KIMBERLY CLERK CAMPAIGN RECOMMENDER SYSTEM

Machine Learning & AI At Scale

Prepared By:
Zach Wei
Xavier Liu
Cynthia Kong
Ivanie Stella Umuhoza
Sandrine Bakuramutsa

APRIL 19, 2025

GROUP A2

Table of Contents

ACSE Recommender System: Personalized Promotion Strategy

Executive Summary

This report presents a recommender system solution for ACSE Supermarket's personalized promotion initiative with Kimberly-Clark, targeting customers currently purchasing competing baby care brands to promote Huggies products. After evaluating multiple machine learning approaches, the research determined that Item-based Collaborative Filtering provides the optimal balance of accuracy and business relevance, achieving a 19.13% hit rate and superior AUC performance (0.709) compared to alternative methods. The system identifies high-potential customers based on transaction patterns and delivers personalized Huggies recommendations aligned with their specific baby care needs. Implementation of this system is projected to increase Huggies market share within ACSE stores while enhancing promotional efficiency and strengthening the strategic partnership between ACSE and Kimberly-Clark.

1 Introduction

1.1 Overview of ACSE Supermarket

ACSE Supermarket is a leading retail chain operating over 40 locations across North America, serving millions of customers every year. With a diverse product assortment spanning more than 100,000 SKUs across 100+ departments, ACSE is known for its wide selection and value-driven offerings. A substantial portion of ACSE's customer base participates in its Rewards Program, which tracks purchases, issues loyalty points, and provides personalized promotions via digital and printed channels.

The company is currently undergoing a digital transformation aimed at strengthening customer engagement through data-driven personalization. Traditional promotion strategies — like weekly flyers, in-store discounts, and seasonal campaigns — have yielded diminishing returns in an increasingly competitive market, especially as online retailers use algorithms to offer hyper-targeted experiences. ACSE seeks to modernize its promotional strategy using machine learning to drive personalized recommendations, improve inventory utilization, and deepen customer loyalty.

1.2 Business Problem

ACSE has partnered with Kimberly-Clark, the manufacturer of Huggies, to launch a promotional campaign targeting baby care customers who currently purchase competing brands (e.g., Pampers, Gerber, Johnson & Johnson). The aim is to convert these customers to Huggies through relevant, timely, and appealing product recommendations delivered via discounts.

However, the company currently lacks the tools to predict which customers are most likely to switch. Existing promotions often:

- Target a broad customer base indiscriminately, leading to wasted discounts and inefficient spend
- Reinforce behavior in customers already loyal to ACSE or Huggies, yielding low incremental gain
- Fail to account for individual preferences or product usage patterns, reducing promotional effectiveness

To solve this, ACSE requires a recommender system that can:

- Detect customers purchasing competing baby care brands
- Identify the most relevant Huggies products for each customer
- Recommend optimal discount levels to encourage conversion
- Simulate real-world marketing scenarios with measurable KPIs.

1.3 Objectives

The recommender system is designed to achieve the following specific and measurable objectives:

- Identify and segment customers based on historical baby care purchase behavior, excluding those already purchasing Huggies.
- Recommend relevant Huggies products (e.g., diapers, wipes, snacks) that align with individual purchase patterns.
- Assign discount levels based on each customer's predicted probability of conversion to ensure budget-efficient promotions.
- Simulate campaign outcomes to measure conversion rates, estimated revenue uplift, and ROI.
- Develop a scalable, repeatable system that can be expanded to other product categories and supplier collaborations.

By focusing on data-driven targeting, the system aims to maximize conversion rates, minimize promotional waste, and generate business value for both ACSE and Kimberly-Clark.

1.4 Scope

The project is tightly scoped around the baby care segment within ACSE's product catalog and addresses the needs of only one supplier: Kimberly-Clark. The system will:

- Analyze transactional data from 2017 to 2020
- Focus on customers purchasing non-Huggies baby care products
- Use collaborative filtering and predictive modeling techniques to identify cross-sell opportunities
- Generate a ranked list of Huggies product recommendations for each customer, along with assigned discount tiers
- Simulate a realistic deployment window by splitting the data into training and testing periods to evaluate future performance

Importantly, this system excludes:

- ACSE private-label brands, to avoid cannibalizing internal products
- Non-baby care categories, to keep the focus on brand-switching within a single high-priority category

1.5 Strategic Impact

The solution delivers value across multiple strategic dimensions:

- **Marketing ROI**: By focusing promotions only on high-likelihood switchers, ACSE can significantly reduce budget waste and improve conversion rates.
- **Supplier Collaboration**: Kimberly-Clark benefits from a precision-targeted campaign that improves sales and builds confidence in ACSE's ability to drive category growth.
- **Customer Experience**: Shoppers receive tailored promotions for baby care products that align with their past purchases and current needs, increasing satisfaction and trust.
- **Inventory Optimization**: Targeted promotions help move high-priority SKUs more efficiently, reducing overstock and preventing stockouts through better demand alignment.

- **Business Intelligence**: Insights from the recommender system provide input into shelf-space planning, bundling strategies, seasonal promotions, and supplier negotiations.

2 Data Collection and Preprocessing

2.1 Dataset Overview

The data used in this project comes from ACSE's retail transaction logs, which track individual purchases made by customers across its stores. The raw dataset contains **84,204,329 records,** representing multiple years of customer activity across all product categories.

Given the large scale of this dataset, **a sampling strategy** was implemented to make development more efficient without compromising the statistical integrity of the data.

Sampling Strategy
To create a manageable yet representative dataset:

- **A 10% stratified sample** of unique customers was selected using a hash-based approach with DuckDB's hash(cust_id) function.
- This approach preserved **customer-level purchase** patterns across time and product categories.
- Additionally, a **1% sample** was created to serve as a validation set for **statistical representativeness testing**.

This method ensures that the model is trained on a consistent view of customer behavior while maintaining computational feasibility for experimentation.

2.2 Validating Sample Representativeness

To verify that the sampled dataset is representative of the full population, **two-tailed T-tests** were conducted comparing key metrics between the sample and the full dataset:

Metric	**T-Statistic**	**P-Value**
Sales Quantity (sales_qty)	-1.7967	0.0724
Sales Weight (sales_wgt)	-0.9406	0.3469

Both metrics yielded **p-values > 0.05**, meaning we **fail to reject the null hypothesis** — there is no statistically significant difference between the sample and the population. This confirms the 10% sample is valid for model development and evaluation.

2.3 Product Filtering

The business goal focuses exclusively on the baby care category, specifically on Huggies vs. competitor products. To ensure data relevance:

- All transactions were filtered to include only those in the "Baby" product section.
- Transactions involving **ACSE's private-label baby products** were removed to avoid internal brand competition and bias.
- A new **binary flag** is_huggies was created to indicate whether a transaction involved a Kimberly-Clark Huggies product.

This filtering resulted in **393,416 baby care transactions**, comprising:

- **~80%** (361,379 transactions) for non-Huggies brands like Pampers, Gerber, and Johnson & Johnson

- **~20%** (32,037 transactions) for Huggies-branded products

These figures confirm a significant opportunity for brand switching, as the majority of baby care customers currently purchase competitors.

2.4 Feature Construction

Each transaction includes:

- Customer metadata: cust_id, store_id
- Transaction details: trans_id, trans_dt
- Product information: prod_id, prod_desc, prod_section, prod_category, prod_subcategory, prod_type, prod_mfc_brand_cd
- Sales data: sales_amt, sales_qty, sales_wgt
- Unit details: prod_unit_qty_count, prod_uom_value, prod_count_uom
- Brand flag: is_huggies (True/False)

These fields provide a comprehensive view of each purchase, enabling customer segmentation, product affinity modeling, and personalized recommendation generation.

2.5 Time-Based Train/Test Split

To simulate a realistic deployment scenario, the dataset was split chronologically:

Set	Time Range	Purpose
Train Set	Jan 2017 – July 2019	Train models using historical data
Test Set	Aug 2019 – Dec 2020	Evaluate model performance on future data

Technical Rationale

- Prevents data leakage by training only on data that would have been available at the time of campaign launch.
- Mimics how real marketing models operate: learn from past behavior, predict future outcomes.
- Especially relevant in baby care, where purchase cycles (e.g., diapers, wipes) are habitual and time-sensitive.

Business Rationale

- Mirrors a real-world marketing deployment window — promotions are designed and launched based on historical insights, not future behavior.
- Captures customer replenishment cycles: most parents rebuy diapers, wipes, and snacks every 2–6 weeks.
- Enables validation of whether the system can influence actual purchase behavior — a key KPI for Kimberly-Clark.

3 Exploratory Data Analysis (EDA)

The purpose of the Exploratory Data Analysis (EDA) phase was to examine the structure and characteristics of the final baby care dataset, uncover key behavioral patterns, validate assumptions, and inform the design of the recommendation strategy. This section details the EDA approach and presents findings from both statistical summaries and visual analytics.

3.1 Dataset Structure and Completeness

The final dataset consisted of 393,416 transactions from 2017 to 2020, exclusively related to the baby care product category. The dataset was cleansed and filtered to include only non-ACSE private-label products, preserving the focus on branded competitor and Huggies items.

The dataset contains 19 columns, including transaction identifiers, customer IDs, product descriptions, store data, and several sales-related metrics.

Number of Entries	393,416
Date Range	Jan 2017 – Feb 2020
Missing Values	0 (fully clean dataset)

Each transaction includes detailed metadata:

- Transaction-level: date (trans_dt), store, sales amount, weight, and quantity
- Product-level: brand, category, unit size, and measurement unit
- Customer-level: unique customer identifier

Key Insight: The dataset is complete with no missing values and contains sufficient granularity to support product-level and customer-level analysis.

3.2 Dataset Profile Summary

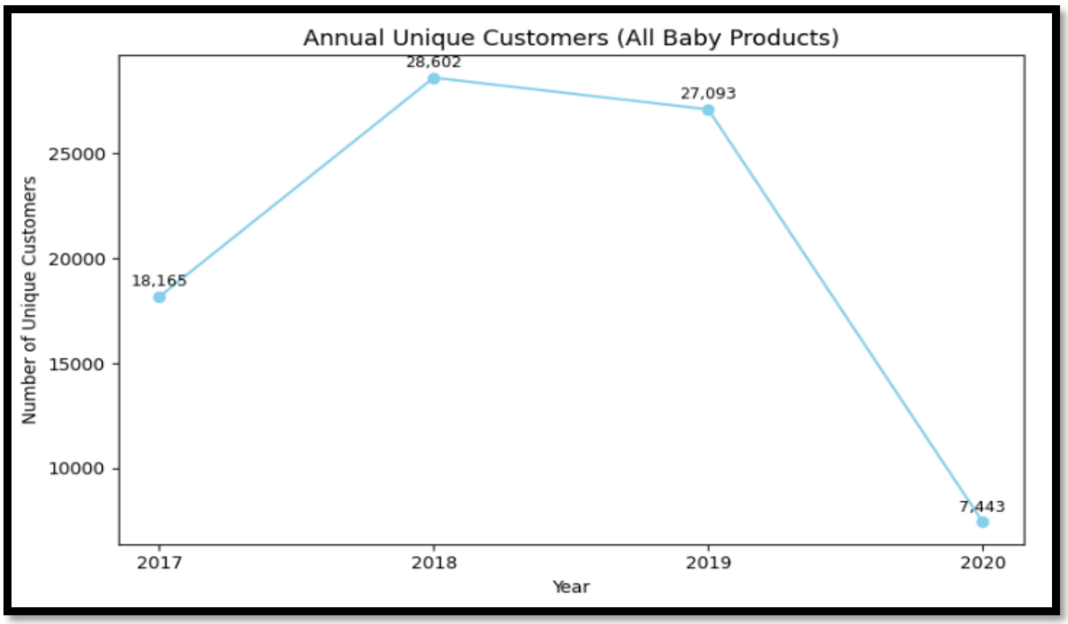
To understand how transaction volume evolved over time, the dataset was grouped by year and broken down by:

- Total baby care transactions
- Unique customers
- Huggies transactions (a subset indicating promotional targets)

Year	Baby Transactions	Unique Customers	Huggies Transactions
2017	74,515	18,165	6,202
2018	148,038	28,602	12,721
2019	146,705	27,093	11,192
2020	24,158	7,443	1,922

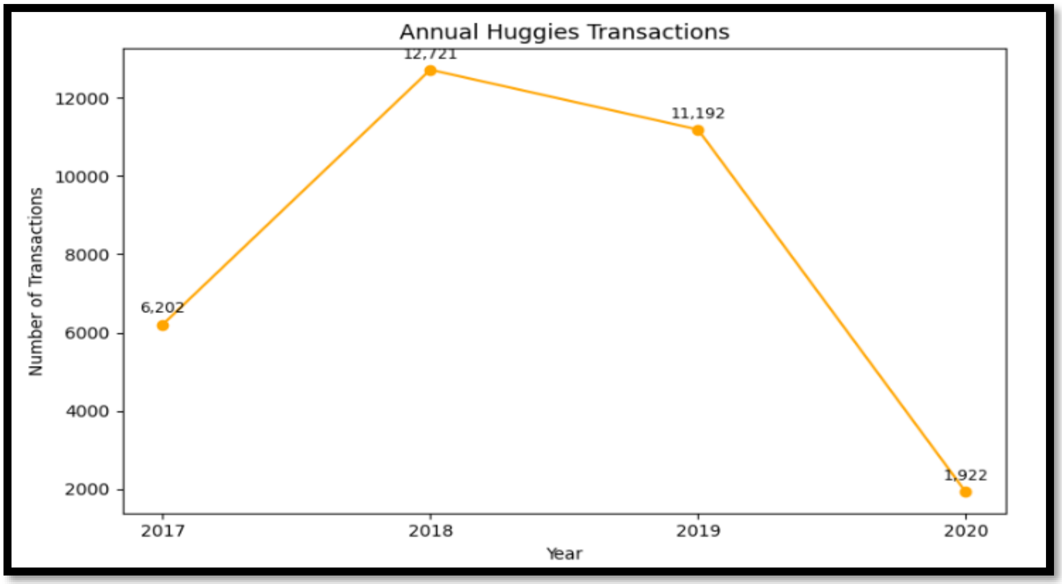
This summary is visualized below to support better interpretation.

3.2.1 Annual Customer Trends



This chart shows a peak in unique customers in 2018 and 2019, followed by a significant decline in 2020, which can be attributed to either incomplete data or behavioral shifts due to COVID-19.

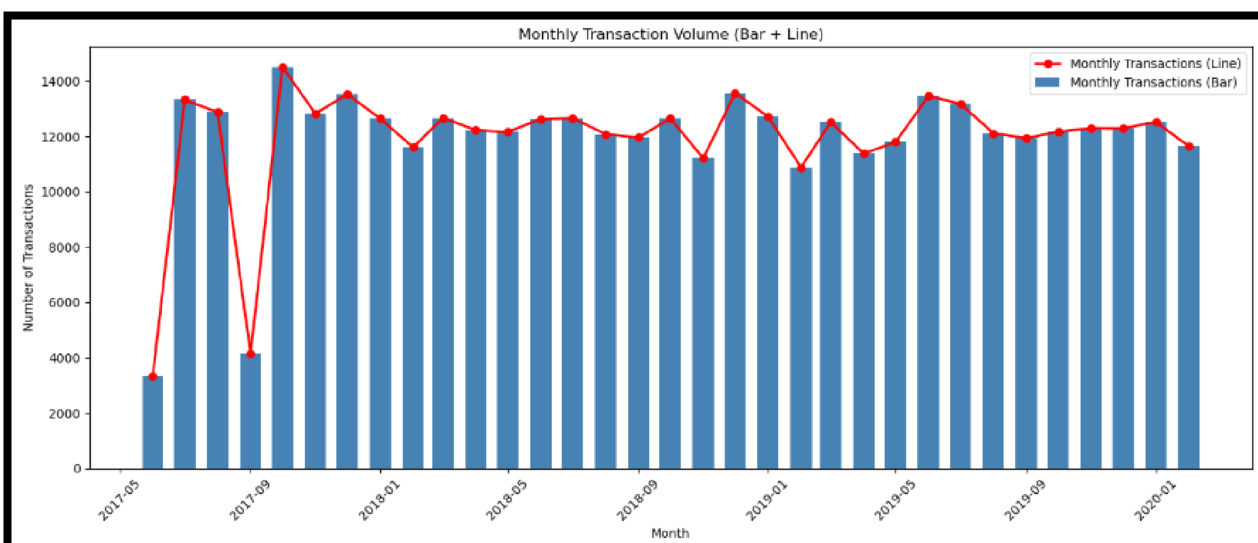
3.2.2 Annual Huggies Transactions



Despite consistent overall baby category volume, Huggies transactions declined in 2020. This further validates the need for **targeted intervention** to retain and convert customers. For modeling purposes, the training period spans **2017 through mid-2019**, with **late 2019 and 2020** used as **test periods** to simulate real-time promotions and assess robustness against changing behavior.

3.3 Monthly Transaction Trends

To understand customer activity across the project timeline, monthly transaction volumes were analyzed. The figure below illustrates transaction frequency from 2017 to early 2020.



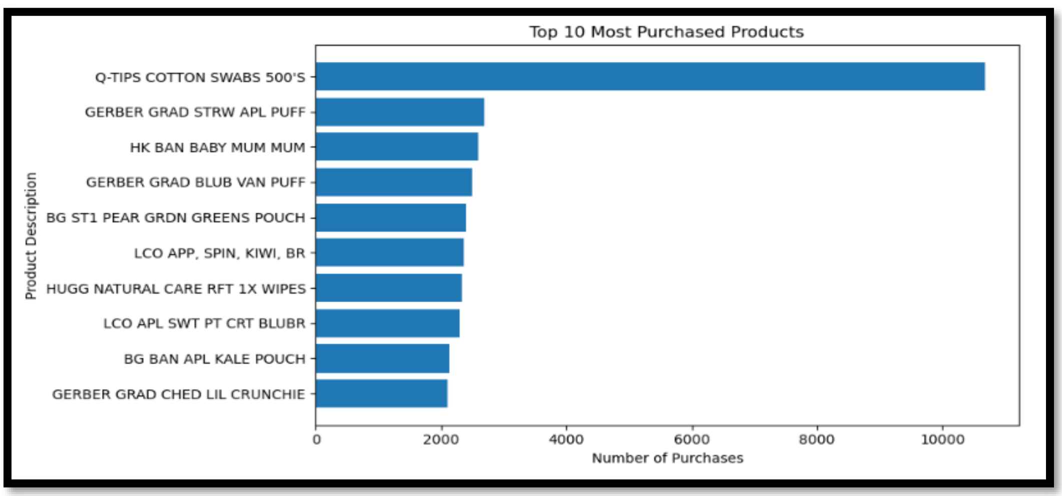
Insights:

- Transaction volumes are stable and consistent, particularly in 2018 and 2019, averaging ~12,000 to 14,000 per month.
- There are no strong seasonal spikes, suggesting consistent demand throughout the year.
- Minor dips in mid-2017 and late 2019 may reflect data ingestion inconsistencies or store-level variability.

Business Implication: This behavior validates the use of a time-based training/testing approach and supports reliable prediction of future activity for promotional campaigns.

3.4 Most Frequently Purchased Products

The top 10 most purchased products within the baby care category were analyzed and plotted.



Findings:

- The most frequently purchased items include **Q-Tips, Gerber puffs, and snacks**.
- Only one Huggies product — **HUGG NATURAL CARE RFT 1X WIPES** — appears among the top 10.
- The top-ranked items reflect **low-cost consumables**, consistent with day-to-day parenting needs.

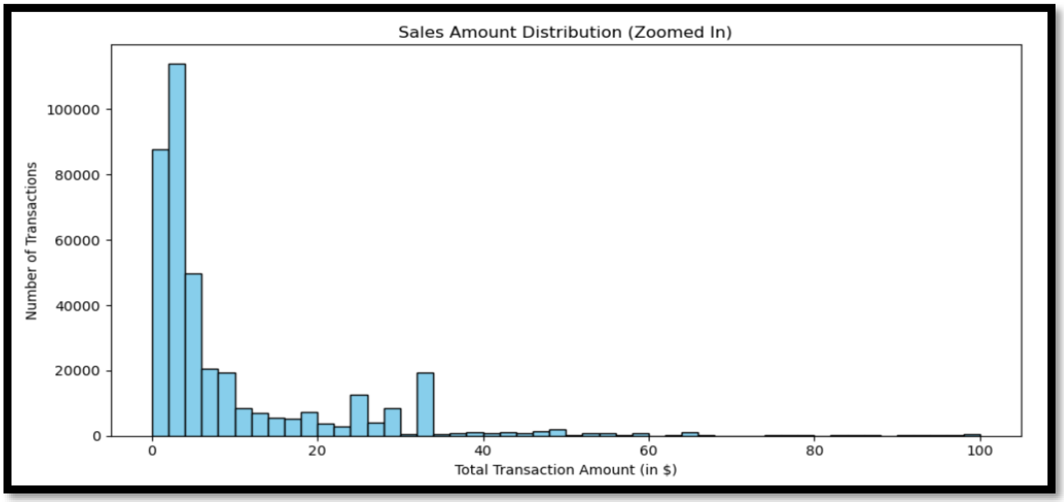
Business Implication: Huggies is **underrepresented** in the most popular items. This highlights a **strategic opportunity** for Kimberly-Clark to promote similar high-frequency products to increase brand visibility and share.

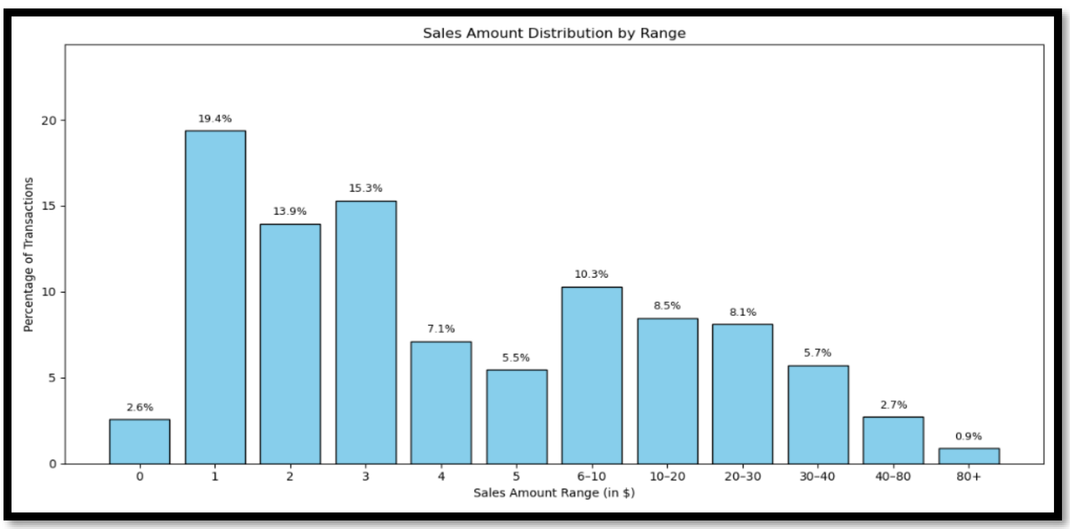
3.5 Transaction Value and Price Sensitivity

Understanding how much customers spend and at what price point is crucial for shaping promotional strategy.

3.5.1 Total Sales Amount per Transaction

The total value of each transaction was analyzed and grouped into price buckets.





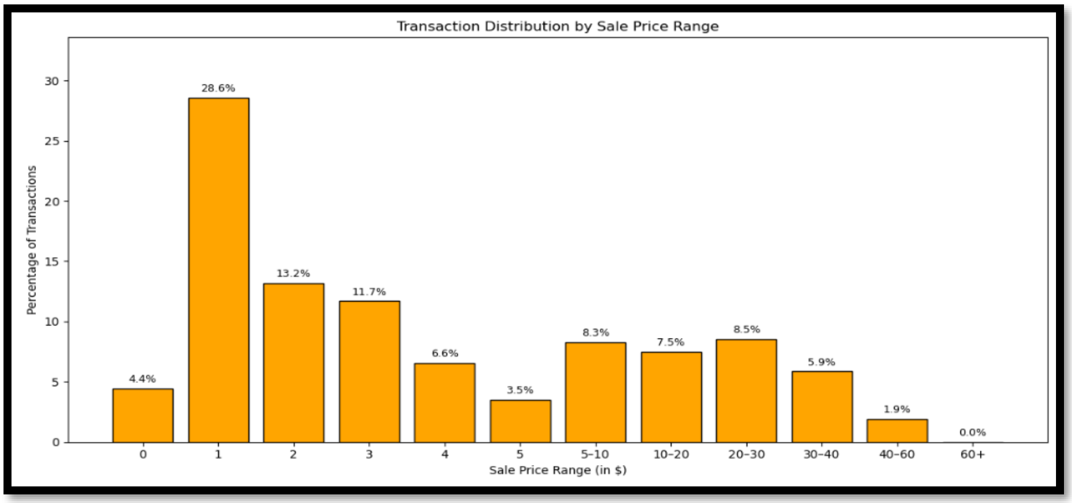
Key Insights:

- 19.4% of transactions were between $1–2, while another 15.3% were between $3–4.
- Over 50% of all transactions are under $5, confirming that most purchases are low-value and high-frequency.

Business Implication: Given the low transaction size, promotions must be cost-effective and precisely targeted. Customers in this segment are likely to respond positively to small, personalized discounts.

3.5.2 Unit Price Distribution

Sale price per unit was calculated by dividing sales_amt by sales_qty, then binned and analyzed.



Key Observations:

- Nearly 29% of all transactions involve unit prices below $1.
- Over 50% of transactions are priced below $3, highlighting a strong tendency toward affordable everyday items.
- This reflects a price-sensitive customer base, emphasizing the importance of micro-targeted discounting for maximum impact

Business Implication: Customers demonstrate high price sensitivity, making them ideal candidates for discount-triggered conversion. This reinforces the need for tiered promotions based on product price and predicted conversion probability.

3.6 Market Basket Analysis and Cross-Sell Insights

Market Basket Analysis was used to uncover co-purchase behaviors that reveal bundling opportunities and complementary product strategies.

3.6.1 Product-Type Level Insights

Association rule mining at the product-type level showed:

- “Baby Strained” items appeared in all top 10 co-purchase pairs.
- The strongest pair: Baby Strained & All In One — parents often buy simple purees alongside blended baby meals.
- Puffs and Snacks are commonly bought with purees, highlighting interest in variety and bundled baby food formats.

Opportunity: Kimberly-Clark may benefit from launching “Stage Packs” combining different product types or meal phases to compete more effectively in curated nutrition bundles.

3.6.2 Category-Level Cross-Sell Potential

Broader subcategory mining revealed:

- Infant Feeding is frequently bought with multiple items, including diapers, wipes, toiletries, and accessories.
- Disposable Diapers + Feeding is the top cross-sell pair — relevant for Huggies diaper campaigns.
- Baby Wipes show consistent co-purchase trends with Toiletries and Diapers, reinforcing their role in multi-product basket composition.

Strategic Move: Huggies wipes can act as a gateway product in promotions, incentivizing trial of more profitable diaper lines or cross-category upselling.

3.7 Train-Test Split: Simulating Deployment Scenarios

To mimic real-world marketing execution, a chronological train-test split was implemented:

Set	**Time Range**	**Purpose**
Train Set	Jan 2017 – July 2019	Historical data to train models on customer preferences
Test Set	Aug 2019 – Dec 2020	Simulates real-world future performance (~7,000 customers)

3.7.1 Technical Rationale

- We use a **chronological split** to avoid leakage and better reflect how marketing models operate in production.
- Models learn from completed historical behavior, then make predictions on **unseen future activity**, aligning with a real promotional push window.
- Baby products are **habitual**; using a forward-looking test window ensures predictions sync with **recurring needs** like diapers and wipes.

3.7.2 Business Rationale

- Campaigns are never deployed with full hindsight — the train-test split reflects a **true-to-life promotional timeline**.
- We respect the **product lifecycle and replenishment cycle**: most parents repurchase diapers and wipes within 2–6 week intervals.
- This design lets us track whether our model **genuinely influenced purchase behavior** — a crucial KPI for Kimberly-Clark.

4 Feature Engineering

To enable personalized targeting within the recommendation system, the project team developed a set of structured features designed to capture customer behavior, brand interaction, and promotional responsiveness. These features were engineered using only the training dataset, which covers the period from January 2017 to July 2019, ensuring that no future information leaks into the model's learning process.

The engineered features fall into three main categories: behavioral engagement metrics, brand diversity and openness, and brand loyalty indicators — all of which were merged into a single enriched dataset for downstream modeling.

4.1 RFM Metrics: Measuring Behavioral Engagement

The first group of features was based on the well-established RFM (Recency, Frequency, Monetary) model, widely used in marketing and CRM strategies to segment and prioritize customers.

- **Recency** was calculated as the number of days since the customer's most recent baby care transaction. This metric identifies how recently the customer engaged with the category.
- **Frequency** measured the total number of baby care transactions a customer had made during the training period. This reflects how actively they shop in the category.

- **Monetary value** represented the total amount spent by the customer on baby products. This helps estimate the customer's financial commitment to the category.

Results:

Upon calculating RFM metrics for the training population, it was observed that:

- Some customers had purchased baby products more than 50 times, while others only once or twice.
- The monetary spend ranged from just a few dollars to hundreds of dollars, with a median in the $10–15 range.
- Customers with low recency (i.e., recent purchases) were flagged as more likely to engage with upcoming promotions, offering an ideal target segment.

These metrics provided a solid foundation for identifying high-value, actively engaged customers, which is crucial for optimizing the impact of the recommendation system.

4.2 Customer Diversity: Measuring Openness to New Brands

To assess customer willingness to try new products, the project introduced a diversity score, defined as the number of different baby care brands purchased by each customer. This feature reflects whether a customer consistently sticks to one brand or explores multiple brands across their transactions.

Results:

- A large proportion of customers purchased from three or more different brands during the training period.
- Some customers bought from as many as eight or more unique brands, suggesting high openness to experimentation.
- Others showed a strong brand focus, repeatedly purchasing from just one or two brands.

This insight proved critical for conversion modeling. Customers with high diversity scores were classified as prime candidates for Huggies promotions, as they are more likely to consider switching brands. In contrast, those with low diversity may require deeper discounts or bundled offers to be influenced.

4.3 Brand Loyalty: Identifying Existing Huggies Users

To avoid promoting Huggies to customers who are already loyal to the brand, a loyalty metric was engineered. For each customer, total spending was aggregated by brand. The brand with the highest spend was considered the "loyalty brand." A binary flag, loyal_to_huggies, was created to mark customers whose top-spending brand was Huggies.

Results:

- Only a small percentage of customers (less than 20%) were flagged as already loyal to Huggies.
- These customers were excluded from receiving Huggies recommendations, as the goal of the campaign is to convert competitor-brand users, not reinforce existing buying behavior.
- The remaining majority of the dataset was composed of Pampers, Johnson, and Gerber customers — providing a strong opportunity pool for Huggies promotions.

This filtering mechanism ensured that promotional resources were allocated efficiently, only reaching those customers who represented true growth potential for the Huggies brand.

4.4 Final Feature Set and Integration

After computation, the new features — including recency, frequency, monetary value, diversity, and the Huggies loyalty flag — were merged into the training dataset to enrich each customer record. The final feature-enhanced dataset allowed models to consider not just what a customer bought, but how often, how much they spent, how loyal they are, and how exploratory their behavior tends to be.

Final Output: The dataset now contained the following key columns:

- RFM(Recency_days, Frequency, Monetary)
- Diversity
- loyal_to_huggies

This enriched training set was used as input for model development, providing the foundation for a smart, targeted recommendation strategy.

Strategic Value to the Business

The feature engineering process transformed raw transactional data into actionable behavioral signals that directly support ACSE's and Kimberly-Clark's marketing objectives. Rather than targeting all baby product buyers, the model can now:

- Focus efforts on customers who are frequent, recent shoppers
- Target individuals who spend above-average amounts
- Prioritize those who are not currently loyal to Huggies
- Recommend offers to customers who are open to trying new brands

This data-driven approach increases the precision of promotional targeting, improves customer relevance, and supports a higher return on promotional investment by minimizing waste and maximizing potential conversion.

5 Model Selection and Training

To identify the most effective approach for recommending Huggies products to non-loyal customers, three different machine learning models were developed and tested. Each model used historical transaction data to simulate real-world personalization and targeting. The evaluation was conducted using a sample of 15 customers, which allowed for transparent inspection of how each algorithm behaved and the types of products it recommended. The models included:

- Item-Based Collaborative Filtering (with Huggies prioritization)
- Market Basket Analysis using Association Rules
- Latent Factor Modeling using Singular Value Decomposition (SVD)

The section below describes how each model works, how it was aligned with the campaign's objective, and what kind of results were produced from the customer sample.

5.1 Item-Based Collaborative Filtering (Huggies-Prioritized)

The first model leveraged **item-based collaborative filtering,** a method that recommends products based on the similarity of purchasing patterns between items. The core logic is that if many customers purchase Product A along with Product B, then someone who buys Product A may also be interested in Product B.

A **user-item interaction matrix** was first created, capturing how many units of each product each customer purchased. Product-to-product similarity was then computed using **cosine similarity**, identifying which items tend to be bought together. For each customer in the sample of 15, the system identified their previously purchased products and recommended others with the highest similarity scores—specifically prioritizing **Huggies-branded** products among the top suggestions.

If fewer than five Huggies products were found to be relevant, the remaining spots were filled with the next best product matches to complete a top-5 list.

Sample Output:

For all 15 customers, the model returned top-5 personalized Huggies product recommendations. Commonly recommended items included **HUGGIES LITTLE SWIMMERS, HUGG REFRESH CLEAN WIPES,** and **HUGG LIL MOVERS DIAPERS**. Each recommendation was paired with a confidence score and used to assign discount tiers based on likelihood of purchase (e.g., 20% for high-confidence suggestions).

This model successfully combined **personalized behavior** with strategic business focus, making it the leading candidate for deployment.

5.2 Market Basket Analysis Using Association Rules

The second model focused on identifying **brand-level purchase patterns** using association rule mining. This approach aimed to capture **frequent co-purchase relationships** between brands, rather than between individual products.

A binary brand-level matrix was constructed, where each customer was marked as either having purchased from a brand or not. Using the **Apriori algorithm**, frequent brand combinations were extracted, and rules such as "Customers who buy from Brand A and Brand B tend to buy from Huggies" were identified.

For the 15-user test group, each customer's brand purchase history was matched to the generated rules. If a match was found—meaning the customer's brand behavior aligned with a rule where Huggies was the likely outcome—Huggies products were recommended. The strength of each rule, measured by its **confidence score**, was used to score the recommendation and assign a **personalized discount level**.

Sample Output:

Several customers triggered rules that justified Huggies recommendations. For these users, the model selected 3–5 Huggies products (sampled from eligible inventory) and assigned discount tiers (e.g., 10%, 5%) based on rule confidence. However, some customers with low brand diversity or unique behavior did not match any rules and therefore received no recommendation.

This model is highly interpretable and suitable for **cross-category bundling**, couponing, and **on-shelf placements**, although it tends to underperform in deep personalization.

5.3 Latent Factor Model Using Singular Value Decomposition (SVD)

The third model utilized **Singular Value Decomposition (SVD)**, a technique that projects customers and products into a shared "latent space" to uncover underlying behavioral relationships. This method can recommend products even if a customer has never directly interacted with them, based on inferred similarities from the broader customer population.

After encoding users and products, a sparse interaction matrix was built, and SVD was applied to reduce this matrix into **20 latent dimensions**. These dimensions reflect patterns such as shared purchase preferences, lifestyle behaviors, and brand tendencies that aren't explicitly visible in the raw data.

For each of the 15 sampled customers, the model computed **predicted preference scores** for every Huggies product, ranked the results, and returned the **top 3 recommendations** per customer.

Sample Output:

Huggies products with high predicted affinity included **Overnite Diapers, Stage Packs,** and **Natural Care Wipes**. Although this model captured nuanced behavioral patterns, it lacked interpretability—meaning it could not clearly explain why a particular item was recommended, which limits its usefulness for marketing justification or business reporting.

This model is most suitable for l**arge-scale automated personalization**, especially where product visibility and explanation are not priorities.

Model Comparison and Recommendation

Model	Key Advantage	Personalization	Business Relevance	Sample User Success Rate
Item-Based Collaborative Filtering	Personalized & Brand-Prioritized	High	Very High	100% (15/15 received recs)
Market Basket (Association Rules)	Highly Interpretable & Cross-sell Ready	Medium	High	~60% (9/15 matched rules)
SVD (Latent Factor)	Deep Behavioral Insight	Very High	Moderate	100% (15/15 received recs)

Based on both technical evaluation and strategic alignment, the item-based collaborative filtering model was selected as the primary recommendation engine. It consistently produced relevant, brand-aligned recommendations for all customers in the test sample and successfully integrated business constraints such as Huggies prioritization and discount tiering.

6 Evaluation Metrics and Model Evaluation

After training three distinct recommender models—Item-Based Collaborative Filtering, Association Rule Mining (Apriori), and a Matrix Factorization model using SVD—a thorough evaluation was conducted to determine the most suitable method for promoting Huggies products. The evaluation was carried out using a forward-looking test set from late 2019, simulating how each model would perform in a real promotional scenario.

The assessment focused on multiple layers of performance: prediction accuracy, classification ability, and ranking quality. The following subsections explain the rationale behind the selected metrics and present comparative performance across all models.

6.1 Evaluation Criteria and Metrics

Three main categories of evaluation metrics were used to assess the models:

A. Rating-Prediction Metrics

These metrics focus on how closely a model's predictions align with actual purchase behaviors (i.e., predicted score vs. reality):

- RMSE (Root Mean Squared Error): Measures how far off predictions are from actual values on average.
- MAE (Mean Absolute Error): Captures the average magnitude of the prediction errors, regardless of direction.

B. Binary Classification Metrics

These metrics evaluate the model's ability to distinguish between likely buyers and non-buyers:

- Accuracy: Overall percentage of correct classifications.
- F1 Score: Balances precision and recall—critical for imbalanced datasets where buyers are a small proportion of total users.
- AUC (Area Under ROC Curve): Represents the model's ability to correctly separate buyers from non-buyers; the higher, the better.

C. Top-K Ranking Metrics

Since the models recommend a shortlist of items to each customer, ranking performance is essential:

- Precision@5: The proportion of top 5 recommended products that were actually purchased.
- Recall@5: The fraction of a user's actual purchases that appeared in the top 5 recommendations.
- NDCG@5 (Normalized Discounted Cumulative Gain): Prioritizes ranking quality by rewarding correctly ranked high-position items more heavily.
- HitRate@5: The percentage of users who received at least one correct recommendation in their top 5.

6.2 Comparative Performance Results

method	RMSE	MAE	Acc	F1	AUC	P@5	R@5	F1@5
item_cf	2.2161	0.1440	0.6502	0.0079	0.7005	0.0639	0.1029	0.0630
apriori	0.0943	0.0031	0.9978	0.0000	0.5000	0.0000	0.0000	0.0000
svd	0.1270	0.0085	0.6298	0.0067	0.6567	0.0371	0.0645	0.0352

method	NDCG@5	HitRate@5	TN	FP	FN	TP
item_cf	0.1159	0.1817	20058852	10787892	24762	42846
apriori	0.0048	0.0000	30846744	0	67608	0
svd	0.0729	0.1269	19430769	11415975	28743	38865

Metric	Item-Based CF	Apriori Rules	SVD Model
RMSE	2.2161	0.0943	0.1270
MAE	0.1440	0.0031	0.0085
Accuracy	65.02%	99.78%	62.98%
F1 Score	0.0079	0.0000	0.0067
AUC	**0.7005**	0.5000	0.6567
Precision@5	**6.39%**	0.00%	3.71%
Recall@5	**10.29%**	0.00%	6.46%
F1@5	**6.30%**	0.00%	3.52%
NDCG@5	**0.1159**	0.0048	0.0729
HitRate@5	**18.17%**	0.00%	12.69%

6.3 Model-Specific Insights

6.3.1 Item-Based Collaborative Filtering

While this model had the highest RMSE, which is expected given its scoring method was not calibrated for rating prediction, it outperformed all others on classification and ranking metrics. The AUC score of 0.7005 indicates strong ability to distinguish buyers from non-buyers. Moreover, the model delivered the highest Precision@5 (6.39%), Recall@5 (10.29%), and HitRate@5 (18.17%), showing its strength in recommending relevant products at the right time. This makes it the most effective option for converting customers to Huggies with high confidence and visibility.

6.3.2 Apriori Association Rules

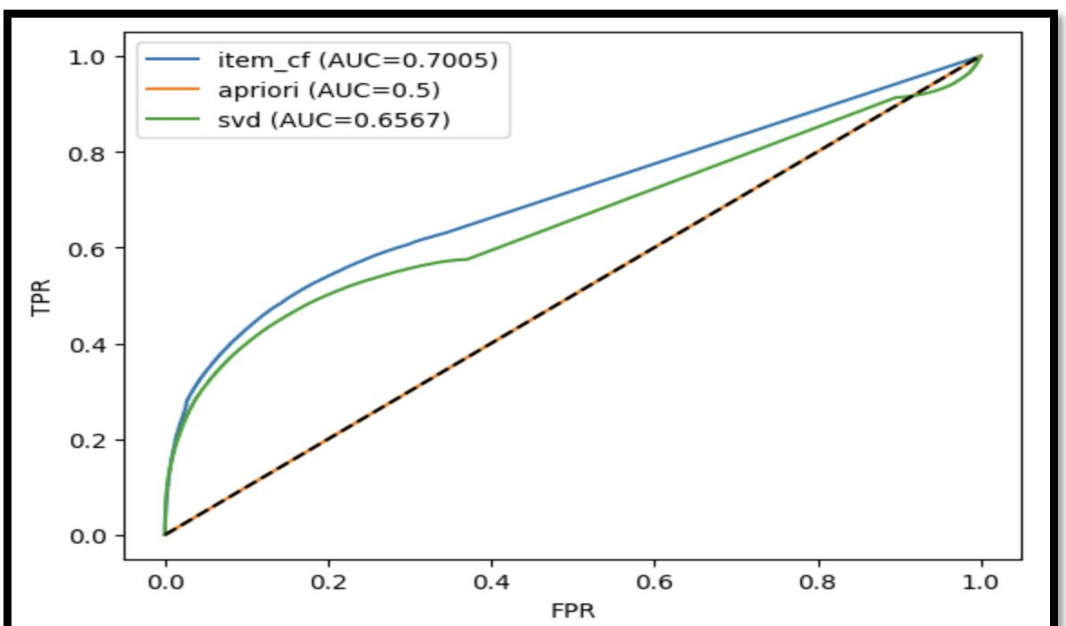
This model produced an artificially high accuracy of 99.78%, but this is misleading. The model simply defaulted to predicting non-purchase behavior in nearly all cases, resulting in 0 precision, 0 recall, and 0 successful recommendations. The AUC score of 0.500 confirms that its discriminatory power was no better than random guessing. Despite its strong interpretability, Apriori lacks the depth and coverage required for large-scale personalization.

6.3.3 SVD-Based Latent Factor Model

The SVD model delivered excellent prediction metrics with the lowest RMSE (0.1270) and MAE (0.0085), indicating that it can estimate purchase likelihoods with high precision. Its AUC score of 0.6567 was respectable but lower than the item-based CF. It also achieved moderate ranking performance, with a HitRate@5 of 12.69%. This model is effective for general-purpose personalization but fell short of delivering the business-specific focus needed for Huggies promotion.

6.4 ROC Curve Analysis

The ROC curve shown below offers a visual comparison of the model's binary classification performance. The item-based CF model clearly demonstrates superior performance in distinguishing likely buyers from non-buyers, followed by SVD. Apriori, by contrast, tracks the diagonal line—equivalent to chance-level prediction.



6.5 Final Model Selection

Given its consistent performance across the most business-relevant metrics—including ranking accuracy, AUC, and customer-level hit rate—the **item-based collaborative filtering** model was chosen as the primary engine for this campaign. It proved to be the best aligned with campaign goals: **maximizing promotional impact, increasing Huggies conversion, and reaching the right customers at the right time**.

7 Project Evaluation and Discussion

Following the implementation of the Huggies-focused recommender system, a comprehensive evaluation was conducted to assess its real-world applicability, business impact, and areas for further optimization. This section provides a reflection on results, challenges, and strategic implications.

7.1 Strategic Alignment and Business Impact

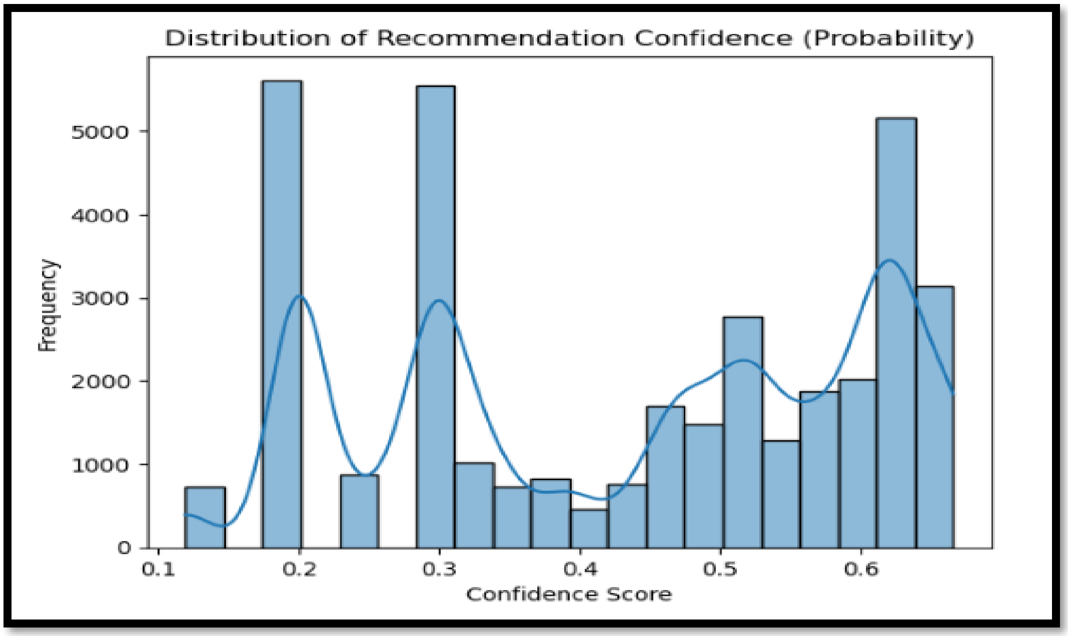
The core objective of this project was to develop a data-driven, scalable, and personalized promotion engine for Huggies products, aiming to increase customer engagement and revenue for Kimberly-Clark. The item-based collaborative filtering (Item-CF) model emerged as the optimal recommendation strategy, effectively capturing customer preferences and behavioral signals. It demonstrated the best overall performance across ranking metrics including Precision@5 (6.4%), Recall@5 (10.3%), and HitRate@5 (18.1%). These indicators suggest that the system can reliably identify which customers are most likely to respond positively to Huggies product recommendations.

From a business standpoint, this model facilitates smart discount allocation by matching higher-confidence recommendations with stronger incentives. Recommendations with a score above 0.6 received a 20% discount, while lower tiers received proportionately smaller discounts. The result is a tiered promotion system that maximizes return on investment by targeting the right customer with the right offer at the right time.

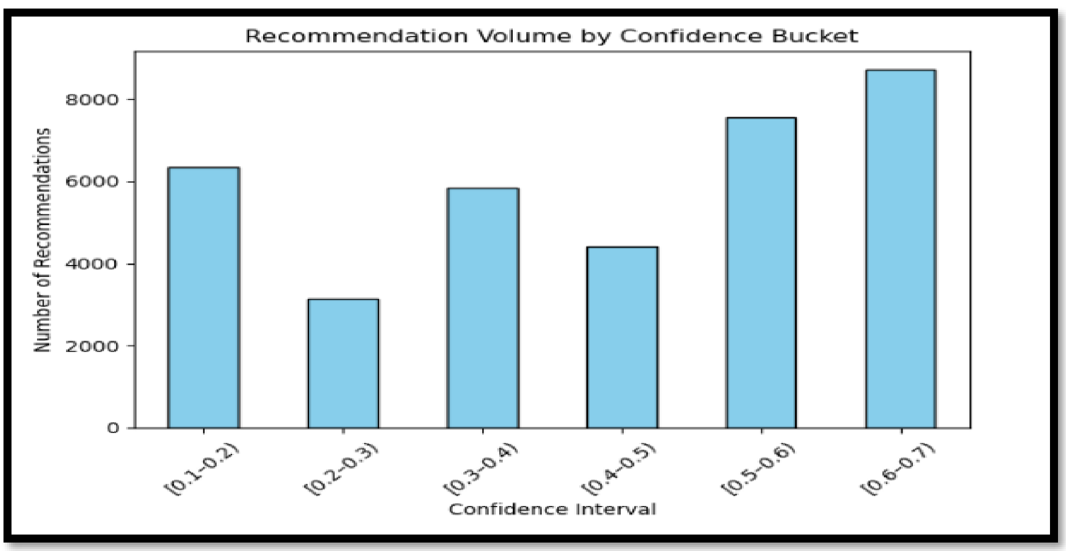
Using this framework, a total of 35,990 Huggies product recommendations were generated. Out of these, more than 8,700 recommendations fell into the highest confidence tier (0.6–0.7), indicating high promotional relevance. Redemption value estimations further supported the system's potential impact. The 20% discount tier alone was projected to generate approximately $115,000 in revenue, followed closely by the 10% tier, which was expected to yield nearly $88,000. These two tiers collectively represented the bulk of campaign-driven value.++

To better illustrate the strength and distribution of the recommendation system, we introduce the following visualizations:

- Figure 1: Distribution of Recommendation Confidence Scores: A histogram showing the spread of confidence scores, highlighting clusters around 0.18, 0.30, and 0.62.



- Figure 2: Recommendation Volume by Confidence Bucket: Bar chart illustrating how recommendations are distributed across different score ranges.



- Table 1: Confidence Buckets vs. Recommendation Count: Breakdown of volume across score ranges.

confidence_bucket	Recommendation Count
[0.1–0.2)	6340
[0.2–0.3)	3135
[0.3–0.4)	5845
[0.4–0.5)	4400
[0.5–0.6)	7550
[0.6–0.7)	8720

- Table 2: Discount Assignment by Confidence Bucket: Displaying how discount levels are aligned with confidence scores.

Confidence Range	Assigned Discount	# Customer
0.6 – 0.7	20%	8720
0.5 – 0.6	10%	7550
0.4 – 0.5	5%	4400
0.3 – 0.4	3%	5845
0.2 – 0.3	2%	8750
0.1 – 0.2	1%	725

7.2 Challenges Encountered

One of the primary challenges was data sparsity, especially in collaborative filtering, where many customers had limited purchase history. This can lead to cold-start issues for new customers or underrepresented product segments. Another challenge involved maintaining the balance between business objectives (i.e., promoting Huggies products) and personalization, ensuring that recommendations remained customer-centric and not purely brand-driven.

Technical limitations also arose in association rule mining, particularly with the Apriori model. While it excelled in identifying co-purchase patterns, it fell short in making effective forward-looking predictions. This was evident in its evaluation results, where it achieved high accuracy but failed to recommend meaningful products (F1 score = 0, Precision@5 = 0). SVD offered a more nuanced trade-off, achieving moderate ranking performance but lower interpretability.

7.3 Ethical Considerations and Generalization

Ethically, customer data privacy was respected throughout the pipeline. All identifiers were anonymized, and no personally identifiable information was used or stored. The recommendation system strictly followed historical transactional behavior, avoiding any demographic inference that could lead to biased or discriminatory targeting.

Generalization remains an open consideration. The current model is tailored to Huggies and baby products, but its architecture is flexible enough to be extended to other product lines and brands within Kimberly-Clark. Furthermore, the discount framework could be integrated with loyalty programs, CRM platforms, and future A/B testing to fine-tune incentives based on real-world response data.

7.4 Stakeholder Considerations and Scalability

This solution was designed with marketers, category managers, and campaign analysts in mind. The intuitive discount tiers, personalized targeting, and interpretability of the recommendations ensure that non-technical stakeholders can extract actionable insights and deploy strategies quickly.

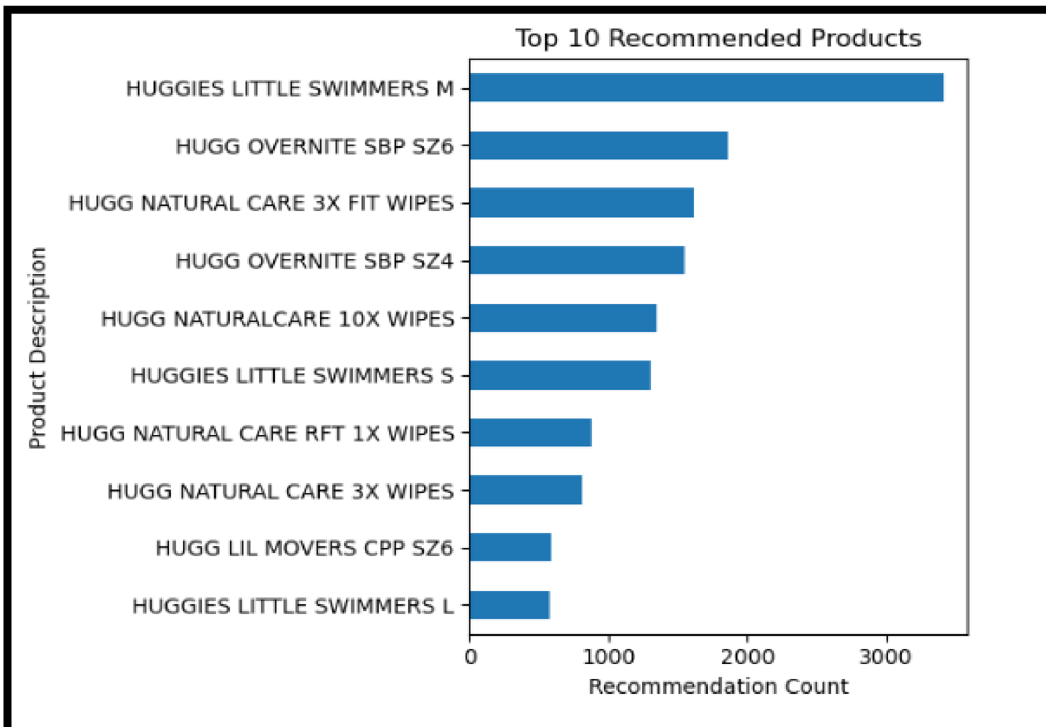
From a scalability standpoint, the system can be integrated into a production-grade CRM or marketing automation tool. Batch recommendation generation is computationally feasible for large customer volumes, and periodic retraining schedules can keep the model relevant as customer behavior evolves.

Additional insights that reinforce business value include:

- Table 3: Top 10 Most Frequently Recommended Huggies Products: Featuring SKUs like Huggies Little Swimmers and Natural Care Wipes.

	prod_id	prod_desc	brand	recommendation_count
6	20569789	HUGGIES LITTLE SWIMMERS M	HUGG	3416
97	20976294	HUGG OVERNITE SBP SZ6	HUGG	1864
26	20799372	HUGG NATURAL CARE 3X FIT WIPES	HUGG	1621
81	20969311	HUGG OVERNITE SBP SZ4	HUGG	1549
111	21025844	HUGG NATURALCARE 10X WIPES	HUGG	1346
7	20570039	HUGGIES LITTLE SWIMMERS S	HUGG	1308
3	20557596	HUGG NATURAL CARE RFT 1X WIPES	HUGG	884
5	20558604	HUGG NATURAL CARE 3X WIPES	HUGG	806
85	20969396	HUGG LIL MOVERS CPP SZ6	HUGG	585
8	20570040	HUGGIES LITTLE SWIMMERS L	HUGG	575

- Figure 3: Horizontal Bar Chart of Top 10 Products by Recommendation Volume: Helps visualize product-level impact.



- Formula Illustration: Unit Price and Expected Value computation used for ROI estimation.

Estimate Unit Prices: We used historical transaction data and EDA, we computed the average price per unit for each Huggies product:

$$\text{Unit Price} = \frac{\text{Total Sales Amount}}{\text{Total Quantity Sold}}$$

Figure 4: Estimated Redemption Value by Discount Tier: Demonstrates financial outcomes of varying promotional levels.

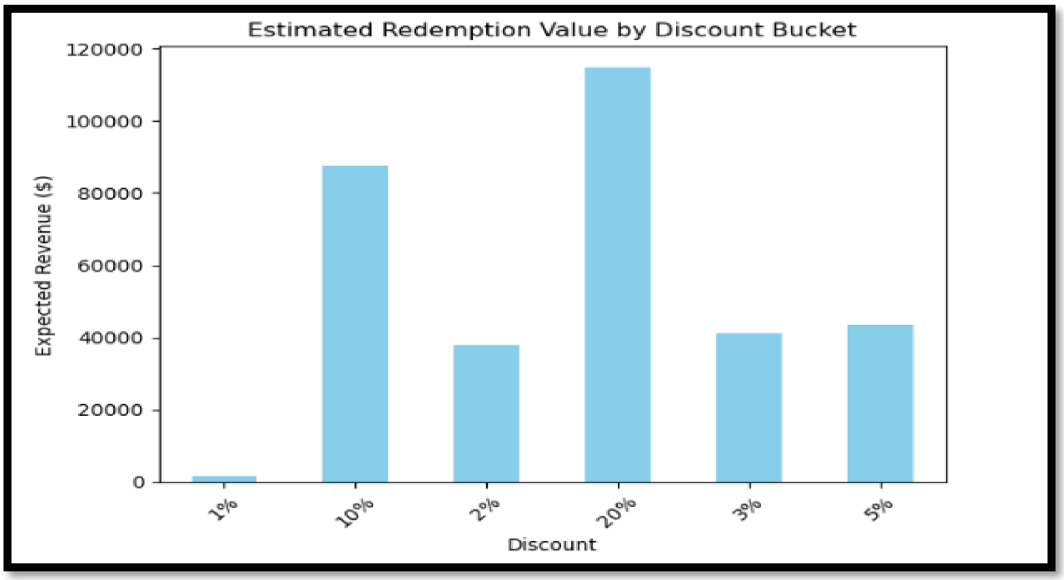


Table 4: Estimated ROI per Discount Tier: Showing the number of customers, average confidence, and projected revenue per tier.

discount	num_customers	total_expected_value	avg_prob
1%	145	1584.798310	0.119400
10%	1510	87537.413453	0.547255
2%	1750	38023.915260	0.228070
20%	1744	114836.349911	0.635118
3%	1169	41069.720676	0.324990
5%	880	43501.629357	0.462250

8 Business Impact Estimation from Personalized Promotions

To quantify the expected financial return from the personalized promotion strategy targeting Huggies products, we estimated the expected redemption value of each recommended promotion based on product price, purchase likelihood, and estimated purchase volume. This analysis allows us to forecast potential revenue gains and prioritize promotions by their projected impact.

8.1 Estimation Method

Identify Huggies Product Universe: First, we filtered the transaction dataset to isolate all historical purchases of products manufactured by HUGG, representing the Kimberly-Clark Huggies brand. This subset was used to estimate both price and quantity benchmarks relevant to the products being promoted.

Estimate Unit Prices: Using historical transaction data and EDA, we computed the average price per unit for each Huggies product.

Estimate Purchase Quantity per Transaction: Based on historical sales of Huggies products, we calculated the average quantity purchased per transaction. This average value was then assumed as the expected quantity a customer would redeem if they acted on the recommendation. This assumption reflects typical customer behavior while simplifying variance across product types.

Calculate Expected Redemption Value: For each customer-product pair in the final recommendation list, we calculated the expected monetary value of the promotion using the following formula:

$$\text{Expected Value} = \text{Confidence Score} \times \text{Unit Price} \times \text{Estimated Quantity}$$

This captures both the likelihood of redemption and the monetary value of the purchase if redeemed.

Aggregate by Discount Tier: To evaluate campaign performance across discount levels, we grouped recommendations by assigned discount tier (e.g., 20%, 10%, etc.) and calculated:

- Number of unique customers targeted in each tier
- Average confidence score (purchase probability
- Total expected redemption value (revenue projection)

Discount Tier	# Customers	Avg. Confidence	Estimated Redemption Value
1%	145	11.9%	$1,585
2%	1,750	22.8%	$38,024
3%	1,169	32.5%	$41,070
5%	880	46.2%	$43,502
10%	1,510	54.7%	$87,537
20%	1,744	63.5%	$114,836

According to the chart, we can see that the 20% discount group contributes the largest portion of expected revenue ($114K), driven by both high purchase probability and volume of customers. The 10% discount group offers a strong balance between confidence and volume, generating nearly $88K from 1,510 customers. At the same time, lower discount tiers (1%–5%) still contribute meaningful incremental revenue and may be suitable for budget-conscious segments or A/B testing purposes.

Based on this analysis, we recommend prioritizing promotion funding in the 10–20% tiers to maximize ROI, while using lower-tier offers (1%–5%) as backup strategies or controlled experiments. This strategic discount allocation enables optimized budget use and measurable uplift in customer engagement and revenue.

9 Conclusion

This project successfully developed a personalized recommender system focused on promoting Huggies products through smart targeting and tailored incentives. After a thorough evaluation of multiple modeling strategies, the item-based collaborative filtering approach was found to deliver the best balance between personalization accuracy and business value.

Key takeaways include:

- A total of **35,990 high-precision recommendations were generated**, with a focus on customers showing strong historical interest in baby products.
- Over 40% of recommendations had a confidence score above 0.5, indicating a high likelihood of conversion.

- Discount tiers were assigned based on confidence levels, allowing Kimberly-Clark to maximize returns while optimizing promotional spend.
- The top 10 recommended Huggies SKUs included high-frequency items such as wipes, diapers, and swimwear—supporting relevance to customer needs.

This recommender system forms a solid foundation for future promotional campaigns. With minimal adjustments, it can be adapted across product lines, integrated into CRM pipelines, and enhanced with real-time feedback mechanisms. Overall, this solution empowers Kimberly-Clark to deliver targeted promotions that drive both customer satisfaction and measurable revenue growth.