MAPPING OF DOMESTIC AND FOREIGN TOURIST VISITS IN EAST JAVA USING THE DBSCAN METHOD

Marita Qori'atunnadyah[1*]

Department of Informatics[1]
Institut Teknologi dan Bisnis Widya Gama Lumajang, Lumajang, Jawa Timur, Indonesia[1]
itbwigalumajang.ac.id[1]
maritaqori@gmail.com[1*]
(*) Corresponding Author

Abstract— *Tourism is important in economic growth and regional development, especially in East Java Province with diverse tourist attractions. However, the mapping of domestic and foreign tourist visit patterns in this province is still limited. For this reason, this study uses the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) method which can group density-based data without determining the number of clusters from the beginning and handle noise. The study aims to map districts/cities in East Java based on the number of tourist visits from 2018 to 2022, using visit data from the East Java Provincial Culture and Tourism Office. The analysis results show that in domestic tourist data, with parameters MinPts = 3 and ε = 1.00, one main cluster is formed consisting of 31 tourist locations and 7 noisy locations. In foreign tourist data, with ε = 0.6 and MinPts = 3, there is one cluster with 30 tourist locations and 8 other locations are categorized as noisy. Noisy locations tend to have higher visits but do not fit into the main cluster. These findings provide important insights for more targeted tourism promotion strategies and efficient resource allocation in East Java.*

Keywords: *clustering, DBSCAN, domestic, foreign, tourism.*

Abstrak— *Pariwisata berperan penting dalam pertumbuhan ekonomi dan pengembangan wilayah, terutama di Provinsi Jawa Timur dengan beragam daya tarik wisata. Namun, pemetaan pola kunjungan wisatawan nusantara dan mancanegara di provinsi ini masih terbatas. Untuk itu, penelitian ini menggunakan metode Density-Based Spatial Clustering of Applications with Noise (DBSCAN) yang dapat mengelompokkan data berbasis kepadatan tanpa menentukan jumlah klaster dari awal serta menangani noise. Penelitian bertujuan memetakan kabupaten/kota di Jawa Timur berdasarkan jumlah kunjungan wisatawan selama 2018 hingga 2022, menggunakan data kunjungan dari Dinas Kebudayaan dan Pariwisata Provinsi Jawa Timur. Hasil analisis menunjukkan bahwa pada data wisatawan nusantara, dengan parameter MinPts = 3 dan ε = 1,00, terbentuk satu klaster utama yang terdiri dari 31 lokasi wisata dan 7 lokasi noise. Pada data wisatawan mancanegara, dengan ε = 0,6 dan MinPts = 3, terdapat satu klaster dengan 30 lokasi wisata dan 8 lokasi lainnya dikategorikan sebagai noise. Lokasi yang termasuk noise cenderung memiliki kunjungan lebih tinggi tetapi tidak cocok dalam klaster utama. Temuan ini memberikan wawasan penting bagi strategi promosi pariwisata yang lebih tepat sasaran dan alokasi sumber daya yang efisien di Jawa Timur.*

Kata Kunci*: pengelompokan, DBSCAN, domestik, asing, pariwisata.*

INTRODUCTION

Tourism has become a strategic sector that continues to be promoted by the government as one of the main drivers of economic growth, job creation, and regional development. In addition, this sector has an important role in improving community welfare and supporting sustainable regional development. Therefore, tourism development is becoming a focus that is increasingly being considered in various government policies. In addition, tourism development is often used as an indicator to measure economic stability and security in a region (Badan Pusat Statistik, 2023).

East Java Province, with its abundant cultural, historical, and natural wealth, is one of the tourist destinations that offers a variety of attractions. Each district and city in this province has unique tourism potential, making it a significant contributor to the regional economy. To optimally utilize this great potential, a clear mapping is needed regarding tourist visit patterns in various destinations in East Java, so that tourism development can be carried out more effectively and on target.

In the period 2018 to 2022, there were quite significant dynamics in the number of tourist visits to various tourist destinations in East Java. Factors such as tourism promotion, infrastructure quality, and the major impact of the COVID-19 pandemic influenced the fluctuation of the number of tourists during this period. Therefore, it is important to group districts/cities based on the number of tourist visits during this period, in order to understand visit trends and identify more accurate tourism development strategies. One of the appropriate methods for this analysis is DBSCAN (Density-Based Spatial Clustering of Applications with Noise), a density-based clustering method that is able to identify high-density data groups while separating noise or outliers.

The DBSCAN method has the advantage of grouping data with irregular cluster shapes and varying sizes, without requiring the determination of the number of clusters from the beginning as in the K-Means method. In the context of East Java tourism, the application of this method will help identify areas with high concentrations of tourist visits, as well as detect areas that may not be included in certain groups but still have the potential to be developed as tourist destinations. Thus, the analysis using DBSCAN is expected to provide a more comprehensive picture of tourist visit patterns and the potential for tourism development in East Java.

This research aims to apply the DBSCAN method in mapping districts/cities in East Java based on the number of tourist visits during the period 2018–2022. Through this analysis, it is expected to identify groups of regions with varying levels of visits, ranging from high, medium, to low. This information will be very valuable in helping local governments formulate more targeted promotional policies, allocate resources more efficiently, and improve the overall attractiveness of tourist destinations. In addition, this research can also provide guidance for tourism industry players to innovate and collaborate in developing the tourism sector in East Java so that it can have a more significant impact on the regional economy.

Previous studies have shown that clustering methods such as Fuzzy C-Means and DBSCAN are effective in regional data analysis. The Fuzzy C-Means method for clustering regions in East Java Province based on the Human Development Index (HDI) indicator, which successfully identified the distribution pattern of human development in the area (Qori'atunnadyah, 2023a). This research emphasizes the advantages of Fuzzy C-Means in handling data with unclear cluster boundaries. Meanwhile, the DBSCAN algorithm to cluster countries based on tobacco control scores. Their results showed that DBSCAN was effective in detecting clusters with varying data densities and was able to identify outliers in the dataset (Riyono et al., 2024). In addition, The performance of the K-Means, K-Medoids, and DBSCAN algorithms in clustering provinces in Indonesia based on indicators of community welfare (Saputri & Arianto, 2023). They found that each algorithm has its own advantages depending on the characteristics of the data used. The findings of these studies underscore the importance of selecting the right clustering method according to the nature of the data and inspire the use of DBSCAN in this research to provide a more comprehensive analysis of tourism potential in East Java.

Previous research entitled "Mapping Domestic and Foreign Tourists in East Java Using the C-Means Clustering" has mapped districts and cities in East Java based on the number of domestic and foreign tourist visits using the C-Means Clustering method, which resulted in three main groups for domestic tourists and five groups for foreign tourists (Qori'atunnadyah, 2024). Although this research provides valuable insights into regional tourism dynamics, the analysis has limitations, especially in capturing outliers or areas with visiting characteristics that do not conform to general patterns. Therefore, this research will use the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) method which has advantages in detecting clusters with irregular shapes and capturing noise or data that is not included in any cluster. By using data on the number of domestic and foreign tourists from 2018 to 2022, it is hoped that the DBSCAN method can provide a more comprehensive analysis of the distribution of tourists and tourism potential in East Java, as well as offering new insights for planning more effective promotional strategies and resource allocation.

Several other studies have also used clustering methods such as C-Means and DBSCAN for various spatial analysis needs. For example, the C-Means methods have been used for clustering regions in East Java based on the Human Development Index (HDI) indicator, which shows its superiority in understanding regional distribution based on development data (Qori'atunnadyah, 2023b). Previous studies have also used the K-Means algorithm for clustering regions based on the teacher-student ratio at the level of education (Qori'atunnadyah, 2022). The C-Means method has also been applied to cluster manufacturing companies based on factors influencing company value (Qori'atunnadyah et al., 2023).

In addition, several other studies have also used clustering methods such as DBSCAN for various spatial analysis needs. For example, the DBSCAN method has been analyzed to detect outlier data, showing its effectiveness in identifying anomalous data (Armiady, 2022). The DBSCAN algorithm has also been used for clustering earthquake data, helping to understand seismic patterns and distributions (Rahman & Wijayanto, 2021). In addition, this method was applied to cluster the number of student visits to campus libraries, which provided insight into visiting patterns and utilization of library facilities (Syafrianto & Riswanto, 2023). DBSCAN was also implemented to cluster student interest data after the COVID-19 pandemic, demonstrating its flexibility in adapting to changes in student interests and behavior in the educational context (Kristianto, 2022). These studies confirm that DBSCAN is an effective tool for various clustering applications in data analysis. This study will expand the application of DBSCAN in the tourism sector, with the hope of providing deeper insights into the dynamics of tourist visits in East Java.

To meet these needs, this study aims to map tourist visit patterns in districts/cities in East Java during the period 2018 to 2022 using the DBSCAN method. With this approach, the study is expected to identify clusters of tourism activities, capture outlier data, and provide insight into regional tourism dynamics. This information will support the development of more targeted promotional strategies and effective resource allocation in the tourism sector in East Java.

MATERIALS AND METHODS

Data on the number of tourist visits in East Java for the period 2018 to 2022 will be obtained from the official publication of the East Java Provincial Culture and Tourism Office. The data used in this study only covers up to 2022 because, at the time of this research, the available and accessible data was only up to that year. If data for 2023 becomes available later, this analysis can be updated to provide more current results. At the stage of calculating the distance between data points in the DBSCAN algorithm, the Euclidean Distance function is used, where differences in scale between variables can affect the grouping results. Therefore, this research applies a normalization process to align the value scales of all variables, using the Z-Score method. The Z-Score equation is stated as follows:

$$zscore = \frac{x - \mu_x}{\sigma_x} \quad (1)$$

Where z is the normalized value, μx is the average of the x variable, and σx is the standard deviation of the x variable.

Clustering or cluster analysis is a group of methods used to find several meaningful groups or clusters in data (Giordani et al., 2020). Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is an algorithm used to group data based on density. Density in DBSCAN refers to the minimum amount of data within a radius of ε. Data that meets this requirement is considered a cluster core if the number of its neighbors and itself within a radius of ε is ≥ MinPts. The values of the radius ε and MinPts are determined independently. Data that has several neighbors and itself within a radius of ε less than MinPts, but has neighbors who are cluster cores, is categorized as boundary data. Meanwhile, if the number of neighbors and itself within a radius of ε is less than MinPts and there are no neighbors who are cluster cores, the data is categorized as noise (Sabor et al., 2021). The following are the steps in DBSCAN:

1. Determine the value of the MinPts and Eps parameters.
2. Randomly determine the value of *p* or the starting point.
3. Calculate Eps or all distances of points that are density reachable to *p* using the following Euclidean distance formula.

$$d_{ij} = \sqrt{\sum_{a}^{p}\left(x_{ia} - x_{ja}\right)^2} \quad (2)$$

   Where x_i is the a-th variable of object *i* (*i*=1, ..., *n*; *a*=1, ..., *p*) and did is the Euclidean distance value.
4. A cluster is formed when the points that meet Eps are more than MinPts and point p is the core point.
5. Repeat steps 3-4 until the process is carried out at all points. If p is a border point and there are no points that are density reachable to p, then the process is continued to another point.

The criteria for identifying data as noise in DBSCAN clustering are based on the density defined by the parameters ε (epsilon) and MinPts (minimum points). In DBSCAN, a point is classified as noise if it does not meet the density threshold required to form or be part of a cluster (Monalisa et al., 2023). Specifically, a point is considered noise if the number of points within its radius ε (neighborhood radius) is less than MinPts. This means that the point does not have enough neighboring points around it to meet the minimum density required for cluster formation (Sabor et al., 2021). In theory, noise points are often considered outliers, because they do not follow the same density pattern as other points in the cluster. Noise points represent isolated points or areas with much lower density compared to the cluster. These noise points can be useful in identifying anomalies or areas of low activity, which can provide valuable insights in a variety of applications.

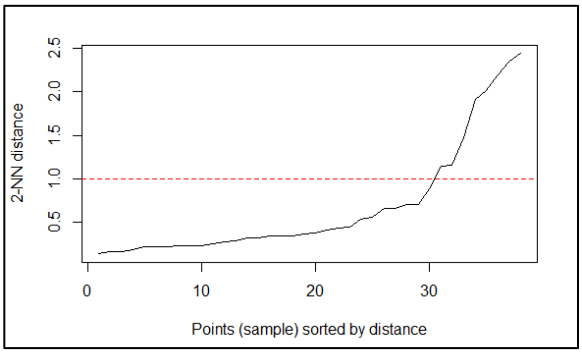
To ensure that the clustering results are appropriate and representative of the data population, the clustering results must be validated. Clustering validation methods consist of internal and external indices. Internal indices are used to measure the quality of the clustering structure without considering external information, while external indices measure the extent to which cluster labels match externally given class labels. One commonly used validation method is the Silhouette Coefficient. The Silhouette Coefficient is calculated using the following equation (Batool & Hennig, 2021):

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad \text{...(3)}$$

The value $s(i)$ represents the Silhouette Coefficient, $a(i)$ is the average distance between point i and all other points within cluster A (the cluster where point i belongs), while $b(i)$ is the average distance between point i and all points in clusters other than A.

RESULTS AND DISCUSSION

In this research, the ε value to be used in the analysis is determined using the k-Nearest Neighbor algorithm with k = 2.



Source : (Research Results, 2024)

Figure 1. 2-NN on Determining the Value of ε (Domestic)

Based on Figure 1, it can be seen that the tipping point is at point 1.00 so the optimum ε value used in further analysis is 1.00. Then, determining the MinPts value can be done by trial and error by observing the Silhouette Coefficient value produced. In the following Table 2, several Silhouette Coefficient values from several experiments are shown.

Table 1. Combination of Input Values ε and MinPts (Domestic)

eps	minPts	silhouette
1	3	0.618844
1	4	0.618844
1	5	0.618844
1	6	0.59859
1	7	0.59859
1	8	0.565245
1	9	0.565245
1	10	0.565245

Source: (Research Results, 2024)

Table 1 shows that at a certain value of ε, increasing the MinPts value produces a pattern of diminishing returns to cluster quality, as measured using the Silhouette Coefficient. Initially, increasing the MinPts value can improve cluster quality, but after reaching a certain point, the increase results in a decrease in quality. In addition, the greater the MinPts value, the less likely it is to form a cluster. This is due to the increase in the minimum requirement for the number of adjacent objects to form a cluster, so more objects are needed for a cluster to form, and vice versa. After the optimum values of ε and MinPts are found, which are indicated by the highest Silhouette Coefficient value, the clustering process is carried out using the DBSCAN algorithm. The optimum values of ε and MinPts (ε = 1.00 and MinPts = 3) are then used as input to run the clustering process with DBSCAN.

Table 2. DBSCAN Output Using MinPts= 3 and ε= 1.00 (Domestic)

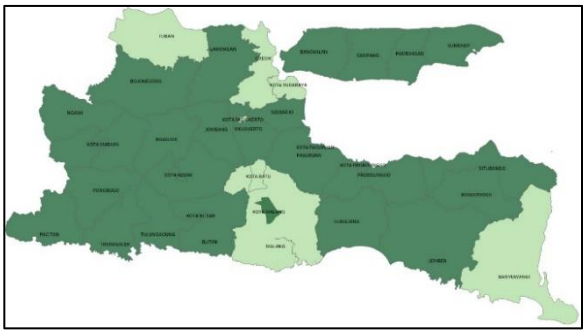
Output	Description	Number	Average tourist visits
0	Noise	7	3,817,966
1	Cluster 1	31	974,583

Source: (Research Results, 2024)

Table 2 shows the results of the DBSCAN algorithm with parameters MinPts = 3 and ε = 1. The table shows that of the total data, 7 data are categorized as noise or outliers. These data do not have enough neighbors within a radius of ε to form a cluster. The average number of domestic tourist visits in the data included in the noise category is 3,817,966, which shows that although considered as noise, this data represents locations with a high number of domestic tourist visits but cannot be included in any cluster.

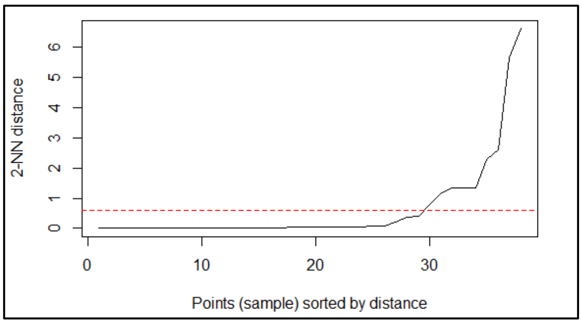
Meanwhile, 31 data were successfully grouped into Cluster 1. The average number of tourist visits in this cluster is 974,583, which shows that the majority of tourist locations in this cluster have a lower number of domestic tourist visits compared to the noise data. Cluster 1 contains areas with relatively similar characteristics of domestic tourist visit density and has sufficient spatial proximity to form a homogeneous group. These results provide an overview that most of the tourist locations in East Java are included in Cluster 1, while only a small part is considered as noise with a higher number of domestic tourist visits.

The cluster map in Figure 2 displays a visualization of tourist visit patterns based on the number of domestic tourist visits in East Java. Each color on the map represents a different cluster group, indicating areas with varying levels of domestic tourist activity. Lighter colors indicate areas that fall into the noise or outlier category, namely areas with high visits but do not conform to the main cluster pattern. Meanwhile, darker colors indicate main clusters with more stable and regular concentrations of visits. This color grouping provides a clear picture of the distribution of tourist visits across the province, making it easier to identify areas with high tourism potential but with visit characteristics that differ from the general pattern.



Source: (Research Results, 2024)
Figure 2. Mapping of Domestic Tourist Visits

In this research on foreign tourist visit data, the ε value to be used in the analysis is determined using the k-nearest Neighbor algorithm with k = 2.



Source: (Research Results, 2024)
Figure 3. 2-NN on Determining the Value of ε (Foreign)

Based on Figure 3, it can be seen that the tipping point is at point 0.6 so the optimum ε value used in further analysis is 0.6. Then, determining the MinPts value can be done by trial and error by paying attention to the Silhouette Coefficient value produced. In the following table 3, several Silhouette Coefficient values from several experiments are shown.

Table 3. Combination of Input Values ε and MinPts (Foreign)

eps	minPts	silhouette
0.6	3	0.71606
0.6	4	0.71606
0.6	5	0.71606
0.6	6	0.71606
0.6	7	0.71606
0.6	8	0.71606
0.6	9	0.71606
0.6	10	0.71606

Source: (Research Results, 2024)

Table 3 shows that at a value of ε = 0.6, the addition of the MinPts value does not result in a significant change in the quality of the cluster, as measured using the Silhouette Coefficient. From the table, it can be seen that for each MinPts value from 3 to 10, the Silhouette Coefficient remains constant at 0.71606. This indicates that the addition of MinPts in this range has no impact on increasing or decreasing the quality of the cluster. Although usually, the addition of MinPts can improve the quality of the cluster up to a certain point, in this case, the increase in the MinPts value does not affect the clustering results. In addition, the greater the MinPts value, the less likely it is to form a new cluster because the minimum number of objects needed to form a cluster increases. Considering the values of ε and MinPts that are consistent in producing optimal Silhouette Coefficient values, the values of ε = 0.6 and MinPts = 3 are used as parameters to run the clustering process with the DBSCAN algorithm.

Table 4. DBSCAN Output Using MinPts= 3 and ε= 1.00 (Foreign)

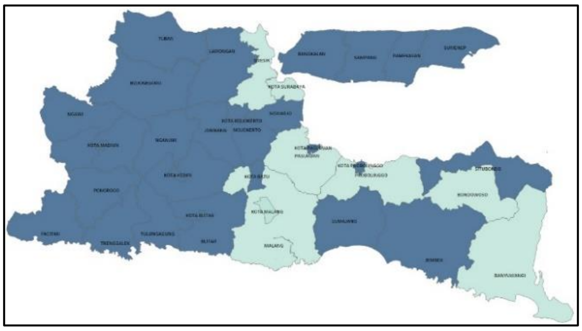
Output	Description	Number	Average tourist visits
0	Noise	8	29,962
1	Cluster 1	30	647

Source: (Research Results, 2024)

Table 4 shows the results of the DBSCAN algorithm application with parameters MinPts = 3 and ε = 1.00. Based on the table results, there are 8 data that are categorized as noise or outliers. This data does not have enough neighbors within a radius of ε to form a cluster. The average tourist visits for data that fall into the noise category is 29,962, which indicates that although this data is classified as noise, the location still has a relatively significant number of tourist visits, but not enough to be included in any cluster.

In addition, 30 data were successfully grouped into Cluster 1. The average tourist visits in this cluster are 647, indicating that the majority of tourist locations in this cluster have a much lower number of visits compared to locations categorized as noise. Cluster 1 includes areas with similar characteristics of tourist visit density, where these locations are spatially close enough to form a homogeneous group. These results indicate that most of the tourist locations in the dataset belong to Cluster 1, while only a small portion is isolated as noise, with a higher number of tourist visits than the cluster average.

The results of this analysis are visualized in the form of a map in Figure 4 to facilitate understanding of the distribution of foreign tourist visits in East Java. The cluster map in Figure 4 shows the distribution of areas based on the clustering results, with different colors for each group. On this map, lighter colors indicate areas that fall into the noise or outlier category, namely areas with a high number of tourist visits but do not follow the main cluster pattern. Conversely, darker colors indicate the main cluster, which consists of areas with lower levels of visits and is consistent with the general pattern. This visualization provides a clearer picture of the pattern of tourist visits to various destinations in East Java, making it easier to identify areas with significant tourism potential but different visit characteristics.



Source: (Research Results, 2024)

Figure 4. Mapping of Foreign Tourist Visits

CONCLUSION

This study successfully mapped the pattern of domestic and foreign tourist visits in East Java using the DBSCAN method with parameters determined through the k-nearest Neighbor algorithm. In domestic tourist data, with a value of $\varepsilon = 1.00$ and MinPts = 3, one main cluster was generated consisting of 31 tourist locations and 7 locations categorized as noise. The average tourist visit in the noise data showed a higher number of visits compared to the main cluster but did not have the appropriate characteristics to be combined in the main cluster. Meanwhile, in foreign tourist data, with parameters $\varepsilon = 0.6$ and MinPts = 3, one main cluster was also generated with 30 tourist locations, and 8 other locations were categorized as noise. The noise locations had a higher average visit compared to the main cluster but did not follow the general pattern of the cluster.

The visualization results in the form of cluster maps provide a clear picture of the distribution pattern of tourist visits in East Java, where lighter colors indicate noisy locations with high visits and darker colors indicate main clusters with lower visits. This shows that most tourist destinations are in the main cluster, while some areas have significant tourism potential but are not grouped in a general pattern. The suggestion from this study is that local governments use the results of this mapping to develop more effective promotional strategies, especially for noisy areas with high potential for visits. In addition, infrastructure development can be focused on main clusters with low visits and noisy areas that require improved facilities. Further studies are also recommended to analyze the factors that influence visits, as well as utilizing mapping technology for regular monitoring of tourist visits.

REFERENCE

Armiady, D. (2022). Analisis Metode DBSCAN (Density-Based Spatial Clustering of Application with Noise) dalam Mendeteksi Data Outlier. *JURIKOM (Jurnal Riset Komputer)*, *9*(6), 2158. https://doi.org/10.30865/jurikom.v9i6.5080

Badan Pusat Statistik. (2023). *Statistik Pariwisata Provinsi Jawa Timur 2022*. https://jatim.bps.go.id/id/publication/2023/07/03/5bab263b9158357b69bc9309/statistik-pariwisata-provinsi-jawa-timur-2022.html

Batool, F., & Hennig, C. (2021). Clustering with the Average Silhouette Width. *Computational Statistics & Data Analysis*, *158*, 107190. https://doi.org/10.1016/j.csda.2021.107190

Giordani, P., Ferraro, M. B., & Martella, F. (2020). *An Introduction to Clustering with R* (Vol. 1). Springer Singapore. https://doi.org/10.1007/978-981-13-0553-5

Kristianto, A. (2022). Implementasi DBSCAN dalam Clustering Data Minat Mahasiswa Setelah Pandemi Covid19. *KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi*, *2*(2). https://doi.org/10.24002/konstelasi.v2i2.5638

Monalisa, S., Juniarti, Y., Saputra, E., Muttakin, F., & Ahsyar, T. K. (2023). Customer segmentation with RFM models and demographic variable using DBSCAN algorithm. *TELKOMNIKA (Telecommunication Computing Electronics* *and Control)*, *21*(4), 742. https://doi.org/10.12928/telkomnika.v21i4.22759

Qori'atunnadyah, M. (2022). Pengelompokkan Wilayah Berdasarkan Rasio Guru-Murid Pada Jenjang Pendidikan Menggunakan Algoritma K-Means. *Journal of Informatics Development*, *1*(2), 33–38.

Qori'atunnadyah, M. (2023a). Fuzzy C-Means for Regional Clustering in East Java Province Based on Human Development Index Indicators. *J Statistika*, *16*(2), 524.

Qori'atunnadyah, M. (2023b). Metode C-Means untuk Pengelompokkan Kabupaten/Kota Provinsi Jawa Timur berdasarkan Indikator Indeks Pembangunan Manusia (IPM). *Journal of Informatics Development*, *1*(2), 51–58. https://doi.org/10.30741/jid.v2i2.1013

Qori'atunnadyah, M. (2024). Mapping Domestic and Foreign Tourists in East Java Using C-Means Clustering. *Jurnal Statistika Dan Aplikasinya*, *8*(1), 54–62. https://doi.org/10.21009/JSA.08105

Qori'atunnadyah, M., Liyundira, F. S., & Indrianasari, N. T. (2023). *Grouping Manufacturing Companies Based on Factors Affecting Firm Value Using C-Means Clustering* (pp. 28–31). https://doi.org/10.2991/978-94-6463-346-7_6

Rahman, R. R. A., & Wijayanto, A. W. (2021). PENGELOMPOKAN DATA GEMPA BUMI MENGGUNAKAN ALGORITMA DBSCAN. *Jurnal Meteorologi Dan Geofisika*, *22*(1), 31. https://doi.org/10.31172/jmg.v22i1.738

Riyono, J., Pujiastuti, C. E., Latifa, A., & Putri, R. (2024). CLUSTERING NEGARA BERDASARKAN SKOR PENGENDALIAN KONSUMSI TEMBAKAU MENGGUNAKAN ALGORITMA DBSCAN. *Jurnal Teknik Informatika Kaputama (JTIK)*, *8*(1).

Sabor, K., Jougnot, D., Guerin, R., Steck, B., Henault, J.-M., Apffel, L., & Vautrin, D. (2021). A data mining approach for improved interpretation of ERT inverted sections using the DBSCAN clustering algorithm. *Geophysical Journal International*, *225*(2), 1304–1318. https://doi.org/10.1093/gji/ggab023

Saputri, F. W., & Arianto, D. B. (2023). PERBANDINGAN PERFORMA ALGORITMA K-MEANS, K-MEDOIDS, DAN DBSCAN DALAM PENGGEROMBOLAN PROVINSI DI INDONESIA BERDASARKAN INDIKATOR KESEJAHTERAAN MASYARAKAT. *Jurnal Teknologi Informasi: Jurnal Keilmuan Dan Aplikasi Bidang Teknik Informatika*, *7*(2), 138–151. https://doi.org/10.47111/jti.v7i2.9558

Syafrianto, A., & Riswanto, E. (2023). Pengelompokkan Jumlah Kunjungan Mahasiswa ke Perpustakaan Kampus Menggunakan Algoritma DBSCAN. *G-Tech: Jurnal Teknologi Terapan*, *7*(1), 75–81. https://doi.org/10.33379/gtech.v7i1.1925