MIKAILALSYS
Journal of Multidisciplinary Sciences

e-ISSN : 2987-2286
p-ISSN : 2987-3924

Index: Harvard, Boston, Sydney, Copernicus, Sinta 4, Scilit, Semantic, Google, etc

https://doi.org/10.58578/mikailalsys.v4i2.10187

Systematic Literature Review: Population Density Mapping Using Data Mining

Muhammad Rafli, Besse Hartati, Gunawan Ari Nursanto
Politeknik Imigrasi dan Pemasyarakatan, Indonesia
rafli1607200@gmail.com; hartatimassaile@yahoo.com

Article Info:

Submitted:	Revised:	Accepted:	Published:
Apr 5, 2026	May 3, 2026	May 15, 2026	May 20, 2026

Abstract

Population density is a critical indicator in regional development planning because it is closely associated with public service distribution, transportation systems, healthcare provision, and environmental management. The rapid growth of digital technology has increased the volume and complexity of demographic data, requiring more effective analytical methods for population density analysis. This study aims to analyze the application of data mining in population density mapping based on studies published between 2021 and 2025. A systematic literature review approach was employed by examining 30 scientific articles obtained from Google Scholar, Semantic Scholar, and Crossref. The review process included article identification, literature screening, data collection, and analysis of findings based on the algorithms, research fields, data sources, and analytical methods used in the selected studies. The findings indicate that the most frequently applied algorithms were K-Means Clustering, DBSCAN, and Density Peaks Clustering. Data Mining and Machine Learning emerged as the dominant research field, representing 50% of the analyzed articles. The primary data sources used in the reviewed studies included public datasets, government data, and spatial imagery. The results also show that clustering was the most commonly applied analytical method in population density analysis. These findings demonstrate that clustering techniques are effective for supporting population density mapping and identifying spatial data distribution patterns relevant to regional decision-making. The study contributes to demographic and regional development research by synthesizing recent evidence on the role of data mining in population density mapping and providing a reference for future studies on spatial demographic analysis.

Keywords: Clustering; Data Mining; DBSCAN; K-Means; Population Density Mapping

INTRODUCTION

Population density is an essential indicator in regional development planning because it is closely related to resource distribution, public service delivery, and community welfare. A high concentration of population within a particular area significantly affects transportation systems, healthcare services, education, environmental management, and social infrastructure. Accurate population distribution data enables governments and policymakers to formulate more effective and evidence-based development strategies. However, the dynamic and heterogeneous nature of population distribution creates significant challenges in producing precise spatial mapping and demographic analysis (Zhang & Li, 2022).

The rapid advancement of digital technology has led to a substantial increase in demographic data generated through population censuses, administrative registration systems, digital governance platforms, and community-based information systems. These demographic datasets exhibit big data characteristics, including high volume, data variety, and rapid data growth. Consequently, conventional analytical approaches are often insufficient for processing and interpreting such complex datasets efficiently. Therefore, the implementation of data mining and machine learning techniques has become increasingly important in supporting demographic clustering and population density analysis ("A Novel Oversampling Method for Imbalanced Datasets Based on Density Peaks Clustering," 2021).

In recent years, clustering algorithms have experienced significant development across various scientific disciplines. Clustering methods are widely utilized to group data based on similar characteristics, enabling more systematic and efficient decision-making processes. The application of clustering is no longer limited to information technology but has expanded into healthcare, agriculture, environmental studies, governance systems, population analysis, and energy management. Through clustering techniques, researchers are able to identify hidden patterns from large-scale and complex datasets, thereby generating meaningful insights for analytical and policy-related purposes (Wang et al., 2023).

The growing implementation of data mining demonstrates that analytical technologies have become a critical component in scientific development and data-driven decision-making. Studies published between 2021 and 2025 indicate that research involving machine learning and data mining has become increasingly dominant across multiple academic fields. These studies utilize highly diverse datasets, including public datasets, government records, spatial imagery, simulation-based data, and experimental datasets. This trend reflects the expanding role of data analytics in addressing complex societal and developmental challenges (Wang et al., 2023).

Population density is also strongly associated with social and environmental risks in modern society. Highly populated regions are generally more vulnerable to disease transmission, traffic congestion, limited green spaces, environmental pollution, and unequal access to public services. Such issues become more severe in rapidly urbanizing areas where population growth exceeds the capacity of spatial planning and infrastructure development. Consequently, population density mapping has become increasingly important for achieving sustainable regional development (Du et al., 2024).

Urbanization is another major factor contributing to changes in population distribution across regions. The migration of people from rural to urban areas places significant pressure on urban infrastructure, transportation systems, housing availability, and social services. This phenomenon not only affects economic conditions but also influences environmental quality and social welfare. In this context, demographic data analysis plays a crucial role in helping policymakers understand population distribution patterns and establish development priorities more accurately (Chen et al., 2022).

The implementation of data mining in population density analysis provides opportunities for producing more accurate and efficient regional mapping systems. Data mining refers to the process of extracting valuable information and hidden patterns from large datasets to support analytical interpretation and decision-making. One of the most widely used methods in data mining is clustering analysis, which groups data based on similarities among objects or variables. Through clustering techniques, regions with similar population density characteristics can be categorized systematically, thereby simplifying spatial analysis and regional planning processes (Han et al., 2021).

Various clustering algorithms have been developed, each possessing distinct characteristics and analytical advantages. Algorithms such as K-Means, DBSCAN, Hierarchical Clustering, and Density Peaks Clustering are frequently employed in spatial and demographic studies. K-Means remains one of the most popular clustering algorithms due to its computational simplicity and efficiency in handling large datasets. Nevertheless, the algorithm is highly sensitive to outliers and requires a predefined number of clusters, which may reduce clustering accuracy in complex datasets (Jain & Kumar, 2021).

In contrast, DBSCAN identifies clusters based on data density and is more effective in recognizing irregular distribution patterns. This algorithm is also capable of detecting noise and outlier data that commonly appear in demographic datasets. However, DBSCAN requires careful parameter selection, particularly regarding epsilon values and minimum points, to achieve optimal clustering performance (Liu et al., 2022).

Recent studies further demonstrate the increasing adoption of density-based clustering approaches due to their flexibility in handling complex datasets. Density Peaks Clustering, for example, identifies cluster centers based on local density and relative distance among data points. This method is considered effective in overcoming the limitations of conventional clustering algorithms, particularly when analyzing non-linear distributions and datasets with varying density characteristics ("A Novel Oversampling Method for Imbalanced Datasets Based on Density Peaks Clustering," 2021).

Despite the growing number of clustering-related studies, most previous research has primarily focused on technical algorithmic performance without comprehensively examining its application in population density analysis and regional development planning. Several studies relied on limited datasets and narrow geographical scopes, making their findings less representative of broader demographic conditions. Furthermore, earlier research often emphasized algorithmic accuracy comparisons without evaluating the practical relevance of clustering results for public policy and regional planning purposes (Rahman et al., 2022).

Based on these conditions, a significant research gap can be identified in the limited integration of clustering analysis with population density studies within the context of regional development. Previous studies have predominantly concentrated on algorithmic model development rather than investigating the practical utilization of clustering techniques as decision-support tools. Moreover, only a limited number of studies have systematically examined recent trends in data mining research, particularly regarding methodological developments, data sources, and interdisciplinary applications (Kim & Park, 2023).

This study attempts to address these gaps by conducting a literature-based investigation of clustering and data mining research published between 2021 and 2025. The study not only identifies the most frequently used algorithms but also analyzes research distribution, data sources, methodological approaches, and technological trends in data mining applications across multiple disciplines. Therefore, this research provides a more systematic understanding of the development of analytical technologies during the selected period.

The novelty of this study lies in its comprehensive approach that combines trend analysis in data mining research with the examination of clustering applications in population density analysis and regional development. Unlike previous studies that primarily focused on technical algorithmic evaluations, this research also investigates the relevance of clustering applications in supporting evidence-based policymaking in the public sector. Furthermore, the study employs a systematic literature review approach to identify research patterns in a more structured and objective manner (Snyder, 2019).

Theoretically, this study is grounded in the concepts of data mining, machine learning, and clustering analysis as the primary analytical framework. Data mining is utilized to extract information from large datasets, while clustering analysis is employed to classify data based on similarities in characteristics and patterns. These approaches facilitate the identification of distribution patterns and support systematic interpretation processes for data-driven decision-making (Han et al., 2021).

Additionally, this study incorporates concepts related to regional development and population density to understand the relationship between demographic distribution and public service requirements. Population density is considered a crucial indicator for evaluating spatial quality, infrastructure distribution, and the effectiveness of public services. Consequently, the utilization of analytical technologies in population mapping is highly relevant for promoting adaptive and sustainable regional development strategies (Du et al., 2024).

This research focuses on analyzing the development of clustering and data mining studies related to population density and regional decision-making processes during the 2021–2025 period. The study emphasizes the identification of commonly used algorithms, research fields, data sources, and methodological approaches employed in scientific publications. Through this analysis, the research aims to provide a comprehensive overview of technological developments in data analytics for demographic management and regional planning.

The objectives of this study are to identify trends in the use of clustering algorithms in data mining research, analyze the distribution of research fields and data sources, and evaluate methodological developments during the 2021–2025 period. In addition, the study seeks to provide recommendations regarding future research opportunities in demographic analysis and regional development based on data mining technologies.

The findings of this study are expected to contribute both theoretically and practically. From a theoretical perspective, this research enriches the academic discussion regarding the development of clustering and data mining within demographic studies. Practically, the findings may serve as a reference for researchers, academics, and policymakers in developing analytical systems that support evidence-based regional planning. Ultimately, this study is expected to contribute to scientific advancement and the implementation of analytical technologies for more effective population management.

METHODS

This study employed a qualitative approach using the Systematic Literature Review (SLR) method. The qualitative approach was selected because the research focused on identifying, interpreting, and analyzing scientific findings related to the application of clustering algorithms and data mining techniques in population density analysis. The study was descriptive in nature, aiming to systematically describe the development of data analysis research published between 2021 and 2025. A literature review approach was considered appropriate because it enabled the researcher to synthesize findings from previous studies and generate a comprehensive understanding of algorithm trends, dominant research fields, analytical methods, and the development of data mining applications across multiple sectors.

The research was conducted from January to May 2025 and required approximately five months to complete. The research stages included topic identification, scientific article retrieval, literature selection, data collection, data analysis, and report preparation. During the research process, the researcher identified scientific publications related to clustering, machine learning, data mining, and population density mapping.

The research design was systematically developed based on the stages of a systematic literature review, including literature identification, article screening, eligibility assessment, data extraction, and synthesis of findings. These stages were implemented to ensure that the selected articles were highly relevant to the research objectives and possessed adequate scientific quality. Through this design, the study aimed to obtain a comprehensive understanding of the development of clustering methods and data analysis techniques within the 2021–2025 publication period.

Research indicators were established as the primary basis for data collection and analysis. These indicators were developed based on theories of data mining, machine learning, and clustering analysis that served as the conceptual foundation of the study. The first indicator concerned the types of clustering algorithms utilized in the selected studies, including K-Means, DBSCAN, Hierarchical Clustering, and Density Peaks Clustering. This indicator was intended to identify the most dominant algorithms used in data analysis research. The second indicator focused on research fields applying clustering and data mining techniques, such as healthcare, environmental studies, information systems, population analysis, agriculture, and energy. The third indicator involved the analytical methods applied in the studies, including clustering, hybrid machine learning, statistical analysis, experimentation, and simulation modeling. The fourth indicator examined the types of data sources used in the articles, including public datasets, government data, spatial imagery, simulation data, and experimental datasets.

The formulation of these indicators was intended to address the previously established research questions. The first research question focused on identifying the most frequently used algorithms in data analysis research during the 2021–2025 period. The second research question examined the dominant research fields applying clustering and data analysis methods. The third research question aimed to identify the most commonly used analytical methods in scientific publications.

The participants in this study were not individuals or respondents but scientific articles used as research data sources. The articles were obtained from national and international journals published between 2021 and 2025. A total of 30 scientific articles meeting the inclusion criteria were selected for analysis. These articles were retrieved from several scientific databases, namely Google Scholar, Semantic Scholar, and Crossref.

The sampling technique employed in this study was purposive sampling. This technique was selected because the articles were intentionally chosen based on specific characteristics relevant to the research objectives. The selected articles were closely related to data mining, clustering techniques, and population density analysis. The use of purposive sampling enabled the researcher to obtain highly relevant data sources and conduct more in-depth analysis.

Several inclusion criteria were established to ensure the quality and relevance of the selected literature. The articles had to be scientific publications written in either Indonesian or English, published between 2021 and 2025, and discuss population density mapping or the implementation of data mining techniques in demographic analysis. In addition, the articles were required to contain information regarding algorithms, analytical methods, data sources, and research findings to support the literature review process.

The article retrieval process was conducted using the keywords "Data Mining on Population Density" and "Mapping Clustering for Population Density." These keywords were used to identify publications relevant to the research topic. The initial search produced a large number of articles from various scientific databases. Subsequently, a preliminary screening process was carried out based on titles and abstracts to determine article relevance. Articles that passed the initial screening stage were then evaluated through full-text review to ensure consistency with the research focus.

The research instruments consisted of documentation sheets and data extraction tables. Documentation sheets were used to record article identity, author names, publication year, research methods, algorithms, data sources, research fields, and findings. Meanwhile, the data extraction table facilitated the organization and classification of information obtained from each article, thereby simplifying the analysis and interpretation processes.

Data validity in this literature review study was maintained through source triangulation and article eligibility evaluation. The researcher compared findings from multiple studies to ensure data consistency and minimize potential bias. Furthermore, the selected articles were evaluated based on publication quality, topic relevance, and completeness of information. This evaluation process was conducted systematically to ensure that the selected literature adequately supported the research objectives.

Data collection techniques involved documentation and systematic literature exploration. The first stage consisted of identifying literature sources through scientific databases. The second stage involved article screening based on titles and abstracts. The third stage focused on full-text evaluation to ensure article relevance to the research objectives. The fourth stage involved data extraction by recording important information from the selected articles. All collected data were subsequently organized into tables to facilitate classification and analytical interpretation.

The data analysis technique used in this study was descriptive qualitative analysis. The analysis was conducted by classifying data according to the predefined research indicators. The analytical stages included data reduction, data presentation, interpretation, and conclusion drawing. During the data reduction stage, the researcher filtered information relevant to the research focus. Subsequently, the collected data were presented in tables and narrative descriptions to simplify interpretation.

In the interpretation stage, the researcher analyzed clustering algorithm trends, research field distribution, data sources, and dominant analytical methods used in studies published between 2021 and 2025. Comparative analysis among the selected articles enabled the researcher to identify patterns in the development of data mining and clustering applications across multiple disciplines. The final stage involved drawing conclusions based on the systematic synthesis of research findings.

Through this research methodology, the study is expected to provide a comprehensive overview of the development of clustering algorithms and data mining applications in population density research and related scientific fields. Furthermore, the systematic literature review approach employed in this study enables the production of a more objective, structured, and scientifically relevant synthesis regarding recent developments in data analysis research.

RESULTS

The findings were obtained through a systematic review of 30 scientific journals published between 2021 and 2025 regarding the application of data mining in population density mapping. The analyzed articles were collected from Google Scholar, Semantic Scholar, and Crossref, focusing on clustering algorithms, analytical methods, data sources, and dominant research domains.

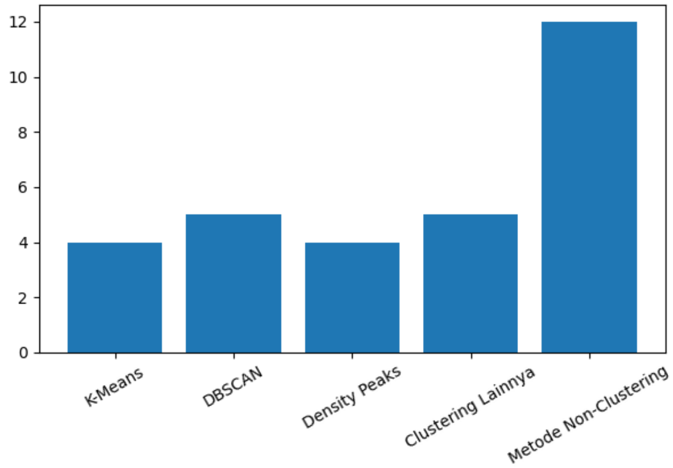

Figure 1. Algorithm Usage

Figure 1. algorithm usage results show that clustering methods were the most frequently applied approaches in population density studies. The dominant algorithms identified were K-Means Clustering, DBSCAN, and Density Peaks Clustering. The prevalence of these methods demonstrates that clustering techniques remain the primary approach for analyzing spatial population distribution patterns.

K-Means Clustering emerged as the most widely used algorithm because of its simplicity, computational efficiency, and ability to classify regions into low-, medium-, and high-density categories. Meanwhile, DBSCAN was extensively implemented because it can detect irregular spatial distribution patterns without requiring a predefined number of clusters. Density Peaks Clustering was applied to improve the accuracy of density center identification and reduce noise influence during the clustering process.

In addition to clustering methods, several studies also utilized Kernel Density Estimation (KDE), Random Forest, Linear Regression, Fuzzy C-Means, and Cellular Automata. Kernel Density Estimation was mainly used to visualize spatial and temporal changes in population density, whereas Random Forest and Linear Regression were employed to predict population growth trends based on socio-economic indicators.

Table 1. Distribution of Research Fields

No	Research Field	Number of Journals
1	Data Mining / Machine Learning	15
2	Information Systems / Computer Technology	3
3	Agriculture / Biology	3
4	Demography	3
5	Healthcare / Medical	2
6	Energy and Environment	1
7	Astronomy / Astrophysics	1
8	Ecology / Zoology	1
9	Herbal Industry / Intellectual Property	1

Source: Compiled by the researcher (2026)

Table 1 shows that the field of Data Mining and Machine Learning dominated the reviewed studies, accounting for 15 journals or 50% of the total publications analyzed. This finding indicates that data analysis technologies are increasingly applied across multiple disciplines, particularly in spatial and demographic studies. In addition, the results demonstrate that clustering and machine learning approaches have become the primary methods for analyzing population distribution patterns and supporting data-driven decision-making processes.

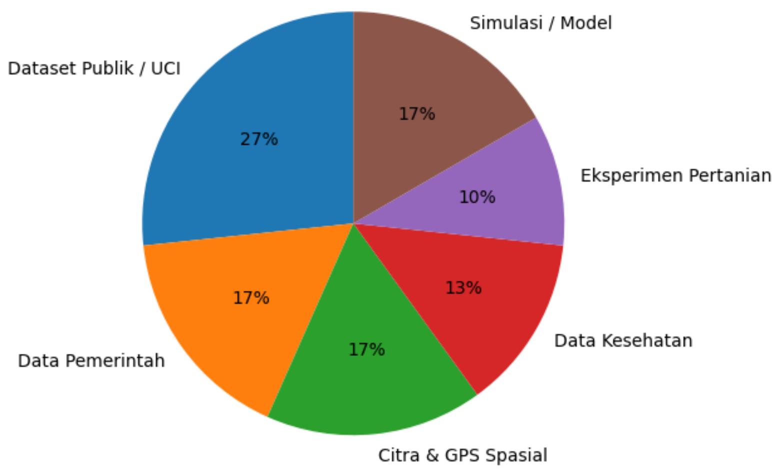

Figure 2. Distribution of Algorithm Usage

Figure 2. shows the distribution of data sources used in the reviewed studies. The results reveal considerable variation in the types of datasets utilized, including public datasets, government databases, satellite imagery, GIS platforms, and GPS-based spatial

data. This finding indicates that population density research increasingly integrates geospatial technologies and diverse data sources to support more accurate spatial analysis and mapping processes.

Table 2. Research Methods Used

No	Research Method	Total
1	Clustering	18
2	Hybrid Machine Learning	4
3	Statistical Analysis	3
4	Experimental Methods	3
5	Simulation/Modeling	2

Source: Compiled by the researcher (2026)

Table 2 indicates that clustering methods dominated the reviewed studies with a total of 18 journals. This demonstrates that clustering techniques remain highly relevant in data-driven research because they effectively identify distribution patterns and relationships among variables. In addition, the findings reveal that clustering approaches are widely utilized in population density mapping due to their ability to classify spatial data into meaningful groups and support more accurate regional analysis.

DISCUSSION

The findings demonstrate that data mining techniques make significant contributions to population density mapping. The dominance of clustering algorithms such as K-Means, DBSCAN, and Density Peaks Clustering indicates that data grouping approaches remain highly effective in identifying spatial population distribution patterns. This finding supports Wang et al. (2022), who emphasized that clustering methods are highly effective for detecting spatial hotspots and grid-based distribution patterns.

K-Means Clustering emerged as the most dominant algorithm due to its simple computational process and efficiency in handling large-scale datasets. The algorithm effectively classifies regions into low-, medium-, and high-density categories, thereby facilitating spatial data interpretation. This result aligns with Delia et al. (2025a), who demonstrated that K-Means is effective for population density classification within administrative regions.

However, K-Means has limitations in handling complex and irregular spatial distributions. Consequently, several studies have shifted toward density-based clustering approaches such as DBSCAN and Density Peaks Clustering. DBSCAN offers greater flexibility because it can identify high-density regions without requiring a predefined number of clusters. Furthermore, the method can detect low-density areas as noise, which are often overlooked by K-Means. This finding is consistent with Lv and Liu (2023), who argued that DBSCAN performs better in handling non-linear spatial distributions.

The implementation of Density Peaks Clustering also reflects recent advancements in population density studies. This method determines cluster centers based on density levels, resulting in higher clustering accuracy. The approach supports the findings of Xu et al. (2023), who reported that density-distance indices improve automatic cluster center identification.

Beyond clustering methods, the study also demonstrates that Kernel Density Estimation, Random Forest, and Cellular Automata play important roles in population density analysis. Kernel Density Estimation is effective for visualizing temporal changes in population density, thereby supporting urbanization and urban growth analysis. Meanwhile, Random Forest and Linear Regression are primarily utilized to predict population growth trends based on socio-economic variables.

The implications of this study indicate that data mining applications can support governments and regional planners in evidence-based decision-making. Information regarding population density patterns can assist in infrastructure planning, healthcare service allocation, educational development, and urbanization control. Furthermore, clustering approaches can also be utilized to identify areas vulnerable to crime, traffic congestion, and disease transmission.

From a theoretical perspective, this study reinforces the concept that clustering techniques are effective approaches in spatial data mining analysis. The findings also indicate that density-based clustering methods are increasingly adopted because they overcome several limitations of conventional clustering techniques.

Nevertheless, this study has several limitations. First, the review analyzed only 30 journals, which limits the overall literature coverage. Second, the study focused exclusively on articles published between 2021 and 2025, meaning that longer-term developments in data mining methodologies were not fully explored. Third, the literature review approach makes the findings highly dependent on the quality and completeness of the analyzed articles.

CONCLUSION

Based on the systematic literature review of 30 scientific journals published between 2021 and 2025, this study demonstrates that data mining techniques provide significant contributions to population density analysis and mapping. The applied methods are not only capable of classifying regions according to population density levels, but also effective in predicting, visualizing, and modeling spatial and temporal population distribution changes. The findings reveal that clustering approaches remain the most dominant methods in population density studies, particularly K-Means Clustering, DBSCAN, and Density Peaks Clustering.

K-Means Clustering emerged as the most widely used algorithm because it is considered simple, efficient, and capable of categorizing regions into low-, medium-, and high-density groups. Meanwhile, DBSCAN demonstrated superior capability in identifying irregular spatial distribution patterns and detecting noise within low-density areas. Density Peaks Clustering was applied to improve the accuracy of cluster center identification based on density levels. In addition, Kernel Density Estimation played an important role in visualizing population density changes, whereas Random Forest and Linear Regression were frequently utilized to predict population growth trends based on socio-economic factors. Cellular Automata and Fuzzy C-Means also provided more adaptive approaches for understanding population density dynamics.

This study successfully addressed the research objectives regarding the most dominant algorithms, the research fields that most frequently apply data mining techniques, and the analytical approaches commonly used in population density mapping. The findings indicate that Data Mining and Machine Learning constituted the most dominant research field, accounting for 50% of the analyzed journals. This result demonstrates that data analysis technologies are increasingly implemented across multiple disciplines, particularly in spatial and demographic studies.

The contribution of this study lies in providing a comprehensive literature synthesis concerning the development of data mining methods for population density mapping during the 2021–2025 period. The study also presents an overview of algorithm trends, data sources, and analytical approaches utilized in spatial-based research. Practically, the findings may serve as a reference for researchers, governments, and regional planners in selecting appropriate analytical methods to support evidence-based decision-making processes.

Nevertheless, this study has several limitations because it only analyzed 30 journals and focused exclusively on publications from 2021 to 2025. Therefore, future studies are recommended to include broader literature coverage, extend the research period, and develop more in-depth comparative analyses of algorithm performance in order to obtain a more comprehensive understanding of the effectiveness of data mining methods in population density mapping.

REFERENCES

Almalki, A., Gokaraju, B., Acquaah, Y., & Turlapaty, A. (2022). Regression analysis for COVID-19 infections and deaths based on food access and health issues. *Healthcare, 10*(2), 324. https://doi.org/10.3390/healthcare10020324

Antonini, F., Gieles, M., Dosopoulou, F., & Chattopadhyay, D. (2023). Coalescing black hole binaries from globular clusters: Mass distributions and comparison to gravitational wave data from GWTC-3. *Monthly Notices of the Royal Astronomical Society, 522*(1), 466–476. https://doi.org/10.1093/mnras/stad972

Barnea, G., & Barnea, N. (2023). On the population density limit to variable renewable energy potential. *Open Access Journal of Applied Science and Technology, 1*(2), 60–67. https://doi.org/10.33140/OAJAST.01.02.03

Cao, J., & Shi, Y. (2021). A novel oversampling method for imbalanced datasets based on density peaks clustering. *Tehnički vjesnik, 28*(6), 1813–1819. https://doi.org/10.17559/TV-20210608123522

Carson, B. D., Orians, C. M., & Crone, E. E. (2023). When does spatial clustering in caterpillars influence the relationship between population density and contact rates? *Research Square*. https://doi.org/10.21203/rs.3.rs-2842538/v1

Delia, F., Rasmita Ngemba, H., Hendra, S., Syahrullah, S., & Trezandy Lapatta, N. (2025). Penerapan Algoritma K-Means Clustering dalam Pengelompokkan Kepadatan Penduduk: Application of K-Means Clustering Algorithm in Population Density Grouping. *Technomedia Journal, 9*(3), 374–386. https://doi.org/10.33050/tmj.v9i3.2270

Du, X., Li, F., Li, X., & Yu, G. (2024). ADPSCAN: Structural graph clustering with adaptive density peak selection and noise re-clustering. *Applied Sciences, 14*(15), 6660. https://doi.org/10.3390/app14156660

Duan, H., Shi, R., Kang, J., Banaschewski, T., Bokde, A. L. W., Büchel, C., Desrivières, S., Flor, H., Grigis, A., Garavan, H., Gowland, P. A., Heinz, A., Brühl, R., Martinot, J.-L., Martinot, M.-L. P., Artiges, E., Nees, F., Papadopoulos Orfanos, D., Paus, T., … IMAGEN Consortium. (2024). Population clustering of structural brain aging and its association with brain development. *eLife*. https://doi.org/10.7554/eLife.94970.1

Forrest, D., & Medyńska-Gulij, B. (2021). Which mapping technique for population density is effective, attractive, and suggestive? *Abstracts of the ICA, 3*, 1–1. https://doi.org/10.5194/ica-abs-3-82-2021

Heningtyas, Y., Rahmi, F., & Muludi, K. (2022). Implementasi Density-Based Clustering pada Segmentasi Citra Betta Fish. *Jurnal Teknoinfo, 16*(1), 8–13. https://doi.org/10.33365/jti.v16i1.1273

Jangada, D., & Zafar, M. Q. (2025). Patent watch and IP mapping for botanicals–polyherbal combinations clustering by condition. *International Journal of Innovative Research in Technology, 12*(5), 1899–1931. https://doi.org/10.64643/IJIRTV12I5-185529-459

Jiang, X., Zhou, Y., & Ma, Y. (2025). Electricity package recommendation integrating improved density peaks clustering and fuzzy group decision-making. *Applied Sciences, 15*(22), 11875. https://doi.org/10.3390/app152211875

Lassoued, Z., & Abderrahim, K. (2023). Extension of the clustering identification by extending the Density Based Spatial Clustering of Applications with Noise approach to multi-input multi-output piece wise affine systems: Application to an industrial robot. *Journal of Control Engineering and Applied Informatics, 25*(2), 25–34. https://doi.org/10.61416/ceai.v25i2.8523

Li, R. (2023). DBSCAN-based line density clustering algorithm for CAD architectural drawings. *Applied and Computational Engineering, 19*(1), 109–115. https://doi.org/10.54254/2755-2721/19/20231018

Lv, Z., & Liu, Q. (2023). Distance-weighted K-means clustering algorithm based on local density. In *Proceedings of the 4th International Conference on Artificial Intelligence and Computer Engineering* (pp. 90–96). ACM. https://doi.org/10.1145/3652628.3652643

Mondal, M. (2024). Spatio-temporal mapping of population density and their characteristics of the Eastcentral Rajmahal Highlands, Jharkhand, India. *International Journal for Multidisciplinary Research, 6*(2). https://doi.org/10.36948/ijfmr.2024.v06i02.15473

Musa, M. A., Shuaib, H., & Ogidan, M. E. (2023). Effect of population density and varieties on the growth and yield of groundnut (*Arachis hypogaea* L.). *Journal of Ecology & Natural Resources, 7*(3). https://doi.org/10.23880/jenr-16000348

Nurmadani, N., & Rakhmawati, F. (2023). Clustering flood prone areas in Deli Serdang Regency using Density-Based Spatial Clustering of Application with Noise (DBSCAN) method. *Desimal: Jurnal Matematika, 6*(2), 261–272. https://doi.org/10.24042/djm.v6i2.19043

Parvin, F., Ali, S. A., Hashmi, S. N. I., & Ahmad, A. (2021). Spatial prediction and mapping of the COVID-19 hotspot in India using geostatistical technique. *Spatial Information Research, 29*(4), 479–494. https://doi.org/10.1007/s41324-020-00375-1

Pertierra, L. R., Hughes, K. A., Vega, G. C., & Olalla-Tárraga, M. Á. (2017). High resolution spatial mapping of human footprint across Antarctica and its implications for the strategic conservation of avifauna. *PLOS ONE, 12*(1), e0168280. https://doi.org/10.1371/journal.pone.0168280

Qi, P., Gimode, D., Saha, D., Schröder, S., Chakraborty, D., Wang, X., Dida, M. M., Malmberg, R. L., & Devos, K. M. (2018). UGbS-Flex, a novel bioinformatics pipeline for imputation-free SNP discovery in polyploids without a reference genome: Finger millet as a case study. *BMC Plant Biology, 18*(1), 117. https://doi.org/10.1186/s12870-018-1316-3

Syaharani, W., & Sriani, S. (2024). Application of k-means algorithm on clustering poor population data for extreme poverty elimination. *Sistemasi: Jurnal Sistem Informasi, 13*(4), 1732–1747. https://doi.org/10.32520/stmsi.v13i4.4384

Thomas, L., Williams, R., & Sandilands, D. (2007). Designing line transect surveys for complex survey regions. *Journal of Cetacean Research and Management, 9*(1), 1–13. https://doi.org/10.47536/jcrm.v9i1.687

Wang, X., Liao, Y., Ji, L., & Zhang, Z. (2023). A sensitive trajectory clustering method based on spatiotemporal density. *Journal of Intelligence and Knowledge Engineering, 1*(4), 101–106. https://doi.org/10.62517/jike.202304418

Wang, X., Zhang, Z., & Luo, Y. (2022). Clustering methods based on stay points and grid density for hotspot detection. *ISPRS International Journal of Geo-Information, 11*(3), 190. https://doi.org/10.3390/ijgi11030190

Wati, R. K., Pratiwi, H., & Sulandari, W. (2024). Perbandingan Algoritma Density-Based Spatial Clustering Algorithm with Noise (DBSCAN) dan Self-Organizing Map (SOM) untuk Clustering Data Gempa Bumi. *Statistika, 24*(2), 152–160. https://doi.org/10.29313/statistika.v24i2.3645

Wywiał, J. L., & Sitek, G. (2021). The influence of clustering population on estimation accuracy of population totals vector. *Journal of Statistical Theory and Practice, 15*(3), 61. https://doi.org/10.1007/s42519-021-00196-x

Xu, X., Liao, H., & Yang, X. (2023). An automatic density peaks clustering based on a density-distance clustering index. *AIMS Mathematics, 8*(12), 28926–28950. https://doi.org/10.3934/math.20231482

Zhang, S., & Li, K. (2023). A novel density peaks clustering algorithm with isolation kernel and K-induction. *Applied Sciences, 13*(1), 322. https://doi.org/10.3390/app13010322