



Segmentasi Kabupaten/Kota Di Jawa Timur Tahun 2019-2024 Berdasarkan Kunjungan Wisatawan Dan TPT Dengan K-Means Clustering

Muhammad Maulana Yusuf ¹, Radithya Hylmi Maulana Putra ², Moch Firman
Fernando ³, Alfian Rizaldy Pratama ⁴, Aviolla Terza Damaliana ⁵

^{1,2,3,4,5}Program Studi Sains Data, Fakultas Ilmu Komputer, Universitas Pembangunan
Nasional Veteran Jawa Timur

E-mail: 23083010094@student.upnjatim.ac.id^{*}

ABSTRACT

The tourism sector plays a crucial role in supporting regional economic growth, yet it does not always align with a reduction in unemployment rates. This study analyzes the segmentation of regencies/cities in East Java Province based on the number of tourist visits and the open unemployment rate (TPT) from 2019 to 2024 using the K-Means Clustering method. The determination of the optimal number of clusters was conducted using the Elbow Method and Silhouette Score, leading to the selection of three clusters. The analysis results show that the first cluster consists of regions with low tourist visits and low TPT, the second cluster includes areas with high tourist visits but also high TPT, and the third cluster reflects regions with moderate values in both indicators. These findings indicate that tourism growth does not always have a direct impact on reducing unemployment. The results of this segmentation can serve as a basis for formulating more targeted and equitable regional development policies.

Keywords: K-Means, Tourism, Unemployment, Clustering, East Java

ABSTRAK

Sektor pariwisata berperan penting dalam mendukung pertumbuhan ekonomi daerah, namun tidak selalu sejalan dengan penurunan tingkat pengangguran. Penelitian ini menganalisis segmentasi kabupaten/kota di Provinsi Jawa Timur berdasarkan jumlah kunjungan wisatawan dan tingkat pengangguran terbuka (TPT) periode 2019-2024 menggunakan metode K-Means Clustering. Penentuan jumlah cluster optimal dilakukan dengan Elbow Method dan Silhouette Score, yang mengarah pada pemilihan tiga klaster. Hasil analisis menunjukkan bahwa klaster pertama terdiri dari wilayah dengan kunjungan dan TPT rendah, klaster kedua mencakup daerah dengan kunjungan tinggi namun TPT juga tinggi, dan klaster ketiga mencerminkan daerah dengan nilai sedang pada kedua indikator. Temuan ini menunjukkan bahwa pertumbuhan pariwisata tidak selalu berdampak langsung terhadap penurunan pengangguran. Hasil segmentasi ini dapat menjadi dasar bagi perumusan kebijakan pembangunan daerah yang lebih terarah dan berkeadilan.

Article History

Received: Juni 2025

Reviewed: Juni 2025

Published: Juni 2025

Plagiarism Checker No
235

Prefix DOI :

[10.8734/Kohesi.v1i2.365](https://doi.org/10.8734/Kohesi.v1i2.365)

Copyright : Author

Publish by : Kohesi



This work is licensed
under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



Kata Kunci: <i>K-Means, Pariwisata, Pengangguran, Clustering, Jawa Timur</i>	
---	--

1. PENDAHULUAN

Pariwisata memiliki peran strategis dalam mendorong pertumbuhan ekonomi dan pemerataan pembangunan daerah. Menurut Yoeti (2008), pariwisata tidak hanya berkontribusi terhadap pendapatan daerah tetapi juga menciptakan lapangan kerja dan menggerakkan sektor ekonomi lainnya. Provinsi Jawa Timur, sebagai salah satu tujuan utama pariwisata di Indonesia, mencatat peningkatan signifikan dalam kunjungan wisatawan domestik selama beberapa tahun terakhir. Di sisi lain, permasalahan sosial seperti Tingkat Pengangguran Terbuka (TPT) juga menjadi perhatian, di mana pada tahun 2022 tercatat TPT di Jawa Timur mencapai 5,49% (BPS, 2023). Tingginya kunjungan wisata seharusnya berkorelasi positif terhadap peningkatan lapangan kerja, namun hal ini tidak selalu tercermin secara merata di seluruh kabupaten/kota.

Kondisi tersebut menunjukkan pentingnya pemetaan karakteristik wilayah untuk mendukung kebijakan pembangunan yang lebih tepat sasaran. Salah satu pendekatan yang dapat digunakan untuk mengungkap pola tersembunyi antar wilayah adalah segmentasi berbasis data menggunakan metode clustering. Metode K-Means Clustering merupakan salah satu algoritma partisi yang paling populer dalam analisis segmentasi wilayah karena kesederhanaan dan efektivitasnya dalam mengelompokkan data berdimensi banyak (Jain, 2010). Selain itu, pendekatan berbasis data seperti ini telah terbukti mampu meningkatkan efektivitas perencanaan daerah melalui pemanfaatan sumber daya yang berbasis pada kebutuhan aktual wilayah (Zhang & Zhang, 2020).

Penelitian sebelumnya oleh Pratama et al. (2022) menunjukkan keberhasilan penggunaan K-Means dalam pengelompokan wilayah berdasarkan indikator ekonomi dan sosial. Namun, masih terbatas kajian yang mengintegrasikan data pariwisata dan ketenagakerjaan secara simultan, khususnya dalam konteks daerah di Indonesia. Oleh karena itu, penelitian ini bertujuan untuk melakukan segmentasi kabupaten/kota di Provinsi Jawa Timur berdasarkan data kunjungan wisatawan dan TPT dari tahun 2019 hingga 2024 menggunakan metode K-Means Clustering, guna mengidentifikasi pola yang relevan bagi pengambilan kebijakan pembangunan daerah.



2. METODE PENELITIAN

2.1 Data dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder bersifat kuantitatif. Data sekunder adalah data yang diperoleh secara tidak langsung melalui pihak lain dan telah dikumpulkan serta dipublikasikan untuk keperluan tertentu. Data ini tidak diperoleh langsung dari pengamatan lapangan, melainkan diakses melalui publikasi resmi lembaga terkait. Data yang dikumpulkan dalam penelitian ini meliputi:

- a. Jumlah kunjungan wisatawan per tahun
- b. Tingkat Pengangguran Terbuka (TPT) dalam persen
- c. Wilayah: Kabupaten/Kota di Provinsi Jawa Timur
- d. Periode: Tahun 2019 hingga 2024

Sumber utama data dalam penelitian ini adalah Badan Pusat Statistik (BPS), lembaga resmi pemerintah yang menyediakan data statistik nasional. Data kunjungan wisatawan diperoleh dari publikasi statistik pariwisata daerah, sedangkan data TPT diperoleh dari publikasi ketenagakerjaan tahunan. Menurut BPS (2022), data yang dihimpun berasal dari survei dan kompilasi administrasi berbagai instansi, yang telah melalui proses validasi untuk menggambarkan kondisi sosial ekonomi masyarakat di berbagai wilayah.

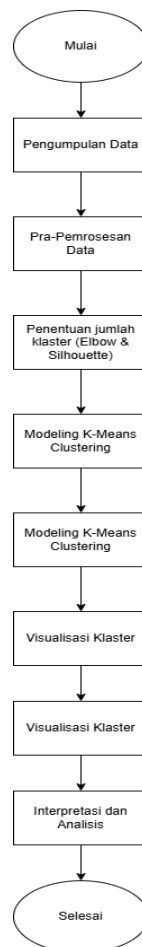
2.2 Tahapan Analisis

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis kluster tidak terawasi (unsupervised clustering), yaitu K-Means Clustering, untuk melakukan segmentasi kabupaten/kota di Provinsi Jawa Timur berdasarkan data kunjungan wisatawan dan tingkat pengangguran terbuka (TPT) pada periode 2019-2024. Metode K-Means dipilih karena kemampuannya dalam mengelompokkan data ke dalam sejumlah kluster berdasarkan kemiripan karakteristik secara efisien dan intuitif. Menurut Siregar & Nursal (2020), K-Means merupakan metode partisi yang populer karena kesederhanaan dan kecepatan proses komputasinya dalam membagi data ke dalam k kelompok berdasarkan jarak Euclidean sebagai fungsi objektif. Kusumadewi (2010) menegaskan bahwa metode ini bekerja optimal pada data dengan skala numerik yang telah dinormalisasi sehingga tiap fitur memiliki kontribusi yang setara. Sementara itu, menurut Wardhani et al. (2021), K-Means efektif untuk segmentasi wilayah berbasis data multivariat karena hasil klasternya dapat dengan mudah diinterpretasikan dan divisualisasikan.

Dalam penelitian ini, data dianalisis menggunakan Google Colab dengan bahasa pemrograman Python. Tahapan yang dilakukan meliputi: pra-pemrosesan data (normalisasi), penentuan jumlah cluster optimal menggunakan metode Elbow dan Silhouette Score,

penerapan algoritma K-Means, serta visualisasi hasil klasterisasi menggunakan scatter plot dan peta wilayah. Penentuan jumlah klaster merupakan langkah penting dalam analisis K-Means karena secara langsung memengaruhi hasil segmentasi wilayah. Dalam penelitian ini, digunakan dua metode evaluasi internal untuk memperoleh jumlah cluster yang paling sesuai, yaitu Elbow Method dan Silhouette Score. Elbow Method digunakan untuk mengamati perubahan nilai *within-cluster sum of squares* (WCSS) pada berbagai nilai klaster (k), dan titik “siku” pada grafik menunjukkan nilai k optimal. Menurut penelitian oleh Herdiana et al. (2025), Elbow Method dapat digunakan untuk menentukan jumlah cluster yang optimal dengan mengidentifikasi titik siku pada grafik WCSS. Sementara itu, Silhouette Score mengukur konsistensi internal masing-masing klaster dengan membandingkan kedekatan antar anggota dalam klaster dan antar klaster lainnya. Penelitian oleh Nugroho et al. (2022) menunjukkan bahwa nilai Silhouette Score yang tinggi mengindikasikan kualitas klaster yang baik dalam analisis K-Means. Berdasarkan hasil kedua metode tersebut, diputuskan bahwa jumlah cluster optimal dalam penelitian ini adalah tiga klaster ($k = 3$). Berikut adalah flowchart penelitian yang menggambarkan tahapan-tahapan yang dilakukan selama proses penelitian berlangsung:

Gambar 1 : Flowchart Proses Analisis





Flowchart ini menggambarkan alur proses analisis data menggunakan metode K-Means Clustering yang digunakan untuk melakukan segmentasi wilayah kabupaten/kota di Jawa Timur berdasarkan data jumlah kunjungan wisatawan dan tingkat pengangguran terbuka (TPT) periode 2019-2024. Berikut tahapan-tahapan dalam flowchart:

1. Mulai
2. Pengumpulan data

Data dikumpulkan dari sumber resmi, yaitu Badan Pusat Statistik (BPS). Data yang digunakan mencakup jumlah kunjungan wisatawan per tahun dan tingkat pengangguran terbuka dari tahun 2019 hingga 2024 untuk masing-masing kabupaten/kota di Jawa Timur.

3. Pemrosesan Data

Tahapan ini meliputi pembersihan data (cleaning), penanganan missing values, konversi tipe data, dan normalisasi agar data berada dalam skala yang seragam dan layak untuk dianalisis dengan metode K-Means.

4. Penentuan Jumlah Cluster (Elbow & Silhouette)

Sebelum menerapkan algoritma K-Means, dilakukan analisis untuk menentukan jumlah cluster optimal menggunakan dua pendekatan: metode Elbow dan koefisien Silhouette. Hasil dari kedua metode ini menjadi dasar dalam menentukan nilai k terbaik.

5. Modeling K-Means Clustering

Algoritma K-Means diterapkan pada data yang telah dipra-pemrosesan. Proses ini melibatkan pembentukan klaster berdasarkan kemiripan pola dari dua variabel utama (kunjungan wisatawan dan TPT).

6. Visualisasi Cluster
7. Interpretasi dan Analisis
8. Selesai

3. HASIL DAN PEMBAHASAN

3.1 Statistika Deskriptif

Sebelum dilakukan proses pengelompokan (clustering), langkah awal dalam analisis ini adalah memahami karakteristik data melalui statistik deskriptif. Statistik deskriptif memberikan gambaran umum tentang persebaran dan kecenderungan nilai-nilai variabel utama, yaitu jumlah kunjungan wisatawan dan tingkat pengangguran terbuka (TPT) di kabupaten/kota Provinsi Jawa Timur selama periode 2019-2024. Analisis ini penting untuk mengidentifikasi adanya penyimpangan nilai (outlier), pola distribusi data, dan kecenderungan umum yang dapat memengaruhi hasil segmentasi wilayah. Tabel berikut menyajikan nilai minimum, maksimum, rata-rata, median, dan sebaran kuartil untuk masing-masing variabel. Berikut tabel statistika deskriptif:

Tabel 1. Pengelompokan kabupaten/kota di Jawa Timur Cluster 0

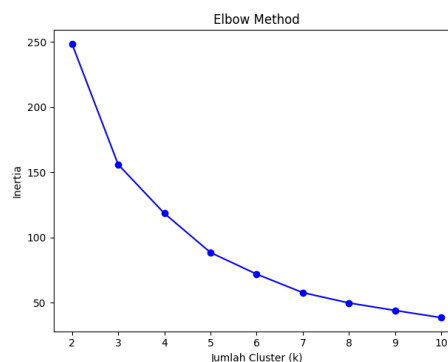
Variabel	Mean	Minimum	Maksimal	Standar Deviasi
Total Kunjungan (Orang)	4.269.537	326.261	33.126.070	4.636.658
TPT(Persen)	4,80	0,91	10,97	1,77

Berdasarkan hasil analisis deskriptif terhadap 228 observasi kabupaten/kota di Jawa Timur pada periode 2019-2024, diperoleh bahwa rata-rata jumlah kunjungan wisatawan mencapai sekitar 4.269.537 orang. Jumlah kunjungan minimum tercatat sebesar 326.261 orang, sedangkan jumlah maksimum mencapai 33.126.070 orang. Nilai standar deviasi sebesar 4.636.658 menunjukkan adanya variasi yang cukup besar antar wilayah, sehingga dapat disimpulkan bahwa distribusi jumlah kunjungan wisatawan cenderung tidak merata. Sementara itu, rata-rata Tingkat Pengangguran Terbuka (TPT) sebesar 4,80% dengan nilai minimum sebesar 0,91% dan maksimum sebesar 10,97%. Standar deviasi TPT sebesar 1,77 menunjukkan adanya perbedaan tingkat pengangguran yang relatif moderat antar kabupaten/kota di Jawa Timur selama periode yang dianalisis.

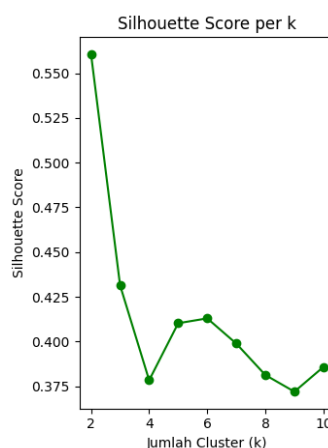
3.2 Penentuan Jumlah Cluster Optimal

Sebelum menerapkan K-Means, dilakukan penentuan jumlah cluster optimal dengan menggunakan Elbow Method dan Silhouette Score. Hasil visualisasi pada Gambar 2 dan Gambar 3 menunjukkan bahwa jumlah cluster terbaik adalah tiga ($k = 3$), ditandai dengan adanya “siku” pada grafik Elbow dan nilai Silhouette yang cukup tinggi.

Gambar 2 : Elbow Method



Gambar 3 : Silhouette Score





Berdasarkan grafik Elbow Method (Gambar 2), terlihat adanya titik belok yang cukup jelas pada $k = 3$, yang mengindikasikan bahwa menambahkan jumlah kluster setelah titik tersebut tidak lagi secara signifikan menurunkan nilai inertia. Hal ini menunjukkan bahwa tiga kluster sudah cukup untuk merepresentasikan struktur data secara efisien. Sementara itu, hasil pada grafik Silhouette Score (Gambar 3) juga menunjukkan bahwa nilai koefisien Silhouette relatif tinggi pada $k = 3$, meskipun nilai tertinggi terdapat pada $k = 2$. Namun, karena tujuan analisis adalah untuk menghasilkan segmentasi yang cukup kaya (tidak terlalu sedikit), maka $k = 3$ dipilih sebagai kompromi terbaik antara stabilitas kluster dan jumlah kluster yang representatif.

3.3 Hasil Klasterisasi K-Means

Penelitian ini menemukan bahwa kabupaten/kota di Jawa Timur dapat dikelompokkan menjadi tiga kluster utama berdasarkan pola kunjungan wisatawan dan tingkat pengangguran terbuka (TPT). Kluster pertama (Cluster 0) menggambarkan wilayah dengan tingkat kunjungan wisatawan yang rendah namun memiliki tingkat pengangguran yang juga rendah. Hal ini menunjukkan bahwa aktivitas ekonomi di wilayah-wilayah tersebut tidak bergantung pada sektor pariwisata, melainkan ditopang oleh sektor ekonomi lokal lainnya, seperti Kabupaten Pacitan Tahun 2019 dan Kabupaten Probolinggo Tahun 2019. Sementara itu, kluster kedua (Cluster 1) menunjukkan karakteristik wilayah dengan tingkat kunjungan wisata yang sangat tinggi, tetapi diiringi oleh tingkat pengangguran yang juga tinggi. Fenomena ini dapat mengindikasikan bahwa meskipun pariwisata berkembang pesat, manfaat ekonominya belum terdistribusi secara merata ke masyarakat, atau terdapat ketimpangan dalam akses terhadap lapangan kerja, seperti Kota Surabaya Tahun 2020 dan Kabupaten Sidoarjo Tahun 2020. Kluster ketiga (Cluster 2) mencerminkan wilayah-wilayah dengan tingkat kunjungan dan TPT yang berada pada kategori sedang. Wilayah-wilayah ini dapat dikatakan berada dalam masa transisi dan berpotensi berkembang lebih lanjut apabila dikelola dengan tepat, seperti Kabupaten Gresik Tahun 2019 dan Kabupaten Bangkalan Tahun 2020. berikut menyajikan hasil pengelompokan kabupaten/kota di Jawa Timur ke dalam tiga kluster berdasarkan indikator jumlah kunjungan wisatawan dan tingkat pengangguran terbuka (TPT) selama periode 2019-2024:

a) Cluster 0: Kunjungan Wisata Rendah Pengangguran Rendah

Tabel 2. Pengelompokan kabupaten/kota di Jawa Timur Cluster 0

Kabupaten/Kota	Tahun
KabupatenPacitan	2019



Kabupaten Ponorogo	2019
Kabupaten Trenggalek	2019
Kabupaten Tulungagung	2019
Kabupaten Blitar	2019
Kabupaten Kediri	2019
Kabupaten Malang	2019
Kabupaten Lumajang	2019
Kabupaten Jember	2019
Kabupaten Banyuwangi	2019
Kabupaten Bondowoso	2019
Kabupaten Situbondo	2019
Kabupaten Probolinggo	2019
Kabupaten Pasuruan	2019
Kabupaten Sidoarjo	2019
Kabupaten Mojokerto	2019
Kabupaten Jombang	2019
Kabupaten Nganjuk	2019
Kabupaten Madiun	2019
Kabupaten Magetan	2019
Kabupaten Ngawi	2019
Kabupaten Bojonegoro	2019



Kabupaten Tuban	2019
Kabupaten Lamongan	2019
Kabupaten Sampang	2019
Kabupaten Pamekasan	2019
Kabupaten Sumenep	2019
Kota Kediri	2019
Kota Blitar	2019
Kota Probolinggo	2019
Kota Pasuruan	2019
Kota Mojokerto	2019
Kota Madiun	2019
Kota Batu	2019
Kabupaten Pacitan	2020
Kabupaten Ponorogo	2020
Kabupaten Trenggalek	2020
Kabupaten Tulungagung	2020
Kabupaten Blitar	2020
Kabupaten Lumajang	2020
Kabupaten Bondowoso	2020
Kabupaten Situbondo	2020
Kabupaten Probolinggo	2020



Kabupaten Nganjuk	2020
Kabupaten Madiun	2020
Kabupaten Magetan	2020
Kabupaten Bojonegoro	2020
Kabupaten Tuban	2020
Kabupaten Sampang	2020
Kabupaten Pamekasan	2020
Kabupaten Sumenep	2020
Kabupaten Pacitan	2021
Kabupaten Ponorogo	2021
Kabupaten Trenggalek	2021
Kabupaten Tulungagung	2021
Kabupaten Blitar	2021
Kabupaten Lumajang	2021
Kabupaten Bondowoso	2021
Kabupaten Situbondo	2021
Kabupaten Probolinggo	2021
Kabupaten Nganjuk	2021
Kabupaten Madiun	2021
Kabupaten Magetan	2021
Kabupaten Ngawi	2021



Kabupaten Bojonegoro	2021
Kabupaten Tuban	2021
Kabupaten Sampang	2021
Kabupaten Pamekasan	2021
Kabupaten Sumenep	2021
Kabupaten Pacitan	2022
Kabupaten Lumajang	2022
Kabupaten Jember	2022
Kabupaten Bondowoso	2022
Kabupaten Situbondo	2022
Kabupaten Probolinggo	2022
Kabupaten Nganjuk	2022
Kabupaten Magetan	2022
Kabupaten Ngawi	2022
Kabupaten Bojonegoro	2022
Kabupaten Tuban	2022
Kabupaten Sampang	2022
Kabupaten Pamekasan	2022
Kabupaten Sumenep	2022
Kota Kediri	2022
Kota Probolinggo	2022
Kota Mojokerto	2022



Kabupaten Pacitan	2023
Kabupaten Ponorogo	2023
Kabupaten Trenggalek	2023
Kabupaten Lumajang	2023
Kabupaten Jember	2023
Kabupaten Banyuwangi	2023
Kabupaten Bondowoso	2023
Kabupaten Situbondo	2023
Kabupaten Probolinggo	2023
Kabupaten Nganjuk	2023
Kabupaten Magetan	2023
Kabupaten Ngawi	2023
Kabupaten Bojonegoro	2023
Kabupaten Tuban	2023
Kabupaten Sampang	2023
Kabupaten Pamekasan	2023
Kabupaten Sumenep	2023
Kota Kediri	2023
Kota Blitar	2023
Kota Probolinggo	2023
Kota Mojokerto	2023



Kota Batu	2023
Kabupaten Pacitan	2024
Kabupaten Ponorogo	2024
Kabupaten Trenggalek	2024
Kabupaten Tulungagung	2024
Kabupaten Lumajang	2024
Kabupaten Jember	2024
Kabupaten Banyuwangi	2024
Kabupaten Bondowoso	2024
Kabupaten Situbondo	2024
Kabupaten Probolinggo	2024
Kabupaten Mojokerto	2024
Kabupaten Jombang	2024
Kabupaten Nganjuk	2024
Kabupaten Madiun	2024
Kabupaten Magetan	2024
Kabupaten Ngawi	2024
Kabupaten Bojonegoro	2024
Kabupaten Tuban	2024
Kabupaten Lamongan	2024



Kabupaten Sampang	2024
Kabupaten Pamekasan	2024
Kabupaten Sumenep	2024
Kota Kediri	2024
Kota Blitar	2024
Kota Probolinggo	2024
Kota Pasuruan	2024
Kota Mojokerto	2024
Kota Madiun	2024
Kota Batu	2024

b) Cluster 1: Kunjungan Wisata Tinggi Pengangguran Tinggi

Tabel 3. Pengelompokan kabupaten/kota di Jawa Timur Cluster

Kabupaten/Kota	Tahun
Kabupaten Sidoarjo	2020
Kota Surabaya	2020
Kabupaten Sidoarjo	2021
Kota Surabaya	2021
Kabupaten Malang	2022
Kabupaten Sidoarjo	2022
Kota Surabaya	2022
Kabupaten Sidoarjo	2023
Kota Surabaya	2023
Kabupaten Sidoarjo	2024



Kota Surabaya 2024

C) Cluster 2: Sedang di kedua indikator

Tabel 4. Pengelompokan kabupaten/kota di Jawa Timur Cluster

Kabupaten/Kota	Tahun
Kabupaten Gresik	2019
Kabupaten Bangkalan	2019
Kota Malang	2019
Kota Surabaya	2019
Kabupaten Kediri	2019
Kabupaten Malang	2019
Kabupaten Jember	2020
Kabupaten Banyuwangi	2020
Kabupaten Pasuruan	2020
Kabupaten Mojokerto	2020
Kabupaten Jombang	2020
Kabupaten Ngawi	2020
Kabupaten Lamongan	2020
Kabupaten Gresik	2020
Kabupaten Bangkalan	2020



Kota Kediri	2020
Kota Blitar	2020
Kota Malang	2020
Kota Probolinggo	2020
Kota Pasuruan	2020
Kota Mojokerto	2020
Kota Madiun	2020
Kota Batu	2020
Kabupaten Kediri	2020
Kabupaten Malang	2020
Kabupaten Jember	2021
Kabupaten Banyuwangi	2021
Kabupaten Pasuruan	2021
Kabupaten Mojokerto	2021
Kabupaten Jombang	2021
Kabupaten Lamongan	2021
Kabupaten Gresik	2021
Kabupaten Bangkalan	2021
Kota Kediri	2021
Kota Blitar	2021



Kota Malang	2021
Kota Probolinggo	2021
Kota Pasuruan	2021
Kota Mojokerto	2021
Kota Madiun	2021
Kota Batu	2021
Kabupaten Ponorogo	2021
Kabupaten Trenggalek	2021
Kabupaten Tulungagung	2022
Kabupaten Blitar	2022
Kabupaten Kediri	2022
Kabupaten Banyuwangi	2022
Kabupaten Pasuruan	2022
Kabupaten Mojokerto	2022
Kabupaten Jombang	2022
Kabupaten Madiun	2022
Kabupaten Lamongan	2022
Kabupaten Gresik	2022
Kabupaten	2022



Bangkalan	
Kota Blitar	2022
Kota Malang	2022
Kota Pasuruan	2022
Kota Madiun	2022
Kota Batu	2022
Kabupaten Tulungagung	2022
Kabupaten Blitar	2022
Kabupaten Tulungagung	2022
Kabupaten Blitar	2022
Kabupaten Kediri	2023
Kabupaten Malang	2023
Kabupaten Pasuruan	2023
Kabupaten Mojokerto	2023
Kabupaten Jombang	2023
Kabupaten Madiun	2023
Kabupaten Lamongan	2023
Kabupaten Gresik	2023
Kabupaten Bangkalan	2023



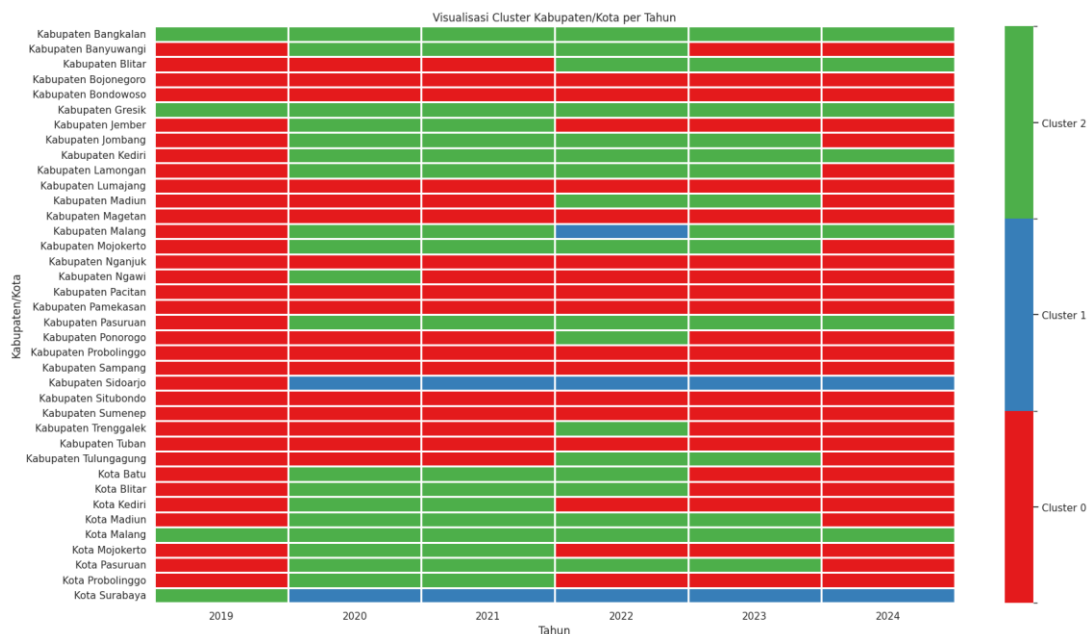
Kota Malang	2023
Kota Pasuruan	2023
Kota Madiun	2023
Kabupaten Blitar	2023
Kabupaten Kediri	2023
Kabupaten Malang	2024
Kabupaten Pasuruan	2024
Kabupaten Gresik	2024
Kabupaten Bangkalan	2024
Kota Malang	2024

Berdasarkan tabel keanggotaan klaster, terlihat bahwa Cluster 0 banyak diisi oleh kabupaten/kota dengan kunjungan wisatawan dan TPT yang rendah, seperti Pacitan dari tahun 2019-2024 pada awal periode, menunjukkan dominasi wilayah dengan basis ekonomi lokal non-pariwisata. Cluster 1 mencakup wilayah dengan kunjungan wisatawan sangat tinggi namun TPT juga tinggi, seperti Kota Surabaya dari tahun 2020-2024, yang cenderung konsisten berada dalam klaster ini di beberapa tahun, mencerminkan ketimpangan distribusi manfaat pariwisata. Sementara itu, Cluster 2 berisi wilayah yang berada di tengah-tengah kedua ekstrem, seperti Gresik dari tahun 2019-2024, yang menunjukkan adanya potensi transisi dan perkembangan ekonomi. Pergeseran beberapa daerah dari satu klaster ke lainnya menunjukkan adanya dinamika pembangunan, efek kebijakan lokal, atau dampak eksternal seperti pandemi yang memengaruhi struktur ekonomi dan pariwisata daerah.

3.3 Visualisasi Hasil Klaster

Visualisasi berikut menampilkan hasil segmentasi kabupaten/kota di Provinsi Jawa Timur ke dalam tiga klaster berdasarkan data jumlah kunjungan wisatawan dan tingkat pengangguran terbuka (TPT) selama periode 2019-2024. Masing-masing warna merepresentasikan satu klaster: merah untuk Cluster 0, biru untuk Cluster 1, dan hijau untuk Cluster 2. Setiap baris menunjukkan satu kabupaten/kota, sedangkan kolom menggambarkan tahun. Warna pada setiap sel mengindikasikan klaster tempat daerah tersebut tergolong pada tahun tersebut.

Visualisasi ini memudahkan dalam mengamati perubahan pola kluster dari waktu ke waktu, serta mengidentifikasi daerah-daerah yang konsisten berada dalam satu kluster maupun yang mengalami perpindahan kluster antar tahun. Berikut gambar visualisasi hasil cluser nya:



Dari visualisasi tersebut terlihat bahwa Cluster 0 (merah) merupakan kluster yang paling dominan di sebagian besar kabupaten/kota dan tahun. Ini menunjukkan bahwa mayoritas wilayah memiliki tingkat kunjungan wisatawan yang rendah dan pengangguran yang rendah, yang mencerminkan daerah dengan aktivitas ekonomi lokal yang stabil namun belum menjadi destinasi wisata utama. Cluster 1 (biru) meskipun lebih sedikit, secara konsisten mencakup kota-kota besar seperti Surabaya dari tahun 2020-2024, dan Sidoarjo dari tahun 2020-2024. Wilayah dalam kluster ini memiliki kunjungan wisatawan yang sangat tinggi namun pengangguran juga tinggi, yang mengindikasikan bahwa manfaat ekonomi dari sektor wisata belum merata dirasakan oleh penduduk lokal. Cluster 2 (hijau) menunjukkan anggota yang lebih bervariasi setiap tahunnya, termasuk daerah seperti Kabupaten Banyuwangi 2020-2022. Wilayah dalam kluster ini berada pada kategori sedang untuk kedua indikator, mencerminkan daerah yang sedang berkembang atau berada dalam fase transisi ekonomi. Perubahan kluster dari tahun ke tahun pada beberapa wilayah mengindikasikan adanya dinamika pembangunan, intervensi kebijakan lokal, atau faktor eksternal seperti pandemi yang memengaruhi kinerja pariwisata dan ketenagakerjaan.

KESIMPULAN

Penelitian ini berhasil mengelompokkan kabupaten/kota di Provinsi Jawa Timur ke dalam tiga kluster berdasarkan data kunjungan wisatawan dan tingkat pengangguran terbuka (TPT)



selama periode 2019-2024 menggunakan metode K-Means Clustering. Pemilihan jumlah cluster ditentukan melalui analisis Elbow Method dan Silhouette Score, di mana $k = 3$ dipilih sebagai jumlah kluster yang optimal berdasarkan keseimbangan antara struktur data dan interpretabilitas hasil.

Hasil segmentasi menunjukkan tiga pola utama. Klaster 0 terdiri dari wilayah dengan kunjungan wisatawan dan TPT yang rendah, mencerminkan daerah dengan ekonomi lokal yang relatif stabil namun belum menjadi destinasi wisata utama. Klaster 1 mencakup wilayah dengan kunjungan wisatawan sangat tinggi namun diiringi tingkat pengangguran yang juga tinggi, mengindikasikan adanya ketimpangan distribusi manfaat ekonomi dari sektor pariwisata. Klaster 2 menampilkan wilayah dengan nilai sedang pada kedua indikator, menunjukkan potensi transisi ekonomi dan perkembangan sektor wisata apabila dikelola secara efektif.

Temuan ini menggarisbawahi pentingnya pendekatan berbasis data dalam perencanaan pembangunan daerah. Segmentasi semacam ini dapat memberikan wawasan strategis bagi pemerintah daerah dalam merancang kebijakan yang lebih kontekstual, baik dalam optimalisasi potensi pariwisata maupun penanganan masalah ketenagakerjaan secara lebih terfokus.

DAFTAR PUSTAKA

- [1] Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651-666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- [2] Zhang, Y., & Zhang, C. (2020). Smart Cities and Big Data Analytics: A Data-Driven Decision-Making Framework. *Journal of Urban Technology*, 27(2), 1-20. <https://doi.org/10.1080/10630732.2020.1738827>
- [3] Pratama, R. A., Sari, R. N., & Wulandari, A. (2022). Segmentasi Wilayah Menggunakan Metode K-Means Berdasarkan Indikator Sosial Ekonomi. *Jurnal Statistika dan Komputasi*, 9(2), 105-112.
- [4] Badan Pusat Statistik Provinsi Jawa Timur. (2023). *Statistik Pariwisata Provinsi Jawa Timur 2023*.
- [5] Siregar, M. F., & Nursal, D. (2020). Segmentasi Wilayah Menggunakan K-Means Clustering untuk Perencanaan Pembangunan Daerah. *Jurnal Ilmiah Informatika KOMPUTA*, 9(2), 107-116. <https://doi.org/10.37676/komputa.v9i2.347>
- [6] Kusumadewi, S. (2010). Analisis Data dengan Teknik Data Mining. *Jurnal Teknologi Informasi DINAMIK*, 15(1), 1-8. <https://doi.org/10.31294/dinamika.v15i1.190>



- [7] Wardhani, D. K., Astuti, D. M., & Huda, M. (2021). Implementasi Metode K-Means untuk Segmentasi Wilayah Berdasarkan Indikator Sosial Ekonomi. *Jurnal Sains dan Informatika*, 7(2), 55-62. <https://doi.org/10.21009/jsi.007.2.05>
- [8] A. R. Pratama, B. S. B. Dewantara, D. M. Sari, dan D. Pramadihanto, Improvement of DBSCAN Algorithm Involving Automatic Parameters Estimation and Curvature Analysis in 3D Point Cloud of Piled Pipe, *Journal of Image and Graphics*, vol. 12, no. 2, pp. 175-185, 2024. doi:10.18178/joig.12.2.175-185.
- [9] Herdiana, I., Kamal, M. A., Triyani, M. N. E., & Renny. (2025). A More Precise Elbow Method for Optimum K-means Clustering. *arXiv preprint arXiv:2502.00851*.
- [10] Nugroho, N., et al. (2022). Penggunaan Metode K-Means dan K-Means++ Sebagai Clustering Data Covid-19 di Pulau Jawa. *Jurnal Teknik Informatika*, 11(3), 502.
- [11] M. T. Jatipaningrum, S. E. Azhari, dan K. Suryowati, Pengelompokan kabupaten dan kota di Provinsi Jawa Timur berdasarkan tingkat kesejahteraan dengan metode K-Means dan Density-Based Spatial Clustering of Applications with Noise, *Jurnal Derivat*, vol. 9, no. 1, pp. 70-81, 2022.
- [12] B. M. Al-Fahmi, E. Rahmawati, dan T. Sagirani, Penerapan K-Means Clustering pada pariwisata Kabupaten Bojonegoro untuk mendukung keputusan strategi pemasaran, *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 9, no. 2, pp. 141-149, 2023. doi:10.25077/TEKNOSI.v9i2.2023.141-149.
- [13] A. Helia, Mulyawan, C. L. Rohmat, dan Fathurrohman, Analisis pengelompokan daya tarik obyek wisata berdasarkan jenisnya menggunakan metode K-Means pada data Pemprov Jabar, *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 3, pp. 2764-2772, 2024.
- [14] A. K. Y. Pratama, Metode K-Means untuk segmentasi wisatawan berdasarkan aspek sosio-ekonomi, *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 4, no. 2, pp. 190-197, 2022. doi:10.28926/ilkomnika.v4i2.503.
- [15] H. Habiballoh, A. Faqih, dan T. Supriati, Implementasi algoritma K-Means dalam mengelompokan kabupaten/kota di Jawa Barat berdasarkan jenis dan jumlah potensi objek daya tarik wisata, *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, vol. 12, no. 2, pp. 1482-1490, 2024. doi:10.23960/jitet.v12i2.4270.
- [16] S. B. Tandilino, Strategi pengembangan pariwisata melalui desa wisata di Kabupaten Timor Tengah Selatan untuk meningkatkan jumlah pergerakan wisatawan, *Jurnal Ilmiah JUPAR*, vol. 1, no. 1, pp. 1-11, 2019.

