

Madani: Jurnal Ilmiah Multidisiplin

Volume 3, Nomor 3, April 2025, P. 242-248

E-ISSN: 2986-6340

Licensed by CC BY-SA 4.0

DOI: <https://doi.org/10.5281/zenodo.15276418>

Penggunaan Metode K-Means Clustering Pemetaan dan Klasterisasi Tempat Wisata di Kabupaten Deli Serdang

Chelsea Beatrice¹, Antonia Ingrid², Kezia Theodora Sinaga³, Albert Samuel Sinaga⁴

^{1, 2, 3, 4}Program Studi Statistika, FMIPA, Universitas Negeri Medan

Email: chelseabeatrice07@gmail.com¹, antoniaingrid1005@gmail.com²,
theodorakezias@gmail.com³, albertsamuels02@gmail.com⁴

Abstrak

Masalah yang sering muncul adalah kurangnya sistem yang jelas untuk mengkategorikan dan mengklasifikasikan tempat wisata berdasarkan karakteristik geografis. Perencanaan rute wisata, pembangunan infrastruktur pendukung, dan strategi promosi yang tepat sasaran semuanya menjadi lebih sulit karena hal ini. Akibatnya, pendekatan berbasis data diperlukan untuk secara lebih sistematis menganalisis dan memetakan tempat wisata di Kabupaten Deli Serdang. Penelitian ini bertujuan untuk melakukan pemetaan dan klasterisasi tempat wisata di Kabupaten Deli Serdang berdasarkan karakteristik spasial dan harga tiket masuk menggunakan metode K-Means Clustering. Metode ini digunakan untuk mengelompokkan lokasi wisata ke dalam beberapa kategori berdasarkan kesamaan atribut, yang meliputi koordinat geografis (latitude dan longitude), harga tiket, serta jarak ke pusat kota sebagai acuan lokasi. Data dianalisis dengan pendekatan statistika spasial menggunakan Python, sementara visualisasi spasial hasil klasterisasi dilakukan dengan bantuan perangkat lunak QGIS guna memudahkan interpretasi wilayah. Hasil analisis menunjukkan bahwa tempat wisata di Kabupaten Deli Serdang dapat dikelompokkan ke dalam tiga cluster utama, masing-masing merepresentasikan kelompok wisata yang sangat dekat, dekat, dan jauh dari pusat kota, serta memiliki perbedaan kisaran harga tiket masuk. Evaluasi kualitas klaster menggunakan Davies-Bouldin Index (DBI) menghasilkan nilai sebesar 1.377, yang mengindikasikan bahwa pembentukan cluster telah cukup baik, meskipun terdapat ruang untuk penyempurnaan. Dengan demikian, penelitian ini berhasil memetakan dan mengelompokkan tempat wisata di Deli Serdang secara spasial menggunakan pendekatan K-Means Clustering, yang diharapkan dapat memberikan kontribusi dalam perencanaan promosi pariwisata dan pengembangan destinasi wisata di wilayah tersebut.

Kata kunci: Statistika spasial, K-Means Clustering, Tempat Wisata, QGIS.

Abstract

A problem that often arises is the lack of a clear system for categorizing and classifying tourist attractions based on geographical characteristics. Tourist route planning, supporting infrastructure development, and targeted promotion strategies all become more difficult because of this. As a result, a data-driven approach is needed to more systematically analyze and map tourist attractions in Deli Serdang Regency. This research aims to map and cluster tourist attractions in Deli Serdang Regency based on spatial characteristics and entrance ticket prices using the K-Means Clustering method. This method is used to group tourist sites into several categories based on the similarity of attributes, which include geographic coordinates (latitude and longitude), ticket prices, and distance to the city center as a reference location. The data were analyzed using a spatial statistics approach using Python, while spatial visualization of the clustering results was done with the help of QGIS software to facilitate interpretation of the area. The results of the analysis show that tourist attractions in Deli Serdang Regency can be grouped into three main clusters, each representing tourist groups that are very close, close, and far from the city center, and have differences in the range of admission prices. Evaluation of cluster quality using the Davies-Bouldin Index (DBI) resulted in a value of 1.377, which indicates that cluster formation has been quite good, although there is room for improvement. Thus, this study succeeded in mapping and clustering tourist attractions in Deli Serdang spatially using the K-Means Clustering approach, which is expected to contribute to tourism promotion planning and tourist destination development in the region.

Keywords: Spatial statistics, K-Means Clustering, Tourist Attractions, QGIS.

Article Info

Received date: 27 March 2025

Revised date: 30 March 2025

Accepted date: 15 April 2025

PENDAHULUAN

Perekonomian sebuah negara dapat ditingkatkan secara strategis melalui sektor pariwisata. Sektor ini tidak hanya meningkatkan pendapatan daerah tetapi juga menciptakan lapangan pekerjaan

dan mengenalkan potensi lokal dan budaya kepada orang lain. Kabupaten Deli Serdang, yang berada di Provinsi Sumatera Utara, memiliki banyak wisata alam dan buatan, termasuk air terjun, agrowisata, tempat religius, dan tempat rekreasi keluarga. Namun, salah satu tantangan tersendiri dalam mengembangkan industri pariwisata di daerah tersebut adalah distribusi informasi yang buruk dan pemetaan potensi wisata yang belum terstruktur dengan baik. (Annisa, dkk., 2023).

Belum adanya sistem yang jelas untuk mengkategorikan dan mengklasifikasikan tempat wisata berdasarkan karakteristik geografis adalah masalah yang sering muncul. Ini membuat perencanaan rute wisata, pembangunan infrastruktur pendukung, dan membuat strategi promosi yang tepat sasaran menjadi sulit. Akibatnya, pendekatan berbasis data diperlukan untuk menganalisis dan memetakan tempat wisata di Kabupaten Deli Serdang secara lebih sistematis. (Anggiani, dkk., 2024)

Algoritma K-Means Clustering adalah salah satu metode yang dapat digunakan untuk menyelesaikan masalah ini. K-Means, metode klasterisasi tak terawasi (*Unsupervised learning*), mengelompokkan data ke dalam sejumlah klaster berdasarkan kemiripan atribut tertentu. Dalam hal ini, tempat wisata dapat dikelompokkan berdasarkan koordinat geografis (*latitude* dan *longitude*), serta atribut lainnya, seperti jenis wisata, jumlah fasilitas, dan tingkat kunjungan. Klasterisasi ini membantu melihat peta wisata. Mereka juga dapat membantu pemerintah daerah, pelaku industri pariwisata, dan masyarakat lokal membuat keputusan strategis. (Jain, 2010)

Penelitian ini bertujuan untuk melakukan pemetaan dan klasterisasi tempat wisata di Kabupaten Deli Serdang dengan menggunakan metode K-Means. Metode ini diharapkan dapat menghasilkan kelompok destinasi wisata yang memiliki karakteristik yang sama. Kelompok-kelompok ini dapat digunakan sebagai dasar untuk membuat strategi pengembangan, promosi, dan distribusi infrastruktur pariwisata yang lebih terarah dan efektif. (Lubis, dkk., 2024).

KAJIAN PUSTAKA

K-Means Clustering

K-Means adalah salah satu algoritma *unsupervised learning* yang paling populer dan banyak digunakan dalam analisis data, terutama untuk clustering atau pengelompokan data ke dalam kelompok-kelompok yang memiliki karakteristik yang sebanding. Metode ini bekerja dengan membagi sekumpulan data ke dalam sejumlah cluster yang telah ditentukan sebelumnya (jumlah cluster = k).

Rumus Euclidean Distance :

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (1)$$

Dengan keterangan

De = Euclidean Distance

i = Banyak objek

x, y = koordinat objek

s, t = koordinat centroid

Proses dimulai dengan memilih k pusat cluster acak, menghitung jarak antara setiap titik data ke pusat cluster tersebut, mengelompokkan data ke dalam cluster terdekat, dan mengubah posisi pusat cluster sesuai dengan posisi rata-rata anggota cluster. Proses ini berulang hingga posisi pusat cluster tidak lagi berubah secara signifikan (Triyani dkk, 2022).

Quantum GIS

Salah satu program Sistem Informasi Geografis (SIG) sumber terbuka yang paling terkenal dan banyak digunakan saat ini adalah Quantum GIS (QGIS). Gary Sherman pada awalnya menciptakan QGIS pada tahun 2002, dan dengan bantuan komunitas sumber terbuka internasional, program ini terus berkembang dengan cepat. Program ini dimaksudkan untuk membantu pengguna mengatur, memeriksa, dan menampilkan data spasial dalam berbagai bentuk (Budiyanto, 2016). QGIS telah digunakan dalam berbagai aplikasi penelitian dan praktis. Misalnya, dalam studi oleh Wicaksono (2023), QGIS digunakan untuk menghasilkan peta tematik dan analisis visual yang mendukung kebijakan tata kota. Hasil penelitian menunjukkan bahwa integrasi QGIS mampu menghasilkan pemetaan spasial yang akurat dan informatif dengan waktu yang relatif efisien.

Devies Bouldin Indeks

Davies-Bouldin Index (DBI) adalah ukuran yang mengukur sejauh mana cluster dalam model saling terpisah. DBI mengevaluasi keterpisahan dan keterpaduan antar cluster. Semakin kecil nilai DBI yang diperoleh, maka akan semakin bagus cluster yang diperoleh.

METODE

Objek penelitian ini menggunakan data tempat wisata di Kabupaten Deli Serdang yang diambil dari Google Maps dengan karakteristik harga tiket yang diambil dari berbagai sumber resmi yang tersedia. Data nantinya akan diolah menjadi 3 kluster berdasarkan jaraknya terhadap pusat kota yaitu sangat dekat, dekat, dan jauh. Penelitian ini akan melakukan pengklusteran dan pemetaan dengan langkah-langkah sebagai berikut :

1. Melakukan pengumpulan dan pengolahan data tempat wisata dan data spasial peta administrasi Kabupaten Deli Serdang.
2. Melakukan proses clustering dan evaluasi model menggunakan validasi Davies-Bouldin Index.
3. Melakukan perancangan peta digital sebagai sistem informasi pemetaan

HASIL DAN PEMBAHASAN

Pengumpulan dan Pengolahan Data

Setelah dilakukan pengambilan data tempat wisata, didapatkan 30 tempat wisata di sekitar Kabupaten Deli Serdang. Kemudian, data teks alamat akan diubah menjadi data koordinat. Hal ini berguna untuk mengidentifikasi pola spasial, menentukan jarak dan kedekatan, serta membantu mendapatkan visualisasi hasil clustering pada peta digital. Berikut adalah tampilan sepuluh data pertama beserta data koordinatnya:

place	latitude	longitude	ticket
Air Terjun Dua Warna	3.258917	98.51315	15000
Air Terjun Sempuren Putih	3.267509	98.6006	5000
Air Terjun Pelangi Indah	3.204489	98.69073	10000
Pulau Siba	3.729176	98.61309	30000
Danau Linting	3.229614	98.72666	10000
Air Terjun Lembah Loknya	3.263684	98.55404	10000
Pemandian Alam Sembahe	3.341543	98.58594	5000
Pemandian Alam Lau Sigembur	3.299482	98.64581	15000
Pantai Salju	3.308302	98.81205	25000
Hillpark Sibolangit	3.28144	98.5563	120000

Tabel 1. Data tempat wisata

Kemudian data harga tiket yang telah dikumpulkan akan disatukan dalam dataset untuk melakukan analisis tingkat harga tiket.

Proses Clustering dan Evaluasi Hasil Clustering

Pada tahap ini, proses dilakukan dengan bantuan Python untuk mengelompokkan cluster dan mencari hasil evaluasi modelnya.

Import Data

Langkah awal dalam tahap analisis data adalah mengimpor dataset yang digunakan. Data yang telah dikumpulkan disimpan dalam format file .xlsx dan kemudian akan dibaca ke dalam direktori kerja lingkungan pemrograman Python.

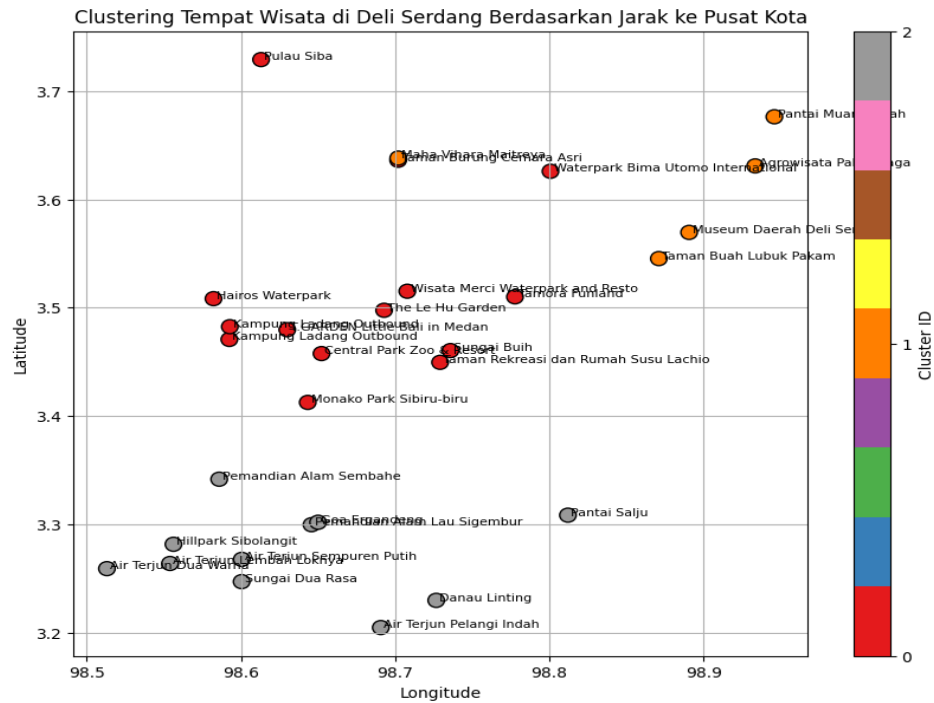
Pengelompokan Data

Jumlah data sebanyak 30 dan jumlah insialisasi centroid k ditentukan secara acak untuk menghasilkan 3 cluster. Pengelompokan cluster berdasarkan jarak dari pusat kota adalah sebagai berikut :

Cluster 0 : Sangat dekat berjumlah 13

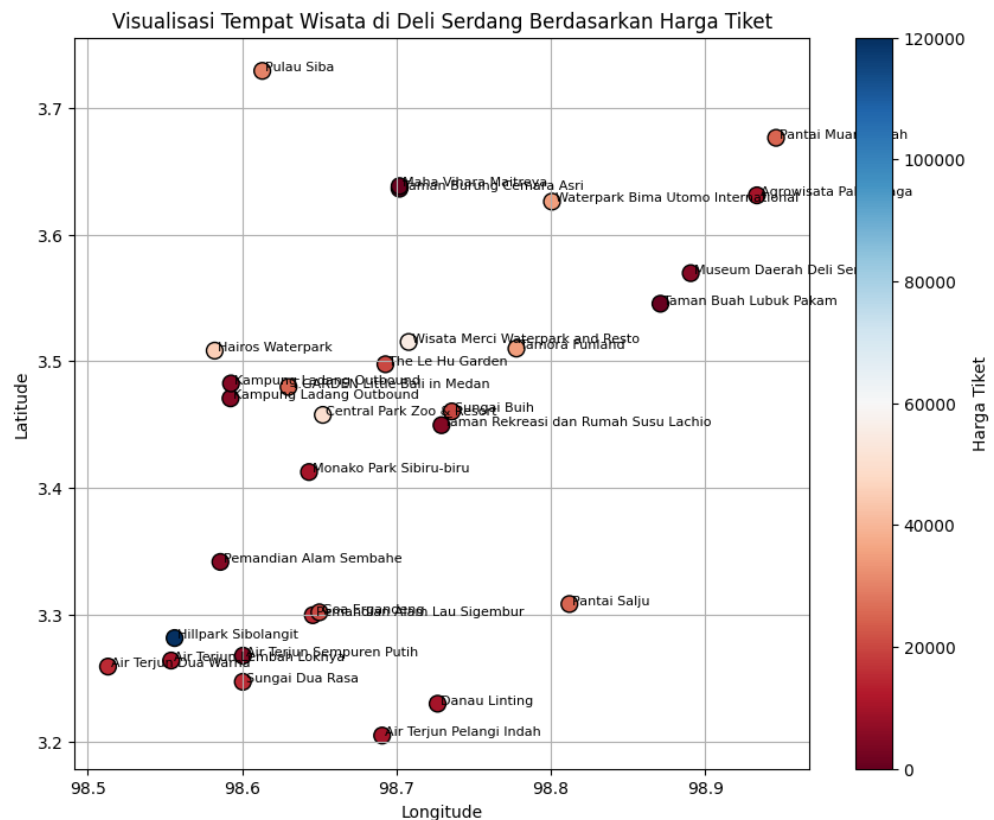
Cluster 1 : Dekat berjumlah 6 tempat

Cluster 2 : Jauh berjumlah 11 tempat

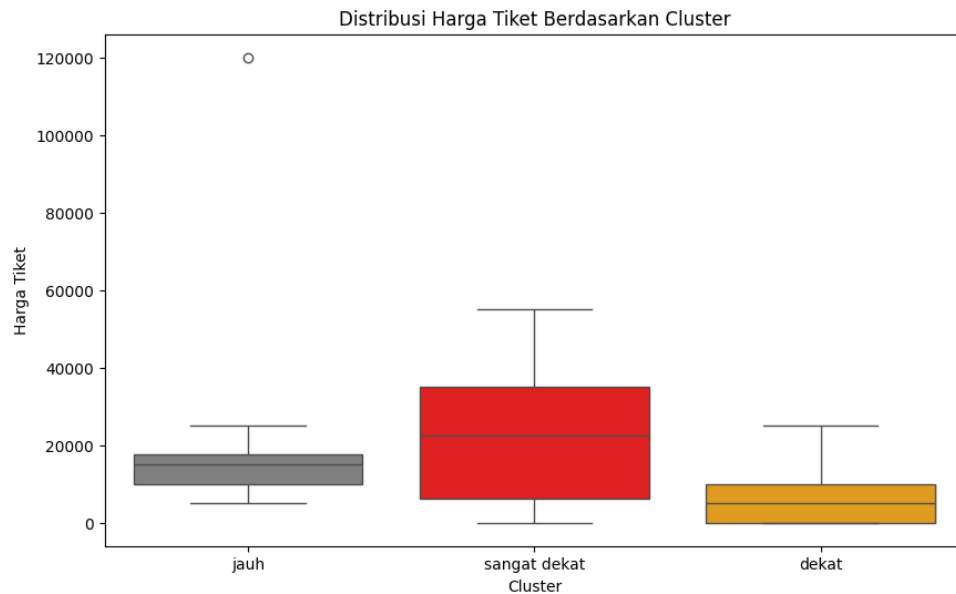


Gambar 1. Plot Cluster

Dari Gambar 1. Ditunjukkan hasil pembagian data menjadi 3 cluster yang ditandai dengan warna berbeda. Cluster 0 ditandai dengan titik berwarna merah, cluster 1 ditandai dengan titik berwarna jingga, dan cluster 2 ditandai dengan titik berwarna abu-abu. Kemudian, dari hasil clustering dianalisis harga tiket setiap tempat wisata untuk mengetahui harga tiket setiap tempat wisata yang telah dikelompokkan berdasarkan cluster. Untuk visualisasi harga tiket dapat dilihat pada Gambar 2 berikut :



Gambar 2. Plot tempat wisata berdasarkan harga tiket



Gambar 3. Distribusi harga tiket berdasarkan cluster

Harga tiket di setiap tempat wisata sangat beragam, hal ini ditunjukkan dari warna tingkat meter harga tiket. Tempat wisata dengan warna merah gelap hingga putih menunjukkan rentang harga tiket gratis (Rp. 0) hingga Rp. 60.000, sementara itu warna biru terang hingga biru gelap menunjukkan rentang harga tiket > Rp. 60.000 hingga Rp. 120.000. Pada Gambar 3 ditunjukkan distribusi harga tiket untuk setiap clusternya. Hasil analisisnya dapat disimpulkan sebagai berikut :

- Cluster “Jauh” memiliki harga tiket rata-rata lebih tinggi dibandingkan cluster “Dekat”, tetapi lebih rendah dari cluster “Sangat Dekat”. Rentang harga relatif sempit dengan satu outlier ekstrim di harga Rp. 120.000.
- Cluster “Sangat Dekat” memiliki variabilitas harga paling tinggi dan kisaran harga yang sangat lebar.
- Cluster “Dekat” distribusinya lebih rapat yang menandakan harga di cluster ini cenderung stabil dengan harga tiket yang lebih rendah dibandingkan dua cluster lain

2.1. Evaluasi Model Cluster

Selanjutnya dilakukan evaluasi kualitas clustering metode Davies-Bouldin Index (DBI) pada data tempat wisata yang telah dianalisis. Didapatkan hasil DBI pada model cluster adalah sebesar 1.3771. Berikut adalah syntax yang dijalankan dalam pemrograman :

```
# Clustering dengan KMeans
kmeans = KMeans(n_clusters=3, random_state=42)
labels = kmeans.fit_predict(X_scaled)

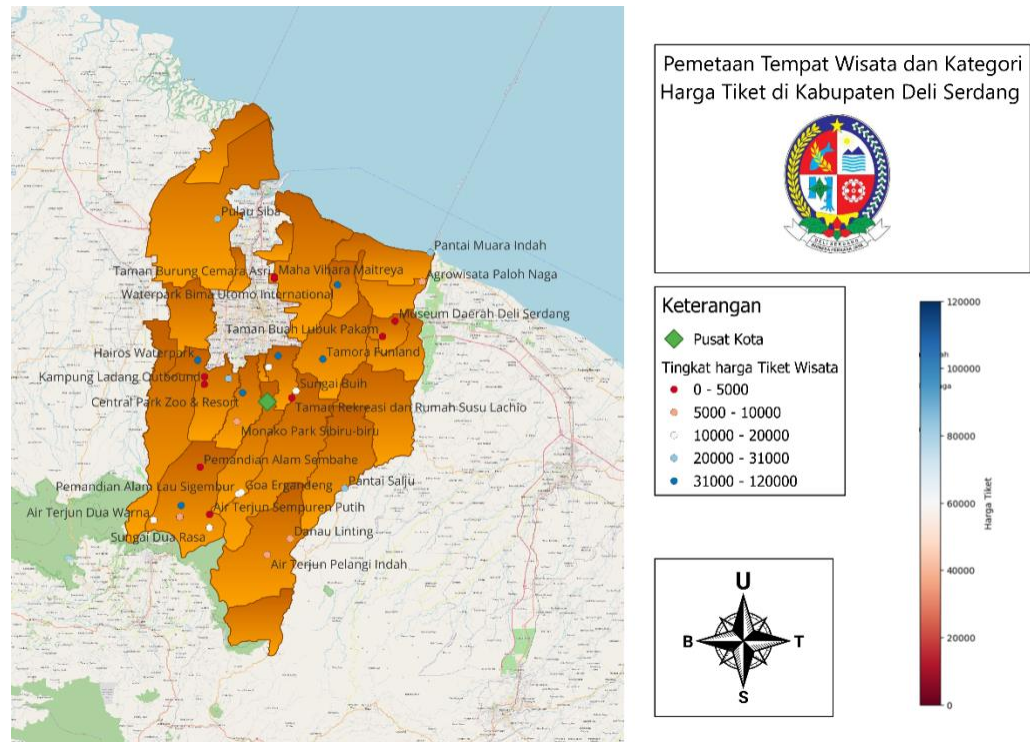
# Hitung Davies-Bouldin Index
db_index = davies_bouldin_score(X_scaled, labels)

# Tampilkan hasil
print("Davies-Bouldin Index:", db_index)
```

Nilai validasi ini dapat menunjukkan bahwa cluster cukup terpisah dan memiliki homogenitas yang cukup baik. Cluster dapat dikatakan bagus jika nilai DBI (non-negatif ≥ 0) kecil, semakin kecil nilai DBI maka akan menunjukkan cluster yang lebih terpisah dan lebih padat.

Perancangan Peta Digital

Proses pada tahap ini dilakukan dengan menggunakan bantuan software Quantum GIS untuk menampilkan hasil visualisasi peta dengan memasukkan data SHP Kabupaten Deli Serdang. Hasil pemetaan dan kategori harga tiket akan menghasilkan peta penyebaran tempat wisata di Kabupaten Deli Serdang. Peta digital dapat dilihat pada gambar 4.



Gambar 4. Peta penyebaran tempat wisata di wilayah Kabupaten Deli Serdang

SIMPULAN

Pemetaan dan klasterisasi tempat wisata di Kabupaten Deli Serdang menggunakan metode k-means clustering berhasil dilakukan dengan pembagian 3 cluster dengan nilai DBI 1.377. Cluster “Dekat” memiliki harga tiket yang lebih rendah dibandingkan dua cluster lain dan cluster “Sangat Dekat” memiliki variabilitas harga yang bervariasi.

Nilai DBI sebesar 1.377 dihasilkan oleh model clustering menunjukkan bahwa, meskipun belum sepenuhnya ideal, pembagian tempat wisata ke dalam tiga cluster memiliki tingkat pemisahan yang cukup baik. Oleh karena itu, hasil clustering dapat dikatakan cukup efektif, namun, kualitas model masih dapat ditingkatkan. Pertimbangkan untuk menggunakan algoritma alternatif seperti Agglomerative Clustering atau DBSCAN. Beberapa algoritma ini lebih baik daripada KMeans dalam menangani bentuk cluster yang tidak linier atau berdensitas tinggi.

REFERENSI

- Annisa, A. I. L. N., & Nasution, A. (2023). Strategi Pengembangan Wisata Sawah dalam Meningkatkan Ekonomi. *Jurnal Regional*, 1(1), 70–80.
- Anggiani, S., Lubis, A., & Siregar, M. A. (2024). Studi Empiris: Dampak Aksesibilitas dan Citra Destinasi terhadap Keinginan Berkunjung Kembali ke Pemandian Alam Sembahe. *Jurnal Ilmiah Manajemen, Ekonomi, & Akuntansi (MEA)*, 8(2), 2307–2314.
- Jain, A. K. (2010). Data Clustering: 50 Years Beyond K-Means. *Pattern Recognition Letters*, 31(8), 651–666.
- Lubis, I., Batubara, M., & Arif, M. (2024). Green Economy-Based Tourism Village Development Strategy in Denai Lama Village, Deli Serdang. *Indonesian Interdisciplinary Journal of Sharia Economics*, 6(3), 2510–2537.
- Triyani, E., Hudjimartsu, S. A., & Primasari, D. (2022). Spasial Clustering Potensi Peternakan Unggas dengan Metode K-Means Berbasis Webgis. *Infotech Journal*, 8(2), 10-18.
- Budiyanto, E. (2016). Sistem informasi geografis dengan Quantum GIS.
- Wicaksono, G. (2023). Pemanfaatan QGIS dalam Pengembangan Kota Digital Berbasis Analisis Spasial: Studi Kasus Surakarta Sebagai Smart City. *Journal of Digital Architectur*, 3(1), 21-26
- Rahmawati, L., Widya Sihwi, S., & Suryani, E. (2016). Analisa Clustering Menggunakan Metode K-Means Dan Hierarchical Clustering (Studi Kasus : Dokumen Skripsi Jurusan Kimia, Fmipa, Universitas Sebelas Maret). *Jurnal Teknologi & Informasi IT Smart*, 3(2), 66.