

## ABSTRAK

Segmentasi pelanggan mengacu pada proses membagi kelompok konsumen yang besar menjadi kelompok-kelompok yang lebih kecil dengan berbagai karakteristik berdasarkan atribut sosial, perilaku, dan atribut-atribut konsumsi mereka. Penelitian ini bertujuan untuk meningkatkan kinerja algoritma *K-Means* yang sensitif terhadap *outlier* dengan menerapkan metode *Interquartile-Range* (IQR) guna mengatasi *outlier*. Penelitian ini menggunakan data pelanggan *E-Commerce* Shopee dari bulan Oktober 2023 hingga Januari 2024 sebanyak 419 data, dengan atribut seperti umur, *recency*, *frequency*, dan *monetary*.

Metode penelitian yang digunakan yaitu Metode *Interquartile-Range* (IQR), metode ini diterapkan untuk mengatasi *outlier* dalam proses segmentasi pelanggan menggunakan algoritma *K-Means*. IQR adalah teknik statistik yang digunakan untuk mengidentifikasi dan menangani nilai-nilai yang ekstrem atau tidak biasa dalam kumpulan data.

Hasil penelitian menunjukkan bahwa penerapan IQR berhasil meningkatkan kinerja *K-Means*, dengan peningkatan *silhouette score* sebesar 3.81% dibandingkan dengan model tanpa IQR. Hal ini menghasilkan struktur *cluster* yang lebih jelas dan akurat, terbagi menjadi empat kelompok yang dapat diidentifikasi dengan warna merah, hijau, biru, dan kuning. Kesimpulan dari penelitian ini adalah bahwa penggunaan metode IQR efektif dalam meningkatkan akurasi segmentasi pelanggan pada algoritma *K-Means*, yang berpotensi meningkatkan strategi pemasaran dan kepuasan pelanggan.

**Kata kunci:** Segmentasi pelanggan, *k-means*, *interquartile-range*, *outlier*, *silhouette score*

## ABSTRACT

*Customer segmentation refers to the process of dividing a large group of consumers into smaller groups with various characteristics based on their social, behavioral and consumption attributes. This research aims to improve the performance of the K-Means algorithm which is sensitive to outliers by applying the Interquartile-Range (IQR) method to overcome outliers. This research uses 419 Shopee E-Commerce customer data from October 2023 to January 2024, with attributes such as age, recency, frequency and monetary.*

*The research method used is the Interquartile-Range (IQR) method, this method is applied to overcome outliers in the customer segmentation process using the K-Means algorithm. IQR is a statistical technique used to identify and treat extreme or unusual values in a data set.*

*The research results show that implementing IQR succeeded in improving K-Means performance, with an increase in silhouette score of 3.81% compared to the model without IQR. This results in a clearer and more accurate cluster structure, divided into four groups that can be identified by the colors red, green, blue and yellow. The conclusion of this research is that the use of the IQR method is effective in increasing the accuracy of customer segmentation in the K-Means algorithm, which has the potential to improve marketing strategies and customer satisfaction.*

**Keywords:** *Customer segmentation, k-means, interquartile-range, outlier, silhouette score*