

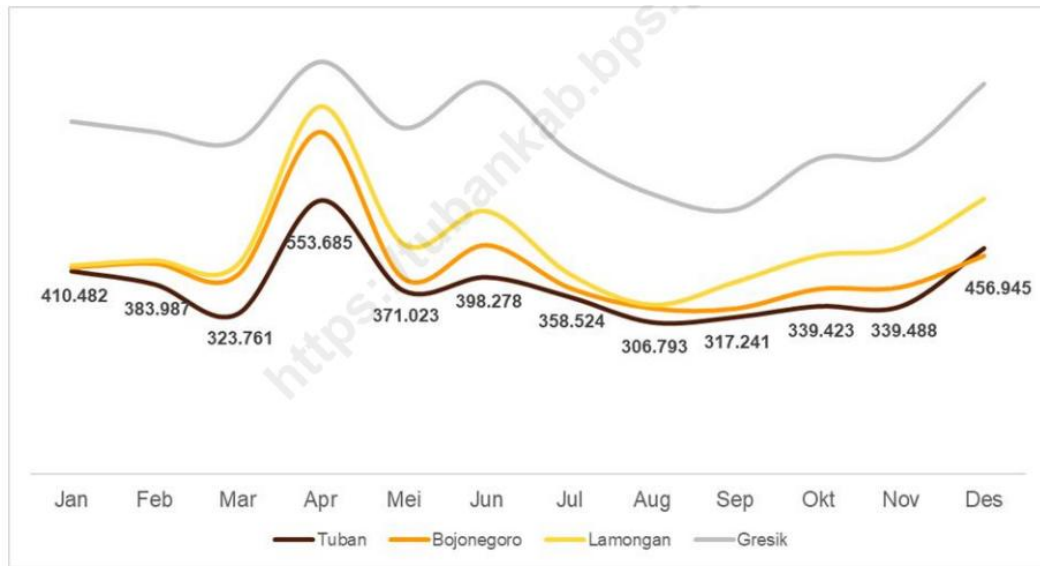
BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara yang memiliki potensi pariwisata yang besar dengan keanekaragaman wisata alam, budaya, sejarah, religi, dan kuliner yang tersebar di berbagai daerah. Sektor pariwisata menjadi salah satu sektor penting dalam pembangunan daerah karena mampu menarik wisatawan serta meningkatkan perekonomian daerah (Metode et al., 2022). Kabupaten Tuban yang terletak di Provinsi Jawa Timur merupakan salah satu daerah yang memiliki beragam potensi wisata. Destinasi wisata di Kabupaten Tuban meliputi wisata alam, sejarah, religi, dan budaya. Beberapa objek wisata yang cukup dikenal antara lain Goa Akbar, Pantai Boom, Makam Sunan Bonang, serta Klenteng Kwan Sing Bio. Keberagaman objek wisata tersebut menunjukkan bahwa Kabupaten Tuban memiliki potensi pariwisata yang cukup besar untuk dikembangkan. Namun demikian, pengelolaan promosi serta penyediaan rekomendasi tempat wisata kepada wisatawan masih belum optimal.

Berdasarkan data Badan Pusat Statistik Kabupaten Tuban (2024) melalui Survei Wisatawan Nusantara (Wisnus) Digital, jumlah wisatawan dari luar Kabupaten Tuban yang berkunjung pada tahun 2024 tercatat sebanyak sekitar 4,56 juta wisatawan. Jumlah tersebut mengalami penurunan sebesar 5,4 persen dibandingkan tahun 2023 yang mencapai sekitar 4,82 juta wisatawan. Dari total wisatawan tersebut, sebanyak 3,39 juta wisatawan berasal dari kabupaten atau kota di Provinsi Jawa Timur, sedangkan sekitar 1,12 juta wisatawan berasal dari luar Provinsi Jawa Timur. Daerah dengan jumlah kunjungan wisatawan terbesar ke Kabupaten Tuban berasal dari Kabupaten Bojonegoro sebesar 693,85 ribu wisatawan, diikuti oleh Kabupaten Lamongan sebesar 677,8 ribu wisatawan dan Kota Surabaya sebesar 659,6 ribu wisatawan. Sementara itu, jumlah kunjungan wisatawan paling sedikit berasal dari Kabupaten Pacitan yaitu sebesar 2,4 ribu wisata.



Gambar 1. 1 Jumlah Wisatawan di Kabupaten Tuban dan Kabupaten sekitar tahun 2024

Berdasarkan Gambar 1.1, terlihat bahwa pola kunjungan wisatawan nusantara di beberapa kabupaten yang berdekatan dengan Tuban, seperti Lamongan, Bojonegoro, dan Gresik menunjukkan pola yang relatif serupa. Jumlah kunjungan wisatawan terendah terjadi pada bulan Maret 2024 yang bertepatan dengan bulan Ramadhan. Pada periode tersebut aktivitas wisata cenderung menurun karena masyarakat lebih banyak memusatkan perhatian pada kegiatan keagamaan. Namun setelah periode tersebut, terjadi peningkatan signifikan pada bulan April 2024 yang bertepatan dengan momentum perayaan Idul Fitri. Periode libur panjang pada saat Idul Fitri menyebabkan peningkatan mobilitas masyarakat sehingga jumlah kunjungan wisatawan mengalami lonjakan di berbagai daerah.

Seiring dengan perkembangan teknologi informasi, berbagai platform digital seperti website, aplikasi mobile, dan media sosial telah banyak menyediakan informasi mengenai tempat wisata. Meskipun demikian, wisatawan sering kali mengalami kesulitan dalam menemukan rekomendasi tempat wisata yang sesuai dengan preferensi, karakteristik, dan kebutuhan mereka. Hal ini disebabkan karena sebagian besar sistem rekomendasi yang tersedia masih bersifat umum dan hanya berdasarkan popularitas atau rating tempat wisata. Oleh karena itu, diperlukan suatu metode yang dapat mengelompokkan destinasi wisata berdasarkan karakteristik

tertentu sehingga dapat memberikan rekomendasi yang lebih relevan bagi wisatawan.

Salah satu pendekatan yang dapat digunakan untuk mengolah dan menganalisis data dalam jumlah besar adalah data mining. Data mining merupakan proses ekstraksi pola atau informasi yang bermanfaat dari kumpulan data yang besar dengan menggunakan teknik statistik, matematika, dan machine learning (Hendrastuty, 2024). Melalui proses data mining, data yang sebelumnya tidak terstruktur dapat diolah menjadi informasi yang lebih bermakna sehingga dapat digunakan sebagai dasar pengambilan keputusan dalam berbagai bidang, termasuk sektor pariwisata (Triandini et al., 2021). Data mining akan menghasilkan model yang baik dalam klasifikasi, prediksi maupun estimasi jika performa machine learning yang digunakan juga handal. Performa machine learning yang baik akan sangat berguna untuk melakukan klasifikasi secara tepat. Tidak sedikit metode klasifikasi mengalami modifikasi agar menjadi handal dalam melakukan klasifikasi sehingga memberikan akurasi yang lebih baik. Namun seiring dengan meningkatnya data-data informasi yang direpresentasikan dalam berbagai format dan platform mengakibatkan peningkatan fitur data dan struktur data menjadi using. Terkadang tidak semua fitur maupun sample data tersebut mendukung untuk klasifikasi secara akurat, sehingga mengurangi performa algoritma klasifikasi (Akbar et al., 2023).

Pada dasarnya clustering merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (similarity) antara satu data dengan data yang lain (Iqbal, 2019). Clustering merupakan salah satu metode data mining yang bersifat tanpa arahan (unsupervised), maksudnya metode ini diterapkan tanpa adanya latihan (training) dan tanpa ada guru (teacher) serta tidak memerlukan target output. Dalam data mining ada dua jenis metode clustering yang digunakan dalam pengelompokan data, yaitu hierarchical clustering dan non-hierarchical clustering (Kurniawan et al., 2022). Clustering bertujuan untuk mengelompokkan objek berdasarkan tingkat kemiripan karakteristik sehingga objek dalam satu kelompok memiliki kesamaan yang tinggi dibandingkan dengan kelompok lainnya (Iqbal, 2019). Salah satu metode clustering yang populer adalah

K-Means Clustering, yaitu metode yang bekerja dengan membagi data ke dalam sejumlah cluster berdasarkan jarak terdekat terhadap pusat cluster (centroid). Metode ini banyak digunakan karena memiliki proses perhitungan yang relatif sederhana serta mampu mengelompokkan data dalam jumlah besar secara efisien.

K-Means merupakan salah satu metode clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster (Pramayuda et al., 2020). Metode ini mempartisi data ke dalam cluster sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam cluster yang lain. K-Means Clustering adalah suatu metode penganalisaan data atau metode Data Mining yang melakukan proses pemodelan tanpa supervisi (unsupervised) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi (Masruro et al., 2005). Terdapat dua jenis data clustering yang sering dipergunakan dalam proses pengelompokan data yaitu Hierarchical dan Non-Hierarchical, dan K-Means merupakan salah satu metode data clustering non-hierarchical atau Partitional Clustering (Sirait, 2017). Metode K-Means Clustering berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain.

Penelitian sebelumnya yang berjudul “Pengelompokan Destinasi Wisata di Batam Berdasarkan Daya Tarik dan Fasilitas Menggunakan Metode K-Means Clustering” bertujuan untuk mengelompokkan destinasi wisata di Kota Batam berdasarkan daya tarik dan fasilitas yang tersedia menggunakan metode K-Means Clustering dengan bantuan aplikasi RapidMiner. Hasil penelitian tersebut menghasilkan tiga cluster yaitu cluster dengan daya tarik dan fasilitas tinggi sebanyak tiga destinasi, cluster menengah sebanyak satu destinasi, dan cluster rendah sebanyak tiga destinasi (Astiti & Harman, 2022). Namun pada penelitian tersebut tidak dilakukan proses evaluasi untuk mengukur kualitas hasil pengelompokan yang dihasilkan.

Penelitian lain dengan judul “Penerapan Metode K-Means untuk Pemetaan Objek Wisata Sebagai Rekomendasi Prioritas Pengembangan Pariwisata” bertujuan untuk memetakan objek wisata serta menentukan rekomendasi prioritas pengembangan wisata menggunakan metode K-Means. Hasil penelitian tersebut menghasilkan tiga cluster dengan jumlah objek wisata pada cluster pertama sebanyak empat objek wisata, cluster kedua sebanyak sepuluh objek wisata, dan cluster ketiga sebanyak delapan puluh satu objek wisata. Proses evaluasi menggunakan metode Silhouette Coefficient menghasilkan nilai sebesar 0,78 yang menunjukkan kualitas cluster yang cukup baik (Rahmiyati et al., 2025). Namun dalam penelitian tersebut tidak dilakukan proses preprocessing data seperti transformasi maupun normalisasi data sebelum proses clustering.

Berdasarkan permasalahan tersebut, penelitian ini menerapkan metode K-Means Clustering untuk mengelompokkan destinasi wisata di Kabupaten Tuban. Sebelum proses clustering dilakukan, data terlebih dahulu dikonversi ke dalam skala ordinal dari 1 hingga 5 untuk variabel harga dan jarak menggunakan teknik discretization. Proses ini bertujuan untuk menyederhanakan data numerik kontinu menjadi kategori ordinal yang memiliki makna urutan atau ranking sehingga dapat mempermudah proses interpretasi hasil clustering yang diperoleh (Jajang et al., 2022). Pada Penelitian sebelumnya belum melakukan evaluasi cluster atau preprocessing data secara optimal.

Melalui penelitian ini diharapkan dapat memberikan rekomendasi pengelompokan tempat wisata yang lebih akurat serta membantu wisatawan dalam menentukan destinasi wisata yang sesuai dengan preferensi mereka. Selain itu, hasil penelitian ini juga diharapkan dapat membantu pemerintah daerah dalam mengidentifikasi potensi wisata yang perlu dikembangkan sehingga dapat meningkatkan jumlah kunjungan wisatawan di Kabupaten Tuban.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana hasil pengelompokan wisata menggunakan metode K-means Clustering ?
2. Berapa nilai evaluasi dari silhouette coefficient menggunakan metode K-means Clustering ?
3. Berapa jumlah cluster optimal dari penerapan algoritma K-Means ?
4. Bagaimana karakteristik setiap cluster wisata yang terbentuk ?

1.3 Tujuan Dan Manfaat

1.3.1 Tujuan Penelitian

Adapun tujuan dari penelitian ini sebagai berikut :

1. Mengelompokkan tempat wisata di Kabupaten Tuban menggunakan algoritma K-Means Clustering berdasarkan karakteristik tertentu, seperti harga, jumlah fasilitas, jarak, rating, dan tiket.
2. Menentukan cluster optimal dan menganalisis karakteristik cluster.
3. Menganalisis karakteristik cluster.

1.3.2 Manfaat Penelitian

Manfaat dari penelitian ini sebagai berikut :

1. Manfaat bagi Pemerintah

Hasil pengelompokan destinasi wisata dapat membantu pemerintah daerah dalam mengidentifikasi karakteristik setiap kelompok wisata berdasarkan indikator seperti harga, rating, jarak, fasilitas, dan durasi kunjungan. Hal ini mendukung perumusan strategi pengembangan pariwisata yang lebih tepat sasaran serta optimalisasi alokasi sumber daya dan promosi.

2. Manfaat bagi Pengelola Wisata

Pengelompokan destinasi memberikan informasi mengenai posisi dan karakteristik objek wisata yang dikelola, sehingga dapat digunakan sebagai dasar evaluasi dan peningkatan kualitas layanan, fasilitas, serta daya tarik wisata guna menarik lebih banyak pengunjung.

3. Manfaat bagi Akademik

Penelitian ini memberikan kontribusi dalam pengembangan ilmu pengetahuan, khususnya di bidang sistem informasi dan pariwisata, serta menjadi referensi terkait penerapan metode K-Means Clustering dalam pengelompokan dan rekomendasi destinasi wisata.

1.4 Batasan Masalah

Batasan dari penelitian ini adalah :

1. Obyek dari penelitian ini berfokus pada Kabupaten Tuban.
2. Metode yang digunakan dalam penelitian ini K-Means Clustering.
3. Variabel yang digunakan pada penelitian ini yaitu jarak dari pusat kota, harga tiket, lokasi, rating, jumlah fasilitas, durasi buka. . Mewakili aksesibilitas, fasilitas, biaya, dan kualitas destinasi.
4. Dataset 187 data
5. Metode elbow dan evaluasi silhouette.

1.5 Sistematika Penulisan

Dalam penelitian ini, sistematika penulisan yang digunakan penulis dalam penyusunan skripsi adalah sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini akan membahas mengenai latar belakang masalah, perumusan masalah, tujuan penelitian, manfaat yang diperoleh dari penerapan solusi yang diusulkan, batasan permasalahan, serta sistematika penulisan

BAB II TINJAUAN PUSTAKA

Pada bab ini akan menjelaskan secara rinci teori yang berkaitan dengan penelitian, meliputi penjelasan dasar mengenai obyek wisata, data mining, clustering, algorithm K-Means, serta penelitian terkait.

BAB III METODE PENELITIAN

Pada bab ini akan membahas mengenai tahapan penelitian, metode yang digunakan dalam penelitian, data yang digunakan, perancangan alur metode, perancangan alur sistem, dan proses evaluasi yang akan digunakan untuk mendukung metode penelitian.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini akan membahas mengenai hasil penelitian dan pembahasan analisis hasil penelitian

BAB V PENUTUP

Berisi Kesimpulan dari hasil penelitian serta saran untuk penelitian selanjutnya.