Jurnal Cyber PR, Volume 6, No. 2, Desember 2026, Hal. 352–366
ISSN (online) 2798 - 0561

Penerapan Algoritma K-MEANS untuk Pengelompokan Potensi PBB-P2 Kabupaten Nabire

Usman Arfan*, Muhammad Thoyyib Ramadhan

STMIK Pesat Nabire
Jl. Poros Samabusa, Sanoba, Nabire, Papua Tengah 98816, Indonesia.
*Email Korespondensi: usmanarfanpesat@gmail.com

Submitted: 06-29-2026, Accepted: 11-04-2026, Published: 11-12-2026

Abstract *- This study aims to classify the administrative profile of Rural and Urban Land and Building Tax (PBB-P2) across 35 villages and urban wards in Nabire Regency using 2025 data obtained from the Regional Revenue Agency of Nabire Regency. The classification was conducted using the K-Means algorithm with three clusters: high, medium, and low. The analyzed variables comprised the number of Tax Due Notification Letters (SPPT), the number of Tax Assessment Register Books (DHKP), and the nominal amount of tax payable. The data were standardized using Z-scores and analyzed with Orange Data Mining. The clustering results comprised six areas in the high cluster, 16 in the medium cluster, and 13 in the low cluster. A silhouette coefficient of 0.45 indicated a moderate degree of cluster separation. However, the number of DHKP books was constant across all areas and therefore provided no discriminatory information for clustering. Accordingly, the results are more appropriately interpreted as an administrative profile of PBB-P2 based on the available data rather than a comprehensive measure of tax revenue potential, taxpayer compliance, arrears, or untapped revenue potential. The findings may provide preliminary information for local government in prioritizing the evaluation and administration of PBB-P2.*
Keyword: K-Means; Clustering; Orange Data Mining; Land and Building Tax; Regional Revenue

Abstrak - Penelitian ini bertujuan mengelompokkan profil administratif Pajak Bumi dan Bangunan Perdesaan dan Perkotaan (PBB-P2) pada 35 desa/kelurahan di Kabupaten Nabire berdasarkan data tahun 2025 dari Badan Pendapatan Daerah Kabupaten Nabire. Pengelompokan dilakukan menggunakan algoritma K-Means dengan tiga kelompok (cluster), yaitu tinggi, sedang, dan rendah. Variabel yang dianalisis meliputi jumlah Surat Pemberitahuan Pajak Terutang (SPPT), jumlah Daftar Himpunan Ketetapan Pajak (DHKP), dan nominal pajak terutang. Data diproses melalui standardisasi Z-score dan dianalisis menggunakan Orange Data Mining. Hasil pengelompokan menghasilkan enam wilayah dalam kelompok tinggi, 16 wilayah dalam kelompok sedang, dan 13 wilayah dalam kelompok rendah. Nilai silhouette coefficient sebesar 0,45 menunjukkan bahwa pemisahan antarkelompok berada pada tingkat moderat. Namun, jumlah DHKP bernilai konstan pada seluruh wilayah sehingga tidak memberikan daya pembeda dalam proses pengelompokan. Oleh karena itu, hasil penelitian lebih tepat diinterpretasikan sebagai profil administratif PBB-P2 berdasarkan karakteristik data yang tersedia, bukan sebagai ukuran menyeluruh atas potensi penerimaan pajak, tingkat kepatuhan wajib pajak, tunggakan, atau potensi penerimaan yang belum tergali. Hasil penelitian dapat menjadi informasi awal bagi pemerintah daerah dalam menentukan prioritas evaluasi dan pengelolaan administrasi PBB-P2.
Kata kunci: K-Means; Data Mining; Orange Data Mining; PBB-P2; Pengelompokan Wilayah

Pendahuluan

Pajak Bumi dan Bangunan Perdesaan dan Perkotaan (PBB-P2) merupakan salah satu sumber pendapatan daerah yang dapat mendukung pembiayaan pembangunan dan penyelenggaraan pemerintahan daerah. Pengelolaan PBB-P2 membutuhkan informasi yang memadai mengenai karakteristik objek pajak dan sebaran kewajiban pajak pada setiap wilayah. Informasi tersebut dapat membantu pemerintah daerah menyusun prioritas pendataan, pengawasan, dan pelayanan perpajakan. (OECD, 2022) menjelaskan bahwa penguatan

DOI: 10.32509/cyberpr.v6i2.7597
CC BY-NC-SA 4.0. Copyright © 2026 The Author(s).

administrasi pajak properti berkaitan dengan pengelolaan data properti, sistem penilaian, dan pemanfaatan informasi untuk mendukung tata kelola fiskal daerah. Sejalan dengan perkembangan digitalisasi administrasi perpajakan, pemanfaatan analitik data juga dapat mendukung proses identifikasi pola dan pengambilan keputusan berbasis data (OECD, 2025).

Salah satu pendekatan yang dapat digunakan untuk menganalisis karakteristik data perpajakan adalah *data mining*, khususnya algoritma K-Means. Algoritma ini mengelompokkan objek berdasarkan kemiripan karakteristik data sehingga objek dengan karakteristik serupa ditempatkan dalam kelompok yang sama. (Ikotun dkk., 2023) menjelaskan bahwa K-Means merupakan salah satu algoritma pengelompokan yang banyak digunakan karena relatif sederhana dan dapat diterapkan pada berbagai jenis permasalahan analisis data. Dalam konteks penelitian ini, pengelompokan dapat digunakan untuk memetakan profil administratif PBB-P2 antardesa atau kelurahan berdasarkan indikator yang tersedia. Namun, kualitas hasil pengelompokan tetap dipengaruhi oleh pemilihan variabel, karakteristik data, standardisasi, dan penentuan jumlah klaster.

Pengelolaan pajak properti di tingkat daerah tidak hanya berkaitan dengan penetapan kewajiban pajak, tetapi juga dengan kualitas data dan kapasitas pemerintah dalam mengidentifikasi potensi penerimaan. (Von Haldenwang, 2017) menunjukkan bahwa pemerintah daerah di Indonesia masih menghadapi sejumlah kendala dalam mengoptimalkan pajak properti, termasuk permasalahan data dan besarnya kewajiban pajak yang belum tertangani. Dalam konteks tersebut, pengolahan data administratif secara sistematis dapat membantu pemerintah memahami karakteristik wilayah sebagai dasar evaluasi lebih lanjut. (OECD, 2020) juga menekankan pentingnya transformasi digital dalam administrasi perpajakan untuk mendukung proses administrasi yang lebih terintegrasi dan berbasis data. Oleh karena itu, analisis pengelompokan data PBB-P2 dapat menjadi salah satu pendekatan untuk memetakan karakteristik administratif wilayah, meskipun hasil pengelompokan tersebut tidak dengan sendirinya menunjukkan realisasi penerimaan atau tingkat kepatuhan wajib pajak.

Penelitian terdahulu telah menerapkan K-Means untuk menganalisis data PBB dengan fokus dan lokasi yang beragam. (Nursaniah dkk., 2024) mengelompokkan data penerimaan PBB berdasarkan kelurahan di Kota Tasikmalaya sehingga menghasilkan gambaran sebaran penerimaan antarwilayah. (Primadeni dkk., 2024) membahas segmentasi wilayah berdasarkan karakteristik PBB di Kabupaten Sragen, sedangkan (Anwar dkk., 2025) menerapkan K-Means untuk menentukan prioritas operasional penagihan PBB di Desa Gunasari. Penelitian lain menempatkan kepatuhan pembayaran sebagai fokus pengelompokan. (Putri dkk., 2022) mengelompokkan data wajib pajak untuk mengetahui tingkat kepatuhan pembayaran PBB di Desa Kendal, Kabupaten Cirebon, sedangkan (Widiya dkk., 2023) mengelompokkan tingkat kepatuhan wajib pajak PBB di Desa Sende, Kecamatan Arjawinangun. Secara umum, kajian-kajian tersebut membentuk tiga orientasi besar dalam penerapan K-Means pada data PBB, yaitu pengelompokan berbasis realisasi penerimaan, segmentasi karakteristik wilayah, dan pengelompokan kepatuhan wajib pajak (Putri dkk., 2022; Widiya dkk., 2023).

Dengan menempatkan kelima kajian tersebut pada satu kerangka perbandingan, perbedaan dengan penelitian ini dapat diuraikan secara eksplisit pada beberapa dimensi. Dari sisi unit analisis, (Nursaniah dkk., 2024) bekerja pada tingkat kelurahan di dalam satu kota, (Primadeni dkk., 2024) pada wilayah kabupaten dengan indikator karakteristik objek pajak, (Anwar dkk., 2025) pada satu desa, sedangkan (Putri dkk., 2022) dan (Widiya dkk., 2023) pada wajib pajak individual. Dari sisi variabel, kajian-kajian tersebut bertumpu pada data realisasi penerimaan atau indikator perilaku pembayaran, sedangkan penelitian ini menggunakan indikator sisi penetapan berupa volume SPPT, ketersediaan DHKP, dan nominal pajak terutang. Dari sisi prosedur analisis, penelitian ini melengkapi proses pengelompokan dengan standardisasi Z-score, evaluasi Silhouette Coefficient, serta penelaahan eksplisit atas kontribusi

setiap variabel dalam pembentukan klaster. Dari sisi hasil dan kegunaan praktis, penelitian ini menghasilkan pemetaan profil administratif untuk seluruh 35 desa dan kelurahan di Kabupaten Nabire sebagai dasar penentuan prioritas evaluasi administrasi, pemutakhiran data, dan perencanaan pelayanan perpajakan. Berdasarkan penelusuran literatur yang dilakukan pada rentang tahun 2022 hingga 2025, kajian yang memadukan indikator sisi penetapan pajak dengan cakupan seluruh desa dan kelurahan dalam satu kabupaten memang belum ditemukan. Kesimpulan kesenjangan tersebut perlu dibaca dalam batas penelusuran yang dilakukan, tetapi cukup menunjukkan bahwa kombinasi unit analisis, indikator, dan tujuan praktis seperti pada penelitian ini belum mendapat perhatian dalam kajian terdahulu.

Potensi PBB-P2 dalam penelitian ini dibatasi sebagai profil administratif wilayah yang dibentuk berdasarkan jumlah Surat Pemberitahuan Pajak Terutang (SPPT), jumlah Daftar Himpunan Ketetapan Pajak (DHKP), dan nominal pajak terutang yang tercatat. Ketiga indikator tersebut dipilih karena masing-masing merekam tahapan penting dalam proses ketetapan pajak properti sehingga secara bersamaan dapat menggambarkan kesiapan administrasi pajak daerah. Jumlah SPPT menunjukkan cakupan penetapan karena setiap SPPT merupakan pemberitahuan resmi kepada wajib pajak atas kewajiban yang telah ditetapkan. Jumlah DHKP menunjukkan ketersediaan sarana pencatatan karena DHKP merupakan daftar rekapitulasi resmi yang menampung seluruh ketetapan pajak pada suatu wilayah. Nominal pajak terutang menunjukkan besarnya kewajiban yang telah ditetapkan dan menjadi dasar bagi penerimaan yang diharapkan. Kerangka yang serupa digunakan dalam literatur administrasi pajak properti yang menempatkan kelengkapan daftar objek pajak, proses penilaian, dan penetapan kewajiban sebagai fondasi potensi penerimaan pajak properti (Norregaard, 2013). Pentingnya pengelolaan data properti dan sistem penilaian juga ditekankan dalam kajian administrasi pajak properti (OECD, 2022). Meskipun demikian, ketiga indikator tersebut hanya merepresentasikan aspek administratif yang tersedia dalam dataset dan tidak mencakup seluruh dimensi potensi penerimaan PBB-P2 karena tidak mengukur realisasi pembayaran, tunggakan, kepatuhan wajib pajak, maupun potensi penerimaan yang belum tergali. Oleh sebab itu, hasil pengelompokan dalam penelitian ini ditafsirkan sebagai profil administratif PBB-P2 dan bukan ukuran komprehensif potensi penerimaan pajak daerah.

Pemanfaatan data administratif juga perlu ditempatkan dalam kerangka pengelolaan informasi yang transparan dan dapat digunakan untuk mendukung keputusan. (Giffari dkk., 2024) membahas peran hubungan masyarakat dalam keterbukaan informasi publik pada PPID Kemendikbudristek. Walaupun penelitian tersebut tidak membahas algoritma pengelompokan pajak, kajiannya memberikan konteks mengenai pentingnya pengelolaan informasi publik dalam lingkungan pemerintahan. Dalam penelitian ini, hasil klaster dapat menjadi informasi pendukung bagi pemerintah daerah untuk memahami perbedaan profil administratif antarwilayah, tetapi penelitian tidak mengukur efektivitas komunikasi publik maupun dampak langsung terhadap penerimaan pajak.

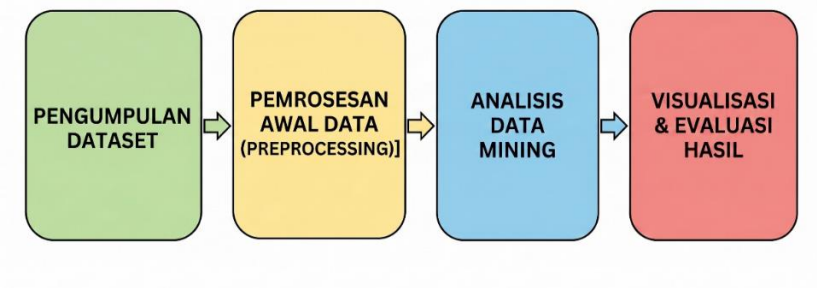
Berdasarkan pemetaan kesenjangan tersebut, penelitian ini menawarkan tiga kontribusi yang saling terkait. Pertama, penelitian ini mengintegrasikan tiga indikator sisi penetapan, yaitu SPPT, DHKP, dan nominal pajak terutang, dalam satu analisis pengelompokan sehingga menghasilkan profil administratif yang tidak dapat diberikan oleh masing-masing indikator secara terpisah. Kedua, penelitian ini memetakan profil administratif pada seluruh 35 desa dan kelurahan di Kabupaten Nabire sehingga cakupan analisis bersifat populasi pada tingkat kabupaten dan bukan sampel pada satu wilayah tertentu. Ketiga, penelitian ini menunjukkan cara penggunaan hasil pengelompokan secara hati-hati sebagai dasar penentuan prioritas pengelolaan administrasi PBB-P2 sekaligus menegaskan batas interpretasinya terhadap penerimaan dan kepatuhan. Dengan ketiga kontribusi tersebut, penelitian ini memperluas

perspektif analisis PBB-P2 yang selama ini bertumpu pada sisi realisasi dan perilaku menuju sisi ketetapan administratif yang menjadi fondasi proses pemungutan pajak daerah.

Berdasarkan uraian tersebut, penelitian ini bertujuan mengelompokkan 35 desa dan kelurahan di Kabupaten Nabire berdasarkan indikator administratif PBB-P2 menggunakan algoritma K-Means. Data tahun 2025 yang bersumber dari BAPENDA Kabupaten Nabire diolah melalui standardisasi Z-score dan dianalisis menggunakan Orange Data Mining. Jumlah klaster ditetapkan sebanyak tiga kelompok, dan kualitas pemisahan klaster dievaluasi menggunakan Silhouette Score. Penelitian ini diharapkan memberikan gambaran awal mengenai distribusi profil administratif PBB-P2 antardesa dan kelurahan serta menyediakan informasi pendukung bagi perencanaan pengelolaan data perpajakan daerah.

Metodologi

Penelitian ini menggunakan pendekatan kuantitatif dengan teknik data mining untuk mengidentifikasi pola potensi Pajak Bumi dan Bangunan Perdesaan dan Perkotaan (PBB-P2) berdasarkan karakteristik wilayah di Kabupaten Nabire. Pendekatan kuantitatif dipilih karena memungkinkan proses analisis dilakukan secara objektif melalui pengolahan data numerik, sedangkan teknik data mining digunakan untuk menemukan pola pengelompokan yang tidak mudah diidentifikasi melalui analisis konvensional. Implementasi analisis dilakukan menggunakan perangkat lunak Orange Data Mining yang menyediakan workflow visual sehingga setiap tahapan analisis dapat dilakukan secara sistematis, terdokumentasi, dan mudah direplikasi. Secara umum penelitian dilaksanakan melalui empat tahapan utama, yaitu pengumpulan data, preprocessing data, implementasi algoritma K-Means, dan evaluasi kualitas hasil clustering. Alur penelitian secara keseluruhan disajikan pada Gambar 1.



Gambar 1. Diagram Tahapan Penelitian

Pengumpulan Dataset

Data penelitian merupakan data sekunder yang diperoleh dari Badan Pendapatan Daerah (BAPENDA) Kabupaten Nabire berupa data Pajak Bumi dan Bangunan Perdesaan dan Perkotaan (PBB-P2) tahun 2025. Unit analisis mencakup 35 desa dan kelurahan di Kabupaten Nabire. Setiap unit analisis direpresentasikan melalui tiga atribut, yaitu jumlah Surat Pemberitahuan Pajak Terutang (SPPT), jumlah Daftar Himpunan Ketetapan Pajak (DHKP), dan nominal pajak terutang. Jumlah SPPT menggambarkan volume dokumen penetapan pajak, sedangkan nominal pajak terutang menggambarkan besarnya kewajiban pajak yang tercatat pada setiap wilayah. Berdasarkan data yang diperoleh, setiap desa dan kelurahan memiliki satu buku DHKP. Dengan demikian, nilai DHKP sama pada seluruh unit analisis sehingga variabel tersebut tidak memiliki variasi yang dapat membedakan karakteristik antarwilayah. Kondisi ini dicatat sebagai karakteristik dataset dan menjadi pertimbangan dalam menginterpretasikan hasil clustering. Dataset lengkap seluruh 35 wilayah disajikan dalam Tabel S1 pada bagian Supplementary Material untuk mendukung transparansi dan keterlacakan data penelitian.

Dataset penelitian mencakup tiga variabel administratif, yaitu jumlah SPPT, jumlah DHKP, dan nominal pajak terutang. Jumlah SPPT dan nominal pajak terutang digunakan sebagai atribut dalam proses clustering, sedangkan jumlah DHKP digunakan sebagai variabel pembanding administratif. Jumlah DHKP pada seluruh 35 desa/kelurahan bernilai satu buku sehingga tidak memiliki variasi antarwilayah dan tidak digunakan untuk membedakan hasil pengelompokan.

Pra-pemrosesan Data (*Preprocessing*)

Tahap preprocessing bertujuan meningkatkan kualitas data sebelum dilakukan proses clustering. Seluruh data terlebih dahulu diperiksa untuk memastikan tidak terdapat data ganda, nilai kosong maupun ketidaksesuaian format data. Selanjutnya dilakukan proses standardisasi menggunakan metode Z-Score Standardization melalui widget Preprocess pada Orange Data Mining (Desyanti dkk., 2024). Standardisasi diperlukan karena kedua atribut yang digunakan dalam clustering, yaitu jumlah SPPT dan nominal pajak terutang, memiliki rentang nilai yang berbeda secara signifikan. Jumlah SPPT bernilai ratusan hingga ribuan, sedangkan nominal pajak terutang dapat mencapai ratusan juta rupiah. Proses standardisasi menghasilkan data dengan skala yang sebanding sehingga setiap atribut memiliki kontribusi yang proporsional dalam proses perhitungan jarak antar objek (Hartono & Lusiana, 2026). Meskipun standardisasi menyetarakan skala variabel yang memiliki rentang nilai berbeda, proses tersebut tidak menciptakan variasi pada variabel yang nilainya konstan. Oleh karena itu, nilai DHKP yang sama pada seluruh wilayah tetap tidak memiliki daya pembeda setelah standardisasi. Interpretasi hasil pengelompokan perlu mempertimbangkan kondisi tersebut, khususnya dalam menjelaskan kontribusi setiap variabel terhadap pembentukan klaster.

Implementasi Algoritma K-Means

Proses clustering dilakukan menggunakan algoritma K-Means yang tersedia pada Orange Data Mining (Sapitri & Afrilia, 2025). Algoritma ini dipilih karena memiliki kemampuan mengelompokkan data berdasarkan tingkat kemiripan karakteristik dengan proses komputasi yang relatif sederhana serta mudah diinterpretasikan. Pada penelitian ini jumlah klaster ditetapkan sebanyak tiga kelompok (K = 3) untuk merepresentasikan kategori potensi PBB-P2, yaitu potensi tinggi, potensi sedang, dan potensi rendah. Penetapan tiga klaster didasarkan pada kebutuhan klasifikasi potensi wilayah yang digunakan sebagai dasar interpretasi kebijakan pengelolaan pajak daerah.

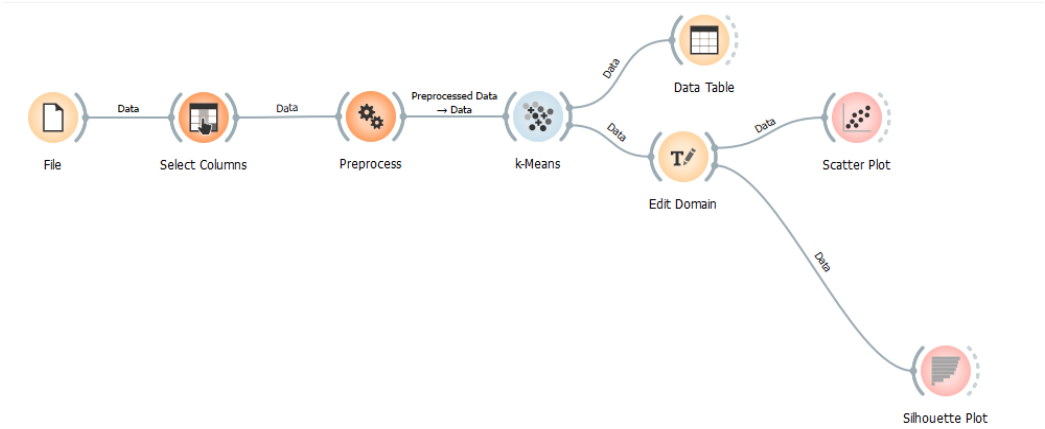
Dalam interpretasi variabel, jumlah DHKP yang bernilai satu buku pada seluruh wilayah diperlakukan sebagai atribut administratif yang tidak memiliki variasi. Oleh karena itu, variabel tersebut tidak dapat digunakan untuk menjelaskan perbedaan antarwilayah. Karakteristik klaster terutama perlu ditafsirkan berdasarkan variabel yang memiliki variasi, yaitu jumlah SPPT dan nominal pajak terutang. Penetapan kategori tinggi, sedang, dan rendah merupakan interpretasi terhadap hasil pengelompokan, bukan klasifikasi berdasarkan ambang batas penerimaan pajak yang telah ditetapkan sebelumnya.

Proses clustering diawali dengan penentuan centroid awal, kemudian dilanjutkan dengan perhitungan jarak setiap objek terhadap centroid menggunakan *Euclidean Distance* sebagaimana persamaan (1).

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (xi - yi)^2}$$

Setiap objek kemudian ditempatkan pada centroid terdekat. Selanjutnya centroid diperbarui berdasarkan rata-rata anggota pada masing-masing klaster. Proses iterasi dilakukan

secara berulang hingga tidak terjadi perubahan keanggotaan klaster sehingga diperoleh hasil pengelompokan yang stabil. Workflow implementasi algoritma pada Orange Data Mining ditunjukkan pada Gambar 2.



Gambar 2. Alur Kerja (Workflow) K-Means pada Orange Data Mining

Evaluasi Kualitas Clustering

Kualitas hasil clustering dievaluasi menggunakan Silhouette Coefficient untuk mengukur tingkat kesesuaian setiap objek terhadap klaster yang ditempatinya dibandingkan dengan klaster lain (Darmawan dkk., 2026). Nilai Silhouette dihitung menggunakan Persamaan (2).

$$S(i) = (b(i) - a(i))/max(a(i), b(i))$$

dengan (a(i)) merupakan rata-rata jarak objek terhadap anggota dalam klaster yang sama, sedangkan (b(i)) merupakan rata-rata jarak objek terhadap klaster terdekat lainnya. Nilai Silhouette berada pada rentang −1 hingga 1, di mana nilai yang semakin mendekati 1 menunjukkan kualitas pengelompokan yang semakin baik karena setiap objek memiliki tingkat kesamaan yang tinggi dengan anggota klasternya serta memiliki perbedaan yang jelas terhadap klaster lainnya. Hasil evaluasi ini digunakan untuk menilai kualitas model clustering yang dihasilkan sebelum dilakukan interpretasi terhadap karakteristik masing-masing kelompok wilayah.

Hasil dan Pembahasan

Preprocessing Data dan Standardisasi

Implementasi algoritma K-Means pada penelitian ini diawali dengan tahap preprocessing untuk menyamakan skala setiap variabel sebelum proses pengelompokan dilakukan. Tahap ini menggunakan metode Z-Score Standardization melalui Orange Data Mining terhadap dua variabel yang digunakan dalam clustering, yaitu jumlah Surat Pemberitahuan Pajak Terutang (SPPT) dan nominal pajak terutang. Sementara itu, jumlah Daftar Himpunan Ketetapan Pajak (DHKP) tetap disajikan sebagai variabel pembanding administratif karena nilainya sama pada seluruh 35 desa dan kelurahan. Standardisasi diperlukan karena variabel numerik yang memiliki rentang nilai berbeda dapat memberikan kontribusi yang tidak seimbang terhadap perhitungan jarak Euclidean apabila tidak ditransformasikan terlebih dahulu. Dengan standardisasi, setiap variabel dapat dipertimbangkan berdasarkan penyimpangannya terhadap nilai rata-rata, bukan semata-mata besarnya skala pengukuran.

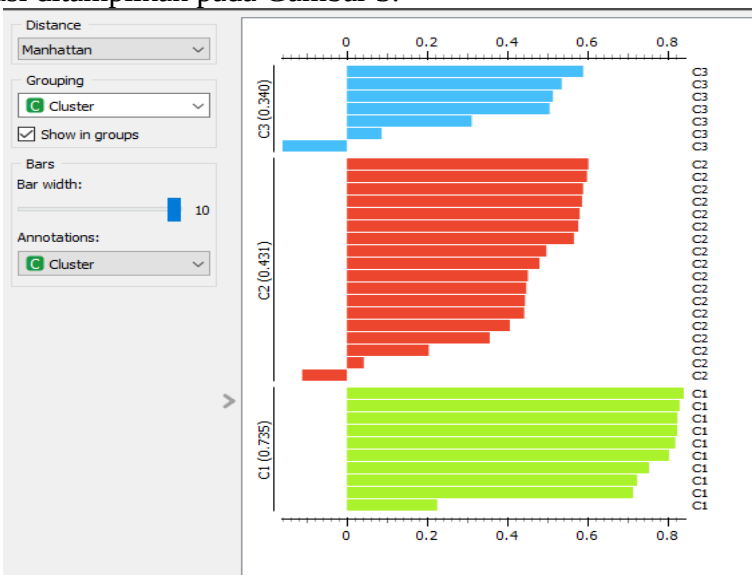
Standardisasi merupakan salah satu tahapan penting dalam penerapan K-Means pada data numerik. (Desyanti dkk., 2024) menerapkan K-Means berbasis Z-Score untuk mengelompokkan tingkat gizi balita, sedangkan (Hartono & Lusiana, 2026) membahas evaluasi

metode penentuan jumlah klaster dan stabilitas pengelompokan. Kedua penelitian tersebut relevan sebagai landasan metodologis untuk memperhatikan tahap pengolahan data dan evaluasi hasil klaster. Dalam penelitian ini, standardisasi digunakan untuk mengurangi pengaruh perbedaan skala jumlah SPPT dan nominal pajak terutang terhadap proses pengelompokan. Jumlah DHKP tetap digunakan sebagai informasi pembanding administratif, tetapi tidak menjadi atribut pembeda karena nilainya konstan pada seluruh wilayah penelitian. Meskipun demikian, standardisasi tidak dengan sendirinya menjamin bahwa setiap variabel memiliki daya pembeda yang sama; kontribusi masing-masing variabel tetap bergantung pada distribusi dan variasi datanya.

Karakteristik dataset penelitian menunjukkan bahwa jumlah DHKP bernilai satu buku pada seluruh 35 desa dan kelurahan. Kondisi ini berarti DHKP tidak memiliki variasi antardaerah sehingga tidak dapat menjelaskan perbedaan profil administratif yang terbentuk. Meskipun DHKP tetap relevan sebagai informasi mengenai administrasi ketetapan pajak, nilai konstan tersebut tidak memberikan informasi pembeda dalam analisis clustering. Dengan demikian, hasil pengelompokan perlu dibaca terutama sebagai segmentasi berdasarkan variasi jumlah SPPT dan nominal pajak terutang. Temuan ini juga menunjukkan bahwa kelengkapan indikator administratif tidak selalu sejalan dengan kemampuan setiap indikator untuk membedakan karakteristik objek yang dianalisis.

Evaluasi Kualitas Clustering

Setelah tahap standardisasi, proses pengelompokan dilakukan menggunakan algoritma K-Means dengan jumlah klaster sebanyak tiga (K = 3). Ketiga klaster tersebut selanjutnya diinterpretasikan sebagai kategori potensi administratif PBB-P2 tinggi, sedang, dan rendah berdasarkan karakteristik variabel yang digunakan. Evaluasi kualitas pengelompokan dilakukan menggunakan *Silhouette Coefficient*, yaitu ukuran evaluasi internal yang mempertimbangkan kedekatan suatu objek terhadap anggota klasternya sendiri dan perbedaannya terhadap klaster terdekat. Hasil evaluasi ditampilkan pada Gambar 3.



Gambar 3. Silhouette Plot Hasil Pengelompokan K-Means

Hasil evaluasi algoritma K-Means menggunakan *Silhouette Coefficient* menunjukkan nilai sebesar 0,45. Nilai ini menunjukkan bahwa struktur klaster memiliki tingkat pemisahan yang moderat. Artinya, terdapat kemiripan karakteristik di antara objek dalam klaster yang sama, tetapi pemisahan antara sebagian klaster belum sepenuhnya tegas. (Rousseeuw, 1987) menjelaskan bahwa *Silhouette Coefficient* mengevaluasi struktur klaster dengan membandingkan kedekatan suatu objek terhadap kelompoknya sendiri dan jaraknya terhadap

klaster terdekat. Oleh karena itu, nilai 0,45 merupakan indikator kualitas struktur pengelompokan, bukan persentase akurasi klasifikasi.

Nilai Silhouette tersebut perlu diinterpretasikan bersama karakteristik variabel yang digunakan. Dalam dataset ini, DHKP memiliki nilai konstan pada seluruh wilayah sehingga tidak memberikan daya pembeda antarklaster. Oleh karena itu, nilai Silhouette sebesar 0,45 menggambarkan kualitas pemisahan berdasarkan konfigurasi data dan variabel yang dianalisis, bukan bukti bahwa seluruh indikator memiliki kontribusi yang sama atau bahwa model telah menjadi konfigurasi terbaik. Evaluasi tambahan diperlukan apabila peneliti ingin membandingkan hasil dari konfigurasi variabel atau jumlah klaster yang berbeda.

Temuan tersebut relevan dengan kajian (Ikotun dkk., 2023), yang menguraikan karakteristik, kelebihan, dan keterbatasan berbagai varian algoritma K-Means. Salah satu aspek yang perlu diperhatikan adalah sensitivitas hasil pengelompokan terhadap karakteristik data dan konfigurasi algoritma. Persamaan kedua kajian terletak pada pengakuan bahwa kualitas klaster tidak hanya ditentukan oleh terbentuknya kelompok, tetapi juga oleh seberapa baik kelompok tersebut terpisah. Perbedaannya, (Ikotun dkk., 2023) menyajikan tinjauan komprehensif mengenai K-Means dalam berbagai bidang, sedangkan penelitian ini menerapkannya pada data administratif PBB-P2 di Kabupaten Nabire.

Dalam konteks penelitian ini, nilai *Silhouette Coefficient* sebesar 0,45 menunjukkan bahwa tiga kelompok yang terbentuk memiliki struktur pengelompokan yang dapat diinterpretasikan, tetapi kualitas pemisahannya masih perlu diperhatikan. Kondisi tersebut mungkin berkaitan dengan kemiripan karakteristik antarwilayah atau distribusi nilai variabel yang digunakan. Namun, penyebab tersebut belum dapat dipastikan hanya berdasarkan nilai Silhouette. Pemeriksaan lebih lanjut terhadap distribusi data, nilai *Silhouette* individual, dan karakteristik *centroid* setiap klaster diperlukan untuk menjelaskan faktor yang memengaruhi kualitas pemisahan kelompok.

Nilai Silhouette Coefficient sebesar 0,45 perlu ditafsirkan dengan mempertimbangkan karakteristik data dan keterbatasan metrik validasi internal. (Ahmed dkk., 2020) menguraikan bahwa K-Means memiliki sejumlah keterbatasan, antara lain kebutuhan menentukan jumlah klaster terlebih dahulu dan sensitivitas terhadap inisialisasi pusat klaster. Selain itu, (Gagolewski dkk., 2021) menunjukkan bahwa indeks validasi internal tidak selalu menghasilkan pengelompokan yang sesuai dengan struktur atau interpretasi substantif yang diharapkan. (Liang dkk., 2020) juga menekankan bahwa sebagian indeks validasi memiliki asumsi tertentu terhadap bentuk, ukuran, kepadatan, dan pemisahan klaster. Sejalan dengan itu, (Hassan dkk., 2024) membahas beragam indeks validasi dan pentingnya memilih ukuran evaluasi sesuai karakteristik data. Dengan demikian, nilai Silhouette 0,45 dalam penelitian ini dapat digunakan sebagai indikasi kualitas struktur pengelompokan, tetapi tidak cukup untuk menyimpulkan bahwa tiga klaster merupakan konfigurasi paling optimal tanpa perbandingan dengan alternatif jumlah klaster dan evaluasi tambahan.

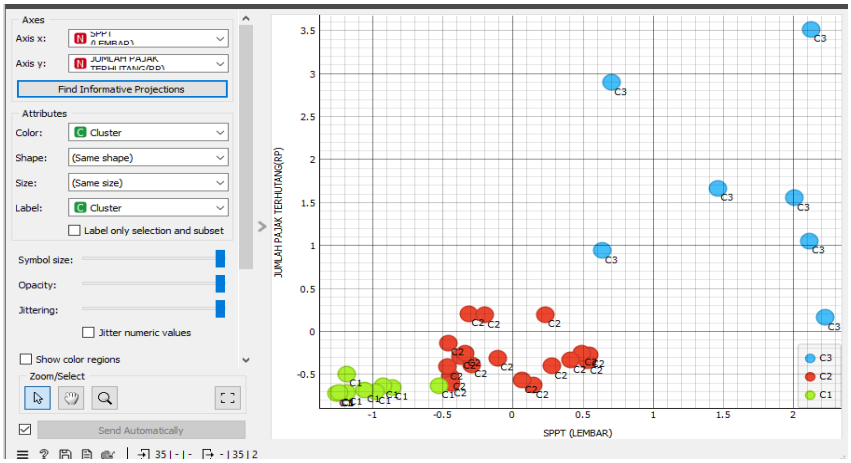
Dari sisi kontribusi metodologis, evaluasi ini memberikan informasi mengenai kualitas model sebelum hasilnya digunakan sebagai bahan pertimbangan pengelolaan PBB-P2. Model tidak cukup dinilai hanya berdasarkan terbentuknya tiga kelompok, tetapi juga perlu dievaluasi berdasarkan kemampuan membedakan karakteristik wilayah. Penelitian selanjutnya dapat membandingkan beberapa konfigurasi jumlah klaster dan mengevaluasinya menggunakan *Silhouette Coefficient* serta metrik validasi internal lainnya. (Hartono & Lusiana, 2026) membahas penggunaan metode Elbow SSE, Silhouette Score, dan Jaccard Stability dalam pemilihan jumlah klaster optimal. Kajian tersebut relevan sebagai landasan bagi penelitian lanjutan untuk menguji apakah tiga klaster merupakan konfigurasi yang paling sesuai dengan struktur data PBB-P2 di Kabupaten Nabire.

Hasil Pengelompokan Wilayah

Tabel 1. Hasil Pengelompokan Wilayah Berdasarkan Algoritma K-Means

Klaster	Kategori Potensi	Jumlah Wilayah	Kelurahan/Desa
C1	Potensi Tinggi	**6**	Kalibobo, Girimulyo, Oyehe, Nabarua, Siriwini, Kali Semen
C2	Potensi Sedang	**16**	Sanoba, Morgo, Karang Tumaritis, Karang Mulia, Samabusa, Kali Harapan, Kalisusu, Wiraska, Bumi Mulia, Karadiri, Margo Jaya, Kimi, Wahario, Bumi Raya, Wadio, Legari Jaya
C3	Potensi Rendah	**13**	Air Mandidi, Wanggar Sari, Wanggar Makmur, Argo Mulyo, Topo Jaya, Gamei Jaya, Lani, Waroki, Maidei, Manunggal Jaya, Makimi, Nipasi, Biha

Penerapan algoritma K-Means menghasilkan tiga kelompok dari 35 desa dan kelurahan di Kabupaten Nabire. Posisi relatif setiap wilayah pada ruang atribut yang digunakan dalam proses pengelompokan, yaitu jumlah SPPT dan nominal pajak terutang yang telah distandardisasi, divisualisasikan pada Gambar 4. Berdasarkan Tabel 1, klaster C1 terdiri atas enam wilayah dengan kategori potensi tinggi, klaster C2 mencakup 16 wilayah dengan kategori potensi sedang, dan klaster C3 mencakup 13 wilayah dengan kategori potensi rendah. Secara proporsional, ketiganya masing-masing mewakili sekitar 17,1%, 45,7%, dan 37,1% dari keseluruhan wilayah penelitian. Distribusi tersebut menunjukkan bahwa klaster dengan kategori sedang memiliki jumlah anggota terbanyak, sedangkan klaster dengan kategori tinggi memiliki jumlah anggota paling sedikit.



Gambar 4. Visualisasi Sebaran Wilayah pada Ruang Atribut Jumlah SPPT dan Nominal Pajak Terutang

Hasil pengelompokan menunjukkan adanya perbedaan karakteristik administratif PBB-P2 antarwilayah. Kelompok tinggi memiliki rata-rata jumlah SPPT sebesar 1.133,33 dan rata-rata nominal pajak terutang sebesar Rp474.517.167,17. Kelompok sedang memiliki rata-rata jumlah SPPT sebesar 501,31 dan nominal pajak terutang sebesar Rp105.012.521,00, sedangkan kelompok rendah memiliki rata-rata jumlah SPPT sebesar 133,23 dan nominal pajak terutang sebesar Rp17.810.272,62. Perbedaan tersebut menunjukkan variasi karakteristik administratif PBB-P2 berdasarkan data yang dianalisis.

Untuk memperjelas perbedaan karakteristik administratif pada setiap kelompok, Tabel 2 menyajikan jumlah wilayah serta rata-rata jumlah SPPT, DHKP, dan nominal pajak terutang berdasarkan hasil pengelompokan.

Tabel 2. Rata-rata Karakteristik PBB-P2 Berdasarkan Hasil Clustering

Kelompok	Jumlah wilayah	Rata-rata SPPT	Rata-rata DHKP	Rata-rata pajak terutang
Tinggi (C1)	6	1.133,33	1,00	Rp474.517.167,17
Sedang (C2)	16	501,31	1,00	Rp105.012.521,00
Rendah (C3)	13	133,23	1,00	Rp17.810.272,62

Berdasarkan Tabel 2, perbedaan karakteristik antarklaster terutama terlihat pada jumlah SPPT dan nominal pajak terutang. Klaster C1 memiliki rata-rata jumlah SPPT dan nominal pajak terutang tertinggi, diikuti C2 dan C3. Sebaliknya, rata-rata jumlah DHKP sama pada seluruh klaster, yaitu satu buku, sehingga variabel tersebut tidak menjelaskan perbedaan karakteristik antarklaster. Hasil ini menunjukkan bahwa kategori tinggi, sedang, dan rendah merepresentasikan profil administratif berdasarkan data yang dianalisis, bukan ukuran definitif potensi penerimaan PBB-P2.

Meskipun terdapat perbedaan karakteristik antar kelompok, interpretasi hasil perlu mempertimbangkan keterbatasan variabel yang digunakan. Jumlah DHKP pada seluruh wilayah bernilai satu buku sehingga variabel tersebut tidak dapat menjelaskan perbedaan antarkelompok. Selain itu, nominal pajak terutang tidak secara langsung menggambarkan realisasi penerimaan, tingkat pembayaran wajib pajak, jumlah tunggakan, maupun potensi pajak yang belum tergali. Oleh karena itu, kategori tinggi, sedang, dan rendah dalam penelitian ini lebih tepat dipahami sebagai profil administratif berdasarkan data yang tersedia, bukan sebagai ukuran definitif potensi penerimaan PBB-P2.

Temuan utama penelitian ini adalah terbentuknya tiga kelompok wilayah dengan profil administratif yang terpisah cukup jelas, di mana rata-rata jumlah SPPT dan nominal pajak terutang bergerak naik secara konsisten dari kelompok rendah, kelompok sedang, hingga kelompok tinggi. Temuan serupa dilaporkan (Nursaniah dkk., 2024) yang mengelompokkan data penerimaan PBB berdasarkan kelurahan di Kota Tasikmalaya dan juga memperoleh kelompok wilayah dengan karakteristik penerimaan yang berbeda. Persamaan kedua penelitian terletak pada penggunaan K-Means untuk memetakan ketimpangan kondisi perpajakan pada tingkat wilayah administratif. Perbedaannya terletak pada variabel dan makna klaster karena penelitian tersebut mengelompokkan nilai penerimaan yang telah dibayarkan, sedangkan penelitian ini mengelompokkan volume penetapan dan kewajiban yang tercatat. Perbedaan tersebut bersumber dari sisi data yang digunakan karena realisasi mencerminkan hasil pembayaran yang dipengaruhi perilaku wajib pajak, sedangkan pajak terutang mencerminkan ketetapan yang dihasilkan oleh proses administrasi. Kontribusinya, penelitian ini melengkapi kajian perpajakan berbasis wilayah dengan perspektif administratif yang tetap dapat digunakan pada wilayah dengan keterbatasan data realisasi.

Hasil pengelompokan pada penelitian ini juga membentuk gradasi bertingkat dengan dominasi kelompok sedang, sebagaimana pola segmentasi wilayah yang dilaporkan (Primadeni dkk., 2024) untuk karakteristik PBB di Kabupaten Sragen. Persamaan keduanya terletak pada penggunaan pengelompokan untuk memahami perbedaan karakteristik wilayah berdasarkan data PBB. Perbedaannya terletak pada indikator yang digunakan dan konteks wilayah karena kajian tersebut berada di Jawa dengan basis data properti yang lebih matang, sedangkan penelitian ini dilakukan di Kabupaten Nabire. Perbedaan konteks tersebut menjadi alasan mengapa bentuk dan komposisi klaster pada kedua wilayah tidak harus sama karena struktur ekonomi lokal, jumlah objek pajak, dan kualitas data administratif turut menentukan hasil pengelompokan. Kontribusinya, penelitian ini memperluas bukti empiris penerapan K-Means pada data PBB-P2 di luar Jawa dengan cakupan seluruh desa dan kelurahan dalam satu kabupaten.

Posisi penelitian ini juga perlu dibedakan dari kajian kepatuhan seperti (Putri dkk., 2022) yang menerapkan clustering pada data wajib pajak untuk mengetahui kelompok kepatuhan pembayaran PBB serta (Widiya dkk., 2023) yang mengelompokkan tingkat kepatuhan wajib pajak PBB menggunakan K-Means. Persamaan ketiganya terletak pada upaya menyederhanakan data perpajakan menjadi kelompok-kelompok yang karakteristiknya relatif serupa. Perbedaannya terletak pada konstruk yang dianalisis karena kajian kepatuhan membutuhkan indikator yang menggambarkan perilaku atau pemenuhan kewajiban pembayaran, sedangkan penelitian ini mengelompokkan wilayah berdasarkan jumlah dokumen administratif dan nominal kewajiban yang ditetapkan.

Perbedaan konstruk tersebut memberikan pelajaran penting bagi interpretasi hasil penelitian ini. Makna sebuah klaster selalu bergantung pada indikator yang membentuknya sehingga wilayah pada klaster berkategori administratif tinggi tidak otomatis memiliki kepatuhan atau realisasi penerimaan yang tinggi, dan wilayah pada klaster berkategori rendah tidak dapat langsung dinyatakan memiliki kepatuhan yang rendah. Perbedaan indikator juga membuat hasil pengelompokan antarpenelitian tidak dapat dibandingkan secara langsung dari sisi kualitas maupun implikasi fiskalnya karena setiap kumpulan indikator membentuk ruang jarak yang berbeda sehingga komposisi anggota, bentuk klaster, dan nilai evaluasi pemisahannya menjadi tidak sebanding. Oleh karena itu, interpretasi pada penelitian ini selalu mengacu pada indikator yang benar-benar dianalisis, yaitu profil administratif ketetapan.

Kontribusi praktis penelitian ini dapat dipahami lebih jelas apabila dibandingkan dengan (Anwar dkk., 2025) yang menggunakan K-Means untuk membantu penentuan prioritas operasional penagihan PBB di Desa Gunasari. Persamaannya terletak pada pemanfaatan hasil pengelompokan sebagai informasi pendukung pengelolaan PBB. Perbedaannya terletak pada tingkat kesiapan output terhadap keputusan penagihan karena penelitian tersebut menggunakan indikator yang mengarah pada status pembayaran, sedangkan variabel pada penelitian ini tidak menunjukkan status pembayaran maupun jumlah tunggakan. Alasan perbedaan tersebut bersifat teknis karena penentuan prioritas penagihan menuntut informasi realisasi pembayaran, umur tunggakan, dan riwayat kepatuhan yang belum tersedia dalam dataset penelitian ini. Kontribusinya, penelitian ini menyediakan segmentasi awal profil administratif wilayah yang dapat menjadi titik masuk bagi analisis prioritas penagihan pada penelitian lanjutan.

Secara keseluruhan, rangkaian perbandingan tersebut menegaskan dua hal bagi interpretasi hasil penelitian ini. Pertama, penerapan clustering pada data PBB dapat diarahkan untuk tujuan yang berbeda bergantung pada unit analisis dan indikator yang tersedia sehingga label tinggi, sedang, dan rendah pada penelitian ini menunjukkan urutan profil administratif ketetapan dan bukan urutan penerimaan, kepatuhan, maupun prioritas penagihan. Kedua, perbedaan indikator membuat hasil pengelompokan antarpenelitian tidak dapat langsung disamakan dari sisi kualitas maupun implikasi fiskalnya karena penelitian berbasis penerimaan memberi informasi mengenai hasil pembayaran, penelitian berbasis kepatuhan memberi informasi mengenai perilaku wajib pajak, sedangkan penelitian ini memberi informasi mengenai cakupan dan beban administrasi ketetapan. Kontribusi penelitian ini adalah menyediakan pemetaan awal profil administratif PBB-P2 untuk 35 desa dan kelurahan di Kabupaten Nabire sebagai titik awal evaluasi administrasi dan pemutakhiran data, dengan tetap menyadari bahwa penetapan prioritas penagihan memerlukan data realisasi pembayaran, tunggakan, dan riwayat kepatuhan wajib pajak.

Interpretasi Karakteristik Klaster dan Implikasi Pengelolaan PBB-P2

Hasil *clustering* dapat dimanfaatkan sebagai informasi awal untuk membedakan kebutuhan pengelolaan PBB-P2 antardaerah. Wilayah dalam klaster dengan nilai administratif relatif tinggi dapat menjadi perhatian dalam pemeriksaan kualitas data dan evaluasi kesesuaian

ketetapan pajak. Wilayah dengan karakteristik menengah dapat dianalisis lebih lanjut untuk mengidentifikasi kebutuhan perbaikan administrasi dan pelayanan, sedangkan wilayah dengan nilai administratif relatif rendah dapat ditinjau untuk memastikan kelengkapan data objek pajak. Rekomendasi tersebut merupakan alternatif tindak lanjut berdasarkan profil administratif dan bukan bukti bahwa setiap wilayah dalam kategori tertentu pasti membutuhkan intervensi yang sama.

Implikasi tersebut sejalan secara konseptual dengan (OECD, 2025), yang membahas peran digitalisasi dan analitik data dalam mendukung pengambilan keputusan administrasi pajak. Persamaannya terletak pada pemanfaatan informasi hasil analisis untuk mendukung proses pengelolaan, bukan sekadar menyimpan data administratif. Perbedaannya, laporan OECD membahas perkembangan digitalisasi administrasi pajak di berbagai yurisdiksi, sedangkan penelitian ini menggunakan data pada tingkat desa dan kelurahan di satu kabupaten. Karena perbedaan cakupan tersebut, penelitian ini tidak dimaksudkan untuk membuktikan efektivitas digitalisasi administrasi pajak secara umum, tetapi menunjukkan penerapan analisis *clustering* sebagai salah satu pendekatan untuk mengolah data administratif lokal.

Implikasi pengelompokan administratif juga perlu ditempatkan dalam konteks pengelolaan pajak properti daerah. (Von Haldenwang, 2017) menyoroti bahwa optimalisasi pajak properti di Indonesia dapat terhambat oleh kualitas data, penerimaan masyarakat, dan kewajiban pajak yang belum tertangani. (OECD, 2020) menempatkan transformasi digital sebagai bagian dari pengembangan administrasi perpajakan yang lebih terintegrasi. Dalam penelitian ini, hasil K-Means dapat dimanfaatkan sebagai informasi awal untuk memeriksa perbedaan profil administratif antardesa dan kelurahan. Namun, penentuan prioritas penagihan dan evaluasi efektivitas penerimaan tetap membutuhkan data tambahan, khususnya realisasi pembayaran, tunggakan, dan riwayat kepatuhan. Oleh sebab itu, kontribusi penelitian ini terletak pada pemetaan awal karakteristik administratif PBB-P2, bukan pada pengukuran langsung potensi penerimaan yang belum tergali.

Dari sisi komunikasi pemerintah, hasil pengelompokan dapat membantu BAPENDA menyusun informasi dan pelayanan yang lebih terarah sesuai kebutuhan wilayah. (Giffari dkk., 2024) membahas peran humas dalam keterbukaan informasi publik pada PPID Kemendikbudristek. Persamaan kedua konteks terletak pada kebutuhan mengelola informasi secara terstruktur agar dapat mendukung pelayanan publik. Perbedaannya, penelitian (Giffari dkk., 2024) berfokus pada peran PPID dalam keterbukaan informasi publik, sedangkan penelitian ini menghasilkan segmentasi administratif PBB-P2 dan tidak mengukur efektivitas komunikasi kepada wajib pajak. Oleh karena itu, hasil *clustering* dapat dipertimbangkan sebagai bahan perencanaan komunikasi dan pelayanan, tetapi belum dapat dijadikan bukti bahwa pengelompokan telah meningkatkan pemahaman atau kepatuhan masyarakat.

Meskipun demikian, pemanfaatan hasil *clustering* untuk menentukan prioritas penagihan memerlukan kehati-hatian. Data jumlah SPPT, DHKP, dan nominal pajak terutang belum menunjukkan apakah kewajiban pajak telah dibayarkan, masih menunggak, atau berpotensi tidak tertagih. Oleh sebab itu, penetapan prioritas penagihan perlu mengintegrasikan informasi tambahan, seperti realisasi pembayaran, tunggakan, riwayat kepatuhan, dan perubahan data objek pajak. Tanpa indikator tersebut, hasil penelitian lebih tepat digunakan sebagai dasar awal untuk evaluasi administrasi, pemutakhiran data, dan perencanaan pelayanan.

Secara keseluruhan, penelitian ini menunjukkan bahwa K-Means dapat digunakan untuk membentuk profil administratif PBB-P2 pada tingkat desa dan kelurahan di Kabupaten Nabire. Kontribusi penelitian terletak pada penerapan segmentasi berbasis tiga indikator administratif dalam konteks wilayah yang diteliti serta interpretasi hasilnya sebagai informasi pendukung perencanaan pengelolaan pajak daerah. Namun, nilai Silhouette sebesar 0,45 dan keterbatasan indikator menunjukkan bahwa hasil ini belum cukup untuk mengukur keseluruhan

potensi ekonomi perpajakan atau efektivitas pemungutan. Penelitian selanjutnya perlu memperluas indikator, memeriksa hubungan antarvariabel, dan membandingkan konfigurasi model agar hasil segmentasi lebih kuat secara metodologis dan lebih operasional bagi pemerintah daerah.

Hasil pengelompokan ini dapat dimanfaatkan sebagai informasi awal untuk mengevaluasi administrasi PBB-P2 pada tingkat desa/kelurahan di Kabupaten Nabire. Namun, penentuan prioritas peningkatan penerimaan pajak perlu didukung indikator tambahan, seperti realisasi pembayaran, tunggakan, jumlah wajib pajak aktif, dan tren penerimaan dari beberapa tahun. Penggunaan indikator tersebut dalam penelitian berikutnya diharapkan dapat menghasilkan pemetaan yang lebih komprehensif mengenai kondisi dan potensi penerimaan PBB-P2.

Simpulan

Penelitian ini menerapkan algoritma K-Means untuk mengelompokkan profil administratif PBB-P2 pada 35 desa/kelurahan di Kabupaten Nabire berdasarkan data tahun 2025. Hasil pengelompokan membentuk tiga kelompok, yaitu kelompok tinggi sebanyak enam wilayah, kelompok sedang sebanyak 16 wilayah, dan kelompok rendah sebanyak 13 wilayah. Nilai silhouette coefficient sebesar 0,45 menunjukkan pemisahan kelompok pada tingkat moderat.

Hasil penelitian menggambarkan perbedaan jumlah SPPT dan nominal pajak terutang antarwilayah. Namun, jumlah DHKP yang konstan pada seluruh wilayah tidak memberikan daya pembeda dalam analisis. Selain itu, variabel yang tersedia belum cukup untuk mengukur potensi penerimaan PBB-P2 secara menyeluruh karena belum mencakup realisasi pembayaran, tunggakan, dan tingkat kepatuhan wajib pajak. Oleh karena itu, hasil pengelompokan perlu dipahami sebagai profil administratif awal.

Penelitian selanjutnya disarankan menggunakan indikator yang lebih beragam, termasuk realisasi penerimaan, tunggakan, jumlah wajib pajak aktif, serta data beberapa tahun. Pemerintah daerah juga dapat menggunakan hasil pengelompokan ini sebagai informasi awal untuk mengevaluasi administrasi PBB-P2, dengan tetap mempertimbangkan verifikasi data dan indikator fiskal lainnya sebelum menetapkan prioritas kebijakan.

Daftar Pustaka

Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means Algorithm: A Comprehensive Survey and Performance Evaluation. *Electronics*, *9*(8), 1295. https://doi.org/10.3390/electronics9081295

Anwar, M., Firmansyah, E., & Sutara, B. (2025). Analisis Clustering Menggunakan Metode K-Means dalam Penentuan Prioritas Operasional Penagihan Pajak Bumi dan Bangunan di Desa Gunasari. *Data Sciences Indonesia (DSI)*, *5*(2), 101–110. https://doi.org/10.47709/dsi.v5i2.7464

Darmawan, D. P. D., Putra, C. M., Yahya, S., Jusman, D., Syahrian, A., & Tundjungsari, V. (2026). Analisis Perbandingan Metode K-Means dan Gaussian Mixture Model dalam Pengelompokan Playlist Musik Berbasis Fitur Audio. *Jurnal Informatika dan Teknologi Komputer (JITEK), 6*(1), 213–227. https://doi.org/10.55606/jitek.v6i1.10754

Desyanti, D., Desriyati, W., & Mesran, M. (2024). Implementasi Metode K-Means Clustering dalam Mengukur Tingkat Gizi Balita Berdasarkan Z-Score. *Building of Informatics, Technology and Science (BITS)*, *6*(3), 1808–1817. https://doi.org/10.47065/bits.v6i3.6177

Gagolewski, M., Bartoszuk, M., & Cena, A. (2021). Are cluster validity measures (in) valid? *Information Sciences, 581,* 620–636. https://doi.org/10.1016/j.ins.2021.10.004

Giffari, M. F., Wiyati, E. K., Andriani, F., & Putri, C. E. (2024). Peran Humas dalam Keterbukaan Informasi Publik pada PPID Kemendikbudristek. *Jurnal Cyber PR, 4*(2), 119–131. https://doi.org/10.32509/cyberpr.v4i2.4634

Hartono, B., & Lusiana, V. (2026). Analisis Metode Elbow SSE, Silhouette Score, dan Jaccard Stability dalam Pemilihan Jumlah Klaster Data yang Optimal. *TIN: Terapan Informatika Nusantara*, *6*(8), 1521–1532. https://doi.org/10.47065/tin.v6i8.9271

Hassan, B. A., Tayfor, N. B., Hassan, A. A., Ahmed, A. M., Rashid, T. A., & Abdalla, N. N. (2024). From A-to-Z review of clustering validation indices. *Neurocomputing*, *601*, 128198. https://doi.org/10.1016/j.neucom.2024.128198

Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, *622*, 178–210. https://doi.org/10.1016/j.ins.2022.11.139

Liang, S., Han, D., & Yang, Y. (2020). Cluster validity index for irregular clustering results. *Applied Soft Computing*, *95*, 106583. https://doi.org/10.1016/j.asoc.2020.106583

Nursaniah, R., Rahaningsih, N., Ali, I., & Nuris, D. (2024). Pengelompokan Data Penerimaan Pajak Bumi Dan Bangunan Berdasarkan Kelurahan Di Kota Tasikmalaya Menggunakan Algoritma K-Means. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(2), 1477–1483. https://doi.org/10.36040/jati.v8i2.9009

Norregaard, J. (2013). *Taxing Immovable Property: Revenue Potential and Implementation Challenges* (IMF Working Paper No. WP/13/169). International Monetary Fund. https://www.imf.org/external/pubs/ft/wp/2013/wp13169.pdf

OECD. (2020). *Tax Administration 3.0: The Digital Transformation of Tax Administration*. OECD. https://doi.org/10.1787/ca274cc5-en

OECD. (2022). *Fiscal Federalism 2022: Making Decentralisation Work*. OECD Publishing. https://doi.org/10.1787/201c75b6-en

OECD. (2025). *Tax Administration Digitalisation and Digital Transformation Initiatives*. OECD Publishing. https://doi.org/10.1787/c076d776-en

Primadeni, F., Pradana, A., & Purwanto, E. (2024). Analisis Clustering Untuk Segmentasi Wilayah Berdasarkan Karakteristik PBB di Kabupaten Sragen. *Jurnal Pendidikan dan Teknologi Indonesia*, *4*, 405–417. https://doi.org/10.52436/1.jpti.500

Putri, M. A., Rahaningsih, N., Basysyar, F. M., & Nurdiawan, O. (2022). Penerapan Data Mining Menggunakan Metode Clustering Untuk Mengetahui Kelompok Kepatuhan Wajib Pajak Bumi dan Bangunan. *Jurnal Accounting Information System (AIMS)*, *5*(2), 145–156. https://doi.org/10.32627/aims.v5i2.496

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

Sapitri, A., & Afrilia, Y. (2025). Implementation of Clustering Method Using K-Means Algorithm for Grouping BPJS Health Patient Medical Record Data. *Journal of Applied Informatics and Computing (JAIC)*, *9*(5), 2391–2398.

Von Haldenwang, C. (2017). The Political Cost of Local Revenue Mobilization: Decentralization of the Property Tax in Indonesia. *Public Finance and Management, 17*(2), 124–151. https://doi.org/10.1177/152397211701700202

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

Supplementary Material

Tabel S1. Dataset Administratif PBB-P2 pada 35 Desa dan Kelurahan di Kabupaten Nabire

No.	Desa/Kelurahan	Jumlah SPPT (lembar)	Jumlah DHKP (buku)	Pajak Terutang (Rp)
1	Air Mandidi	151	1	14.163.617
2	Sanoba	354	1	174.437.821
3	Kalibobo	1.266	1	794.222.429
4	Morgo	363	1	63.066.162
5	Karang Tumaritis	397	1	171.831.583
6	Girimulyo	1.260	1	333.259.495
7	Karang Mulia	709	1	312.026.450
8	Oyehe	734	1	678.342.112
9	Nabarua	1.221	1	426.946.468
10	Siriwini	1.017	1	447.851.801
11	Samabusa	559	1	172.453.580
12	Kali Harapan	345	1	88.055.081
13	Kalisusu	301	1	110.736.358
14	Wiraska	577	1	60.145.075
15	Wanggar Sari	298	1	59.164.303
16	Bumi Mulia	431	1	78.170.644
17	Karadiri	626	1	73.691.811
18	Wanggar Makmur	306	1	37.260.755
19	Margo Jaya	526	1	19.085.357
20	Argo Mulyo	31	1	3.230.579
21	Topo Jaya	7	1	479.174
22	Gamei Jaya	3	1	397.580
23	Kimi	333	1	81.592.850
24	Wahario	656	1	89.071.204
25	Lani	126	1	17.808.846
26	Bumi Raya	676	1	84.945.419
27	Kali Semen	1.302	1	166.480.698
28	Wadio	671	1	71.362.243
29	Waroki	77	1	9.046.084
30	Maidei	308	1	22.066.030
31	Manunggal Jaya	107	1	5.111.194
32	Makimi	11	1	2.126.768
33	Nipasi	31	1	43.290.625
34	Biha	276	1	17.387.989
35	Legari Jaya	497	1	29.528.698

Catatan: Data bersumber dari Badan Pendapatan Daerah Kabupaten Nabire tahun 2025. Jumlah DHKP pada seluruh wilayah bernilai satu buku sehingga variabel tersebut tidak menunjukkan variasi antarwilayah.