

ABSTRAK

Analisis cluster merupakan metode yang penting dalam memahami pola dan hubungan dalam data, khususnya dalam konteks sosial ekonomi. Penelitian ini menggunakan dua pendekatan, yaitu K-Means dan DBSCAN (Density-Based Spatial Clustering of Applications with Noise), untuk menganalisis kondisi sosial ekonomi di Kabupaten Aceh Utara. Data yang digunakan mencakup variabel tingkat pendidikan (TPD), tingkat kesehatan (TK), kepadatan penduduk (KP), dan data perumahan (PR) dari 27 kecamatan selama periode 2021-2023.

Metode K-Means, yang berbasis pada pembentukan centroid, memungkinkan pengelompokan data berdasarkan kedekatan nilai variabel. Namun, metode ini memiliki keterbatasan dalam mendeteksi noise dan outlier. Di sisi lain, DBSCAN menawarkan keunggulan dalam mengidentifikasi cluster dengan kepadatan tinggi dan dapat mendeteksi noise, tanpa memerlukan penentuan jumlah cluster sebelumnya. Dengan menggunakan Microsoft Excel untuk perhitungan manual dan Python untuk implementasi sistem, penelitian ini bertujuan untuk memberikan wawasan yang lebih mendalam mengenai kondisi sosial ekonomi di daerah tersebut.

Hasil analisis menunjukkan bahwa kedua metode memberikan informasi yang berbeda namun saling melengkapi. K-Means cenderung lebih efisien dalam pengelompokan data yang terstruktur, sementara DBSCAN lebih efektif dalam menangani data yang memiliki noise. Penelitian ini diharapkan dapat menjadi acuan bagi pengambil keputusan dalam merumuskan kebijakan yang lebih tepat sasaran di Kabupaten Aceh Utara.

Kata kunci: *Analisis Cluster, K-Means, DBSCAN, Kondisi Sosial Ekonomi, Kabupaten Aceh Utara.*

ABSTRACT

Cluster analysis is an important method for understanding patterns and relationships in data, especially in a socio-economic context. This research uses two approaches, K-Means and DBSCAN (Density-Based Spatial Clustering of Applications with Noise), to analyze the socio-economic conditions in Aceh Utara Regency. The data used includes variables such as education level (TPD), health level (TK), population density (KP), and housing data (PR) from 27 districts over the period 2021-2023.

The K-Means method, which is based on the formation of centroids, allows for data clustering based on the proximity of variable values. However, this method has limitations in detecting noise and outliers. On the other hand, DBSCAN offers the advantage of identifying clusters with high density and can detect noise without the need to predefine the number of clusters. By using Microsoft Excel for manual calculations and Python for system implementation, this study aims to provide deeper insights into the socio-economic conditions of the region.

The analysis results indicate that both methods offer different yet complementary information. K-Means tends to be more efficient in clustering structured data, while DBSCAN is more effective in handling data with noise. This research is expected to serve as a reference for policymakers in formulating more targeted policies in Aceh Utara Regency.

Keywords: *Cluster Analysis, K-Means, DBSCAN, Socio-Economic Conditions, Aceh Utara Regency.*