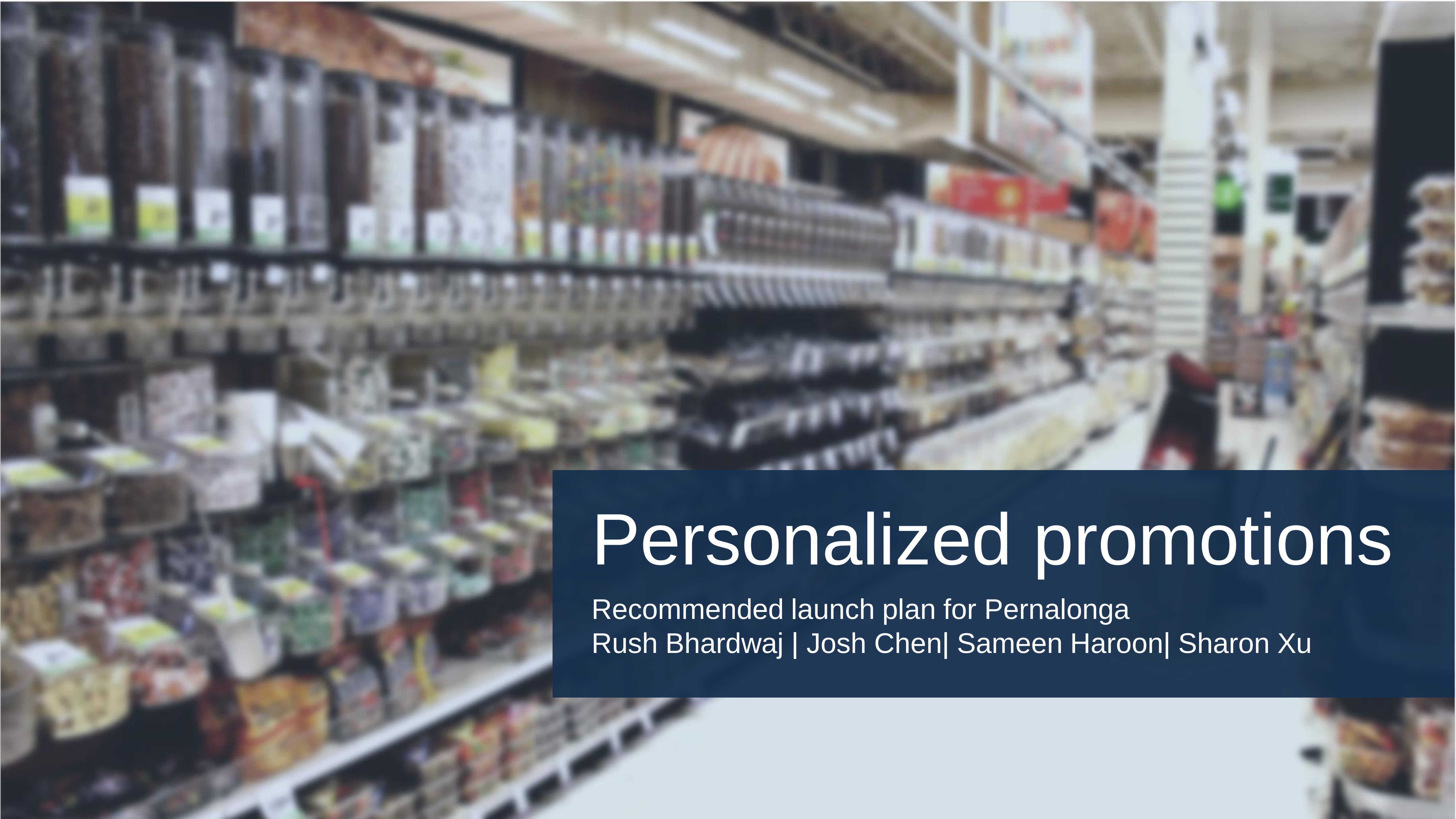


Personalized promotions

Recommended launch plan for Pernalonga

Rush Bhardwaj | Josh Chen| Sameen Haroon| Sharon Xu

Our goal is to help Pernalonga unlock value from personalized promotions

From: Mass promotions

30%
of all sales from promotions

10%
avg. discount across purchases

100%
of customers receive promotion

To: Data-driven personalization

63%
of customers want personalization

78%
feel they get little personalization

86%
say personalization influences purchase

Sources: Rethinking Retail (2013), InfoSys. Based on survey of 1000 customers and 50 retailers across the United States; RILA. (2018, 2 7). *New research explores the changing consumer and the new definition of retail;* Segment (2017), The 2017 State of Personalization Report

To help drive this, we have used data-driven analytics to identify priority customers, products and stores for pilot launch

Right person

299 priority customers

- high likelihood of using promotions
- drive substantial value

Right place

20 stores

- primary stores for target customers
- growth opportunity

Right product

15 product categories

- 2 products per customer
- high margin packaged goods sold in April

To identify target customers, we focused on customers with high probable conversion and high business value

Rules based-approach used to identify target segment…

Find customers with high probability of accepting offer (top 40%)
- Proxied by percent of product purchases with discount

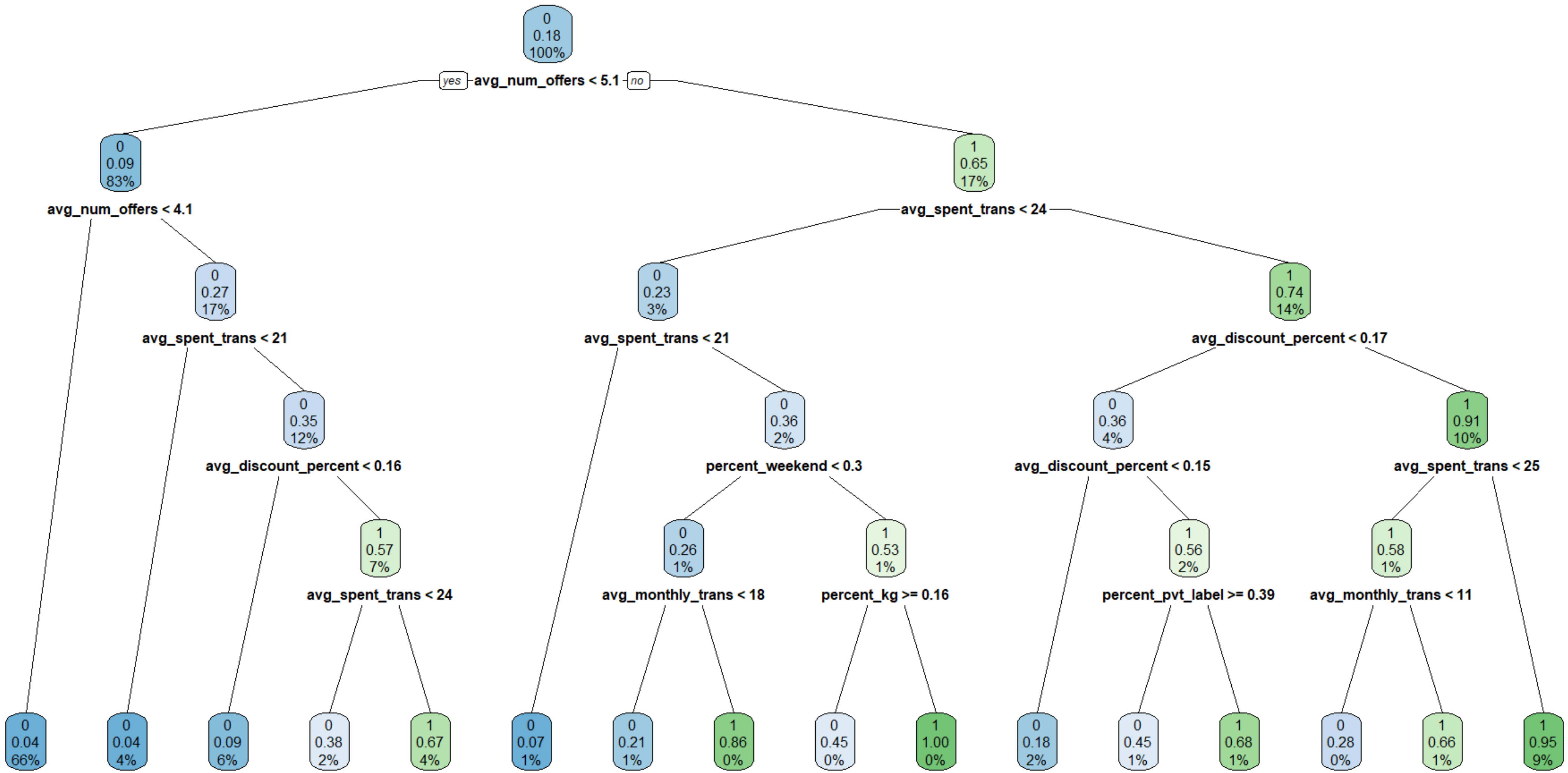
Find customers with high profit generated per transaction (top 50%)
- Relied on product cost estimates

Ensure customer are still active
- Have purchased in the last 30 days

…and evaluated using predictive modeling on relevant features

- Recency of purchase
- Frequency of transactions
- Monetary value in terms of average margin
- Price sensitivity in terms of discount rate and number of offers used
- Loyalty, capturing both length of relationship and number of stores
- Shopping behavior, including tendency to purchase fresh products, private label products, and to shop on weekends

We validated this segmentation using decision-tree modeling to ensure that promotion use and monetary value were consistent differentiators



To estimate customer profitability, we used historical price data and business assumptions to estimate product costs

Assumptions	***Packaged goods***	***Fresh goods***
Classification	Measured in ‘CT’ units	Measured in ‘KG’ units
Cost variation	Cost varies by store	Cost varies by store and by season
Min. margin	5% margin at minimum	0% margin i.e. sold at cost
Cost estimate	95% of effective minimum price by store over 2 year period	Effective minimum price by store by season over 2 year period

Based on our target customers, we further curated target stores and built out recommender systems

Primary stores

- with over 50% of customer purchases
- exclude top 10% of stores in YoY growth

Eligible products

- focus on higher margin packaged goods
- only consider products present in all 20 stores

Final products

- Tested three different recommender systems
- Identified products to promote using best system

In order to produce product recommendations, we tested three main recommender system techniques

User-based collaborative filtering

- Finds users who are similar based on similar rating patterns
- Recommends items that those similar users liked
- “Users who are similar to you also liked…”

Item-based collaborative filtering

- Finds users who liked an item
- Recommends other items that those users or similar users also liked
- “Users who liked this item also liked…”

Singular Value Decomposition (SVD)

- Converts recommendation problem into an optimization problem
- Predicts ratings for items given a user
- Can handle scalability and sparsity better

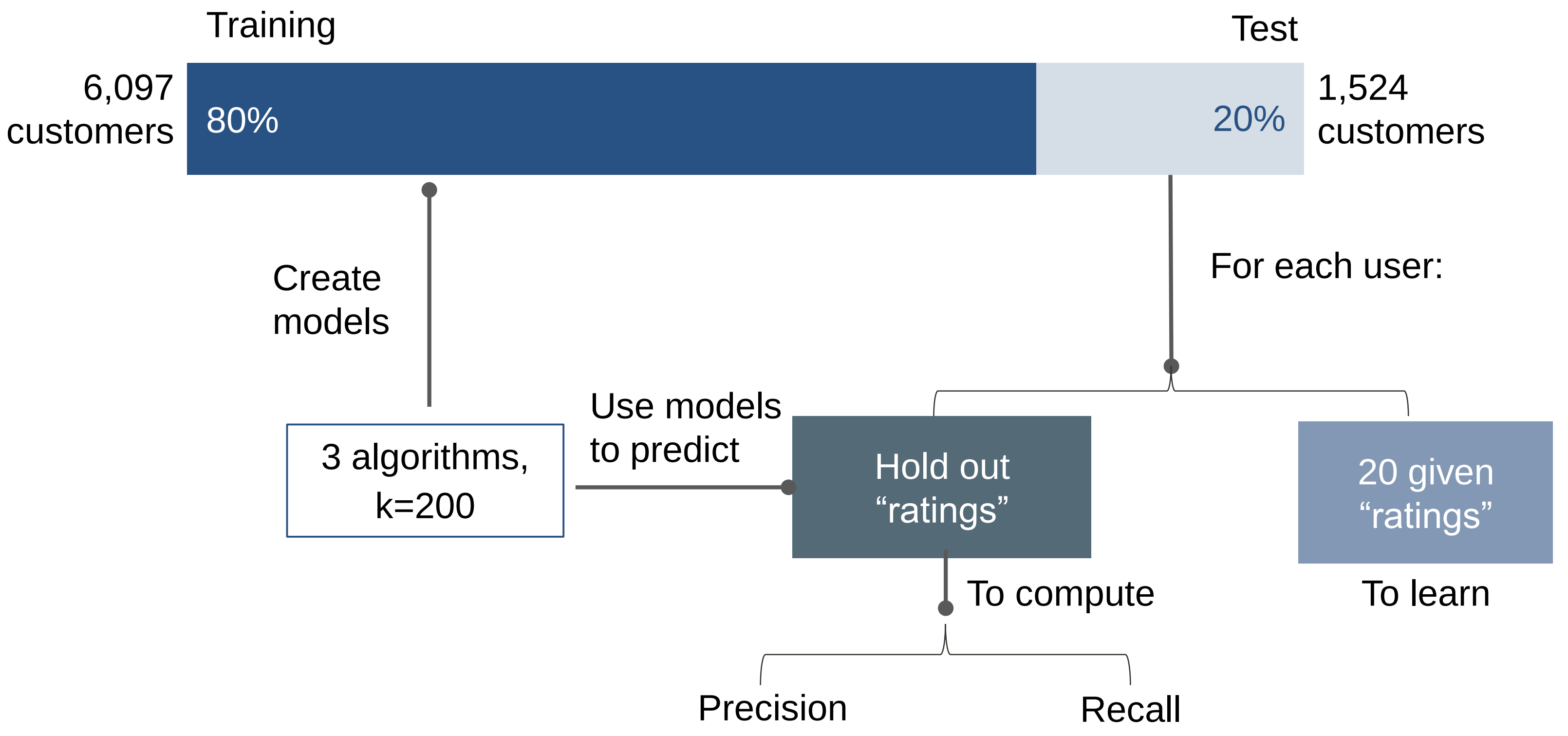
Random item recommender system used as baseline

For evaluation, we used non-target customer data to build recommender systems and evaluated results on precision

Prepare input data

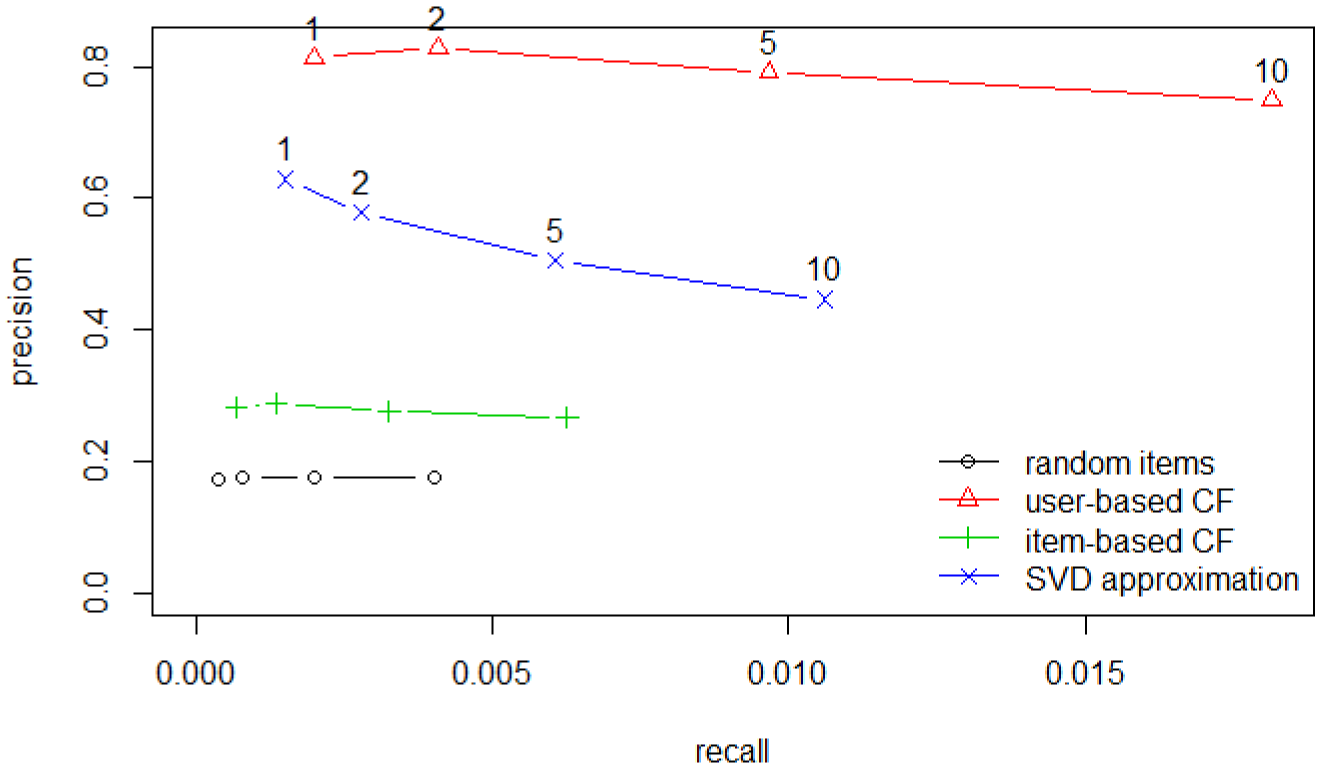
- Use normalized sales quantity as rating
- Group ‘rating’ of 20+ sales
- Obtain matrix of user-product ratings

Build out and evaluate multiple models



Final product recommendations developed using user-based recommendation system associated with best performance

All algorithms perform better than random, user-based collaborative filtering with highest precision



2 products per customer chosen from top 5 using business sense, examples shown below

Product category	Number recommended
Kleenex	178
Mineral waters	110
Canned tuna	87
Oil	37
Canned vegetables	33
Yogurt drink	33
Beer with alcohol	30
Ice tea	26
Canned sausages	15
Yogurt traditional	11

To measure the success this effort, we recommend determining incremental impact through difference-in-difference testing

Proposed approach

Determine test vs. control groups

- Ideal: random assignment of target customers into test and control at store-level, where test receives personalized offer
- Given existing constraints: Targeted customers as test group, control can be customers at 20 stores who aren't targeted

Measure sales using difference-in-difference

- Look at sales on promoted products before/after plan launch in test and control groups
- Compare impact across test & control groups

Example impact estimation

		Test	Control
Group definition		Priority promotion Customers	Customers outside segment visiting the same 20 stores
Total revenue from promoted products	Before promotion	$100	$100
	After promotion	$250	$130
	Difference before/after promotion	$150	$30
	Difference-in-difference	$120	

Recommended next steps

Pressure-test assumptions, especially related to product cost and margin estimates

Determine the right price, i.e. finalize optimal promotion amounts for each customer offer

Identify main marketing channel and adjust estimates of promotion uptake

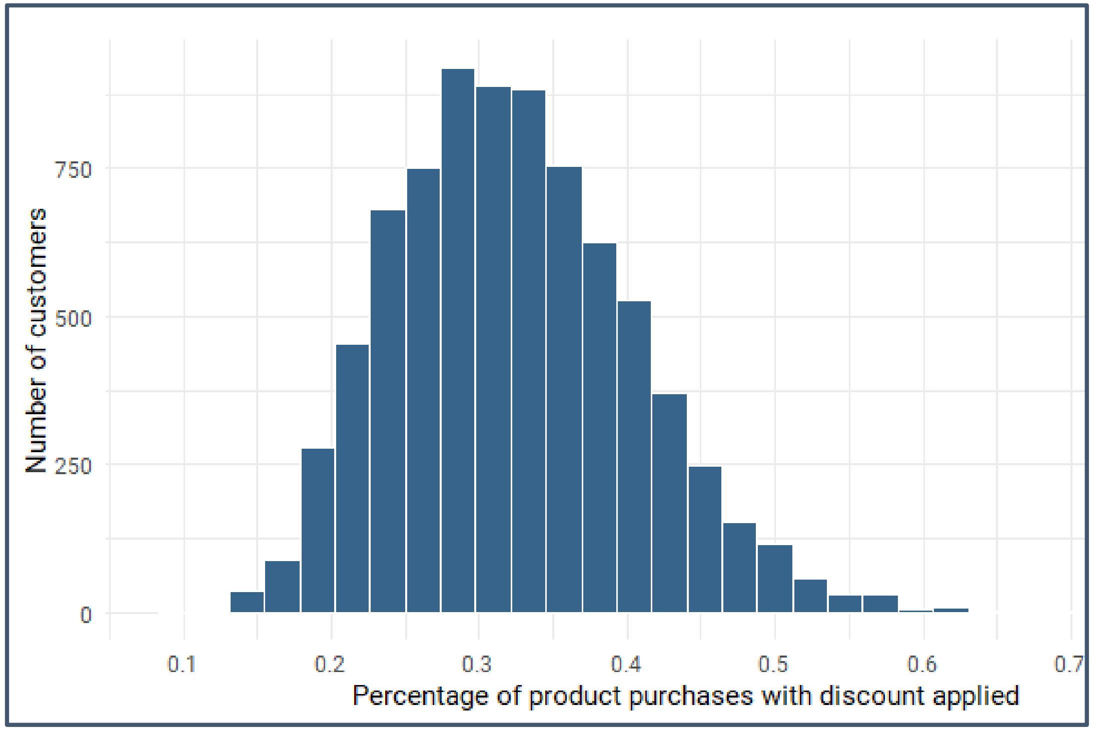
Estimate expected business impact versus baseline promotion from historical patterns

Once impact is proven, scale personalized offers to more customers and stores

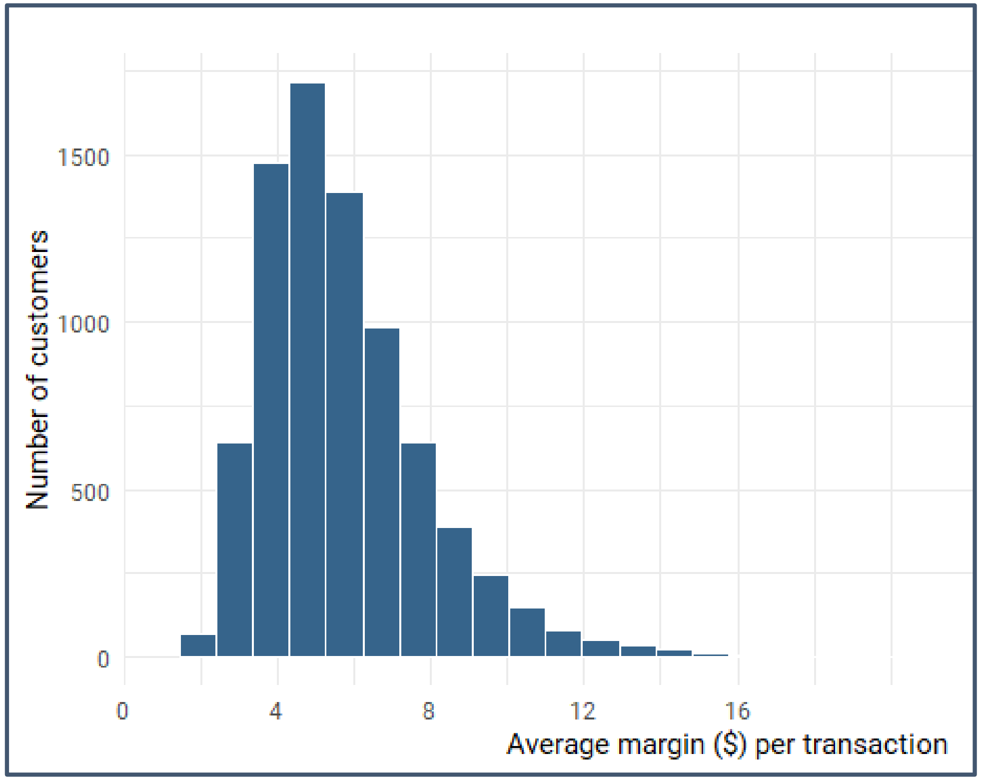
Appendix

Backup: Distribution of primary variables used to build out customer segments

Top 40% customers in terms of discount use…



…and top 50% in terms of profit per transaction



Backup: Distribution of sales quantity used to produce normalized ratings estimates, with quantities at 20+ grouped to avoid long tail