

Klasterisasi Pola Perjalanan Wisatawan Nusantara Menggunakan DBSCAN Berbasis SIG

Clustering Domestic Tourist Spatial Dynamics Using GIS-Based DBSCAN

Widya Aprilia^{1,*}, Setyoningsih Wibowo², Ramadhan Renaldy³

^{1,2,3}Program Studi Informatika, Universitas PGRI Semarang, Semarang, Indonesia
¹widyaaprila04@gmail.com; ²setyoningsihwibowo@upgris.ac.id; ³ramadhanrenaldy@upgris.ac.id;
*corresponding author

Informasi Artikel	ABSTRAK
Diserahkan : 2 Mei 2026 Diterima : 19 Mei 2026 Direvisi : 7 Agustus 2026 Diterbitkan : 31 Agustus 2026 Kata Kunci: Dinamika spasial DBSCAN SIG Wisatawan Nusantara Clustering berbasis densitas Keywords: Spatial dynamics DBSCAN GIS Domestic tourists Density-based clustering This is an open access article under the CC-BY-SA license. 	Penelitian ini menganalisis dinamika spasial Wisatawan Nusantara (Wisnus) periode 2019 sampai 2024 dengan menerapkan algoritma <i>Density-Based Spatial Clustering of Applications with Noise</i> (DBSCAN) yang terintegrasi dengan Sistem Informasi Geografis (SIG). Penelitian membagi periode observasi menjadi tiga fase: pra-pandemi, masa pandemi COVID-19, dan fase pemulihan, yang masing-masing menunjukkan karakteristik distribusi spasial berbeda secara signifikan. Peneliti menggunakan data kunjungan Wisatawan Nusantara tingkat provinsi sebagai dasar analisis dan menambahkan atribut spasial berupa koordinat geografis. Peneliti mengekstrak fitur temporal dalam bentuk statistik deskriptif, seperti total kunjungan, rata-rata, standar deviasi, dan tren pertumbuhan. Peneliti menerapkan standardisasi <i>z-score</i> agar kontribusi setiap variabel seimbang. Peneliti mengoptimalkan parameter DBSCAN melalui <i>grid search</i> dan memperoleh kombinasi <i>Eps</i> 1,4 dan <i>MinPts</i> 2 dengan nilai <i>Silhouette Score</i> 0,430, <i>Davies-Bouldin Index</i> 1,889, dan <i>Calinski-Harabasz Index</i> 7,555. Hasil <i>clustering</i> memperlihatkan pergeseran signifikan pola distribusi spasial selama pandemi dan peningkatan konsentrasi kunjungan pada fase pemulihan. Integrasi DBSCAN dengan SIG menghasilkan visualisasi <i>choropleth</i> informatif yang mendukung pengambilan keputusan berbasis data dalam perencanaan destinasi wisata nasional yang adaptif. ABSTRACT <i>This study analyzes the spatial dynamics of Domestic Tourists (Wisnus) for the period 2019 to 2024 by applying the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm integrated with Geographic Information Systems (GIS). The study divides the observation period into three phases: pre-pandemic, the COVID-19 pandemic, and the recovery phase, each showing significantly different spatial distribution characteristics. Researchers used provincial-level domestic tourist visit data as the basis for analysis and added spatial attributes in the form of geographic coordinates. Temporal features were extracted in the form of descriptive statistics, such as total visits, average, standard deviation, and growth trends. Researchers applied z-score standardization to balance the contribution of each variable. DBSCAN parameters were optimized through grid search, obtaining a combination of Eps 1.4 and MinPts 2 with a Silhouette Score of 0.430, Davies-Bouldin Index of 1.889, and Calinski-Harabasz Index of 7.555. The clustering results show a significant shift in spatial distribution patterns during the pandemic and an increased concentration of visits in the recovery phase. The integration of DBSCAN with GIS produces informative choropleth visualizations that support data-driven decision-making in adaptive national tourism destination planning.</i>

I. Pendahuluan

Pariwisata merupakan sektor ekonomi strategis yang berkontribusi signifikan terhadap pertumbuhan ekonomi regional, penciptaan lapangan kerja, dan mobilitas antarwilayah. BPS mencatat perjalanan wisatawan nusantara sebelum pandemi mencapai lebih dari 700 juta perjalanan per tahun [1]. Besarnya kontribusi sektor ini menuntut pengelolaan dan perencanaan pariwisata yang optimal, sebab pergerakan wisatawan terbukti berdampak langsung terhadap peningkatan pertumbuhan ekonomi dan Indeks Pembangunan Manusia (IPM) di wilayah tujuan [2]. Oleh karena itu, perumusan kebijakan berbasis data menjadi prasyarat krusial untuk memastikan pemerataan dampak positif tersebut.

Dinamika perjalanan wisatawan periode 2019-2024 mengalami perubahan kompleks akibat pandemi COVID-19 yang diumumkan WHO pada 11 Maret 2020 [3], yang tidak hanya menekan jumlah kunjungan tetapi juga mengubah pola perilaku wisatawan domestik. Pengelolaan destinasi wisata dituntut mampu menghadapi krisis sekaligus menawarkan pengalaman yang aman dan nyaman [4], dengan digitalisasi sebagai faktor kunci dalam memperkuat daya saing melalui pemasaran dan pengalaman interaktif [5]. Pergeseran pola kunjungan ini berdampak langsung pada sektor ekonomi lokal, khususnya Usaha Mikro, Kecil, dan Menengah (UMKM) yang menjadi tulang punggung ekosistem pariwisata daerah. Oleh karena itu, pemetaan distribusi wisata yang akurat menjadi sangat krusial; melalui pemetaan berbasis data, pemangku kebijakan dapat mengidentifikasi wilayah potensial guna mengalirkan dukungan sarana, promosi, dan investasi bagi UMKM setempat. Upaya penataan ini pada akhirnya berkontribusi langsung terhadap pemulihan ekonomi lokal dan peningkatan kesejahteraan masyarakat sekitar destinasi secara berkelanjutan.

Data yang berdimensi ruang dan waktu menimbulkan tantangan analitis tersendiri karena pengaruh variabel sering kali tidak homogen antardaerah maupun antarwaktu [6], sehingga pendekatan konvensional tanpa analisis spasial mengakibatkan pembangunan yang timpang antara pusat kota dan wilayah pinggiran. Pendekatan spasial-temporal terbukti mampu mengungkap pola distribusi yang tidak terdeteksi melalui analisis agregat biasa [7]. Perkembangan SIG dan data mining membuka peluang analisis distribusi wisatawan yang lebih komprehensif [8], di mana teknik clustering sebagai pendekatan unsupervised learning [9], khususnya DBSCAN, menonjol karena kemampuannya mendeteksi kluster berbentuk arbitrer dan mengenali *noise* tanpa penentuan jumlah kluster di awal [10].

Penelitian terdahulu telah mengeksplorasi berbagai pendekatan analisis dalam konteks pariwisata. Dalam analisis spasial berbasis clustering, DBSCAN dengan reduksi dimensi PCA terbukti mampu menghasilkan kluster destinasi wisata berkualitas baik dan mendukung kebijakan pemerataan pengembangan pariwisata di skala provinsi [11], sementara DBSCAN juga mengungguli K-Means dalam mengelompokkan opini masyarakat terkait kebijakan pariwisata pada masa pandemi [12]. Pada cakupan yang lebih luas, pendekatan data mining terbukti mampu mengungkap disparitas distribusi spasial-temporal efisiensi pariwisata yang persisten antarwilayah [8]. Kondisi serupa juga ditemukan pada kawasan wisata vulkanik, di mana pengelolaan konvensional tanpa analisis spasial menyebabkan pengembangan terkonsentrasi di pusat kota, sementara wilayah pinggiran kurang optimal [7].

Pandemi COVID-19 turut mendorong studi manajemen krisis pariwisata. Pengelolaan destinasi wisata selama pandemi menuntut adaptasi kebijakan secara bertahap mulai dari prakrisis, respons krisis, hingga pascakrisis untuk menjamin keberlangsungan operasional wisata [13]. Pandemi juga menekankan kebutuhan analisis spasial prediktif, di mana penurunan kunjungan hingga 50% mendorong pentingnya pemetaan berbasis data untuk mendukung pemulihan destinasi pasca-pandemi [7]. Hal ini semakin relevan mengingat efisiensi dan distribusi spasial pariwisata sangat dipengaruhi oleh dinamika infrastruktur transportasi, struktur industri, dan kebijakan regional [14].

Metode ST-DBSCAN yang mengintegrasikan dimensi ruang dan waktu secara simultan [15] menunjukkan potensi besar dalam analisis spasial-temporal, namun penerapan clustering berbasis DBSCAN yang diintegrasikan dengan SIG pada data Wisatawan Nusantara Indonesia di skala nasional untuk periode transisi pandemi masih terbatas. Berdasarkan kesenjangan tersebut, penelitian ini bertujuan menganalisis dinamika spasial Wisatawan Nusantara periode 2019-2024 menggunakan algoritma DBSCAN yang diintegrasikan dengan SIG guna menghasilkan pemetaan informatif yang mendukung perencanaan pariwisata berbasis bukti.

II. Metode

A. Problem Analysis

Distribusi Wisatawan Nusantara periode 2019-2024 menampilkan dinamika spasial yang kompleks, ditandai dengan fluktuasi signifikan akibat pandemi COVID-19 yang berdampak berbeda pada setiap provinsi [3]. Data BPS menunjukkan penurunan drastis kunjungan pada 2020–2021, diikuti pemulihan bertahap yang tidak merata secara geografis [1]. Ketimpangan laju pemulihan ini secara langsung memengaruhi ketahanan ekonomi daerah, mengingat volume kunjungan wisatawan terbukti memberikan kontribusi nyata terhadap pertumbuhan ekonomi regional serta peningkatan Indeks Pembangunan Manusia (IPM) di wilayah tujuan [2].

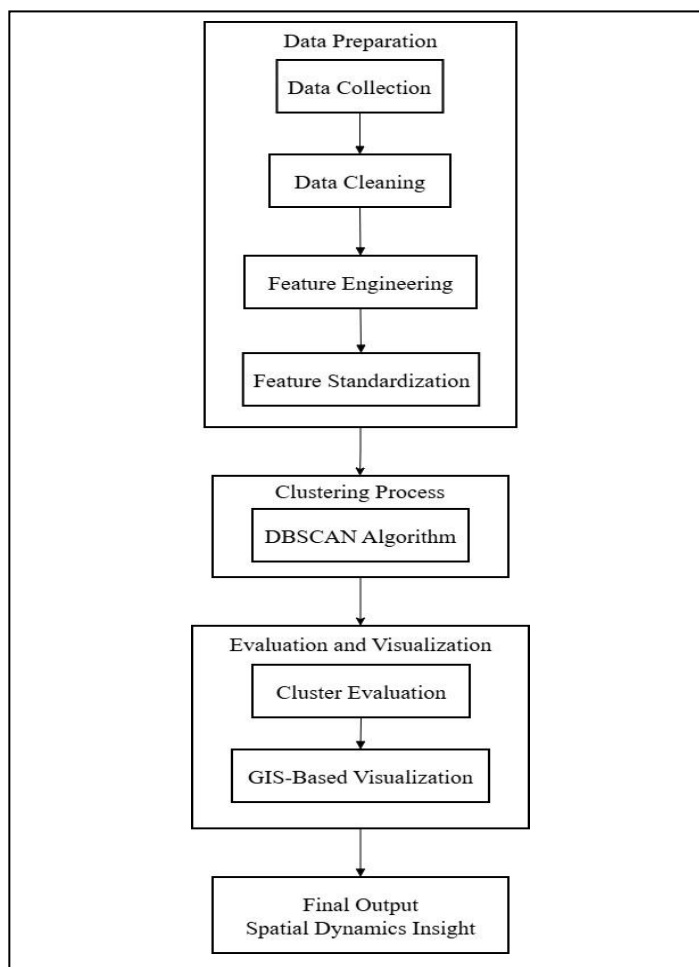
Permasalahan utama penelitian ini adalah belum teridentifikasinya pola kepadatan spasial berbasis waktu secara nasional yang mampu merepresentasikan perubahan distribusi wisatawan secara adaptif pada tiga fase temporal tersebut. Karakteristik data spasial dengan variasi kepadatan heterogen dan distribusi nonlinier menuntut pendekatan clustering yang tidak memerlukan asumsi jumlah kluster dan mampu mendeteksi *noise* secara otomatis [10]. DBSCAN dipilih karena memenuhi kedua kebutuhan tersebut sekaligus memiliki efisiensi komputasi yang baik untuk dataset berskala nasional [9].

B. System Design

Desain sistem penelitian dirancang sebagai arsitektur pemrosesan data bertahap yang mencakup pengumpulan data, preprocessing, ekstraksi fitur, standarisasi, proses clustering, evaluasi model, dan

visualisasi berbasis SIG. Data kunjungan Wisatawan Nusantara tingkat provinsi tahun 2019-2024 bersumber dari BPS [1], dilengkapi dengan atribut spasial berupa koordinat geografis (*latitude dan longitude*).

Ekstraksi fitur statistik mencakup total kunjungan, nilai rata-rata (*mean*), standar deviasi, serta tren pertumbuhan kunjungan selama periode penelitian. Pendekatan ekstraksi fitur ini sejalan dengan praktik terbaik data mining spasial [8], di mana representasi fitur yang komprehensif meningkatkan kualitas hasil clustering. Alur pemrosesan sistem penelitian ini terdiri dari tiga fase utama: *Data Preparation (Raw Data, Data Cleaning, Feature Engineering, Standardization)*, *Clustering Process (DBSCAN Algorithm)*, dan *Validation and Visualization (Cluster Evaluation, GIS-Based Visualization)*, sebagaimana ditunjukkan pada Gambar 1. Hasil akhir berupa dataset berlabel klaster serta peta choropleth interaktif yang merepresentasikan dinamika spasial distribusi Wisatawan Nusantara.



Gambar 1. Desain Sistem Pemrosesan Data Spasial

Tahapan pertama adalah Data Preparation, yang mencakup Data Collection untuk mengumpulkan data mentah kunjungan Wisatawan Nusantara dari BPS, diikuti dengan Data Cleaning untuk memastikan kualitas data, Feature Engineering untuk mengekstraksi atribut statistik temporal, serta Feature Standardization menggunakan metode z-score agar setiap variabel memberikan kontribusi yang seimbang. Setelah data siap, proses berlanjut ke tahap Clustering Process yang menerapkan DBSCAN Algorithm untuk mengidentifikasi pola kepadatan spasial tanpa memerlukan penentuan jumlah klaster secara a priori. Tahap berikutnya adalah Evaluation and Visualization, di mana dilakukan Cluster Evaluation menggunakan metrik seperti Silhouette Score, Davies-Bouldin Index, dan Calinski-Harabasz Index, yang kemudian dilanjutkan dengan GIS-Based Visualization untuk menghasilkan pemetaan choropleth guna memudahkan interpretasi pergeseran distribusi wisatawan secara geografis. Seluruh rangkaian proses ini bermuara pada Final Output berupa Spatial Dynamics Insight yang memberikan pemahaman mendalam mengenai dinamika distribusi kunjungan Wisatawan Nusantara pada periode 2019-2024.

C. Data Preparation

Karena setiap fitur memiliki skala yang berbeda, dilakukan normalisasi menggunakan metode *z-score* untuk mencegah dominasi variabel tertentu dalam perhitungan jarak *Euclidean*. Proses standarisasi dirumuskan sebagai (1) [16]:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

di mana x merupakan nilai asli, μ adalah rata-rata, dan σ Adalah simpangan baku. Sebagaimana ditekankan dalam [9], standarisasi merupakan langkah kritis dalam preprocessing data mining untuk memastikan setiap variabel memberikan kontribusi yang seimbang dalam proses clustering berbasis jarak.

1) Data Collection

Data yang digunakan dalam penelitian ini bersumber dari Badan Pusat Statistik Indonesia berupa data Jumlah Perjalanan Wisatawan Nusantara Menurut Provinsi Tujuan (Perjalanan) pada periode 2019-2024 [1]. Data ini mencakup catatan pergerakan atau mobilisasi penduduk Indonesia yang melakukan perjalanan ke berbagai wilayah di dalam negeri untuk tujuan wisata, rekreasi, maupun kunjungan lainnya. Cakupan waktu dari tahun 2019 hingga 2024 dipilih untuk melihat tren pergerakan wisatawan sebelum, selama, hingga pascapandemi COVID-19. Selain itu, data ini mengintegrasikan dimensi spasial berupa provinsi tujuan dan dimensi temporal berupa periode tahunan, sehingga membentuk *dataset* spasial-temporal yang representatif untuk analisis dinamika distribusi wisatawan di skala nasional.

2) Data Cleaning

Tahap pembersihan data dilakukan untuk memastikan kualitas dan konsistensi dataset sebelum digunakan dalam proses analisis. Proses ini mencakup penanganan nilai hilang, penghapusan duplikasi data, serta validasi konsistensi nilai antarperiode. Data yang tidak memenuhi kriteria kelengkapan dan konsistensi dieliminasi untuk menghindari bias dalam hasil clustering [9].

Selain penanganan nilai hilang dan duplikasi, tahap pembersihan data juga mencakup penyelarasan unit analisis spasial $N = 34$ provinsi (mengabaikan entitas agregat 'INDONESIA'). Meskipun terjadi dinamika pemekaran wilayah administratif di Indonesia seperti pembentukan 3 provinsi baru pada Juli 2022 (Papua Selatan, Papua Tengah, Papua Pegunungan) dan 1 provinsi baru pada Desember 2022 (Papua Barat Daya) yang menjadikan total 38 provinsi pada periode 2022-2024 data pada provinsi-provinsi baru tersebut dikonsolidasikan kembali ke provinsi induknya. Langkah konsolidasi ini dilakukan untuk menjaga konsistensi deret waktu (*time-series continuity*) serta kesetaraan perbandingan unit observasi dari tahun 2019 hingga 2024.

3) Feature Engineering

Rekayasa fitur dilakukan untuk mengekstrak informasi yang relevan dari data mentah guna meningkatkan kualitas representasi data dalam proses clustering. Fitur yang dibangun mencakup statistik agregat distribusi kunjungan wisatawan per provinsi, tren perubahan antarperiode, serta indikator pemulihan pasca-pandemi. Pendekatan ini penting mengingat data pariwisata yang berdimensi ruang dan waktu memerlukan representasi fitur yang mampu menangkap heterogenitas spasial-temporal [6].

Guna menjamin keterulangan penelitian (*replicability*), proses rekayasa fitur secara spesifik menghasilkan matriks masukan multivariat berdimensi 5 ($D = 5$) untuk setiap unit provinsi (i), yang diekstrak ke dalam array NumPy X dan didefinisikan sebagai berikut:

$$v_i = [\text{latitude}_i, \text{longitude}_i, \text{total_kunjungan}_i, \text{std_kunjungan}_i, \text{growth_rate}_i] \quad (2)$$

Adapun rincian kelima atribut input yang diekstrak dari objek DataFrame (df_last) tersebut adalah:

- latitude dan longitude: Titik koordinat pusat (*centroid*) geografis provinsi ke- i sebagai komponen spasial murni.
- total_kunjungan: Akumulasi total volume perjalanan wisatawan nusantara pada fase temporal analisis.
- std_kunjungan: Standar deviasi yang mencerminkan tingkat variabilitas atau fluktuasi pergerakan wisatawan bulanan.
- growth_rate: Laju pertumbuhan pergerakan wisatawan antar-periode untuk menangkap dinamika tren pemulihan.

4) Feature Standardization

Standarisasi fitur dilakukan menggunakan metode *z-score* sebagaimana diuraikan pada bagian *system design*. Seluruh atribut fitur multivariat dalam matriks $X \in \mathbb{R}^{34 \times 5}$ (latitude, longitude, total_kunjungan, std_kunjungan, dan growth_rate) ditransformasikan secara otomatis menggunakan kelas `StandardScaler().fit_transform(X)` dari pustaka `scikit-learn` pada Python [9]. Proses ini memastikan bahwa seluruh fitur berada pada skala yang sebanding sehingga tidak terjadi dominasi atribut tertentu dalam

perhitungan jarak pada algoritma DBSCAN [15]. Tanpa standardisasi, fitur dengan rentang nilai besar akan mendominasi perhitungan jarak dan menyebabkan hasil clustering yang tidak akurat [9].

D. Clustering Process

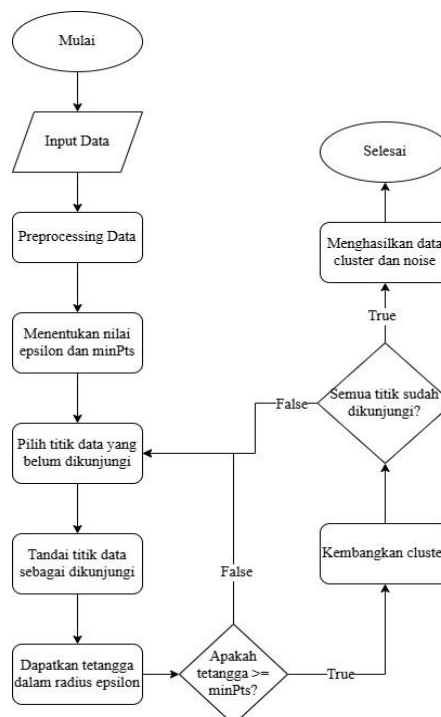
1) DBSCAN Algorithm

Proses clustering menggunakan algoritma DBSCAN yang telah terbukti efektif untuk analisis data spasial berskala besar [10]. DBSCAN mengelompokkan data berdasarkan kepadatan lokal menggunakan dua parameter utama: ϵ atau *Epsilon (Eps)* sebagai radius pencarian tetangga dan *MinPts (Minimum Points)* sebagai jumlah minimum titik untuk membentuk suatu kluster. Titik yang tidak memenuhi kriteria kepadatan minimum diklasifikasikan sebagai *noise*, suatu keunggulan yang relevan untuk dataset distribusi wisatawan dengan variasi kepadatan tinggi [11].

Pengelompokan dilakukan terhadap matriks fitur multivariat $X_{\text{scaled}} \in \mathbb{R}^{34 \times 5}$ yang telah di-standarisasi menggunakan *Z-score* pada Tahap 2.3.4. Kalkulasi jarak antar-objek menggunakan metrik jarak Euclidean (*Euclidean distance*). Eksperimen algoritma DBSCAN dijalankan menggunakan pustaka *scikit-learn* pada Python dengan menetapkan parameter *eps* (ϵ) = [1,4] dan *min_samples* (*MinPts*) = [2]. Penentuan parameter ini dioptimalkan untuk memisahkan kluster pola perjalanan utama dari titik-titik data penciran (*noise*).

Proses *clustering* dalam penelitian ini dieksekusi secara terpisah pada setiap fase temporal (Pra-Pandemi (2019), Pandemi (2020-2021), dan Pemulihan (2022-2024)). Pendekatan ini diterapkan untuk mengidentifikasi pergeseran kerapatan spasial secara eksplisit akibat fluktuasi mobilitas wisatawan.

Pemilihan parameter optimal dieksekusi melalui mekanisme *grid search* secara independen pada masing-masing fase temporal. Pemilihan parameter optimal dilakukan dengan rentang nilai ϵ dan *MinPts* yang telah ditentukan. Dasar penentuan batasan nilai *Epsilon (Eps)* didasarkan pada karakteristik *dataset* yang telah ditransformasikan menggunakan *StandardScaler* (Z-Score) ke dalam distribusi normal. Transformasi statistik ini menyebabkan jarak Euclidean standar antar-koordinat provinsi secara teoretis berpusat pada rentang 1,0 hingga 2,0. Melalui batasan teoretis tersebut, *grid search* menguji nilai radius secara presisi dan menetapkan nilai ϵ optimal yang dinamis sesuai densitas data per fase, yaitu $\epsilon = 2,1$ pada fase Pra-Pandemi, $\epsilon = 1,5$ pada fase Pandemi, dan $\epsilon = 1,4$ pada fase Pemulihan (dengan *MinPts* = 2 konstan pada seluruh fase). Selaras dengan hal tersebut, hasil *clustering* pada Sistem Informasi Geografis (SIG) disajikan secara *multi-temporal* berdasarkan parameter optimal dari masing-masing fase. Pendekatan ini selaras dengan praktik clustering spasial terdahulu [12]. Pengembangan lebih lanjut dari DBSCAN dalam konteks spasial-temporal telah diwujudkan melalui ST-DBSCAN [15] yang memungkinkan pengelompokan dengan mempertimbangkan dimensi waktu secara eksplisit, meskipun pada penelitian ini dimensi temporal direpresentasikan melalui fitur statistik agregat.



Gambar 2. Proses DBSCAN [17]

E. Evaluation and Visualization

1) Cluster Evaluation

Evaluasi kualitas kluster dilakukan menggunakan metrik validasi internal yang saling melengkapi. *Silhouette Score* mengukur seberapa baik setiap titik data sesuai dengan klasternya dibandingkan dengan kluster terdekat lainnya, sementara *Davies-Bouldin Index* mengevaluasi jarak antarkluster [18]. Berikut rumus untuk *silhouette score* sebagai (3) [16]:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (3)$$

di mana $a(i)$ adalah rata-rata jarak intra-kluster dan $b(i)$ Adalah jarak terdekat antarkluster. Nilai mendekati 1 menunjukkan pemisahan kluster yang baik [16].

Davies-Bouldin Index (DBI) mengukur rasio rata-rata jarak intra-kluster terhadap jarak antar-kluster, di mana nilai lebih kecil mengindikasikan kluster yang lebih kompak dan terpisah [19]. *Davies-Bouldin Index* dirumuskan pada Persamaan (4) [16]:

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} \frac{s_i + s_j}{d_{ij}} \quad (4)$$

Nilai DBI yang lebih kecil menunjukkan kluster yang lebih kompak dan terpisah dengan baik.

Calinski-Harabasz Index (CHI) merupakan metrik evaluasi clustering yang mengukur rasio variansi antarkluster terhadap variansi dalam kluster, di mana nilai yang lebih tinggi menunjukkan struktur kluster yang lebih jelas. Secara matematis, *Calinski-Harabasz Index* dirumuskan sebagaimana ditunjukkan pada Persamaan (5) [16]:

$$CH = \frac{Tr(B_k)/(k-1)}{Tr(W_k)/(n-k)} \quad (5)$$

Penggunaan metrik validasi seperti CHI juga banyak digunakan dalam penelitian sebelumnya untuk membandingkan performa metode clustering seperti K-Means dan DBSCAN [19].

2) GIS-Based Visualization

Hasil clustering diintegrasikan dengan data batas wilayah provinsi Indonesia dalam format *GeoJSON*, di mana visualisasi dilakukan menggunakan metode *choropleth* untuk mewarnai setiap provinsi berdasarkan label kluster. Pendekatan ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa evaluasi spasial-temporal dapat memetakan evolusi efisiensi pariwisata secara nasional guna mendukung pembangunan berkelanjutan [8]. Visualisasi berbasis spasial secara signifikan meningkatkan interpretabilitas hasil analisis kluster bagi pemangku kebijakan pariwisata, sebagaimana ditunjukkan dalam studi penggunaan algoritma DBSCAN untuk memetakan kepadatan destinasi wisata secara regional [11] maupun melalui analisis clustering berbasis densitas lainnya [12]. Untuk mendukung hal tersebut, dalam penelitian ini dibangun visualisasi interaktif menggunakan library kartografi Python yang memungkinkan eksplorasi spasial multitemporal secara mendalam.

III. Hasil dan Pembahasan

A. Clustering Results

Untuk mengevaluasi kualitas hasil klasterisasi pada setiap fase temporal, penelitian ini menerapkan proses optimasi parameter menggunakan metode *grid search* terhadap kombinasi nilai *Epsilon* dan *MinPts*. Evaluasi dilakukan menggunakan tiga metrik validasi internal, yaitu *Silhouette Score*, *Davies-Bouldin Index* (DBI), dan *Calinski-Harabasz Index* (CHI). Ketiga metrik tersebut digunakan untuk mengukur tingkat kohesi, separasi antar-kluster, serta stabilitas struktur spasial yang dihasilkan oleh algoritma DBSCAN. Hasil evaluasi kuantitatif parameter optimal pada masing-masing fase temporal ditunjukkan pada Tabel 1.

Tabel 1. Evaluasi Kuantitatif Kualitas Kluster DBSCAN Per Fase Temporal

Fase Temporal	Parameter Optimal	Silhouette Score	Davies-Bouldin Index	Calinski-Harabasz Index	Jumlah Kluster	Noise Points
Pra-Pandemi (2019)	$\epsilon = 2,1$; MinPts = 2	0,589	0,500	25,680	2	0
Pandemi (2020-2021)	$\epsilon = 1,5$; MinPts = 2	0,543	0,527	19,840	2	0

Fase Temporal	Parameter Optimal	Silhouette Score	Davies-Bouldin Index	Calinski-Harabasz Index	Jumlah Klaster	Noise Points
Pemulihan (2022-2024)	$\varepsilon = 1,4$; MinPts = 2	0,430	1,889	7,555	2	2

Berdasarkan Table 1, hasil optimasi parameter menggunakan metode grid search memperlihatkan dinamika perubahan kerapatan lokal sebaran wisatawan yang sangat fluktuatif di setiap fase temporal. Pada fase Pra-Pandemi (2019), algoritma DBSCAN menghasilkan kinerja spasial paling prima dengan *Silhouette Score* tertinggi sebesar 0,589 dan nilai *Davies-Bouldin Index (DBI)* terkecil mencapai 0,500. Capaian metrik ini membuktikan secara empiris bahwa sebelum pandemi, pergerakan wisatawan nusantara terdistribusi ke dalam struktur klaster yang sangat solid, kompak, dan terpisah dengan batas yang tegas tanpa memicu kemunculan wilayah pencilan (noise = 0). Memasuki fase Pandemi (2020–2021), terjadi penyusutan radius kerapatan optimal menjadi $\varepsilon = 1,5$ seiring dengan pembatasan mobilitas masyarakat secara nasional, namun model tetap mempertahankan kohesi kelompok yang relatif kuat dengan nilai *Silhouette Score* 0,543 dan DBI 0,527.

Pergeseran struktur spasial paling drastis terekam pada fase Pemulihan (2022–2024). Pada fase ini, *grid search* menetapkan nilai parameter optimal $\varepsilon = 1,4$ dan *MinPts* = 2, yang menghasilkan nilai *Silhouette Score* sebesar 0,430, *Davies-Bouldin Index* sebesar 1,889, dan *Calinski-Harabasz Index* sebesar 7,555. Munculnya 2 *noise points* (provinsi pencilan) pada fase ini secara substansial mengonfirmasi adanya fenomena anomali trajektori pemulihan pariwisata daerah pascapandemi, di mana terdapat beberapa provinsi tertentu yang arus kunjungannya pulih jauh lebih lambat atau timpang dibandingkan kelompok utamanya.

Secara keseluruhan, hasil eksekusi DBSCAN secara terpisah per fase temporal ini membuktikan bahwa penurunan nilai ε dari 2,1 (Pra-Pandemi) hingga 1,4 (Pemulihan) mencerminkan fragmentasi dan pergeseran tingkat kerapatan pergerakan wisatawan secara nasional. Nilai *Silhouette Score* pada fase Pemulihan (0,430), meskipun berada sedikit di bawah ambang batas ideal 0,500, tetap mencerminkan tingkat pemisahan dan kohesi yang memadai serta interpretabilitas yang tinggi untuk data pariwisata berfluktuasi tinggi. Hasil ini diperkuat dengan nilai DBI dan CHI yang secara konsisten mengonfirmasi bahwa eksekusi model DBSCAN per fase temporal berhasil mengidentifikasi struktur kepadatan spasial yang stabil, terpisah, dan representatif pada *dataset* Wisatawan Nusantara tingkat provinsi.

Nilai *Silhouette Score* sebesar 0,430 menunjukkan bahwa struktur klaster memiliki tingkat pemisahan dan kohesi yang memadai. Meskipun berada sedikit di bawah ambang batas ideal 0,500 nilai ini tetap mencerminkan interpretabilitas yang tinggi dalam konteks data pariwisata nasional, mengingat kemampuan DBSCAN dalam memisahkan klaster dari *noise* secara efektif sebagaimana diterapkan dalam studi pengelompokan indikator wilayah lainnya [17]. Hasil ini diperkuat dengan nilai DBI sebesar 1,889. Rendahnya nilai tersebut mencerminkan bahwa klaster yang terbentuk memiliki kepadatan internal yang tinggi dan terpisah secara spasial dengan baik, yang krusial dalam analisis pengelompokan daerah [20]. Selain itu, nilai CHI sebesar 7,555 memperlihatkan rasio yang baik antara variasi antarklaster dan intraklaster, di mana nilai CHI yang lebih tinggi menunjukkan kualitas klaster yang lebih optimal. Penggunaan kombinasi metrik evaluasi ini sejalan dengan penelitian sebelumnya yang melakukan komparasi berbagai metrik validasi untuk menjamin objektivitas hasil clustering [16]. Secara lebih rinci, perbandingan kinerja antara algoritma DBSCAN dengan beberapa metode clustering lainnya pada fase pemulihan disajikan dalam Tabel 2.

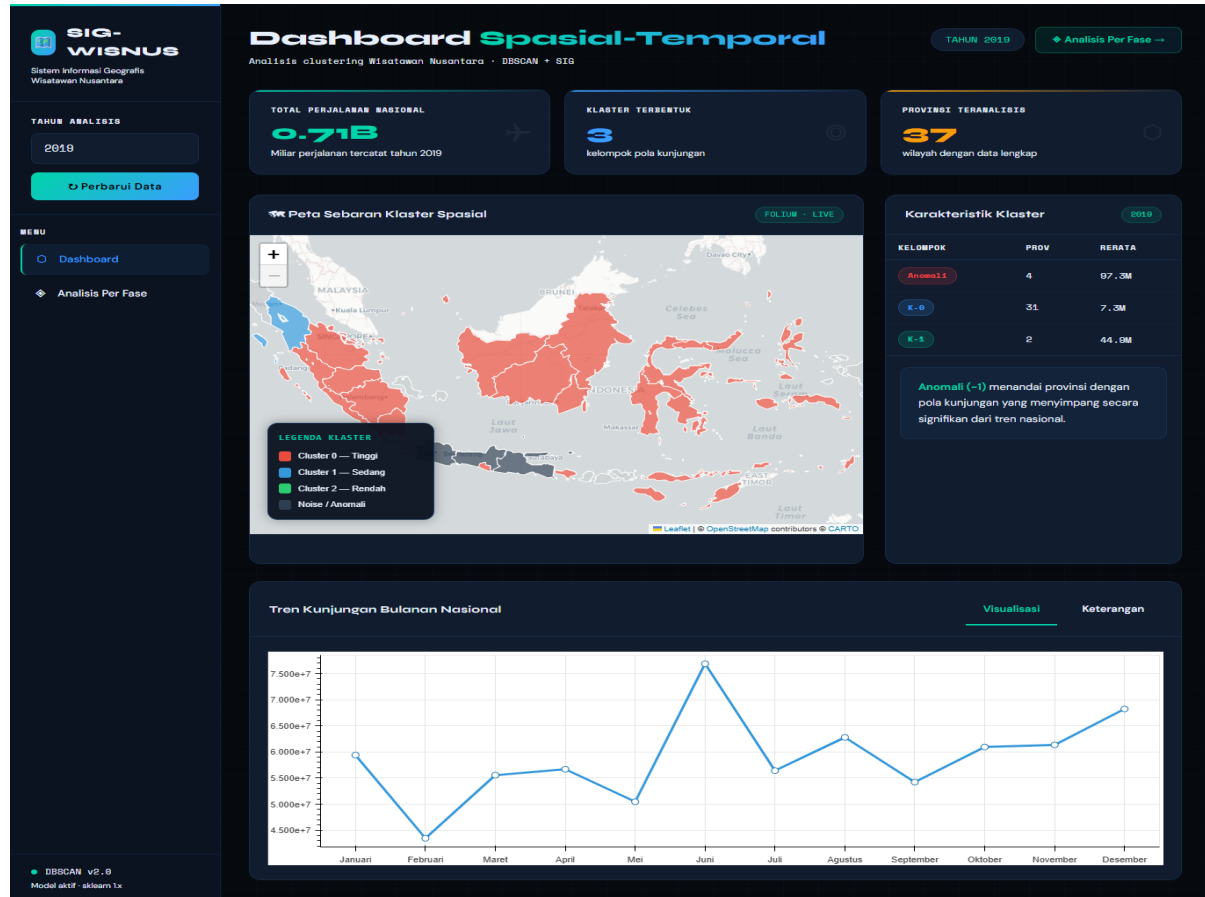
Tabel 2. Perbandingan Kinerja Algoritma Clustering

Algoritma	Silhouette	Davies-Bouldin	Calinski-Harabasz	Keterangan
DBSCAN	0,430	1,889	7,555	Terpilih, deteksi outlier otomatis
K-Means	0,385	0,887	27,319	Sensitif outlier, perlu tentukan k
Hierarchical	0,374	0,861	26,320	Kompleksitas tinggi, perlu tentukan k

Berdasarkan Tabel 2, terlihat bahwa DBSCAN mengungguli algoritma berbasis *centroid* dalam ketiga metrik evaluasi. Rendahnya nilai *Silhouette Score* algoritma pembanding membuktikan keunggulan DBSCAN dalam mengekstrak struktur kelompok yang bersih dari distorsi pencilan. Meskipun indeks *Davies-Bouldin* pada K-Means tampak lebih kecil secara parsial, hal tersebut disebabkan oleh karakteristik bawaan algoritma konvensional tersebut yang secara paksa menarik seluruh titik data ekstrem (*outlier*) ke dalam klaster terdekat, sehingga memicu bias interpretasi regional.

B. Spatial Interpretation

Visualisasi *choropleth* berbasis SIG mengungkapkan pola dinamika spasial yang berbeda pada setiap fase temporal. Pada periode pra-pandemi (2019), kluster dengan kepadatan tinggi terkonsentrasi pada provinsi dengan aktivitas ekonomi yang berkembang dan infrastruktur pariwisata yang matang. Implementasi visualisasi ini diwujudkan dalam bentuk dashboard sistem informasi geografis seperti yang ditunjukkan pada Gambar 3.



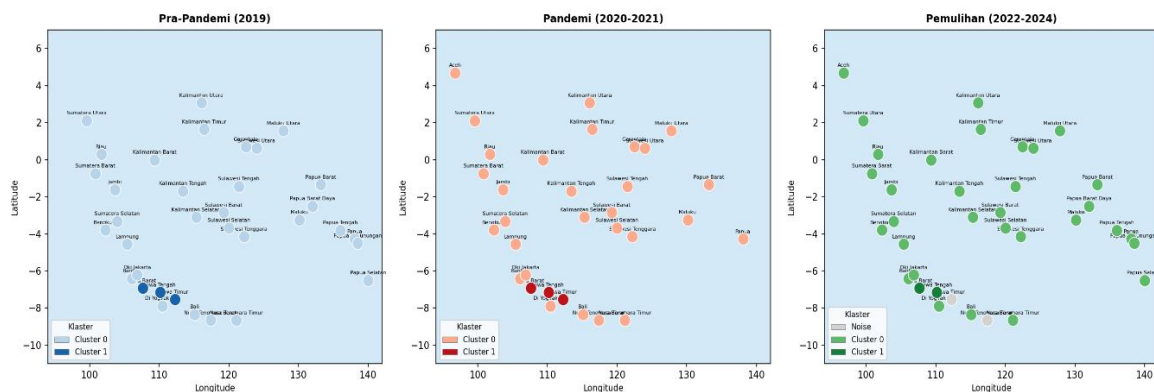
Gambar 3. Antarmuka Dashboard Spasial-Temporal Analisis Kluster Wisatawan Nusantara menggunakan DBSCAN dan SIG.

Gambar 3 menampilkan antarmuka sistem yang mengintegrasikan hasil pengolahan data DBSCAN dengan elemen visual. Dashboard tersebut menyajikan informasi komprehensif mulai dari total perjalanan nasional, jumlah kluster yang terbentuk, hingga pendeteksian data anomali (*noise*) yang ditandai dengan label -1. Keberadaan label ini menegaskan kemampuan DBSCAN dalam mengisolasi provinsi dengan pola kunjungan yang menyimpang secara signifikan dari tren nasional. Hal ini memberikan tingkat presisi yang lebih tinggi dibandingkan dengan metode pengelompokan konvensional yang sering kali memaksa data pencilan masuk ke dalam kluster tertentu. Melalui visualisasi tersebut, terlihat bahwa kluster kepadatan rendah cenderung berada di provinsi dengan aksesibilitas terbatas. Sebagaimana dinyatakan dalam [8], distribusi spasial dan efisiensi pariwisata sangat dipengaruhi oleh faktor-faktor pendukung regional dan strategi pembangunan berkelanjutan di setiap wilayah.

Pembahasan mendalam terhadap metrik kuantitatif pada fase Pemulihan (2022–2024) membuktikan secara empiris keberadaan 2 *noise points* regional, yang secara spesifik mengisolasi Provinsi Jawa Timur dan Provinsi Nusa Tenggara Barat (NTB) pada aspek koordinat spasial sebaran *scatter plot*. Terisolasinya Provinsi Jawa Timur disebabkan oleh volume kunjungan wisatawan nusantara yang melesat sangat tinggi pascapandemi, membuat jarak statistik *Euclidean* wilayah ini melompat menjauhi kluster Jawa lainnya (Cluster 1). Sementara itu, anomali pada Nusa Tenggara Barat (NTB) dipicu oleh akselerasi *sport tourism* skala internasional (seperti Sirkuit Mandalika) yang mendorong lonjakan perjalanan wisata domestik secara nonlinier, memisahkan karakteristik data spasialnya dari kluster kepulauan luar Jawa (Cluster 0).

Catatan penting juga perlu diberikan pada karakteristik dataset sekunder dari BPS selama periode pemulihan pariwisata Indonesia. Pada beberapa tahun amatan, pencatatan perjalanan wisatawan nusantara di beberapa wilayah terutama pada kluster provinsi kepulauan dan perbatasan (seperti wilayah Papua) mengalami kendala teknis pencatatan atau data hilang (*missing data*) yang ditandai dengan strip (-). Di dalam arsitektur sistem informasi, adanya keterbatasan data riil lapangan ini ditangani secara otomatis melalui proses imputasi nol (`.fillna(0)`). Imputasi ini menyebabkan wilayah-wilayah dengan catatan data kosong tersebut memiliki nilai agregat volume kunjungan yang tergolong rendah. Akibatnya, sistem memasukkan provinsi tersebut ke dalam Cluster 0 bersama dengan daerah lain seperti Provinsi Maluku, Kalimantan Utara, serta wilayah luar Jawa lainnya yang secara riil memiliki karakteristik volume kunjungan rendah yang serupa. Kemampuan DBSCAN menampung variasi anomali data hilang maupun data rendah ini ke dalam satu kelompok murni (Cluster 0) tanpa mengganggu atau membiaskan pusat kerapatan data padat di Pulau Jawa membuktikan ketangguhan model *density-based* ini dibandingkan model konvensional.

Dinamika Spasial Kluster Wisatawan Nusantara per Fase Temporal



Gambar 4. Pemetaan Spasial Hasil Clustering Wisatawan Nusantara Menggunakan SIG-DBSCAN Fase Pra-Pandemi, Pandemi, dan Pemulihan

Selama periode pandemi (2020-2021), terjadi pergeseran signifikan distribusi kluster yang ditandai dengan penurunan konsentrasi pada beberapa wilayah utama dan peningkatan relatif pada destinasi alternatif domestik berskala lebih kecil. Fenomena ini sesuai dengan temuan global yang menyatakan bahwa pandemi COVID-19 secara fundamental mengubah pola mobilitas wisatawan [3], [4], mendorong diversifikasi destinasi domestik sebagai respons terhadap pembatasan perjalanan jarak jauh. Hasil penelitian menunjukkan bahwa kluster alternatif mulai muncul seiring dengan perkembangan sektor pariwisata spesifik. Hal ini sejalan dengan temuan Wang dkk. Yang menyatakan bahwa meskipun terdapat perbedaan indeks pengembangan antarwilayah, kecepatan pertumbuhan di wilayah yang sebelumnya tertinggal menunjukkan potensi besar untuk memperkuat posisi destinasi baru melalui evolusi spasial-temporal yang dinamis [14].

Pada fase pemulihan (2022-2024), terlihat konsolidasi kembali kepadatan kunjungan pada provinsi dengan daya tarik wisata kuat, di mana pemanfaatan data multidimensi dan segmentasi strategis menjadi kunci dalam efektivitas intervensi kebijakan di tingkat regional [11]. Temuan ini dikonfirmasi oleh hasil penelitian ini yang menunjukkan bahwa distribusi kluster pasca-pandemi tidak sepenuhnya identik dengan pola pra-pandemi [5].

C. Discussion

Pendekatan *density-based clustering* terbukti efektif dalam menganalisis dinamika spasial wisatawan pada skala nasional. Algoritma DBSCAN dinilai lebih sesuai untuk data yang bersifat spasial dan heterogen karena bekerja berdasarkan kepadatan data, dengan keunggulan utama pada kemampuannya mendeteksi kluster dengan bentuk yang tidak beraturan serta memisahkan *noise* atau *outlier* secara otomatis tanpa penentuan jumlah kluster secara *a priori* [10], [11]. Karakteristik ini sangat relevan untuk distribusi Wisatawan Nusantara yang tidak merata dan berfluktuasi selama 2019-2024 [1], sebagaimana metode berbasis kepadatan terbukti lebih adaptif untuk data dengan distribusi nonlinier dan kompleks dibandingkan dengan metode berbasis *centroid* seperti K-Means [16].

Berdasarkan evaluasi internal, nilai *Silhouette Score* sebesar 0,430 mengindikasikan bahwa setiap titik data telah ditempatkan dengan baik dalam klusternya sendiri dibandingkan dengan kluster lain [20]. Nilai *Davies-Bouldin Index* (DBI) sebesar 1,889 mencerminkan kompaksi intra-kluster yang tinggi serta pemisahan antar-kluster yang jelas, di mana nilai DBI yang lebih rendah menunjukkan kualitas kluster yang lebih baik [20]. Nilai *Calinski-Harabasz Index* (CHI) sebesar 7,555 memperlihatkan rasio variansi antarkluster terhadap

variansi dalam klaster yang mendukung struktur pengelompokan yang terbentuk [17]. Kombinasi ketiga metrik ini menunjukkan konsistensi performa model, sejalan dengan rekomendasi penggunaan metrik validasi ganda dalam evaluasi hasil *clustering* [19], [20].

Integrasi algoritma *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) dengan Sistem Informasi Geografis (SIG) memungkinkan visualisasi *choropleth* yang informatif dan interpretatif. Representasi visual ini memudahkan identifikasi wilayah dengan konsentrasi tertentu serta membantu memahami pergeseran distribusi antar fase temporal dalam pemetaan efisiensi ekologis [8]. Pendekatan serupa telah diterapkan dalam analisis klaster kebijakan pariwisata, di mana penggunaan DBSCAN terbukti mampu menghasilkan pengelompokan topik yang lebih spesifik dan mendalam dibandingkan dengan metode konvensional [12]. Hasil analisis tersebut memiliki nilai praktis yang signifikan bagi pemangku kepentingan, terutama dalam mendukung perumusan kebijakan ketahanan sektor pariwisata yang adaptif terhadap perubahan perilaku konsumen [5], serta memperkuat manajemen krisis destinasi dalam menghadapi adaptasi kebiasaan baru di era pascapandemi [13].

Keberhasilan pengidentifikasian anomali spasial (Jawa Timur dan NTB) serta penanganan objektif terhadap variansi data luar Jawa memberikan sumbangsih praktis yang berharga bagi perumusan kebijakan pariwisata nasional. Integrasi DBSCAN dan SIG mampu menyajikan realitas dinamika daerah secara transparan tanpa bias pemaksaan data. Temuan ini menjadi basis rekomendasi bagi Kementerian Pariwisata dan Ekonomi Kreatif untuk merumuskan intervensi kebijakan taktis regional yang asimetris, seperti pembukaan rute penerbangan perintis baru untuk mengatasi keterbatasan aksesibilitas pada wilayah-wilayah di Cluster 0 (seperti Papua, Maluku, Kalimantan Utara, dan sekitarnya), penguatan digitalisasi promosi pemasaran pada destinasi tertinggal, hingga manajemen pengendalian beban kapasitas pada wilayah pertumbuhan pariwisata yang terlalu agresif.

Sebagai perbandingan, pengembangan pendekatan berbasis algoritma *Ordering Points To Identify the Clustering Structure* (OPTICS) telah dilakukan untuk mengatasi keterbatasan klasterisasi pada dataset dengan tingkat kepadatan yang tidak merata [18]. Meskipun metode tersebut memiliki keunggulan dalam eksplorasi struktur klaster multiskala, algoritma DBSCAN dipilih dalam penelitian ini karena efisiensi komputasinya, kesederhanaan parameter, serta kemampuan deteksi *noise* yang lebih langsung [16]. Selain itu, pengembangan lebih lanjut dengan menerapkan metode *Spatial Temporal - Density Based Spatial Clustering of Applications with Noise* (ST-DBSCAN) memiliki potensi besar untuk meningkatkan presisi deteksi pergeseran klaster yang melibatkan dimensi ruang dan waktu secara simultan [15].

Secara keseluruhan, model yang dihasilkan menunjukkan performa yang stabil dan representatif dalam mengidentifikasi dinamika spasial wisatawan. Temuan ini memperkuat argumen bahwa metode *density-based clustering* yang terintegrasi dengan SIG memiliki potensi signifikan sebagai instrumen analisis spasial. Pendekatan ini terbukti efektif dalam mendukung perencanaan pariwisata nasional yang adaptif terhadap dinamika temporal serta perubahan pola mobilitas masyarakat [7], [9].

IV. Kesimpulan dan saran

Penelitian ini telah berhasil memetakan dinamika spasial wisatawan Nusantara sepanjang periode 2019-2024 melalui integrasi algoritma DBSCAN dan Sistem Informasi Geografis (SIG). Melalui optimasi parameter dengan *grid search*, ditemukan bahwa kombinasi $\epsilon = 1,4$ dan $MinPts = 2$ merupakan konfigurasi yang paling representatif untuk data ini. Kualitas struktur klaster tersebut dikonfirmasi melalui nilai *Silhouette Score* sebesar 0,430 dan *Davies-Bouldin Index (DBI)* sebesar 1,889. Selain itu, nilai *Calinski-Harabasz Index (CHI)* sebesar 7,555 secara kolektif menunjukkan bahwa model memiliki performa yang stabil dan valid dalam memetakan kelompok data pariwisata. Temuan utama mengungkapkan adanya pergeseran pola distribusi spasial yang signifikan, terutama dipicu oleh krisis kesehatan global, yang ditandai dengan penurunan konsentrasi pada destinasi utama dan diversifikasi menuju destinasi alternatif. Pada fase pemulihan, meskipun terjadi konsolidasi kembali kunjungan pada wilayah dengan daya tarik kuat, distribusi pascapandemi menunjukkan perubahan struktural dalam preferensi wisatawan dibandingkan dengan periode pra-pandemi. Integrasi DBSCAN dengan SIG menghasilkan visualisasi *choropleth* yang informatif untuk mendukung pengambilan keputusan berbasis data bagi para pemangku kepentingan di sektor pariwisata. Keunggulan pendekatan ini terletak pada kemampuan algoritma dalam menangani distribusi data nonlinier dan mendeteksi *noise* secara otomatis tanpa asumsi jumlah klaster di awal. Penelitian ini turut memperluas literatur mengenai penerapan DBSCAN untuk pemetaan wilayah di Indonesia berdasarkan indikator strategis. Meskipun demikian, keterbatasan dalam hal granularitas data yang masih pada tingkat provinsi menjadi catatan penting. Oleh karena itu, penelitian lanjutan disarankan untuk menerapkan metode Spatial-Temporal (ST-DBSCAN) pada skala data yang lebih mikro, serta mengeksplorasi faktor aksesibilitas, strategi digital pariwisata, dan manajemen krisis destinasi guna memperdalam pemahaman terhadap dinamika mobilitas wisatawan di masa depan.

Daftar Pustaka

- [1] Badan Pusat Statistik, “Jumlah Perjalanan Wisatawan Nusantara Menurut Provinsi Tujuan (Perjalanan),” Badan Pusat Statistik. Accessed: Mar. 02, 2026. [Online]. Available: <https://www.bps.go.id/id/statistics-table/2/MjIwMSMy/jumlah-perjalanan-wisatawan-nusantara-menurut-provinsi-tujuan--perjalanan-.html>
- [2] S. Haryono, “Socio e-Kons Analisis Kontribusi Pariwisata Terhadap Pertumbuhan Ekonomi dan Dampaknya Terhadap IPM di Pulau Jawa,” vol. 18, p. 1, 2026, doi: <https://doi.org/10.30998/0ce7zp51>.
- [3] A. Syaifudin, H. Hendarmawan, and E. Novianti, “Impact of COVID-19 on the Global Tourism Economy,” *J. Tour. Educ.*, vol. 2, no. 1, pp. 1–9, 2022, doi: 10.17509/jote.v2i1.41629.
- [4] Meitolo Hulu and Lidia Ariyanti, “Kepariwisata di masa krisis: kunjungan wisata pada masa pandemi covid-19 di pantai Pangandaran, Indonesia,” *J. Tour. Econ.*, vol. 5, no. 2, pp. 178–185, Dec. 2022, doi: 10.36594/jtec/2ffbbq71.
- [5] D. Mayasari and I. Zubaidah, “The analysis of consumer behavior and tourism resilience post-pandemic covid-19 in Belitung Regency,” *J. Pariwisata Pesona*, vol. 10, no. 1, pp. 40–49, Jul. 2025, doi: 10.26905/jpp.v10i1.13474.
- [6] A. Iriany, N. A. Sovia, W. Ngabu, and D. Ariyanto, “Analisis Spasial Temporal Indeks Ketahanan Pangan Menggunakan Geographically Temporally Weighted Regression dan Klaster Fuzzy di Nusa Tenggara Timur,” *MaTHunesa J. Ilm. Mat.*, vol. 13, no. 03, pp. 9–17, 2025.
- [7] E. R. S. Moningkey, N. Timang, and K. Santa, “Spatial Clustering Of Tourism Investment Potential In Tomohon City,” *Acad. Open*, vol. 11, no. 1, pp. 1–19, 2026, doi: 10.21070/acopen.11.2026.13038.
- [8] C. Feng and H. R. Wang, “Measurement, Temporal and Spatial Evolution of Tourism Eco-Efficiency in China Based on Data Mining,” *Appl. Ecol. Environ. Res.*, vol. 23, no. 5, pp. 8895–8912, 2025, doi: 10.15666/aecr/2305_88958912.
- [9] J. Han, J. Pei, and H. Tong, *Data Mining: Concepts and Techniques*, 4th ed. Cambridge: Morgan Kaufmann, 2023. [Online]. Available: https://books.google.com/books/about/Data_Mining.html?id=NR1oEAAAQBAJ
- [10] A. M. A. Tambunan, U. Khaira, and B. F. Hutabarat, “Implementasi Algoritma DBSCAN untuk Klasterisasi Koperasi Berdasarkan Indikator Keuangan,” *NARATIF J. Ilm. Nas. Ris. Apl. dan Tek. Inform.*, vol. 7, no. 1, pp. 38–55, 2025.
- [11] K. Emilia, K. M. Hindrayani, and A. R. Pratama, “Penerapan Algoritma Density-Based Spatial Clustering of Application with Noise untuk Analisis Multidimensi Data Destinasi Wisata Provinsi Jawa Tengah,” in *Prosiding Seminar Nasional Sains dan Teknologi Seri V Fakultas Sains dan Teknologi, Universitas Terbuka*, Fakultas Sains dan Teknologi, Universitas Terbuka, 2026, pp. 110–124.
- [12] R. Wulandari and W. Yustanti, “Analisis Text Clustering Kebijakan Pembukaan Daerah Wisata pada Masa Pandemi Berbasis Densitas Spasial (DBSCAN),” *J. Emerg. Inf. Syst. Bus. Intell.*, vol. 3, no. 2, pp. 1–10, 2022, doi: 10.26740/jeisbi.v3i2.45509.
- [13] L. W. Putri, A. Z. Rahman, and Kismartini, “Manajemen Krisis Kepariwisata pada Objek Wisata Umbul Pongkok dalam Menghadapi Adaptasi Kebiasaan Baru di Era Pandemi COVID-19,” *J. Manaj. dan Kebijak. Publik*, vol. 12, no. 2, 2023, doi: <https://doi.org/10.14710/jppmr.v12i2.37770>.
- [14] H. Wang *et al.*, “Analysis of spatial-temporal pattern, dynamic evolution and influencing factors of health tourism development in China,” *Sci. Rep.*, vol. 13, no. 1, p. 15436, Sep. 2023, doi: 10.1038/s41598-023-42462-x.
- [15] S. A. Hasibuan, K. M. Hindrayani, and S. S. M. Wara, “Analisis Daerah Rawan Kebakaran di Kota Surabaya Menggunakan Metode Spatial Temporal Density Based Spatial Clustering of Applications with Noise,” in *Prosiding Seminar Nasional Sains dan Teknologi Seri V Fakultas Sains dan Teknologi, Universitas Terbuka*, Fakultas Sains dan Teknologi, Universitas Terbuka, 2026, pp. 653–663.
- [16] G. A. Rahman, I. Hasbi, A. Tenriawaru, L. Surimi, and A. N. Alfian, “Penggunaan Algoritma DbSCAN Dalam Pengelompokan Kabupaten/Kota Di Sulawesi Tenggara Berdasarkan Indikator Pendidikan,” *Simtek J. Sist. Inf. dan Tek. Komput.*, vol. 10, no. 1, pp. 184–193, 2025, doi: 10.51876/simtek.v10i1.1546.
- [17] A. D. Rabbani, K. M. Hindrayani, and M. Nasrudin, “Perbandingan K-Means, DBSCAN, dan OPTICS untuk Klasterisasi Pasien Anemia Berdasarkan Parameter Hematologi,” in *Data Science, Informatics, Statistics, and Electrical Research Conference*, 2025, p. 01.
- [18] Y. Hasan, “Pengukuran Silhouette Score dan Davies-Bouldin Index pada Hasil Cluster K-Means dan DbSCAN,” *KAKIFIKOM (Kumpulan Artik. Karya Ilm. Fak. Ilmu Komputer)*, vol. 06, no. 01, 2024.
- [19] S. Mutiah, Y. Hasnataeni, A. Fitrianto, E. Erfiani, and L. M. R. D. Jumansyah, “Perbandingan Metode

- Klastering K-Means dan DBSCAN dalam Identifikasi Kelompok Rumah Tangga Berdasarkan Fasilitas Sosial Ekonomi di Jawa Barat,” *Teorema Teor. dan Ris. Mat.*, vol. 9, no. 2, p. 247, Sep. 2024, doi: 10.25157/teorema.v9i2.16290.
- [20] A. D. Laksmi, D. Kusnandar, and N. Imro’ah, “Hierarchical clustering multiscale bootstrap pada capaian dimensi IPM tahun 2020 yang terdampak COVID-19,” *Bul. Ilm. Mat. Stat. dan Ter.*, vol. 12, no. 2, pp. 185–194, 2023.