



Pengelompokan Lokasi Tempat Wisata di Indonesia Menggunakan Algoritma K-Means

Stephen Pratama Kurnia

Universitas Pelita Bangsa

Nicky Pascal Tambunan

Universitas Pelita Bangsa

Serius Ndruru

Universitas Pelita Bangsa

Jl. Inspeksi Kalimalang Tegal Danas Arah Deltamas, Cibatu, Cikarang

Korespondensi penulis: kurnia16j@gmail.com

Abstrak. *Tourism is one of the important sectors in Indonesia's economy with great potential for further development. However, Indonesia's vast territory poses challenges in terms of effective management and promotion of tourist destinations. This study aims to cluster tourist locations in Indonesia based on geographical proximity using the K-Means algorithm. The data used consists of geographical coordinates (latitude and longitude) of various tourist destinations across Indonesia. The clustering process was carried out using five clusters with the K-Means algorithm, and the results were visualized using a distribution map. The results show that the K-Means algorithm is capable of grouping tourist locations according to Indonesia's major geographical regions, such as Bali and Nusa Tenggara, Java, Sumatra, Kalimantan, as well as Maluku and Papua. These findings are expected to serve as a useful basis to support tourism planning and development in Indonesia.*

Keywords: *Clustering, K-Means, Tourist Locations, Geographical Data, Indonesia.*

Abstrak. Pariwisata merupakan salah satu sektor penting dalam perekonomian Indonesia yang memiliki potensi besar untuk terus dikembangkan. Namun, luasnya wilayah Indonesia menjadi tantangan dalam hal pengelolaan dan promosi destinasi wisata secara efektif. Penelitian ini bertujuan untuk mengelompokkan lokasi tempat wisata di Indonesia berdasarkan kedekatan geografis menggunakan algoritma K-Means. Data yang digunakan berupa koordinat geografis (latitude dan longitude) dari berbagai lokasi wisata di Indonesia. Pengelompokan dilakukan menggunakan lima cluster dengan memanfaatkan algoritma K-Means, dan hasilnya divisualisasikan dalam bentuk peta sebaran. Hasil penelitian menunjukkan bahwa algoritma K-Means mampu mengelompokkan lokasi wisata sesuai dengan wilayah geografis utama di Indonesia, seperti Bali dan Nusa Tenggara, Jawa, Sumatera, Kalimantan, serta Maluku dan Papua. Temuan ini diharapkan dapat menjadi dasar yang berguna dalam mendukung perencanaan dan pengembangan sektor pariwisata di Indonesia.

Kata Kunci: *Pengelompokan, K-Means, Tempat Wisata, Data Geografis, Indonesia.*

PENDAHULUAN

Pariwisata merupakan salah satu sektor penting dalam perekonomian lokal¹. Keanekaragaman budaya, keindahan alam, serta kekayaan sejarah yang tersebar di seluruh wilayah Indonesia menjadikan sektor pariwisata memiliki potensi besar untuk terus dikembangkan. Namun, luasnya wilayah Indonesia yang terdiri dari ribuan pulau sering kali menjadi tantangan tersendiri dalam hal pengelolaan dan promosi destinasi wisata secara efektif.

Salah satu langkah strategis yang dapat dilakukan adalah dengan mengelompokkan lokasi-lokasi wisata berdasarkan kedekatan geografis. Dengan adanya pengelompokan ini, pemerintah maupun pihak terkait dapat lebih mudah dalam merumuskan kebijakan, mengembangkan infrastruktur, serta mempromosikan destinasi wisata secara lebih terarah dan efisien.

Teknik data mining, khususnya algoritma K-Means Clustering, dapat digunakan untuk melakukan pengelompokan lokasi wisata berdasarkan data koordinat geografis. Algoritma ini mampu mengelompokkan data ke dalam beberapa cluster yang merepresentasikan kedekatan lokasi, sehingga pola persebaran destinasi wisata di Indonesia dapat terlihat lebih jelas.

Berdasarkan hal tersebut, penelitian ini dilakukan dengan tujuan untuk mengelompokkan lokasi-lokasi tempat wisata di Indonesia menggunakan algoritma K-Means. Hasil dari penelitian ini diharapkan dapat memberikan gambaran visual dan informasi yang bermanfaat dalam mendukung pengembangan sektor pariwisata di Indonesia.

KAJIAN TEORITIS

I. Data Mining

Data mining merupakan proses untuk menemukan pola, hubungan, atau informasi yang berguna dari sekumpulan data besar secara otomatis atau semi otomatis. Menurut Han, Kamber, dan Pei (2012), data mining adalah proses penemuan pengetahuan yang tersembunyi dalam data yang besar dengan memanfaatkan teknik statistik, matematika, dan machine learning. Dalam konteks penelitian ini, data mining digunakan untuk mengelompokkan lokasi tempat wisata berdasarkan data geografis yang tersedia.

II. Clustering

Clustering atau pengelompokan adalah salah satu metode dalam data mining yang bertujuan untuk membagi data ke dalam beberapa kelompok atau cluster, sehingga data dalam satu cluster memiliki karakteristik yang mirip satu sama lain dan berbeda dengan data pada cluster lainnya. Clustering banyak digunakan dalam berbagai bidang, seperti pemasaran, segmentasi pelanggan, pengelompokan dokumen, hingga analisis spasial atau geografis.

III. Algoritma K-Means

K-Means adalah salah satu algoritma clustering yang paling populer dan banyak digunakan karena kesederhanaan serta efisiensinya. Algoritma ini berusaha membagi data ke dalam K cluster berdasarkan kedekatan data dengan pusat cluster (centroid). Langkah kerja algoritma K-Means menurut Hartigan & Wong (1979) adalah sebagai berikut:

1. Tentukan jumlah cluster (K) yang diinginkan.
2. Inisialisasi posisi centroid secara acak.
3. Hitung jarak setiap titik data ke centroid.
4. Kelompokkan titik data ke centroid terdekat.
5. Hitung ulang posisi centroid berdasarkan rata-rata posisi data dalam setiap cluster.
6. Ulangi langkah 3-5 hingga tidak ada perubahan signifikan (konvergen).

Dalam penelitian ini, K-Means digunakan untuk mengelompokkan lokasi tempat wisata di Indonesia berdasarkan atribut koordinat *latitude* dan *longitude*, sehingga diperoleh cluster-cluster yang merepresentasikan kedekatan geografis.

METODE PENELITIAN

I. Jenis Penelitian

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan eksploratif, yang bertujuan untuk mengelompokkan lokasi tempat wisata di Indonesia berdasarkan data koordinat geografis menggunakan algoritma K-Means Clustering.

II. Data dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari kaggle, dataset ini berisi informasi tempat wisata di Indonesia. Adapun atribut utama yang digunakan adalah:

- Latitude: Koordinat lintang lokasi wisata.
- Longitude: Koordinat bujur lokasi wisata.
- Informasi tambahan seperti nama wisata, provinsi, dan deskripsi digunakan untuk interpretasi hasil.

III. Alat dan Perangkat Lunak

- Python.
- Library: pandas, matplotlib, dan sklearn.\
- Visual Studio Code

IV. Tahapan Penelitian

Adapun tahapan penelitian dilakukan sebagai berikut:

1. Pengumpulan Data

Dataset lokasi tempat wisata dikumpulkan dan diperiksa kelengkapan serta kebenaran datanya, khususnya pada atribut koordinat geografis.

2. Pra-pemrosesan Data

Data dibersihkan apabila terdapat nilai kosong atau tidak valid pada atribut latitude dan longitude.

3. Penerapan Algoritma K-Means

- Menentukan jumlah cluster (**K**) yang diinginkan, dalam penelitian ini dicoba dengan 5 cluster.
- Menggunakan algoritma K-Means untuk mengelompokkan data berdasarkan atribut latitude dan longitude.
- Proses clustering dilakukan dengan fungsi KMeans dari pustaka scikit-learn.

4. Visualisasi Hasil

Hasil pengelompokan divisualisasikan menggunakan diagram scatter untuk mempermudah interpretasi pola sebaran lokasi wisata berdasarkan cluster.

5. Interpretasi Hasil

Cluster yang terbentuk dianalisis berdasarkan wilayah geografis dan distribusi tempat wisata di masing-masing cluster.

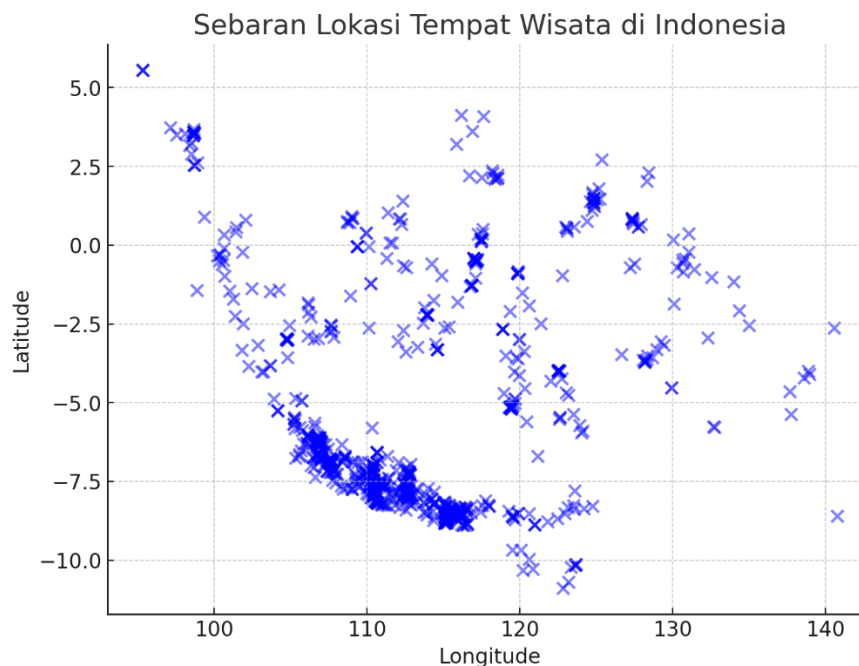
V. Parameter Pengujian

Parameter yang digunakan dalam penerapan algoritma K-Means adalah:

- Jumlah Cluster (**K**) = 5
- random_state = 42 (untuk menjaga konsistensi hasil)
- n_init = 10 (jumlah inisialisasi ulang centroid)

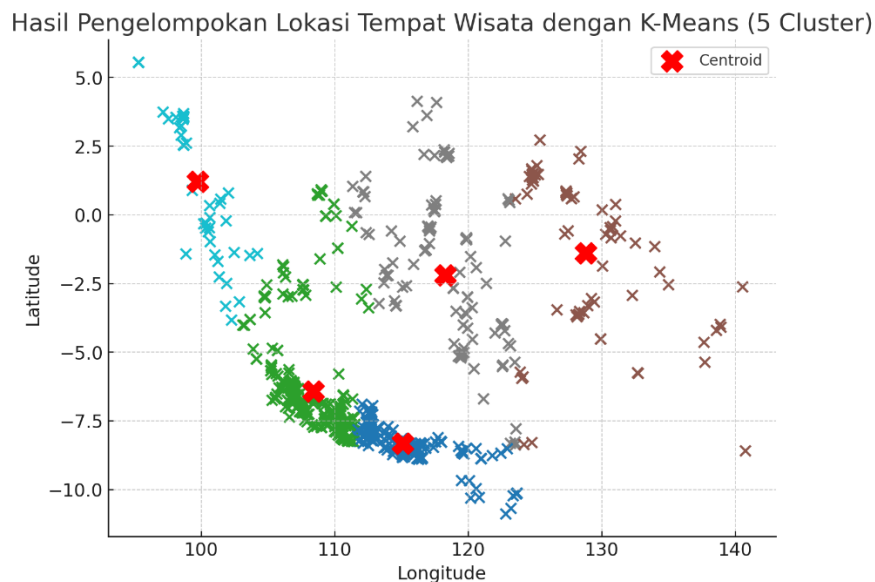
HASIL PENELITIAN DAN PEMBAHASAN

A. Visualisasi Sebaran Lokasi Tempat Wisata.



Berdasarkan hasil visualisasi awal, diketahui bahwa lokasi tempat wisata di Indonesia tersebar di berbagai pulau besar seperti Sumatera, Jawa, Bali, Kalimantan, Sulawesi, Nusa Tenggara, Maluku, dan Papua. Penyebaran yang luas ini menjadi dasar perlunya dilakukan pengelompokan untuk mempermudah analisis dan pengelolaan.

B. Hasil Pengelompokan dengan Algoritma K-Means.



Pengelompokan dilakukan menggunakan algoritma K-Means dengan jumlah cluster sebanyak 5, yang dipilih berdasarkan pertimbangan geografis wilayah Indonesia. Berikut adalah hasil visualisasi pengelompokan:

- Setiap titik mewakili lokasi tempat wisata, dengan warna yang berbeda menunjukkan cluster yang berbeda.
- Titik centroid ditandai dengan simbol 'X' berwarna merah, yang menunjukkan pusat dari masing-masing cluster.

Hasil pengelompokan menunjukkan bahwa algoritma K-Means berhasil memetakan lokasi wisata ke dalam beberapa kelompok besar berdasarkan kedekatan geografis.

C. Interpretasi Clustering

Berdasarkan analisis hasil clustering, diperoleh interpretasi sebagai berikut:

CLUSTER	WILAYAH DOMINAN	KETERANGAN
Cluster 0 (Biru Tua)	Bali, Jawa Timur (sebagian), NTB, dan NTT	Mewakili wilayah Bali dan Kepulauan Nusa Tenggara.
Cluster 1 (Hijau)	Jawa Barat, Jawa Tengah, Yogyakarta, Sumatera bagian selatan, Kalimantan Barat	Wilayah Jawa bagian barat dan tengah, serta sebagian Sumatera dan Kalimantan.
Cluster 2 (Coklat)	Maluku, Papua, Papua Barat	Mewakili wilayah Indonesia bagian timur.
Cluster 3 (Abu-abu)	Kalimantan bagian tengah dan timur	Wilayah Kalimantan yang terpisah dari cluster lain.
Cluster 4 (Biru Muda)	Sumatera bagian utara	Sumatera bagian utara yang memiliki konsentrasi wisata tersendiri.

D. Pembahasan

Hasil clustering ini menunjukkan bahwa persebaran tempat wisata di Indonesia secara umum mengikuti pola geografis kepulauan Indonesia. Algoritma K-Means berhasil memisahkan wilayah-wilayah yang memiliki konsentrasi tempat wisata yang berdekatan ke dalam satu kelompok.

Manfaat dari hasil pengelompokan ini antara lain:

- **Perencanaan Pariwisata Terarah**
Setiap cluster dapat dijadikan acuan dalam pengembangan dan promosi pariwisata yang lebih spesifik sesuai karakteristik wilayahnya.
- **Pemetaan Potensi Pariwisata**
Dengan memahami konsentrasi lokasi wisata, pemerintah daerah maupun pihak swasta dapat lebih fokus dalam meningkatkan fasilitas dan aksesibilitas pada cluster tertentu.
- **Efisiensi Promosi**
Promosi pariwisata dapat dilakukan berdasarkan cluster untuk meningkatkan efektivitas dan efisiensi, misalnya promosi terintegrasi untuk Bali dan Nusa Tenggara yang masuk dalam satu cluster.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa pengelompokan lokasi tempat wisata di Indonesia menggunakan algoritma K-Means mampu membagi data ke dalam lima cluster yang secara geografis mewakili wilayah-wilayah besar di Indonesia, seperti Bali dan Nusa Tenggara, Jawa bagian barat-tengah, Kalimantan, Sumatera,

serta wilayah timur Indonesia seperti Maluku dan Papua. Hasil pengelompokan ini menunjukkan bahwa algoritma K-Means efektif dalam memetakan persebaran lokasi wisata berdasarkan kedekatan geografis, sehingga dapat menjadi dasar yang berguna dalam perencanaan, pengembangan, dan promosi pariwisata secara terfokus dan efisien.

DAFTAR PUSTAKA

- Wasesha, D. A., & Syafrianto. (2025). *Penerapan algoritma K-Means untuk mengelompokkan kunjungan wisatawan pada dua puluh tempat wisata di Jakarta*. *Technologia : Jurnal Ilmiah*, 16(1), 33 <https://doi.org/10.31602/tji.v16i1.16959>.
- Al-Fahmi, B. M., Rahmawati, E., Sagirani, T. (2023). *Penerapan K-Means Clustering pada pariwisata Kabupaten Bojonegoro untuk mendukung keputusan strategi pemasaran*. *Jurnal Nasional Teknologi dan Sistem Informasi*, 9(2), 141–149. <https://doi.org/10.25077/TEKNOSI.v9i2.2023.141-149>.
- Arul, S. I. U., & Sanjaya, I. (2024). *Penerapan algoritma K-Means clustering pada sentimen pengunjung Desa Wisata Hanjeli*. *Jurnal CoSciTech (Computer Science and Information Technology)*, 5(1), 126–133. <https://doi.org/10.37859/coscitech.v5i1.6955>.
- Simorangkir, S., Sany, E., Feraldi, M. (2025). *Penerapan metode K-Means untuk pengelompokan data kunjungan wisata pada Dinas Kebudayaan dan Pariwisata Provinsi Jambi*. *Akademika: Jurnal Ilmu*, 17(2) <https://doi.org/10.53564/akademika.v17i2.1496>.
- Rahmiyati, D., Nuraeni, & Siswantoro, E. B. (2025). *Penerapan metode K-Means untuk pemetaan objek wisata sebagai rekomendasi prioritas pengembangan pariwisata*. *Jurnal Ilmu Komputer dan Teknologi*, 6(1), 17–24. <https://doi.org/10.35960/ikomti.v6i1.1748>.
- Dalipang, M. P., Wibisono, L. K., & Marchelin (2025). *Tourist cluster analysis for the development of Ke'te' Kesu' tourism object*. *Ekbis Journal*, 26(1), 387. <https://jurnalekonomi.unisla.ac.id/index.php/ekbis/article/view/2642>.
- Muqorobin, Supatminingsih, S. & Wahyuningsih, D, H,. (2023). *Metode klasterisasi untuk pengelompokan potensi wisata guna mendukung digitalisasi pariwisata di Indonesia*. *Jurnal Sains dan Teknologi*, 2023. <https://doi.org/10.23887/jstundiksha.v13i2.84743>.
- Liyanto, E. W., Homaidi, A., Lutfi, A . (2024). *Implementasi K-Means clustering menggunakan RapidMiner dalam pengelompokan data kunjungan wisatawan asing di Provinsi Jawa Timur*. *Jurnal Teknik Elektro dan Informatika*, 19(2), 205–216. <https://doi.org/10.30587/e-link.v19i2.8224>.
- Puspita, D., & Sasmita. (2020). *Application of K-Means algorithm in grouping of city tourism Pagar Alam*. *Sinkron Jurnal & Penelitian Teknik Informatika*, <https://doi.org/10.33395/sinkron.v7i1.11220>.
- Purnomo, B. S., & Prasetyaningrum, P. T. (2022). *Penerapan data mining dalam mengelompokkan kunjungan wisatawan di Kota Yogyakarta menggunakan metode K-Means*. *Jurnal Computer Science & Tech*.

<https://doi.org/10.54840/jcstech.v1i1.38>.

- Wamulkan, U. A. H., Utami, N. W., Anggara, I. N. Y. (2021). *Bali tourist visits clustered via TripAdvisor reviews using K-Means*. Pilar Nusa Mandiri. https://www.researchgate.net/publication/377548691_BALI_TOURIST_VISITS_CLUSTERED_VIA_TRIPADVISOR_REVIEWS_USING_K-MEANS_ALGORITHM.
- Savitri, N., et al. (2021). *Pengelompokan kunjungan wisata Kabupaten Kulon Progo tahun 2019 menggunakan K-Means clustering*. Just IT. <https://doi.org/10.24853/justit.12.1.%25p>.
- Qori'atunnadiyah, M. (2024). *Mapping domestic and foreign tourists in East Java using C-Means clustering*. Jurnal Statistika dan Aplikasinya. <https://doi.org/10.21009/JSA.08105>.
- Usman, A., et al. (2022). *Analisis sentimen objek wisata Geopark Ciletuh dengan algoritma K-Means clustering*. Prosiding SENDIKO: Seminar Nasional Bidang Ilmu Komputer, hlm. 1–8. <https://prosiding.unipma.ac.id/index.php/sendiko/article/view/3790>.
- Astiti, S., Harman, R., & Darmansah, D. (2024). *Pengelompokan destinasi wisata di Batam berdasarkan daya tarik dan fasilitas menggunakan metode K-Means clustering*. Kesatria: Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen), 5(4), 2005–2012. <https://tunasbangsa.ac.id/pkm/index.php/kesatria/article/view/522>.
- Helia, A., Mulyawan., Rohmat, C. L., & Fathurrohman. (2024). *Analisis pengelompokan daya tarik objek wisata berdasarkan jenisnya menggunakan metode K-Means pada data Pemprov Jabar*. JATI: Jurnal Mahasiswa Teknik Informatika, 8(3). <https://doi.org/10.36040/jati.v8i3.8369>.
- Wulandari, S., Astuti, R., & Basysyar, F. M. (2024). *Implementasi teknik data mining pada data kunjungan wisatawan menggunakan K-Means clustering di hotel bintang Kota Cirebon*. JATI: Jurnal Mahasiswa Teknik Informatika, 8(1). <https://doi.org/10.36040/jati.v8i1.8470>.
- Purnomo, S. B., & Prasetyaningrum, P. T. (2022). *Penerapan data mining dalam mengelompokkan kunjungan wisatawan di Kota Yogyakarta menggunakan metode K-Means*. Journal of Computer Science and Technology (JCS-TECH), 1(1), 27–32. <https://doi.org/10.54840/jcstech.v1i1.38>
- Satria, E., Tambunan, H. S., Saragih, I. S., Damanik, I. S., & Sitanggang, F. T. E. (2019). *Penerapan Clustering dalam Mengelompokkan Jumlah Kunjungan Wisatawan Mancanegara Dengan Metode K-Means*. Seminar Nasional Riset Information Science. <https://tunasbangsa.ac.id/seminar/index.php/senaris/article/view/52>.
- Beatrice, C., Ingrid, A. I., Sinaga, K. T., & Sinaga, A. S. (2024). *Penggunaan metode K-Means clustering pemetaan dan klasterisasi tempat wisata di Kabupaten Deli Serdang*. Madani: Jurnal Ilmiah Multidisiplin, 2(1). <https://doi.org/10.5281/zenodo.15276418>.