

ABSTRAK

Peningkatan penggunaan *e-commerce* telah mengubah cara konsumen berbelanja dan menuntut pelaku usaha untuk memahami perilaku pelanggan secara lebih mendalam. Toko Buku Togamas Affandi Yogyakarta sebagai salah satu pelaku bisnis ritel turut menghadapi tantangan dalam mengidentifikasi dan mengelompokkan pelanggan secara efektif. Selama ini, proses identifikasi pelanggan masih dilakukan secara manual dan belum memanfaatkan potensi data transaksi yang tersedia. Ketiadaan segmentasi pelanggan yang sistematis menyebabkan strategi pemasaran yang dijalankan kurang terarah dan belum sepenuhnya berbasis data. Oleh karena itu, dilakukan penerapan segmentasi pelanggan berbasis data historis guna mendukung perumusan strategi pemasaran yang lebih efektif dan efisien serta meningkatkan hubungan jangka panjang dengan pelanggan.

Penelitian ini menggunakan pendekatan CRISP-DM yang terdiri dari enam tahapan utama: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment* dalam bentuk penyusunan rekomendasi strategi pemasaran. Data yang dianalisis merupakan data transaksi pelanggan dari *platform e-commerce* Shopee pada periode 1 Januari hingga 20 Juni 2025. Setelah dilakukan pembersihan data, dihitung nilai *Recency*, *Frequency*, dan *Monetary* (RFM), serta dilakukan standarisasi menggunakan *Min-Max Scaling*. Proses segmentasi dilakukan dengan algoritma *K-Means Clustering*, sedangkan jumlah *Cluster* optimal ditentukan dengan metode *Elbow* dan *Silhouette Coefficient*.

Hasil penelitian menunjukkan pelanggan yang dikelompokkan menjadi dua segmen utama, yaitu *Recent Customers*, dan *Dormant Customers*. Evaluasi model menggunakan nilai *Sum of Squared Error* (SSE) menunjukkan hasil terbaik saat *Cluster* berjumlah dua, dengan penurunan nilai yang signifikan. Selain itu, nilai *Silhouette Coefficient* sebesar 0,64 menunjukkan bahwa segmentasi yang dihasilkan cukup baik dan representatif. Setiap segmen dianalisis untuk menyusun profil pelanggan dan menghasilkan rekomendasi strategi pemasaran yang lebih personal dan tepat sasaran. Melalui proses analisis, dihasilkan pula pemetaan lokasi pelanggan berdasarkan distribusi transaksi, yang dapat digunakan oleh toko untuk menetapkan wilayah prioritas pemasaran secara lebih strategis. Temuan ini memberikan kontribusi nyata terhadap pengambilan keputusan bisnis berbasis data.

Kata Kunci: Segmentasi Pelanggan, RFM, *K-Means*, Pemasaran

ABSTRACT

The increasing use of e-commerce has transformed consumer shopping behavior and demands that businesses better understand customer behavior. Toko Buku Togamas Affandi Yogyakarta, as one of the retail business actors, faces challenges in effectively identifying and segmenting its customers. Until now, customer identification has been carried out manually and has not utilized the potential of available transaction data. The absence of systematic customer segmentation has resulted in marketing strategies that are less targeted and not fully data-driven. Therefore, this study was conducted to implement customer segmentation based on historical transaction data to support the formulation of more effective and efficient marketing strategies, as well as to enhance long-term customer relationships.

This study adopts the CRISP-DM approach, which consists of six main stages: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment in the form of marketing strategy recommendations. The data analyzed includes customer transaction data from the Shopee e-commerce platform during the period from January 1 to June 20, 2025. After the data was cleaned, Recency, Frequency, and Monetary (RFM) values were calculated, followed by standardization using the Min-Max Scaling method. The segmentation process was carried out using the K-Means Clustering algorithm, with the optimal number of Clusters determined using the Elbow Method and Silhouette Coefficient.

The research results show that customers can be grouped into two main segments: Recent Customers and Dormant Customers. Model evaluation using the Sum of Squared Error (SSE) showed the best results when the number of Clusters was two, with a significant drop in value. In addition, the Silhouette Coefficient score of 0,64 indicates that the segmentation produced is reasonably good and representative. Each segment was analyzed to build customer profiles and generate more personalized and targeted marketing strategy recommendations. The study also produced a customer location map based on transaction distribution, which can be used by the store to determine marketing priority areas more strategically. These findings provide a practical contribution to data-driven business decision-making.

Keywords: Customer Segmentation, RFM, K-Means, Marketing