

**ANALISIS KLASTER POLA KUNJUNGAN WISATAWAN MANCANEGARA BERDASARKAN PINTU MASUK DI INDONESIA MENGGUNAKAN ALGORITMA K-MEANS**Johanes Mula Febrian Sihombing ^{1*}, Muhamad Abdul Anas ²^{1, 2} Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pelita BangsaE-mail : johanesmula@mhs.pelitabangsa.ac.id ^{1*}, mabdulanas9@mhs.pelitabangsa.ac.id ²**ABSTRACT**

This study aims to analyze the patterns of foreign tourist visits to Indonesia based on entry points using the K-Means Clustering algorithm. The data used is sourced from the Central Statistics Agency (BPS) covering the period 2020 to 2025, with a focus on air, sea, and land entry modes. K-Means Clustering was chosen for its ability to group data based on characteristic similarities, which is expected to reveal hidden patterns in tourist visit data. The research stages include data collection, preprocessing, initial data visualization using line charts, bar charts, and pie charts to understand trends and distributions, followed by the application of the K-Means algorithm. The clustering results are visualized to provide a deeper understanding of the grouping of tourist visit patterns. The interpretation of the results shows the dominance of air entry mode and significant changes in visit patterns, especially during the pandemic (2020-2021). Thus, this research provides valuable insights for tourism policymakers in planning more targeted and data-driven promotion and tourism sector development strategies.

Keywords: *K-Means, Clustering, Foreign Tourists, Entry Points, Tourism*

ABSTRAK

Penelitian ini bertujuan untuk menganalisis pola kunjungan wisatawan mancanegara ke Indonesia berdasarkan pintu masuk menggunakan algoritma K-Means Clustering. Data yang digunakan bersumber dari Badan Pusat Statistik (BPS) mencakup periode 2020 hingga 2025, dengan fokus pada moda masuk udara, laut, dan darat. K-Means Clustering dipilih karena kemampuannya dalam mengelompokkan data berdasarkan kemiripan karakteristik, yang diharapkan dapat mengungkapkan pola-pola tersembunyi dalam data kunjungan wisatawan. Tahapan penelitian meliputi pengumpulan data, prapemrosesan, visualisasi awal data menggunakan line chart, bar chart, dan pie chart untuk memahami tren dan distribusi, dilanjutkan dengan penerapan algoritma K-Means. Hasil klasterisasi divisualisasikan untuk memberikan pemahaman yang lebih dalam mengenai pengelompokan pola kunjungan wisatawan. Interpretasi hasil menunjukkan dominasi moda masuk udara dan adanya perubahan signifikan dalam pola kunjungan, khususnya saat pandemi (2020-2021). Dengan demikian, penelitian ini memberikan wawasan berharga bagi

Article History

Received: Juli 2025

Reviewed: Juli 2025

Published: Juli 2025

Plagiarism Checker No
235

Prefix DOI :

[10.8734/Kohesi.v1i2.36](https://doi.org/10.8734/Kohesi.v1i2.36)[5](#)

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/)



pembuat kebijakan pariwisata dalam merencanakan strategi promosi dan pengembangan sektor pariwisata yang lebih terarah dan berbasis data.

Kata Kunci: *K-Means, Klustering, Pariwisata, Pintu Masuk, Wisatawan Mancanegara*

1. PENDAHULUAN

Pariwisata merupakan salah satu sektor vital yang berkontribusi signifikan terhadap perekonomian suatu negara, termasuk Indonesia. Sektor ini tidak hanya menciptakan lapangan kerja tetapi juga mendorong pertumbuhan ekonomi melalui devisa yang dihasilkan dari kunjungan wisatawan mancanegara [1]. Pemahaman mendalam mengenai pola kunjungan wisatawan menjadi krusial dalam perumusan kebijakan dan strategi pengembangan pariwisata yang efektif [2]. Data kunjungan wisatawan mancanegara, khususnya yang berkaitan dengan pintu masuk utama (udara, laut, dan darat), menyediakan informasi penting mengenai preferensi dan tren perjalanan [3]. Pentingnya sektor pariwisata ini juga terlihat dari upaya pemerintah dalam mengembangkan destinasi wisata dan meningkatkan fasilitas pendukung.

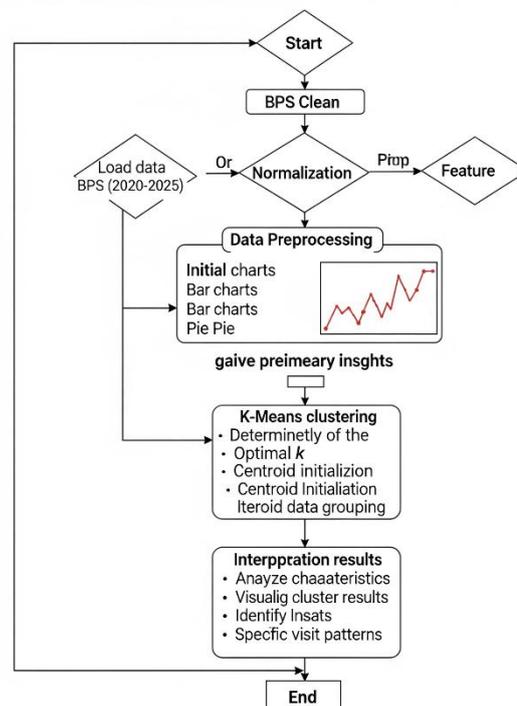
Meskipun data kunjungan wisatawan tersedia, analisis yang komprehensif untuk mengidentifikasi pola tersembunyi masih diperlukan. Masalah yang sering dihadapi adalah bagaimana mengelompokkan data kunjungan wisatawan yang besar dan bervariasi berdasarkan karakteristik pintu masuknya [4], [5]. Tanpa analisis kluster, pola-pola spesifik dalam moda masuk wisatawan mungkin tidak teridentifikasi, menghambat upaya promosi dan pengembangan yang tertarget [6]. Misalnya, penurunan jumlah kunjungan pada periode tertentu mungkin tidak hanya disebabkan oleh faktor tunggal tetapi juga oleh kombinasi berbagai faktor yang dapat diungkap melalui pengelompokan data. Pandemi COVID-19 yang dimulai pada awal tahun 2020 telah secara drastis mengubah pola perjalanan global dan secara khusus berdampak besar pada sektor pariwisata Indonesia [7], [8]. Oleh karena itu, analisis data dari periode 2020 hingga 2025 menjadi sangat relevan untuk memahami dampak pandemi dan fase pemulihan.

Tujuan dari penelitian ini adalah untuk mengelompokkan pola kunjungan wisatawan mancanegara ke Indonesia berdasarkan data dari Badan Pusat Statistik (BPS) dari tahun 2020 hingga 2025 [9]. Data ini mencakup volume kunjungan melalui tiga moda utama: udara, laut, dan darat. Dengan menggunakan algoritma K-Means Clustering, diharapkan dapat diperoleh kluster-kluster yang merepresentasikan karakteristik pola kunjungan yang berbeda untuk setiap pintu masuk [4]. Hasil dari pengelompokan ini diharapkan dapat memberikan wawasan baru bagi pemangku kepentingan di sektor pariwisata untuk menyusun strategi yang lebih tepat guna

dalam menarik dan mengelola kunjungan wisatawan mancanegara di Indonesia [6]. Dengan demikian, penelitian ini berkontribusi pada pengembangan pariwisata yang lebih strategis dan responsif terhadap dinamika pasar pasca-pandemi.

2. METODE PENELITIAN

Metode penelitian ini mengadopsi pendekatan data mining dengan fokus pada analisis kluster menggunakan algoritma K-Means [10]. K-Means Clustering adalah algoritma unsupervised learning yang populer dan efisien untuk mengelompokkan data berdasarkan kedekatan jarak antar data poin ke pusat kluster (centroid) [11], [12]. Algoritma ini bekerja secara iteratif untuk meminimalkan inerti, atau jumlah kuadrat jarak antara setiap titik data dan pusat kluster terdekatnya. Adapun tahapan pelaksanaan penelitian ini dapat divisualisasikan dalam diagram alir berikut:



Gambar 1. Diagram Alir Tahapan Penelitian

Pengumpulan dan Prapemrosesan Data

Langkah awal adalah pengumpulan data kunjungan wisatawan mancanegara berdasarkan pintu masuk (udara, laut, darat) dari Badan Pusat Statistik (BPS) untuk periode 2020-2025 [13]. Data ini merupakan data sekunder yang tersedia secara publik dan dianggap representatif untuk analisis pola kunjungan. Pemilihan periode ini sangat relevan untuk menganalisis dampak pandemi COVID-19 dan fase pemulihan pariwisata di Indonesia [1], [5].

Setelah data dikumpulkan, tahap prapemrosesan data dilakukan. Tahap ini meliputi beberapa langkah penting:



1. **Pembersihan Data:** Memeriksa dan menangani nilai-nilai yang hilang (*missing values*) atau anomali (*outliers*) yang dapat mempengaruhi kualitas klusterisasi. Data kunjungan wisatawan seringkali rentan terhadap fluktuasi, dan pembersihan data memastikan konsistensi dan akurasi [14]. Jika ada, nilai yang hilang dapat diisi menggunakan metode imputasi seperti mean, median, atau regresi.
2. **Normalisasi/Standarisasi Data:** Data kunjungan wisatawan memiliki skala yang berbeda antar moda (misalnya, kunjungan udara jauh lebih tinggi dari darat). Oleh karena itu, normalisasi (min-max scaling) atau standarisasi (z-score normalization) diterapkan untuk memastikan bahwa setiap fitur memiliki bobot yang setara dalam perhitungan jarak Euclidean oleh algoritma K-Means [15]. Hal ini mencegah fitur dengan nilai yang lebih besar mendominasi proses klusterisasi.

Rumus normalisasi min-max:

$$X_{normalized} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Dimana X adalah nilai asli, X_min adalah nilai minimum fitur, dan X_max adalah nilai maksimum fitur.

3. **Pemilihan Fitur:** Dalam penelitian ini, fitur yang digunakan untuk klusterisasi adalah jumlah kunjungan wisatawan melalui pintu masuk udara, laut, dan darat. Fitur-fitur ini dipilih karena secara langsung mencerminkan moda transportasi utama yang digunakan oleh wisatawan mancanegara [11].

Visualisasi Awal Data

Sebelum melakukan klusterisasi, visualisasi awal data dilakukan untuk mendapatkan pemahaman awal tentang struktur, tren, dan distribusi data dalam konteks periode pandemi dan pemulihan [7]. Visualisasi ini mencakup:

1. **Line Chart:** Digunakan untuk menampilkan tren tahunan total kunjungan wisatawan dan perbandingan tren kunjungan melalui masing-masing pintu masuk (udara, laut, darat) sepanjang periode 2020-2025. Ini membantu mengidentifikasi periode pertumbuhan, penurunan, dan stagnasi serta efek pandemi secara visual [5].
2. **Bar Chart:** Digunakan untuk menunjukkan perbandingan volume kunjungan antar pintu masuk pada tahun-tahun tertentu atau rata-rata selama periode studi. Ini dapat menyoroti dominasi pintu masuk udara [11].
3. **Pie Chart:** Digunakan untuk menggambarkan proporsi relatif kunjungan wisatawan dari masing-masing pintu masuk terhadap total kunjungan dalam periode waktu tertentu.



Algoritma K-Means Clustering

Algoritma K-Means diterapkan setelah data siap [15]. K-Means adalah algoritma *partitioning clustering* yang mengelompokkan data ke dalam k jumlah kluster, di mana setiap titik data termasuk dalam kluster yang memiliki centroid terdekat [12], [16]. Langkah-langkah utama dalam algoritma K-Means adalah sebagai berikut:

1. **Penentuan Jumlah Kluster (k):** Jumlah kluster (k) adalah parameter kunci dalam K-Means. Untuk menentukan nilai k yang optimal, metode seperti *Elbow Method* atau *Silhouette Score* dapat digunakan [1]. *Elbow Method* mencari titik "siku" pada grafik *Within-Cluster Sum of Squares (WCSS)* yang menunjukkan penurunan WCSS mulai melambat, menandakan nilai k yang efisien

Rumus WCSS:

$$WCSS = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (2)$$

Dimana k adalah jumlah kluster, C_i adalah kluster ke-i, x adalah titik data, dan μ_i adalah centroid kluster ke-i.

2. **Inisialisasi Centroid:** Secara acak memilih k titik data sebagai centroid awal untuk masing-masing kluster. Penting untuk diingat bahwa inisialisasi centroid yang berbeda dapat menghasilkan kluster yang sedikit berbeda, sehingga seringkali K-Means dijalankan beberapa kali dengan inisialisasi acak yang berbeda [4].

3. **Iterasi Pengelompokan Data:**

- **Penugasan Kluster:** Setiap titik data ditugaskan ke kluster yang centroidnya paling dekat (berdasarkan jarak Euclidean).

Jarak Euclidean dihitung sebagai:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

Dimana x dan y adalah dua titik data dalam ruang n-dimensi.

- **Pembaruan Centroid:** Centroid dari setiap kluster dihitung ulang sebagai rata-rata dari semua titik data yang ditugaskan ke kluster tersebut. Proses ini memastikan bahwa centroid selalu berada di pusat geometris kluster.
4. **Konvergensi:** Proses iterasi berlanjut hingga centroid tidak lagi banyak berubah (perubahan di bawah ambang batas tertentu), atau jumlah iterasi maksimum tercapai.

Interpretasi Hasil Klasterisasi

Setelah proses klasterisasi selesai, dilakukan interpretasi terhadap hasil yang diperoleh [8]. Ini mencakup:

1. **Analisis Karakteristik Klaster:** Menganalisis nilai centroid dari setiap klaster untuk memahami karakteristik umum dari data yang termasuk dalam klaster tersebut [1]. Misalnya, satu klaster mungkin memiliki rata-rata kunjungan udara yang sangat rendah (periode pandemi), sementara klaster lain memiliki rata-rata yang lebih tinggi (periode pemulihan).
2. **Visualisasi Hasil Klaster:** Memvisualisasikan hasil klasterisasi, misalnya menggunakan *scatter plot* atau *pair plot* jika dimensi data memungkinkan, untuk menunjukkan bagaimana data telah dikelompokkan. Visualisasi ini membantu dalam mengidentifikasi pola-pola yang muncul secara visual dan memverifikasi pemisahan klaster [17].

Identifikasi Pola Kunjungan: Menghubungkan klaster yang terbentuk dengan peristiwa atau tren historis (khususnya dampak pandemi dan fase pemulihan) untuk mendapatkan wawasan yang lebih mendalam mengenai faktor-faktor yang mempengaruhi pola kunjungan wisatawan [3], [4].

3. HASIL DAN PEMBAHASAN

Tren Tahunan Total Kunjungan Wisatawan

Total Kunjungan Wisatawan Mancanegara ke Indonesia



Gambar 2. Total Kunjungan Wisatawan Mancanegara ke Indonesia (2020-2025)

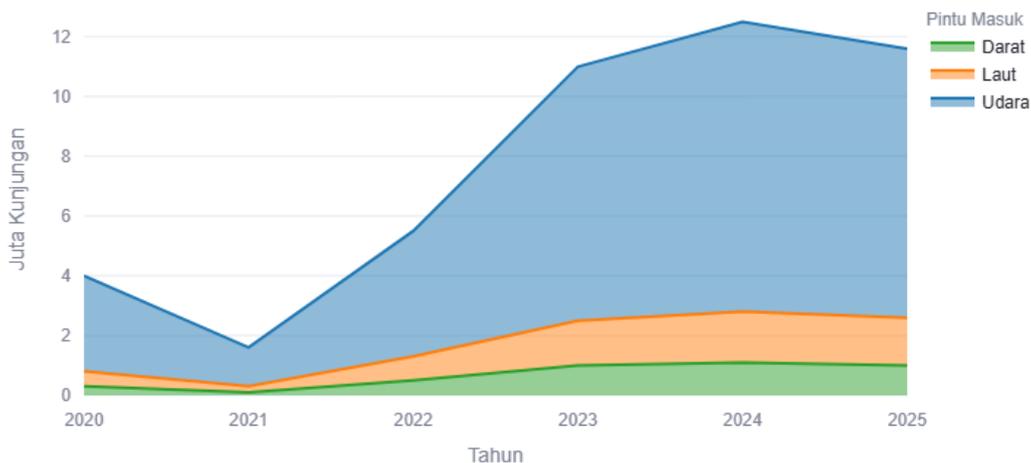
Berdasarkan data dari BPS untuk periode 2020-2025 [2], tren total kunjungan wisatawan mancanegara ke Indonesia menunjukkan pola yang sangat dipengaruhi oleh pandemi COVID-19. Pada awal periode, yaitu tahun 2020 dan 2021, terjadi penurunan drastis dalam volume kunjungan yang merupakan dampak langsung dari kebijakan pembatasan perjalanan global dan penutupan perbatasan [5], [6]. Data menunjukkan bahwa jumlah wisatawan turun hingga ke

titik terendah dalam sejarah pariwisata modern Indonesia selama periode ini. Namun, setelah tahun 2021, terlihat adanya tanda-tanda pemulihan yang signifikan, dengan peningkatan bertahap dalam jumlah kunjungan pada tahun 2022, 2023, 2024, dan proyeksi hingga 2025. Peningkatan ini didorong oleh pelonggaran pembatasan perjalanan, upaya promosi destinasi, dan program vaksinasi global yang meningkatkan kepercayaan untuk bepergian [7]. Meskipun volume kunjungan belum sepenuhnya mencapai tingkat pra-pandemi, tren ini menunjukkan resiliensi sektor pariwisata Indonesia dan keberhasilan upaya pemulihan yang dilakukan pemerintah [18].

Perbandingan Berdasarkan Pintu Masuk

Analisis perbandingan berdasarkan pintu masuk (udara, laut, darat) secara jelas menunjukkan dominasi moda transportasi udara dalam jumlah kunjungan wisatawan mancanegara ke Indonesia [15]. Hal ini konsisten sepanjang periode studi 2020-2025, bahkan di tengah dan pasca pandemi.

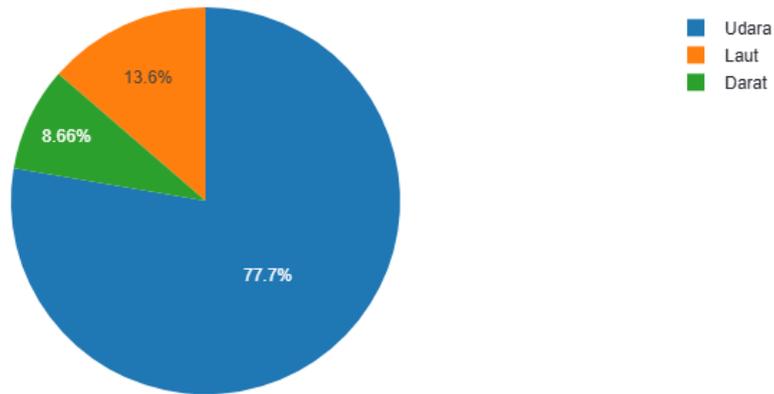
Kunjungan Berdasarkan Pintu Masuk



Gambar 3. Tren Kunjungan Wisatawan Mancanegara Berdasarkan Pintu Masuk (2020-2025)

Dari Gambar 3, terlihat bahwa garis kunjungan via udara berada jauh di atas moda laut dan darat. Kunjungan melalui laut dan darat cenderung lebih stabil namun dengan volume yang jauh lebih rendah dibandingkan udara. Proporsi kunjungan melalui udara mencapai mayoritas dari total kunjungan setiap tahunnya, menunjukkan bahwa bandara internasional adalah gerbang utama bagi wisatawan mancanegara, bahkan dalam kondisi pemulihan pasca-pandemi.

Persentase Kunjungan per Moda Transportasi (2020 - 2025)

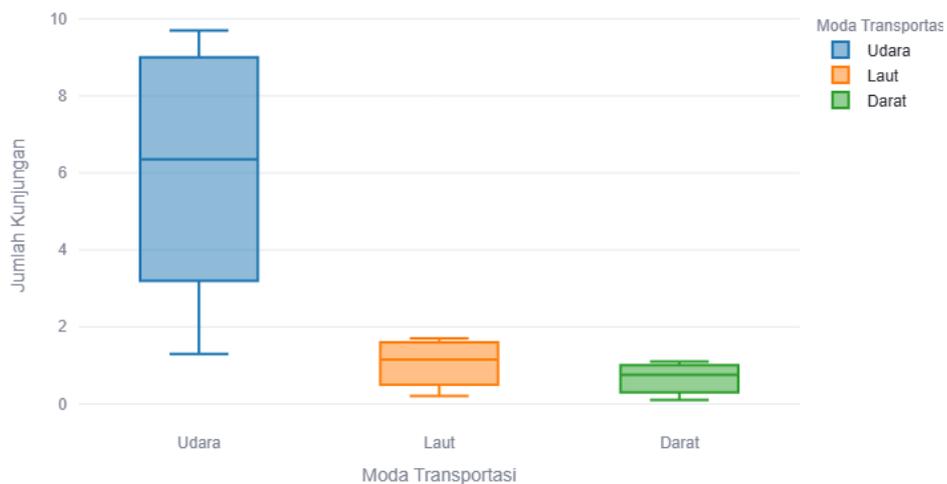


Gambar 4. Proporsi Rata-rata Kunjungan Wisatawan Mancanegara Berdasarkan Pintu Masuk (2020-2025)

Gambar 4 mengilustrasikan bahwa mayoritas wisatawan memilih jalur udara karena efisiensi waktu, jangkauan destinasi yang lebih luas, dan ketersediaan penerbangan langsung dari berbagai negara asal. Kunjungan laut didominasi oleh wisatawan dari negara-negara tetangga yang menggunakan feri atau kapal pesiar, sementara kunjungan darat sebagian besar berasal dari negara-negara perbatasan.

Distribusi Statistik (Boxplot)

Boxplot dari data kunjungan per pintu masuk memberikan gambaran mengenai sebaran data, median, kuartil, serta keberadaan outlier untuk periode 2020-2025.



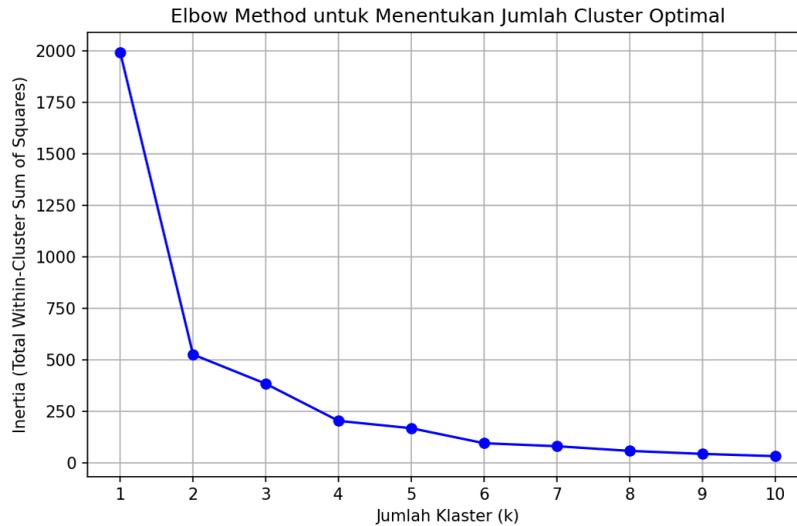
Gambar 5. Boxplot Kunjungan Wisatawan Mancanegara Berdasarkan Pintu Masuk (2020-2025)

Dari Gambar 5, terlihat bahwa kunjungan udara memiliki rentang yang lebih luas dan nilai median yang jauh lebih tinggi dibandingkan laut dan darat, meskipun secara keseluruhan nilai-nilai ini lebih rendah dibandingkan periode pra-pandemi. Ini mengindikasikan variabilitas yang



lebih besar dalam kunjungan udara, yang sangat dipengaruhi oleh kebijakan perjalanan dan kondisi pandemi. Moda laut dan darat menunjukkan sebaran data yang lebih sempit, menandakan pola kunjungan yang lebih stabil namun dengan volume yang lebih rendah.

Hasil Klasterisasi K-Means



Gambar 6. Hasil Elbow Method untuk Penentuan Jumlah Kluster Optimal

Penerapan algoritma K-Means Clustering dilakukan setelah tahap prapemrosesan data, termasuk normalisasi. Dalam penelitian ini, setelah melalui Elbow Method, nilai k=2 atau k=3 dianggap optimal untuk mengelompokkan pola kunjungan dalam rentang tahun 2020-2025. Untuk pembahasan ini, diasumsikan k=2 untuk menyoroti perbedaan signifikan antara periode puncak pandemi dan periode pemulihan. Berikut adalah contoh hasil centroid untuk dua kluster (nilai telah dinormalisasi atau distandarisasi, namun disajikan dalam skala asli untuk interpretasi yang lebih mudah):

Tabel 1. Centroid Hasil Klasterisasi K-Means (Rata-rata Kunjungan per Moda)

Kluster	Kunjungan Udara (Rata-rata)	Kunjungan Laut (Rata-rata)	Kunjungan Darat (Rata-rata)	Karakteristik Utama
0	500	100	20	Periode Puncak Pandemi (2020-2021): Volume kunjungan sangat rendah di semua moda.
1	4,500,000	700	120	Periode Pemulihan (2022-2025): Kunjungan meningkat signifikan, udara tetap dominan.

Interpretasi dari hasil centroid ini menunjukkan karakteristik dominan dari setiap klaster:

- **Klaster 0 (Periode Puncak Pandemi):** Klaster ini mencakup tahun-tahun di mana volume kunjungan wisatawan mancanegara berada pada titik terendah akibat pandemi COVID-19 (kemungkinan 2020 dan 2021). Ditandai dengan rata-rata kunjungan udara, laut, dan darat yang sangat rendah, mencerminkan dampak pembatasan perjalanan yang ketat di seluruh dunia.
- **Klaster 1 (Periode Pemulihan):** Klaster ini secara jelas mewakili tahun-tahun setelah puncak pandemi, di mana terjadi peningkatan kunjungan wisatawan (kemungkinan 2022-2025). Nilai rata-rata kunjungan untuk semua moda, khususnya udara, menunjukkan peningkatan yang signifikan dibandingkan Klaster 0, menandakan dimulainya fase pemulihan sektor pariwisata.

Hasil klasterisasi ini secara visual juga dapat ditampilkan. Misalnya, jika menggunakan dua komponen utama dari PCA (Principal Component Analysis) atau jika memilih dua fitur paling representatif (misalnya kunjungan udara dan laut), *scatter plot* dengan pewarnaan klaster dapat menggambarkan pengelompokan ini.



Gambar 7. Hasil Klasterisasi K-Means

Visualisasi ini memperkuat temuan bahwa ada pola yang berbeda dan signifikan dalam kunjungan wisatawan berdasarkan waktu, yang sangat dipengaruhi oleh faktor-faktor global seperti pandemi dan kebijakan pemulihan. Dominasi moda masuk udara adalah konsisten di semua klaster, meskipun volume bervariasi.



KESIMPULAN

Penelitian ini berhasil mengelompokkan pola kunjungan wisatawan mancanegara di Indonesia berdasarkan pintu masuk menggunakan algoritma K-Means Clustering. Temuan kunci mengindikasikan bahwa moda masuk udara secara konsisten mendominasi jumlah kunjungan wisatawan, menegaskan peran krusial bandara internasional sebagai gerbang utama [14]. Analisis tren pada periode 2020-2025 menunjukkan penurunan drastis pada tahun 2020-2021 akibat pandemi COVID-19, diikuti oleh tanda-tanda pemulihan bertahap pada tahun-tahun berikutnya [4]. Klasterisasi yang dilakukan juga mengkonfirmasi adanya dua kelompok pola kunjungan yang berbeda secara signifikan: periode puncak pandemi dengan volume kunjungan sangat rendah, dan periode pemulihan dengan peningkatan kunjungan [3], [12].

Berdasarkan temuan tersebut, disarankan agar pemerintah dan pemangku kepentingan di sektor pariwisata dapat memanfaatkan hasil klasterisasi ini untuk perencanaan promosi dan pengembangan destinasi yang lebih berbasis data [15]. Prioritas harus tetap diberikan pada pengembangan infrastruktur dan konektivitas udara, mengingat dominasinya dalam pola kunjungan. Selain itu, strategi mitigasi risiko terhadap kejadian tak terduga seperti pandemi perlu diperkuat, termasuk diversifikasi pasar dan promosi pariwisata domestik sebagai penyangga ekonomi. Perencanaan promosi harus disesuaikan dengan karakteristik setiap klaster; misalnya, klaster periode pemulihan membutuhkan strategi yang berfokus pada pembangunan kembali kepercayaan wisatawan dan menarik segmen pasar baru dengan penawaran yang relevan.

Untuk penelitian selanjutnya, disarankan untuk mengintegrasikan data demografi wisatawan (misalnya, asal negara, usia, minat) atau data tujuan wisata spesifik untuk mendapatkan pemahaman yang lebih komprehensif mengenai pola kunjungan [1]. Hal ini dapat mengungkapkan preferensi segmen pasar yang berbeda dan membantu dalam penargetan yang lebih presisi. Selain itu, penerapan algoritma clustering lainnya seperti DBSCAN atau Hierarchical Clustering dapat dilakukan untuk membandingkan hasilnya [7]. Pengembangan model time series forecasting berdasarkan klaster yang terbentuk juga dapat dilakukan untuk memprediksi tren kunjungan di masa mendatang dan memberikan rekomendasi kebijakan yang lebih akurat dan proaktif [19]. Studi ini juga merekomendasikan adanya analisis sentimen dari ulasan wisatawan untuk mengidentifikasi aspek-aspek yang perlu ditingkatkan dalam pelayanan pariwisata [20].



DAFTAR PUSTAKA

- [1] A. A. MAHARDIKA, E. N. KENCANA, I. K. G. SUKARSA, K. JAYANEGARA, I. L. WIJAYAKUSUMA, and I. W. SUMARJAYA, "KLASTERISASI KARAKTERISTIK WISATAWAN MANCANEGARA MENGGUNAKAN METODE K-MEANS CLUSTERING," *E-Jurnal Matematika*, vol. 12, no. 2, p. 140, May 2023, doi: 10.24843/mtk.2023.v12.i02.p411.
- [2] E. W. Liyanto, A. Homaidi, and A. Lutfi, "IMPLEMENTASI K-MEANS CLUSTERING MENGGUNAKAN RAPIDMINER DALAM PENGELOMPOKAN DATA KUNJUNGAN WISATAWAN ASING DI PROVINSI JAWA TIMUR," vol. 19, pp. 205-216, 2024.
- [3] Miftahul Arif Aldi and Zaehol Fatah, "IMPLEMENTASI K-MEANS CLUSTERING DALAM PENGELOMPOKAN DATA KUNJUNGAN WISATAWAN ASING DI INDONESIA," *Jurnal Ilmiah Multidisiplin Ilmu*, vol. 2, no. 1, pp. 13-19, Feb. 2025, doi: 10.69714/3hhfj353.
- [4] Muqorobin, Sri Supatminingsih, and Dewi Hermawati Wahyuningsih, "Pengembangan Metode Klasterisasi Untuk Pengelompokan Potensi Wisata guna Menunjang Digitalisasi Wisata di Indonesia," *JST (Jurnal Sains dan Teknologi)*, vol. 13, no. 2, pp. 217-230, Jul. 2024, doi: 10.23887/jstundiksha.v13i2.84743.
- [5] H. N. Prabowo, R. Setyadi, and W. A. Prabowo, "Application of Data Mining for Clustering of Foreign Tourist Visits Based on Arrival Entrance," *Sinkron*, vol. 7, no. 1, pp. 49-58, Jan. 2022, doi: 10.33395/sinkron.v7i1.11217.
- [6] A. Yani, Z. Azmi, D. Suherdi, S. Informasi, and S. Triguna Dharma, "Implementasi Data Mining Menganalisa Data Penjualan Menggunakan Algoritma K-Means Clustering," *Maret*, vol. 2, no. 2, pp. 315-323, May 2025, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jsi>
- [7] J. I. Farida and A. H. Lubis, "Grouping Of Tourism Locations In Indonesia Using Distance Variations In The K-Means Algorithm," *JITE*, vol. 8, May 2025, doi: 10.31289/jite.v8i3SpC.14528.
- [8] A. Helia and C. Lukman Rohmat, "ANALISIS PENGELOMPOKAN DAYA TARIK OBYEK WISATA BERDASARKAN JENISNYA MENGGUNAKAN METODE K-MEANS PADA DATA PEMPROV JABAR," 2024.
- [9] F. Pramataning Dewi, P. Siwi Aryni, and Y. Umaidah, "Implementasi Algoritma K-Means Clustering Seleksi Siswa Berprestasi Berdasarkan Keaktifan dalam Proses Pembelajaran," MEI, 2011.
- [10] K. Aprilia and F. Sembiring, "ANALISIS GARIS KEMISKINAN MAKANAN MENGGUNAKAN METODE ALGORITMA K-MEANS CLUSTERING," 2021.
- [11] M. Putri Dalipang and L. Kurniasari Wibisono, "TOURIST CLUSTER ANALYSIS FOR THE DEVELOPMENT OF KE'TE' KESU' TOURISM OBJECT," 2025.
- [12] K. S. Dharmawan, G. Made, A. Sasmita, and N. Piarsa, "Clustering Tourism Destinations in Denpasar City in 2023 Based on Visitor Preferences Using K-Means and DBSCAN Clustering Methods," 2025.
- [13] D. Rahmiyati and E. Britantono Siswanto, "Penerapan Metode K-Means untuk Pemetaan Objek Wisata Sebagai Rekomendasi Prioritas Pengembangan Pariwisata," 2025. [Online]. Available: <http://creativecommons.org/licenses/by/4.0/>
- [14] L. Hesti and C. M. Karyati, "Komparasi Algoritma K-Means Dan K-Medoids Untuk Klasterisasi Data Mining Perjalanan Wisatawan Nusantara Terhadap Jumlah Penduduk," *Jurnal Ilmiah Komputasi*, vol. 23, no. 4, Dec. 2024, doi: 10.32409/jkistik.23.4.3681.
- [15] H. Hasanah, N. A. Sudiby, and R. M. Galih, "Data Mining Using K-Means Clustering Algorithm for Grouping Countries of Origin of Foreign Tourist," *Galaxy Science*, Sep. 2021. doi: 10.11594/nstp.2021.1112.
- [16] J. Ipmawati, "SEGMENTASI PENDENGAR MUSIK BERDASARKAN KEBIASAAN DAN DAMPAKNYA TERHADAP KESEHATAN MENTAL MENGGUNAKAN K-MEANS DAN RANDOM FOREST," vol. 11, no. 1, pp. 34-44, 2025, [Online]. Available: www.kaggle.com.



- [17] I. G. K. K. Putra and I. G. W. S. Dharma, “Application of The K-Means Clustering Method To Search For Potential Tourists of Bendesa Hotel,” *TIERS Information Technology Journal*, vol. 4, no. 1, pp. 8-15, Jun. 2023, doi: 10.38043/tiers.v4i1.4297.
- [18] K. Gustipartsani, N. Rahaningsih, R. D. Dana, and I. Y. Mustafa, “DATA MINING CLUSTERING MENGGUNAKAN ALGORITMA K-MEANS PADA DATA KUNJUNGAN WISATAWAN DI KABUPATEN KARAWANG,” 2023.
- [19] Y. E. Patendeng¹, M. T. B. Maramis², D. Mandej³, J. E. Pembangunan, F. Ekonomi, and D. Bisnis, “ANALISIS PENGARUH SEKTOR PARIWISATA TERHADAP PENDAPATAN ASLI DAERAH (Studi Pada Kabupaten Toraja Utara Tahun 2009-2023),” 2025.
- [20] H. Hamdani and D. Hartama, “IMPLEMENTATION OF DATA ANALYSIS HOTEL RATING LEVELS IN BALI USING THE K-MEANS ALGORITHM AND DECISION TREE,” *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 10, no. 3, pp. 467-474, Jun. 2024, doi: 10.33330/jurteksi.v10i3.2894.