Pemetaan Potensi Pajak Reklame Menggunakan K-Means WebGIS

Ema Deloris Silaban 1, Paniel Sam P. Sitorus 2, Dudes Manalu 3
1,2,3 Program Studi Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas HKBP Nommensen Pematangsiantar, Indonesia

ARTICLE INFORMATION
Received: June 00, 2026
Revised: July 00, 2026
Available online: August 00, 2026

KEYWORDS
K-Means, Web GIS, Clustering, Pajak Reklame, BPKPD

CORRESPONDENCE
Phone: +62 813-6001-7666
E-mail: manalu.dudes@gmail.com

A B S T R A C T

Purpose: This study develops a Web GIS-based advertisement tax potential mapping system using K-Means Clustering at the Regional Financial and Revenue Management Agency of Simalungun Regency, with Bandar District as the case study. Design/methods/approach: The study applies Research and Development with the Waterfall model through requirements analysis, system design, implementation, and testing. The system uses PHP Native, MySQL, Leaflet.js, latitude, longitude, and nominal tax value. K-Means with k=3 and Euclidean Distance groups advertisement objects into high, medium, and low potential categories. Findings/results: The system processed 150 advertisement tax records and produced three clusters: 3 high-potential objects, 18 medium-potential objects, and 129 low-potential objects. The process converged at the third iteration. Manual calculation and PHP output showed 100% conformity, and black-box testing confirmed that the main system features worked properly. Conclusions: The proposed system converts conventional tabular records into interactive geospatial information that helps officers prioritize advertisement tax supervision more objectively.

INTRODUCTION

Digitalisasi layanan publik menjadi bagian penting dalam penguatan tata kelola pemerintah daerah. Pada sektor pajak daerah, kualitas data menentukan ketepatan pengawasan, penagihan, dan pengambilan keputusan. Pajak reklame memiliki karakter spasial karena objeknya tersebar pada jalan utama, pusat perdagangan, dan area publik. Karena itu, pengelolaan data reklame tidak cukup hanya menggunakan daftar tabel, tetapi juga membutuhkan peta digital yang dapat menunjukkan lokasi dan sebaran objek pajak secara langsung.

Badan Pengelola Keuangan dan Pendapatan Daerah Kabupaten Simalungun membutuhkan sistem yang dapat membantu membaca potensi pajak reklame secara lebih cepat dan objektif. Kecamatan Bandar dipilih sebagai studi kasus karena memiliki aktivitas ekonomi dan perdagangan yang cukup padat. Data reklame pada wilayah tersebut memuat informasi koordinat dan nilai pajak yang dapat dianalisis untuk membedakan objek dengan potensi tinggi, sedang, dan rendah.

Penelitian sebelumnya telah membahas Sistem Informasi Geografis untuk pemetaan reklame, pemanfaatan Leaflet.js pada Web GIS, dan penerapan K-Means untuk clustering data pajak atau data geospasial. Namun, sebagian penelitian masih memisahkan fungsi peta dan fungsi analitik. Sistem pemetaan sering hanya menampilkan lokasi, sedangkan penelitian K-Means sering berhenti pada tabel hasil pengelompokan tanpa menyajikannya dalam Web GIS interaktif.

Berdasarkan kesenjangan tersebut, penelitian ini mengintegrasikan algoritma K-Means Clustering dengan Web GIS untuk memetakan potensi pajak reklame. Kontribusi utama penelitian adalah model aplikasi yang menggabungkan latitude, longitude, dan nilai pajak sebagai dasar klasterisasi, kemudian menampilkan hasilnya ke dalam peta tematik berbasis marker warna. Hasil sistem diharapkan membantu BPKPD menentukan prioritas pengawasan lapangan secara lebih berbasis data.

Table 1. State of the Art and Research Gap

Study	Focus	Method/Technology	Gap Addressed
Putri et al. [1]	Web-based billboard mapping	Web GIS	Focus on mapping without potential clustering
Setiawan [2]	Geographic information for billboard points	Leaflet.js	Presents location, not tax-potential grouping
Putri et al. [3]	Taxpayer compliance grouping	Data mining clustering	Not integrated with Web GIS visualization
Naim and Yahya [4]	WebGIS customer distribution	K-Means and WebGIS	Different domain and object type
This study	Advertisement tax potential mapping	K-Means + Web GIS	Combines spatial and economic attributes in one system

Source: Related works synthesis, 2026.

METHOD

Research Design and Dataset

This research applies Research and Development with the Waterfall model. The stages consist of requirements analysis, system design, implementation, testing, and evaluation. The study was conducted at the Regional Financial and Revenue Management Agency of Simalungun Regency using advertisement tax data from Bandar District. The dataset consists of 150 advertisement objects. The main variables are latitude, longitude, and nominal tax value. Latitude and longitude represent the spatial position of the object, while tax value represents its economic weight.

Table 2. Research and System Development Summary

Stage	Activity	Output	Technology/Method
Analysis	Observation, interview, and advertisement data review	System needs and data variables	R&D
Design	UML, database structure, and map interface design	System model and table structure	MySQL, UML
Implementation	Application coding and clustering algorithm	Web GIS and clustering process	PHP Native, Leaflet.js

Testing	Manual calculation vs PHP output and black-box test	Algorithm validation and functional status	Euclidean Distance, Black-box

Source: Research design, 2026.

K-Means Clustering Procedure

K-Means is used with k=3 to divide advertisement objects into high, medium, and low potential zones. The process starts by retrieving advertisement records from the database, selecting initial centroids, calculating the distance from every record to each centroid, assigning data to the nearest cluster, updating the centroid based on the average of cluster members, and repeating the process until the centroid no longer changes. The distance calculation uses Euclidean Distance with three attributes: latitude, longitude, and tax value.

$$d(i,c) = \sqrt{(lat_i - lat_c)^2 + (lon_i - lon_c)^2 + (tax_i - tax_c)^2} \quad (1)$$

Equation (1) calculates the distance between each advertisement object and every centroid. The smallest distance determines the cluster membership of each object.

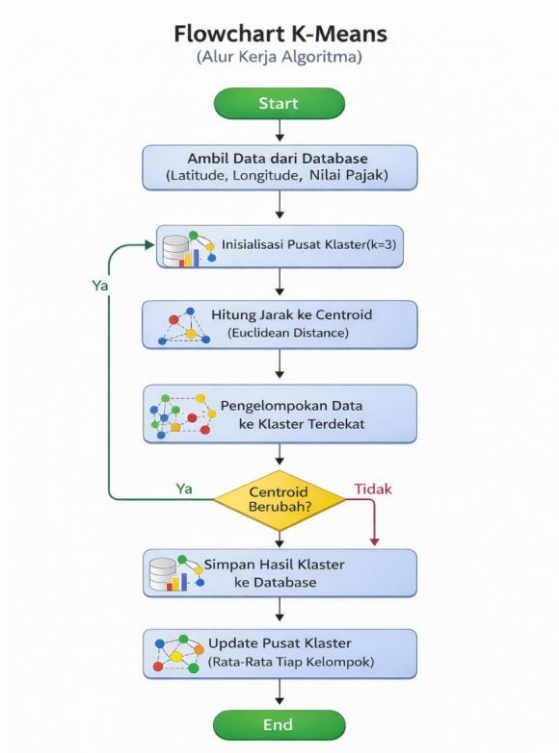

Figure 1. K-Means Workflow in the Proposed System

Web GIS Implementation

The Web GIS is developed using PHP Native as the backend, MySQL as the relational database, and Leaflet.js with OpenStreetMap as the map interface. The main tables include admin, advertisement object, and clustering result. After the clustering process is executed, every advertisement point is displayed as a colored marker. Red represents high potential, yellow represents medium potential, and green represents low potential. The system also provides login, data input, clustering calculation, popup information, region filter, Fly To navigation, and report output.

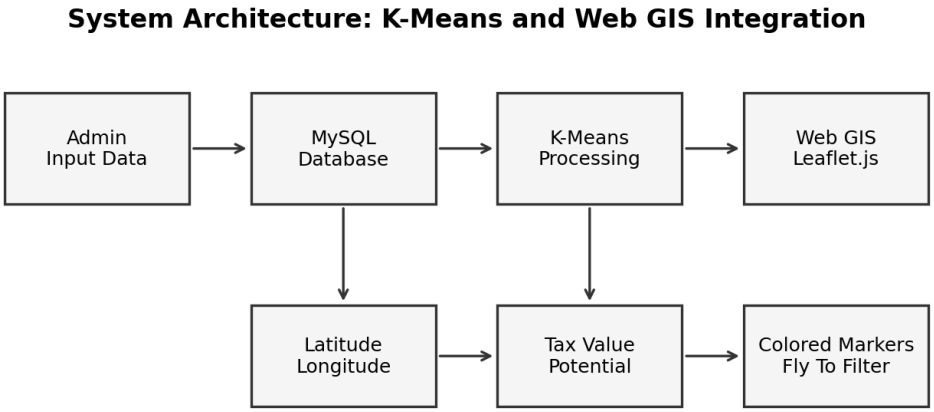

Figure 2. Architecture of K-Means and Web GIS Integration

RESULTS AND DISCUSSION

Clustering Result

The system processed 150 advertisement tax records and classified them into three potential categories. The clustering process reached convergence at the third iteration, meaning that the centroid and cluster membership became stable. The final result shows that 3 objects belong to the high-potential cluster, 18 objects to the medium-potential cluster, and 129 objects to the low-potential cluster. This distribution

indicates that most advertisement objects in the dataset have relatively low tax values, while only a small number of objects contribute high economic potential.

Table 3. Final Cluster Distribution and Centroid Summary

Zone	Number of Objects	Average Tax	Latitude Centroid	Longitude Centroid
High (C1)	3	Rp13,087,595.67	3.158451	99.328209
Medium (C2)	18	Rp4,244,302.50	3.159782	99.326866
Low (C3)	129	Rp1,187,606.59	3.157825	99.326364
Total	150	-	-	-

Source: Final K-Means processing result, 2026.

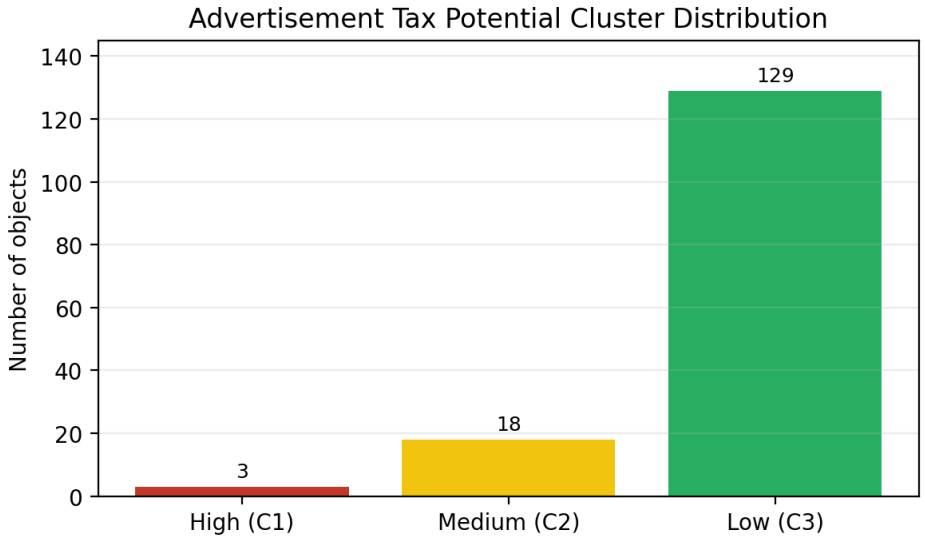

Figure 3. Distribution of Advertisement Objects by Cluster

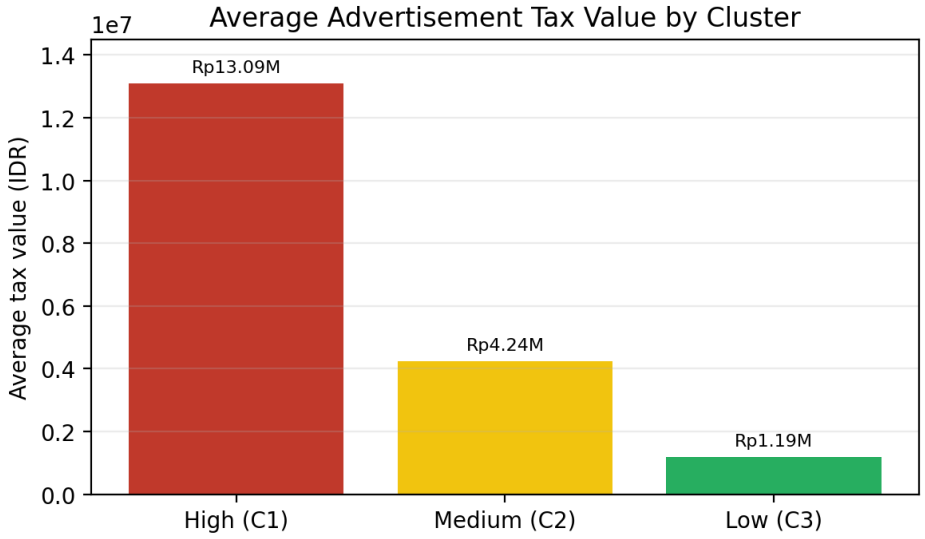

Figure 4. Average Tax Value by Cluster

The final centroid values show clear economic separation among clusters. The high cluster has an average tax value above Rp13 million, far higher than the medium and low clusters. This makes the high cluster a reasonable first priority for field supervision, data verification, and follow-up monitoring. The medium cluster can be monitored periodically because its economic contribution is moderate, while the low cluster can be managed through routine administrative supervision.

Validation and Functional Testing

Validation was carried out by comparing manual K-Means calculation and PHP system output. All validation samples produced the same cluster labels, so the conformity level reached 100%. This result shows that the implementation of Euclidean Distance, cluster allocation, and centroid updating in the PHP system follows the designed K-Means process. Functional testing was conducted using black-box testing to verify whether each feature produced the expected output.

Table 4. System Testing Summary

Test Aspect	Scenario	Result	Status
Algorithm comparison	Manual calculation vs PHP system	Cluster output is the same	100% valid
Login admin	Wrong password	Access is rejected and error message appears	Success
Data input	Add advertisement data	Data stored in MySQL	Success
K-Means calculation	Click Hitung Cluster button	Labels and marker colors updated	Success
Region filter	Select Bandar District	Map performs Fly To navigation	Success
Final report	Open clustering result	Cluster table and map information displayed	Success

Source: Manual-system validation and black-box testing, 2026.

Cluster-Based Supervision Strategy

The cluster output can be translated into operational priorities for advertisement tax supervision. High-potential objects need the fastest verification because their average tax value is significantly higher than other groups. Medium-potential objects require scheduled monitoring and data completeness checking. Low-potential objects can be managed through routine supervision, database maintenance, and periodic updates. This interpretation helps the agency move from equal treatment of all objects to risk-based and value-based supervision.

Table 5. Cluster-Based Supervision Strategy

Cluster	Main Characteristic	Recommended Action	Priority
High (C1)	Few objects with the largest average tax value	Immediate verification, intensive supervision, and payment status review	Very high
Medium (C2)	Moderate number of objects with medium tax value	Scheduled monitoring, data update, and reminder for renewal	Medium
Low (C3)	Largest number of objects with lower average tax value	Routine database maintenance and periodic field checking	Standard

Source: Interpretation of clustering result, 2026.

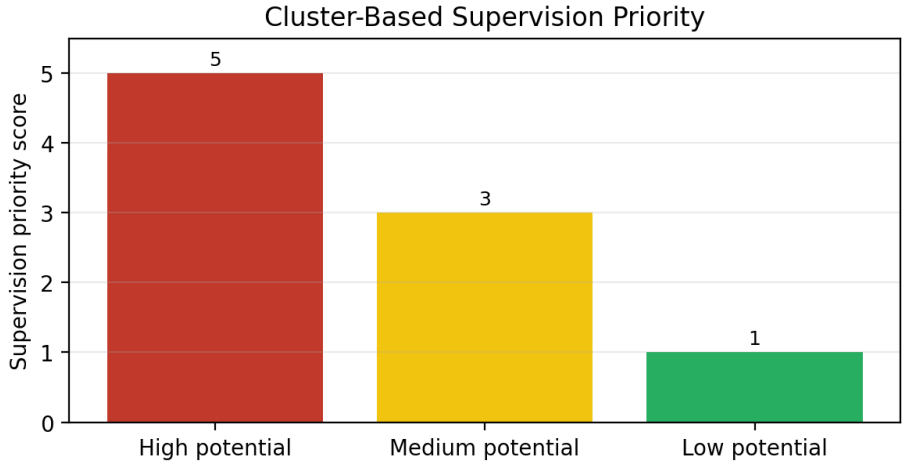

Figure 5. Cluster-Based Supervision Priority

Operational Scenario

In practical use, the administrator can input advertisement object data, run the K-Means calculation, and view the result directly on the Web GIS interface. The map marker color gives a quick visual signal for decision making. A red marker indicates a high-potential object that should be checked first, a yellow marker indicates a medium-potential object that requires scheduled monitoring, and a green marker indicates a low-potential object for routine control. The Fly To feature makes it easier to focus on Bandar District, while the report feature supports documentation and coordination with field officers.

This scenario shows that the system functions not only as a data storage tool but also as an analytical support tool. The combination of clustering and Web GIS allows BPKPD to read the relationship between location and tax potential in a single interface. This matters because supervision resources are usually limited. A priority map helps officers choose which points should be visited first, which points need data verification, and which points can be monitored through routine administration.

Discussion

The integration of K-Means and Web GIS adds value to conventional advertisement tax management. Instead of reading objects only as tabular records, officers can view high, medium, and low potential objects through thematic markers. This improves the speed of identifying priority points, especially for field verification and supervision planning.

The findings also confirm the relevance of combining spatial and economic attributes. Latitude and longitude preserve the geographic context, while tax value separates the objects according to economic potential. The cluster output therefore supports not only technical visualization but also managerial decision-making. High-potential objects can receive closer supervision because small errors in this group may create larger revenue risk.

Compared with previous Web GIS studies, this system does not stop at showing map points. It processes the data automatically through K-Means and updates marker categories based on clustering results. Compared with tax clustering studies that only produce tables, this research presents the output through an interactive map. That combination makes the result easier to interpret for non-technical users.

The limitation of the study is its scope. The dataset covers only Bandar District and has not yet integrated official real-time payment records through an external API. The cluster output also represents point-based potential mapping, not a formal administrative zoning boundary. Future development can add GeoJSON boundary layers, payment API integration, notification for expiring advertisement permits, and comparison with other clustering methods such as K-Medoids, DBSCAN, or Hierarchical Clustering.

CONCLUSIONS

This study successfully developed a Web GIS-based K-Means system to map advertisement tax potential in Bandar District, Simalungun Regency. The system processed 150 advertisement objects using latitude, longitude, and tax value, and grouped them into three categories: 3 high-potential objects, 18 medium-potential objects, and 129 low-potential objects. The clustering process converged at the third iteration, and comparison between manual calculation and PHP system output reached 100% conformity. The system features, including admin login, data input, K-Means calculation, colored markers, popup information, Fly To navigation, and result reporting, worked according to black-box testing. The system can help BPKPD read advertisement tax distribution more objectively, interactively, and based on geospatial data. Future work should expand the dataset to all districts, integrate official payment data, add GeoJSON administrative layers, and apply additional clustering evaluation metrics.

ACKNOWLEDGMENT

The authors express gratitude to Universitas HKBP Nommensen Pematangsiantar and the Regional Financial and Revenue Management Agency of Simalungun Regency for supporting the data collection, observation, and research process.

REFERENCES

[1] M. A. Putri, N. Rahaningsih, F. M. Basysyar, and O. Nurdiawan, "Penerapan Data Mining Menggunakan Metode Clustering Untuk Mengetahui Kelompok Kepatuhan Wajib Pajak Bumi dan Bangunan," Jurnal Accounting Information System (AIMS), vol. 5, no. 2, pp. 145-156, 2022.

[2] A. Setiawan, "Implementasi Leaflet.js Dalam Sistem Informasi Geografis Berbasis Web," Jurnal Teknik Informatika, 2021.

[3] Y. Naim and A. Yahya, "Implementasi Metode K-Means Dalam Penyebaran Pelanggan Koran Fajar Berbasis WebGIS," Jurnal Ilmiah Sistem Informasi dan Teknik Informatika (JISTI), vol. 5, no. 2, pp. 25-32, 2022, doi: 10.57093/jisti.v5i2.124.

[4] A. S. H. P. E. Muzaki, "Implementasi algoritma K-Means untuk clustering data potensi pajak daerah," Jurnal Sistem dan Teknologi Informasi (JUSTIN), vol. 12, no. 1, pp. 45-52, 2024.

[5] D. S. Nurfadhilah, A. Setiawan, and R. Zulkifli, "Implementasi Geospatial Clustering Menggunakan Algoritma K-Means Untuk Penentuan Lokasi Strategis Promosi Kampus," Jurnal Janitra Informatika dan Sistem Informasi, vol. 5, no. 2, pp. 108-116, 2025, doi: 10.59395/xfctnc12.

[6] S. Wahyuni, "Implementation of the K-Means Algorithm for Clustering Tax Compliance of Land and Building Taxpayers in Medan City," Journal of Computer Engineering, System and Science, vol. 10, no. 1, 2025.

[7] Wahyudi, "Pengelompokan Kabupaten di Indonesia untuk Pemetaan Pendapatan Daerah Menggunakan Algoritma K-Means," MALCOM: Indonesian Journal of Machine Learning and Computer Science, vol. 5, no. 3, pp. 1143-1151, 2025, doi: 10.57152/malcom.v5i3.2206.

[8] F. A. I. S. Aji, "Penerapan metode clustering pada analisis realisasi PAD dengan algoritma K-Means," 2021.

[9] Direktorat Jenderal Perimbangan Keuangan, Laporan Kerja Direktorat Jenderal Perimbangan Keuangan 2022. Jakarta: Kementerian Keuangan Republik Indonesia, 2022.

[10] Bupati Simalungun, Peraturan Bupati Nomor 30 Tahun 2024 tentang Petunjuk Teknis Pelaksanaan Pengelolaan Pajak Reklame dan Pajak Air Tanah. Simalungun, 2024.