ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Spatial Clustering of Dormitory Density in Mueang District, Buriram Province Using the DBSCAN Algorithm

Jiravadee Yoyram[1], Prem Enkvetchakul[2], Jongkol Janruang[3], and Kittikoon Boonkate[4*]

[1] Faculty of Sciences, Buriram Rajabhat University, Buriram, 31000, Thailand
[2] Faculty of Sciences, Buriram Rajabhat University, Buriram, 31000, Thailand
[3] Faculty of Science and Liberal Arts, Rajamangala University of Technology Isan, Nakhonratchasima, 30000, Thailand
[4] Faculty of Sciences, Buriram Rajabhat University Buriram, Thailand
* Correspondence: kittikoon.bk@bru.ac.th

Citation:
Yoyram, J.; Enkvetchakul, P.; Janruang, J.; Boonkate, K. Spatial clustering of dormitory density in Mueang district, Buriram province using the DBSCAN algorithm. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(3), e258117. https://doi.org/10.55164/ajstr.v28i3.258117.

Article history:
Received: March 6, 2025
Revised: April 15, 2025
Accepted: April 19, 2025
Available online: April 26, 2025

Publisher's Note:
This article has been published and distributed under the terms of Thaksin University.

Abstract: Spatial distribution studies the geographical arrangement of elements within a given area, crucial for urban planning, resource allocation, and spatial decision-making. This research aimed to examine the spatial distribution of dormitories in Mueang District, Buriram Province, using the DBSCAN algorithm, and to cluster dormitory neighborhoods within the area. The methodology followed a five-step data mining process: 1) data gathering from 387 points obtained via website extraction, 2) data preparation, 3) clustering using DBSCAN, 4) visualization of results through a distribution map and clustering outcomes, and 5) validation of clustering. The findings revealed that dormitory clusters in Mueang District exhibited a dense and notable pattern. Data was classified into three clusters: Cluster 1 included Nai Mueang and Isan Subdistricts, Cluster 2 covered Samed Subdistrict, and Cluster 3 encompassed Krasang and Ban Bua Subdistricts. The Silhouette Coefficient of 0.495 indicated good clustering performance, while the Davies-Bouldin Index (DBI) of 2.785 showed acceptable results, demonstrating DBSCAN's effectiveness in clustering dormitories. The algorithm's flexibility in parameter adjustment allows for results that align with the specific context of the area, making it suitable for spatial research. It also reduces time and costs in data collection and analysis. These clusters provide valuable insights for students selecting dormitories and entrepreneurs investing in private dormitories, and hold potential for future development.

Keywords: Dormitory distribution; DBSCAN algorithm, Spatial distribution

1. Introduction

Machine Learning (ML) and Artificial Intelligence (AI) are increasingly becoming central in developing web and mobile applications called digital platforms. That means the main application task is a specialized task requiring mathematically provable computations. The technologies must support engineering simulations, automated theorem proving, scientific computing, and AI-powered math solvers, ensuring accuracy and verifiability in complex calculations. Machine Learning and Artificial Intelligence are essential tools for data analysis in various fields. For example, Convolutional Neural Networks are used for image classification of the cassava leaf to develop a Line Bot that can identify and provide information about Cassava Leaf Diseases [15], and DBSCAN (density-based spatial clustering of application with noise) is applied in spatial analysis to decide for planning and developing the main content of the city such as hotel distribution [1][5], tourism planning [11][12][16] and analyzing an urban space [6]. These tools assist in analyzing complex data and making decisions effectively.

To demonstrate the applicability of DBSCAN in urban planning, several studies have successfully employed this clustering algorithm to support spatial decision-making. For instance, Tu et al. applied DBSCAN using point-of-interest (POI) data to analyze the spatial structure of "production–living–ecological" zones in Wuhan, China, which helped inform sustainable urban development strategies [9]. Similarly, Fauzan et al. utilized DBSCAN to cluster hotel locations in Bali, enabling tourism stakeholders to better plan post-pandemic economic recovery by identifying key zones of accommodation demand [5]. Furthermore, Camilo Alberto Caudillo-Cos and Jorge Alberto Ruiz-Pérez conducted a study titled Defining Urban Boundaries through DBSCAN and Shannon's Entropy: The Case of the Mexican National Urban System, in which they proposed a novel method for delineating urban boundaries in Mexico. This approach integrates the DBSCAN algorithm with Shannon's entropy to analyze data from multiple and diverse sources. [2].

In the context of this study, the application of DBSCAN to cluster dormitories in Mueang District, Buriram Province, follows a similar rationale. The ability of DBSCAN to identify high-density residential zones based on spatial distribution provides valuable insights for students seeking accommodation and for city planners or private entrepreneurs aiming to develop or invest in dormitory infrastructure. The algorithm's robustness in handling noise and irregular densities makes it particularly effective in modeling real-world urban patterns.

1.1 Problem Statement

Buriram Rajabhat University serves as a key educational institution driving local development. It holds a prominent position within Thailand's northeastern region, contributing to local advancement and global competitiveness**.** The university boasts 11,514 regular students [10], and the surrounding vicinity is home to a significant population of students, faculty, and members of the general public in the academic year 2022. It causes a substantial portion of students who originate from various regions to face challenges when selecting suitable accommodation or dormitories. Consequently, dormitories play a crucial role in addressing the accommodation needs of students, with preferences typically leaning towards options offering security, proximity to the university for ease of commuting, and affordable rental rates. Private dormitories have seen a surge in demand among university students, largely due to insufficient on-campus housing provided by the university. As the student population grows, there has been a corresponding rise in entrepreneurs venturing into the private dormitory sector. The location selection process for dormitory construction is a critical decision for these entrepreneurs.

The dormitory zones are important information for students who must find a suitable dormitory. This is because a suitable dormitory can help students reduce the cost and time spent at university. Therefore, examining the spatial distribution of dormitories in Mueang District, Buriram Province, is presented as a new approach to finding and providing information about dormitory zones for students. A new approach used the DBSCAN algorithm as a main model to support scientific computing and AI-powered math solvers. Moreover, the advantages of the DBSCAN algorithm are that it can effectively solve problems from large amounts of data, its performance is based on the overlapping and covering of points of interest, and its fast clustering speed, which can effectively handle noise points and divide high-density areas into clusters to form easier-to-recognize spatial clusters, it is applying to both convex and non-convex sample sets [9]. However, implementing DBSCAN algorithms must combine with other platforms to show the algorithms' results in visualization and results analysis, such as the Shiny web application [18] and ArcGIS [9].

To substantiate the selection of DBSCAN as the primary clustering algorithm in this study, a comparative evaluation was conducted against other well-established clustering methods, namely K-Means and Hierarchical Clustering. While widely used, K-Means assumes spherical cluster shapes and requires the number of clusters to be defined in advance—limitations that make it less suitable for spatial datasets with irregular distributions and unknown group structures [3, 12]. Hierarchical clustering, though advantageous in not requiring a pre-defined number of clusters, becomes computationally intensive and less efficient when applied to large-scale datasets [11]. In contrast, DBSCAN offers several key advantages that align well with the characteristics of the dataset used in this study. It can effectively identify clusters of arbitrary shapes and automatically detect noise or outliers without prior knowledge of the number of clusters. Given the spatial nature of dormitory data in Mueang District, Buriram Province—characterized by varying densities and noise—DBSCAN demonstrates superior flexibility and accuracy. Its robustness in handling non-linear spatial patterns makes it particularly well-suited for uncovering meaningful dormitory clusters within this urban context [9], [13].

Based on the advantages of DBSCAN algorithms, this paper presents a new approach to analyzing dormitory distribution and clustering dormitories to identify dormitory zones. In addition, the Shiny web application displays the results of DBSCAN algorithms and is active with users. The results of the proposed approach can be used as information to support students in finding and making decisions about their suitable dormitory based on spatial distribution.

2. Materials and Methods

The research framework of the proposed approach for analysis of dormitory distribution and clustering to find the dormitory zone in Mueang District, Buriram Province. The research methodology includes the following details:

2.1 Research Framework

Clustering dormitories in Mueang District, Buriram Province, based on their density using the DBSCAN algorithm, involves extracting and retrieving location data from websites through Web Scraping. This data encompasses dormitories, apartments, mansions, and housing in the Mueang District area of Buriram Province. Subsequently, the data undergoes clustering using the DBSCAN algorithm, which hinges on the distance or proximity between points (Eps) and the minimum number of points (MinPts). The R-Studio program with R language and Shiny Web Application Framework are utilized to visualize the spatial distribution on maps and present the clustering results. The efficacy of clustering is validated using the Silhouette Coefficient concept. The research framework is illustrated in the five layers, which are shown in Figure 1.

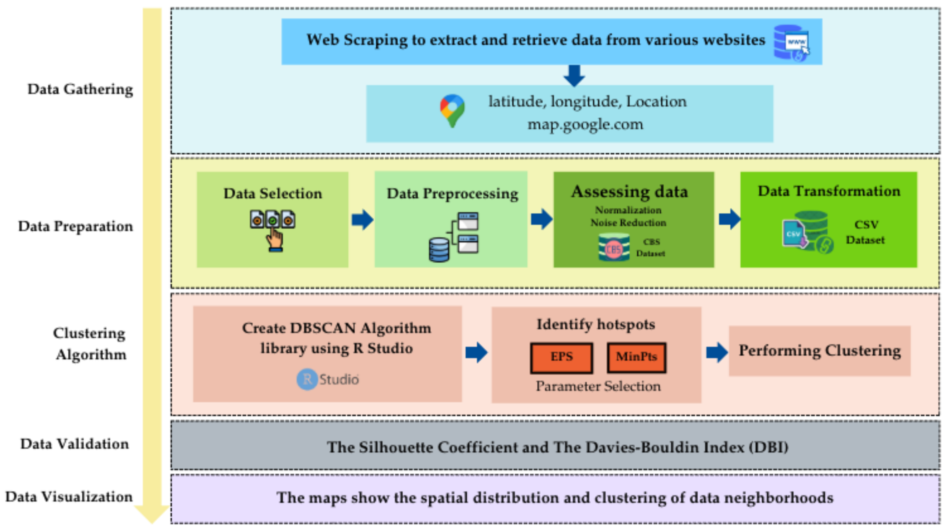


Figure 1. Research Framework

According to Figure 1, the research workflow for spatial clustering of dormitories in Mueang District, Buriram Province, using the DBSCAN algorithm, is an application of the data mining process consisting of 5 operational layers [3][4]. The main point of these steps is in a Clustering algorithm layer called the DBSCAN algorithm, which is used to identify the dormitory zones for supporting students of Buriram Rajabhat University. Finally, in the data visualization layer, the dormitory zones are displayed on the web using the

Shiny web application. Interestingly, users can set the distance (Eps) and the minimum number of points (MinPts) by themselves when the shiny web application is used as a tool to develop the user interface of the system [12], [16].

The evaluation of the proposed approach consists of two steps. First, researchers cross-check all datasets with local data sources in the areas. This cross-checking is a sub-process in data preparation and must be performed to ensure validating data accuracy. This means the dataset is corrected because it is verified before being used in the clustering model. Second, the Silhouette Coefficient and the Davies-Bouldin Index (DBI) is a statistical models used to evaluate the efficacy of clustering.

2.2 Research Methodology

Based on the Research Framework in Figure 1, this research proposes a new approach to examine dormitories' spatial distribution and cluster the dormitory zone in Mueang District, Buriram Province, using the DBSCAN algorithm. The proposed approach consists of five 5 steps, and the details of each process are shown below.

1) Step 1: Data Gathering

The data-gathering process entails examining the dormitory distribution in Mueang District, Buriram Province, using Web Scraping to extract and retrieve data from various websites. This process encompasses gathering the coordinates and names of each location and recording/organizing the data into an Excel spreadsheet format, delineating it into rows and columns, as illustrated in Table 1.

Table 1. Data Gathering Details

Field	Description
latitude	Latitude coordinates
longitude	Longitude coordinates
Location	Name of the place

2) Step 2: Data Preparation

Data preparation involves refining raw data to align with research objectives, comprising three key steps :

2.1) Data Selection or Acquisition: Following data extraction from websites via Web Scraping based on Google Maps, the data equals 390. The comprehensive data should encompass latitude and longitude coordinates along with location names. The dataset used in this study was selected from January 2024. The data were collected from publicly accessible sources using web scraping. No personal or sensitive information was included, and the process complied with ethical guidelines for academic use. The example of Data Selection or Data Acquisition Process is shown in Figure 2

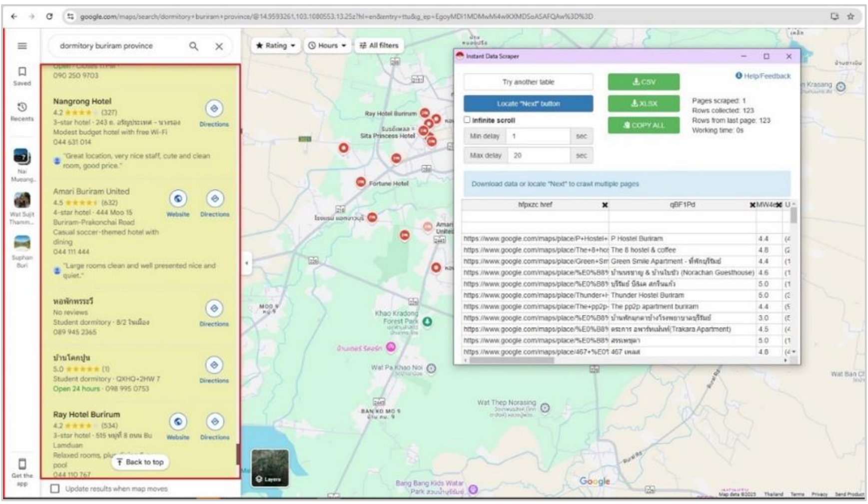


Figure 2. Example of Data Selection or Data Acquisition Process

2.2) Data Preprocessing: This step entails validating data accuracy, including latitude and longitude coordinates and location names. Researchers ensure data integrity by cross-checking all entries with local data sources in the areas. Then, data cleaning was carried out to clear the attributes with missing values, such as latitude or latitude which is loose or wrong. This is because errors from scraping via map.google.com may arise, including inaccurate latitude and longitude coordinates that deviate from the actual location. Additionally, some retrieved data may be incomplete or duplicated, such as duplicate place names or missing information, which affects data preparation, the accuracy of the clustering process, and the quality of results produced by the DBSCAN algorithm. For example, the dormitory location is incorrect when researchers cross-check all entries with local data sources in the areas. There are 3 mistakes in the locations deleted.

Based on Data Preparation, the dataset can be uploaded to the R-Studio program to analyze the distribution of dormitory locations and understand the data. For example, the distribution of dormitory is shown in Figure 3.

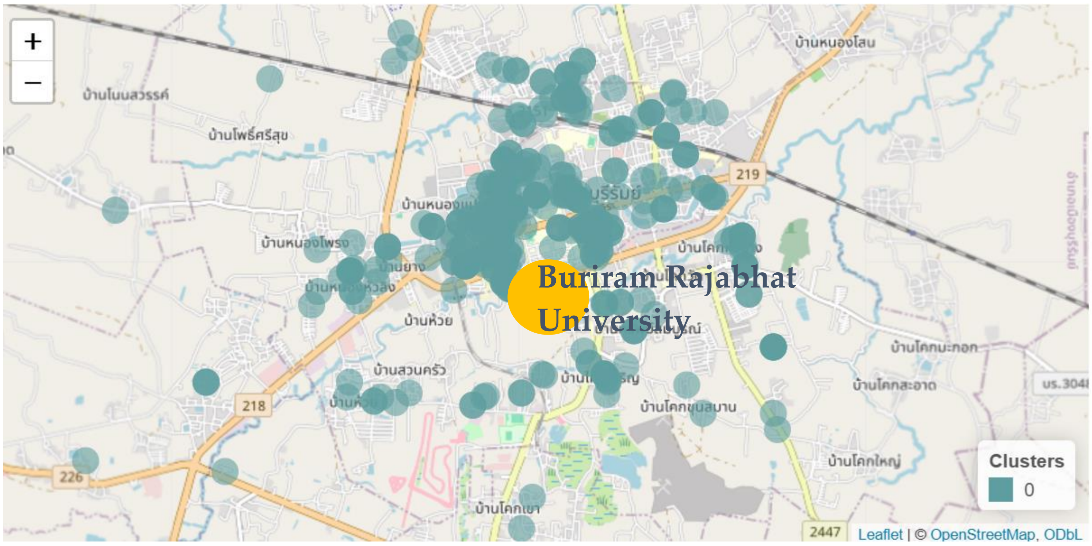


Figure 3. Example of distribution of dormitory location

According to Figure 3, dormitories are located in a high concentration in the urban area near Buriram Rajabhat University. The dataset distribution on the map shows that the dormitory location is a disordered pattern without clear planning, called unstructured distribution. Moreover, noises, outliers, or small data groups on Google Maps occur around urban areas near Buriram Rajabhat University (cycle with yellow color). This means that the dataset used as a sampling of this paper is suitable for the DBSCAN algorithm. This is because DBSCAN is an algorithm for datasets that are separated as clusters with varying shapes and sizes of cluster.

The scope of the spatial study area encompasses dormitory data within Mueang District, Buriram Province. The data scope includes dormitories, apartments, mansions, and housing. Final data was collected on January 1, 2024, referencing locations through Google Maps. The data was recorded and stored in an Excel spreadsheet comprising 387 entries. An example of data collection is illustrated in Table 2.

Table 2. Example of Dormitory Location Coordinate Data Gathering in Mueang District, Buriram Province

latitude	longitude	location
103.0942767	14.9956900	AP Apartment Buriram
103.0942767	14.9882406	Serm-Suni 4 Place Dormitory
103.0919406	15.0113636	Boonmani Dormitory
.	.	.
.	.	.
103.0818726	15.0134048	Baitong Mansion
103.0702891	14.9836372	Night for You Residence

2.3) Data Transformation: Data transformation involves converting data into a format conducive to analysis by storing files in CSV format. The CSV format consists of three fields: latitude, longitude, and the place's name (location).

3) Step 3: Clustering Algorithm

In the clustering algorithm step, the DBSCAN algorithm is chosen as the main model to find the pattern of spatial data. This is because the clustering algorithm process involves organizing spatial data using the DBSCAN algorithm, which is particularly suitable for datasets exhibiting specialized non- linear distribution patterns that cannot be effectively clustered by the K-means method [7]. The DBSCAN algorithm operates by considering the distance (Eps) and the minimum number of points (MinPts) [4][11]. Users can adjust parameter values to tailor the outcomes to the specific requirements and characteristics of the data available in the area. This customization is enabled by creating a library using clustering tools, including the RStudio program in conjunction with the R language and the Shiny Web Application Framework, an application widely utilized for spatial development tasks or spatial data analysis. Users can set the distance (Eps) and the minimum number of points (MinPts) to ensure alignment and suitability with the data in the area [13][16].

4) Step 4: Validation

Validation entails evaluating the efficacy of clustering the data by employing

4.1) The Silhouette Coefficient, introduced by Peter J. Rousseeuw in 1986 as a metric for data grouping, considering the compactness of points within clusters and the separation between clusters. The Silhouette Coefficient acts as an internal criterion for assessing the quality of cluster formation [10], utilizing the class and their interrelations. It is calculated from the inter-class cohesion and can be computed using equation (1).

$$S(i) = \frac{a(i)-b(i)}{Max\{a(i),b(i)\}} \tag{1}$$

By definition, $a(i)$ represents the number of connections between class i and any other class within the same cluster while

$b(i)$ denotes the number of connections between class i and any other class across clusters.

The range of the Silhouette coefficient is from -1 to 1, where a value close to 1 indicates that a class is highly suitable for its cluster, signifying high cohesion. Conversely, a value approaching -1 indicates low cohesion, suggesting errors in clustering. Therefore, it is highly probable that any class would need to modify its Silhouette coefficient (S(i)) calculation within each cluster to obtain the average value $\bar{s}$ for assessing intra-cluster suitability, which falls between -1 and 1. Evaluation criteria for the Silhouette coefficient are detailed in Table 3.

Table 3. Evaluation Criteria for Silhouette Coefficient

Silhouette Coefficient (S)	Interpretation
The range between 0.71 - 1	The structure of the cluster is considered excellent.
The range between 0.51 - 0.71	The structure of the cluster is acceptable.
The range between 0.26 - 0.50	The clustering should be improved.
Less than 0.25	The structure of the cluster has no relationship.

Source: Silhouettes, a graphical aid to the interpretation and validation of cluster analysis [14]

4.2) Davies-Bouldin Index

Developed to assess the quality of clustering, the Davies-Bouldin Index (DBI) utilizes Euclidean Distance as a key metric for calculation. The Davies-Bouldin Index clustering process can be computed using the following equation [17].

$$\text{Davies Bouldin} = \frac{1}{n}\sum_{i=1}^{n} \max_{i \neq j} \left(\frac{\sigma_i+\sigma_j}{d(c_i,c_j)}\right) \quad (2)$$

n: The number of clusters., ci: The average distance within cluster i., cj: The average distance within cluster j., σi: The average distance from points in cluster i to its centroid ci. Σj: The average distance from points in cluster j to its centroid cj., d(ci,cj): The distance between the centroids ci and cj.

The Davies-Bouldin Index (DBI) calculation involves determining the Intra-cluster Dispersion ratio. A lower DBI value indicates better clustering quality, as the clusters exhibit less dispersion and are well-separated. Conversely, a higher DBI value suggests that the clusters overlap or cannot be effectively distinguished.

5) Step 5: Visualization

Visualization presents the results in graphical format, including maps illustrating the spatial distribution and clustering of data neighborhoods.

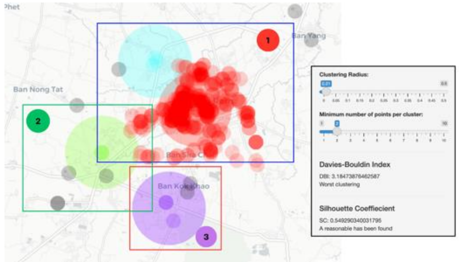
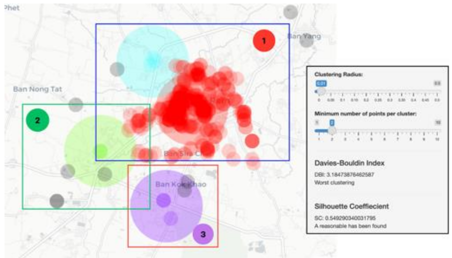
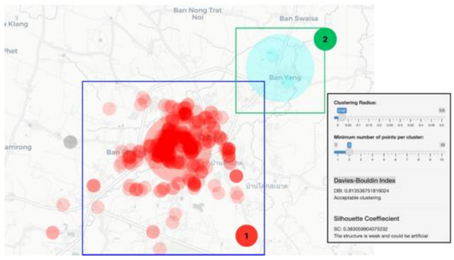
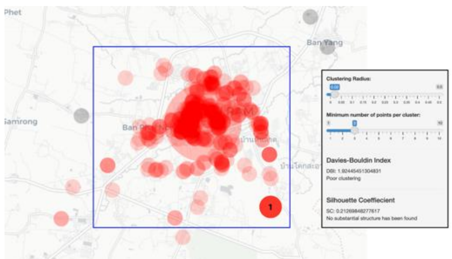
3. Results and Discussion

The spatial clustering of dormitory density using DBSCAN algorithm yielded the following outcomes:

3.1 Results

1) Examining the distribution of dormitories in Mueang District, Buriram Province, using the DBSCAN algorithm with the assistance of the RStudio in conjunction with the R language and the Shiny Web Application involved the visualization of maps depicting the spatial distribution of dormitories in the area. In addition, the clustering results of dormitory neighborhoods were presented. By comparing different parameter values for distance or proximity between points (Eps) and the minimum number of points (MinPts), variations in the spatial distribution and clustering of dormitory neighborhoods in Mueang District, Buriram Province, were observed, as depicted in Table 4.

Table 4. Comparison of Data Distribution across Varying Eps and MinPts Values

Eps	MinPts	Clusters	Davies-Bouldin Index/ Silhouette Coefficient/	Visualization
0.01	2	3	3.185/ 0.549	 (A)
0.01	3	3	2.785/0.495	 (B)
0.02	2	2	0.914/0.383	 (C)
0.02	3	3	1.924/0.213	 (D)

The data analysis from Table 4 reveals the spatial distribution of dormitories in Mueang District, Buriram Province, along with the clustering outcomes :

(1) Setting Eps at 0.01 and MinPts at 2 resulted in the data clustering into 3 clusters. Cluster 1 comprised areas in Nai Mueang and Isan Subdistricts containing 368 data points. Cluster 2 covered the Krasang and Ban Bua Subdistricts, containing 6 data points. Cluster 3 covered areas in the Samed Subdistrict, with 5 data points as depicted in Figure **(**A**)**. The evaluation of the Davies-Bouldin Index (DBI) shows a value of 3.185, indicating poor clustering due to the high value. Additionally, the Silhouette Coefficient (SC) is 0.549, which suggests that the clustering structure is acceptable to some extent but can still be improved.

(2) With Eps set to 0.01 and MinPts to 3, the data formed 3 clusters. Cluster 1 included areas in Nai Mueang and Isan Subdistricts, containing 366 data points. Cluster 2 covered the Krasang and Ban Bua

Subdistricts, containing 5 data points. Cluster 3 covered areas in the Samed Subdistrict, with 6 data points as depicted in Figure (B). The clustering quality assessment using the Davies-Bouldin Index (DBI) yielded a value of 2.785, which is considered high and indicates low clustering performance. This suggests significant overlap among clusters and a lack of clear separation. Additionally, the Silhouette Coefficient (SC) was 0.495, a moderate value that implies the clustering structure is not sufficiently well-defined.

(3) When Eps was set to 0.02 and MinPts to 2, the data clustered into 2 clusters. Cluster 1 included areas in Nai Mueang and Isan Subdistricts, containing 381 data points. While Cluster 2 comprised areas in the Ban Yang Subdistrict, containing 2 data points, as illustrated in Figure (C), The clustering quality assessment using the Davies-Bouldin Index (DBI) yielded a value of 0.914, which is considered acceptable. This indicates that the data points within each cluster are reasonably compact, and the clusters are fairly well-separated. However, the Silhouette Coefficient (SC) was found to be 0.383, which is considered low to moderate, suggesting that the clustering structure is not particularly strong or well-defined.

(4) When Eps was set to 0.02 and MinPts to 3, the data clustered into 1 cluster. Comprising areas in Nai Mueang and Isan Subdistricts contain 381 data points, as illustrated in Figure (D). The clustering quality assessment using the Davies-Bouldin Index (DBI) yielded a value of 1.924, which is relatively high and indicates poor clustering performance. However, the Silhouette Coefficient (SC) was 0.213, indicating the lowest level with an unrelated group structure.

The findings from Table 4 illustrate how varying Eps and MinPts values influence the clustering structure and quality in the spatial distribution of dormitories. As demonstrated, the number of clusters, cluster membership, noise points, and clustering quality metrics such as DBI and Silhouette Coefficient change significantly depending on the parameter settings. Table 5 presents a consolidated overview of clustering outcomes across different Eps and MinPts combinations to highlight these variations further and enable more precise comparisons. This comparison facilitates the identification of parameter settings that yield the most reliable and interpretable clustering structure, both algorithmically and from a human-judgment perspective.

Table 5. Comparison of Data Distribution across Varying Eps and MinPts Values

Eps	MinPts	Clusters	members	noise	Silhouette Coefficient	DBI	Human (Clusters)	Clustering Quality
0.01	2	3	379	8	0.549	3.185	377	Moderate
0.01	3	3	377	10	0.495	2.785	377	Best
0.02	2	2	383	4	0.383	0.914	377	Moderate
0.02	3	1	381	6	0.213	1.924	377	Lowest

Based on the results presented in Table 5, the clustering quality varies significantly depending on the combination of Eps and MinPts values. When Eps was set to 0.01 and MinPts to 2, the clustering yielded the highest Silhouette Coefficient (0.549), indicating that the cluster structure was relatively well-defined. However, this configuration also produced the highest Davies-Bouldin Index (DBI) at 3.185, which implies poor compactness and significant overlap between clusters. Therefore, while the clusters were distinguishable, they lacked cohesion.

In contrast, Eps = 0.01 and MinPts = 3 parameter settings provided a more balanced outcome. Although the Silhouette Coefficient slightly decreased to 0.495, the DBI dropped to 2.785. This suggests that the clusters were both reasonably compact and moderately well-separated. As a result, this configuration can be considered the most appropriate overall, offering a good trade-off between intra-cluster cohesion and inter-cluster separation.

When Eps was increased to 0.02, and MinPts remained at 2, the clustering quality improved compactness, as reflected by the lowest DBI value of 0.914. However, the Silhouette Coefficient decreased to 0.383, indicating a less distinct clustering structure. This setting may be suitable when compact grouping is prioritized over structural clarity.

Finally, the configuration with Eps = 0.02 and MinPts = 3 yielded the weakest clustering performance. It produced only a single cluster with a very low Silhouette Coefficient of 0.213 and a relatively high DBI of

1.924, suggesting no meaningful or substantial cluster structure was identified. In summary, the parameter setting of Eps = 0.01 and MinPts = 3 demonstrates the best overall clustering quality, as it achieves a reasonable balance between the compactness and clarity of clusters, as shown in Figure 4.

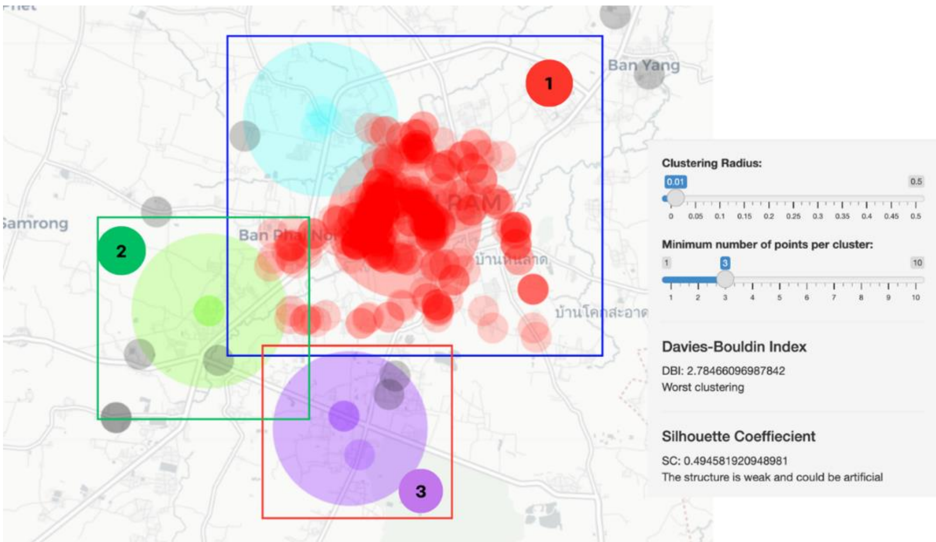


Figure 4. Clustering of Dormitory Neighborhoods in Mueang District, Buriram Province

According to Figure 4, the DBSCAN algorithm performs cluster analyses for each dataset element. The result shows that Cluster 1 (the red color) is a group of dormitories near the university (Buriram Rajabhat University). In cluster 1, the points of interest (POIs) are dormitories distributed around Nai Mueang and Isan Subdistricts. It is a bigger group than another group. This means that the suitable dormitory for students of Buriram Rajabhat University is the dormitory that occurred in cluster 1. The university can use information about dormitory groups to support students when they need to find and plan suitable dormitories. Additionally, students can utilize dormitory group information as a data source to support their decision-making when selecting a rental dormitory.

The densest cluster of dormitory neighborhoods was found in Nai Muang and Isan Subdistricts, with a total of 366 points**.** These points were primarily situated around Jira Road and near educational institutions, bus stations, restaurants, convenience stores, and shopping malls, as shown in Figure 5**.**

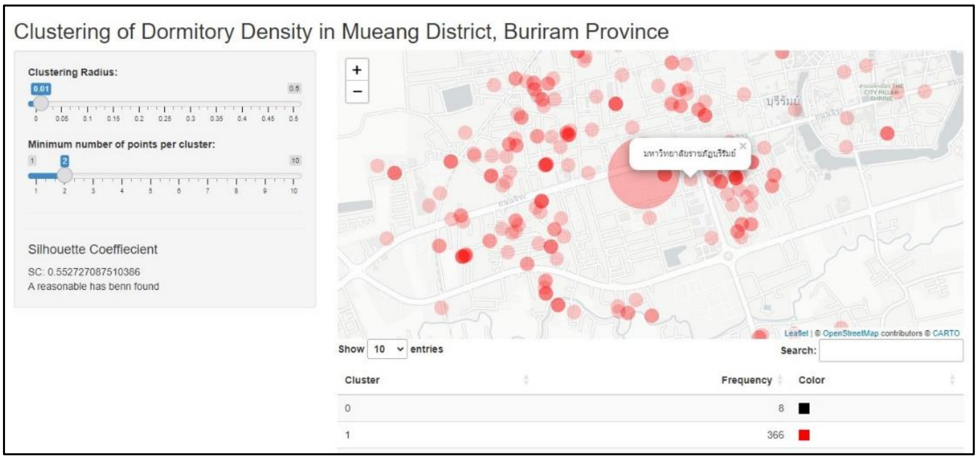


Figure 5 Clustering of High-Density Dormitory Neighborhoods in Mueang District, Buriram Province

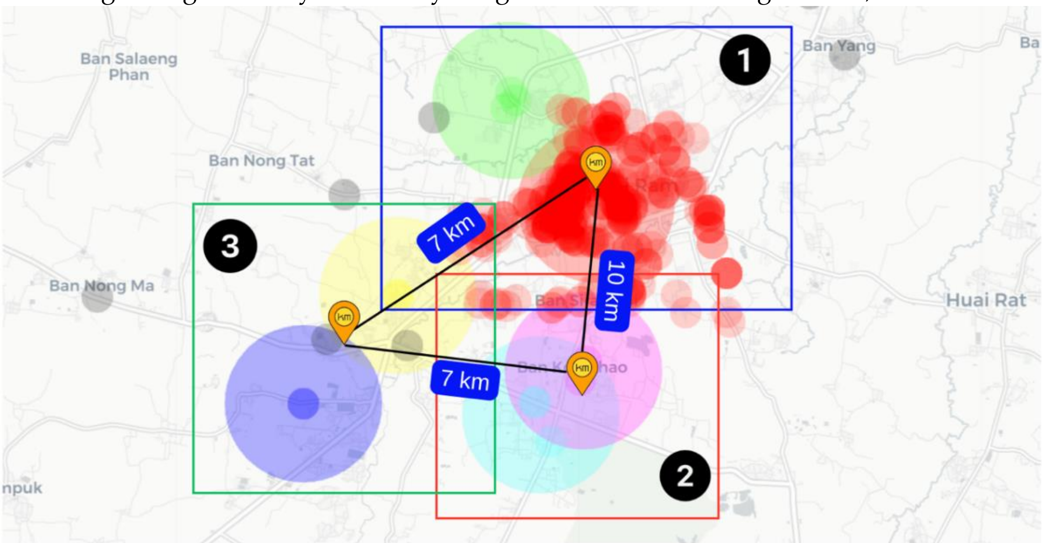


Figure 6. Example of Dormitory Neighborhood Cluster Analysis based on distance

According to Figure 6, the pattern of dormitory distribution was also found that the dormitory neighborhoods have expanded to areas outside the city, specifically in Krasang and Ban Bua Subdistricts, which are approximately 7 kilometers from the city center and 10 kilometers from Samed Subdistrict.

However, the three clusters are not significantly distant, and amenities such as convenience stores, restaurants, and public roads in Krasang, Ban Bua, and Samed Subdistricts are increasing. These clusters can potentially develop into additional dormitory neighborhoods in the future. The point where new dormitories are being constructed can lead to further analysis, providing insights into the potential development of these dormitory areas.

3.2 Discussion

Dormitory neighborhood clustering in Mueang District, Buriram Province, was analyzed using data extracted from Google Maps (https://maps.google.com/) on January 1, 2024. This data served as supplementary information for decision-making processes concerning selecting university dormitories for students at Buriram Rajabhat University and for entrepreneurs planning to invest in private dormitory businesses and selecting suitable construction locations. The clustering process identified three distinct neighborhood clusters: Cluster 1 covers areas in the Nai Muang and Isan Subdistricts, Cluster 2 encompasses areas in the Samed Subdistrict, and Cluster 3 comprises areas in the Krasang and Ban Bua Subdistricts. Additionally, it was found that the distances between the three neighborhood clusters were not significantly distant, and amenities such as convenience stores, restaurants, and public roads emerged in Krasang, Ban Bua, and Samed Subdistricts. These neighborhood clusters can potentially develop into additional dormitory neighborhoods in the future. The points where these new dormitories are being constructed can lead to further analysis, providing insights into the potential development of these dormitory areas. This concept aligns with the research of Fauzan et al. [5], who applied the DBSCAN algorithm to analyze the spatial clustering of hotels in Bali island, demonstrating that clustering hotels near tourist attractions helps managers prioritize economic recovery strategies post-COVID-19 effectively. It is also consistent with the research of Nyoman et al.[8], who conducted spatial data analysis using the DBSCAN algorithm before deciding to develop tourist destination maps based on Bali's tourism strategy, leading to equitable development and cost reduction in spatial development towards becoming tourist destinations. The clustering results offer meaningful insights for urban planning, particularly in identifying high-density dormitory zones that can guide infrastructure development and public service allocation. These insights may assist local authorities and private developers in optimizing the location and accessibility of student accommodations in the Mueang District. Furthermore, the results are consistent with previous studies that applied DBSCAN to hotel distribution in Bali [5] and urban functional areas in Wuhan [8], affirming the algorithm's utility in spatial analysis. However, some limitations should be acknowledged. The dormitory data were collected via web scraping from Google Maps, which may result in biases such as incomplete or outdated records. Despite efforts to clean and validate the dataset, minor inaccuracies may still affect the clustering precision.

4. Conclusion

The paper presents a new approach to identifying dormitory clusters, which uses the DBSCAN clustering algorithm as a model to calculate the clustering density with the shiny web application framework. The research results conclude that the DBSCAN algorithm can serve as a tool for presenting spatial distribution data and clustering dormitory neighborhoods in Mueang District, Buriram Province. This algorithm allows parameter adjustments to align results with the context and needs of the area, facilitating data storage and analysis in spatial research, thereby efficiently reducing time and costs and examining dormitory distribution in Mueang District, Buriram Province, using the DBSCAN algorithm revealing neighborhood clusters characterized by dense and intriguing data distributions. Data clustering resulted in three neighborhoods: Cluster 1 encompassed areas in Nai Muang and Isan Subdistricts, Cluster 2 included areas in Samed Subdistrict, and Cluster 3 comprised areas in Krasang and Ban Bua Subdistricts. Validating the efficiency of clustering dormitory neighborhoods in Mueang District, Buriram Province, indicated good clustering performance, with an acceptable group structure, as reflected by a Silhouette Coefficient of 0.495 and a Davies-Bouldin Index (DBI) value of 2.785. The conclusion shows that the DBSCAN cluster analysis model is effectively characterized to identify dormitory clusters. Moreover, combining DBSCAN clustering algorithms with the shiny web application framework is a more interesting active visualization that can be applied to support information when making decisions. The proposed approach can be used as a model to develop an information system to support students who would like to find and make decisions about suitable dormitories based on spatial distribution.

5. Recommendations

The spatial clustering analysis is based on dormitory density in Mueang District, Buriram Province, utilizing the DBSCAN algorithm. Policy Recommendations: The DBSCAN algorithm is a robust tool for presenting spatial clustering data regarding dormitory density in Mueang District, Buriram Province. Tailoring parameter adjustments to suit the contextual and regional requirements facilitates its application in spatial research. This, in turn, can streamline data storage and analysis processes, thereby effectively reducing time and costs. Operational Research Recommendations: The DBSCAN algorithm facilitates the visualization and analysis of neighborhood distribution and groupings concerning dormitory data. This analytical capability can be harnessed to develop decision support systems and conduct predictive analyses using artificial intelligence methodologies, particularly for urban development and expansion initiatives.

6. Acknowledgments

We sincerely thank the Information Technology and Statistics & Information Science Programs (Faculty of Science, Buriram Rajabhat University) for providing essential research tools. We also sincerely thank the BRU GoScopus II project organizers for their invaluable content guidance and unwavering support in developing and publishing our research article. Their contributions significantly aided this project's success.

References

[1] Arifoğulları, Ö.; Orman, G. K. (2023, July 5–7). *Experimental Study of Hotel Clustering* [Conference session]. IMECS 2023: International MultiConference of Engineers and Computer Scientists, Hong Kong. https://www.iaeng.org/publication/IMECS2023/IMECS2023_pp82-87.pdf

[2] Caudillo-Cos, C. A.; Ruiz-Pérez, J. A. Defining urban boundaries through DBSCAN and Shannon's entropy: The case of the Mexican National Urban System. *Cities*. **2024**, *149*, Article 104969. https://doi.org/10.1016/j.cities.2024.104969

[3] Cherfi, H.; Napoli, A.; Toussaint, Y. Towards a Text Mining Methodology Using Association Rules Extraction. *Soft Computing.* **2006,** *10*(5), 431–441. https://doi.org/10.1007/s00500-005-0504-x

[4] Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. (1996, August 2–4). *A density-based algorithm for discovering clusters in large spatial databases with noise* [Conference session]. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, OR, United States. https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf.

[5] Fauzan, A.; Novianti, A.; Ramadhani, R. M. A.; Adhiwibawa, M. A. S. Analysis of hotels spatial clustering in Bali: Density-Based Spatial Clustering of Application Noise (DBSCAN) algorithm approach. *EKSAKTA* **2022**, *3*(1), 25–38. https://doi.org/10.20885/EKSAKTA.vol3.iss1.art4

[6] Gammanee, S.; Rattananonsathien, N.; Klaprabchone, K. Model development to increase skills in using artificial intelligence to add value to work and workability of workers in Kanchanaburi Province. *Journal of Administrative and Management Innovation.* **2023**, *11*(1), 45–53.

[7] Janvattanavong, P.; Muenchaisri, P. (2020, November 3). Software remodularization using Tabu search [Conference session]. ICICSE 2020: International Conference on Internet Computing for Science and Engineering, Singapore. https://doi.org/10.1145/3424311.3424321

[8] Jaya, N. M.; Wiryasa, N. M. A.; Sudarsana, D. K.; Salain, P. D. P. (2018, October 25–26). *Mapping the potential for tourism strategic areas to improve the equality of development in Bali* [Conference session]. ICAnCEE 2018: International Conference on Applied Engineering, Yogyakarta, Indonesia. https://doi.org/10.1051/matecconf/201927602008

[9] Tu, X.; Fu, C.; Huang, A.; Chen, H.; Ding, X. DBSCAN spatial clustering analysis of urban "production–living–ecological" space based on POI data: A case study of central urban Wuhan, China. *International Journal of Environmental Research and Public Health*. **2022**, *19*(9), 5153. https://doi.org/10.3390/ijerph19095153

[10] Office of Academic Affairs and Registration. (2023, August). Report on the number of students. Buriram Rajabhat University. https://oapr.bru.ac.th/wp-content/uploads/2023/08/%E0%B8%88%E0%B8%B3%E0%B8%99%E0%B8%A7%E0%B8%99-1-66.pdf

[11] Paesrivarotai, O.; Tanaksaranond, G. (2021, June 23–25). *Study of active fire evolution in Northern Thailand using Kernel Density Estimation (KDE)* [Conference session]. The 26th National Convention on Civil Engineering (NCCE 26), Online Conference, Thailand. https://www.researchgate.net/publication/354506941.

[12] Pensiri, F.; Visutsak, P.; Chaowalit, O. (2021, November 15–17). *Clustering tourist using DBSCAN algorithm* [Conference session]. 2021 Asia-Pacific Conference on Applied Mathematics and Statistics (AMAS 2021), Bangkok, Thailand. https://doi.org/10.1063/5.0082995

[13] Phobphimail, S.; Janruang, J. (2021, August 19–20). *Applied DBSCAN clustering algorithm to identify tourism and hospitality zone: Case study of Sikhoraphum District at Surin Province* [Conference session]. The 10th National and the 3rd International Conference on Management Sciences: Innovation Management for Enhancing the Local Economy, Ubon Ratchathani University, Thailand.

[14] Rousseeuw, P. J.; Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics.* **1987**, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

[15] Yoyram, J.; Enkvetchakul, P.; Boonkate, K. Deep Learning with Convolutional Neural Networks for Cassava Leaf Diseases Via Line Bot: A Case Study of Buriram Provincial Protection Service Group, Buriram Provincial Agriculture Office. *Harbin Gongcheng Daxue Xuebao/Journal of Harbin Engineering University*. **2023**, *44*(7). https://harbinengineeringjournal.com/index.php/journal/article/view/691

[16] Yoyram, J.; Janruang, J.; Karnka, S. The study of the distribution and clustering of tourist attraction zones and facilities in Buriram Province using the DBSCAN algorithm. *Journal of Information Science Research and Practice*. **2024**, *42*(3), July–September. https://doi.org/10.14456/jiskku.2024.17

[17] Moonpen, U.; Mungsing, S.; Banditwattanawong, T. Clustering algorithm optimization model for essential attribute analysis of tour package forms inbound tourism market in Thailand. *The Journal of KMUTNB*. **2020**, *30*(4), 656–667. https://doi.org/10.14416/j.kmutnb.2020.04.006