



Theo Dawson



Virginia Maguire

Problem Statement

Context

- CCSWB has successfully transferred 92% of Food Service on Premise (FSOP) customers to use myCoke, the e-commerce ordering tool.
- The myCoke experience can be customized to each customer, which CCSWB has not yet done, but plans to in the near future.

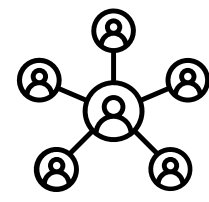


Opportunity

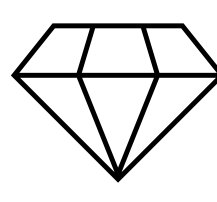
Use advanced analytics to better understand customers to understand how to best interact with them through myCoke.

Project Objective

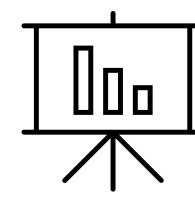
Our goal was to set the foundation for analytics in e-commerce by:



Providing detailed customer segmentation

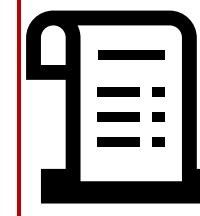


Proving the value of 1 modeling use-case



Building a process to host insights

Data



Sales

Customer, Product, Date of all sales
400M+ rows



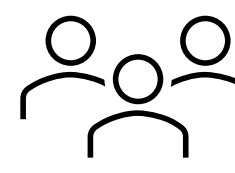
Product

Brand, Packaging, Pack size for each product



Orders

Used to identify myCoke orders



Customer

Trade Channel, Location, Status for each store/customer

Aggregation

Data was aggregated at a quarterly cadence for each customer.

Scope

- 2021 Q1 – 2024 Q1
- FSOP & Small Stores
- Eligible myCoke customers
- Primary volume customers

Segmentation

Our segmentation focuses on order efficiency and growth potential, allowing us to differentiate customers on the value they provide CCSWB now and in the future.

Order Efficiency

- Number of Orders
- Avg EQ per Order
- Avg Revenue & Profit per Order
- Avg Number of Products per Order
- myCoke Order Percentage
- Percentage of EQ Ordered in Quarter
- Percentage of EQ from Top 5 Brands

Growth Potential

(Percentage Change)

- Number of Order
- Avg Size (EQ) of Order
- Avg Revenue and Profit per Order
- myCoke usage

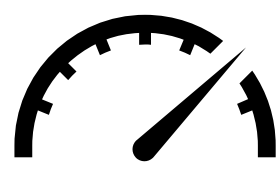
We separate new and established customers to generate deeper insights into both groups by changing the time granularity of the variables:

	New	Established
Quarters with CCSWB	<= 8	> 8
Order Efficiency Variable Level	Quarter	Rolling Year
Growth Comparison Period	Prior Quarter	Prior Rolling Year

Using K-Means, we end up with 14 clusters for FSOP that we can bucket into 4 groups. Small Store results were similar, but omitted.

High Potential or Value

- High Volume per Order
- Increasing volume per order
- Ordering much more frequently



New & Stable

- Brand New Customers
- Ordering more often & order size is similar
- Ordering less often & each order is more profitable
- Modestly increasing volume per order
- Placing orders in most quarters



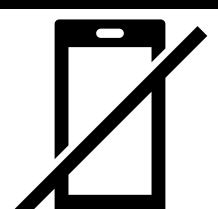
Customers to Watch

- Ordering more often & each order is less profitable
- Ordering less often & each order is smaller
- Placing orders infrequently
- No orders in period



Limited myCoke Usage

- Limited or no myCoke usage



Churn Prediction

A customer is considered **churned** if they **have not ordered in 12 consecutive months**.

To give CCSWB an opportunity to intervene, we **predict the likelihood of churning up to 1 year** before the customer does.

We focus on predicting **churn of FSOP myCoke customers**.

1 Identifying Key Variables

Feature importance, covariance, & mutual information were used to identify key variables:

Avg Days in % of Volume Quarters Segmentation myCoke
Week with from Top Without Orders Labels usage in
Order Brands in Prior Year Prior Year

2 Model Training

The data is split temporally with the last 3 quarters used for testing. For each model tried, we report the test accuracy & recall:

Model	Accuracy	Recall
Baseline (Average Churn Risk)	11%	100%
Logistic Regression	81%	88%
Random Forest	87%	86%
XGBoost	90%	82%
Isolation Forest	84%	74%

3 Model Selection

XGBoost performs the best, balancing accuracy & recall.

8X

6%

Better than baseline accuracy More accurate than other models

4 Model Use Case

This model will help the e-commerce team identify high-risk customers and prioritize site visits to engage with the customer. This helps CCWB retain customers and increase revenue.

Business Impact

The insights from our project inform support the CCSWB business by providing insights into

26K

myCoke customers, with revenue, in 2023 totaling over

\$350M

By accurately predicting churn, CCSWB could retain up to

1.1K

customers and whose annual revenue is estimated to by

\$1.3M

"Leveraging advanced analytics to segment customers allows us **to tailor initiatives with precision**, ensuring **each interaction drives engagement** with our brands and products."
- Eugenio Lozano, E-Commerce Lead

"No longer are the days of the peanut butter spread approach. Using advanced analytics, we can **make decisions based on facts** and are able to make the **most impactful to the business in an optimized way** while we continue our digital expansion."
- Jermey Parent, E-Commerce Analyst