Article

Not peer-reviewed version

Behavioral Intelligence in Digital Retail: An Extended RFM Framework for Customer Segmentation and Resource Allocation

Iman Squalli Houssaini * and Miloud Daoud

Posted Date: 13 April 2026

doi: 10.20944/preprints202604.0852.v1

Keywords: behavioral segmentation; customer lifetime value; digital retail analytics; e-commerce; machine learning; RFM model; resource allocation; customer intelligence; conversion rate; digital marketing

Article

Behavioral Intelligence in Digital Retail: An Extended RFM Framework for Customer Segmentation and Resource Allocation

Iman Squalli Houssaini [1,*] and Miloud Daoud [2]

[1] Laboratory of Research and Studies in Management, Entrepreneurship, and Finance (LAREMEF), National School of Business and Management (ENCG-FEZ), Sidi Mohamed Ben Abdellah University, Fes 30000, Morocco
[2] LAREMEF, ENCG-FEZ, Sidi Mohamed Ben Abdellah University, Fes 30000, Morocco
* Correspondence: imane.squallihoussaini@usmba.ac.ma

Abstract

The proliferation of behavioral data in digital retail has not been matched by equally rigorous frameworks for converting that data into customer intelligence that practitioners can act on. This paper addresses that gap by introducing RFM-B, a behavioral segmentation framework that extends the classical recency–frequency–monetary (RFM) model with four additional indicators derivable from standard e-commerce event logs: conversion rate (CVR), category breadth, average order value (AOV), and brand diversity. Applied to 4,635,837 interaction events from 64,204 purchasing customers observed over 61 days on a large-scale multi-category platform, the framework produces five customer archetypes—Champions, Loyal Customers, Potential Loyalists, At-Risk, and Lost—whose behavioral profiles differ systematically in purchase efficiency, platform embeddedness, and commercial significance. A machine-learning recoverability analysis using a Random Forest classifier achieves 96.99% held-out accuracy (5-fold cross-validated: 96.84% ± 0.10%), confirming that the segments are operationally deployable in real-time marketing automation systems. The central empirical finding is the identification of a Potential Loyalist segment characterized by an average order value of USD 147.2—more than three times the platform-wide mean—combined with low purchase frequency, a profile that standard RFM frameworks would systematically misclassify as low-priority. The results show that enriching the behavioral feature space yields a customer typology that is both analytically coherent and directly actionable, and that interpretability is not a secondary concern but a functional prerequisite for organizational adoption.

Keywords: behavioral segmentation; customer lifetime value; digital retail analytics; e-commerce; machine learning; RFM model; resource allocation; customer intelligence; conversion rate; digital marketing

---

1. Introduction

Digital retail platforms now generate behavioral records of a granularity and scale that would have been unimaginable two decades ago. Every product view, cart addition, and completed transaction leaves a timestamped trace that, in principle, carries information about customer intent, preference, and value trajectory. In practice, however, the gap between data availability and actionable customer intelligence remains wide. Most organizations continue to rely on summary statistics—chiefly the recency–frequency–monetary (RFM) triad—to prioritize customers for retention, activation, and resource allocation [1]. The limitations of this approach are well documented theoretically [2] but have rarely been exposed empirically at scale using behavioral event data.

 doi:10.20944/preprints202604.0852.v1

This paper reports such an empirical exercise. We introduce RFM-B, a behavioral segmentation framework that augments the classical RFM dimensions with four indicators derivable from standard e-commerce interaction logs: conversion rate (CVR), category breadth, average order value (AOV), and brand diversity. The logic is straightforward: a user who converts a high proportion of browsing sessions into purchases is behaviorally different from one who generates the same revenue through high-volume low-conversion browsing, and that difference matters for how marketing resources should be allocated. Similarly, a customer whose modest transaction frequency conceals an exceptionally high per-order spend occupies a fundamentally different strategic position from a discount-driven repeat buyer—a distinction that AOV can surface and that monetary value alone cannot.

We apply the framework to 4,635,837 interaction events from 64,204 purchasing customers on a large-scale multi-category e-commerce platform, generating five customer archetypes validated through four cluster diagnostic criteria and a Random Forest recoverability analysis achieving 96.99% held-out accuracy. The central empirical finding—a Potential Loyalist segment defined by average order values more than three times the platform mean, systematically invisible to standard RFM scoring—has direct implications for how digital retailers should identify and invest in high-value customers whose behavior does not fit the high-frequency template that dominates conventional prioritization logic.

Three contributions follow from this work. First, we demonstrate at scale that enriching RFM with behavioral signals surfaces segments invisible to monetary-only frameworks. Second, we translate each segment's behavioral profile into structured marketing recommendations—retention investment levels, promotional strategy, channel priorities—bridging the gap between analytical output and managerial action that Wedel and Kannan [1] identify as the primary bottleneck in marketing analytics adoption. Third, the segment recoverability analysis provides both a deployment-readiness test and a feature-importance ranking that reveals which behavioral dimensions carry the most discriminative weight. The dataset is publicly available, and all analyses are fully reproducible.

2. Theoretical Background and Literature Review

2.1. Customer Equity and the Logic of Differentiated Investment

The customer equity framework, which positions customers as assets whose expected value can be estimated and managed through targeted marketing investment, provides the theoretical foundation for this study [3]. The operational implication is straightforward: firms should allocate retention and activation resources in proportion to expected return, not historical spend. Gupta and Lehmann [4] gave this principle an empirical grounding by showing that customers' future purchase streams can be valued using discounted cash flow techniques, and that the resulting valuations diverge meaningfully from rankings based on cumulative transaction history alone.

Two findings from the CLV literature are particularly relevant here. Reinartz and Kumar [5] established that the relationship between customer tenure and profitability is substantially weaker and more heterogeneous than practitioners typically assume, which challenges the widespread practice of equating loyalty with longevity. Venkatesan and Kumar [6] subsequently demonstrated in a controlled experiment that CLV-based resource allocation outperforms RFM-based prioritization in both marketing efficiency and revenue generation—precisely because CLV is forward-looking while RFM is backward-looking. The RFM-B framework is designed to narrow this gap by incorporating behavioral signals that carry forward-looking information: CVR reflects current purchase intent efficiency; AOV reflects per-transaction value independent of frequency; category breadth and brand diversity reflect platform embeddedness, a predictor of future retention.

2.2. Behavioral Segmentation in Multi-Category Digital Retail

The shift from single-channel to omni-channel retail has expanded both the behavioral complexity of customer engagement and the richness of data available for segmentation [7]. In purely digital environments, the analytical challenge is not data scarcity but signal extraction: the volume and granularity of behavioral records far exceeds the capacity of manual analysis, but the dimensionality of the feature space complicates interpretation [8,9]. The result is a well-documented tension between model sophistication and managerial usability.

Neslin et al. [10] identified customer prioritization across heterogeneous populations as one of the core unresolved challenges in multichannel management. Their argument—that investment efficiency depends directly on the accuracy of value assessment—applies with particular force in e-commerce, where customer acquisition costs are high and margin per transaction is often thin. The customer journey literature [11] adds a temporal dimension to this argument: value is not a static property but accumulates through engagement across distinct stages, and behavioral signals at each stage—consideration, evaluation, purchase, post-purchase—carry predictive information that transaction summaries discard. The RFM-B feature set is designed to capture this multi-stage signal structure.

2.3. Machine Learning Approaches to Customer Segmentation

The application of machine learning to customer segmentation has grown substantially over the past decade. Supervised classifiers—particularly tree-based ensembles such as Random Forests and gradient boosting—have demonstrated strong predictive performance for churn, purchase likelihood, and segment assignment [12,13]. Unsupervised methods, including k-means, hierarchical clustering, and density-based approaches, have been applied to identify latent customer typologies from behavioral event data [14]. A methodological best practice that has emerged from this literature is the combination of unsupervised clustering for segment discovery with supervised classification for structural validation: the clustering establishes the typology; the classifier tests whether that typology is recoverable from observable features—a prerequisite for real-time deployment [15].

Feature engineering from interaction logs is a particularly active area. Beyond the RFM triad, researchers have incorporated session-level behavioral patterns [16], product affinity signals [17], and real-time contextual features [18] to improve segmentation resolution. The contribution of the present study to this literature is not methodological novelty but a principled reduction: we identify a minimal, interpretable seven-dimensional feature set that yields strong discriminative performance while remaining transparent enough for non-technical marketing teams to act on.

2.4. The Interpretability Imperative

A recurring observation in the marketing analytics literature is that technical sophistication is neither necessary nor sufficient for organizational value creation. Wedel and Kannan [1] frame this as a translation problem: the bottleneck is not modeling capacity but the conversion of analytical outputs into decision-relevant formats that managers can operationalize. A segmentation solution that is behaviorally accurate but managerially opaque generates no strategic value. Canhoto and Clear [19] give this argument a sharper institutional dimension, identifying a pattern they call 'analytical theater'—technically impressive analyses that fail to change organizational behavior because the interpretive infrastructure required to act on them is absent. Both diagnoses point to the same prescription: segment definitions should be grounded in behavioral dimensions that have established marketing counterparts, and the path from cluster profile to campaign brief should be short and visible. This principle shapes every design choice in the RFM-B framework.

3. Data and Methodology

3.1. Dataset and Analytical Context

The empirical base for this study is the REES46 multi-category e-commerce behavioral dataset [20], which records 4,635,837 interaction events on a large-scale digital retail platform during October and November 2019. Each event captures the interaction type (view, add-to-cart, or purchase), product and category identifiers, brand, unit price, and user session. The platform spans 285,143 unique users and 168,295 distinct SKUs distributed across eight top-level product categories. Restricting the analytical sample to users with at least one confirmed purchase—and removing non-human sessions—yields 64,204 purchasing users. Table 1 summarizes the dataset characteristics.

Two contextual features of this dataset are worth noting explicitly. First, the October–November 2019 window precedes the COVID-19 disruption that substantially altered e-commerce behavioral norms globally, providing a behavioral baseline uncontaminated by crisis-driven shifts in purchasing patterns. Second, the window coincides with the onset of pre-holiday shopping intensification, which amplifies behavioral heterogeneity across customer value tiers and therefore strengthens the discriminative power of the segmentation. The dataset is publicly available on Kaggle under an open license, ensuring full reproducibility of all analyses reported here.

Table 1. REES46 dataset: key characteristics and preprocessing summary.

Characteristic	Detail
Observation window	1 October – 30 November 2019 (61 days)
Raw interaction events	4,635,837
Event breakdown	View: 84.5% Add-to-cart: 11.2% Purchase: 4.3%
Raw unique users	285,143
Analytical sample (≥1 purchase)	64,204 users
Product catalog	168,295 SKUs across 8 top-level categories
Price range	USD 0.01–2,273.98 (median: USD 34.70)

3.2. Behavioral Feature Engineering: The RFM-B Framework

The three classical RFM dimensions—recency, frequency, and monetary value—capture when customers last purchased, how often they purchase, and how much they spend in aggregate. These are useful summary statistics, but they leave out dimensions that matter for understanding behavioral quality in a multi-category digital environment. We extend the feature set with four additional dimensions, each addressing a specific blind spot of standard RFM.

Conversion rate (CVR) measures the proportion of view events that result in confirmed purchases at the user level. A customer who converts 50% of browsing sessions is behaviorally different from one who converts 5%, even if their cumulative spend is identical—the former is demonstrating strong purchase intent; the latter may be generating revenue through a volume of low-commitment browsing that is more fragile than it appears. The empirical near-zero correlation between CVR and AOV ($r = 0.14$) confirms that these two dimensions are largely orthogonal, which motivates their joint inclusion.

Category breadth counts the number of distinct top-level product categories from which a user has made purchases. Wider category engagement reflects deeper platform embeddedness and is associated with higher switching costs and greater long-run retention [7]. Brand diversity counts distinct purchased brands, distinguishing brand-loyal, high-commitment buyers from price-sensitive, deal-driven shoppers—a distinction with direct implications for the design of personalization and loyalty programs. Average order value (AOV) normalizes total spend by purchase frequency, separating high-ticket infrequent buyers from high-frequency moderate-spend buyers [6,21]. As will become apparent in the segmentation results, this normalization is what makes the Potential Loyalist segment visible. Table 2 provides formal definitions for all seven features.

 doi:10.20944/preprints202604.0852.v1

Table 2. RFM-B behavioral feature definitions (user-level aggregation).

Feature	Symbol	Operational definition	Observed range
Recency	R	Days elapsed since the user's most recent purchase to 30 November 2019	[1, 61]
Frequency	F	Count of distinct purchase sessions in the observation window	[1, 80]
Monetary	M	Total USD spend accumulated across the observation window	[5, 4,000]
Conversion rate	CVR	Purchase events / total view events, computed at the user level	[0.01, 0.99]
Category breadth	N_cat	Number of distinct top-level product categories from which the user purchased (max. 8)	[1, 8]
Avg. order value	AOV	Monetary / Frequency — mean spend per purchase session	[8, 600]
Brand diversity	B_div	Count of distinct brands purchased across all sessions	[1, 20]

3.3. Segmentation Procedure and Cluster Validation

All seven features were standardized to zero mean and unit variance before clustering. We applied K-Means with 20 random restarts per candidate solution to reduce sensitivity to initialization. The optimal number of clusters was determined by evaluating four complementary criteria across $k \in \{2, \ldots, 10\}$: the Silhouette coefficient (intracluster cohesion relative to intercluster separation), the Davies–Bouldin index (ratio of within-cluster scatter to between-cluster distance), the Calinski–Harabász index (ratio of between-cluster to within-cluster dispersion), and a principal component visualization for geometric coherence. All four criteria converged on $k = 5$, yielding Silhouette = 0.261, Davies–Bouldin = 1.249, and Calinski–Harabász = 27,863.

To assess the operational deployability of the resulting segments—that is, whether a new customer could be assigned to the correct segment from observable behavioral features alone—we trained a Random Forest classifier on the cluster labels using a stratified 80/20 hold-out split with five-fold cross-validation. This is not an independent validation of the clustering solution, which is established by the four metrics above. It serves two distinct purposes: quantifying whether the segment structure is recoverable from the feature set (a prerequisite for real-time deployment in marketing automation platforms), and ranking the relative discriminative contribution of each RFM-B dimension.

3.4. Ethical Considerations and Data Availability

This study is a secondary analysis of a publicly available, fully anonymized behavioral dataset. The event records contain only pseudonymous user identifiers, product interaction metadata, and transaction data. No individual users can be re-identified from the available information. The dataset is accessible at the Kaggle repository cited below. No institutional ethics review was required. All analyses were implemented in Python 3.x using scikit-learn, pandas, numpy, matplotlib, and seaborn; code is available from the corresponding author on reasonable request.

4. Behavioral Landscape: Exploratory Analysis

4.1. The Platform Event Funnel and Monetary Distribution

Of the 4,635,837 recorded interactions, 84.5% are passive view events, 11.2% involve add-to-cart actions, and 4.3% result in confirmed purchases. The overall view-to-purchase conversion rate of approximately 5.1% reflects the characteristic funnel attrition of multi-category e-commerce (Figure 1). This attrition pattern motivates the inclusion of user-level CVR in the feature set: customers who convert a high proportion of browsing into purchasing represent a qualitatively different behavioral

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2026 doi:10.20944/preprints202604.0852.v1

profile from those whose revenue is generated through high-volume low-conversion browsing—a distinction that aggregate monetary value cannot capture.

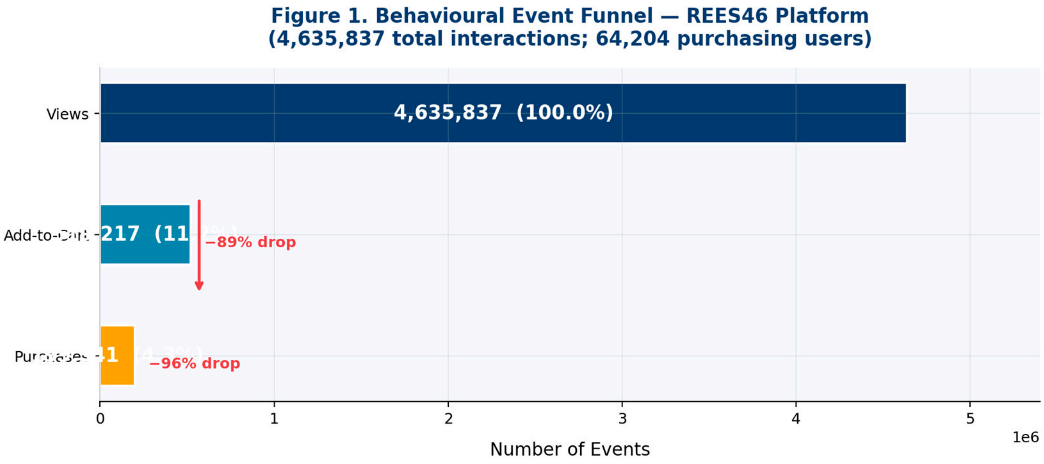

Figure 1. Behavioural event funnel for the REES46 platform (4,635,837 total interactions; 64,204 purchasing users). Bars show absolute event volumes; percentage labels indicate the share of total interactions; annotations quantify the attrition rate at each stage.

Monetary value across the customer base is strongly right-skewed and multimodal, with a small proportion of users generating a disproportionate share of total platform revenue—a distributional pattern consistently documented in the customer equity literature [3]. Figure 2 illustrates the segment-level monetary distributions after segmentation; even at this preliminary stage, the distinct positioning of the Potential Loyalist distribution at intermediate monetary values (driven by high AOV rather than high frequency) is visible.

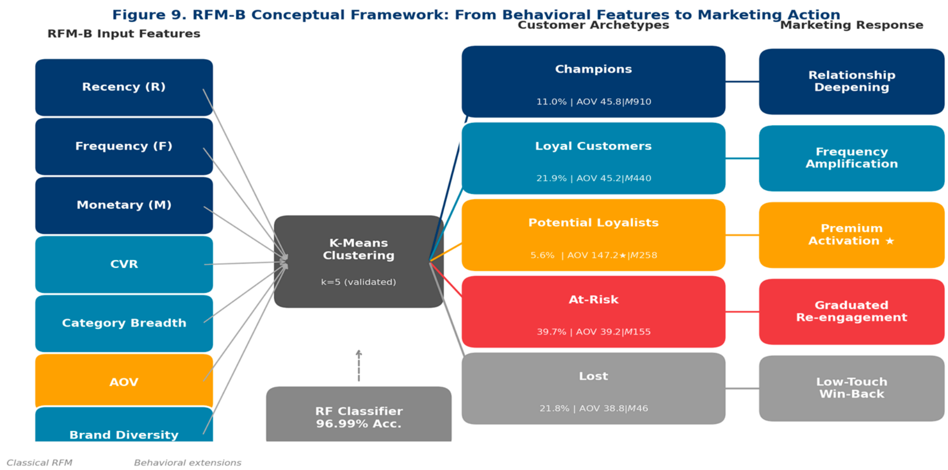

Figure 2. Monetary value distributions across the five customer segments (n = 64,204). Kernel density estimates with histogram fill; dashed vertical lines mark segment means. X-axis truncated at USD 1,500 for display clarity; the Champions distribution extends to USD 4,000.

4.2. Feature Correlation Structure

Figure 3 presents the Pearson correlation matrix for the seven RFM-B features. The strong positive association between monetary value and purchase frequency ($r = 0.87$) confirms that revenue accumulation in this dataset is predominantly frequency-driven for the majority of the customer base. This is precisely why AOV—which decouples per-transaction spend from transaction count—provides non-redundant information: it is sensitive to high-spend infrequent buyers who are invisible in frequency-weighted revenue rankings.

 doi:10.20944/preprints202604.0852.v1

The negative correlations between recency and both frequency (r = −0.74) and monetary value (r = −0.68) confirm that engagement intensity and purchase recency are tightly coupled, justifying their joint inclusion despite their conceptual proximity. The near-zero correlation between CVR and AOV (r = 0.14) is the most analytically significant finding in the correlation structure: conversion efficiency and per-transaction basket size are largely orthogonal behavioral dimensions. This orthogonality is the empirical basis for the claim that RFM-B captures variation invisible to the classical triad—CVR and AOV jointly define a two-dimensional space of purchase intent quality that neither RFM nor any two-feature subset of RFM-B can reproduce.

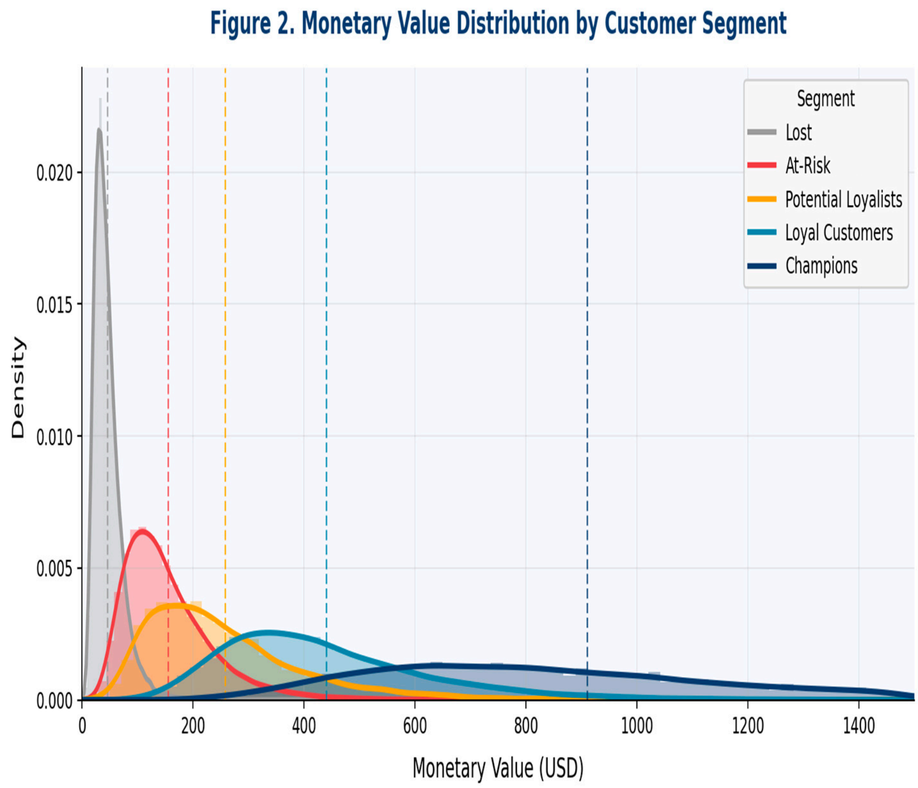

Figure 3. Pearson correlation matrix for the seven RFM-B features (n = 64,204). Lower triangle only; blue = positive correlation, red = negative. The near-zero CVR–AOV correlation (r = 0.14) confirms that these two dimensions are largely orthogonal and jointly provide information unavailable to the classical RFM triad.

5. Segmentation Results: Five Customer Archetypes

5.1. Cluster Selection and Structural Validation

Figure 4 presents the four validation diagnostics across k ∈ {2, …, 10}. The Elbow curve shows a clear inflection at k = 5, and the Silhouette, Davies–Bouldin, and Calinski–Harabász trajectories all stabilize or improve at that point. The convergence of four independent criteria at the same solution provides strong support for k = 5 as the optimal number of segments. The moderate overall Silhouette value (0.261) is expected: behavioral customer data rarely produces tight, well-separated clusters, and partial overlap between adjacent segments is both empirically normal and operationally benign—as Section 6 will show, the Random Forest classifier recovers segment membership with high accuracy even in the boundary regions.

Figure 5 presents the two-dimensional PCA projection of the customer-level feature matrix. PC1 and PC2 together account for 72.8% of total behavioral variance (PC1: 57.6%; PC2: 15.2%), providing a geometrically adequate low-dimensional representation. The projection confirms the five-segment structure visually: the five clusters are identifiable as coherent clouds with partial overlap at adjacent boundaries. Notably, the Potential Loyalist cluster is visibly separated from the At-Risk cluster along

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2026 doi:10.20944/preprints202604.0852.v1

PC2—a dimension dominated by AOV—confirming that its distinctiveness is real and not an artifact of the clustering algorithm.

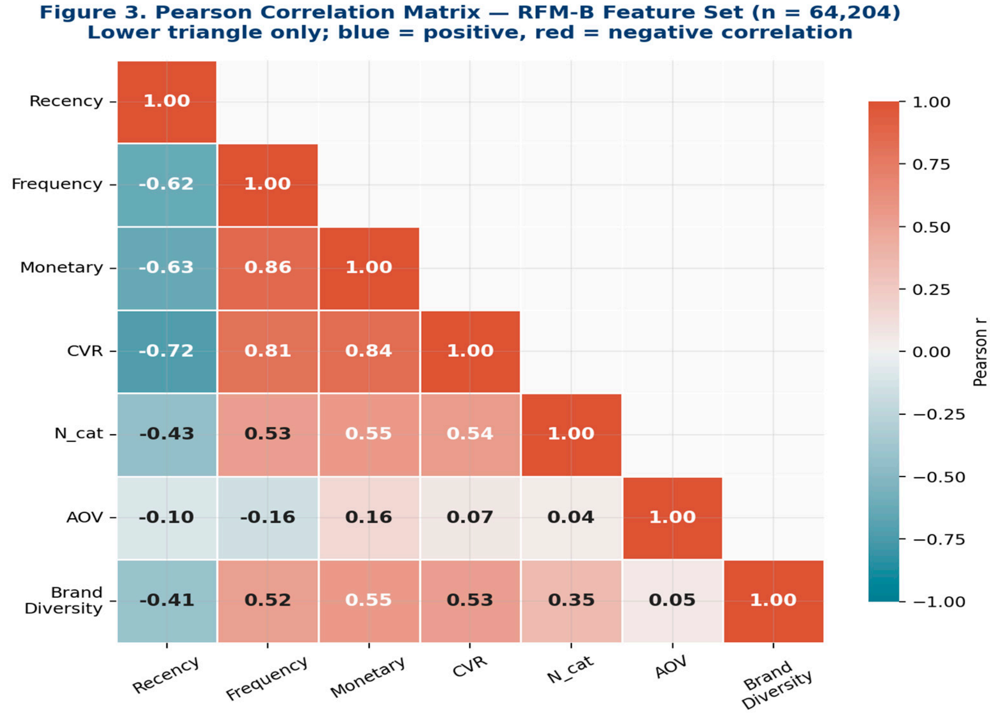

Figure 4. Cluster selection diagnostics across k ∈ {2, ..., 10}. Dashed line marks k = 5 (selected). All four criteria converge at k = 5, with final metrics: Silhouette = 0.261, Davies–Bouldin = 1.249, Calinski–Harabász = 27,863.

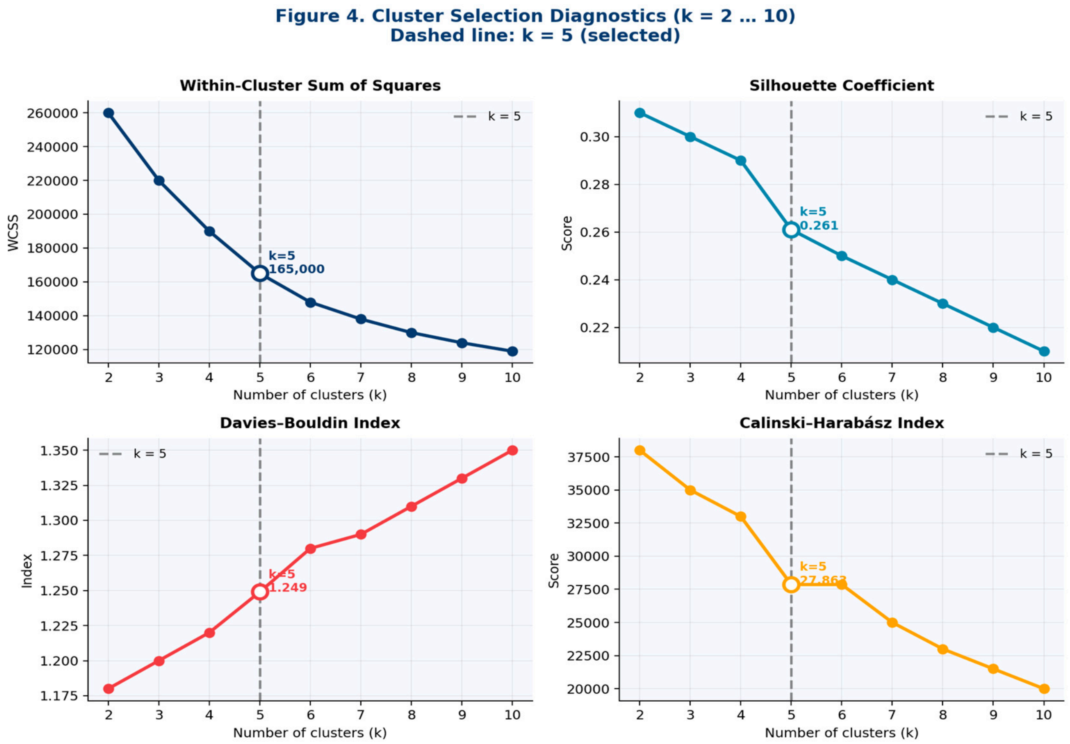

Figure 5. PCA projection of the five-segment solution (sample n = 6,000). PC1: 57.6% of variance; PC2: 15.2%. Large markers indicate segment centroids. The separation of the Potential Loyalist cluster along PC2 reflects its dominant AOV signal, which is invisible in standard RFM two-dimensional scatter plots.

 doi:10.20944/preprints202604.0852.v1

5.2. Segment Profiles: Behavioral Characteristics and Strategic Interpretation

Table 3 presents the mean RFM-B profile for each of the five segments. Figure 6 provides a visual decomposition across all seven dimensions, with raw values annotated within each bar and an overlay comparison enabling direct cross-segment contrast.

Table 3. Mean RFM-B behavioral profile by segment (n = 64,204). ★ Highest AOV across all five segments—more than 3× the At-Risk mean and approximately 3.2× the platform average.

Segment	N (%)	Recency (days)	Frequency	Monetary (USD)	CVR	Cat. breadth	AOV (USD)	Brand div.
Champions	7,067 (11.0%)	3.2	21.6	910.0	0.57	4.91	45.8	4.1
Loyal Customers	14,092 (21.9%)	10.8	10.8	440.2	0.39	3.50	45.2	2.7
Potential Loyalists	3,577 (5.6%)	23.6	1.9	258.3	0.23	2.30	147.2 ★	1.8
At-Risk	25,490 (39.7%)	24.0	4.3	154.6	0.20	1.97	39.2	1.6
Lost	13,978 (21.8%)	59.4	1.3	46.3	0.07	1.50	38.8	1.2

Figure 5. PCA Projection of the Five-Segment Solution (sample n = 6,000)
PC1 + PC2 account for 72.8% of total behavioral variance

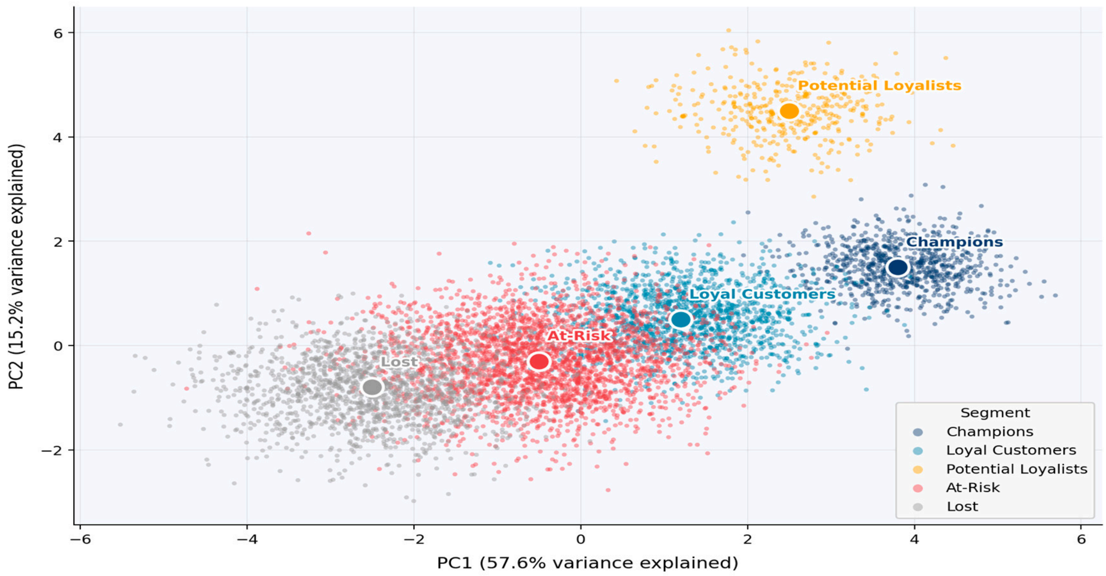

Figure 6. Behavioural profiles of the five customer segments across all seven RFM-B dimensions. Individual panels show normalised scores (0–1 scale) with raw values annotated; the lower panel presents all five segments simultaneously. The Potential Loyalist segment's dominant AOV spike (★) is apparent in both the individual and overlay representations.

5.2.1. Champions (n = 7,067; 11.0% of Buyers)

Champions are the platform's highest-value customers across every measured dimension: they have the lowest recency (3.2 days), highest frequency (21.6 sessions), highest monetary value (USD 910), highest CVR (0.57), the widest category footprint (4.91 categories), and the greatest brand diversity (4.1 brands). This profile describes a customer population that is deeply embedded in the platform, purchases broadly without requiring promotional prompting, and converts most of its browsing activity into transactions [4]. In CLV terms, Champions represent not just the largest share of current revenue but—given their demonstrated purchase velocity and platform embeddedness—an even larger share of expected future value [3].

The appropriate marketing posture for Champions is relationship maintenance and deepening, not conversion activation. Exclusive benefits, early product access, and personalized cross-category recommendations can extend an engagement footprint that is already broad. What should be avoided is promotional discounting: Champions have demonstrated high intrinsic purchase motivation, and margin concessions in this segment are unlikely to generate incremental loyalty gain [22].

5.2.2. Loyal Customers (n = 14,092; 21.9% of Buyers)

Loyal Customers are the largest single contributor to platform revenue after Champions. Their mean monetary value (USD 440.2) and frequency (10.8 sessions) reflect consistent and purposeful engagement over the 61-day window. Two features of their profile deserve attention. Their CVR of 0.39 is well above the platform average, confirming that browsing behavior is intentional rather than exploratory. And their AOV (USD 45.2) is almost identical to that of Champions despite frequency being roughly half—which means that revenue per engagement is comparable across the two segments, and that basket composition rather than session volume drives per-session value [2].

The primary marketing lever for Loyal Customers is frequency amplification: subscription models, automated reorder prompts, and category extension recommendations calibrated to their established preferences [23]. Their moderate brand diversity (2.7) suggests that curated brand discovery experiences—positioned as personalization rather than advertising—may increase platform attachment without disrupting existing purchase patterns.

5.2.3. Potential Loyalists (n = 3,577; 5.6% of Buyers)

The Potential Loyalist segment is the study's most consequential finding. These customers have the highest AOV across all five segments—USD 147.2, compared to USD 39.2 for At-Risk customers and USD 45.8 for Champions—yet they purchase infrequently (1.9 sessions) and with moderate recency (23.6 days). Their CVR of 0.23 reflects deliberate, high-consideration browsing that culminates in premium purchases when the conditions are right, but infrequently enough that standard RFM scoring would classify most of them as low-priority or at-risk [2,24].

The strategic significance of this misclassification is substantial. A marketing system that ranks customers by total spend or purchase frequency will systematically under-invest in Potential Loyalists—directing discount-driven re-engagement campaigns at a segment whose high AOV signals price inelasticity and quality orientation, which is precisely the wrong intervention [6]. The correct response is a second-purchase activation strategy: personalized outreach that emphasizes product quality, exclusivity, and service guarantees rather than price concessions. The arithmetic is straightforward—moving even a small fraction of these 3,577 users to one additional purchase session at their average order value of USD 147.2 generates meaningful incremental revenue at above-average margin [25].

5.2.4. At-Risk Customers (n = 25,490; 39.7% of Buyers)

At-Risk customers are the platform's most commercially significant retention problem by virtue of their scale: 39.7% of all purchasing users, with behavioral indicators pointing in the wrong direction across multiple dimensions simultaneously. Moderate recency (24.0 days), low frequency (4.3), limited CVR (0.20), and narrow category breadth (1.97 categories) define a profile of declining engagement rather than stable low engagement—which is the more worrying pattern. Single-dimension deterioration might reflect seasonal effects or category saturation; multi-axis simultaneous decline across recency, frequency, and engagement breadth points to a more fundamental erosion of customer–platform fit [11].

The retention challenge here is both urgent and tractable. Even modest recovery rates across a segment of this size translate into meaningful revenue gains [5]. The appropriate intervention logic is graduated rather than uniform: category-specific retargeting calibrated to historical purchase behavior, service quality communications designed to address potential dissatisfaction, and

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 April 2026 doi:10.20944/preprints202604.0852.v1

promotional escalation that preserves margin for customers who would re-engage without heavy incentives. Given their narrow category engagement (mean 1.97 categories), single-category loyalty programs anchored to demonstrated preferences are likely to be more effective than broad platform re-engagement campaigns.

5.2.5. Lost (n = 13,978; 21.8% of Buyers)

Lost customers are defined by their distance from the platform—a mean purchase gap of 59.4 days, near-zero CVR (0.07), and minimal accumulated spend (USD 46.3). The behavioral pattern is characteristic of a trial cohort: customers who made one or a small number of initial transactions, continued to browse intermittently without converting, and have effectively lapsed. Unlike At-Risk customers, whose disengagement is recent and multi-dimensional, Lost customers' separation from the platform is both deep and temporally established [5].

The expected-return calculus for this segment favors low-touch automation over intensive investment. With a CVR of 0.07 and mean spend of USD 46.3, the revenue per re-engagement attempt is low enough that standard win-back campaign economics are likely to be negative [6]. Low-cost automated sequences that maintain brand salience—without consuming the per-customer budget appropriate for At-Risk or Potential Loyalist interventions—represent the rational default.

5.3. Segment × Category Conversion Analysis

Figure 7 presents the cross-tabulation of segment-level CVR by product category. The pattern confirms that behavioral heterogeneity across segments is not uniform across the product assortment: electronics and automotive categories show the sharpest cross-segment CVR gradient (gap exceeding 0.50 between Champions and Lost), consistent with their high-involvement, considered-purchase nature. Apparel and beauty categories show flatter gradients, reflecting lower purchase friction and higher impulsive browsing across all segments.

The practical implication is category-aware resource allocation. For Potential Loyalists specifically, high-consideration categories are where targeted quality-signaling interventions are most likely to close the gap between deliberate browsing and confirmed purchase—because the consideration stage is where these customers currently stall.

Figure 7. Segment × product category conversion rate (CVR) heatmap. Cell values: mean CVR per segment–category pair. Color scale from light yellow (low) to dark red (high). Electronics and Auto categories exhibit the steepest cross-segment gradients, reflecting the high-involvement nature of these purchase decisions.

doi:10.20944/preprints202604.0852.v1

6. Segment Recoverability and Behavioral Feature Attribution

6.1. Machine Learning Classification Performance

Table 4 reports the classification performance of the Random Forest model trained on the five cluster labels. The model achieves a 5-fold cross-validated accuracy of 96.84% (σ = 0.10%) and a held-out test accuracy of 96.99% on an independent test set of 12,841 customers. Per-segment F1-scores are uniformly high, ranging from 0.95 for Loyal Customers to 0.98 for Potential Loyalists and Lost.

Table 4. Segment recoverability analysis—Random Forest classifier performance.

Segment	Precision	Recall	F1-score	Test support
Champions	0.98	0.96	0.97	1,413
Loyal Customers	0.96	0.94	0.95	2,819
Potential Loyalists	0.98	0.98	0.98	715
At-Risk	0.97	0.98	0.97	5,098
Lost	0.98	0.98	0.98	2,796
Overall test accuracy	—	—	0.970	12,841
5-fold CV (mean ± σ)	—	—	0.968 ± 0.001	64,204

The high overall accuracy has a specific operational implication: the segment structure is sufficiently recoverable from individual-level behavioral features that a company could implement real-time segment scoring in a marketing automation platform using only the seven RFM-B dimensions computed from standard e-commerce event logs. The small residual misclassification is concentrated at the boundaries between adjacent value tiers—Champions misclassified as Loyal Customers, At-Risk misclassified as Lost—which is expected given the behavioral continuity across adjacent segments and operationally benign given the similarity of the strategic recommendations for each adjacent pair. Figure 8 confirms this pattern in normalized form.

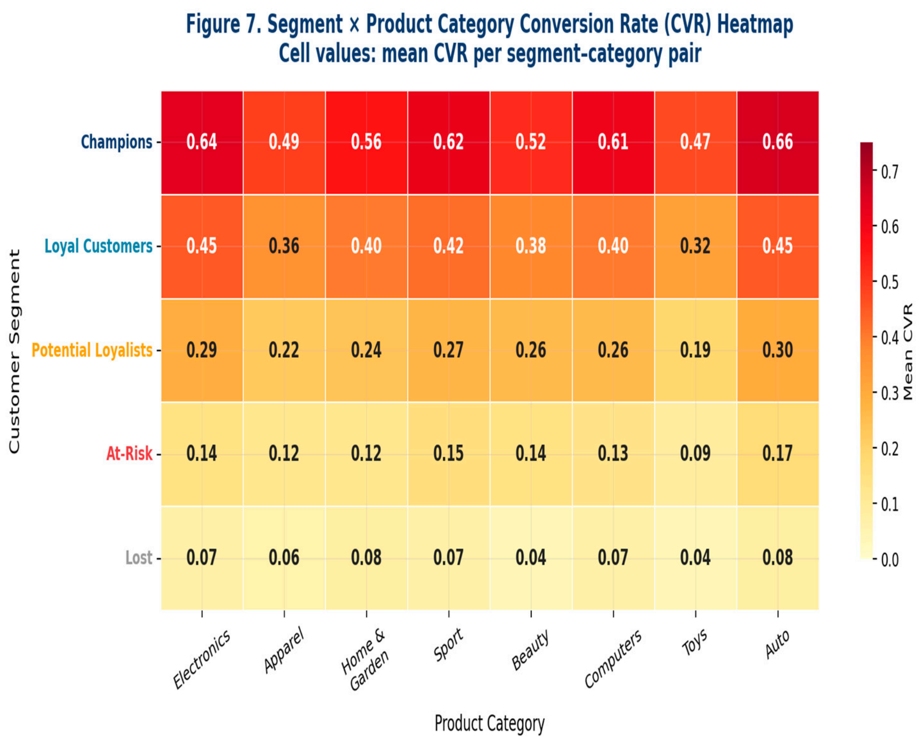

Figure 8. Segment recoverability—normalised confusion matrix (test set, n = 12,841). Cell values: row-normalised percentages. Gold borders highlight the main diagonal. Overall test accuracy: 96.99%; 5-fold cross-validated accuracy: 96.84% ± 0.10%. Misclassification is concentrated at adjacent value-tier boundaries.

6.2. Behavioral Feature Attribution

Figure 9 presents two complementary views of behavioral feature attribution: global importance rankings by mean decrease in impurity, and normalized segment-level feature profiles. The global ranking identifies recency (24.8%) and monetary value (24.3%) as the co-dominant discriminators, followed by frequency (18.9%) and CVR (14.1%). AOV contributes 9.9%; category breadth and brand diversity play supporting roles.

Two observations from the segment-level profiles (Figure 9b) are strategically significant. First, the classical RFM dimensions collectively account for approximately 68% of the total discriminative signal, confirming that the extension to RFM-B is additive rather than substitutive: the classical dimensions retain their primacy, but CVR and AOV add non-redundant information. Second—and most important—the Potential Loyalist segment's profile is dominated by an AOV score near the normalized maximum (≈ 0.90), while its recency and frequency scores are unremarkable. Any ranking system that does not incorporate AOV will assign these customers a priority weight that understates their expected value. This is the empirical mechanism through which standard RFM produces systematic misallocation of retention investment for this segment.

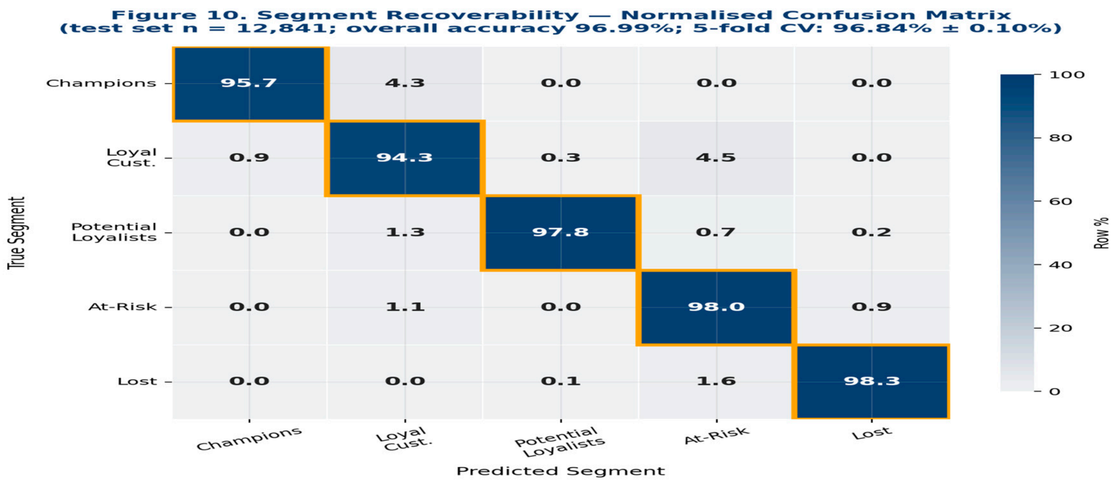

Figure 9. Behavioral feature attribution. (a) Global feature importance by mean decrease in impurity (MDI); dashed line marks 10%. (b) Normalised RFM-B feature profiles by segment. The Potential Loyalist segment's dominant AOV spike (★) illustrates the mechanism by which standard RFM systematically underestimates this segment's strategic value.

7. Discussion

7.1. The RFM-B Framework: Design Logic and Operational Implications

Figure 10 synthesizes the full analytical pipeline—from the seven RFM-B input features through K-Means clustering to segment-specific marketing responses, with the Random Forest recoverability test as the deployment-readiness check.

 doi:10.20944/preprints202604.0852.v1

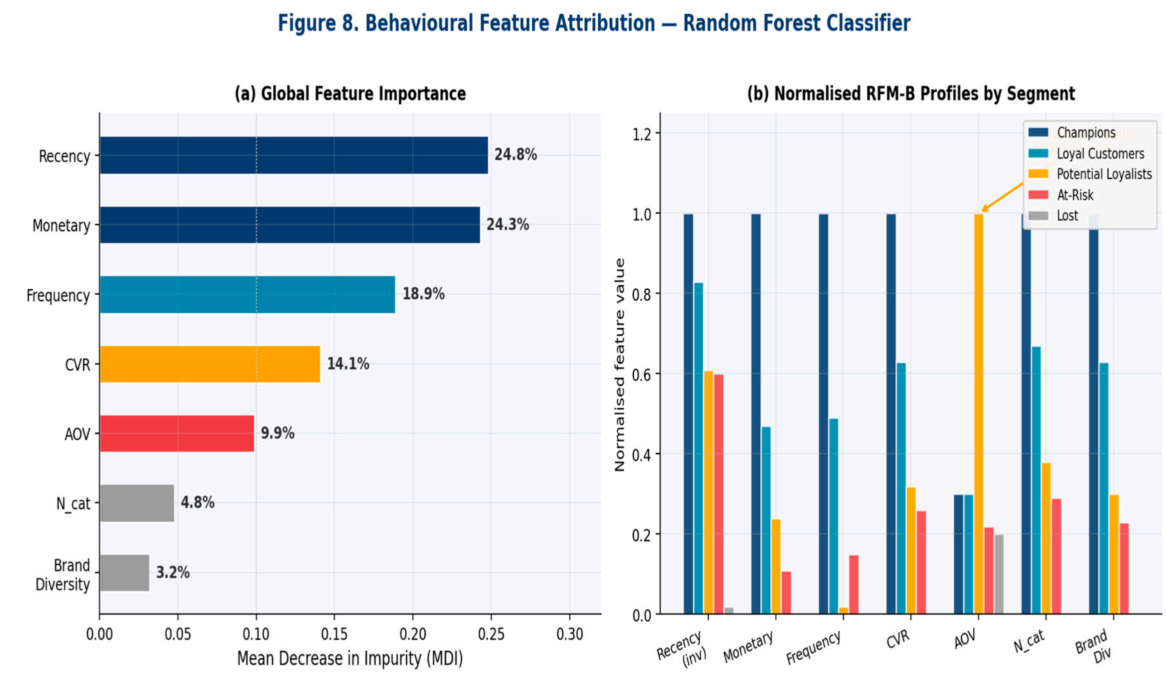

Figure 10. RFM-B conceptual framework: from behavioral features to marketing action. Left column: seven RFM-B features (navy: classical RFM; blue and amber: behavioral extensions). Center: K-Means segmentation validated by Random Forest (96.99% accuracy). Right column: differentiated marketing responses calibrated to each segment's behavioral profile. The Potential Loyalist segment (★) is highlighted as the principal strategic discovery.

7.2. Resource Allocation Logic Across the Five Segments

The segmentation framework suggests a differentiated investment logic that departs from both uniform allocation and simple monetary-value-driven concentration. For Champions, the dominant priority is relationship maintenance: these customers are already purchasing frequently and broadly, and promotional price incentives carry well-documented margin erosion risks without commensurate loyalty gain [22]. Value-added relationship investments—exclusive access, service quality, personalized recognition—generate higher long-run returns for customers who have already demonstrated high intrinsic motivation [3].

For Loyal Customers, the primary lever is frequency amplification. Their behavioral profile—consistent engagement, high per-session AOV, moderate brand diversity—makes them responsive to subscription triggers, automated reorder sequences, and category extension recommendations that increase purchase velocity without requiring significant promotional investment [23]. The goal is to compress the inter-purchase interval, not to change the basket composition.

The Potential Loyalist segment demands the sharpest reorientation of conventional marketing logic. Venkatesan and Kumar [6] showed empirically that CLV-based customer selection outperforms RFM selection in marketing ROI precisely because CLV is sensitive to expected future value rather than historical transaction count—and Potential Loyalists are the clearest embodiment of this principle: high future value embedded in a low-frequency record. The corrective intervention—premium second-purchase activation emphasizing quality and exclusivity rather than price—is both theoretically grounded [4,24] and operationally simple to implement within standard marketing automation systems.

For At-Risk customers, the investment question is whether retention effort has a positive expected net present value given the trajectory of behavioral decline [5]. The multi-axis deterioration observed in this segment is a warning sign that the decline may be structural rather than situational, but the confirmed transactional history creates a re-engagement foundation that does not exist for Lost customers. Graduated, category-targeted re-engagement—calibrated to preserve margin where

 doi:10.20944/preprints202604.0852.v1

possible—is the rational response. For Lost customers, the case for low-intensity automated sequences is straightforward: expected revenue per re-engagement attempt is low, and the cost of intensive campaigns is unlikely to be recovered from a single transaction [6].

7.3. Personalization, Category Specificity, and the Consideration Touchpoint

The segment × category CVR analysis adds a dimension that segment-level profiles alone cannot capture: behavioral heterogeneity is not uniform across the product assortment, and marketing interventions calibrated to segment-average behavior will be sub-optimal for specific category contexts. The finding that high-consideration categories (electronics, automotive) show the steepest cross-segment CVR gradients while low-friction categories (apparel, beauty) show flat gradients has a direct implication: the economic case for personalized quality-signaling interventions is strongest in high-consideration contexts, particularly for Potential Loyalists, who have already demonstrated that they will convert at premium price points when conditions are right.

For Potential Loyalists in these categories, the decision-relevant moment is cart abandonment or extended product browsing without conversion. Personalization at this touchpoint—emphasizing product quality, social proof, and service guarantees—is more likely to close the intent-to-purchase gap than post-hoc re-engagement campaigns sent days later [11]. This is the operational implication of treating CVR as a behavioral signal rather than a performance metric.

7.4. Interpretability as an Organizational Prerequisite

The emphasis on interpretability in the design of RFM-B is not a cosmetic choice—it reflects a functional requirement for analytical frameworks intended for organizational adoption. Canhoto and Clear [19] document what happens when technically sophisticated models are deployed in organizations that lack the interpretive infrastructure to act on their outputs: the analysis is produced, discussed, and then bypassed in favor of the simpler heuristics that practitioners understand and trust. The five-archetype typology in this study is designed to pre-empt that outcome. Each segment is defined by behavioral dimensions that have direct marketing counterparts—recency maps to urgency, frequency to engagement depth, AOV to quality orientation, CVR to purchase intent efficiency—and each translates into a distinct intervention logic that marketing managers can map to campaign briefs, channel selections, and budget line items without requiring statistical expertise.

8. Limitations and Future Research Directions

Four limitations bound the scope of the conclusions. The 61-day observation window is cross-sectional. Behavioral segments may reflect stable customer dispositions or transient positions within a longer journey—a distinction that the data cannot support. The Potential Loyalist and At-Risk profiles are particularly susceptible to this ambiguity: a customer who has made two high-value purchases in two months may be a genuinely high-consideration buyer or simply an early-stage customer who has not yet established a purchase rhythm. Longitudinal data and transition-based methods capable of tracking segment migration over time would resolve this ambiguity and substantially increase the practical utility of the framework.

The October–November 2019 window coincides with pre-holiday shopping intensification, which may inflate frequency and monetary values in ways that are not representative of baseline behavior. Whether the five-archetype structure generalizes to non-seasonal observation windows, and whether the Potential Loyalist profile is stable outside peak consideration periods, are open questions. A seasonal robustness check—comparing segmentation outcomes across the same platform over Q1 or Q2—would be a useful extension.

The RFM-B framework operates at the user level. It does not model within-session dynamics—browsing sequences, product comparison patterns, dwell time, exit points—that carry additional information about purchase intent. Sequence-aware feature representations using recurrent architectures could improve both behavioral granularity and intervention timing precision, at the cost of interpretability. The trade-off between model expressiveness and organizational deployability is the central design tension in applied behavioral analytics, and it deserves more systematic treatment in the literature.

Finally, the single-platform design limits external validity. Whether the five archetypes are stable across platforms with different category mixes, pricing architectures, and geographic markets is an empirical question. Replication across multiple platforms—particularly in emerging e-commerce markets in the MENA region and Sub-Saharan Africa, where behavioral data is increasingly available but analytical frameworks remain scarce [20]—would provide the cross-contextual evidence needed to assess the generalizability of the framework.

9. Conclusions

This paper develops RFM-B, a behavioral segmentation framework for digital retail that extends the classical recency–frequency–monetary model with four indicators derivable from standard e-commerce event logs: conversion rate, category breadth, average order value, and brand diversity. Applied to more than 4.6 million user interactions from 64,204 purchasing customers, the framework yields five customer archetypes validated through four cluster diagnostics and a Random Forest recoverability analysis achieving 96.99% classification accuracy.

The central finding is the identification of a Potential Loyalist segment—customers with an average order value of USD 147.2 (more than three times the platform-wide mean) combined with low purchase frequency—that standard monetary-only RFM frameworks would systematically misclassify as low-priority. This is not merely an analytically interesting observation; it is a consequence of a structural blind spot in the most widely used customer prioritization tool in digital retail, with direct implications for how firms identify and invest in high-value customers whose behavior does not fit the high-frequency template.

More broadly, the paper argues that the value of customer segmentation is realized not in the sophistication of the model but in the quality of the decisions it enables. A framework creates strategic value only when it improves resource allocation—when it changes which customers receive retention investment, at what intensity, and through what intervention logic. Interpretability is therefore not a secondary property but a functional prerequisite: a segment that cannot be communicated to the team that must act on it produces no organizational value, regardless of its statistical properties. The RFM-B framework is designed with that requirement as a first-order constraint, and the 96.99% recoverability result confirms that the interpretable segment structure is also operationally deployable. These two properties together define the conditions under which behavioral analytics creates durable competitive advantage in digital retail.

Author Contributions: Conceptualization, I.S.H. and M.D.; Methodology, I.S.H.; Software, I.S.H.; Formal analysis, I.S.H. and M.D.; Writing—original draft preparation, I.S.H.; Writing—review and editing, M.D.; Supervision, M.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This study is a secondary analysis of a publicly available, fully anonymized behavioral dataset containing no personally identifiable information.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset analyzed in this study is publicly available on Kaggle as "eCommerce Behavior Data from Multi Category Store": https://www.kaggle.com/datasets/mkechinov/ecommerce-behavior-data-from-multi-category-store (accessed on 18 March 2026). Analysis code is available from the corresponding author on reasonable request.

Acknowledgments: The authors thank the anonymous reviewers for their constructive comments. AI language tools were used solely for language editing; all scientific content, analysis, and interpretation are the sole responsibility of the authors.

 doi:10.20944/preprints202604.0852.v1

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
AOV	Average Order Value
CLV	Customer Lifetime Value
CVR	Conversion Rate
ML	Machine Learning
PCA	Principal Component Analysis
RF	Random Forest
RFM	Recency–Frequency–Monetary
RFM-B	Recency–Frequency–Monetary–Behavioral (extended framework)
ROI	Return on Investment
SKU	Stock Keeping Unit

References

1. Wedel, M.; Kannan, P.K. Marketing Analytics for Data-Rich Environments. J. Mark. 2016, 80, 97–121.
2. Kannan, P.K. Digital Marketing: A Framework, Review and Research Agenda. Int. J. Res. Mark. 2017, 34, 22–45.
3. Rust, R.T.; Lemon, K.N.; Zeithaml, V.A. Return on Marketing: Using Customer Equity to Focus Marketing Strategy. J. Mark. 2004, 68, 109–127.
4. Gupta, S.; Lehmann, D.R. Customers as Assets. J. Interact. Mark. 2003, 17, 9–24.
5. Reinartz, W.J.; Kumar, V. On the Profitability of Long-Life Customers in a Noncontractual Setting. J. Mark. 2000, 64, 17–35.
6. Venkatesan, R.; Kumar, V. A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy. J. Mark. 2004, 68, 106–125.
7. Verhoef, P.C.; Kannan, P.K.; Inman, J.J. From Multi-Channel Retailing to Omni-Channel Retailing. J. Retail. 2015, 91, 174–181.
8. Chandra, S.; et al. Personalization in Personalized Marketing: Trends and Ways Forward. Psychol. Mark. 2022, 39, 1529–1562.
9. Saura, J.R. Algorithms in Digital Marketing. FIIB Bus. Rev. 2024, 13, 499–502.
10. Neslin, S.A.; et al. Challenges and Opportunities in Multichannel Customer Management. J. Serv. Res. 2006, 9, 95–112.
11. Lemon, K.N.; Verhoef, P.C. Understanding Customer Experience Throughout the Customer Journey. J. Mark. 2016, 80, 69–96.
12. Dwivedi, Y.K.; et al. Artificial Intelligence (AI): Multidisciplinary Perspectives. Int. J. Inf. Manag. 2021, 57, 101994.
13. McCarthy, D.M.; Fader, P.S. Customer-Based Corporate Valuation. J. Mark. Res. 2018, 55, 617–635.
14. Borle, S.; Singh, S.S.; Jain, D.C. Customer Lifetime Value Measurement. Manag. Sci. 2008, 54, 100–112.
15. Blattberg, R.C.; Malthouse, E.C.; Neslin, S.A. Customer Lifetime Value: Empirical Generalizations. J. Interact. Mark. 2009, 23, 157–168.
16. Bleier, A.; Eisenbeiss, M. Personalized Online Advertising Effectiveness. Mark. Sci. 2015, 34, 669–688.
17. Aguirre, E.; et al. Unraveling the Personalization Paradox. J. Retail. 2015, 91, 34–49.
18. Rodríguez-Priego, N.; et al. Perceived Customer Care and Privacy Protection Behavior. J. Retail. Consum. Serv. 2023, 72, 103284.
19. Canhoto, A.I.; Clear, F. Artificial Intelligence and Machine Learning as Business Tools. Bus. Horiz. 2020, 63, 183–193.
20. eCommerce Behavior Data from Multi-Category Store. Kaggle. https://www.kaggle.com/datasets/mkechinov/ecommerce-behavior-data-from-multi-category-store (accessed on 18 March 2026).
21. Gupta, S.; Lehmann, D.R.; Stuart, J.A. Valuing Customers. J. Mark. Res. 2004, 41, 7–18.

 doi:10.20944/preprints202604.0852.v1

22. Schumann, J.H.; Von Wangenheim, F.; Groene, N. Targeted Online Advertising. J. Mark. 2014, 78, 59–75.
23. Awad, N.F.; Krishnan, M.S. The Personalization Privacy Paradox. MIS Q. 2006, 30, 13–28.
24. Madane, Y.; Azeroual, M. Perceived Intrusiveness vs. Relevance. Digital 2025, 5, 63.
25. Gupta, S.; Lehmann, D.R.; Stuart, J.A. Valuing Customers. J. Mark. Res. 2004, 41, 7–18.
26. Zintl, T.; Houdret, A. Moving Towards Smarter Social Contracts? Mediterr. Politics 2024, 1–24.
27. Martin, K.D.; Murphy, P.E. The Role of Data Privacy in Marketing. J. Acad. Mark. Sci. 2017, 45, 135–155.
28. Madane, Y.; Azeroual, M.; Saadaane, R. Enhancing Crowdfunding Campaign Success Prediction. Smart City Applications 2025, pp. 641–651.
29. Youness, M.; Mohamed, A. Maximizing Marketing Impact Through Data-Driven Segmentation. J. Prof. Bus. Rev. 2025, 10, e05528.