# **Executive Summary**

In response to rising competition and a saturated retail landscape, ACSE Supermarket is shifting away from traditional mass promotions and exploring personalized strategies to increase customer engagement and promotional return on investment. In partnership with ACSE, our analytics consulting team (Group A6) developed a personalized promotion system explicitly designed for Coca-Cola, a market-leading brand in the carbonated soft drink (CSD) category. The goal is to drive incremental sales by identifying non-Coca-Cola customers with high potential for brand switching, while avoiding cannibalization of ACSE’s private-label offerings.

Our solution is a two-stage, product-level recommender system that balances domain knowledge, machine learning, and practical deployment constraints. The first stage filters customers based on behavioral indicators, churn logic, and private-label exclusions. A logistic regression model, trained on features like spend, purchase frequency, and brand diversity, estimates each customer’s probability of purchasing Coca-Cola. Customers with a predicted probability of 0.5 or higher proceed to Stage 2.

In Stage 2, we assign product-level purchase probabilities using a second logistic regression model trained on features such as brand affinity, global popularity, and customer-product co-occurrence patterns. These product-specific probabilities enable ACSE to recommend Coca-Cola products, such as Diet Coke, Coke Zero, and Fridgemate, based on each customer’s unique behavior and brand alignment. The final model achieved strong recall on held-out data (77% for actual buyers) and enables precise, SKU-level personalization at scale.

Beyond model performance, our EDA revealed several key business insights that informed the design of promotions. These include Coca-Cola’s co-purchase patterns with party staples (e.g., chips, condiments, hamburger buns), seasonal sales spikes during summer and holidays, and churn dynamics among ACSE brand switchers. These findings were used to inform the development of bundle-based offers, timing strategies, and eligibility filters.

Ultimately, our system empowers ACSE to move from static mass discounting to intelligent, customer-level promotional targeting. It is scalable, interpretable, and generalizable across future campaigns. We recommend immediate pilot deployment supported by A/B testing to validate lift and ROI, followed by expansion to other suppliers and categories.

# **Introduction & Background**

ACSE Supermarket is a large-scale North American retailer offering over 100,000 products across more than 100 categories through its network of 40+ stores. As competition intensifies and consumer attention becomes harder to capture, ACSE is rethinking its approach to promotions. Traditionally reliant on broad in-store discounts and weekly flyers, the company is now exploring personalized promotions as a way to drive growth, build customer loyalty, and collaborate more effectively with suppliers.

In this context, ACSE has partnered with our analytics consulting group to develop a recommender system that can power personalized promotions at the customer level. By leveraging transaction history data, the system is expected to identify the right customers to target and the most effective products to promote, ultimately driving a higher return on promotional investments.

As Group A6, our team was assigned the Coca-Cola campaign, which focuses on growing Coca-Cola sales by converting buyers of competing carbonated soft drink (CSD) brands. Importantly, ACSE prohibits us from targeting customers who currently purchase ACSE’s own private-label CSDs, ensuring the campaign does not cannibalize internal sales.

The objective of this project is to:

1. Identify customers who purchase non-Coca-Cola, non-ACSE CSD products.
2. Recommend Coca-Cola products that align with their purchase behavior.
3. Build a robust, scalable, **generalizable** system that supports these personalized recommendations efficiently and accurately.

To meet this objective, we designed and implemented a two-part recommender system. The first component filters customers using domain-informed logic to identify ideal promotion targets based on brand-switching potential. The second component utilizes collaborative filtering and predictive modeling techniques to rank customers based on their likelihood of responding positively to Coca-Cola promotions.

This report outlines our approach, modeling choices, system performance, and business implications. Our solution is designed to be generalizable across future supplier campaigns and to serve as a template for ACSE’s evolving promotional strategy.

# **Data & Sampling Overview**

Our analysis is based entirely on first-party transaction data provided by ACSE Supermarket. This dataset serves as the foundation for building our recommender system, providing granular information on customer purchases, product attributes, and sales metrics across all ACSE locations.

**2.1 Data Sources**

The primary dataset consists of historical transaction records and includes the following key fields:

* **Transaction-level data:** trans\_id, trans\_dt, store\_id, cust\_id
* **Product-level data:** prod\_id, prod\_desc, prod\_category, prod\_subcategory, prod\_mfc\_brand\_cd, prod\_type, prod\_section
* **Sales metrics:** sales\_amt, sales\_qty, sales\_wgt
* **Customer-product interactions:** Derived from aggregating transactions to form customer-level behavioral vectors

This data enables us to reconstruct customer behavior over time and model their brand preferences within the carbonated soft drink (CSD) category.

**2.2 Data Sampling**

To ensure the scalability and manageability of our recommender system pipeline, we implemented a stratified sampling strategy to work with a smaller, yet representative, subset of ACSE's full transaction dataset. This allowed us to efficiently prototype and evaluate our models without compromising on data quality or population representativeness.

We employed a stratified decile-based sampling method, ensuring proportional representation across different customer spending levels. First, we aggregated the total spend per customer across the entire dataset. Customers were then assigned to one of 10 deciles based on total spend, using the NTILE() function. Within each decile, we randomly sampled approximately 5% of customers. This maintains representativeness across the entire spend distribution. After identifying the sampled customers, we retrieved all historical transactions associated with each customer to maintain a complete behavioral context. This method balances computational efficiency with statistical rigor, reducing the data volume while maintaining population diversity. We originally started with 15% of customers within each decile, but due to memory constraints, we had to revise our approach and take a smaller sample.

**2.3 Statistical Validation of Sample Representativeness**

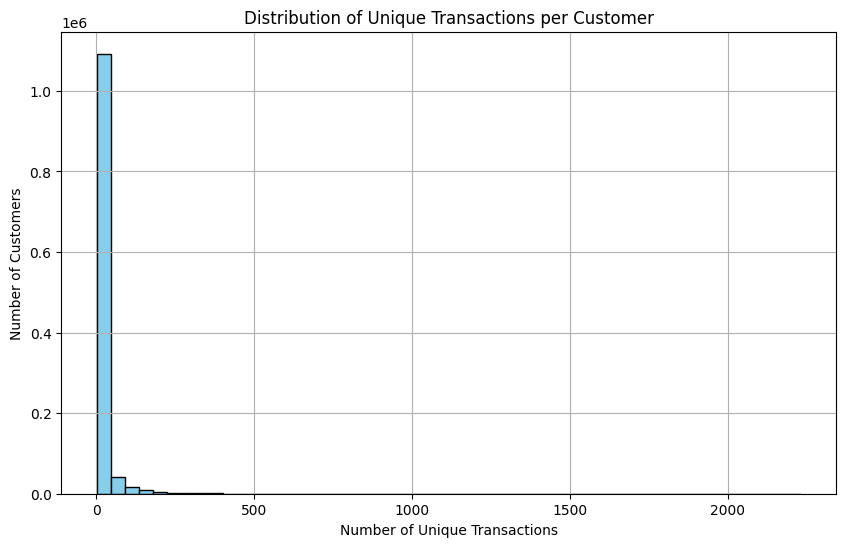
To ensure that our stratified sample is representative of the full dataset, we conducted both Cohen’s d effect size analysis and independent-sample t-tests to compare key metrics between the sample and the population. The two core variables we tested were Sales Amount per Transaction and Sales Quantity per Transaction.

The Cohen’s d results showed a negligible effect size (d) of 0.001 for both Sales Amount and Sales Quantity when comparing our sample to the original dataset. This production of negligible effect sizes for both metrics indicates the minimal difference in the distributions between the sample and population.

For our T-Test, the p-values exceed 0.05 with values of 0.75 and 0.45 for Sales Amount and Sales Quantity, respectively. This further confirms that the sample and population means are statistically indistinguishable, indicating that our sample is representative of the broader customer base.

**2.4 Final Subsetting: Focusing on Targetable Customers**

To further refine the sampled dataset for modeling purposes, we focused on customers who are both recently active and demonstrate meaningful engagement. If we were to proceed without subsetting, we would observe a long-tailed customer distribution (**Figure** **1)**, indicating that our sample contains many customers who have made only one or two transactions. These customers should be excluded from the analysis, as their limited transaction history does not offer meaningful insights into their purchase behavior when building a recommendation system.



**Figure 1:** *Long-Tail Customer Distribution*

To address this, we identified customers who made at least one purchase within the past six months, using the most recent transaction date in the dataset as the reference point. From this group, we retained only those who had made at least five total purchases historically, ensuring we were working with customers whose behavioral patterns are well-defined and informative for modeling. Finally, we extracted all available transactions for this filtered set of customers, preserving their complete purchase history. This process yielded a curated dataset comprising 120,272 customers and approximately 3.5 million transactions, providing a high-quality, campaign-relevant foundation for training and evaluation.

# **Data Cleaning & Exploratory Data Analysis (EDA)**

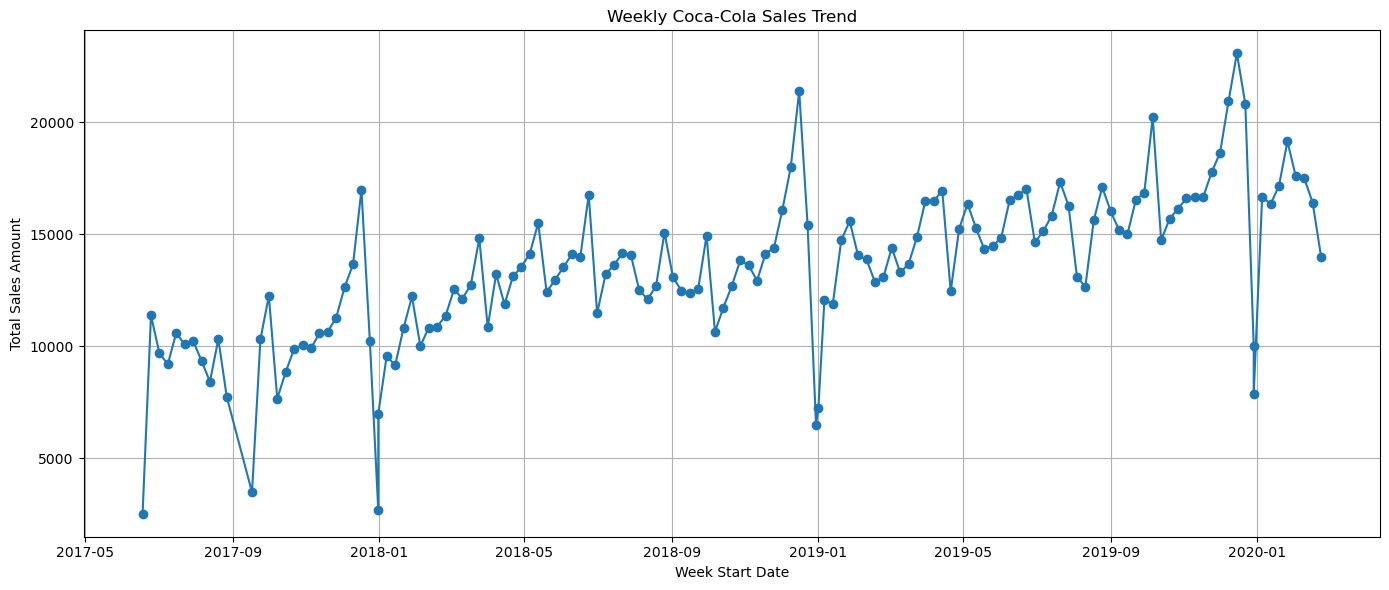
To design a recommender system that not only aligns with the business objectives of the Coca-Cola campaign but also meets operational constraints, we first conducted an in-depth exploratory data analysis (EDA). Our analysis focused on uncovering the shopping behavior of Coca-Cola buyers, their overlap with competing brands, and identifying the most valuable customer segments for personalized promotion. These findings guided key modeling decisions and targeting logic.

**3.1 Coca-Cola Is the Dominant Brand in the Category**

Among the 50 carbonated soft drink (CSD) brands sold at ACSE, Coca-Cola is the clear market leader by a significant margin. However, this dominance means a large portion of customers are already brand loyal—these individuals are not the target of this promotion. Instead, our goal was to uncover opportunities for brand switching, specifically among customers who buy other CSD brands but not Coca-Cola or ACSE's private-label soft drinks.

**3.2 Coca-Cola’s Seasonality**

Analysis of Coca-Cola’s weekly sales trend over the four-year period reveals a clear and consistent seasonal pattern (**Figure 2**), with noticeable sales spikes occurring in the late spring and early summer months, and again during the end-of-year holiday season. These surges are likely driven by warmer weather, increased social gatherings, and holiday-related shopping behavior—occasions where soft drinks are commonly consumed. Correspondingly, there are sharp declines immediately following the holidays and early in Q1, suggesting a post-holiday demand dip.

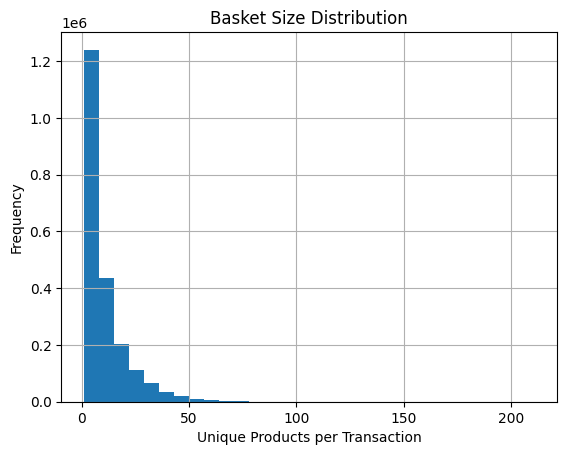


**Figure 2:** *Coca-Cola Seasonality*

Despite short-term volatility, the long-term trend reveals a steady increase in Coca-Cola sales, indicating a growing customer base and brand loyalty over time. These seasonal dynamics highlight optimal windows for personalized promotions, particularly leading into high-demand months, to maximize lift and customer engagement.

**3.2 Basket Patterns Indicate Behavioral Stability and Category Breadth**

Coca-Cola customers exhibit stable transaction behavior with an average of ~70 purchases per customer, indicating frequent shopping and rich transaction history. Additionally, they have an average basket size of ~9.6 items (**Figure 3**), suggesting cross-category engagement and potential for bundled promotions.



**Figure 3:** *Basket Size Distribution for Customers*

Top co-purchased product categories included vegetables, fruit, milk, snacks, eggs, and other grocery staples. While this indicates that Coca-Cola sits alongside household essentials in the basket, these categories are broadly purchased by nearly all customers and offer limited insight into promotional targeting. That is, items like bananas or tomatoes, though frequently co-purchased with Coca-Cola, appear across almost all customer baskets, making them too generic to inform differentiated marketing decisions.

**3.3 Surface-Level Frequency vs. High-Lift, Informative Products**

To move beyond surface-level popularity, we calculated lift scores, which highlight products purchased more often than expected among Coca-Cola buyers. This allowed us to filter out ubiquitous items from the previous section and instead spotlight products that signal more specific affinities.

Some of the high-lift, high-importance products we observed using this method included various other Coca-Cola portfolio products like Glaceau Smartwater, Sprite, Fanta, and Fresca. Additionally, we observed Ruffles Chips, Doritos, Wonder Hamburger Buns, Heinz Ketchup, French’s Mustard, and ACSE Ground Beef, amongst others. These products are often associated with barbecues or parties that often involve having soft drinks.

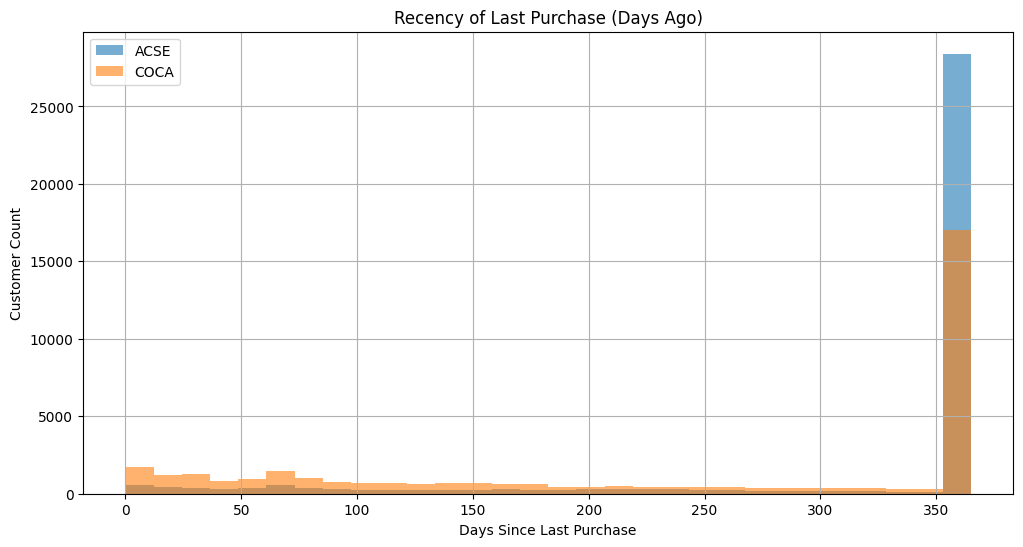
These products are not just frequent; they are contextually meaningful and offer clear implications for cross-promotion or bundling. For example, the frequent pairing of Coca-Cola with grilling staples informed our emphasis on using product lift as a modeling feature and prioritizing these items within our recommender system.

**3.4 Brand Dynamics: Coca-Cola, ACSE, and Other Competitors**

To further our understanding of Coca-Cola in relation to other carbonated soft drinks, we explored the overlap across Coca-Cola, ACSE, and other CSD brands. Out of 122,072 customers, 29,181 purchased Coca-Cola, 14,935 purchased ACSE-branded CSDs, 40,389 purchased non-Coca-Cola, non-ACSE CSD competitors, and 7,634 purchased both Coca-Cola and ACSE - a non-trivial overlap.

Given ACSE’s requirement to exclude its own customers from the campaign, this overlap was critical. We applied strict filters to remove customers who purchased ACSE-branded CSD within the past 90 days and customers who previously purchased Coca-Cola but appear to have churned to ACSE within the past 3 months (~2,000 customers).

We decided on this exclusion time frame based on our recency analysis. Our recency analysis allowed us to quantify how recently customers purchased Coca-Cola or ACSE-Branded CSDs, allowing us to operationalize churn logic and define exclusion windows for promotional targeting.



**Figure 4:** *Number of Days Since Last Purchase for Coca-Cola and ACSE Customers*

As shown in the figure, most customers made their last Coca-Cola or ACSE purchase nearly one year ago, suggesting a long tail of dormant or infrequent buyers. However, there's a noticeable concentration of recent purchases within the past 90 days, particularly for ACSE. This supports our decision to exclude any customers with an ACSE purchase in the last 90 days, as they are either current loyalists or recent switchers away from Coca-Cola, making them unsuitable for a Coca-Cola promotion.

Similarly, the flat distribution of Coca-Cola recency within the first 180 days, followed by a spike at the 365-day mark, provides a clear rationale for labeling customers as "churned from Coca-Cola" if they haven’t purchased the brand in over 180 days. When paired with a recent ACSE purchase (within 90 days), this defines our Coke to ACSE switcher segment, which we also exclude from targeting.

# **Methodology**

Our recommender system leverages a hybrid approach that integrates multiple machine learning and rule-based methods to generate personalized product recommendations. Each method contributes unique strengths—some excel in discovering latent patterns across customer behavior, while others provide interpretable probability estimates or surface frequently co-purchased items. Below, we explain the four key methods we evaluated and/or incorporated: k-Nearest Neighbors (k-NN), Logistic Regression, Naive Bayes, and Association Rule Mining.

**4. 1 k-Nearest Neighbors (k-NN)**

k-Nearest Neighbors (k-NN) is a collaborative filtering method that identifies customers with similar behavioral profiles based on their transaction histories. In a recommendation context, each customer is represented as a vector of brand or product-level purchases. The algorithm calculates similarity, commonly using cosine similarity, between customers and identifies the "nearest neighbors" who have made similar purchase decisions. Recommendations are generated by finding products that those neighbors have purchased but the target customer has not, under the assumption that similar users will have similar preferences.

**4.2. Logistic Regression**

Logistic Regression is a supervised classification model that estimates the probability of a binary outcome, such as whether a customer will purchase a specific product. In a recommender system, each customer-product pair is treated as a training example, with engineered features such as customer spending habits, product popularity metrics, and recency or churn indicators. Logistic regression then learns the relationship between these features and the observed purchase behavior. The output is a probability that can be thresholded to generate a yes/no recommendation or used to rank potential product suggestions based on their likelihood of purchase.

**4.3 Naive Bayes**

Naive Bayes is another probabilistic classifier, based on Bayes’ Theorem, that assumes independence between features. Despite its simplicity, Naive Bayes is effective in high-dimensional and sparse data environments, making it a useful approach in certain recommendation scenarios. It calculates the probability of a customer purchasing a product based on the prior probability of product purchases and the conditional probability of each feature given the product was purchased. This approach is fast and interpretable, but the independence assumption can limit its accuracy in more complex behavioral settings.

**4.4 Association Rule**

Association Rule Mining is a rule-based approach used to uncover patterns in transaction data. Algorithms like Apriori identify frequent itemsets and generate rules of the form “if product A is purchased, then product B is likely to be purchased.” These rules are evaluated using support, confidence, and lift. In recommendation contexts, this method is beneficial for uncovering co-purchase relationships. While support indicates how common an item pair is, lift is particularly useful because it measures how much more likely two items are to be purchased together than if they were purchased independently. This helps distinguish between generic, widely-purchased items (e.g., milk or bananas) and truly informative product associations (e.g., Coca-Cola and chips or ketchup).

Each of these methods represents a different philosophy: collaborative similarity, supervised learning, probabilistic inference, and pattern mining. Evaluating their mechanics helped inform our modeling choices and refine the logic behind our final recommendation system.

# **Model Development & Evaluation**

To guide model development, we adopted a top-down, two-stage approach designed for both computational efficiency and practical deployment. Our primary goal was first to identify which customers were most likely to purchase Coca-Cola products, using behavioral and transactional features. Once we established this high-propensity customer set, we then explored their historical purchasing patterns, beginning at the brand level to understand which brands most frequently co-occurred with Coca-Cola, and eventually narrowing the analysis down to product-level affinities. To validate this framework, we began by building and testing models that focused exclusively on the first-stage classification task—predicting whether a given customer should be targeted or not. Only after this classification step proved reliable did we proceed to generate specific customer-product recommendations based on the output of the top-tier model.

**5.1 Coca-Cola Purchase Probability**

**5.1.1 Naive Bayes + k-NN**

We first tested a Naive Bayes classifier, which models the probability that a customer would purchase Coca-Cola based on their features, assuming feature independence. The model was fast to train and interpretable, making it attractive for prototyping. However, performance was limited. While it achieved decent accuracy (0.79), it struggled with recall for Class 1 (Coke buyers), achieving only 40%. This indicates that while the model reliably identified customers unlikely to purchase (Class 0), it missed a significant portion of potential Coca-Cola buyers. This imbalance—reflected in the macro F1 score of 0.68—suggests that Naive Bayes may be overly simplistic for this setting, where interactions between features are likely important. Its practical use may be more appropriate for early filtering rather than final decision-making.

**5.1.2 XGBoost**

Next, we implemented a gradient-boosted decision tree model (XGBoost), which aggregates multiple weak learners to minimize classification error. XGBoost achieved a similar performance to Logistic Regression across all metrics, with an overall accuracy of 0.83, Class 1 recall of 0.56, and a Class 1 F1 score of 0.62. What sets XGBoost apart from the others is its ability to capture complex nonlinear relationships and interactions between features. It also handled class imbalance more effectively, likely due to its internal regularization and ensemble structure. While the model is more computationally intensive and less interpretable than Logistic Regression, it represents a strong candidate for maximizing predictive performance, particularly in settings where accuracy and conversion are prioritized over transparency.

**5.1.3 Logistic Regression + k-NN**

Next, we trained a Logistic Regression model with L2 regularization and hyperparameter tuning via GridSearch. This model estimates the probability of a customer purchasing Coca-Cola using a linear combination of input features. It performed similarly to XGBoost in all metrics and outperformed Naive Bayes in all metrics as well, including accuracy (0.83), Class 1 precision (0.72), recall (0.55), and F1 score (0.63). Logistic Regression offers a good balance of interpretability and performance, with clearly understandable coefficients and relatively fast training time. It also accommodates correlated features better than Naive Bayes, which improved its ability to identify potential buyers. While not as powerful as tree-based models, Logistic Regression is a solid candidate for production environments where model transparency and calibration are important.

**5.1.3 Association Rule**

We also tried association rule, which leverages patterns in transaction data to make recommendations. The accuracy of the model stood at 0.34. The system achieved a high recall of 0.85, indicating that it successfully captured the majority of actual Coca-Cola-related purchase opportunities. However, the precision was extremely low at 0.04, meaning that most of the recommendations did not result in actual Coca-Cola purchases. The F1 score, balancing precision and recall, was also low at 0.08, highlighting the trade-off and overall weakness in recommendation quality. This suggests that while the model is effective at not missing relevant cases (high recall), it lacks the precision needed for practical, targeted recommendations. Additionally, the computation of association rules is time-consuming, especially when handling large-scale transaction data and extensive item combinations.

**5.2 Final Recommender System (Logistic Regression + k-NN)**

To further refine our personalized promotion strategy, we implemented a two-stage product-level recommender system tailored for the Coca-Cola brand. While earlier models focused on brand-level predictions, we will use our knowledge of their ability to predict who is most likely to purchase a Coca-Cola product to then create an advanced method that predicts the likelihood of a customer purchasing specific Coca-Cola products. This final system not only identifies which customers are most likely to engage with the Coca-Cola brand but also assigns product-specific purchase probabilities to each customer, allowing us to personalize promotional campaigns down to the SKU level. This top-down framework was chosen to ensure both computational efficiency and practical business applicability.

**Stage 1: Brand-Level Filtering**

The first stage begins by estimating each customer’s likelihood of purchasing a Coca-Cola product using behavioral and transactional features. These include total spend, purchase frequency, number of brands/categories purchased, and customer similarity scores derived from k-nearest neighbors (k-NN) on brand purchase profiles. To exclude ineligible customers, we apply churn and private-label exclusion rules, removing customers who have recently switched to private-label competitors (e.g., ACSE) or whose last Coca-Cola purchase was significantly older than their most recent private-label purchase. A logistic regression model with class-weight balancing and grid search optimization is trained to classify whether a customer is a viable target for Coca-Cola.

**Stage 2: Product-Level Recommendations**

For each eligible customer identified in Stage 1 (customers with a predicted probability of ≥ 0.5), we construct a personalized product-level recommendation set using a combination of behavioral features and product-specific statistics:

1. Brand Affinity: The proportion of Coca-Cola customers who purchased a given product, measuring how representative the product is of typical brand behavior.
2. Global Popularity: The proportion of all customers who purchased the product, indicating its market reach and saturation.
3. Matched Brands: The number of non-Coca-Cola brands a customer has purchased that frequently co-occur with specific Coca-Cola SKUs.
4. Additional Features: Customer-level metrics such as purchase frequency, category diversity, and total spend.

Using these features, we trained a logistic regression classifier to compute a purchase probability for each customer-product pair, effectively quantifying the likelihood that a specific customer would buy a specific Coca-Cola SKU. This is a major improvement over simpler rule-based systems or naive brand-level models, as it produces SKU-specific probabilities that adapt to each customer’s profile.

Evaluation of this product-level model yielded the following performance on held-out data:

Accuracy: 0.80

Recall (Class 1 – actual product buyers): 0.77

F1 Score (Class 1): 0.06

Despite low precision due to the high volume of non-buyers, the model successfully captured the majority of true product buyers and produced differentiated probability scores across SKUs and customers, improving over simpler rule-based methods.

**Example Output:**

| **Customer ID** | **Product ID** | **Product Description** | **Recommended** | **Purchase Probability** | **Brand Affinity** | **Global Popularity** |
| --- | --- | --- | --- | --- | --- | --- |
| 1000003274 | 20318694002 | COCA-COLA FRIDGEMATE | 1 | 0.8696 | 0.1754 | 0.0031 |
| 1000003274 | 20318694001 | Coca-Cola Zero | 1 | 0.6186 | 0.1054 | 0.0019 |
| 1000003274 | 20318694003 | Coca-Cola Diet | 1 | 0.5694 | 0.0953 | 0.0017 |
| 1000003274 | 20316026002 | COCA-COLA | 0 | 0.4825 | 0.0780 | 0.0014 |

\*Only products with a predicted purchase probability above 50% were marked as recommended (indicated by a 1 in the “Recommended” column).

We selected the logistic regression + k-NN framework because it offered a flexible, modular, and interpretable pipeline that generalizes well across brands. Its probabilistic output makes it easy to apply threshold-based decision-making, while the use of purchase behavior, churn logic, and similarity metrics ensures tailored recommendations. The system is deployable at scale, easily updatable, and fully data-driven—capable of adapting to new products, brands, and customer behaviors as they emerge.

**5.3 Generalizability of Chosen Model**

Our final recommender system is designed to be broadly generalizable across brands, product categories, and retail contexts. Its flexibility begins with how the target brand is specified- rather than hardcoding product identifiers, the system accepts a brand name as an input parameter and dynamically filters transactions related to that brand. This means the system can be seamlessly reused for different manufacturers (Coca-Cola, Pepsi, etc.), across various product categories (e.g., snacks, beverages, or cleaning supplies), and in any environment where customer transactions with product information is available.

At the product level, recommendations are generated without reliance on static product lists. Instead, the model computes product-specific metrics such as brand affinity (the proportion of brand customers who purchased the product) and global popularity (the proportion of all customers who purchased the product). These metrics are automatically recalculated for each brand, ensuring the system remains future-proof for newly launched products without requiring manual updates.

The system utilizes a comprehensive set of features derived directly from transaction-level data. These include customer-level behavioral metrics, such as the number of unique brands purchased, total spend, transaction frequency, and similarity scores that measure how similar a customer is to others in terms of brand preferences. Since these features are not specific to any particular product or retailer, they can be applied universally as long as the input data adheres to a standardized schema.

To support strategic targeting decisions, the recommender includes configurable churn and exclusion logic. Customers who have recently bought from a private-label competitor (ACSE) or who appear to have churned from the target brand are automatically excluded from recommendation eligibility. These controls are managed through parameters, providing marketers with complete flexibility over who is considered a viable target.

The system is also modular in terms of modeling. It supports both logistic regression and k-nearest neighbors (KNN) classifiers out of the box and could be easily extended to incorporate more advanced models such as XGBoost or LightGBM. This plug-and-play architecture allows for rapid experimentation and tuning, making the pipeline suitable for both production and prototyping environments.

Finally, the recommender does not depend on business-specific rules or labels. The labeling logic, which determines whether a customer purchased a specific product, is built entirely from historical transaction patterns. This makes the system fully data-driven and portable across domains, as it learns directly from observed customer behavior rather than relying on heuristics.

# **Recommendations**

To support Coca-Cola’s goal of increasing market share through personalized promotions, we recommend deploying the two-stage product-level recommender system developed in this project. This system is uniquely suited to identify high-propensity customers and match them with Coca-Cola products they are most likely to purchase. Based on model performance, business constraints, and EDA insights, we propose the following strategic recommendations for implementation.

**6.1 Use Predicted Probabilities to Power Personalized Promotions**

This system generates product-specific purchase probabilities for each eligible customer, enabling ACSE to personalize offers at the product level rather than deploying one-size-fits-all brand discounts. These probabilities reflect a combination of customer behavior (spend, frequency, etc.), product affinity (brand affinity and global popularity), and customer-product co-occurrence history. We recommend using these outputs to generate targeted digital coupons or app-based offers for Coca-Cola products with a predicted probability of 0.5 or higher. This approach ensures that promotions are directed only to customers with a meaningful likelihood of conversion, minimizing waste and maximizing ROI.

**6.2 Prioritize High-Affinity Coca-Cola Products in Promotions**

Our system identified products like Coca-Cola Fridgemate, Coke Zero, and Diet Coke as consistently high-affinity products with strong predictive lift. These products not only demonstrate high brand alignment but are also widely purchased across various customer segments. We recommend focusing promotional spend on these SKUs to drive immediate lift while maintaining brand equity.

In addition, these products should be prioritized in “Buy X, Get Y” or “$1 Off” personalized promotions within the ACSE app or loyalty program. Such precision ensures customers are nudged toward products they are both likely and able to purchase.

**6.3 Bundle Coca-Cola Offers with Thematically Relevant Items**

Our lift and co-purchase analysis revealed that Coca-Cola products are frequently purchased alongside chips (especially Ruffles and Doritos), hamburger buns, condiments like Heinz Ketchup and French’s Mustard, and other grilling essentials. These items collectively represent typical components of social events such as barbecues, picnics, and game-day gatherings—occasions where Coca-Cola is not only consumed but also strongly associated with its brand identity. The consistency of these pairings across customer baskets suggests a clear behavioral signal that can be strategically leveraged in promotion design.

We recommend developing bundle-based personalized promotions that capitalize on these naturally co-occurring items. Example offers could include:

* “Buy Coca-Cola and Ruffles, Get $1 Off”
* “Bundle Coke with Heinz Ketchup for Summer Savings”
* “Party Starter Packs” customized by each customer’s past basket behavior

These bundles serve a dual purpose: they are likely to resonate with customer needs by providing convenience and value, and they also encourage larger basket sizes by promoting multiple high-margin items together. Additionally, because they are rooted in historical co-purchase behavior, these promotions feel contextually relevant, making them more compelling than generic product discounts. This strategy aligns well with ACSE’s goal of increasing transaction value while strengthening Coca-Cola’s brand presence during key consumption moments.

**6.4 Time Promotions to Seasonal Peaks - Where Applicable**

While Coca-Cola demonstrated steady year-round demand, our EDA revealed sales spikes during the spring, summer, and holiday seasons. Personalized promotions for high-affinity SKUs can be strategically timed to these peaks to maximize incremental lift. Even though Coca-Cola itself isn’t highly seasonal, its co-purchased items are, making time-aligned promotions more powerful when bundled. We recommend automating seasonal triggers into the recommendation engine to adjust SKU prioritization as holidays or warm-weather months approach.

**6.5 Enforce Customer Eligibility Filters in All Personalized Offers**

To ensure promotional precision and safeguard ACSE’s brand and revenue objectives, we recommend strict enforcement of customer eligibility filters when deploying personalized Coca-Cola offers. Specifically, customers who have purchased ACSE-branded carbonated soft drinks (CSDs) within the past 90 days should be excluded from promotional targeting. Additionally, customers who have churned from Coca-Cola to ACSE, defined as those whose most recent purchase was from ACSE but not from Coca-Cola within the past 180 days, should also be excluded from the campaign. These filters help prevent cannibalization of ACSE’s private label and ensure that Coca-Cola promotions are directed only toward high-potential customers with true brand-switching potential. This logic has already been operationalized within our model pipeline and should be preserved during implementation.

**6.6 A/B Test Personalization Against Standard Promotions**

Before rolling out the recommender system at scale, we advise running a structured A/B test to evaluate the lift and ROI generated by personalized product-level promotions. In this experiment, the test group would receive personalized Coca-Cola offers based on the model’s predicted probabilities. In contrast, the control group would receive conventional mass promotions, such as a generic “$1 off any Coca-Cola” discount. Key performance indicators such as conversion rate, redemption rate, average basket size, and incremental sales lift should be tracked over time to quantify the added value of personalization. The results of this test can validate the business case for scaling the system across other brands and categories and provide critical insights into the optimal promotional design for different customer segments.

Together, these recommendations offer a comprehensive strategy for operationalizing the Coca-Cola recommender system in a way that maximizes both marketing effectiveness and business value. By leveraging customer-level probabilities, product-specific affinities, and behavioral context, ACSE can deliver targeted promotions that are far more precise than traditional mass discounting. The system’s ability to identify high-propensity customers, exclude those with low conversion potential, and match users to the Coca-Cola products they are most likely to buy creates a strong foundation for ROI-driven personalized marketing. If implemented thoughtfully—with eligibility filters, seasonality logic, and controlled A/B testing—the recommender system can serve as a blueprint for future campaigns, not only for Coca-Cola but also for other brand partners seeking more innovative, data-influenced promotions.