Application of the k-Means Algorithm in Grouping Skincare Purchasing Patterns in Tokopedia E-Commerce

Nuzula Afini

Data Science Study Program, Faculty of Business and Information Technology
University of Technology Yogyakarta
Jl. Ringroad Utara Jombor Sleman Yogyakarta
E-mail : nuzula.afini@gmail.com

ABSTRACT

This study analyzes the effectiveness of skincare product promotion strategies on Tokopedia e-commerce using the k-Means algorithm to cluster purchase data. The k-Means algorithm helps identify skincare purchasing patterns in the Tokopedia e-commerce dataset. Data is processed through stages such as data import, exploratory data analysis (eda), data preparation, data visualization, advanced exploration, and data preparation for the k-Means algorithm. Evaluation using inertia values and Silhouette Score shows that clustering with two clusters (k = 2) provides better separation compared to three or four clusters. The results of the analysis show that discount prices and promotions have a significant impact on increasing skincare product sales, especially for brands such as Skintific, Wardah, and Npure. Product sales are stable at over 1000 units per month with peak sales in January. The optimal price range for skincare products is between IDR 79,500 to IDR 569,000, with optimal discounts in the range of 1-30% and cashback below IDR 30,000. The main factor influencing clustering is the discount percentage. Products with high discount prices (above 78%) tend to fall into cluster 0, while products with discount prices below 30% tend to fall into cluster 1. Products with high initial prices also fall into cluster 0 more often. A large discount strategy has proven effective in attracting buyers despite high final prices, while low discounts on affordable products also increase sales. Recommendations for Tokopedia e-commerce include periodic evaluation of promotional strategies using a data-driven approach, improving service quality, developing region-specific promotional strategies, and utilizing temporal analysis for more effective promotional planning. The results of this analysis provide an overview of consumer profiles and preferences.

Keywords: k-Means, skincare purchasing patterns, Tokopedia e-commerce, data-driven.