

ABSTRAK

Angka buta huruf merupakan proporsi penduduk usia lima belas tahun ke atas yang tidak dapat membaca dan menulis huruf latin dan huruf lainnya terhadap total jumlah penduduk usia lima belas tahun ke atas. Banyaknya penduduk yang masih mengalami buta huruf di suatu negara merupakan salah satu indikator yang menunjukkan negara tersebut masih belum maju. Provinsi Jawa Timur merupakan salah satu provinsi di Indonesia dengan angka buta huruf yang tinggi sebesar 6,8% pada tahun 2022. Penelitian ini bertujuan untuk melakukan pengelompokan kabupaten/kota di Provinsi Jawa Timur berdasarkan faktor yang mempengaruhi angka buta huruf. Metode pengelompokan yang digunakan pada penelitian ini adalah metode *K-Means Clustering* dan *K-Medoids Clustering*. Metode perbandingan yang digunakan yaitu metode *Silhouette Coefficient Index*. Berdasarkan nilai *silhouette Coefficient Index*, diperoleh bahwa pada metode *K-Means* dan *K-Medoids* jumlah *cluster* optimum yaitu lima buah *cluster*. Berdasarkan nilai *Silhouette Coefficient Index* juga diperoleh bahwa metode yang lebih baik pada penelitian ini adalah metode *K-Means* karena nilai *Silhouette Coefficient Index* yang diperoleh lebih tinggi dari metode *K-Medoids*.

Kata Kunci : Angka Buta Huruf, *K-Means*, *K-Medoids*.

ABSTRACK

The illiteracy rate is the proportion of the population aged fifteen years and over who cannot read and write Latin letters and other letters to the total population aged fifteen years and over. The large number of people who are still illiterate in a country is an indicator that shows that the country is still not developed. East Java Province is one of the provinces in Indonesia with a high illiteracy rate of 6.8% in 2022. This research aims to group districts/cities in East Java Province based on factors that influence the illiteracy rate. The grouping methods used in this research are the K-Means Clustering and K-Medoids Clustering methods. The comparison method used is the Silhouette Coefficient Index method. Based on the Silhouette Coefficient Index value, it was found that in the K-Means and K-Medoids methods the optimum number of clusters was five clusters. Based on the Silhouette Coefficient Index value, it was also found that the better method in this research was the K-Means method because the Silhouette Coefficient Index value obtained was higher than the K-Medoids method.

Keywords: *Illiteracy Rate, K-Means, K-Medoids.*