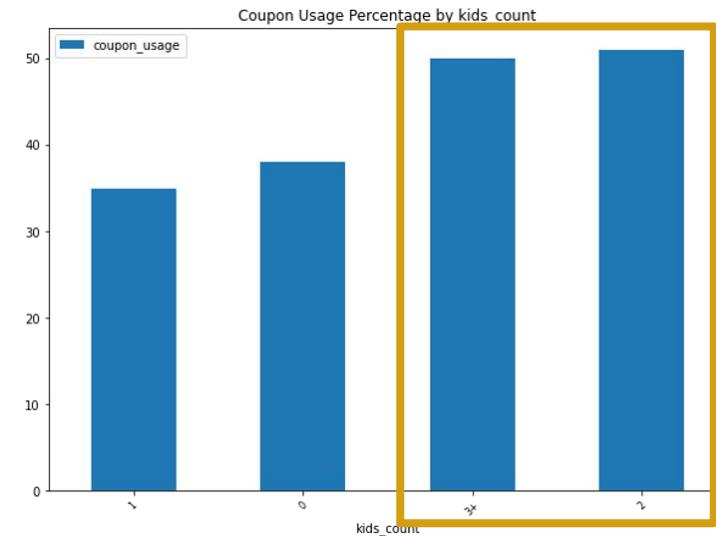
- Largest coupon usage within 45-54 year old age group
- 19-24 age group used coupons the least

Households with more members and more kids used coupons more frequently



HOUSEHOLD SIZE

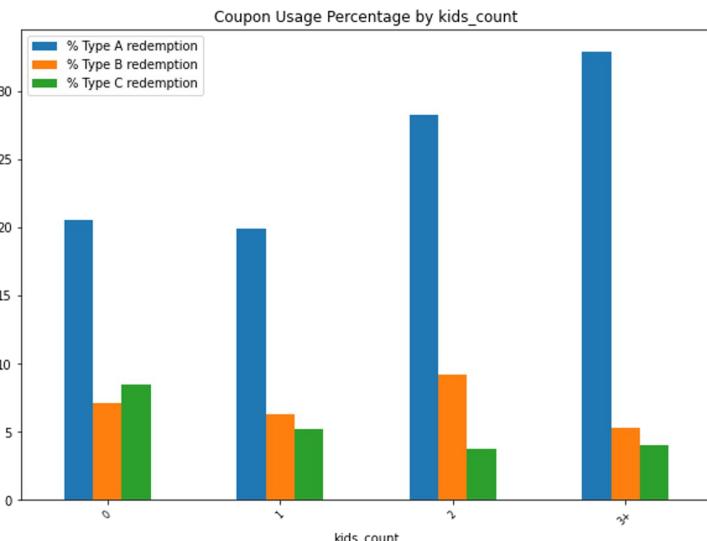
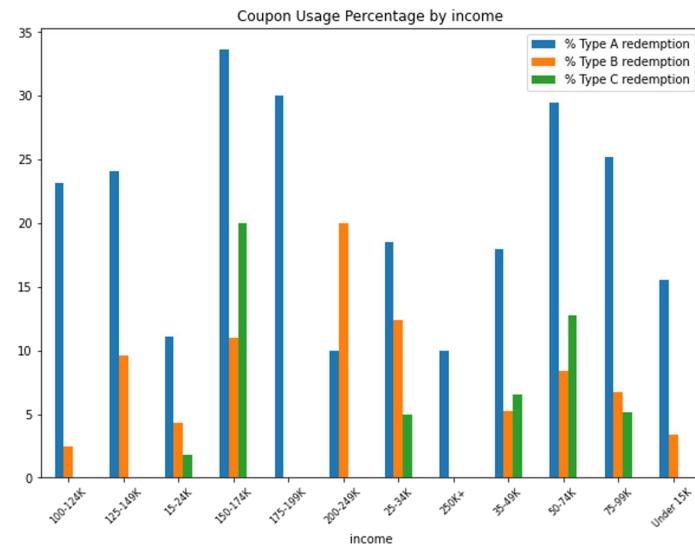
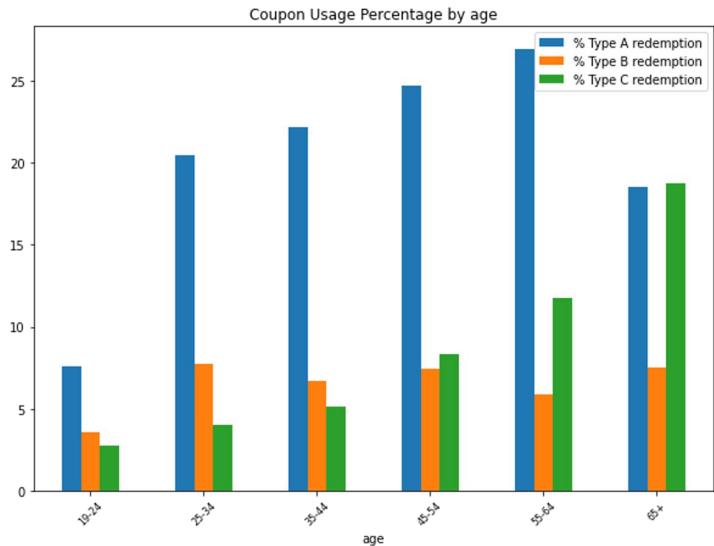
- Households with 4+ members used coupons 15% more often than those with <4 members



KIDS COUNT

- Households with more kids redeemed coupons more
- Likely because larger households have higher financial costs

Campaign Type Effectiveness by Demographics: Type A most effective



- Trends were generally consistent across all 3 campaign types
- Type A resulted in the largest redemption rates

Demographics-Based Recommendations



- Continue focused marketing directed towards key demographics
- Integrate coupons with marketing strategies directed towards children
 - Kroger Kids Fresh Friends
 - Game Better with Butterfinger and Kroger
 - Designing coupons for brands popular with children



Demographics-Based Recommendations



- Design coupons for upper class customers
 - Premium/organic products
 - Higher quality brands
- Advertise towards middle aged customers
 - Loyalty programs
 - Targeted ads on relevant sites



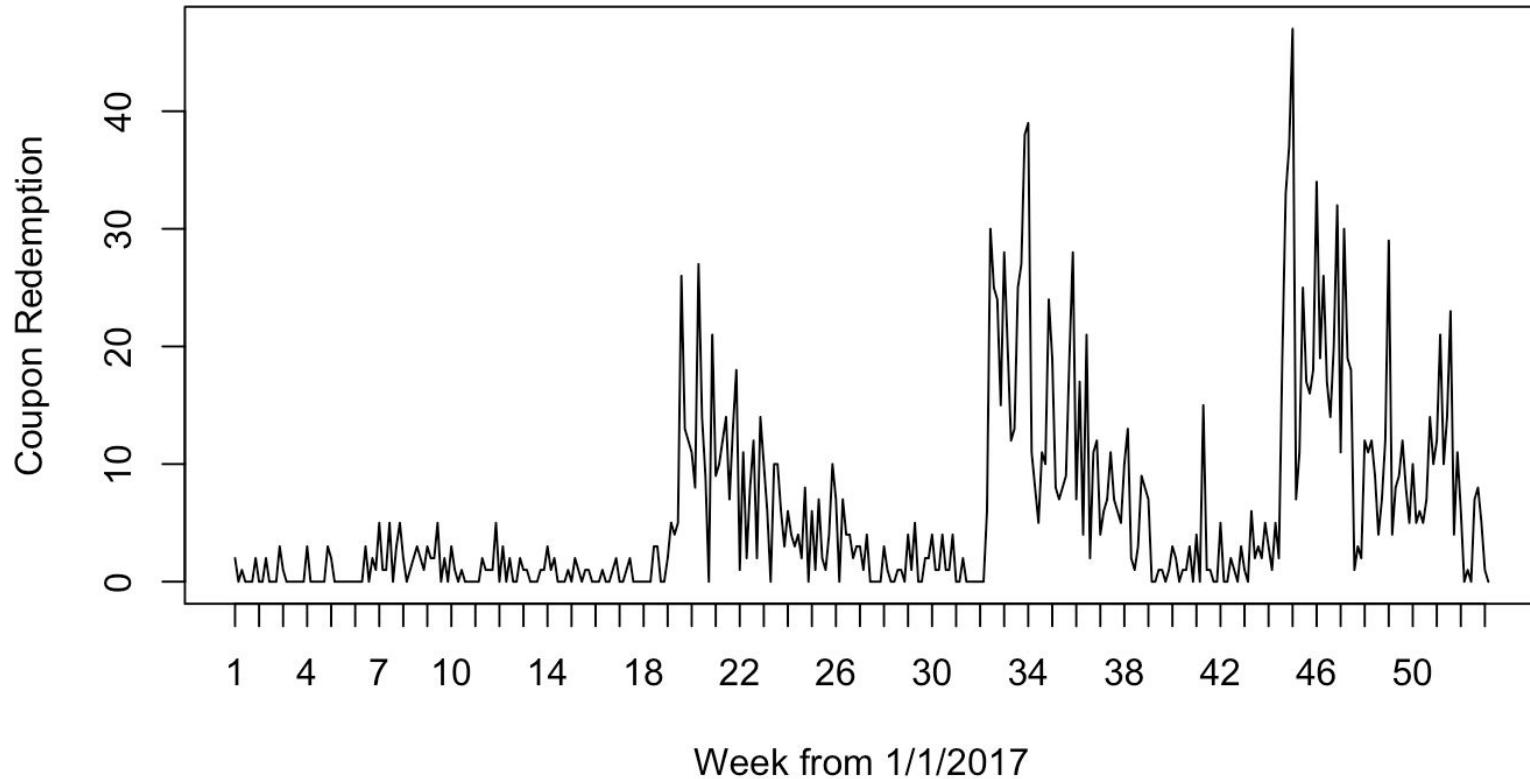


Trend & Seasonality

Weekly coupon usage has an increasing trend with seasonality



1/1/2017-1/1/2018 Weekly Coupon Redemption

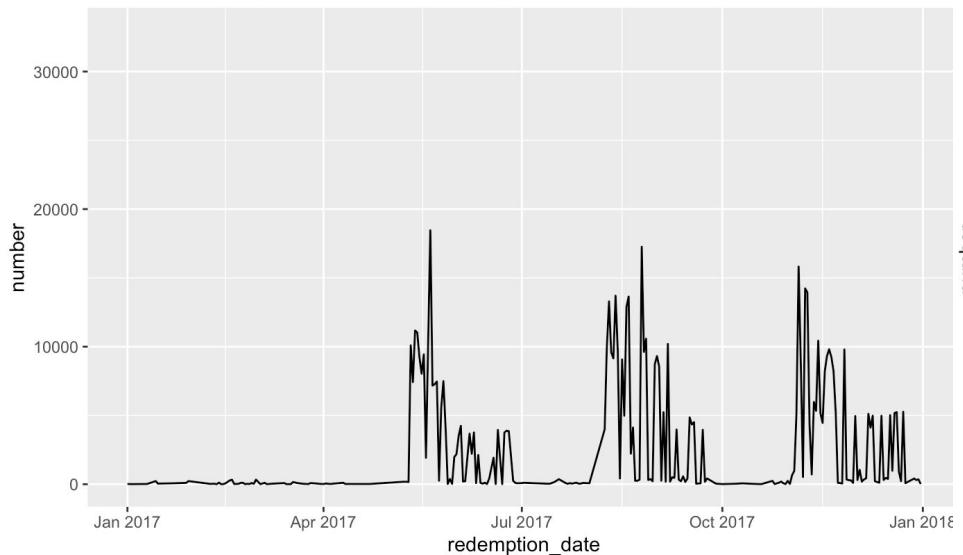


- Three peaks: mid April, mid July, early October
- Great fluctuations

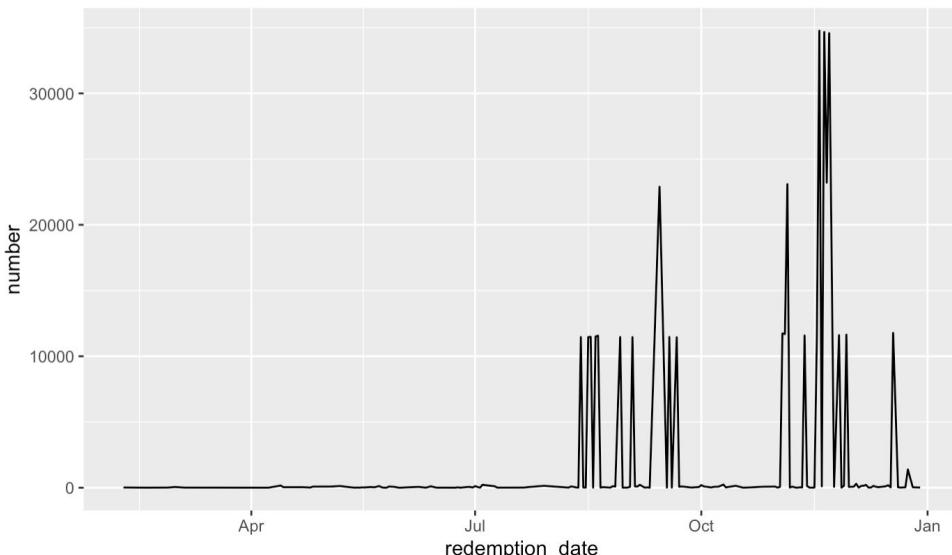
Breakdown by Department: Grocery Department build up the largest coupon usage



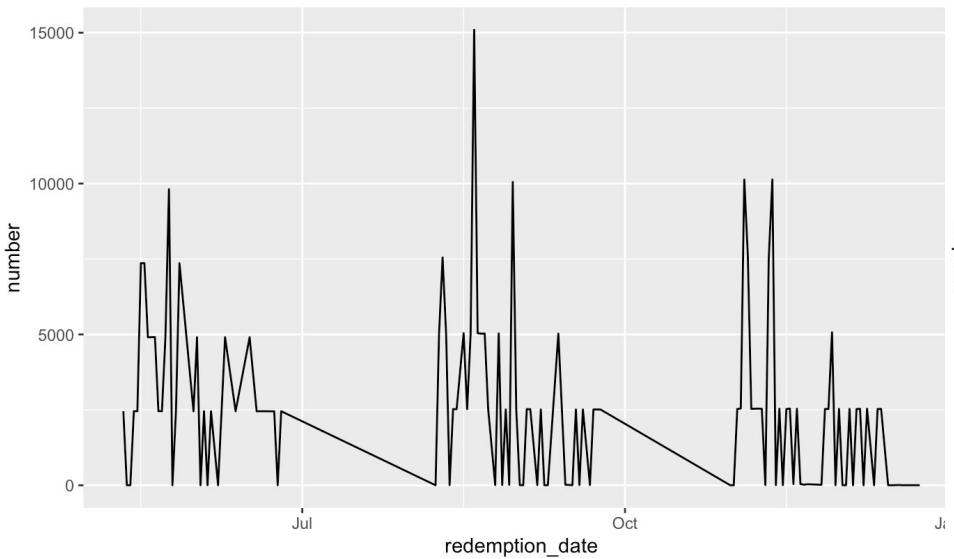
Coupon Redemption Count for Grocery Department



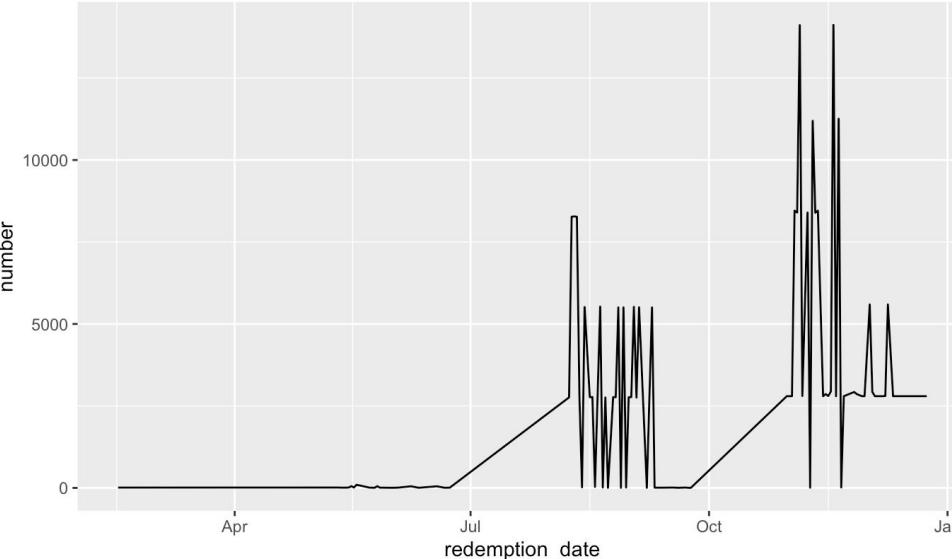
Coupon Redemption Count for Drug-General Merchandise



Coupon Redemption Count for Meat



Coupon Redemption Count for Produce



Breakdown by Customer Demographics: Customers were grouped into low-, middle-, and high-income class based on household composition and income



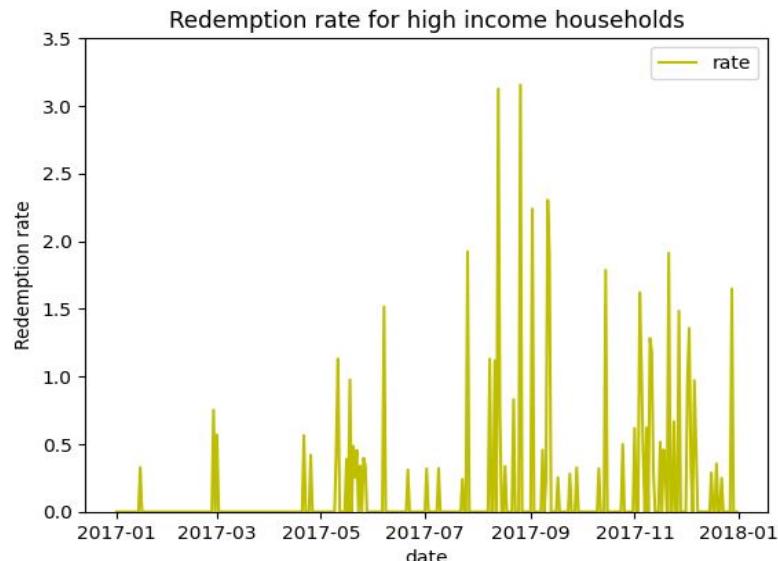
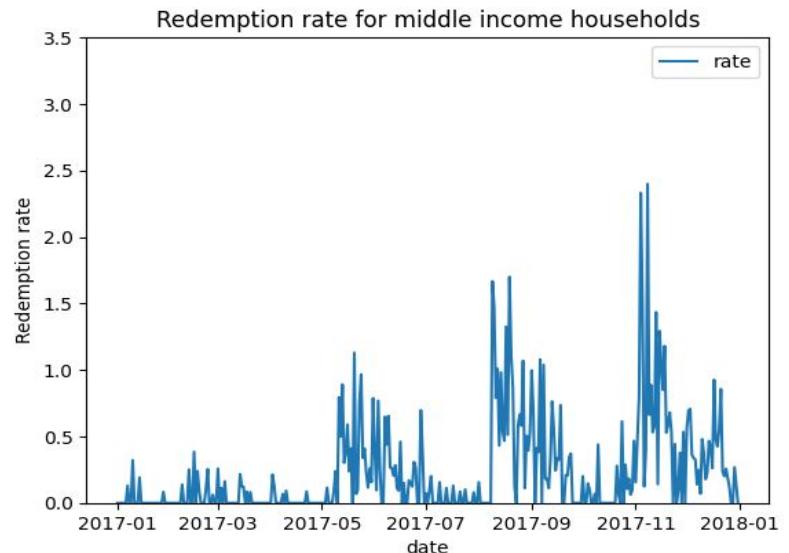
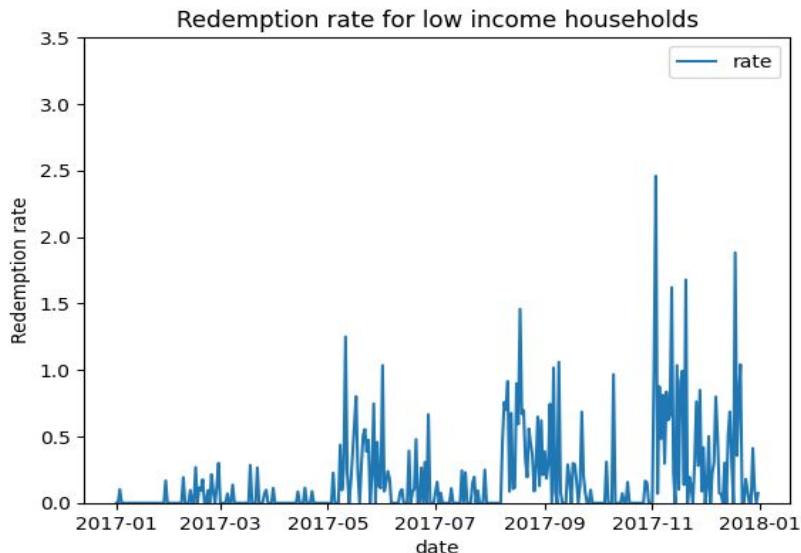
“For a single individual, a middle-class income ranges from \$30,000 - \$90,000 per year.

For a couple it starts at \$42,430 up to \$127,300; for a family of three, \$60,000 - \$180,000; and four \$67,100 - \$201,270.”¹

Customer Demographics			
	Lower	Middle	Upper
1 Adult Kids	<\$49k	\$49k-\$149k	>\$149k
1 Adult No Kids	<\$34k	\$34k-\$99k	>\$99k
2 Adults Kids	<\$49k	\$49k-\$149k	>\$149k
2 Adults No Kids	<\$49k	\$49k-\$124k	>\$124k

¹[Pew Research Center](#)

Coupon redemption redemption rate is affected by household income



Low and Middle Income:

- Seasonality

High Income:

- Possible outliers that are randomly distributed

Outliers: Reasons behind high redemption rate for high income households



household_id	coupon_upc	campaign_id	redemption_date	product_category	product_type
936	1197	10000085429	13	2017-08-26	YOGURT
937	1197	50308412076	13	2017-08-26	CONVENIENT BRKFST/WHLSM SNACKS
938	1197	50813790076	13	2017-08-26	CONVENIENT BRKFST/WHLSM SNACKS
939	1197	51600070033	13	2017-08-26	YOGURT
940	1197	51980010075	13	2017-08-26	PICKLE/RELISH/PKLD VEG
941	1197	52740022050	13	2017-08-26	MISC. DAIRY
942	1197	52840050076	13	2017-08-26	BEEF
943	1197	53700040075	13	2017-08-26	CHEESE
944	1197	54300020076	13	2017-08-26	BREAKFAST SAUSAGE/SANDWICHES
945	1197	55410000076	13	2017-08-26	CRACKERS/MISC BKD FD
946	1197	56414410013	13	2017-08-26	SOUP CRACKERS (SALTINE/OYSTER)
947	1197	57047091041	13	2017-08-26	BACON
948	1197	57631000076	13	2017-08-26	GRAPES

The high redemption rate is due to the shopping behavior of one single household on a certain day.

For the below 4 time series models, model with linear trend and seasonality provides the best accuracy

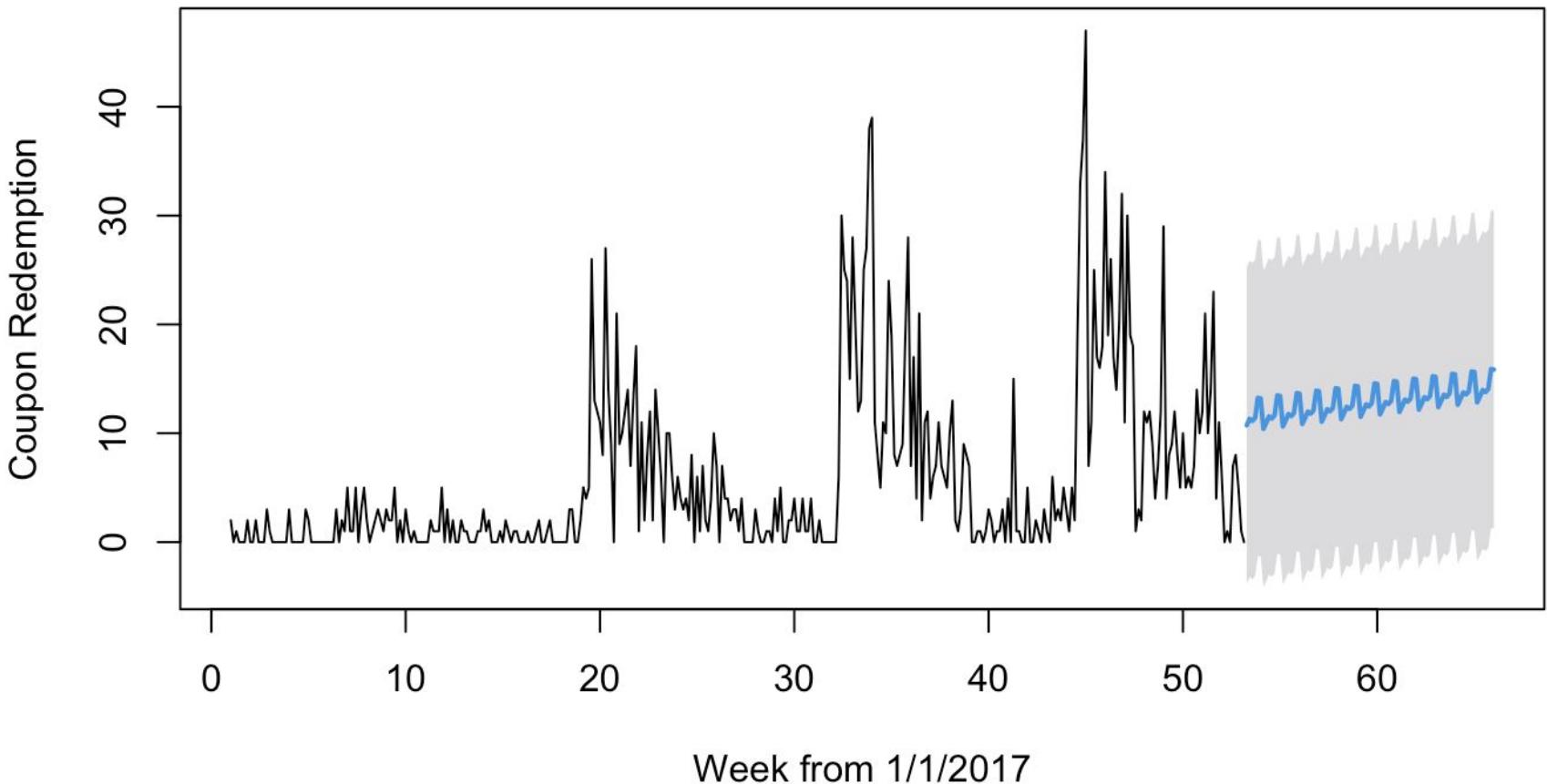


Model	ME	RMSE	MAE
Linear Trend with Seasonality	2.451	10.358	7.492
Polynomial Trend with Seasonality	6.387	11.722	8.209
ETS(A,N,N)	10.082	14.232	10.297
ARIMA	9.999	14.178	10.256

Apply the model to the entire dataset and predict the number of coupon redemption for the next quarter (90 days)



Forecasts from Linear regression model





Customer Behavior



Four new variables were created to account for household buying behavior



Sales_no_disc

how much each household spent on average on a basket without discount

Number_of_purchases

number of times a household made a purchase in the year

Unique_stores_visited

number of unique stores the household visited

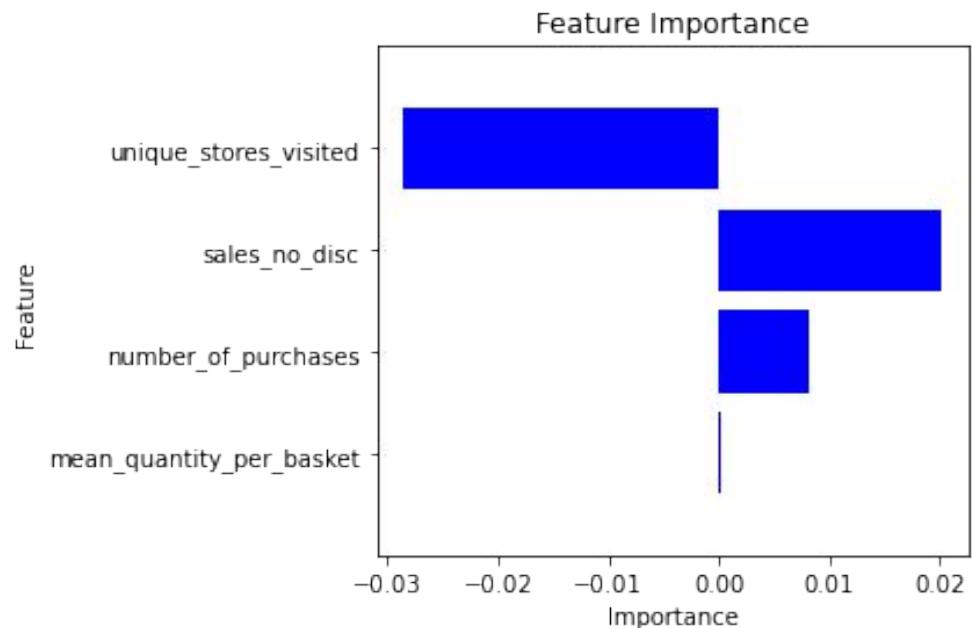
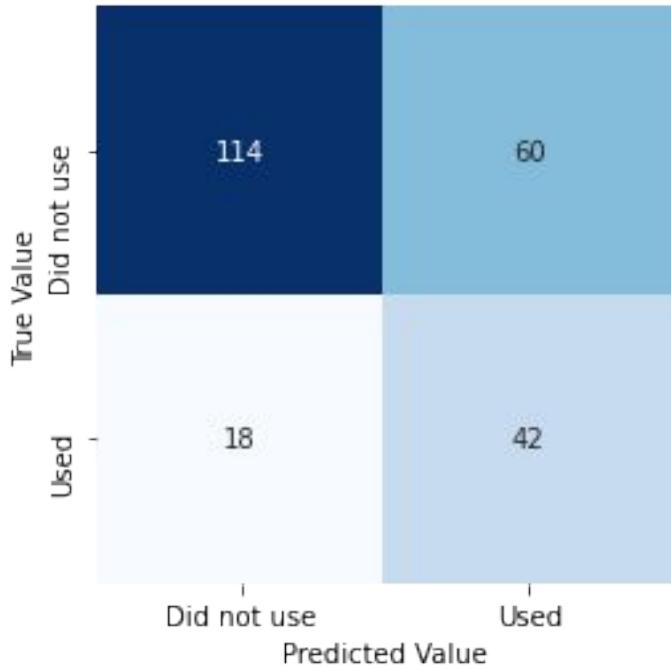
Mean_quantity_per_basket

average quantity of products per basket bought for a household

Correlation Table

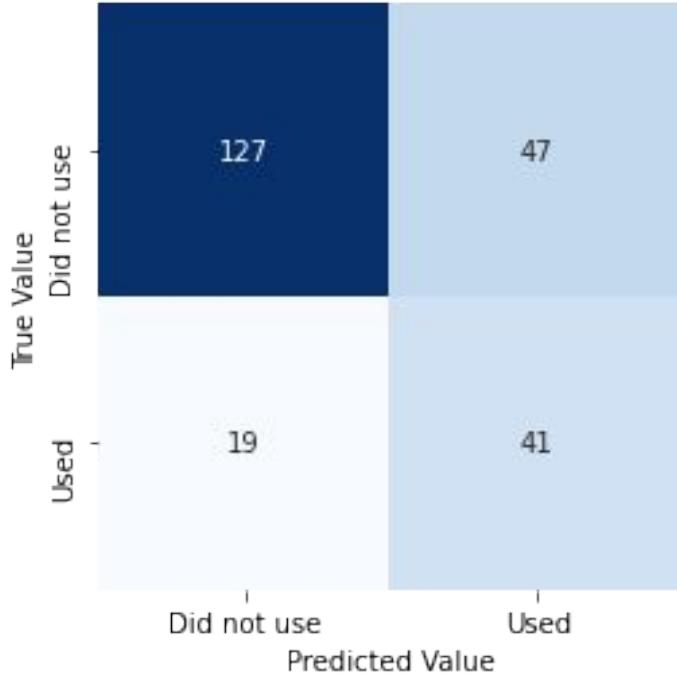
	sales_no_disc	number_of_purchases	unique_stores_visited	mean_quantity_per_basket	coupon_usage
sales_no_disc	1.00	-0.24	-0.14	0.17	0.20
number_of_purchases	-0.24	1.00	0.26	0.14	0.20
unique_stores_visited	-0.14	0.26	1.00	0.09	0.05
mean_quantity_per_basket	0.17	0.14	0.09	1.00	0.22
coupon_usage	0.20	0.20	0.05	0.22	1.00

Logistic Regression using buying behavior is better at predicting coupon usage



- 66.67% accuracy
- Better result in predicting coupon usage than when using demographics characteristics

Does including product category bought improve the model?



- 71.79% accuracy
- 5.12% improvement
- Product category treated as dummy variable
- 296 more variables

Answer: Yes, but the cost in increasing the complexity does not outweigh the benefits

What is our final recommendation?



Personalized Loyalty Program

Due to the new demographic statistics, Kroger should consider creating a loyalty program which provides coupons that are tailored to the consumer's preferences.

Seasonality

- Given the high spending tendencies in certain seasons, Kroger should offer more coupons in certain seasons.
- For more loyal customers there is opportunity to create special **bundles**.

Children

- Families with **kids** tend to spend more.
- Create coupons that **combine** adult necessities with benefits for kids

Income Targeting

- Coupons targeted at households in the 125-200k income range should seek to increase per-trip spending
- Create coupons to increase foot traffic for low and high income households

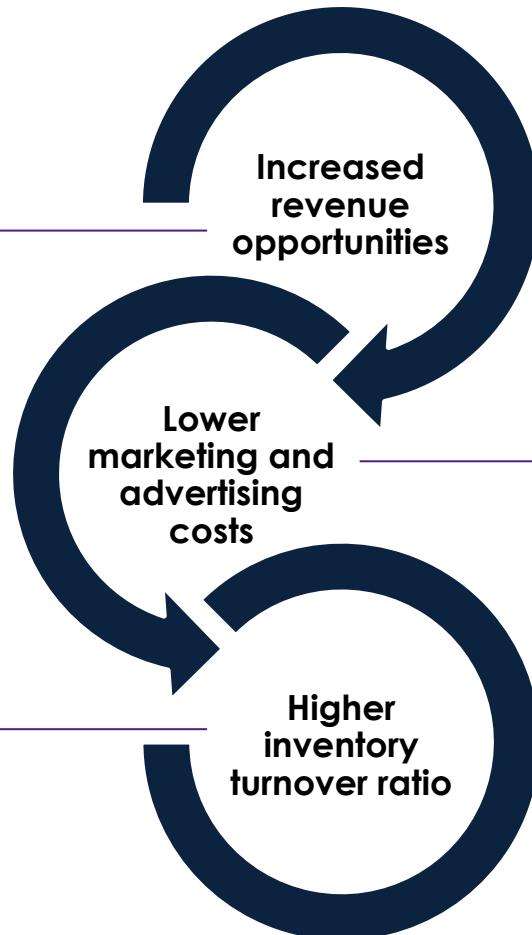
What impact will this have on the company?



Creating specialized coupons directly correlates with increased sales and therefore higher revenue



By selling more product, inventory will be off the shelves quicker, reducing spoilage rates.



By having specific demographics targeted, marketers can focus their efforts on certain audiences instead of broadcasting to everyone.



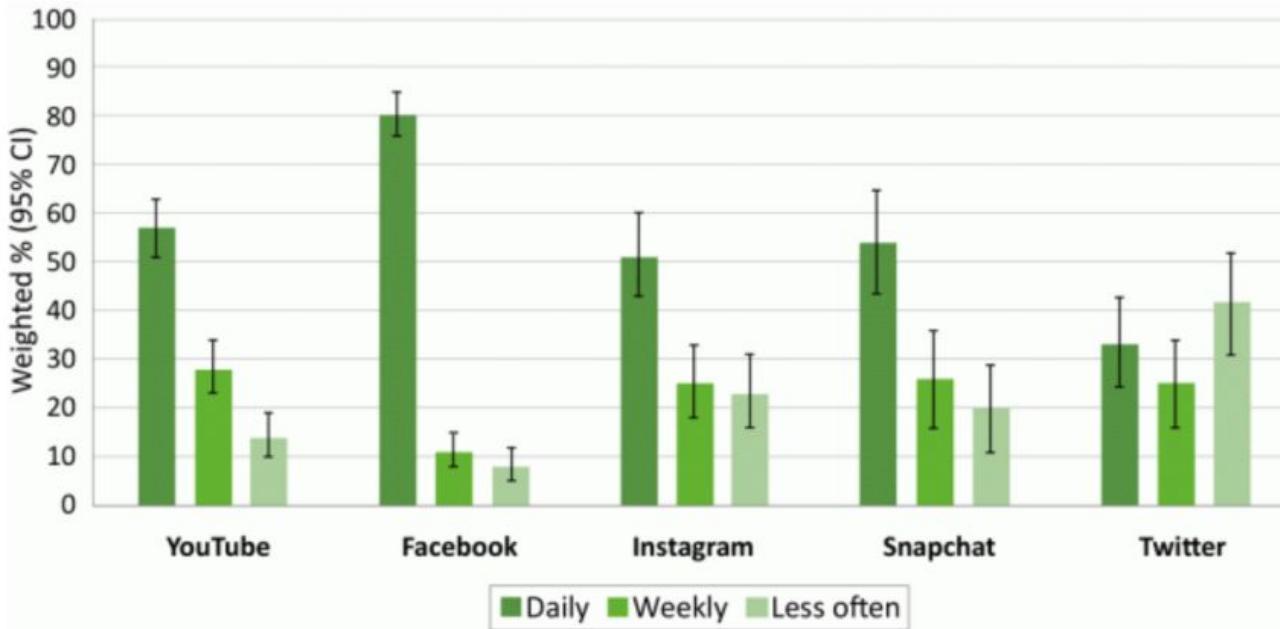
3 Sources of Coupon



Parents	79%	47%	88%
---------	-----	-----	-----

Recommendation: Place coupon promotions on Facebook, Instagram, and YouTube

Coupon Source



- 80% of Parents use Facebook daily
- 55% of Parents use YouTube daily
- 50% of Parents use Instagram daily

Recommendation: Promote daily coupons specials on Facebook



Thank you!
Questions?

