

## ANALISIS PREFERENSI WISATAWAN TERHADAP OBJEK WISATA DI KALIMANTAN MENGGUNAKAN CLUSTERING

Muhammad Raihan Efendi<sup>1</sup>, Fajar Maulana<sup>2</sup>, Fikry Saputro Ardiansyah<sup>3</sup>, Surya Rama Aditya<sup>4</sup>, Mohammad Rajib Balhaj Masdi<sup>5</sup>, Dimas Risma Alfian<sup>6</sup>, Muhammad Agung Febriyanto<sup>7</sup>, Tiara Ardilla Bazzury<sup>8</sup>, Mei Lestari<sup>9</sup>

1,2,3,4,5,6,7,8,9 Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer

Universitas Indraprasta PGRI

Jalan Raya Tengah No. 80, Kelurahan Gedong, Pasar Rebo, Jakarta Timur

[raihanefendi17@gmail.com](mailto:raihanefendi17@gmail.com)<sup>1</sup>, [fajarmaulana0611@gmail.com](mailto:fajarmaulana0611@gmail.com)<sup>2</sup>, [ikriardiyyah@gmail.com](mailto:ikriardiyyah@gmail.com)<sup>3</sup>,  
[ramma9748@gmail.com](mailto:ramma9748@gmail.com)<sup>4</sup>, [mohammadrajib111001@gmail.com](mailto:mohammadrajib111001@gmail.com)<sup>5</sup>, [dimassatria884@gmail.com](mailto:dimassatria884@gmail.com)<sup>6</sup>,  
[muhammad.agungf@gmail.com](mailto:muhammad.agungf@gmail.com)<sup>7</sup>, [tiarardillaa@gmail.com](mailto:tiarardillaa@gmail.com)<sup>8</sup>, [mei.lestari6@gmail.com](mailto:mei.lestari6@gmail.com)<sup>9</sup>

### Abstrak

Perkembangan industri di Indonesia sudah berkembang dengan pesat, hal ini ditandai dengan banyaknya tamu asing yang datang ke Indonesia untuk melakukan kunjungan sebagai objek daya tarik wisata. Selain karna tempat wisata yang indah Indonesia pun dikenal memiliki penduduk yang sangat ramah sehingga menarik para wisatawan asing. Pariwisata di Kalimantan memiliki potensi besar karena keindahan alam dan kekayaan budayanya, namun belum dioptimalkan secara maksimal. Tantangan utama dalam pengembangannya adalah kurangnya pemahaman terhadap preferensi wisatawan, yang menyebabkan ketidakseimbangan dalam distribusi kunjungan wisata dan kurang efektifnya strategi promosi. Oleh karena itu, penelitian ini bertujuan untuk menganalisis preferensi wisatawan terhadap objek wisata di Kalimantan. Tujuan *clustering* ini adalah menemukan preferensi wisatawan terhadap objek wisata Kalimantan. Studi ini memakai Metode *Clustering* untuk mengelompokkan objek wisata berdasarkan preferensi. Hasil analisis ini menunjukkan bahwa objek wisata di Kalimantan dapat dibagi menjadi beberapa *cluster*, yang digunakan untuk merekomendasikan pengembangan objek wisata. Dengan demikian, studi ini membantu pengembang pariwisata meningkatkan kunjungan wisatawan dan kualitas layanan yang ditawarkan.

**Kata Kunci:** *Clustering*, Preferensi Wisatawan, Kalimantan, Objek Wisata, Analisis.

### Abstract

*The development of the industry in Indonesia has developed rapidly, this is marked by the number of foreign guests who come to Indonesia to visit as an object of tourist attraction. Tourism in Kalimantan has great potential because of its natural beauty and cultural richness, but it has not been optimized. The main challenge in its development is the lack of understanding of tourist preferences, which leads to imbalances in the distribution of tourist visits and ineffective promotional strategies. Therefore, this study aims to analyze tourist preferences for tourist attractions in Kalimantan. The purpose of this clustering is to find tourist preferences for Kalimantan tourist attractions. This study uses the Clustering Method to group tourist attractions based on preferences. The results of this analysis show that tourist attractions in Kalimantan can be divided into several clusters, which are used to recommend the development of tourist attractions. Thus, the study helps tourism developers improve tourist visits and the quality of services offered.*

**Keywords:** *Clustering*, *Tourist Preferences*, *Kalimantan*, *Tourist Attractions*, *Analysis*.

### PENDAHULUAN

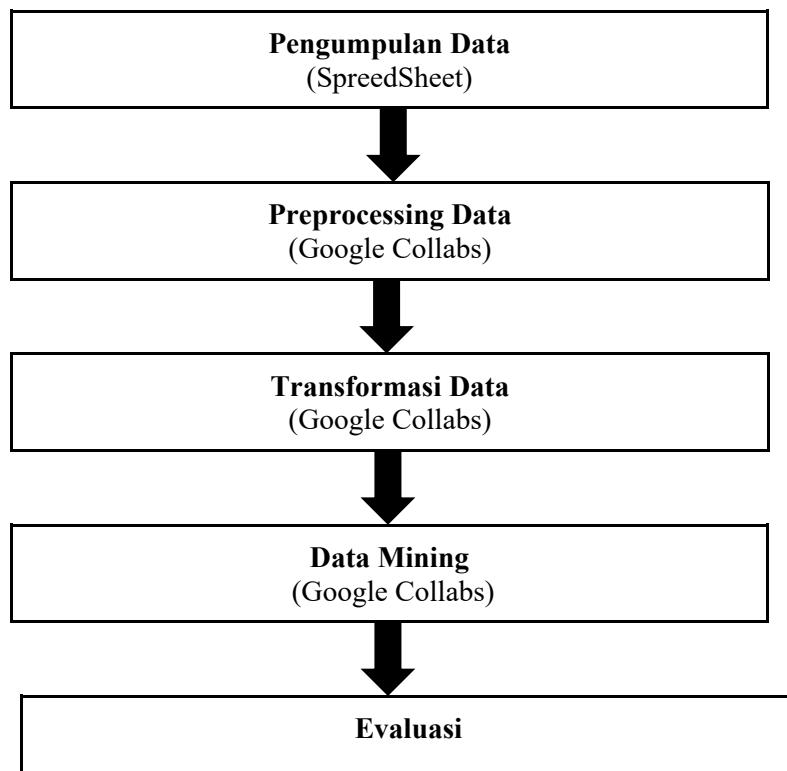
Pada era globalisasi, media sosial saat ini sudah sangat umum dan banyak digunakan untuk kepentingan masyarakat. Dalam implementasinya, media sosial lebih banyak digunakan untuk kegiatan jual beli, menyampaikan informasi, bahkan sebagai media untuk mengekspresikan diri. Sebab itu dengan kemajuan teknologi di era globalisasi ini banyak tamu asing yang datang ke Indonesia untuk berwisata. Selain itu, yang membuat para tamu asing merasa senang datang ke Indonesia, keramah - tamahan penduduknya juga merupakan salah satu faktor yang sangat mendukung [1]. Indonesia memiliki tempat wisata yang cantik dan indah, maka itu yang membuat wisatawan asing tertarik datang ke Indonesia dan ditambah dengan adanya para penduduk yang ramah tamah sehingga menjadi daya tarik tersendiri untuk para tamu asing. Pertumbuhan media sosial sangat cepat tidak hanya penggunaanya yang terus meningkat, namun semakin banyaknya

media sosial yang ditawarkan melalui aplikasi mobil ataupun *website* [2]. Maka wisatawan terhadap objek wisata di Kalimantan merupakan topik yang penting dalam pengembangan sektor pariwisata di Indonesia. Kalimantan, dengan kekayaan alam dan budaya yang melimpah, memiliki potensi besar untuk menarik wisatawan domestik maupun mancanegara. Namun, Segala bentuk dukungan sudah semestinya diperlukan dalam hal pengelolaan tempat yang memiliki potensi atau sudah menjadi lokasi pariwisata [3]. Mengoptimalkan potensi ini, diperlukan pemahaman yang mendalam mengenai preferensi dan perilaku wisatawan. Metode *Clustering*, khususnya algoritma K-means, dapat digunakan untuk mengelompokkan objek wisata berdasarkan karakteristik dan preferensi yang dimiliki oleh wisatawan.

Bidang pariwisata merupakan salah satu contoh sektor ekonomi unggulan dalam pembangunan suatu negara, terutama bagi negara Indonesia [4]. Dalam konteks ini, penelitian ini bertujuan untuk mengidentifikasi dan menganalisis preferensi wisatawan terhadap berbagai objek wisata di Kalimantan. Dengan menggunakan data kunjungan dan survei terhadap wisatawan, analisis *clustering* akan membantu dalam mengelompokkan objek wisata yang memiliki kesamaan dalam hal daya tarik, fasilitas, dan pengalaman yang ditawarkan. Hasil dari analisis ini diharapkan dapat memberikan rekomendasi strategis bagi pengembang pariwisata dan pemerintah daerah dalam merancang program promosi serta pengembangan objek wisata yang lebih sesuai dengan keinginan pasar.

Penentuan nilai *cluster* pada sistem rekomendasi ini didasarkan pada jarak minimum atau jarak terdekat antar objek data, berdasarkan jumlah wisatawan pada tempat wisata tersebut, jarak menuju tempat wisata tersebut, dan rating dari tempat wisata tersebut [5]. Maka diharapkan dapat tercipta pemerataan kunjungan ke berbagai objek wisata di Kalimantan, sehingga tidak hanya objek wisata populer yang mendapatkan perhatian, tetapi juga objek-objek lain yang memiliki potensi namun kurang dikenal. Dengan demikian, penelitian ini tidak hanya berkontribusi pada pengembangan ilmu pengetahuan di bidang pariwisata tetapi juga pada peningkatan ekonomi lokal melalui sektor pariwisata yang berkelanjutan.

## METODE PENELITIAN



Gambar 1. Tahapan Metode Penelitian

#### A. Pengumpulan Data

Data yang digunakan dalam penelitian ini berasal dari internet, ulasan wisatawan, dan informasi dari platform pariwisata. Data yang dikumpulkan mencakup informasi demografis wisatawan, preferensi mereka terhadap objek wisata di Kalimantan, serta faktor-faktor yang mempengaruhi keputusan mereka, seperti biaya, dan lokasi. Akan diseleksi untuk memastikan validitas dan relevansi dengan tujuan penelitian. Data yang tidak lengkap, redundan, atau tidak relevan akan hapus agar tidak terdapat nilai kosong (*Missing Value*).

#### B. Preprocessing Data

Proses *cleansing* data dilakukan untuk menghilangkan data yang tidak konsisten, atau menghapus atribut yang tidak diperlukan [6]. Seperti inkonsistensi, duplikasi, nilai kosong dan *outliner*.

#### C. Transformasi Data

Pada tahap transformasi data, format data yang telah dibersihkan dan diintegrasikan akan disesuaikan untuk analisis [7]. Kemudian, data diubah ke dalam format yang sesuai untuk dianalisis lebih lanjut. Proses ini melibatkan pengkodean variabel kategorikal, agregasi data, serta mengubah data dari non numerik menjadi numerik.

#### D. Data Mining

*Clustering* adalah di mana data dikelompokkan ke dalam beberapa *cluster* atau kelompok sehingga data dalam satu *cluster* memiliki tingkat kemiripan yang tinggi, sedangkan tingkat kemiripan antar *cluster* rendah [8]. Data *mining* dilakukan dengan menggunakan algoritma *Clustering K-Means*. Proses ini mencakup jumlah *cluster* optimal dan pembagian data berdasarkan kesamaan preferensi wisatawan. *Cluster* yang terbentuk menggambarkan kelompok wisatawan dengan karakteristik serupa, Berdasarkan rumus K-Means:

$$d(x_i, c_j) = \sqrt{\sum_{n=1}^p (x_{in} - c_{jn})^2}$$

#### E. Evaluasi

Tahap evaluasi dilakukan untuk menilai kualitas hasil *clustering*. Hasil evaluasi memastikan bahwa *cluster* yang terbentuk cukup representatif dan memberikan wawasan yang dapat digunakan untuk pengembangan pariwisata.

### HASIL DAN PEMBAHASAN

Untuk menganalisis preferensi wisatawan yang mengunjungi Kalimantan menggunakan Metode K-Means *Clustering*. Penelitian ini mengelompokkan wisatawan berdasarkan kunjungan wisata, tingkat kepuasan, dan kategori wisata yang banyak dikunjungi. Hasilnya menunjukkan adanya Tiga klaster utama wisatawan berdasarkan tempat wisata dan lokasi wisata . Dengan adanya analisis ini dapat digunakan untuk mengembangkan strategi pemasaran yang lebih efektif dan meningkatkan pelayanan di destinasi wisata Kalimantan. Tabel data wisata Kalimantan menampilkan berbagai tempat wisata yang tersebar di beberapa provinsi, seperti Kalimantan Timur, Kalimantan Selatan, Kalimantan Barat, dan Kalimantan Tengah. Tempat wisata ini terdiri dari berbagai kategori, termasuk taman hiburan, budaya nusantara, cagar alam, dan taman nasional. Beberapa destinasi terkenal antara lain Pantai Lamaru, Pasar Terapung Muara Kuin, Pulau Derawan, dan Taman Nasional Tanjung Puting. Data ini memberikan gambaran tentang ragam destinasi wisata yang dapat menjadi daya tarik utama di Kalimantan.

**Tabel 1.** Data Wisata Kalimantan

No	Tempat Wisata	Kategori	Provinsi
1.	Pantai Lamaru	Taman Hiburan	Kalimantan Timur
2.	Pasar Terapung Muara Kuin	Budaya Nusantara	Kalimantan Selatan
3.	Pulau Derawan	Cagar Alam	Kalimantan Timur
4.	Bukit Kelam	Cagar Alam	Kalimantan Barat
5.	Museum Wasaka	Budaya Nusantara	Kalimantan Selatan
6.	Taman Nasional Kutai	Taman Nasional	Kalimantan Timur
7.	Taman Nasional Tanjung Putting	Taman Nasional	Kalimantan Tengah
8.	Taman Wisata alam Sungai Wain	Cagar Alam	Kalimantan Timur
9.	Taman Nasional Betung Kerihun	Taman Nasional	Kalimantan Barat
10.	Taman Nasional Danau Sentarum	Taman Nasional	Kalimantan Barat

### Mengubah Data ke Numerik

Cara ini menggunakan metode LabelEncoder untuk mengubah data kategorikal menjadi numerik. Kolom **TEMPAT\_WISATA** dan **LOKASI**, yang awalnya berisi data teks (kategori), dikonversi ke format numerik agar dapat diproses oleh algoritma pembelajaran mesin. Proses ini dilakukan dengan memetakan setiap kategori unik ke nilai angka. Setelah transformasi, data yang sudah diproses ditampilkan untuk memastikan perubahan, bersama kolom lain seperti **RATING**. Langkah ini penting untuk mempersiapkan data sebelum digunakan dalam analisis atau model pembelajaran mesin.

```
# LabelEncoder untuk mengubah kolom kategorikal
le_tempat_wisata = LabelEncoder()
le_lokasi = LabelEncoder()

data['TEMPAT_WISATA'] = le_tempat_wisata.fit_transform(data['TEMPAT_WISATA'])
data['LOKASI'] = le_lokasi.fit_transform(data['LOKASI'])

# menampilkan data yang sudah diproses
print(data[['TEMPAT_WISATA', 'LOKASI', 'RATING']].head())
```

**Gambar 2.** Mengubah Data Numerik

### 1. Clustering Data

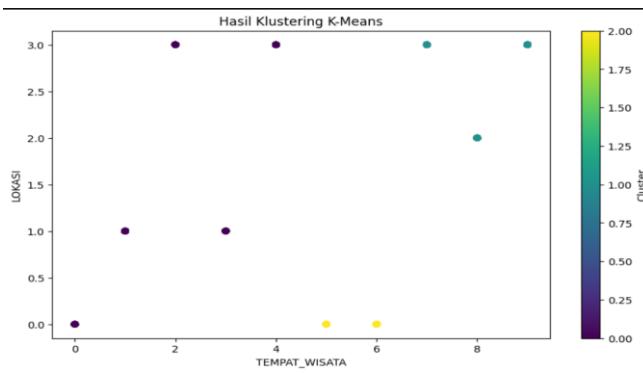
Cara ini menggunakan algoritma K-Means untuk mengelompokkan data ke dalam 3 cluster berdasarkan fitur **TEMPAT\_WISATA** dan **LOKASI**. Model K-Means diinisialisasi dengan jumlah cluster tertentu dan dilatih menggunakan metode **fit\_predict**, yang kemudian menyimpan hasil cluster di kolom baru bernama **CLUSTER**. Proses ini mempermudah segmentasi data berdasarkan kesamaan fitur tertentu.

```
[ ] # nentuin jumlah cluster
n_clusters = 3 # bisa ngubah jumlah cluster sesuai kebutuhan
kmeans = KMeans(n_clusters=n_clusters, random_state=30)

# ngelatih model K-Means
data['CLUSTER'] = kmeans.fit_predict(data[['TEMPAT_WISATA', 'LOKASI']])
```

**Gambar 3.** Clustering Data

Berikut adalah hasil visual grafik dari hasil data *Clustering K-Means* dengan berdasarkan tempat wisata dan lokasi wisata pada kalimantan.



Gambar 4. Grafik Clustering Data

## 2. Centroid

Hasil analisis data *Clustering* K-Means menghasilkan tiga kelompok (*cluster*) Tertinggi, Menengah dan Terendah berdasarkan dua variabel utama, yaitu 'Tempat Wisata' dan 'Lokasi'. Berikut adalah penjelasan lebih lanjut mengenai masing-masing *cluster* berdasarkan nilai centroidnya:

Tabel 2. Tabel Centroid dari setiap cluster

Cluster	Tempat Wisata	Lokasi
0	1.902439	1.557491
1	8.060606	2.666667
2	5.531034	3.330669

*Cluster 0 (Tertinggi):*

*Cluster* ini menunjukkan tempat wisata yang terletak di lokasi strategis dengan nilai centroid lokasi paling rendah (lebih dekat atau mudah dijangkau). Meskipun tingkat popularitasnya rendah, lokasi menjadi faktor utama yang membuat *cluster* ini dianggap unggul, misalnya untuk destinasi yang fokus pada aksesibilitas.

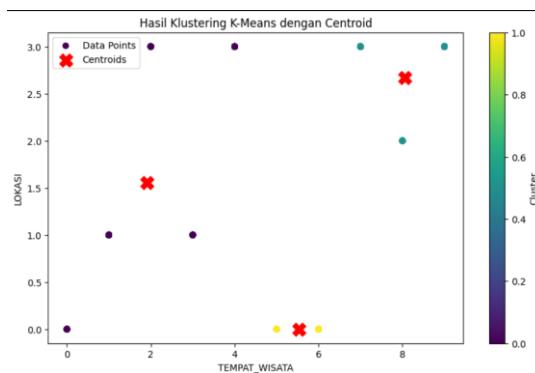
*Cluster 1 (Menengah):*

*Cluster* ini mencakup tempat wisata yang sangat populer dengan nilai centroid tertinggi untuk variabel 'Tempat Wisata'. Namun, dari segi lokasi, tempat-tempat dalam *cluster* ini cenderung kurang strategis dibandingkan *Cluster 0*, sehingga memerlukan usaha lebih untuk diakses.

*Cluster 2 (Terendah):*

*Cluster* ini menunjukkan tempat wisata yang berada di tingkat menengah dari segi popularitas. Namun, nilai centroid lokasi yang paling tinggi menunjukkan bahwa tempat wisata dalam cluster ini berada di lokasi yang paling kurang strategis dan sulit diakses, sehingga dianggap sebagai *cluster* dengan performa terendah.

Berikut adalah hasil visual grafik dari hasil *Clustering* K-Means dengan centroid berdasarkan tempat wisata dan lokasi wisata pada kalimantan.



Gambar 5. Grafik Centroid

### 3. Hasil

Berikut adalah hasil pengelompokan tempat wisata di Kalimantan berdasarkan tingkat daya tariknya. Data ini dikelompokkan ke dalam beberapa *cluster* untuk memberikan gambaran lebih jelas mengenai popularitas dan karakteristik masing-masing destinasi wisata. Pengelompokan ini dapat membantu dalam menyusun strategi pengembangan dan promosi pariwisata di wilayah tersebut.

**Tabel 3. Cluster 0 (Tertinggi)**

No	Tempat Wisata	Kategori	Provinsi	Cluster
1.	Pantai Lamaru	Taman Hiburan	Kalimantan Timur	0
2.	Pasar Terapung Muara Kuin	Budaya Nusantara	Kalimantan Selatan	0
3.	Pulau Derawan	Cagar Alam	Kalimantan Timur	0
4.	Bukit Kelam	Cagar Alam	Kalimantan Barat	0
5.	Museum Wasaka	Budaya Nusantara	Kalimantan Selatan	0

**Tabel 4. Cluster 1 (Menengah)**

No	Tempat Wisata	Kategori	Provinsi	Cluster
1.	Taman Nasional Kutai	Taman Nasional	Kalimantan Timur	1
2.	Taman Nasional Tanjung Puting	Taman Nasional	Kalimantan Tengah	1
3.	Taman Wisata alam Sungai Wain	Cagar Alam	Kalimantan Timur	1

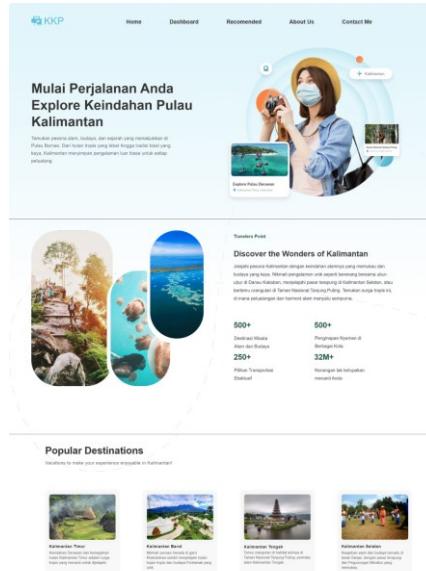
**Tabel 5. Cluster 2 (Terendah)**

No	Tempat Wisata	Kategori	Provinsi	Cluster
1.	Taman Nasional Betung Kerihun	Taman Nasional	Kalimantan Barat	2
2.	Taman Nasional Danau Sentarum	Taman Nasional	Kalimantan Barat	2

Berdasarkan hasil *Clustering*, data tempat wisata dikelompokkan ke dalam tiga *cluster* berdasarkan kategori, lokasi, dan karakteristiknya. *Cluster 0* (tertinggi) mencakup tempat wisata populer seperti Pantai Lamaru dan Pulau Derawan, yang memiliki daya tarik budaya dan alam yang kuat. *Cluster 1* (menengah) terdiri dari taman nasional seperti Taman Nasional Kutai dan Tanjung Puting, yang merupakan area konservasi alam. *Cluster 2* (terendah) mencakup tempat wisata seperti Taman Nasional Betung Kerihun, yang memiliki tingkat kunjungan relatif lebih rendah. Pengelompokan ini membantu memahami potensi daya tarik dan prioritas pengembangan wisata di setiap *cluster*.

#### 4. Tampilan

Selama proses analisis ini, kami mengembangkan antarmuka yang sederhana dan mudah digunakan untuk memahami preferensi wisatawan Kalimantan. Tampilan antarmuka website wisata dirancang untuk menarik perhatian pengguna dengan visual yang estetis dan informatif. Gambar destinasi wisata ditampilkan secara menonjol untuk memberikan gambaran jelas tentang daya tarik lokasi.



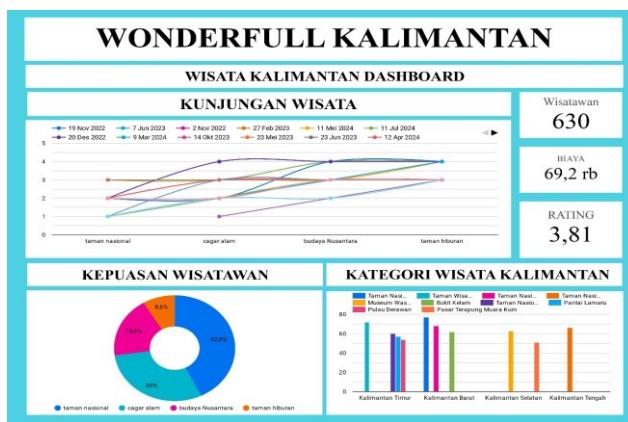
Gambar 6. Tampilan UI

Gambar 6 menunjukkan tampilan ini, yang memungkinkan pengguna untuk mengeksplorasi dan menganalisis data tentang daya tarik, fasilitas, dan pengalaman wisata yang diinginkan pengunjung.



Gambar 7. Halaman Penjelasan

Gambar 7 menunjukkan halaman yang didedikasikan untuk menjelaskan keindahan alam dan keragaman budaya Pulau Kalimantan. Halaman ini menyajikan informasi mendalam tentang warisan budaya yang kaya, tradisi unik, dan pesona alam setiap provinsi Kalimantan. Tujuan halaman ini adalah untuk mendukung promosi pariwisata yang lebih baik dan menginspirasi wisatawan untuk melihat kekayaan Pulau Kalimantan melalui visualisasi yang menarik dan informasi yang mudah dipahami.



Gambar 8. Dashboard data

Data pariwisata untuk wilayah Kalimantan ditunjukkan pada gambar ini. Jumlah pengunjung ke taman nasional, cagar alam, lokasi budaya, dan taman hiburan ditunjukkan pada grafik batang. Wilayah ini menerima 630 pengunjung, dengan nilai rata-rata 3,81 dan biaya rata-rata 69,2 ribu rupiah per orang.

Persentase pengunjung yang merasa puas dengan setiap jenis objek wisata ditunjukkan pada diagram lingkaran; 42,8% dari mereka menilai taman nasional sebagai objek wisata terbaik mereka. Masing-masing dari empat wilayah Kalimantan menunjukkan tempat wisata yang paling populer di grafik terakhir.

## SIMPULAN

Dalam penelitian ini, Metode *Clustering K-Means* digunakan untuk membagi objek wisata di Kalimantan menjadi beberapa klaster berdasarkan daya tarik, fasilitas, dan pengalaman yang ditawarkan. Hasil analisis menunjukkan bahwa strategi pengembangan yang lebih fokus dapat diterapkan pada destinasi di setiap provinsi Kalimantan, seperti Kalimantan Timur, Kalimantan Barat, Kalimantan Selatan, dan Kalimantan Tengah.

Menurut penelitian ini, destinasi di Kalimantan memiliki 3 *cluster* yaitu *cluster 0 (Cluster Tertinggi)* dengan 5 destinasi wisata, diikuti dengan *cluster 1 (Cluster Menengah)* terdapat 3 destinasi wisata, dan *cluster 2 (Cluster Terendah)* dengan 2 destinasi wisata. Studi ini memberikan informasi tentang cara meningkatkan strategi pemasaran dan layanan pariwisata untuk lebih sesuai dengan preferensi wisatawan. Pengembangan destinasi baru sebaiknya meningkatkan fasilitas, menyesuaikan kampanye promosi dengan karakteristik *cluster*, dan mengeksplorasi area yang memiliki potensi namun kurang dikenal.

## DAFTAR PUSTAKA

- [1] D. Pramayuda, M. R. Lubis, I. S. Saragih, and S. Tunas Bangsa, "Pengelompokan Tamu Asing Ke Indonesia Berdasarkan Provinsi Dengan Algoritma K-Means," Oct. 2020.
- [2] Y. W. Syaifudin and R. A. Irawan, "IMPLEMENTASI ANALISIS CLUSTERING DAN SENTIMENT DATA TWITTER PADA OPINI WISATA PANTAI MENGGUNAKAN METODE K-MEANS," 2018.
- [3] E. Satria *et al.*, "Prosiding Seminar Nasional Riset Information Science (SENARIS) Penerapan Clustering dalam Mengelompokkan Jumlah Kunjungan Wisatawan Mancanegara Dengan Metode K-Means," p. 462, 2019, [Online]. Available: <https://www.bps.go.id/dynamictable/2018/07/30/1548/jumlah-kunjungan->
- [4] Arinaldi Vinsa, Pramuditya Soesanto R, and Dwi Anggana H, "Perancangan Sistem Rekomendasi Paket Wisata Alam Kabupaten Rembang Jawa Tengah Dengan Metode K-Means Design of a Recommendation System for Nature Tourism Packages in Rembang Regency, Central Java with the K-Means Method," Jun. 2023.
- [5] A. Utama, W. Imama Sabilla, R. Wakhidah, P. Studi Teknik Informatika, P. Negeri Malang, and J. Soekarno Hatta No, "Sistem Rekomendasi Tempat Wisata di Malang Raya Menggunakan Metode K-Means Clustering," vol. 16, no. 1, 2024.

- [6] Sitorus Michael, Fitron Depriansa, and Agung Segara Wisesa C, “4662-Article Text-18159-1-10-20230215,” 2022.
- [7] S. Astiti and R. Harman, “Pengelompokan Destinasi Wisata di Batam Berdasarkan Daya Tarik dan Fasilitas Menggunakan Metode K-Means Clustering,” Oct. 2024.
- [8] 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.