

## OPTIMASI KLASTERISASI PENERIMAAN PAJAK BUMI DAN BANGUNAN MENGUNAKAN ALGORITMA K-MEDOIDS

Febri Abdi Annur Dhuha<sup>1\*</sup>, Ade Irma Purnamasari<sup>2</sup>, Denni Pratama<sup>3</sup>, Edi Tohidi<sup>4</sup>,  
Edi Wahyudin<sup>5</sup>

<sup>1,2</sup>Teknik Informatika, STMIK IKMI Cirebon, Indonesia

<sup>3</sup>Manajemen Informatika, STMIK IKMI Cirebon, Indonesia

<sup>4,5</sup>Komputerisasi Akuntansi, STMIK IKMI Cirebon, Indonesia

\*E-mail: [febriabdi04@gmail.com](mailto:febriabdi04@gmail.com)

Submitted: 26 November 2025	Accepted : 14 January 2026	Published: 02 Februari 2026
-----------------------------	----------------------------	-----------------------------

### ABSTRAK

Pajak Bumi dan Bangunan (PBB) merupakan komponen strategis dalam Pendapatan Asli Daerah (PAD) yang berperan penting dalam mendukung penyelenggaraan pembangunan dan pelayanan publik. Namun, heterogenitas data dan variasi karakteristik objek pajak menyebabkan pemerintah daerah mengalami kesulitan dalam memetakan potensi fiskal dan tingkat kepatuhan wajib pajak secara akurat. Penelitian ini bertujuan menganalisis efektivitas algoritma K-Medoids dalam mengelompokkan wajib pajak di Kecamatan Tanjung berdasarkan atribut numerik, yaitu luas tanah, luas bangunan, *NJOP* tanah, *NJOP* bangunan, dan nilai PBB tahun berjalan. Metode penelitian meliputi tahapan pengumpulan data, pra-pemrosesan, transformasi logaritmik, normalisasi, implementasi algoritma K-Medoids, serta evaluasi hasil kluster menggunakan metrik *Silhouette Coefficient* dan *Davies–Bouldin Index*. Proses komputasi dilakukan menggunakan *Python* dengan pustaka *pyClustering* dan *scikit-learn*. Hasil penelitian menunjukkan terbentuknya empat kluster wajib pajak dengan karakteristik berbeda: kluster aset besar berkontribusi rendah, kluster premium berkontribusi tinggi, kluster ekonomi rendah dengan pola pembayaran tidak stabil, dan kluster ekonomi menengah dengan kepatuhan cukup baik. Evaluasi kualitas model menghasilkan *Silhouette Coefficient* sebesar 0,4204 dan *Davies–Bouldin Index* sebesar 0,7893, yang menunjukkan struktur kluster cukup baik dan stabil. Temuan ini memberikan kontribusi empiris dalam mendukung optimalisasi pengelolaan PBB berbasis analitik, serta dapat digunakan sebagai dasar penyusunan strategi penagihan berbasis prioritas dan formulasi kebijakan fiskal yang lebih tepat sasaran.

*Kata Kunci:* Data Mining, K-Medoids, Clustering, Pajak Bumi dan Bangunan, Analisis Fiskal.

### ABSTRACT

*Land and Building Tax (PBB) is a strategic component of Local Own-Source Revenue (PAD) that plays an important role in supporting development and public services. However, data heterogeneity and variations in tax object characteristics make it difficult for local governments to accurately map fiscal potential and taxpayer compliance levels. This study aims to analyze the effectiveness of the K-Medoids algorithm in grouping taxpayers in Tanjung Regency based on numerical attributes, namely land area, building area, land NJOP, building NJOP, and current year PBB value. The research methods included data collection, pre-processing, logarithmic transformation, normalization, implementation of the K-Medoids algorithm, and evaluation of cluster results using the Silhouette Coefficient and Davies–Bouldin Index metrics. The calculation process was performed using Python with the pyClustering and scikit-learn libraries. The results show the formation of four taxpayer clusters with different characteristics: a large asset cluster with low contribution, a premium cluster with high contribution, a low-income cluster with unstable payment patterns, and a middle-income cluster with fairly good*

---

*compliance rates. The model quality evaluation produced a Silhouette Coefficient of 0.4204 and a Davies–Bouldin Index of 0.7893, indicating a fairly good and stable cluster structure. These findings provide empirical support for the optimization of analytics-based property tax management and can be used as a basis for developing priority-based collection strategies and more targeted fiscal policy formulation.*

---

*Keywords: Data Mining, K-Medoids, Clustering, Land and Building Tax, Fiscal Analysis.*

---

## **PENDAHULUAN**

Pajak Bumi dan Bangunan (PBB) merupakan instrumen fiskal yang memiliki kontribusi signifikan terhadap Pendapatan Asli Daerah (PAD). Sebagai pajak yang dikenakan atas kepemilikan dan pemanfaatan tanah serta bangunan, PBB berperan penting dalam menopang pembiayaan penyelenggaraan pemerintahan daerah dan pembangunan infrastruktur publik (Ayu et al., 2022). Sejalan dengan kebijakan desentralisasi fiskal, kewenangan pengelolaan PBB-P2 dialihkan kepada pemerintah daerah yang secara empiris terbukti meningkatkan kemandirian fiskal, meskipun implementasinya masih menghadapi tantangan berupa rendahnya kesadaran masyarakat dan efektivitas administrasi perpajakan yang belum optimal (Hasan et al., 2025; Selni et al., 2024). Penelitian Adhim (2023) juga menegaskan bahwa efektivitas dan efisiensi dalam proses pemungutan PBB-P2 memiliki hubungan langsung dengan peningkatan PAD, sementara Psdku (2023) menunjukkan bahwa fluktuasi penerimaan PBB tetap berperan signifikan dalam menjaga stabilitas fiskal daerah.

Meskipun memiliki peran strategis, pengelolaan PBB di banyak daerah masih menghadapi kendala substantif. Ketidaktepatan pendataan, ketidaksesuaian data lapangan dengan basis data pajak, serta rendahnya pemahaman masyarakat mengenai kewajiban perpajakan menjadi penyebab kurang optimalnya penerimaan pajak (Ayu et al., 2022; Nurhani et al., 2024). Aspek operasional seperti keterbatasan sumber daya manusia dalam pelayanan pajak turut memperburuk efektivitas administrasi perpajakan (Akbar, 2021). Di sisi lain, integrasi teknologi informasi melalui sistem informasi perpajakan terkomputerisasi terbukti mampu meningkatkan efisiensi pendataan, pemrosesan, dan pengelolaan dokumen pajak (Meilinda & Jayanti, 2022; Rizal et al., 2022). Karena itu, penerapan pendekatan analitik yang memanfaatkan data besar menjadi kebutuhan penting dalam penguatan kebijakan perpajakan daerah.

Pemanfaatan teknik Data Mining memungkinkan identifikasi pola kepatuhan wajib pajak dan pengelompokan kelompok prioritas pembayaran secara sistematis (Widiya et al., 2023). Pendekatan ini memberikan basis analitis bagi kebijakan berbasis bukti, terutama dalam memahami faktor yang memengaruhi kepatuhan, seperti pengetahuan, kesadaran, pendapatan, dan faktor psikologis (Ramadhani & Pusposari, 2024; Sari et al., 2025; Zheng et al., 2024). Digitalisasi sistem perpajakan melalui e-filing dan layanan daring juga terbukti meningkatkan efisiensi pemungutan dan memperluas partisipasi Masyarakat (Darajat & Sofianty, 2023; Mulyani et al., 2023). Namun, analisis data PBB tidak terlepas dari tantangan keberadaan outlier yang dapat menurunkan akurasi model clustering, terutama pada algoritma berbasis centroid seperti K-Means (Gantara & Ali, 2023). Untuk itu, metode berbasis kepadatan seperti HDBSCAN dan DBSCAN telah digunakan untuk menangani data heterogen dan pola kepadatan kompleks (Fatemi et al., 2022, 2023; Isaenko et al., 2024).

Dalam konteks pengelompokan data fiskal, algoritma K-Medoids dipandang lebih stabil karena menggunakan medoid sebagai pusat kluster sehingga lebih tahan terhadap pengaruh data ekstrem dan nilai pencilan (Farissa et al., 2021; Orisa & Faisol, 2021). Keunggulan K-Medoids juga telah divalidasi pada berbagai domain seperti energi, kesehatan, pelanggan, dan kecerdasan buatan (Herviany et al., 2021; Martins et al., 2024; Ramadevi & Das, 2024;

Sulistiyawati & Sadikin, 2021). Dalam konteks pajak daerah, metode ini terbukti efektif untuk mengelompokkan wilayah atau wajib pajak berdasarkan indikator sosial-ekonomi dan perilaku pembayaran pajak (Azhar & Zakiah, 2025; Rahmawati & Fauzan, 2024; Sartono & Kudus, 2025; Wibowo et al., 2022). Validitas kelompok hasil klusterisasi kemudian dapat diverifikasi menggunakan *Silhouette Coefficient* dan *Davies–Bouldin Index* (Septian & Nohe, 2023).

Berdasarkan kondisi tersebut, penelitian ini memiliki urgensi dalam menyediakan model klusterisasi wajib pajak PBB yang akurat dan stabil sebagai dasar penguatan strategi pengelolaan pajak daerah di Kecamatan Tanjung. Tujuan penelitian ini adalah menerapkan algoritma K-Medoids dalam mengelompokkan wajib pajak berdasarkan karakteristik pembayaran dan nilai objek pajak, serta mengevaluasi kualitas kluster yang terbentuk menggunakan indeks validasi. Rencana pemecahan masalah dilakukan melalui serangkaian proses pengolahan data, penerapan metode klusterisasi, serta analisis hasil kluster sebagai landasan perumusan kebijakan intensifikasi dan ekstensifikasi pajak yang lebih efektif, adaptif, dan berkelanjutan.

### METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan rancangan eksploratif untuk menganalisis pola wajib pajak PBB-P2 di Kecamatan Tanjung pada tahun 2025. Objek penelitian mencakup seluruh data wajib pajak yang tercatat di Kantor Kecamatan Tanjung, sedangkan data yang digunakan berasal dari dokumentasi sekunder yang diperoleh melalui ekstraksi basis data administrasi pajak. Proses pengumpulan data dilakukan melalui *data retrieval* dan *preprocessing* yang mencakup pemeriksaan konsistensi nilai, pembersihan duplikasi, imputasi nilai hilang, serta deteksi outlier untuk memastikan integritas dataset. Variabel penelitian didefinisikan secara operasional meliputi NJOP, total pembayaran, tunggakan aktif, tingkat kepatuhan wajib pajak, dan kategori objek pajak, yang seluruhnya dinormalisasi menggunakan *Min-Max Normalization*. Analisis dilakukan menggunakan algoritma K-Medoids yang dipilih karena ketahanannya terhadap data ekstrem; proses analisis meliputi perhitungan matriks jarak, inialisasi medoid, optimasi dissimilarity, dan pembentukan kluster, sedangkan validasi kualitas kluster dilakukan menggunakan *Silhouette Coefficient* dan *Davies–Bouldin Index*. Seluruh analisis dikerjakan menggunakan perangkat lunak Python dengan pustaka *NumPy*, *Pandas*, dan *scikit-learn* guna menghasilkan pemetaan kluster wajib pajak yang informatif sebagai dasar rekomendasi kebijakan pengelolaan PBB di Kecamatan Tanjung.

### HASIL DAN PEMBAHASAN

#### Pengumpulan dan Pemuatan Data

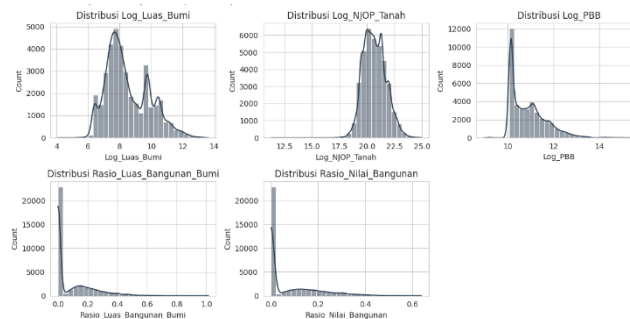
Tabel 1 Deskripsi Data Awal

No	NOP	Nama Wajib Pajak	Desa	Luas Bumi (m <sup>2</sup> )	Luas Bangunan (m <sup>2</sup> )	NJOP Tanah (Rp)	NJOP Bangunan (Rp)	PBB 2025 (Rp)	Colom 1
1	33.29.110.001.00 1-0001.0	KASMINAH ELI NURWATI	SARIREJA	3.471.0	0.0	35.751.300	0.0	134.14 8	NaN
2	33.29.110.001.00 1-0002.0	SUYUD KARNETI	SARIREJA	6.942.0	0.0	71.502.600	0.0	268.29 4	NaN
3	33.29.110.001.00 1-0003.0	NARSIAH BT MICHRUD	SARIREJA	370.0	0.0	38.110.000	0.0	25.000	NaN
4	33.29.110.001.00 1-0005.0	DAREM	SARIREJA	252.0	0.0	25.956.000	0.0	25.000	NaN

No	NOP	Nama Wajib Pajak	Desa	Luas Bumi (m <sup>2</sup> )	Luas Bangunan (m <sup>2</sup> )	NJOP Tanah (Rp)	NJOP Bangunan (Rp)	PBB 2025 (Rp)	Colom 1
5	33.29.110.001.00 1-0006.0	WARNO	SARIREJA	1.260.0	0.0	129.780.000	0.0	48.697	NaN
...	....	....	...	...	...	...	...	...	...
37.5 97									

Pada Tabel 1, dataset awal Pajak Bumi dan Bangunan (PBB) Kecamatan Tanjung yang digunakan dalam penelitian ini terdiri dari 37.597 baris dengan 252 atribut yang menggambarkan detail, setiap entri mencakup informasi identitas wajib pajak (NOP, nama, desa), serta variabel ekonomi seperti luas bumi, luas bangunan, nilai jual objek pajak (NJOP), dan besaran PBB tahun 2025. Secara umum, sebagian besar data memiliki nilai NJOP Bangunan = 0, yang mengindikasikan bahwa mayoritas objek pajak merupakan tanah kosong atau belum memiliki bangunan. Hal ini sesuai dengan karakteristik daerah suburban yang sedang berkembang. Ukuran file sebesar 72,3 MB dan jumlah atribut yang sangat besar menunjukkan bahwa dataset ini berskala besar dan kompleks, sehingga diperlukan proses *data preprocessing* untuk membersihkan nilai hilang (*NaN*) dan atribut yang tidak relevan sebelum dilakukan analisis *clustering*. Adapun struktur data yang dominan numerik memungkinkan penerapan metode *unsupervised learning* berbasis jarak seperti K-Medoids, karena algoritma ini bekerja efektif pada data kuantitatif multidimensional. Tahapan selanjutnya akan berfokus pada proses pra-pemrosesan data untuk memastikan kualitas dan konsistensi dataset yang digunakan.

### Pra-Pemrosesan Data (*Data Preprocessing*)



Gambar 1 Transformasi Logaritmik

Hasil grafik proses pra-pemrosesan data pada Gambar 1 di bawah yang dilakukan untuk menyiapkan dataset Pajak Bumi dan Bangunan (PBB) sebelum proses *clustering* menggunakan algoritma K-Medoids. Tahap ini merupakan salah satu bagian paling krusial dalam penelitian berbasis *Data Mining* karena kualitas hasil analisis sangat bergantung pada kebersihan dan keseragaman data yang digunakan. Setelah proses pembersihan data (*cleaning*) dilakukan, jumlah data yang semula sebanyak 37.597 baris berkurang menjadi 37.545 baris akibat penghapusan nilai kosong (*missing values*) dan duplikasi pada atribut Nomor Objek Pajak (NOP). Langkah ini memastikan bahwa setiap entri data bersifat unik dan representatif terhadap kondisi sebenarnya di Kecamatan Tanjung.

Menunjukkan hasil transformasi logaritmik terhadap variabel numerik utama, yaitu *Luas Bumi (LUAS BUMI SPPT)*, *Nilai Jual Objek Pajak Tanah (NJOP TANAH)*, dan *PBB Tahun 2025*. Sebelum transformasi, data memiliki rentang nilai yang sangat lebar (*skewed*), yang dapat menyebabkan pada proses *clustering*. Transformasi logaritmik berhasil menstabilkan variasi data sehingga distribusi menjadi

lebih simetris dan mendekati normal. Hal ini terlihat dari bentuk histogram yang lebih seimbang, di mana puncak distribusi berada di sekitar nilai tengah. Kondisi ini sangat penting agar jarak antar data yang dihitung oleh algoritma K-Medoids tidak terlalu dipengaruhi oleh perbedaan skala yang ekstrem antaratribut.

Selain transformasi logaritmik, dilakukan pula feature engineering dengan membuat dua fitur turunan: *Rasio Luas Bangunan terhadap Bumi* dan *Rasio Nilai Bangunan*. Kedua rasio ini berfungsi untuk menggambarkan proporsi fisik dan ekonomi antara tanah dan bangunan, sehingga dapat memperkaya informasi dalam proses pengelompokan wajib pajak. Berdasarkan hasil distribusi yang ditampilkan pada grafik bagian bawah, sebagian besar nilai rasio berkisar antara 0 hingga 0,2, menandakan bahwa sebagian besar objek pajak memiliki lahan luas dengan bangunan kecil atau bahkan tanpa bangunan, yang tipikal untuk area perdesaan seperti Kecamatan Tanjung.

Tabel 2 Data Setelah Standard Scaling

No	Log_Luas_Bumi	Log_NJOP_Tanah	Log_PBB	Rasio_Luas_Bangunan_Bumi	Rasio_Nilai_Bangunan
1	1.43	1.29	1.23	-0.65	-0.66
2	1.94	1.98	2.13	-0.65	-0.66
3	-0.21	-0.91	-0.98	-0.65	-0.66
4	-0.49	-1.29	-0.98	-0.65	-0.66
5	0.69	0.29	-0.10	-0.65	-0.66

Tabel 2 memperlihatkan hasil normalisasi data menggunakan *Standard Scaler*. Setiap fitur dinormalisasi menjadi skala standar ( $mean = 0$ ,  $standard deviation = 1$ ), memastikan bahwa semua variabel memiliki bobot yang seimbang dalam perhitungan jarak *Euclidean*. Normalisasi ini mencegah dominasi satu variabel terhadap variabel lain yang mungkin memiliki rentang nilai lebih besar. Hasilnya menunjukkan bahwa setiap fitur memiliki nilai rata-rata mendekati nol dan sebaran relatif kecil, yang menandakan proses standarisasi berhasil dilakukan.

### Implementasi Algoritma K-Medoids

Hasil implementasi algoritma K-Medoids pada penelitian ini menunjukkan efektivitas metode tersebut dalam mengelompokkan data wajib pajak berdasarkan kemiripan karakteristik ekonomi dan nilai objek pajak di Kecamatan Tanjung. Proses analisis dilakukan menggunakan pendekatan *unsupervised learning* dengan empat kluster yang terbentuk secara optimal melalui perhitungan jarak *Euclidean* dan pembaruan medoid secara iteratif hingga mencapai kondisi konvergen. Setiap kluster merepresentasikan kategori wajib pajak yang berbeda, mulai dari kelompok dengan nilai lahan tinggi namun tanpa bangunan, hingga kelompok wajib pajak dengan kontribusi besar terhadap PAD. Keberhasilan K-Medoids dalam menghasilkan kluster yang stabil membuktikan keunggulannya dalam menangani dataset besar dan heterogen tanpa temuan ini memperkuat posisi K-Medoids sebagai metode *clustering* yang andal dalam mendukung kebijakan fiskal berbasis data, khususnya dalam optimalisasi penerimaan Pajak Bumi dan Bangunan di tingkat daerah.

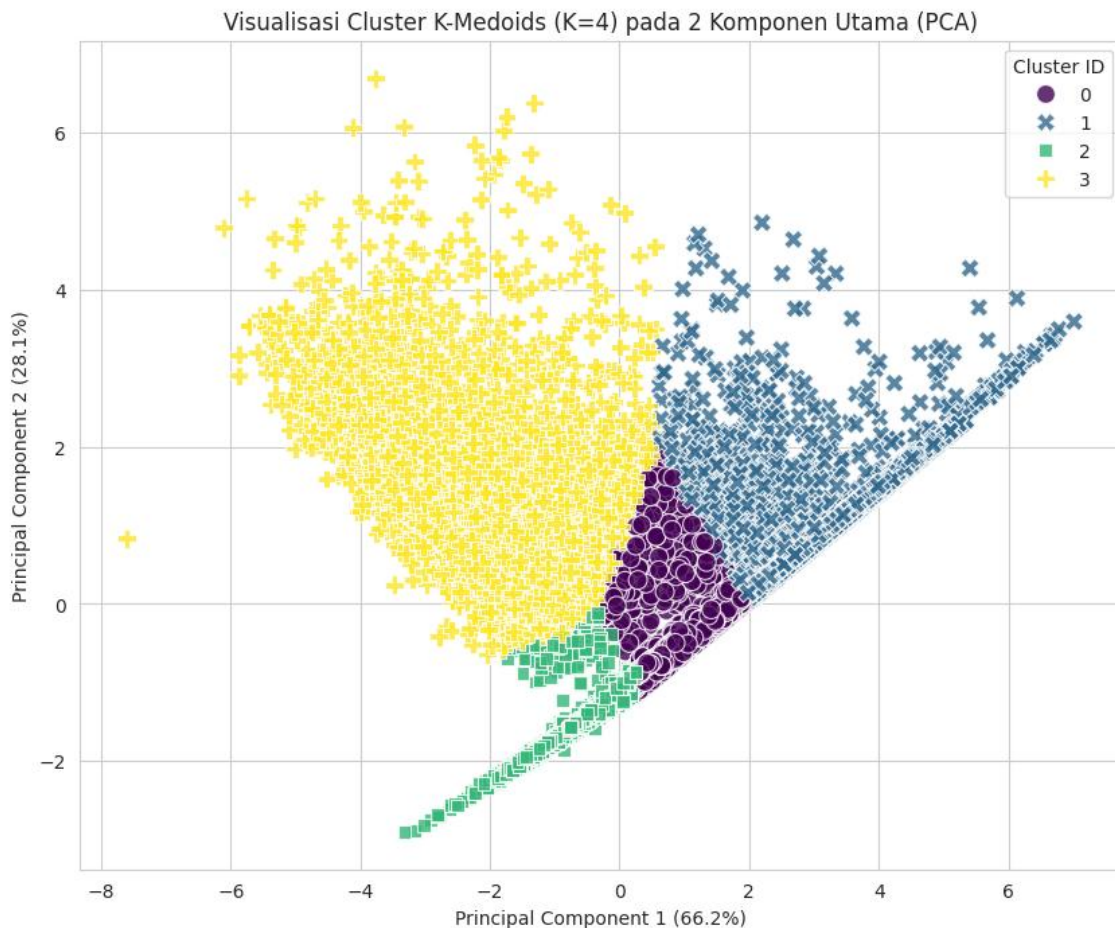
### Evaluasi Dan Visualisasi

1. Silhouette Coefficient Rata-rata: 0.4204 (Ideal: Mendekati 1)
2. Davies-Bouldin Index: 0.7893 (Ideal: Mendekati 0)

Gambar 2 Evaluasi

Hasil penelitian ini merupakan proses evaluasi pada Gambar 2 kinerja algoritma K-Medoids setelah proses klusterisasi terhadap data Pajak Bumi dan Bangunan (PBB) Kecamatan Tanjung berhasil dilakukan. Evaluasi ini bertujuan untuk menilai sejauh mana kualitas hasil pengelompokan dapat merepresentasikan struktur alami dari data menggunakan dua metrik utama, yaitu *Silhouette Coefficient* dan *Davies–Bouldin Index*. Berdasarkan hasil eksperimen, nilai *Silhouette Coefficient* rata-rata sebesar 0,4204, yang mengindikasikan bahwa hasil klusterisasi tergolong cukup baik dan setiap data sudah memiliki kedekatan yang cukup jelas dengan klusternya masing-masing. Nilai ini menunjukkan bahwa sebagian besar data berada pada kluster yang sesuai dan hanya sedikit yang tumpang tindih antar-kluster.

Sementara itu, nilai *Davies–Bouldin Index* (DBI) sebesar 0,7893, yang relatif mendekati nol, mengindikasikan bahwa kluster yang terbentuk memiliki tingkat pemisahan yang baik antar satu sama lain. Nilai DBI yang lebih kecil menunjukkan bahwa jarak antar-kluster cukup jauh dan variasi dalam kluster relatif rendah, menandakan bahwa model berhasil menemukan struktur kelompok yang stabil dan tidak terlalu overlap. Dalam konteks akademik.



Gambar 3 Visualisasi PCA

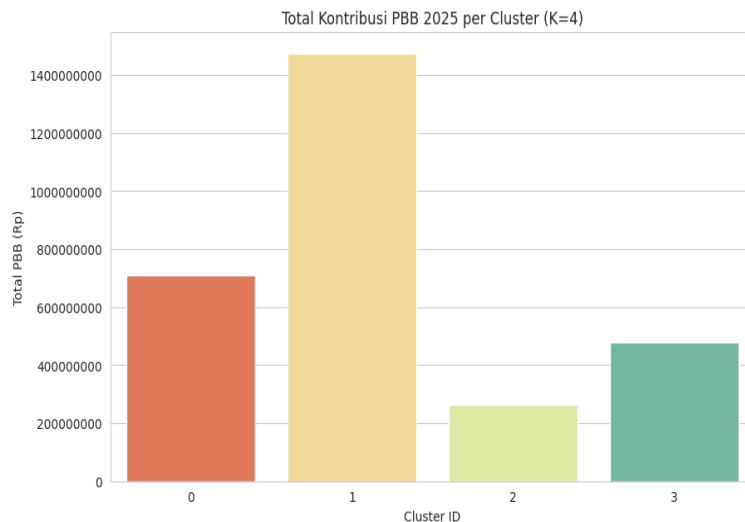
Gambar 3 Visualisasi *Principal Component Analysis* (PCA) di bawah tersebut menampilkan hasil visualisasi pengelompokan data Pajak Bumi dan Bangunan (PBB) menggunakan algoritma

K-Medoids dengan jumlah kluster ( $K$ ) = 4, yang diproyeksikan ke dalam dua komponen utama hasil *Principal Component Analysis* (PCA). Sumbu horizontal (*Principal Component 1*) menjelaskan 66,2% variasi data, sedangkan sumbu vertikal (*Principal Component 2*) menjelaskan 28,1%, sehingga total varian yang terwakili mencapai 94,3% dari keseluruhan informasi dataset. Setiap warna merepresentasikan kelompok wajib pajak dengan karakteristik ekonomi dan nilai objek pajak yang berbeda. Terlihat bahwa keempat kluster terpisah secara cukup jelas, menunjukkan bahwa model K-Medoids berhasil mengidentifikasi segmentasi wajib pajak berdasarkan kemiripan atribut seperti luas tanah, luas bangunan, dan nilai jual objek pajak. Pola sebaran yang relatif kompak dan tidak tumpang tindih secara signifikan menunjukkan validitas struktural hasil pengelompokan, yang juga didukung oleh nilai *Silhouette Coefficient* sebesar 0,4204, dan *Davies–Bouldin Index* sebesar 0,7893, menandakan kualitas kluster yang baik dan stabil secara matematis.

### Analisis Profil *Cluster* dan Rekomendasi

Cluster	Median Luas Bumi (m <sup>2</sup> )	Median Luas Bangunan (m <sup>2</sup> )	Median NJOP Tanah (Rp)	Median NJOP Bangunan (Rp)	Median PBB 2025 (Rp)	Jumlah WP	Persentase WP (%)
0	11.400,00	0,00	162.740.000,00	0,00	67.573,00	10.177	27,11
1	37.200,00	0,00	410.880.000,00	0,00	164.641,00	6.325	16,85
2	1.840,00	0,00	42.120.000,00	0,00	25.000,00	9.551	25,44
3	2.000,00	490,00	563.760.000,00	153.300.000,00	32.914,00	11.492	30,61

Tabel ini menyajikan profil statistik median untuk setiap kluster yang terbentuk. Kluster 1 memiliki nilai median tertinggi dalam hampir semua variabel ekonomi, khususnya NJOP Tanah Rp 410.880.000 dan median PBB Rp 164.641, yang menunjukkan bahwa kelompok ini terdiri dari wajib pajak dengan aset besar namun cenderung tidak memiliki bangunan, sehingga dikategorikan sebagai lahan kosong berpotensi tinggi. Sebaliknya, Kluster 3 menampilkan kombinasi luas tanah dan bangunan yang relatif seimbang serta median PBB Rp 32.914, mengindikasikan kelompok rumah tangga menengah yang stabil dalam kepatuhan pajak. Kluster 0 dan 2 memiliki karakteristik serupa sebagai lahan kosong kecil hingga sedang, dengan nilai kontribusi rendah namun populasi wajib pajak cukup besar. Distribusi ini mencerminkan adanya ketimpangan dalam struktur nilai aset antar wajib pajak dan menggarisbawahi pentingnya kebijakan diferensiasi tarif untuk menciptakan keadilan fiskal di tingkat daerah.



Gambar 4 Total Kontribusi PBB 2025

Gambar 4 kontribusi total PBB tahun 2025 yang dihimpun dari setiap kluster hasil pengelompokan K-Medoids. Kluster 1 memberikan kontribusi tertinggi, mencapai sekitar Rp145 juta, jauh melampaui kluster lain. Hal ini menegaskan bahwa meskipun jumlah wajib pajaknya relatif sedikit (16,85%), kelompok ini menyumbang porsi pendapatan terbesar karena nilai objek pajak yang tinggi. Kluster 0 dan 3 memberikan kontribusi menengah dengan total sekitar Rp70 juta dan Rp50 juta, menunjukkan potensi penerimaan yang stabil namun belum optimal. Sementara itu, kluster 2 memiliki kontribusi paling rendah sekitar Rp25 juta, yang mencerminkan kelompok wajib pajak kecil dengan kemampuan ekonomi rendah. Pola ini memperlihatkan bahwa distribusi beban pajak masih terpusat pada segmen ekonomi atas, sehingga perlu kebijakan fiskal yang lebih inklusif untuk memperluas basis penerimaan pajak dari kelompok kecil dan menengah.

Akhir analisis profil *Cluster* dan rekomendasi kebijakan fiskal menggunakan algoritma K-Medoids pada data Pajak Bumi dan Bangunan (PBB) Kecamatan Tanjung. Berdasarkan tabel median nilai asli dan hasil klasifikasi, diperoleh empat kluster utama yang merepresentasikan karakteristik wajib pajak (WP) dengan atribut seperti luas bumi, luas bangunan, nilai jual objek pajak (NJOP) tanah dan bangunan, serta besaran PBB tahun 2025.

1. *Cluster* 0 diidentifikasi sebagai kelompok *lahan kosong potensial (WP investasi lahan)* dengan median PBB sebesar Rp 67.573 dan total 10.177 wajib pajak atau 27,11% dari populasi. Ciri khas klaster ini adalah luas lahan besar (11.400 m<sup>2</sup>) tetapi tanpa bangunan, menandakan potensi aset yang belum produktif. Rekomendasi kebijakan bagi kelompok ini adalah melakukan verifikasi status lahan dan mempertimbangkan penerapan *PBB progresif untuk lahan tidur* agar dapat mendorong pemanfaatan lahan secara ekonomi dan meningkatkan kontribusi terhadap PAD (Pendapatan Asli Daerah).
2. *Cluster* 1 juga termasuk kategori *lahan kosong potensial* dengan median PBB yang jauh lebih tinggi yaitu Rp 164.641 dan jumlah wajib pajak sebanyak 6.325 (16,85%). Nilai NJOP tanah yang tinggi (Rp 410,88 miliar) menunjukkan bahwa kelompok ini terdiri atas pemilik lahan bernilai tinggi namun belum dimanfaatkan optimal. Pemerintah dapat memfokuskan kebijakan pengawasan terhadap klaster ini, misalnya dengan melakukan penilaian ulang NJOP dan mengenakan tarif pajak tambahan untuk lahan tidak produktif.
3. *Cluster* 2 mencerminkan kelompok *wajib pajak kecil* dengan median PBB terendah yakni Rp 25.000 dan mencakup 9.551 wajib pajak (25,44%). Karakteristiknya menunjukkan luas tanah

yang kecil (1.840 m<sup>2</sup>) dan tanpa bangunan, menggambarkan masyarakat dengan kapasitas ekonomi rendah. Rekomendasi yang relevan bagi klaster ini adalah peningkatan sosialisasi dan penyuluhan pajak, serta pemberian insentif atau keringanan tarif untuk mendukung kepatuhan pembayaran.

4. Sementara itu, *Cluster* 3 mewakili wajib pajak menengah (WP standar/rumah tinggal) dengan median PBB Rp 32.914 dan jumlah wajib pajak terbesar yaitu 11.492 orang (30,61%). Klaster ini memiliki karakteristik yang paling seimbang, dengan kombinasi lahan dan bangunan yang proporsional (luas tanah 2.000 m<sup>2</sup> dan luas bangunan 490 m<sup>2</sup>). Pemerintah daerah dapat memfokuskan pelayanan digital, kemudahan pembayaran, dan penguatan sistem monitoring untuk mempertahankan tingkat kepatuhan kelompok ini yang relatif stabil.

Hasil ini menunjukkan bahwa algoritma K-Medoids berhasil mengelompokkan data PBB secara efektif dan menghasilkan profil klaster yang jelas antara kelompok wajib pajak berpotensi tinggi, sedang, dan rendah. Rekomendasi kebijakan yang dihasilkan bersifat aplikatif — dari verifikasi aset dan penilaian NJOP ulang hingga pemberdayaan kelompok ekonomi kecil. Temuan ini menegaskan relevansi analisis *Data Mining* berbasis *clustering* sebagai dasar *evidence-based policy* dalam optimalisasi penerimaan pajak daerah, yang menekankan pentingnya pendekatan analitik dalam perencanaan fiskal daerah.

### **Pembahasan**

Hasil pengolahan data menunjukkan bahwa proses pra-pemrosesan—meliputi pembersihan data, transformasi logaritmik, dan standardisasi—berhasil menstabilkan sebaran nilai dan meningkatkan kualitas input untuk analisis clustering. Pola distribusi yang sangat lebar pada data awal, terutama pada variabel luas tanah, NJOP, dan nilai PBB, dapat dikendalikan melalui transformasi sehingga algoritma dapat bekerja lebih proporsional.

Implementasi K-Medoids menghasilkan empat klaster utama yang mencerminkan segmentasi wajib pajak berdasarkan karakteristik ekonomi dan nilai objek pajak. Keberhasilan ini didukung oleh nilai evaluasi Silhouette Coefficient dan Davies–Bouldin Index yang menunjukkan pemisahan klaster cukup baik. Secara visual, hasil PCA juga memperlihatkan batas antarklaster yang jelas, menandakan struktur data yang memang memiliki pola pembeda alami.

Interpretasi profil klaster menunjukkan adanya kelompok wajib pajak beraset tinggi, wajib pajak menengah, kelompok pemilik lahan luas tanpa bangunan, serta kelompok dengan kemampuan ekonomi rendah. Pola ini logis jika dikaitkan dengan literatur mengenai pengelolaan PBB yang menekankan adanya variasi kemampuan bayar dan tingkat pemanfaatan lahan di setiap wilayah. Kelompok beraset besar terbukti memberikan kontribusi PBB tertinggi, sementara kelompok kecil membutuhkan pendekatan pembinaan untuk meningkatkan kepatuhan.

Secara keseluruhan, hasil penelitian mendukung pemahaman bahwa penggunaan metode clustering mampu membantu pemerintah daerah memetakan potensi fiskal secara lebih akurat. Temuan ini selaras dengan prinsip pengelolaan pajak modern yang mengutamakan analisis berbasis data untuk menyusun strategi intensifikasi maupun ekstensifikasi pajak yang lebih tepat sasaran.

### **KESIMPULAN**

Penelitian ini menunjukkan bahwa pengolahan data melalui pembersihan, transformasi logaritmik, dan standardisasi berhasil meningkatkan kualitas dataset sehingga lebih siap digunakan dalam proses clustering. Penerapan algoritma K-Medoids mampu membentuk empat

klaster wajib pajak yang mencerminkan perbedaan karakteristik ekonomi dan nilai objek pajak di Kecamatan Tanjung. Hasil evaluasi menggunakan Silhouette Coefficient dan Davies–Bouldin Index menunjukkan bahwa struktur klaster yang terbentuk cukup baik, stabil, dan mampu merepresentasikan pola alami yang terdapat dalam data.

Interpretasi profil klaster memperlihatkan adanya kelompok wajib pajak beraset tinggi, kelompok lahan luas tanpa bangunan, kelompok menengah, serta kelompok ekonomi rendah dengan nilai PBB minimal. Temuan ini menegaskan bahwa pendekatan data mining dapat memberikan gambaran segmentasi wajib pajak yang lebih akurat dan bermanfaat sebagai dasar penyusunan strategi pengelolaan PBB yang lebih efektif, baik dalam aspek intensifikasi, ekstensifikasi, maupun perumusan kebijakan fiskal berbasis data.

## SARAN

Penelitian ini masih memiliki beberapa keterbatasan, sehingga diperlukan pengembangan lebih lanjut. Pertama, penggunaan variabel tambahan seperti riwayat pembayaran, status kepemilikan, dan klasifikasi jenis lahan dapat meningkatkan akurasi klaster yang terbentuk. Kedua, metode klasterisasi lain seperti DBSCAN, HDBSCAN, atau model berbasis pembelajaran mesin terintegrasi dapat digunakan sebagai pembanding untuk memperoleh struktur pengelompokan yang lebih komprehensif. Ketiga, penelitian selanjutnya dapat memperluas cakupan wilayah atau melakukan analisis temporal untuk melihat perubahan perilaku wajib pajak dari tahun ke tahun. Terakhir, integrasi hasil klasterisasi dengan sistem informasi pajak daerah diharapkan dapat membantu pengambilan keputusan yang lebih efektif dalam pengelolaan PBB.

## DAFTAR PUSTAKA

- Adhim, P. (2023). Analisis Pengaruh Efektifitas Dan Efisiensi Terhadap Pendapatan Asli Daerah Kabupaten Sanggau Dengan Kontribusi Pajak Bumi Dan Bangunan Pedesaan Dan Perkotaan Sebagai Variabel Mediasi. *Equator Journal of Management and Entrepreneurship (Ejme)*, 11(03), 162–174. <https://doi.org/10.26418/ejme.v11i03.62784>
- Akbar, S. (2021). Evaluasi Pelayanan Pembayaran Pajak Bumi Dan Bangunan Terhadap Pencapaian Target Pendapatan Pajak Di Kota Jambi. *Tanah Pilih*, 1(1), 1–10. <https://doi.org/10.30631/tpj.v1i1.673>
- Ayu, M., Kuswarak, & Ikbah, M. (2022). Penyuluhan Mengenai Pendataan Dan Pemutakhiran Data Pajak Bumi Dan Bangunan (PBB-P2) Di Kelurahan Sukadanaham. *I-Com Indonesian Community Journal*, 2(3), 689–694. <https://doi.org/10.33379/icom.v2i3.1902>
- Azhar, R., & Zakiah, K. (2025). Dampak Berita Penggelapan Dana Pajak Oleh Rafael Alun Terhadap Wajib Pajak. *Jurnal Riset Manajemen Komunikasi*, 87–94. <https://doi.org/10.29313/jrmk.v5i1.5785>
- Darajat, R. P., & Sofianty, D. (2023). Pengaruh Penerapan Digitalisasi Pajak Terhadap Kepatuhan Wajib Pajak Orang Pribadi PBB Di Kota Bandung Dengan Sosialisasi Perpajakan Sebagai Variabel Moderating. *Bandung Conference Series Accountancy*, 3(1). <https://doi.org/10.29313/bcsa.v3i1.6499>
- Farissa, R. A., Mayasari, R., & Umaidah, Y. (2021). Perbandingan Algoritma K-Means Dan K-Medoids Untuk Pengelompokan Data Obat Dengan Silhouette Coefficient Di Puskesmas Karangsembung. *Journal of Applied Informatics and Computing*, 5(2), 109–116. <https://doi.org/10.30871/jaic.v5i1.3237>
- Fatemi, M., Feng, E., Sharma, C., Azher, Z., Goel, T., Ramwala, O. A., Palisoul, S., Barney, R. E.,

- Perreard, L., Kolling, F., Salas, L. A., Christensen, B. C., Tsongalis, G. J., Vaickus, L., & Levy, J. (2023). Inferring Spatial Transcriptomics Markers From Whole Slide Images to Characterize Metastasis-Related Spatial Heterogeneity of Colorectal Tumors: A Pilot Study. *Journal of Pathology Informatics*, 14, 100308. <https://doi.org/10.1016/j.jpi.2023.100308>
- Fatemi, M., Feng, E., Sharma, C., Azher, Z. L., Goel, T., Ramwala, O. A., Palisoul, S. M., Barney, R. E., Perreard, L., Kolling, F., Salas, L. A., Christensen, B. C., Tsongalis, G. J., Vaickus, L., & Levy, J. J. (2022). *Inferring Spatially Resolved Transcriptomics Data From Whole Slide Images for the Assessment of Colorectal Tumor Metastasis: A Feasibility Study*. <https://doi.org/10.1101/2022.11.24.517856>
- Gantara, N. P., & Ali, I. (2023). Penerapan Metode K-Means Clustering Pada Penjualan Barang Di Sports Station. *E-Link Jurnal Teknik Elektro Dan Informatika*, 18(1), 28. <https://doi.org/10.30587/e-link.v18i1.5339>
- Hasan, C. A., Abdussamad, Z., & Nani, Y. N. (2025). Implementasi Kebijakan Pengelolaan Pajak Bumi Dan Banguna Perdesaan Perkotaan (PBB-P2) Di Badan Pendapatan Daerah Kabupaten Gorontalo. *Riggs Journal of Artificial Intelligence and Digital Business*, 4(3), 5662–5671. <https://doi.org/10.31004/riggs.v4i3.2848>
- Herviany, M., Delima, S. P., Nurhidayah, T., & Kasini, K. (2021). Perbandingan Algoritma K-Means Dan K-Medoids Untuk Pengelompokan Daerah Rawan Tanah Longsor Pada Provinsi Jawa Barat. *Malcom Indonesian Journal of Machine Learning and Computer Science*, 1(1), 34–40. <https://doi.org/10.57152/malcom.v1i1.60>
- Isaenko, N., Colombaroni, C., Fusco, G., & Lahijanian, Z. (2024). Statistical and Clustering-Based Assessment of Variable Speed Limits Effects on Motorway Performance From Real-World Observations. *Future Transportation*, 4(2), 409–428. <https://doi.org/10.3390/futuretransp4020020>
- Martins, A. S. C., de Araujo, L. R., & Penido, D. R. R. (2024). K-Medoids clustering applications for high-dimensionality multiphase probabilistic power flow. *International Journal of Electrical Power and Energy Systems*, 157, 109861.
- Meilinda, E., & Jayanti, W. E. (2022). Peran Waterfall Sebagai Metode Pengembangan Perangkat Lunak Pada Sistem Informasi Pendataan Pajak Bumi Dan Bangunan. *Jurnal Teknologi Informasi Mura*, 14(2), 144–155. <https://doi.org/10.32767/jti.v14i2.1853>
- Mulyani, F., Sartika, D., & Ilyas, A. (2023). Faktor-Faktor Yang Mempengaruhi Kepatuhan Wajib Pajak Dalam Membayar PBB-P2 Di Kelurahan Batuang Taban XX Kecamatan Lubuk Begalung Kota Padang. *Jurnal Ekonomi Dan Bisnis Dharma Andalas*, 25(1), 153–170. <https://doi.org/10.47233/jebd.v25i1.746>
- Nurhani, Fatmawati, L. S., & Yusuf, M. (2024). Penegakan Hukum Terhadap Pelanggaran Wajib Pajak Bumi Dan Bangunan Pada Badan Pendapatan Daerah Kota Kendari. *Arus Jurnal Sosial Dan Humaniora*, 4(1), 97–115. <https://doi.org/10.57250/ajsh.v4i1.346>
- Orisa, M., & Faisol, A. (2021). Analisis Algoritma Partitioning Around Medoid Untuk Penentuan Klusterisasi. *Jurnal Teknologi Informasi Dan Terapan*, 8(2), 86–90. <https://doi.org/10.25047/jtit.v8i2.258>
- Psdku, T. K. (2023). Analisis Efektivitas Dan Kontribusi Pajak Bumi Dan Bangunan Terhadap Pendapatan Asli Daerah Kabupaten Maluku Barat Daya. *Kupna*, 3(2), 82–92. <https://doi.org/10.30598/kupna.v3.i2.p82-92>
- Rahmawati, O., & Fauzan, A. (2024). Provincial Clustering Based on Education Indicators: K-Medoids Application and K-Medoids Outlier Handling. *Barekeng Jurnal Ilmu Matematika Dan Terapan*, 18(2), 1167–1178. <https://doi.org/10.30598/barekengvol18iss2pp1167-1178>
- Ramadevi, M., & Das, A. (2024). Noise-resistant face recognition using hybrid deep learning and

- K-Medoids clustering approach. *Heliyon*, 10(4), e15210. <https://doi.org/10.1016/j.heliyon.2024.e15210>
- Ramadhani, W., & Pusposari, D. (2024). Faktor-Faktor Yang Mempengaruhi Kepatuhan Wajib Pajak Dalam Membayar Pajak Bumi Dan Bangunan Perdesaaan Dan Perkotaan (Pbb-P2) (Studi Pada Wajib Pajak Orang Pribadi Di Kota Bontang). *Tiara*, 2(3), 383–398. <https://doi.org/10.21776/tiara.2024.2.3.133>
- Rizal, A. R., Indriati, R., & Andriyanto, T. (2022). Sistem Informasi Pajak Bumi Bangunan Desa. *Jurnal Tecnoscienza*, 7(1), 1–14. <https://doi.org/10.51158/tecnoscienza.v7i1.790>
- Sari, L. A., Hakim, A. R., & Rusgiyono, A. (2025). Penggunaan Index Calinski-Harabasz Pada Clustering K-medoids Algorithm Untuk Penggolongan Kabupaten/ Kota Di Provinsi Jawa Tengah Berdasarkan Karakteristik Penduduk Miskin. *Jurnal Gaussian*, 14(1), 179–187. <https://doi.org/10.14710/j.gauss.14.1.179-187>
- Sartono, A. Z., & Kudus, A. (2025). Pengelompokan Perangkat Daerah Di Jawa Barat Berdasarkan Jenis Pajak Menggunakan Metode K-Medoids. *Bandung Conference Series Statistics*, 5(1), 61–70. <https://doi.org/10.29313/bcss.v5i1.16792>
- Selni, N. I., Stefanus, K. Y., & Udju, H. R. (2024). Pengaturan Kewenangan Pengelolaan Pajak Bumi Dan Bangunan Dan Implikasinya Terhadap Pendapatan Asli Daerah Kota Kupang. *JHB*, 13(05), 1–7. <https://doi.org/10.47709/jhb.v13i05.4600>
- Septian, R., & Nohe, D. A. (2023). Penerapan Algoritma K-Medoids Pada Pengelompokan Wilayah Provinsi Di Indonesia Berdasarkan Indikator Pendidikan. *Eksponensial*, 14(2), 85. <https://doi.org/10.30872/eksponensial.v14i2.1150>
- Sulistiyawati, A. A. D., & Sadikin, M. (2021). Penerapan Algoritma K-Medoids Untuk Menentukan Segmentasi Pelanggan. *Sistemasi*, 10(3), 516. <https://doi.org/10.32520/stmsi.v10i3.1332>
- Wibowo, R. A., Nisa, K., Venelia, H., & Warsono, W. (2022). Robust Clustering of Covid-19 Pandemic Worldwide. *Barekeng Jurnal Ilmu Matematika Dan Terapan*, 16(2), 687–694. <https://doi.org/10.30598/barekengvol16iss2pp687-694>
- Widiya, P., Rahaningsih, N., & Suarna, N. (2023). Pengelompokan Tingkat Kepatuhan Wajib Pajak Bumi Dan Bangunan Menggunakan Algoritma K-Means. *Jati (Jurnal Mahasiswa Teknik Informatika)*, 7(2), 1358–1362. <https://doi.org/10.36040/jati.v7i2.6836>
- Zheng, Q., Xu, Y., Liu, H., Shi, B., Wang, J., & Dong, B. (2024). A Survey of Tax Risk Detection Using Data Mining Techniques. *Engineering*, 34, 43–59. <https://doi.org/10.1016/j.eng.2023.07.014>