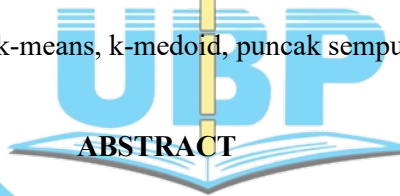


## ABSTRAK

Setiap pengelola objek wisata melakukan pendataan kunjungan wisatawan, pendataan ini sangat diperlukan dalam menentukan target promosi. Puncak sempur merupakan salah satu objek wisata alam di Karawang yang melakukan pendataan kunjungan wisatawan melalui buku pengunjung, namun tidak mudah dalam menentukan daerah kurang peminat jika dilihat dari buku daftar pengunjung. Permasalahan ini dapat diselesaikan dengan menggunakan salah satu teknik *data mining* yaitu Algoritma k-means *clustering*. Algoritma yang digunakan dalam penelitian ini dapat membantu pengelola wisata alam puncak sempur untuk mengetahui daerah mana saja yang harus menjadi target promosi sehingga dapat membuat strategi yang pasti. Variabel yang dibutuhkan dalam *data mining* ini berupa asal daerah pengunjung dan jumlah pengunjung, yang di mana data tersebut dapat ditemui pada buku daftar pengunjung. Data yang didapat terdiri dari 1.797 wisatawan dan 32 daerah akan dihitung manual dan bahasa python menggunakan algoritma k-means dan dibandingkan dengan perhitungan algoritma k-medoid, hasil perhitungan algoritma k-means dan k-medoid berbeda sehingga di lakukan evaluasi untuk mengetahui tingkat akurasi dengan menggunakan metode *Silhouette coefficient*. Hasil evaluasi menunjukan algoritma k-means lebih baik akurasi daripada algoritma k-medoid untuk daerah asal wisatawan yang tergolong ramai (C0) berjumlah 1 daerah, daerah asal wisatawan yang tergolong sedang (C1) berjumlah 1 daerah, dan daerah asal wisatawan yang tergolong sepi (C2) berjumlah 28 daerah.

**Kata Kunci:** *Clustering, Data mining, k-means, k-medoid, puncak sempur.*



*The manager of every tourist attraction collects data on tourist visits, this data collection is very necessary for determining promotion targets. Puncak Sempur is one of the natural appeals in Karawang that collects data on tourist visits through the visitor book, but it is not easy to determine which areas are less interested if seen from the visitor list book. This problem can be solved by using one of the data mining techniques, namely the k-means clustering algorithm. The algorithm used in this study can help the manager of Puncak Sempur natural tourism to find out which areas should be the target of promotion so that they can make a definite strategy. The variables needed in this data mining are the origin of the visitor area and the number of visitors, where the data can be found in the visitor list book. The data obtained consists of 1.797 tourists and 32 regions will be calculated manually and python language using the k-means algorithm and compared with the calculation of the k-medoid algorithm, the results of the calculation of the k-means and k-medoid algorithms are different so an evaluation is carried out to determine the level of accuracy with using the Silhouette coefficient method. The evaluation results show that the k-means algorithm has better accuracy than the k-medoid algorithm for the tourist origin area which is classified as crowded (C0), totaling 1 area, the tourist origin area classified as moderate (C1) totaling 1 regions, and the tourist origin area being classified as quiet (C1). C2) totaling 32 regions.*

**Keyword:** *Clustering, Data mining, k-means, k-medoid, puncak sempur.*