Data and Metadata. 2025; 4:1206
doi: 10.56294/dm20251206

ORIGINAL

Geospatial Clustering of Potential Tourist Locations Using the K-Means Algorithm: A Case Study of Unesco Global Geopark (Sukabumi, Indonesia)

Clustering Geoespacial de Ubicaciones Turísticas Potenciales Usando el Algoritmo K-Means: Un Estudio de Caso del Geoparque Global de la UNESCO (Sukabumi, Indonesia)

Nunik Destria Arianti[1], Rahmat Hidayat[1], Adhitia Erfina[1], Muhamad Muslih[1], Riyan Mirdan Faris[2]

[1]Nusa Putra University. Information Systems. Sukabumi, Indonesia.
[2]Nusa Putra University, Management. Sukabumi, Indonesia.

Cite as: Destria Arianti N, Hidayat R, Erfina A, Muslih M, Mirdan Faris R. Geospatial Clustering of Potential Tourist Locations Using the K-Means Algorithm: A Case Study of Unesco Global Geopark (Sukabumi, Indonesia). Data and Metadata. 2025; 4:1206. https://doi.org/10.56294/dm20251206

Submitted: 14-04-2025 **Revised:** 12-07-2025 **Accepted:** 15-10-2025 **Published:** 16-10-2025

Editor: Dr. Adrián Alejandro Vitón Castillo

Corresponding Author: Nunik Destria Arianti

ABSTRACT

Sukabumi Regency, located in southern West Java, Indonesia, is home to abundant natural tourism resources. However, many of these sites remain underutilized due to limited infrastructure, insufficient promotion, and the lack of data-driven planning. Tourism is essential for local economic development, making the identification of high-potential areas critical for growth strategies. This study applied the CRISP-DM methodology to classify tourism destinations in Sukabumi using data from 17 tourist sites. Variables such as average visitor numbers, ticket prices, monthly growth rates, and area usage percentages were analyzed. After data preprocessing and normalization, the K-Means algorithm was employed for clustering. The Elbow Method determined the optimal number of clusters, and cluster quality was assessed using the Silhouette Score and Davies-Bouldin Index. Three distinct clusters were identified: developed, developing, and emerging tourism sites. The Silhouette Score of 0,226 and Davies-Bouldin Index of 1,323 indicated moderate cluster cohesion and separation. A thematic map visualized the spatial patterns, showing clear geographical distinctions between clusters. Cluster 0 (red) represented low-performing destinations, Cluster 1 (green) included high-traffic, developed sites, and Cluster 2 (blue) contained mid-level destinations with growth potential. The results provide valuable insights for regional tourism development. The study offers a data-driven foundation for targeted promotion, resource allocation, and sustainable planning. The findings demonstrate that geospatial clustering can effectively support tourism strategies tailored to local needs, contributing to inclusive economic growth in Sukabumi Regency.

Keywords: Tourism Clustering; Potential Mapping; Spatial Analysis; Economic Development; Machine Learning.

RESUMEN

El distrito de Sukabumi, ubicado en el sur de Java Occidental, Indonesia, alberga abundantes recursos naturales para el turismo. Sin embargo, muchos de estos sitios siguen siendo subutilizados debido a la infraestructura limitada, la promoción insuficiente y la falta de planificación basada en datos. El turismo es esencial para el desarrollo económico local, lo que hace que la identificación de áreas de alto potencial sea crucial para las estrategias de crecimiento. Este estudio aplicó la metodología CRISP-DM para clasificar los destinos turísticos en Sukabumi utilizando datos de 17 sitios turísticos. Se analizaron variables como el número promedio de visitantes, los precios de las entradas, las tasas de crecimiento mensual y los porcentajes de uso del área.

Después del preprocesamiento y normalización de los datos, se empleó el algoritmo K-Means para la agrupación. El Método del Codo determinó el número óptimo de clústeres, y la calidad de los clústeres se evaluó mediante el Silhouette Score y el Davies-Bouldin Index. Se identificaron tres clústeres distintos: sitios turísticos desarrollados, en desarrollo y emergentes. El Silhouette Score de 0,226 y el Davies-Bouldin Index de 1,323 indicaron una cohesión y separación moderadas de los clústeres. Un mapa temático visualizó los patrones espaciales, mostrando distinciones geográficas claras entre los clústeres. El Clúster 0 (rojo) representó destinos de bajo rendimiento, el Clúster 1 (verde) incluyó sitios desarrollados con alta afluencia de turistas, y el Clúster 2 (azul) contenía destinos de nivel medio con potencial de crecimiento. Los resultados proporcionan valiosos conocimientos para el desarrollo del turismo regional. El estudio ofrece una base basada en datos para la promoción dirigida, la asignación de recursos y la planificación sostenible. Los hallazgos demuestran que el agrupamiento geoespacial puede respaldar eficazmente estrategias turísticas adaptadas a las necesidades locales, contribuyendo al crecimiento económico inclusivo en el distrito de Sukabumi.

Palabras clave: Agrupamiento Turístico; Cartografía del Potencial; Análisis Espacial; Desarrollo Económico; Aprendizaje Automático.

INTRODUCTION

Tourism is an important sector in regional economic development, especially in developing countries, as it can create job opportunities, increase community income, and encourage the development of supporting infrastructure.[1] In the context of sustainable development, tourism also serves as a driver of economic growth while enhancing the competitiveness of a region.[2] Sukabumi, one of the regencies in West Java Province, has diverse tourism potential, ranging from natural attractions, culture, to culinary tourism. The presence of the UNESCO Global Geopark Ciletuh-Palabuhanratu is one of the main attractions that can boost the number of tourist visits. However, tourism development in this region is still not optimal. Several challenges faced include inadequate infrastructure, limited digital promotion, and the absence of a data-driven decision support system.[2] This condition makes tourism development strategies tend to be based on general assumptions, resulting in ineffective resource allocation and promotion.[3]

Historically, efforts to develop tourism in Sukabumi have been carried out gradually, but have not been accompanied by a systematic method to map and classify tourist destinations.[4] This has led to difficulties in determining development and promotion priorities. One approach that can be used to address this issue is clustering, which is a data grouping technique based on specific characteristics such as visit frequency, availability of facilities, and spatial distribution.[5] Through clustering, local governments and stakeholders can develop more targeted development strategies and support effective tourism promotion.[6]

This study applies the K-Means algorithm to analyze and map the tourism potential in Sukabumi. This algorithm has proven effective in identifying patterns and groups of destinations based on popularity and spatial characteristics.[7,8] The novelty of this research lies in the integration of data mining techniques and geospatial analysis to produce a decision support model that can be replicated in other regions.[9,10]
Based on this, the study aims to develop a data-based tourism potential mapping model that can support more efficient development planning and promotion strategies. This model is expected to help local governments and stakeholders optimize resource management and enhance the contribution of the tourism sector to economic growth and regional competitiveness.[10]

Related works

Tourism has a strategic contribution to regional economic growth, especially in developing countries.[11] In addition to creating job opportunities and increasing community income, this sector also accelerates infrastructure development and encourages local economic diversification.[12,13,14,15] In the context of Indonesia, the application of data-driven strategies has proven to enhance the effectiveness of decision-making and policy efficiency in the tourism sector.[16,17] This indicates that proper tourism management requires an evidence-based approach that can map potential and guide policies that are responsive to market dynamics.[13]

One approach commonly used in regional planning is clustering, which allows for the grouping of tourist destinations based on geographic, economic, and socio-cultural characteristics. Muqorobin et al.[16] and Salsabila et al.[18] demonstrated that this method is effective in identifying visit patterns, availability of facilities, and supporting infrastructure to create a strategic tourism development map(. These findings align with Gura et al.[15], who emphasized the role of clustering in integrating the concept of the circular economy into regional planning, ensuring that destination development is sustainable and based on local potential.[15]

The integration of clustering methods with Geographic Information Systems (GIS) further enhances the

accuracy of spatial analysis and the visualization of regional potential. Research by Adya Hermawat has proven that spatial-cluster analysis can improve the effectiveness of destination promotion strategies. Meanwhile, a studies emphasize the importance of utilizing digital technologies and big data, such as thematic dashboards and co-word mapping, to monitor tourism trends and enhance location-based experiences.[19] This integration of technology not only facilitates analysis but also enables the creation of more adaptive and targeted policies.

In managing complex tourism data, the CRISP-DM (Cross Industry Standard Process for Data Mining) framework provides a systematic approach from business understanding to implementation. A studie demonstrate that CRISP-DM can help organize destination data, resulting in an adaptive and responsive decision support system that can adjust to changes in tourism conditions.[17,18] This is highly relevant in the context of tourism management in Sukabumi, which requires a data-driven approach and structured analysis.

Policy aspects also play a crucial role in tourism development. A study emphasize the importance of multi-actor collaboration in designing cluster-based tourism strategies that connect governments, businesses, and local communities.[20] This approach supports the concept of creative tourism where tourist preferences are integrated with local arts and cultural potential.[21] Thus, tourism development is not only economically oriented but also focuses on the preservation of cultural identity.

Clustering is also used across sectors, demonstrating the flexibility of this method in various contexts. A study applied cluster analysis to evaluate integrated rural sanitation in China and successfully identified regional segmentation efficiently.[22] This strengthens the validity of the clustering method for spatial planning, including in the tourism sector, which involves diverse data and high complexity.

The COVID-19 pandemic brought significant changes to the dynamics of global tourism. A study emphasize the importance of smart tourism approaches and sustainable clustering in building a resilient and crisis-responsive tourism system.[23] This transformation strengthens the urgency of utilizing digital technology in the promotion and management of destinations. A study found a significant increase in the use of computers (61,56 %) and social media (77,92 %) as the primary means of searching for destination information during the pandemic in the Pichincha Province, Ecuador.[24] This behavioral shift indicates that digital-based promotion strategies are key to reaching modern travelers who increasingly rely on technology.

Based on the study, it is clear that the integration of clustering, GIS, and digital technology holds great potential in the development of sustainable and evidence-based tourism. This forms a crucial foundation for this research to develop a tourism potential mapping model in Sukabumi that can support strategic planning, digital promotion, and more effective decision-making.

METHOD

Type of Study

This study is a non-experimental, observational, descriptive–analytical research with a cross-sectional approach. The primary goal is to map and classify tourist destinations in Sukabumi Regency based on their characteristics at a single point in time. This supports evidence-based decision-making for sustainable tourism development. The analytical framework used is CRISP-DM (Cross Industry Standard Process for Data Mining), which consists of six structured stages: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment.[25] The K-Means clustering algorithm serves as the core model for identifying groups of destinations with similar tourism potential.

Universe and Sample

Universe (Target Population)

All officially registered tourist destinations in Sukabumi Regency, including those located within the UNESCO Global Geopark Ciletuh-Palabuhanratu area.

Sampling Frame

The registry and administrative data provided by the Tourism Office of Sukabumi Regency.

Inclusion Criteria

Destinations with complete records of:

- Average monthly visitors.
- Entrance ticket price (HTM/ETP).
- Monthly growth rate of visits (%).
- Total area and utilized area for tourism.

Accurate geographic coordinates for spatial mapping.

Exclusion Criteria

Destinations with incomplete, inconsistent, or duplicate data that could not be corrected during cleaning.

Final Sample Size

17 tourist destinations met the criteria and were included in the analysis. These destinations represent various types, such as natural attractions, cultural sites, and adventure tourism areas.

Variables

Primary Variables

Table 1. Primary and Derived Variables Used in the Study

Variable	Description
Average Monthly Visitors	Total annual visitors divided by 12 months
Entrance Ticket Price (HTM)	Standard admission fee per visitor (IDR)
Monthly Visit Growth Rate (%)	Change in visitors between initial and final month of observation
Total Area (ha)	Physical size of the destination
Utilized Area (ha)	Area actively used for tourism activities

Derived Variables

Table 2. Research Variables (Primary and Derived)

Variable	Formula
Percentage of Area Utilized (%)	(Utilized Area / Total Area) × 100
Visitor Growth Rate (%)	((Final month – Initial month) / Initial month) × 100

Data Collection and Processing

Data Collection

The dataset was obtained from the Tourism Office of Sukabumi Regency and consisted of:

Administrative Data

Monthly visitor counts, entrance ticket prices, and destination area information.

Spatial Data

Geographic coordinates and administrative boundaries for mapping and spatial clustering.

Expert Input

Semi-structured interviews with the Head of the Tourism Office to identify gaps in destination management and validate results. Figure 1 shows the geographic layout of Sukabumi Regency and spatial distribution of tourist destinations. This map is essential for understanding accessibility and clustering patterns.[12,13,14]

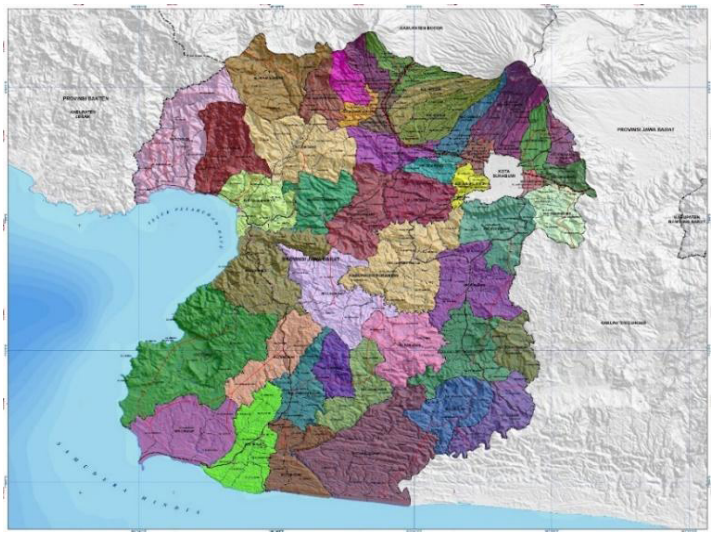

Figure 1. Map of Sukabumi Regency Area

Figure 2 displays the distribution of destinations by type, which guided the cluster modeling process. Beaches and waterfalls dominate, with 19 and 14 locations respectively, followed by estuaries (6 locations). Other types, such as caves, geoparks, cultural heritage sites, and plantations, are fewer but still significant.[5]

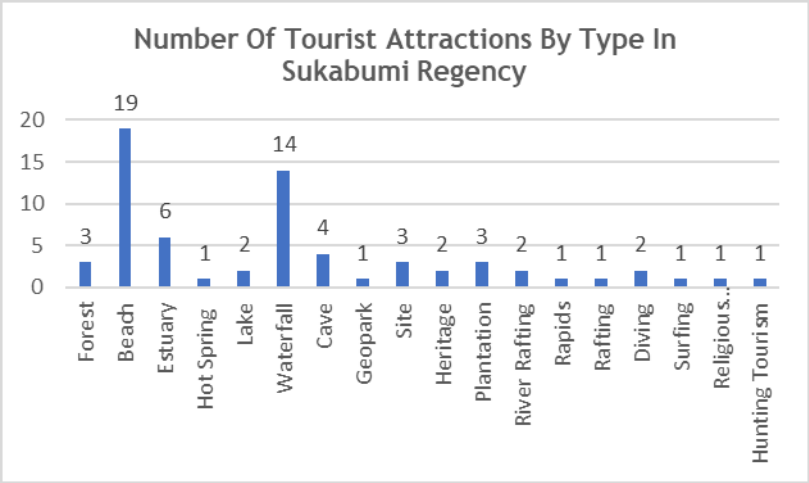

Figure 2. Number of Tourist Attractions by Type in Sukabumi Regency

Data Processing

The CRISP-DM methodology was applied step-by-step as follows:

Business Understanding

Identification of challenges, such as:

- Weak digital promotion and limited exposure of destinations.
- Absence of integrated data-driven systems.
- Ineffective resource allocation for tourism development.

The objective was to design a clustering model to optimize tourism planning and promotional strategies.

Data Understanding

Exploration of raw data to ensure completeness, detect missing values, and verify logical consistency (e.g., utilized area cannot exceed total area). Descriptive statistics were generated to understand variable distributions.

Data Preparation

- Cleaning: removal of duplicate records and correction of inconsistencies.
- Normalization: min-max scaling applied to standardize variables before clustering.
- Derivation: calculation of percentage utilization and growth rates using formulas provided earlier.
- Final Dataset: 17 destinations × 5 variables ready for modeling.

Modeling

- Algorithm.
- Cluster Determination.
- Cluster Results.

Table 3. Cluster Interpretation

Cluster	Characteristics
Advanced Cluster	High visitation, high growth rate, premium ticket pricing
Developing Cluster	Moderate visitation, potential for improvement
Pioneering Cluster	Low visitation, limited facilities and growth

Evaluation

The clustering model achieved:

- Silhouette Score.
- Davies-Bouldin Index (DBI).

These metrics indicate moderate separation between clusters. Validation was performed through:

1. Internal performance metrics.
2. Stakeholder discussions with tourism officials to align results with regional planning needs.

Deployment

Final outputs were visualized through thematic maps to assist policymakers and stakeholders in identifying:

- Priority development zones.
- Strategic promotional targets.
- Infrastructure investment areas.

Ethical Standards

Data Privacy

Only aggregated, non-identifiable administrative data were used.

Permissions

Written approval was obtained from the Tourism Office of Sukabumi Regency to access datasets.

Informed Consent

Verbal consent was obtained for interviews with tourism officials.

Data Handling

Data were securely stored on password-protected devices and used exclusively for academic purposes.

Compliance

This study complies with institutional ethical guidelines for research using secondary, non-identifiable data. If required, an exemption letter can be provided to the journal.

RESULTS

Overview of Tourism Data

A total of 17 tourist destinations met the inclusion criteria for this study. These destinations represent diverse categories, including beaches, waterfalls, geoparks, cultural sites, and adventure tourism areas. The core descriptive statistics—covering average monthly visits, entrance ticket prices, visit growth rates, and area utilization percentages—are summarized in table 4 and table 5. These datasets served as the foundation for clustering analysis.

Table 4. Average Visits to Tourist Destinations in Sukabumi Regency

No	Tourist Attraction Type	Number of Destinations	Average Visits
1	Beach Tourism	19	4758,5
2	Waterfalls (Curug)	14	4658
3	Caves	4	5923
4	Forest Tourism	3	2933
5	Geopark	3	5890
6	Hot Springs	2	8585,5
7	Whitewater Rafting	2	6078
8	Wave Rafting	2	1684
9	Religious Tourism	1	5811
10	Marine Tourism (Diving)	1	6896
11	Hunting / Outdoor Tourism	1	9774
12	Tea/Flower Plantation	2	2577
13	Historical Sites	2	5800
14	Recreation Area	1	4500
15	Estuary	1	3000
16	Lake / Reservoir	2	6234
17	Surfing	1	9166

Table 5. Entrance Ticket Prices of Tourist Destinations

No.	Tourist Attraction Type	Entrance Ticket Price (IDR)
1	Beach Tourism	10 000
2	Waterfalls (Curug)	7000
3	Caves	10 000
4	Forest Tourism	5000
5	Geopark	5000
6	Hot Springs	5000
7	Whitewater Rafting	10 000
8	Wave Rafting	7000
9	Religious Tourism	5000
10	Marine Tourism (Diving)	7500
11	Hunting / Outdoor Tourism	5000
12	Tea / Flower Plantation	10 000
13	Historical Sites	15 000
14	Recreation Area	5000
15	Estuary	10 000
16	Lake / Reservoir	7000
17	Surfing	10 000

Data Preparation

The four attributes used for clustering average monthly visits, entrance ticket prices, visit growth rates, and area utilization were standardized to ensure uniformity in the clustering process.

The final dataset prepared for analysis is shown in table 6.

Table 6. Final Dataset for Clustering Analysis

No	Tourist Attraction Type	Average Visits	Entrance Ticket Price (IDR)	Average Visits	Visit Growth (%)	Area Size (%)
1	Beach Tourism	4758,5	10 000	7 770	8,06	56,04
2	Waterfalls (Curug)	4658	7000	1360	1,68	67,07
3	Caves	5923	10 000	5890	6,89	60,79
4	Forest Tourism	2933	5000	5691	7,86	69,02
5	Geopark	5890	5000	6234	5,74	66,06
6	Hot Springs	8585,5	5000	6765	3,41	78,31
7	Whitewater Rafting	6078	10 000	966	3,18	61,30
8	Wave Rafting	1684	7000	4926	7,96	84,52
9	Religious Tourism	5811	5000	6078	2,94	49,15
10	Marine Tourism (Diving)	6896	7500	8822	2,25	35,43
11	Hunting / Outdoor Tourism	9774	5000	2185	7,33	48,04
12	Tea / Flower Plantation	2577	10 000	1269	9,60	36,84
13	Historical Sites	5800	15 000	7449	1,53	79,72
14	Recreation Area	4500	5000	2933	5,85	32,81
15	Estuary	3000	10 000	5811	8,41	67,58
16	Lake / Reservoir	6234	7000	5551	6,71	62,86
17	Surfing	9166	10 000	6 920	7,63	79,16

Determination of Optimal Clusters

The Elbow Method was applied to determine the most efficient number of clusters. As shown in figure 3, the inflection point was observed at K = 3, indicating that three clusters were optimal for this analysis.

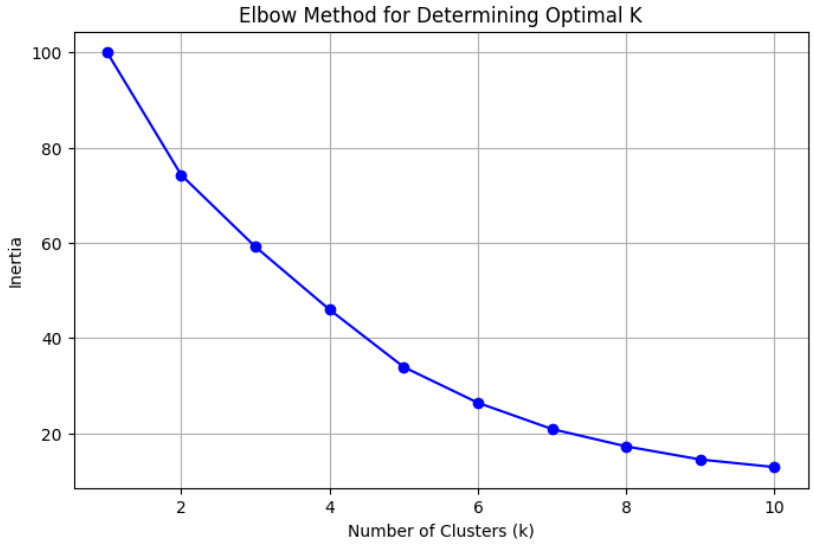

Figure 3. Elbow Method for Determining Optimal Number of Clusters

Internal validation confirmed this decision, with a Silhouette Score of 0,226 and Davies-Bouldin Index (DBI) of 1,323, indicating moderate cluster separation.

Clustering Results

The K-Means algorithm grouped the 17 destinations into three distinct clusters:

- Cluster 0: lower visits and limited facilities.
- Cluster 1: high visitation and premium services.
- Cluster 2: moderate characteristics and growth potential.

Cluster assignments for each destination are presented in table 7.

Table 7. Cluster Assignment of Tourist Destinations

Index	Tourist Spot Name	Cluster
0	Pelabuhan Ratu Beach	1
1	Sawer Waterfall	2
2	Ciletuh Geopark	1
3	Ujung Genteng Beach	0
4	Situ Gunung	0
5	Cimarinjung Waterfall	1
6	Cibangban Beach	2
7	Buniayu Cave	1
8	Cikaso Rafting	2
9	Cikundul Hot Springs	2
10	Karang Hawu Beach	2
11	Goalpara Tea Plantation	2
12	Sodong Waterfall	0
13	Cikaso Forest	2
14	Gunung Tilu Religious Tourism	0
15	Cikembang Beach	1
16	Lalay Cave	0
17	Cimaja Wave Rafting	2

The spatial distribution of these clusters is illustrated in figure 4, showing geographic segmentation across Sukabumi Regency.

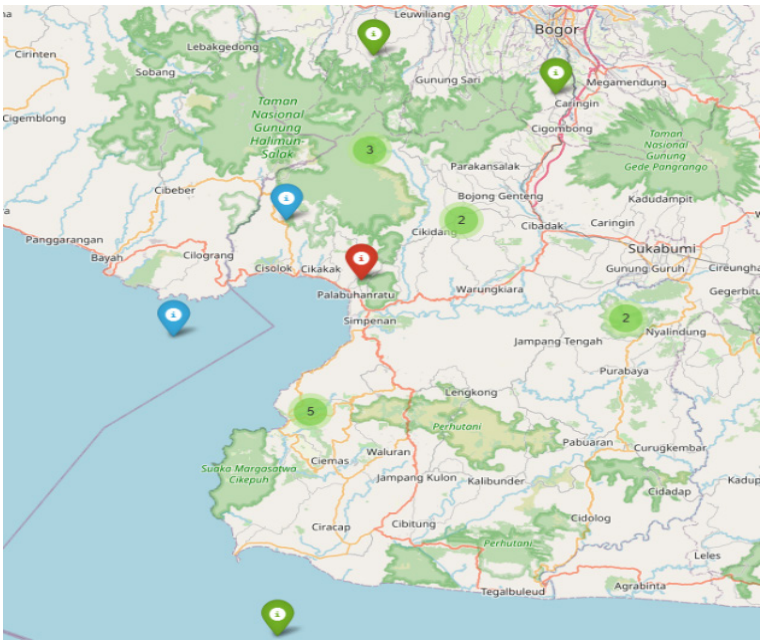

Figure 4. Spatial Distribution of Tourism Clusters

Visualization of Cluster Patterns

The scatter plot in figure 5 provides a visual representation of cluster separation using average visits and ticket prices. Each color represents one cluster, demonstrating distinct groupings.

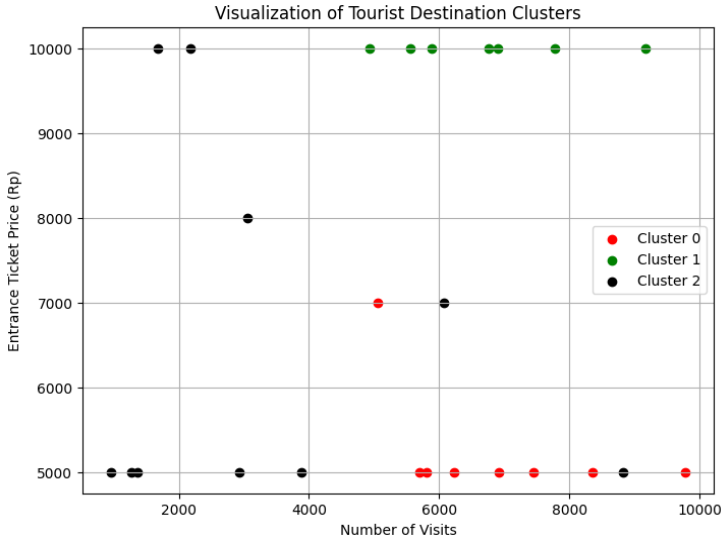

Figure 5. Scatter Plot of Clustering Results

Model Evaluation

The quality of the clustering model was evaluated using internal metrics. As shown in figure 6, the Silhouette Score and DBI confirm that the three-cluster model is a balanced representation of the dataset.

```
Centroid (standar):
[[ 0.66065913  0.59952369  0.22777046 -0.79091194]
 [ 0.39591087  0.52071787  0.09358436  1.26200156]
 [-0.80566491 -0.84412146 -0.24772542 -0.25067154]]
Silhouette Score: 0.226
Davies-Bouldin Index (DBI): 1.323
```

Figure 6. Clustering Evaluation Metrics

Deployment and Thematic Map

The final clustering results were visualized on a thematic map to support decision-making for tourism development. Figure 7 shows the spatial distribution of clusters, providing a clear visual tool for stakeholders.

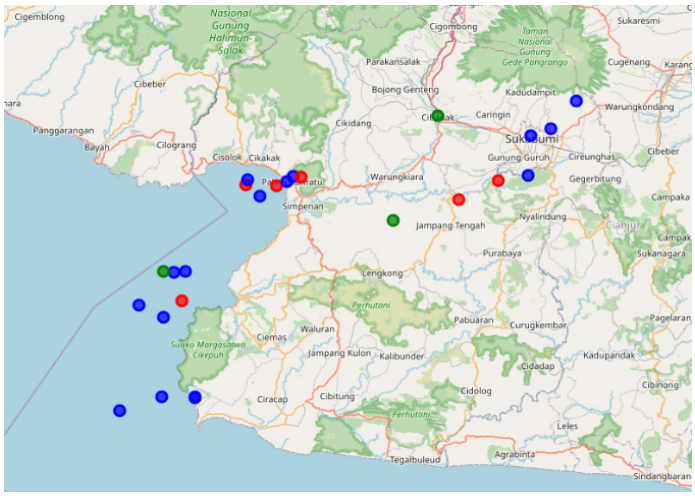

Figure 7. Thematic Map of Tourism Clusters in Sukabumi Regency

DISCUSSION

This study demonstrates the use of the K-Means algorithm in clustering tourist destinations in Sukabumi Regency with the aim of improving resource allocation and more effective tourism development. The results confirm that K-Means is capable of separating tourist destinations into three clusters based on criteria such as the number of visits, ticket prices, and destination area size. The advantage of K-Means lies in its ability to handle data with non-linear distributions and class imbalances, which is highly relevant in the tourism context, where low-performing destinations require more attention in terms of promotion and facility development. Additionally, handling missing data at 38,1 % with median and mode imputation, as well as using SMOTE to address class imbalance, demonstrates the reliability of this algorithm in facing real-world data challenges. The feature selection process through PCA and RFE successfully identified 12 important features directly impacting destination performance, such as service system type and age group, which can be the focus of tourism policies. The practical implications of these findings are crucial for tourism planning policies, as identifying the right destinations enables more targeted resource allocation and workforce planning, especially for underdeveloped destinations. However, the results also reveal limitations in traditional models like Logistic Regression and SVM, which showed lower performance in handling data with class imbalances and non-linear complexities. Future research could explore how this classification can be integrated into a broader tourism system, including its impact on the tourist experience and its application in real-time decision support systems.

CONCLUSIONS

This study successfully demonstrates the potential application of the K-Means algorithm in clustering tourist destinations in Sukabumi Regency to support more targeted tourism planning and development. The primary objective of this study was to identify clusters of destinations based on characteristics such as the number of visits, ticket prices, and area size, providing clearer insights into destinations that require more attention in development and resource allocation. By utilizing clustering techniques, the findings offer a strong foundation for data-driven tourism policies, helping decision-makers to plan interventions more efficiently at underdeveloped destinations. Although this study provides a comprehensive overview, the limitations of traditional models indicate the need for a more dynamic approach in tourism data analysis. Therefore, further research is needed to explore the application of clustering techniques in a broader context, as well as to examine their impact on the tourist experience and human resource planning in the tourism sector.

BIBLIOGRAPHIC REFERENCES

1. Merylova I. Revitalizing industrial territories in Ukraine: Spatial model of the tourism cluster in the Prydniprovsk region. Front Archit Res. 2024;13(6):1342–62. https://www.sciencedirect.com/science/article/pii/S2095263524000761

2. Mutiara S, Auliasari K, Zahro HZ. Penerapan Metode K-Means Untuk Pengelompokan Populasi Ternak Di Kabupaten Sumbawa Provinsi Ntb. JIKA (Jurnal Inform. 2025;9(1):118–25.

3. Fundeanu DD. Innovative Regional Cluster, Model of Tourism Development. Procedia Econ Financ. 2015;23:744–9. https://www.sciencedirect.com/science/article/pii/S2212567115005018.

4. Bittencourt JCN, Costa DG, Portugal P, Peixoto MLM, Vasques F. On the spatiotemporal knowledge-driven vulnerability assessment of urban areas: A clustering-based approach. Int J Disaster Risk Reduct. 2025;127:105681. https://doi.org/10.1016/j.ijdrr.2025.105681.

5. Ajustina F, Nisa LF. Implementasi Prinsip Ekonomi Syariah Terhadap Pengembangan Ekonomi Umat Untuk Meningkatkan Pertumbuhan Ekonomi. J Ilm Ekon dan Manaj. 2024;2(6):626–37. https://doi.org/10.61722/jiem.v2i6.1575.

6. Seftia M, Efan E, Arif A, Teknologi I, Alam P. Optimalisasi Strategi Wisata Di Kota Pagar Alam Menggunakan Algoritma K-Means Clustering. Jurnal Informatika (JURI) e-ISSN. 2024;01.

7. Jauhari A, Anamisa DR, Mufarroha FA. Analysis of Clusters Number Effect Based on K-Means Method for Tourist Attractions Segmentation. J Phys Conf Ser. 2022;2406(1).

8. Rahmiyati D, Nuraeni, Siswantoro EB. Penerapan Metode K-Means untuk Pemetaan Objek Wisata sebagai Rekomendasi Prioritas Pengembangan Pariwisata. J Ilmu Komput dan Teknol. 2025;6(1):17–24.

9. Ramadanti E, Muslih M, Arianti ND. Design of Web-Based Virtual Tourism Information System At Geopark Ciletuh Sukabumi. J Ris Inform. 2022;4(3):259–68.

10. Arsezen-Otamisa P, Yuzbasioglu N. Analysis of Antalya Tourism Cluster Perceived Performance with Structural Equation Model. Procedia - Soc Behav Sci. 2013;99:682–90. https://www.sciencedirect.com/science/article/pii/S1877042813039840.

11. Sabolak Y, Zebua F. Analisis Keberlanjutan Dan Peran Csr Dalam Pengembangan Ekonomi Lokal. J Ilmu Ekon dan Bisnis. 2024;1(1):20–6.

12. Xu J, Wang P. Study on distribution characteristic of tourism attractions in international cultural tourism demonstration region in South Anhui in China. PLoS One. 2022;17(6):1–16. http://dx.doi.org/10.1371/journal.pone.0269948

13. Utama A, Sabilla WI, Wakhidah R. Sistem Rekomendasi Tempat Wisata di Malang Raya Menggunakan Metode K-Means Clustering. J Inform dan Multimed. 2024;16(1):1–15.

14. Wiangkham A, Kieanwatana K, Vongvit R. Journey into virtual reality: Identifying behavioral intentions to use virtual reality in tourism through spectral clustering. J Open Innov Technol Mark Complex. 2025;11(1):100442. https://www.sciencedirect.com/science/article/pii/S2199853124002361.

15. Gura KS, Nica E, Kliestik T, Puime-Guillén F. Circular economy in territorial planning strategy: Incorporation in cluster activities and economic zones. Environ Technol Innov. 2023;32.

16. Muqorobin, Wahyuningsih DH, Supatminingsih S. Pengembangan Metode Klasterisasi Untuk Pengelompokan Potensi Wisata guna Menunjang Digitalisasi Wisata di Indonesia. JST (Jurnal Sains dan Teknol. 2024;13(2):217–30.

17. Ermawati A, Panunjul NM, Sopanah A, Suhermin. Penerapan Cluster Analysis untuk Pemilihan Prioritas Variabel Penentu Kinerja UKM Industri Pariwisata di Jawa Timur. SENIMA (Seminar Manajemen). 2024;9(1):347–62. https://proceeding.unesa.ac.id/index.php/senima/article/view/3717.

18. Nafri TC, Pinasty S, Dini SK, Nurcahayani H. Studi Klasterisasi Usaha Pertanian Perorangan di Kabupaten Bantul Tahun 2023 dengan Pendekatan Hirarki. Emerg Stat Data Sci J. 2025;3(1):462–71.

19. Parra-Sanchez DT, Viviescas-Jaimes PA. Application of digital technologies in nature-based tourism: A scientometric analysis. J Digit Econ. 2024;3:249–59.

20. Shen Y, Zhang X. Journal of Innovation. J Innov Knowl. 2023;8(3):100384. https://doi.org/10.1016/j.jik.2023.100384.

21. Liu T, Kou ITE. Determinants for the development of creative tourism: A stakeholder perspective. Heliyon. 2024;10(13).

22. Li Y, Zhao Y, Xiang X, Zhang Q, Yuan F, Wang Z, et al. Rural integrated sanitation assessment at community level in Chongqing, China—A cluster analysis. Ecol Indic. 2024;165.

23. Casado-Aranda LA, Sánchez-Fernández J, Bastidas-Manzano AB. Tourism research after the COVID-19 outbreak: Insights for more sustainable, local and smart cities. Sustain Cities Soc. 2021;73.

24. Mendoza Poma RM, Sampedro Arrieta MA, Álvarez Lema FA, Abarca Zaquinaula MA. Analysis of technological tools for tourism in the Pichincha province. Data Metadata. 2024;3.

25. Amodio ML. Effects of seasonal variation on MAP response of fresh-cut artichokes. Acta Hortic. 2024;(1386):175–6.

FINANCING

The author expresses their deepest gratitude to the Ministry of Higher Education, Science, and Technology of Indonesia for the funding support provided through the research grant scheme to a lecturer at Universitas Nusa Putra. This support has been crucial in the execution of the research, from data collection to analysis and the formulation of recommendations. The author also extends their thanks to the Department of Tourism of Sukabumi Regency for the cooperation and data contributions provided, which have greatly facilitated the smooth conduct of this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Nunik Destria Arianti, Rahmat Hidayat.
Data curation: Rahmat Hidayat.
Formal analysis: Adhitia Erfina.
Research: Rahmat Hidayat, Adhitia Erfina.
Methodology: Rahmat Hidayat, Adhitia Erfina.
Project management: Muhamad Muslih.
Resources: Nunik Destria Arianti, Rahmat Hidayat.
Software: Muhamad Muslih, Riyan Mirdan Faris.
Supervision: Nunik Destria Arianti, Muhamad Muslih.
Validation: Nunik Destria Arianti, Muhamad Muslih.
Display: Riyan Mirdan Faris.
Drafting - original draft: Nunik Destria Arianti, Rahmat Hidayat, Adhitia Erfina.
Writing - proofreading and editing: Adhitia Erfina.