PERSONALIZING OFFERS AT PERNALONGA

PROPOSED MARKETING PLAN

PREPARED BY:

RUSH BHARDWA, JOSH CHEN,
SAMEEN HAROON, SHARON XU

TABLE OF CONTENTS

Introduction

Project Overview

The goal of this project is to help enable personalized promotions at Pernalonga stores in Lunitunia. With over 400 stores and 10,000 products, Pernalonga is heavily dependent on promotions to facilitate sales, and is looking to drive business gains from more customer-specific offers. Numerous studies have shown that personalized promotions can significantly increase the profit of retailers, with 86% of respondents in a retail-based survey indicating that personalization impacts their purchase behavior to some extent[1]. There is also an increasing demand from customers for personalized shopping experience, with over 63% of consumers indicating that they are interested in personalized recommendations.[2] The challenge facing retailers is how best to operationalize the lofty goal of increased personalization.

For Pernalonga, in particular, there is a need to provide the right customers with the correct product recommendations at the right discount level at the right stores. Given the rich transactional data possessed by Pernalonga, we plan to use a data-driven approach to develop personalized promotion strategies to grow Pernalonga's revenue. To validate the business case for such strategies and identify best practices early, the goal of this project is to develop a marketing plan for a pilot personalized promotion campaign

Proposed approach

To help design a personalized marketing test plan for Pernalonga, we will look to find priority target customers for such offers who mainly shop at 20 stores. Customers and stores will be identified using business understanding, such as by finding customers who are likely to accept offers and drive revenue in stores that could benefit from increased traffic. Eligible products for promotion will further be selected based on relevant factors such as promotion timing, inventory availability in target stores, and product type. The top 2 products to promote for each customer in our target segment will be suggested using the best recommender system we found on Pernalonga's data, and we recommend that campaign deployment include test/control measurement to validate the lift in sales from personalized promotion prior to larger-scale implementation.

Methodology

Data exploration and cleaning

Prior to building out a personalized marketing plan, we first explored and cleaned the data we received. Pernalonga provided 5GB of transactional data with information about individual transactions including the store, customer, products purchased, price paid, and discounts applied. Each row in the dataset

[1] Rethinking Retail (2013), InfoSys. Based on survey of 1000 customers and 50 retailers across the United States.
[2] RILA. (2018, 2 7). *New research explores the changing consumer and the new definition of retail.* Retrieved from Retail Industry Leaders Association: https://www.rila.org/news/topnews/Pages/New-Research-Explores-the-Changing-Consumer-and-the-New-Definition-of-Retail.aspx

represents a unique product in the transaction, this will variously be referred to as a 'product purchase' or line item in the remainder of the report.

After initial exploration, it became clear that the transaction id failed to uniquely identify each customer visit, since there were fewer transaction ids than customers in the data. To create a truly unique identifier, we combined the transaction date with the unique customer id and the store id. After generating a new transaction id, we determined that the data contains a total of 2.96M unique transactions with 10,770 unique products purchased by 7920 customers. The data included transactions occurring over 24 months (from 2016-2017) and across 421 store locations.

While no missing values were detected, several other issues were identified and corrected. For instance, three product ids in the transactional data did not correspond to product ids in the product table. Since these products were only linked to 510 out of 2.9 million unique transactions., i.e. made up a very small percentage of the data, we decided to remove these transactions. We also found that there were a few instances (8 rows in data) where the discount amount was higher than the product sale amount (i.e. discount exceeded the maximum amount a customer could owe for a purchase). Since this shouldn't occur in a business setting and represented a handful of observations, we excluded these transactions.

In addition, we noticed that bananas and bags appeared in significantly larger amount of transactions than other products. Within the 2 years, excessively high numbers of bananas and bags were sold, well above average product quantities. Initially, our team believed that this quantity was a mistake in the data, however, as we investigated the data and the business context, we concluded that while these products did follow abnormal patterns, the data was most likely accurate. We believe that the category "bags" probably referred to bags purchased at checkout. In some states and countries, plastic bags must be paid for by law. This would explain the high quantity sold, and since such purchases aren't tied to customer product preferences, we removed purchases for the high-volume bag SKU from the dataset. After some research, we also discovered that in Pernalonga's region, one of the largest exporters of bananas shipped 80% of bananas to the mainland where Pernalonga's stores operate, accounting for the large quantity sold. This appears to represent a region-specific supply and consumption pattern for this product. As such, no changes were made to the data on this front.

Transactions with Cod Fish also gave us some pause because the transactions regularly had very high numbers of discounts applied. While most transactions only had a handful of discounts at most, Cod Fish had up to 76 discounts applied to a single transaction. We considered this as a team and considered that the Cod may be sold by a set amount of weight (such as 100 grams). If this were the case and a customer bought a couple pounds worth of Cod on sale, it may appear as if that customer were applying many discounts. We confirmed this theory by examining the quantity of Cod sold per transaction and discovering that Cod was often sold in quantities of 50-100. We chose to leave the transactions with Cod in the dataset.

When exploring product data, we also noticed that there were 244,503 unique prices in the dataset but only 10,770 unique products. Most of this price variation was found to be associated with produce goods, mapping to up to 3 price changes per store per day on average for certain product ids associated with bananas. In contrast, packaged goods (denoted by goods measured in 'CT" units) had at most 58 unique prices across all stores in the 2-year period. Although the price fluctuations for bananas seemed especially high, we opted to accept this data as given. In addition, we noticed that 10 products lacked English descriptions and 3 products lacked Portuguese descriptions. While this decreased our ability to look at these products from a business perspective, we did not believe the lack of description reduced the integrity of those rows.

Looking at stores, we saw that some branches had not been in operation for many days. In some instances, the stores had opened towards the end of the 24 months and represented newer stores. While we had fewer transactions associated with these stores (i.e. below 500 transactions), we opted to keep them in our dataset since the available transaction data on customers/products could still be useful. However, we noticed two stores where the first and last transaction occurred on the same day. We dropped data for those stores. The final cleaned data produced after this process served as the basis of all subsequent data processing.

Estimating base product costs

Prior to identifying target customers for personalized promotions, knowing which customers and products drive high versus low margins for Pernalonga stores should be considered. In order to approximate this using available data, we sought to get a baseline estimate of the unit cost (i.e. cost of good sold) for each product by leveraging information on effective unit prices. In particular, we assumed that the minimum average unit price a product was sold for during the two-year period in a specific store corresponded to its base unit cost, with some adjustments for different product types.[3]

For products measured in 'KG' units, which presumably correspond to fresh goods, we observed that there was substantial variability in unit prices set week-to-week and across stores. We also wanted to be mindful of the especially high variability in unit price associated with such goods (described earlier in data exploration and exploration), and the fact that costs for fresh produce can vary by whether the product is on- or off-season. As such, for KG products, we assumed the base unit cost for each SKU corresponded to the minimum price it was sold for in each store over the 2-year period, by calendar season.[4] A more refined future estimation could define season specific to the product type in question (i.e. the agricultural season for specifics fruits, vegetables, etc.), but the current approach felt sufficient for our needs. The minimum unit price a KG product was sold for was assumed to be its base unit cost because it feels reasonable to believe that highly perishable goods can be sold at-cost when spoilage is likely. Store-level minimum prices were used to account for potential differences in cost by store.

[3] Pressure-testing this assumption with client feedback will be necessary before the final test plan is deployed, as substantial discrepancies could influence customer/product recommendations and expected impact

[4] A basic mapping of months of the year to calendar seasons was used (i.e. Dec-Feb: Winter, Mar-May: Spring, Jun-Aug: Summer, Sep-Nov: Winter), with the minimum found for the season across both years of data.

For products measured in ‘CT’ units, which should primarily correspond to packaged goods, we simply looked to get the minimum unit price a product was sold for (i.e. the paid amount divided by the quantity sold) in a single transaction across the two year-period. Once again, unique unit costs were calculated by product and store, to capture any potential differences in cost by store. In addition, since it is less likely that packaged goods will be sold at-cost even with substantial discounting, we assumed that the minimum margin associated with each good for Pernalonga would be 5%. As such, 95% of the minimum price sold for each product by store across the two-year period was assumed to be it’s base unit cost.

In generating these calculations, we noticed that there were instances where the effective unit price paid for a product was zero, but it had non-zero discounts applied. We reasoned that this could potentially be due to cross-product promotions, for instance, if a supplier makes multiple products and wants to promote a new product line by free giveaways with a purchase of a more established product. With this possible explanation, and the low observation count for such instances (approximately 3000 rows out of 29M), we chose to keep these transactions in the overall data. However, such cases were removed from the ‘minimum’ price detection to avoid having products being assigned a cost of zero. In effect, there were 200 product SKUs that had only been sold under such circumstances i.e. had been given away for free – these were removed from the data.

Ultimately, base cost estimates for each unit of a product were generated based on the minimum by season and store for KG goods, and 95% of the minimum for the product and store for CT goods. Although costs may change from year to year, since our data includes transactions across 2 years, we chose to find the minimum across the full 2 years.

Identifying target customers

An effective personalized promotion is one that gets a customer into the store shopping, such that they buy more than just the one product associated with an offer. As such, ideal customers for this personalization plan are those who are likely to accept a promotion offer, and whose shopping is tied to high purchase value (in terms of margin) once they are in a store buying products.

To find customers who are likely to accept promotion offers, we looked for customers who had relatively high promotion use. While more sophisticated approaches are possible such as examining product units purchased with/without discounts for customers in comparable sets to identify price sensitivity, the current rule-based approach is both intuitive and reasonably effective based on our analysis. The specific metric used to determine high promotion use is the percentage of product purchases (i.e. line items across all transactions) where a user applied a discount. We recognize that this doesn't factor in the number of units of each product bought or the number of discounts applied, but we believe users with a high proportion of discount use across items will be a good target for promotions. There is also the possibility that some of this may map to customers who purchase at stores where products are frequently discounted. To avoid overfitting on this (presumably more niche) circumstance, we will be looking to get folks in the top 40% percentile for discount redemption.

To further find customers who are associated with higher monetary value to Pernalonga, we restricted the target customer list to those who were above the median i.e. in the top 50th percentile for average margin per transaction. Here, margin was calculated using the difference between product paid amount and base product cost estimations detailed in the previous section. In addition, as a sanity-check to make sure promotions would be offered to relevant customers, we further ensured each customer in this list had purchased once in the last month.

Given our decision to use a rule-based approach for identifying top segments, to pressure-test that the resulting segmentation was meaningful, we prepared additional customer metrics mapping to relevant characteristics from a business perspective. The business characteristic and specific metric are detailed below. Once our target customers using high promotion and margin eligibility are identified, we used decision trees to find the attributes from the list below that best explained the difference between target vs. non-target segments. Our goal is to ensure the rule-based segmentation consistently maps to additional metrics tied to high promotion use/likelihood and high monetary value[5], as well as to look for any other interesting patterns tied to the segmentation.

- *Recency*: the most recent transaction date i.e. ensure customer purchased from us in the last 2 months
- *Frequency*: the average number of transactions they have per month
- *Monetary Value*: the average margin as a percent of amount paid by that customer across all transactions, where margin is tied to product base cost estimations, the average amount they spend per transaction
- *Price sensitivity:* the average discount rate of purchases by that customer (applied across all transactions), the average number of offers applied per transaction
- *Loyalty:* number of stores shopped at, how long they have been a customer, measured by number of unique months in which they bought from us i.e. transactions occurred
- *Shopping behavior*: percentage of their product purchases that are 'kg' goods, percentage of their product purchases that are private label goods, the percentage of customer transactions that occur on a weekend

Identifying top 20 stores

With the target customer segment identified, to find the top 20 stores for promotions, we first sought to find a subset of stores that would benefit from increased traffic. Any approach that looked to find 'underperforming' stores (e.g. by finding stores that had no annual sales growth) was found to overly limit the sample of customers available for testing. As such, we instead focused on excluding top performing stores that would have little to gain from a personalized promotion campaign. Top performing in this case was defined based as stores in the top 10th percentile for Year-over-Year (YoY) sales growth between 2016-2017. For this, only stores with more than a year of sale were included – this only excluded 10 stores which had just one year of sales. We also tried segmenting stores based on high

[5] Once we exclude the actual metrics used for the segmentation

average days between customer transactions, but that substantially reduced our segment size for analysis and hence, was not pursued.

To identify top 20 stores for promotion testing, we looked at stores that were in the subset based on YoY growth (i.e. excluded the top 10%) and were the primary store for customers in our target segment for promotion testing (based on the criterion described in the previous section). For each customer, their primary store was identified as the store that captured at least 50% of their total transactions with Pernalonga over the past 2 years. Hence, the top 20 stores that were in the subset, had the highest volume of target customers and were the primary store for each customer were selected as the ideal stores for targeting.

Identifying eligible products

To ensure the recommendation is effective, we set criteria to select eligible items that best serve our purpose. We removed all products in KG unit because they are commonly perishables such as produce or meat. Perishables are normally discounted to facilitate sales before the produce decays but not for attracting customers or generating higher revenue. In addition, produce goods usually have a relatively low margin compared to packaged goods. Therefore, providing discounts on perishable goods is not the best promotion strategy. In addition, since the plan is to launch these offers in April, we removed products that are off-season in April from potential products to recommend. More specifically, we filtered out products that were never sold in April.

Lastly, to ensure the targeted customers can use their offers, we only consider products that exist in all 20 stores that were most frequently visited by the priority promotion customers. This curation produced a total of 2084 eligible products for promotion.

Obtaining product recommendations

To market target customers with personalized promotions at our 20 stores, we need to find the specific products that they are likely to purchase. Within eligible promotion products, we would like to push offers for products that customers have not purchased before but are relevant to their shopping preferences. To find the best mechanism for suggesting relevant products, we opted to build and evaluate widely-used recommender system algorithms and build a final model based on the best performer for Pernalonga data. From here, the final model will be used to predict the top 2 recommended products for each customer in our target set.

To implement recommender systems that map products to users, additional data processing is needed. For this, we first created a preference profile for all customers using a rating matrix to map each customer to each product (within products that appeared in our eligible set). The "ratings" in the matrix are calculated using normalized total sales quantity across all the transactions for each customer-product pair[6]. As such, only products that were purchased by a customer have a non-zero rating. After experimentation, all sales quantities greater than 20 were grouped together under 20 to help model

[6] Since the product set is limited to 'count' unit goods, no equivalent volume calculations were needed

user preference. The distribution of total sales quantity per customer-product pair after setting 20 as the maximum amount is shown in the figure below.

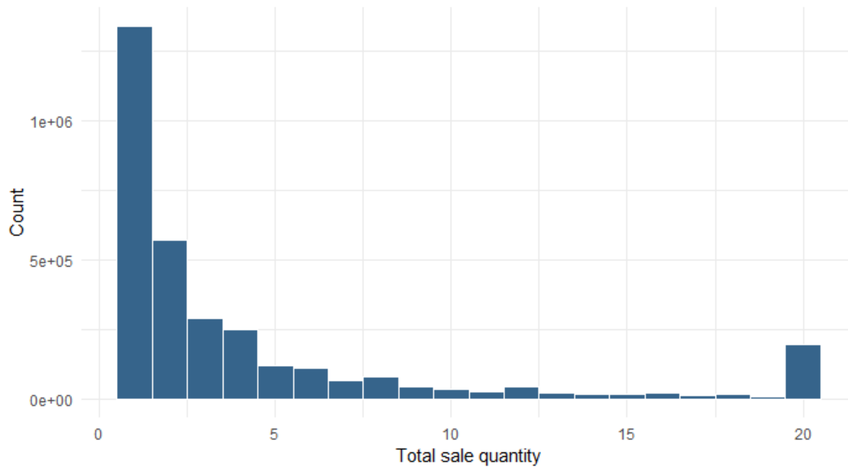

Figure 1: Distribution of total sales quantity for each customer-product paid

Broadly speaking, there are two types of recommender systems: content-based filtering and collaborative filtering. Content-based filtering relies on consumer preferences on product attributes to make recommendations. Since we only have transaction-related information for products, we cannot use such a technique. On the other hand, collaborative filtering techniques look for patterns in user ratings to produce recommendations based on rating similarity across users or items. User-based collaborative filtering (UBCF) creates item recommendations based on consumption patterns of similar users while Item-based collaborative filtering (IBCF) technique uses similarity between items to generate recommendations. Singular Value decomposition (SVD) is less intuitive than UBCF and IBCF, but is a model-based approach that involves matrix factorization to make recommendations, and can produce superior results when the rating matrix is sparse.

To build the most effective recommender system for Pernalonga, we decided to first evaluate the three main algorithms in collaborative filtering mentioned: UBCF and IBCF and SVD. The model creation and evaluation process is shown in Figure 2. In order to ensure separation between model training and evaluation, an 80%-20% split of non-target customers into training and test set is implemented. Cosine similarity measure is used to determine the similarities between customers or products. Once the models are created on the training set, for each customer in the test set, 20 ratings were used to learn the preferences of the customer and the other ratings were held out to predict the top-N recommendations and to compute precision and recall for each algorithm.

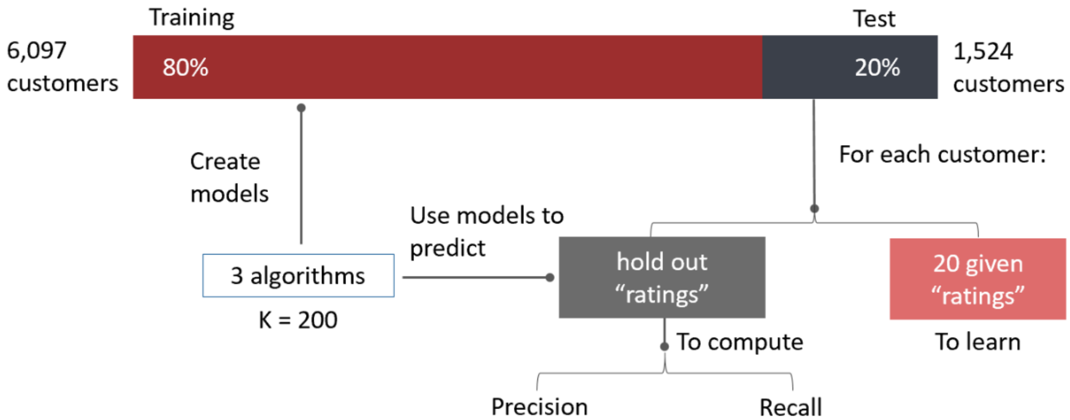

Figure 2: Model buildout and evaluation process for recommender algorithms

In order to make precision i.e. the recommendation accuracy rate, we decided to use the precision-recall curve to evaluate the models. Top-1/ 2/ 5/ 10 recommendations' classification performance results for all 3 algorithms are shown in Figure 3. All the 3 algorithms produced better results than recommending random items, and UBCF clearly outperformed the other 3 algorithms with a precision rate above 0.8 for top 2 recommendations while SVD took the second place.

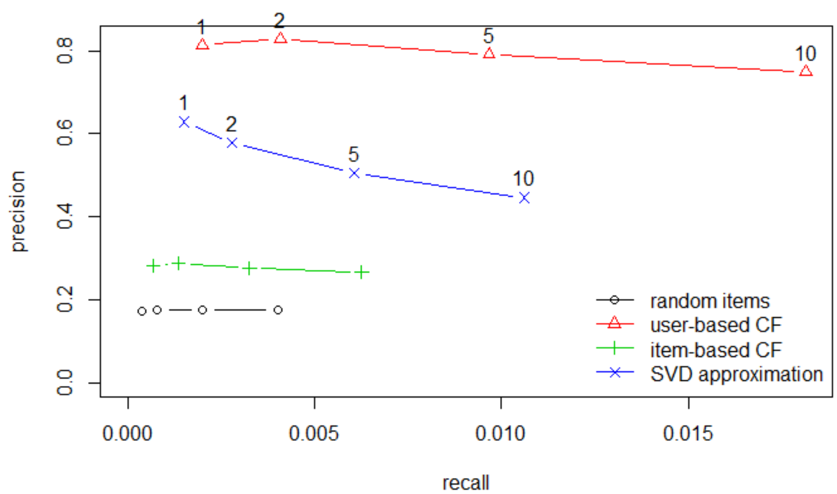

Figure 3: Precision-recall curve showing performance of recommender algorithms

Based on these results, we decided to use UBCF to build the final recommender system, we used the whole matrix to train the model and obtained the top 5 recommended products for each target customer through prediction, which were associated with a precision of approximately 0.80. We opted to use the top 5 recommended products for each target customer instead of the top 2 to further use business logic to screen out product categories that felt less suitable for promotions, such as fresh produce that had not been screened out by the 'ct' product unit criterion mentioned earlier. Based on this approach, we developed the final recommendation of the top 2 products for each customer through a combination of model recommendations and business understanding

Results and recommendations

Target customers, stores and products

From our analysis, we identified a total of 1,453 customers among the overall 7920 customers in our data (i.e. about 18%) with relatively high promotion use and purchase value in terms of margin. Evaluating these segments on additional characteristics as described earlier, we found that they were in line with our segmentation goals, as shown in *figure 4* below. More specifically, the primary factors determining whether we target a customer with a promotion are the average number of discount offers they use per transaction, the average amount they spend per transaction and the typical discount percentage applied across their transactions.

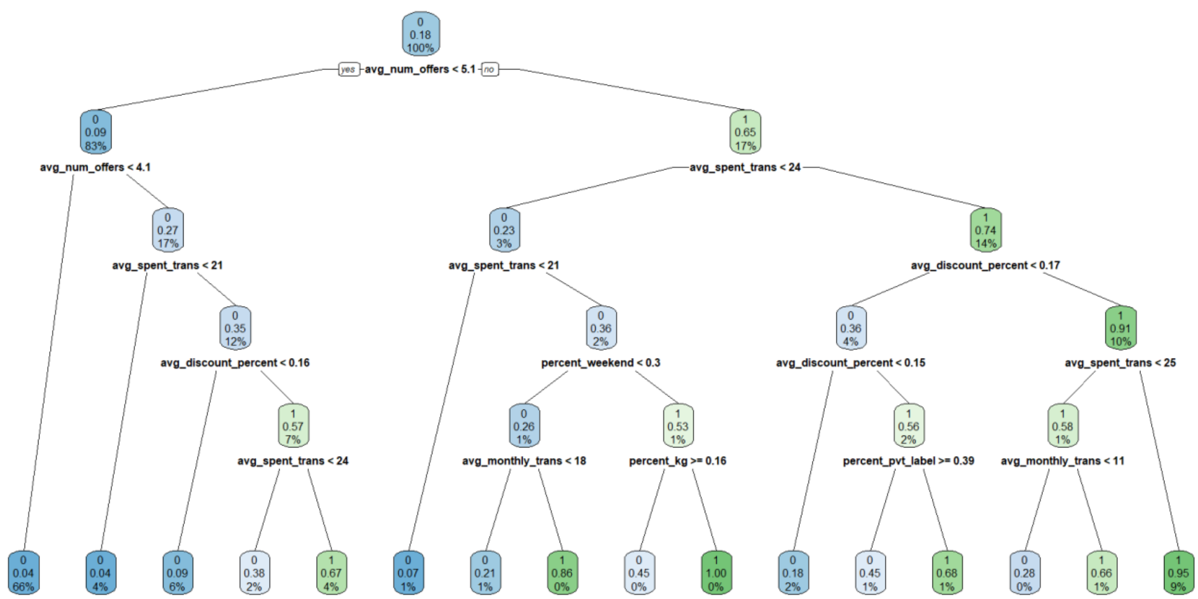

Figure 4: Key attributes of target customers based on decision tree classifier

For the 20 target stores, we evaluated them in terms of total gross profits of all sales transactions for packaged goods. Figure 5 shows the top 50 out of 489 stores in terms of total gross profits: 8 out of 10 top stores are target stores we identified and 17 of the 20 stores are in the top 50. This is in line with our business understanding because these stores are identified as primary stores for our target customers, whose shopping are tied to high purchase value (in terms of margin).

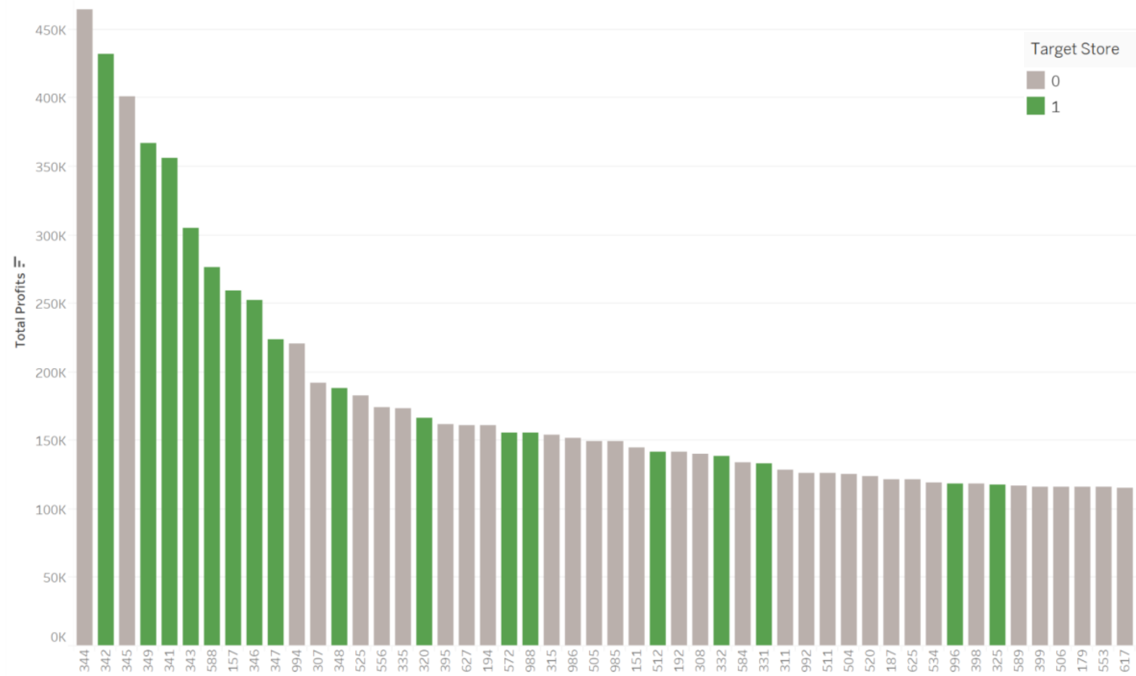

Figure 5: Top 50 stores in total gross profits

For the overall personalized promotion campaign, we found 299 target customers who mainly shop at 20 stores within our criterion. For these customers, recommended products mapped to 15 categories, as detailed in Table 1. Kleenex and mineral waters appear to be the most commonly recommended product categories. It is also worth noting that the current recommendation results include 30 product ids that contain alcohol, and region-specific legislation may limit promotion of these. In such instances, recommender results can be adjusted to exclude products from this category.

Product category	Number of items 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
Salt	6
Rice	5
Paper napkins	4
Standard wines	2
Olive oil	2

Table 1: Overview of product categories mapping to recommended products

In addition, the comparison of recommended products and others' average sales quantity per transaction in is shown in Figure 6. All the recommended products' average sales quantity per transaction are greater than the mean of those for other products. In other words, our recommender system based on customer preferences of products using their purchasing history successfully recommended the products that have relatively high volume per transaction and might give people more incentives to use discount offers for purchasing. Our goal of attracting target customers into the store using personalized promotion is likely to be achieved.

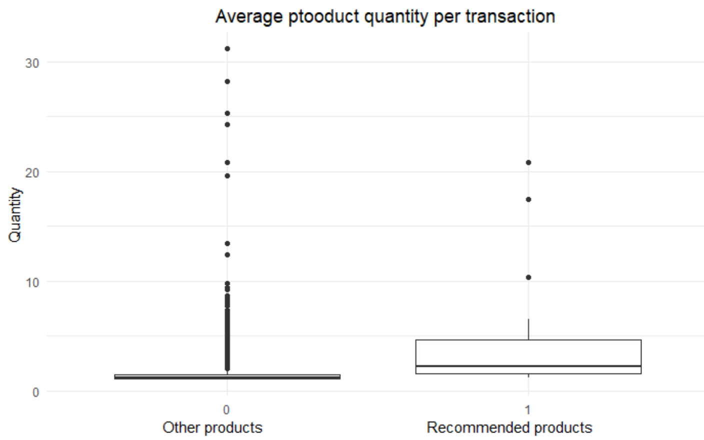

Figure 6: Boxplots for products' average sales quantity per transaction

Proposed deployment plan

For deploying this personalized marketing plan, we recommend conducting an A/B test comparing the spending of customers on the set of promoted products to understand the effectiveness of the targeted promotion plan. From the segmentation of promoted customers, we selected the 299 priority promotion customers to be the test group. For the A/B test to be valid, the control group should act similarly to the test group. Ideally, if our test group had been sufficiently large, we would recommend randomly splitting customers in the target segment in each store, such that the two random splits mapped to test/control groups where one set received promotions and the other didn't. However, given the limited sample size distributed across 20 stores for initial testing, we suggest using a slightly different approach. Specifically, we believe people shopping at the same 20 stores should share basic commonalities with our test group since the available products are the same and they are in the same geography. Hence, customers who are not in the target segment for promotions but have more than 50% of their purchases in our target 20 stores can serve as the control group.

For evaluation, we recommend a method similar to the difference-in-difference approach in statistics. We compare the within-group spending difference before and after the promotion was offered, as illustrated in *table 2*. Specifically, we can calculate the difference of spending on the set of promoted products in these two periods. It allows us to control for the potential change in spending behavior induced by seasonal changes or external events. Then we compare the within-group spending change between our test group and control group. We can then evaluate whether the promotion encouraged the customers to purchase more often, and whether it resulted in higher revenue.

		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	

Table 2: Example approach for determining campaign impact

If the difference-in-difference revenue is positive, we can conclude that the recommendation facilitated more revenue. The actual dollar amount indicates the actual effectiveness of the recommendation. It should be noted that the benefit should be scaled up when the plan is implemented as the number of customers receiving promotions increases, although the scaling will be non-linear if we eventually extend promotion offers to customers who are not in the first priority segment identified in this report.

Potential business impact

To understand the potential business impact of the personalized promotions at Pernalonga, we first sought to quantify the expected volume and amount redeemed based on personalized offers for this specific campaign. The specific approach is outlined in *figure 5*, but the goal was to use the number of times each product will be offered across target customers as well as product-level information such as average quantity purchase and average discount applied across the target 20 stores to estimate potential volume and discount redemption. This information was aggregated across all products. However, since it is unlikely for all offers to be accepted, so we multiplied the potential volume and discounts redeemed by 0.70 – i.e. calibrated our predictions with the assumption that 30% of our recommendations would not be implemented. The 70% offer acceptance rate is a lower estimate based on the precision associated with our recommendations, although additional uncertainty is introduced by whether or not customers will respond well to the specific channels used for offering promotion.

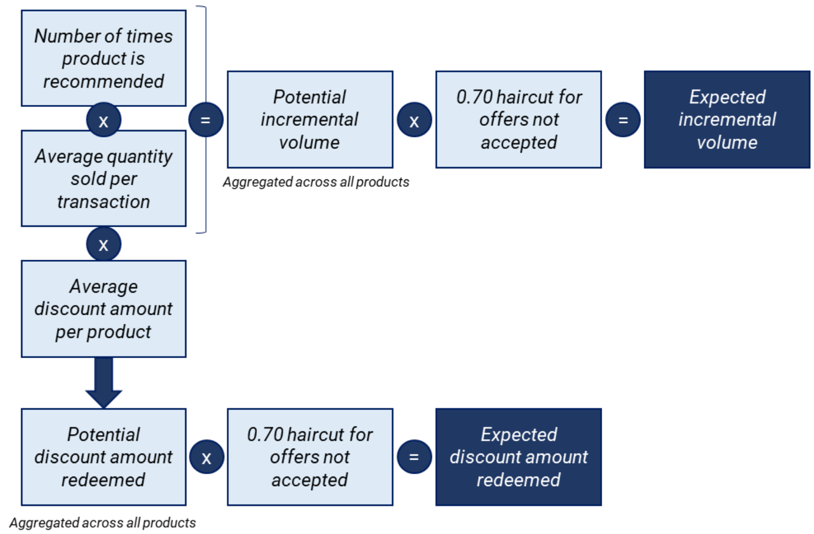

Figure 5: Overview of approach to estimate expected product volume and discounts redeemed

From the above approach, rough estimates suggest that serving 299 customers at 20 stores with 598 offers across 15 categories should lead to about $231.98 in redeemed discounts and 3,445 unit sales. The gross revenue is estimated to increase by $1651.423. It seems reasonable to believe that such volume can be driven within a 2-week window in April given typical store volumes.

Rank	Product Category	Product ID	Incremental Volume
1	Kleenex	999255351	1417
2	Beer and Alcohol	999262736	563
3	Kleenex	999235995	496
4	Mineral Water	999401572	126
5	Yogurt Drink	357541010	115

Table 3: Top recommended products by incremental volume

Conclusion and Next Steps

Based on our findings, personalized promotions can unlock substantial returns for Pernalonga, beginning with the proposed offer campaign in April. Prior to launch, however, pressure-testing some of the assumptions that went into customer, store, product selection (especially around margin and product eligibility) is critical to ensure the campaign is set up for success. Once the final details around targeting and product recommendations are settled, additional iteration will also be needed to finalize the optimal promotion amounts associated with each offer – unless historical averages are deemed acceptable. Moreover, given the variability in offer acceptance tied to different marketing channels (e.g. emails versus in-app notifications versus text messages), recalibrating expectations of recommendation uptake may be necessary once the (primary) channel for communicating offers is identified.

Given the uncertainty associated with quantifying impact prior to launch, the best measure of campaign performance should be based on the attributed uptick in customer visits and purchases tied to promoted products, as suggested in our proposed deployment plan. Upon successful beta testing, the goal will be to scale up promotional offers to additional customers by expanding to more stores, and testing into offers for broader customer segments as well. We look forward to continued work to support Pernalonga in these endeavors.