

Analisis Segmentasi Wisatawan Menggunakan Algoritma K-Means Clustering untuk Personalisasi Strategi Pemasaran Pariwisata

Doddy Handoko¹, Yoga Azhari, Nurmulia Risky Risalman³, Edi Muslim^{4*} Agil Ramadan⁵, Muhammad Habibullah Wijaya⁶

Universitas Baturaja, Ogan Komering Ulu, Baturaja Timur, Indonesia

doddyhandoko24@gmail.com, yogaazhari215@gmail.com,
Nurmuliarskyrisalman0@gmail.com,
edim69539@gmail.com, 5Agilterata21@gmail.com,
habibwijaya233@gmail.com

ABSTRACT

The tourism sector in the Special Region of Yogyakarta faces significant challenges in maintaining repeat visits and tourist loyalty amidst increasingly competitive market conditions. This study aims to analyze tourist segmentation using the K-Means Clustering algorithm to formulate personalized and effective marketing strategy recommendations. Through a profound understanding of tourist characteristics, destination managers are expected to optimize regional revenue potential sustainably. Tourist visitation and reservation data from 2020 to 2025 were analyzed using the K-Means Clustering method, with cluster validity verified via the Elbow Method and Silhouette Score. The analysis results successfully categorized tourists into three distinct clusters: Cluster 0 represents loyal domestic tourists with a high frequency of visits but moderate spending; Cluster 1 consists of international tourists with low visit frequencies but highly diverse national origins; and Cluster 2 is an exclusive segment characterized by the highest spending levels and the longest length of stay. Based on these distinct traits, tactical marketing strategy recommendations were developed using the 4P Marketing Mix framework. The evaluation results indicate that the proposed personalized marketing strategies are highly relevant and adaptive for implementation to strengthen tourist engagement and boost economic growth within the tourism sector.

ABSTRAK

Sektor pariwisata di Daerah Istimewa Yogyakarta menghadapi tantangan yang signifikan dalam mempertahankan tingkat kunjungan berulang serta loyalitas wisatawan di tengah persaingan pasar yang semakin kompetitif. Penelitian ini bertujuan untuk menganalisis segmentasi wisatawan menggunakan algoritma *K-Means Clustering* guna merumuskan rekomendasi strategi pemasaran yang terpersonalisasi dan efektif. Melalui pemahaman mendalam terhadap karakteristik wisatawan, pengelola destinasi diharapkan dapat mengoptimalkan potensi pendapatan daerah secara berkelanjutan. Data kunjungan dan reservasi wisatawan dari tahun 2020 hingga 2025 dianalisis menggunakan metode *K-Means Clustering*, dengan pengujian validitas klaster melalui metode *Elbow* dan *Silhouette Score*. Hasil analisis berhasil mengelompokkan wisatawan ke dalam tiga klaster utama: Klaster 0 merupakan kelompok wisatawan domestik loyal dengan frekuensi kunjungan tinggi namun pengeluaran moderat; Klaster 1 terdiri dari wisatawan mancanegara dengan frekuensi kunjungan rendah namun memiliki keberagaman asal negara; dan Klaster 2 merupakan segmen eksklusif dengan tingkat pengeluaran tertinggi serta durasi tinggal paling lama. Berdasarkan karakteristik tersebut, rekomendasi strategi pemasaran taktis dikembangkan menggunakan kerangka bauran pemasaran (*Marketing Mix 4P*). Hasil evaluasi menunjukkan bahwa usulan strategi pemasaran terpersonalisasi ini sangat relevan dan adaptif untuk diimplementasikan guna memperkuat keterikatan wisatawan serta meningkatkan pertumbuhan ekonomi sektor pariwisata.

Article History

Submitted: 15 Juni 2026

Accepted: 22 Juni 2026

Published: 23 Juni 2026

Key Words

Tourist Segmentation, K-Means Clustering, 4P Marketing Mix, Personalized Strategy, Yogyakarta Tourism

Sejarah Artikel

Submitted: 15 Juni 2026

Accepted: 22 Juni 2026

Published: 23 Juni 2026

Kata Kunci

Segmentasi Wisatawan, *K-Means Clustering*, *Marketing Mix 4P*, Personalisasi Strategi, Pariwisata Yogyakarta.

PENDAHULUAN

Industri pariwisata di Daerah Istimewa Yogyakarta saat ini menghadapi tantangan besar dalam mempertahankan angka kunjungan berulang (*tourist retention rate*) serta loyalitas wisatawan akibat fluktuasi kunjungan pasca-pandemi dan ketatnya kompetisi pasar. Pihak otoritas dan pelaku pariwisata menyadari bahwa strategi pemasaran massal sudah tidak lagi efektif karena karakteristik wisatawan yang sangat majemuk, sementara ketiadaan analisis data historis kunjungan yang mendalam menjadi hambatan utama dalam merancang promosi yang tepat sasaran. Sebagai solusinya, penelitian ini mengusulkan penerapan algoritma *K-Means Clustering* untuk mengelompokkan wisatawan berdasarkan kemiripan pola tertentu, seperti anggaran pengeluaran, durasi kunjungan, dan aktivitas reservasi. Melalui segmentasi berbasis data ini, pengelola destinasi dapat merumuskan bauran strategi pemasaran pariwisata yang terpersonalisasi guna meningkatkan keterikatan (*engagement*) serta retensi pelancong secara optimal.

Secara empiris, keandalan *K-Means Clustering* telah terbukti di berbagai sektor, seperti dalam mengoptimalkan strategi promosi *e-commerce* hingga mendongkrak penjualan sebesar 30%, memetakan segmentasi perilaku belanja, meninjau preferensi publik via analisis sentimen media sosial, serta mengefisiensikan manajemen inventaris [4]. Kebaruan (*novelty*) dari riset ini terletak pada implementasi komprehensif *K-Means* di sektor pariwisata daerah yang selama ini masih jarang dieksplorasi. Dengan demikian, tujuan utama dari penelitian ini adalah melakukan klusterisasi wisatawan menggunakan algoritma *K-Means* yang divalidasi dengan metode *Elbow* dan *Silhouette Score*. Hasil pengelompokan tersebut kemudian diintegrasikan dengan kerangka kerja *Marketing Mix 4P* (*Product, Price, Place, Promotion*) untuk menghasilkan rekomendasi strategi pemasaran yang aplikatif dan terpersonalisasi demi mendongkrak daya saing pariwisata Yogyakarta secara berkelanjutan.

METODE PENELITIAN

Penelitian ini menerapkan metode *K-Means Clustering* untuk menganalisis segmentasi wisatawan dalam rangka mengoptimalkan strategi pemasaran pariwisata terpersonalisasi di Yogyakarta. Alur pengerjaan riset ini diawali dengan tahapan pengumpulan data historis kunjungan dan karakteristik wisatawan, yang kemudian dilanjutkan ke proses *preprocessing* data untuk mengeliminasi informasi yang tidak relevan maupun data yang tidak lengkap. Setelah dataset dipastikan bersih dan siap, algoritma *K-Means Clustering* diimplementasikan untuk mengelompokkan wisatawan ke dalam beberapa klaster berdasarkan kesamaan pola perjalanan serta preferensi mereka. Output dari pembentukan klaster tersebut kemudian dibedah secara mendalam menggunakan kerangka kerja bauran pemasaran atau *Marketing Mix 4P* (*Product, Price, Place, Promotion*) guna menyusun rekomendasi strategi pemasaran yang spesifik dan terpersonalisasi. Seluruh rangkaian metodologi ini diadopsi secara sistematis mulai dari pengumpulan data awal, pembersihan, proses klusterisasi, evaluasi performa model, hingga tahap perancangan rekomendasi taktis.

Proses komputasi riset dimulai dengan mengeksplorasi data sekunder berupa catatan reservasi dan kunjungan wisatawan di destinasi pariwisata Yogyakarta. Tahap *preprocessing* dijalankan secara ketat meliputi pembersihan data (*data cleaning*) guna mengatasi nilai yang hilang (*missing values*), redundansi data (*duplication*), serta

inkonsistensi format penulisan agar validitas pengujian terjaga. Tahap berikutnya adalah penentuan jumlah klaster optimal (\$K\$) dengan memanfaatkan pendekatan *Elbow Method* yang berfungsi mengidentifikasi jumlah kelompok wisatawan paling representatif berdasarkan pola variansi data kunjungan mereka. Selanjutnya, algoritma *K-Means Clustering* dijalankan untuk membagi wisatawan ke dalam kelompok-kelompok homogen berdasarkan kesamaan karakteristik tertentu. Guna menjamin akurasi dan kekuatan pengelompokan, dilakukan evaluasi matematis menggunakan nilai *Silhouette Index* (*Silhouette Score*) untuk memastikan bahwa klaster yang terbentuk telah memiliki kohesi internal yang tinggi dan pemisahan antar-klaster yang optimal.

Hasil segmentasi wisatawan yang telah tervalidasi kemudian dianalisis karakteristik uniknya untuk melahirkan strategi bauran pemasaran (*Marketing Mix*) 4P yang bersifat personal.

- **Dari aspek Produk (*Product*)**, paket wisata, fasilitas destinasi, dan program hiburan dikembangkan secara khusus agar selaras dengan preferensi dari setiap klaster wisatawan.
- **Pada aspek Harga (*Price*)**, penyesuaian tarif tiket masuk, harga paket tur, serta bundling layanan dilakukan dengan mempertimbangkan daya beli dan pola pengeluaran (*spending habits*) masing-masing kelompok.
- **Untuk aspek Tempat (*Place*)**, saluran distribusi informasi dan platform pemesanan tiket dioptimalkan baik melalui agen travel digital (OTA) maupun kanal resmi milik destinasi guna mempermudah aksesibilitas wisatawan.
- **Terakhir, pada aspek Promosi (*Promotion*)**, kampanye pemasaran dirancang secara variatif dan segmentatif menggunakan media digital maupun konvensional agar pesan promosi dapat tersampaikan secara efisien dan tepat sasaran ke target klaster yang dituju

HASIL DAN PEMBAHASAN

Implementasi K-Means Clustering

Proses pengelompokan wisatawan dilakukan melalui dua tahapan utama, yaitu mencari jumlah klaster yang paling ideal (*optimal number of clusters*) serta mengeksekusi algoritma klasterisasi tersebut. Penentuan nilai \$K\$ terbaik dievaluasi secara matematis menggabungkan metode grafik *Elbow* dan nilai *Silhouette Score*

Tabel 4.1 Hasil Pengujian Validitas Nilai *Silhouette Score*

Jumlah Klaster (K)	Nilai Silhouette Score	Status Validasi
<i>K=2</i>	0,5401	Moderat
<i>K=3</i>	0,6526	Optimal (tertinggi)
<i>K=4</i>	0,6090	Baik
<i>K=5</i>	0,6111	Baik
<i>K=6</i>	0,5673	Moderat
<i>K=7</i>	0,5600	Moderat
<i>K=8</i>	0,5756	Moderat
<i>K=9</i>	0,5283	Lemah
<i>K=10</i>	0,5291	Lemah

Berdasarkan Tabel 4.1, jumlah klaster paling optimal yang dipilih untuk riset ini adalah tiga klaster ($K=3$). Nilai *Silhouette Score* menunjukkan angka tertinggi, yakni **0,6526** pada saat data dibagi menjadi 3 bagian. Hal ini mengindikasikan tingkat kepadatan internal objek di dalam klaster yang tinggi (*good cohesion*) sekaligus batas pemisahan antar-klaster yang sangat jelas (*good separation*). Data riil wisatawan kemudian diklasifikasikan menggunakan tiga variabel utama: frekuensi kunjungan (*total visits*), rata-rata durasi tinggal (*average room nights*), dan rata-rata nominal transaksi finansial (*average final payment*).

Analisis Karakteristik Segmentasi Wisatawan

Melalui pengolahan data akhir pariwisata Yogyakarta, ditemukan pembagian kelompok wisatawan yang teridentifikasi secara jelas:

- **Klaster 0 (Segmen Wisatawan Loyal / *Stable Mid-Range*):** Merupakan kelompok wisatawan yang didominasi oleh pelancong domestik dengan tingkat kunjungan sangat stabil dan tinggi (rata-rata mencapai 371 kali kunjungan), durasi tinggal tergolong menengah (rata-rata 3,73 malam), serta pengeluaran biaya liburan pada tingkat moderat (rata-rata Rp2.795.015,00).
- **Klaster 1 (Segmen Wisatawan Heterogen / *Mass Tourist*):** Merupakan kelompok dengan cakupan wilayah geografis paling luas (mencakup 35 negara/daerah asal), namun mencatat performa kunjungan yang rendah (rata-rata 6,57 kunjungan), durasi tinggal paling singkat (rata-rata 1,76 malam), serta alokasi anggaran belanja pariwisata terendah (rata-rata Rp1.585.194,00).
- **Klaster 2 (Segmen Eksklusif / *Premium Tourist*):** Merupakan kelompok dengan kuantitas wisatawan paling sedikit dan frekuensi kedatangan rendah (rata-rata hanya 4 kunjungan), namun memiliki nilai ekonomi sangat tinggi. Kelompok ini mencatat durasi liburan terlama (rata-rata mencapai 12,25 malam) dan mencatatkan rata-rata pengeluaran finansial paling tinggi di antara semua segmen, yakni menembus angka Rp12.476.455,00. Segmen premium ini umumnya dihuni oleh pelancong dari negara-negara jauh seperti Amerika Serikat, Inggris, Australia, dan China.

Personalisasi Strategi Pemasaran Berbasis Bauran Pemasaran (4P)

Rekomendasi taktis disusun menggunakan integrasi hasil klasterisasi dengan pilar *4P Marketing Mix* agar strategi yang dijalankan tepat sasaran dan terpersonalisasi sesuai kebiasaan tiap segmen:

1. Strategi untuk Klaster 0 (Segmen Wisatawan Loyal)

- **Product:** Menyediakan paket liburan berdurasi jangka menengah (3-4 malam) yang mencakup akomodasi standar plus dengan fasilitas tambahan gratis seperti akses Wi-Fi cepat dan sarapan khas lokal.
- **Price:** Menerapkan skema harga kompetitif kelas menengah, promo pendaftaran awal (*early bird booking*), serta opsi fleksibilitas sistem pembayaran.
- **Place:** Mendistribusikan penawaran melalui integrasi agen perjalanan digital (OTA) dan jalur reservasi langsung (*direct booking*).
- **Promotion:** Menjalankan program poin loyalitas wisatawan untuk meningkatkan angka retensi kunjungan, didukung kampanye media sosial aktif, *email marketing*,

serta kolaborasi paket hemat (*bundling*) bersama restoran lokal atau penyedia aktivitas wisata daerah.

2. Strategi untuk Klaster 1 (Segmen Wisatawan Heterogen)

- **Product:** Menawarkan paket efisiensi wisata berdurasi pendek (2-3 malam) dengan standarisasi kenyamanan fasilitas dasar destinasi pariwisata yang memenuhi ekspektasi umum pelancong lintas wilayah.
- **Price:** Menetapkan strategi harga yang bersaing, pemberian diskon khusus pada masa sepi kunjungan (*low-season discounts*), serta harga grup rombongan.
- **Place:** Mengoptimalkan jangkauan pasar yang sensitif terhadap harga dengan memaksimalkan kerja sama melalui OTA, *marketplace*, dan kanal platform digital massal lainnya.
- **Promotion:** Menggunakan program rujukan (*referral programs*), pemasaran digital yang hemat biaya, serta penyediaan paket *bundling* destinasi yang terintegrasi dengan moda transportasi lokal guna memberikan nilai tambah.

3. Strategi untuk Klaster 2 (Segmen Eksklusif / Premium)

- **Product:** Menyusun paket menginap dan berwisata premium jangka panjang (4-5 malam ke atas), lengkap dengan layanan yang sangat personal (*personalized services*) seperti layanan penyambutan eksklusif, cinderamata premium, serta akses prioritas ke fasilitas destinasi mewah.
- **Price:** Menerapkan strategi harga premium yang mencerminkan prestise dan tingginya kualitas layanan (*value of service*), program keuntungan khusus anggota (*member-only benefits*), dan skema peningkatan kelas akomodasi (*room upgrade schemes*).
- **Place:** Mengarahkan distribusi informasi dan jalur pemesanan eksklusif melalui kemitraan dengan agen perjalanan premium internasional serta platform reservasi khusus kelas atas.
- **Promotion:** Memfokuskan kampanye pada keanggotaan VIP privat, pemasaran langsung secara personal (*personalized direct marketing*), serta menyelenggarakan acara atau festival seni-budaya khusus yang sesuai dengan gaya hidup kelas premium guna memaksimalkan potensi pendapatan dari segmentasi ini.

KESIMPULAN DAN SARAN

Berdasarkan analisis data kunjungan wisatawan di Daerah Istimewa Yogyakarta, riset ini membuktikan bahwa pemetaan segmentasi pasar memegang peran yang sangat krusial dalam menyusun strategi pemasaran pariwisata yang efektif dan terpersonalisasi. Penerapan algoritma *K-Means Clustering* yang divalidasi melalui kombinasi *Elbow Method* dan *Silhouette Score* sukses menghasilkan tiga klaster wisatawan paling optimal dengan nilai validitas tinggi sebesar **0,6526**. Setiap klaster yang terbentuk memaparkan karakteristik perilaku perjalanan yang unik, yang kemudian ditransformasikan menjadi basis rancangan bauran pemasaran (*Marketing Mix 4P*). Strategi yang direkomendasikan mencakup pengelolaan program loyalitas dan tarif kompetitif untuk wisatawan domestik setia berdurasi tinggal moderat (Klaster 0), promosi bernilai ekonomis tinggi untuk pelancong massal (Klaster 1), serta pengadaan paket wisata eksklusif berbiaya tinggi guna menangkap pasar wisatawan mancanegara premium (Klaster 2). Hasil studi ini tidak hanya memperluas wawasan mengenai dinamika tren pelancong, melainkan juga menyajikan peta jalan taktis yang adaptif bagi pengelola destinasi untuk menggenjot daya

saing pariwisata daerah di pasar global. Sebagai langkah strategis ke depan, pemangku kebijakan dan pelaku industri pariwisata Yogyakarta disarankan untuk mulai mengadopsi sistem basis data digital yang terintegrasi demi mewujudkan mekanisme segmentasi wisatawan otomatis berbasis data (*data-driven*). Evaluasi berkala terhadap efisiensi program bauran pemasaran di setiap segmen juga mutlak diperlukan untuk mengimbangi pergeseran tren preferensi liburan yang bersifat dinamis. Terakhir, pengelola pariwisata disarankan untuk meningkatkan fokus alokasi pengembangan fasilitas maupun layanan premium yang menyasar kelompok wisatawan eksklusif (Klaster 2). Kendati persentase jumlah kunjungannya paling sedikit, segmen ini terbukti mendominasi dari aspek durasi tinggal terlalu lama serta volume pengeluaran finansial tertinggi, sehingga menjadi kunci utama dalam memaksimalkan potensi pendapatan sektor pariwisata secara berkelanjutan.

DAFTAR PUSTAKA

- [1] Muhammad, H., Haryadi, H., & Sularno, S. (2024). Segmentasi Pelanggan Berdasarkan Pendapatan Dan Skor Pengeluaran Menggunakan Metode K-Means. *Jurnal Sains dan Teknologi (JSIT)*, 4(3), 325–337. <https://doi.org/10.47233/jsit.v4i3.2468>
- [2] Adryan, Y., Rizqi, F., Bayu, I., Yosafat, R., & Saprudin. (2024). Analisis Segmentasi Pelanggan Mall Menggunakan Metode K-Means Clustering Dan Elbow Untuk Optimasi Strategi Pemasaran. *JRIIN: Jurnal Riset Informatika dan Inovasi*, 2(9), 1698–1708.
- [3] Geni, S. P. L., Safitri, H., Merry, F., Wati, M., & Havaluddin, H. (2025). Penerapan K-Means Clustering untuk Segmentasi Konsumen E-Commerce Berdasarkan Pola Pembelian. *JUKI: Jurnal Komputer dan Informatika*, 7(1), 89–99. <https://ioinformatic.org/index.php/JUKI/article/view/1049>
- [4] Hadiwibowo, A. W. T. (2023). Analisis Segmentasi Pengunjung Mall Menggunakan Algoritma K-Means Clustering. *JUKOMIKA (Jurnal Ilmu Komputer dan Informatika)*, 6(2), 42–51.
- [5] Siagian, R. I. P., Pratama, E., Lubis, F. A., Priscilia, S. A., & Ramadhani, F. (2025). Segmentasi Pelanggan dengan Algoritma K-Means untuk Strategi Pemasaran yang Efektif. *JATI: Jurnal Mahasiswa Teknik Informatika*, 9(4), 5615–5620.
- [6] Pakpahan, M., Sahid, S., Simanullang, M., & Winanda, R. (2026). Analisis Hierarchical Clustering untuk Segmentasi Pelanggan pada Dataset Mall Customers. *Jurnal Sistem Informasi dan Informatika*, 4(1), 58–63. <https://doi.org/10.47233/jiska.v4i1.2615>
- [7] Yunita, I., Ali, P. R., Kartawidjaja, M. A., & Sukwadi, R. (2025). Segmentasi Pelanggan Menggunakan K-Means Clustering: Menganalisis Metrik RFM untuk Strategi Pemasaran. *Jurnal Digitech*, 5(2).
- [8] Januzaj, Y., Beqiri, E., & Luma, A. (2023). Determining the Optimal Number of Clusters using Silhouette Score as a Data Mining Technique. *International Journal of Online and Biomedical Engineering*, 19(4), 174–182.
- [9] Prayitno, E., Perdana, I. J., Iskandar, E., Winarno, B. H., & Subagyo, A. A. (2024). Optimalisasi Profitabilitas Ritel Melalui Segmentasi Pelanggan dengan K-Means Clustering. *Jurnal Informasi Interaktif*, 9(3), 113–120.
- [10] Fauzan, R. M., & Alfian, G. (2024). Segmentasi Pelanggan E-Commerce Menggunakan Fitur Recency, Frequency, Monetary (RFM) dan Algoritma Klasterisasi K-Means. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 9(1), 1–9. <https://doi.org/10.14421/jiska.2024.9.1.1-9>

- [11] Rahmawati, S., & Putra, I. (2024). Penerapan Algoritma K-Means dalam Segmentasi Anggota Koperasi Berdasarkan Pola Simpanan dengan Analisis RFMP untuk Meningkatkan Loyalitas. *Bulletin of Computer Science Research*. <http://www.hostjournals.com/bulletincsr/article/view/843>