

Effective customer segmentation using the recency, frequency, and monetary framework, combined with density-based clustering

Pham Van Hau, Thomas McIver, Kirsten Robertson, Park Thaichon

ABSTRACT

Customer segmentation underpins Customer Relationship Management (CRM) and growth, yet purchase patterns are often sparse and irregular. This study addresses this gap by integrating an extended RFMI framework (Recency, Frequency, Monetary, and Interpurchase) with the density-based HDBSCAN algorithm and applying it to an H&M transactions dataset. The approach detects non-spherical structure and permits a 'noise' label for irregular shoppers. This study derives RFMI features, standardises inputs, and estimates segments with HDBSCAN. The solution yields five segments: Low-Value Inactive, Low-Value Dormant, Mid-Value Occasional, Loyal Mid-tier, and Premium Champions, plus a small noise group (n = 21,011). Each group displayed unique recency, frequency, monetary value, and interpurchase profiles. Product analysis shows shared preferences for upper- and lower-body garments, with high-value customers engaging more broadly. Managerial implications include sharper retention allocation, targeted reactivation, and assortment/promotion design aligned to segmentation and price sensitivity. The RFMI+HDBSCAN pipeline offers a scalable alternative that improves segment fidelity in real-world retail data.

KEYWORDS

customer segmentation; density-based clustering; behavioural frameworks
