

SEGMENTASI DAN ANALISIS KETIDAKPATUHAN PEMBAYARAN PAJAK REKLAME KOTA BANDAR LAMPUNG MENGGUNAKAN ALGORITMA K-PROTOTYPES CLUSTERING

Rizky Ahmad Rifai

Program Studi Matematika, Fakultas Sains, Institut Teknologi Sumatera, Lampung Selatan, Indonesia
e-mail: rizky.122160002@student.itera.ac.id

Werry Febrianti

Program Studi Matematika, Fakultas Sains, Institut Teknologi Sumatera, Lampung Selatan, Indonesia
e-mail: werry.febrianti@ma.itera.ac.id*

Abstrak

Pajak reklame menjadi salah satu komponen yang memiliki kontribusi penting pada Pendapatan Asli Daerah (PAD) Kota Bandar Lampung. Terdapat permasalahan berupa tingginya jumlah tunggakan pajak reklame yang mencerminkan adanya ketidakpatuhan pembayaran pada objek pajak. Penelitian ini bertujuan untuk melakukan segmentasi dan analisis terhadap ketidakpatuhan pembayaran pajak reklame dengan memanfaatkan algoritma *K-Prototypes Clustering*. Data pada penelitian ini berasal dari arsip Badan Pendapatan Daerah Kota Bandar Lampung pada masa tagihan April 2025 dengan total 949 objek pajak. Tahapan penelitian meliputi *pre-processing* data (transformasi data kategorik dan normalisasi variabel numerik dengan metode *z-score*), analisis deskriptif, penerapan algoritma *K-Prototypes Clustering*, serta evaluasi hasil menggunakan *Average Silhouette Width* (ASW). Hasil pengelompokan menghasilkan tiga *cluster* utama yang merepresentasikan kategori wajib pajak dengan tingkat ketidakpatuhan berbeda, ditinjau dari variabel pokok tunggakan, keterlambatan, bunga, tahun masa pajak, bulan masa pajak, dan UPTD. Evaluasi menggunakan ASW menghasilkan nilai sebesar 0,422 yang menunjukkan algoritma *K-Prototypes Clustering* sudah cukup baik dalam mengelompokkan objek pajak reklame.

Kata Kunci: Pajak Reklame, *K-Prototypes Clustering*, Segmentasi, ASW.

Abstract

Advertising tax constitutes a significant component contributing to the Local Own-Source Revenue of Bandar Lampung. A persistent problem identified is the high volume of advertising tax arrears, reflecting non-compliance in payment among tax objects. This research aims to perform segmentation and analysis of advertising tax non-compliance by employing the *K-Prototypes Clustering* algorithm. The data for this study were sourced from the archives of the Regional Revenue Agency of Bandar Lampung (Badan Pendapatan Daerah) for the billing period of April 2025, encompassing a total of 949 tax objects. The research methodology includes data *pre-processing* (categorical data transformation and numerical variable normalization using the *z-score* method), descriptive analysis, application of the *K-Prototypes Clustering* algorithm, and result evaluation utilizing the *Average Silhouette Width* (ASW). The clustering results yield three primary clusters, which represent taxpayer categories with varying degrees of non-compliance, analyzed based on the variables: principal arrears, payment delay, interest, tax year, tax month, and Technical Implementation Unit (UPTD). The evaluation using ASW produced a value of 0.422, indicating that the *K-Prototypes Clustering* algorithm is sufficiently effective in grouping the advertising tax objects.

Keywords: Advertising Tax, *K-Prototypes Clustering*, Segmentation, ASW.

PENDAHULUAN

Negara Kesatuan Republik Indonesia merupakan salah satu negara yang dalam sistem pemerintahannya menganut asas desentralisasi, oleh sebab itu setiap daerah memiliki hak dan kewenangan dalam penyelenggaraan otonomi daerah (Widodo et al., 2024). Salah satu bentuk hak

otonomi adalah pengelolaan terhadap pendapatan yang dimiliki oleh daerah tersebut. Pendapatan Asli Daerah (PAD) merupakan salah satu pendapatan yang dikelola oleh daerah, yang terdiri atas pajak daerah, retribusi daerah, hasil pengelolaan kekayaan daerah yang dipisahkan, dan lain-lain PAD yang sah.

Pada tahun 2024, target PAD Kota Bandar Lampung mencapai 66,16%. Pajak daerah memiliki besaran yang paling tinggi dari seluruh komponen

PAD. Realisasi pajak daerah Kota Bandar Lampung tahun 2024 mencapai Rp555.732.059.619,00 atau sekitar 94,68% dari target pajak daerah tahun anggaran 2024 sebesar Rp586.955.121.520,00. Angka tersebut memiliki persentase sebesar 77,55% dari keseluruhan nilai realisasi PAD Kota Bandar Lampung. Berdasarkan data tersebut, dapat diasumsikan bahwa pajak daerah memiliki peran yang sangat besar dalam pertumbuhan pendapatan di Kota Bandar Lampung.

Pajak reklame merupakan salah satu bagian dari pajak daerah yang pengelolaannya berada di tingkat pemerintah kota atau kabupaten dan berperan penting dalam meningkatkan Pendapatan Asli Daerah (PAD). Pada tahun 2024, pendapatan yang dihasilkan dari pajak reklame sebesar Rp31.509.004.033,00 dan meningkat sebesar 2,96% dibandingkan tahun 2023. Berdasarkan data tersebut, pajak reklame memiliki peran yang cukup penting dalam meningkatkan pajak daerah. Masalah yang kemudian muncul adalah masih banyaknya objek pajak di Kota Bandar Lampung yang memiliki tunggakan pembayaran yang jumlahnya mencapai Rp4.002.658.919.

Algoritma *K-Prototypes Clustering* merupakan algoritma pengelompokan yang menggabungkan algoritma *K-Means Clustering* dan *K-Modes Clustering*. Masalah pada segmentasi objek pajak reklame yang memiliki tunggakan, algoritma *K-Prototypes Clustering* lebih relevan dibandingkan kedua algoritma sebelumnya. *K-Prototypes Clustering* lebih relevan digunakan karena dapat menangani pengelompokan yang terdiri atas data numerik dan data kategorikal (Annas et al., 2022). Data yang terdapat pada masalah ini relevan dengan kondisi tersebut, antara lain data besaran pokok tunggakan (data numerik), data masa pajak (data kategorikal), dan sebagainya.

Beberapa studi terdahulu telah menggunakan algoritma *K-Prototypes Clustering* dan terbukti efektif digunakan dalam masalah yang mengharuskan adanya pembentukan segmentasi. Studi pertama pada masalah mengenai kurang meratanya infrastruktur digital di Kabupaten Bekasi, algoritma *K-Prototypes Clustering* berhasil melakukan pengelompokan pada desa-desa di Kabupaten Bekasi berdasarkan infrastruktur digital yang dimiliki tiap desa dengan baik (Purba et al., 2025). Studi kedua dilakukan untuk melakukan segmentasi data gempa

tektonik yang terjadi di Pulau Sulawesi pada tahun 2017 hingga 2020 (Annas et al., 2022). Permasalahan segmentasi tersebut dapat diselesaikan dengan algoritma *K-Prototypes Clustering* berdasarkan kedalaman, kekuatan, dan kategori jarak. Studi ketiga yang pernah dilakukan selanjutnya adalah membentuk segmentasi kabupaten atau kota di Provinsi Sulawesi Selatan berdasarkan indikator kesejahteraan rakyat. Studi tersebut berhasil membentuk dua segmentasi dengan menggunakan karakteristik tingkat partisipasi angkatan kerja, sumber penerangan listrik, persentase penduduk, angka partisipasi sekolah, angka partisipasi murni, sumber air minum, dan pengeluaran per kapita (Rais et al., 2024). Algoritma *K-Prototypes Clustering* juga pernah digunakan dalam melakukan segmentasi nasabah pada suatu perusahaan *multifinance* (Pradhana et al., 2022).

Penelitian ini akan membentuk segmentasi objek pajak reklame Kota Bandar Lampung berdasarkan ketidakpatuhan pembayarannya dan mengevaluasi hasil dari segmentasi yang terbentuk dengan menggunakan metode *Average Silhouette Width*.

KAJIAN TEORI

PAJAK REKLAME

Pajak reklame merupakan salah satu bentuk pajak daerah yang pengelolaan dan pemungutannya berada di bawah wewenang pemerintah kabupaten atau kota. Pajak reklame adalah pajak yang dikenakan atas penyelenggaraan reklame, yaitu benda, alat, perbuatan, atau media lain yang bentuk dan corak ragamnya dirancang untuk tujuan komersial memperkenalkan, menganjurkan, mempromosikan, atau menarik perhatian umum terhadap sesuatu (Undang-Undang Republik Indonesia Nomor 1 Tahun 2022 Tentang Hubungan Keuangan Antara Pemerintah Pusat dan Pemerintahan Daerah, 2022).

NORMALISASI DATA

Normalisasi data merupakan metode statistika yang digunakan untuk menangani data-data yang memiliki perbedaan nilai cukup besar dan tidak normal. Normalisasi data mengubah setiap nilai X menjadi nilai baku Z . Salah satu metode normalisasi data yang sering digunakan adalah metode z -score.

Secara matematis metode z-score dituliskan seperti pada Persamaan (1) (Primadeni et al., 2024).

$$Z = \frac{X - \mu}{\sigma} \quad (1)$$

di mana X = nilai data, μ = rata-rata data, dan σ = simpangan baku data.

Data yang telah ditransformasi akan memiliki distribusi dengan mean 0 dan simpangan baku 1, membuat variabel dengan skala yang berbeda akan sebanding. Metode ini sangat berguna dalam algoritma yang sensitif terhadap skala seperti K-NN, regresi, atau *clustering*.

K-PROTOTYPES CLUSTERING

K-Prototypes Clustering merupakan pengembangan metode pengelompokan yang menggabungkan algoritma *K-Means Clustering* dan *K-Modes Clustering*. *K-Prototypes Clustering* digunakan untuk menangani pengelompokan dengan data campuran, yaitu data numerik dan data kategorik. Perhitungan jarak pada algoritma *K-Prototypes Clustering* merupakan penggabungan antara jarak *Euclidean* dan jarak *simple matching*.

$$d(b_i, a_t) = \sqrt{\sum_{j=1}^l (b_{ij} - a_{tj})^2} + \gamma \cdot \sum_{j=1}^m \delta(a_{tj}, b_{ij}) \quad (2)$$

di mana,

$$\delta(a_t, b_i) = \begin{cases} 0, & a_t = b_i \\ 1, & a_t \neq b_i \end{cases} \quad (3)$$

dengan perhitungan nilai γ seperti Persamaan (4).

$$\gamma = \frac{1}{l} \sum_{i=1}^l \sigma_i \quad (4)$$

di mana a_t = nilai data, b_i = pusat *cluster* (*centroid*), l = banyaknya variabel bertipe numerik, m = banyaknya variabel bertipe kategorik, dan γ = koefisien penimbang.

Berikut merupakan langkah-langkah dalam melakukan segmentasi data campuran menggunakan algoritma *K-Prototypes Clustering*.

1. Menentukan nilai k .
2. Menentukan nilai koefisien penimbang γ
3. Menentukan pusat *cluster* atau *centroid* awal.

4. Menghitung jarak objek tiap *centroid* menggunakan Persamaan (2).
5. Mengelompokan objek berdasarkan jarak minimum
6. Menentukan *centroid* baru dan membandingkan dengan *centroid* sebelumnya.

AVERAGE SILHOUETTE WIDTH

Algoritma *K-Prototypes Clustering* memerlukan validasi untuk mengevaluasi hasil dari *cluster* yang dibentuk. Metode *Average Silhouette Width* merupakan salah satu metode yang dapat digunakan sebagai evaluasi pada analisis hasil *clustering*. Metode *Average Silhouette Width* menilai sejauh mana objek lebih cocok dengan *cluster* yang telah ditempati daripada *cluster* lain berdasarkan *Silhouette score* (Davies Raihannabil, 2024). Nilai ASW berada pada rentang $[-1,1]$, semakin dekat nilai ASW ke +1 maka hasil *clustering* akan semakin baik (Purba et al., 2025).

METODE

DATA PENELITIAN

Pada penelitian ini, data yang digunakan adalah data tunggakan atau piutang objek pajak reklame pada masa tagihan April 2025. Yang berasal dari Badan Pendapatan Daerah (Bapenda) Kota Bandar Lampung.

Berdasarkan data yang penelitian yang didapatkan, terdapat 10 kolom data, dengan hanya 3 kolom data yang penulis gunakan untuk penelitian di laporan ini. Data-data tersebut antara lain UPTD, masa pajak, dan pokok tunggakan. Selain itu data yang digunakan dalam penelitian ini berjumlah sebanyak 949 baris data. Tabel 1 menampilkan variabel penelitian yang digunakan dalam penelitian ini.

Tabel 1. Variabel Penelitian

Variabel	Keterangan	Tipe Data	Satuan
X_1	UPTD	Kategorik	-
X_2	Tahun masa pajak	Kategorik	-
X_3	Bulan masa pajak	Kategorik	-
X_4	Keterlambatan	Numerik	Bulan
X_5	Pokok Tunggakan	Numerik	Rupiah
X_6	Bunga Tunggakan	Numerik	Rupiah

RANCANGAN PENELITIAN

Langkah-langkah yang dilakukan dalam menyelesaikan permasalahan untuk mencapai tujuan penelitian ini sebagai berikut.

1. Melakukan *input* data penelitian
2. Melakukan transformasi data variabel kategorik dan normalisasi data variabel numerik
3. Menentukan jumlah *cluster* k
4. Menentukan nilai koefisien penimbang γ
5. Menginisialisasi *centroid* awal
6. Melakukan proses *clustering* menggunakan algoritma *K-Prototypes Clustering*. Proses ini terdiri dari perhitungan jarak tiap data ke *centroid* dan mengelompokannya berdasarkan jarak terdekat hingga didapatkan hasil *clustering* yang konvergen
7. Mengevaluasi hasil *clustering* menggunakan metode *Average Silhouette Width*
8. Membuat visualisasi dan analisis segmentasi pada hasil *clustering*

HASIL DAN PEMBAHASAN

PROSES K-PROTOTYPES CLUSTERING

Proses pertama yang dilakukan adalah mentransformasi data kategorik, transformasi dilakukan dengan mengubah kategori setiap data menjadi angka sederhana. Selain itu dilakukan juga normalisasi data numerik menggunakan Persamaan (1). Data yang telah ditransformasi dan dinormalisasi selanjutnya digunakan dalam perhitungan algoritma *K-Prototypes Clustering*.

Pada penelitian ini digunakan nilai k (jumlah *cluster*) sebanyak $k = 3$, agar data objek pajak reklame dapat dibagi ke dalam segmentasi ketidakpatuhan rendah, sedang, dan tinggi. Selain itu, dilakukan juga perhitungan nilai koefisien penimbang (γ). Ketiga variabel numerik memiliki simpangan baku $\sigma_4 = \sigma_5 = \sigma_6 = 1,00053$. Berikut perhitungan nilai γ .

$$\gamma = \frac{1}{3}(1,00053 + 1,00053 + 1,00053) = 1,00053$$

Inisialisasi *centroid* awal yang digunakan pada iterasi pertama algoritma *K-Prototypes Clustering* ditampilkan pada Tabel 2.

Tabel 2. Centroid Iterasi Pertama

Centroid	C1	C2	C3
x_1	2	4	2
x_2	1	5	2
x_3	3	1	1
x_4	-0,274	2,577	0,430
x_5	-0,105	-0,042	5,420
x_6	-0,182	0,718	5,375

Pada akhir iterasi pertama, 949 data telah berhasil dikelompokkan dalam tiga *cluster*. *Cluster* 1 terdiri atas 845 data, *cluster* 2 sebanyak 87 data, dan *cluster* 3 sebanyak 17 data. Langkah berikutnya adalah menentukan *centroid* untuk iterasi kedua. Perhitungan *centroid* dilakukan dengan menentukan modus variabel kategorik dan mean variabel numerik.

Hasil dari perhitungan *centroid* iterasi kedua menghasilkan tiga *centroid* seperti yang ditampilkan pada Tabel 3.

Tabel 3. Centroid Iterasi Kedua

Centroid	C1	C2	C3
x_1	2	4	2
x_2	1	5	2
x_3	3	1	1
x_4	-0,272	2,560	0,430
x_5	-0,099	-0,098	5,420
x_6	-0,171	0,614	5,375

Jika dibandingkan *centroid* iterasi pertama dan iterasi kedua, terlihat ada perbedaan untuk variabel numerik pada *centroid cluster* 1 dan *cluster* 2. Oleh karena itu algoritma *K-Prototypes Clustering* dilanjutkan ke iterasi kedua.

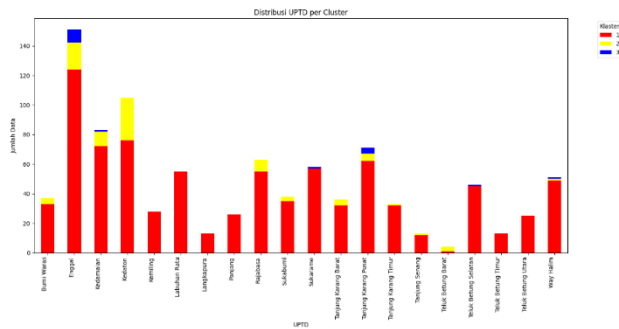
Pengelompokan iterasi kedua menghasilkan jumlah data yang sama seperti pada iterasi pertama untuk setiap *cluster*. Hal tersebut mengindikasikan bahwa algoritma *K-Prototypes Clustering* telah konvergen pada iterasi kedua. Perhitungan yang dilakukan untuk *centroid* iterasi ketiga juga menunjukkan tidak terdapat adanya perubahan di ketiga *centroid*. Oleh karena itu algoritma *K-Prototypes Clustering* dihentikan pada iterasi kedua.

EVALUASI HASIL CLUSTERING

Nilai *Average Silhouette Width* (ASW) yang didapatkan dalam pengelompokan objek pajak reklame Kota Bandar Lampung menggunakan algoritma *K-Prototypes Clustering* dengan 3 *cluster*

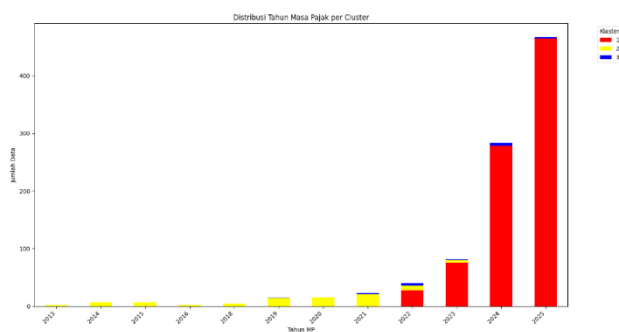
sebesar 0,422. Berdasarkan nilai tersebut, algoritma *k-Prototypes Clustering* dengan $k = 3$ sudah cukup baik dalam mengelompokkan data objek pajak.

VISUALISASI DAN ANALISIS SEGMENTASI



Gambar 1. Visualisasi Hasil *Clustering* Variabel X_1

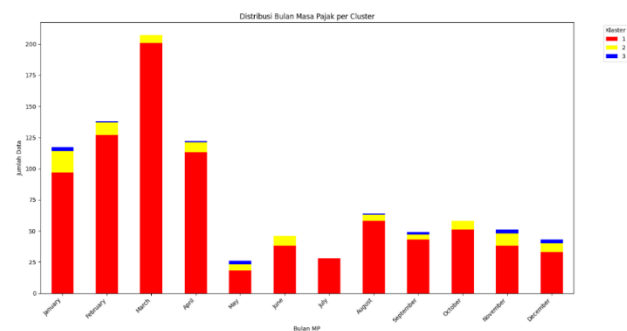
Visualisasi hasil *clustering* pada variabel X_1 (UPTD) menunjukkan bahwa hampir seluruh UPTD didominasi oleh *cluster* 1 (merah), baik pada UPTD dengan jumlah data besar maupun kecil. Contohnya, UPTD Enggal memiliki 124 objek pajak dan UPTD Kedaton memiliki 76 objek pajak yang termasuk *cluster* 1. Sementara pada *cluster* 2 dan *cluster* 3 memiliki distribusi yang tidak merata dan hanya muncul pada beberapa UPTD tertentu dengan jumlah relatif kecil; misalnya *cluster* 2 hanya muncul di UPTD seperti Bumi Waras, Enggal, Kedamaian, Kedaton, Rajabasa, Sukabumi, Teluk Betung Barat, Tanjung Senang, Way Halim, dan seluruh UPTD wilayah Tanjung Karang, dengan jumlah tertinggi hanya 29 objek pajak. Kondisi ini menunjukkan adanya kelompok data dengan karakteristik berbeda dari mayoritas, sekaligus menegaskan bahwa objek pajak reklame yang memiliki tunggakan di Kota Bandar Lampung cenderung homogen dan terkonsentrasi pada *cluster* 1.



Gambar 2. . Visualisasi Hasil *Clustering* Variabel X_2

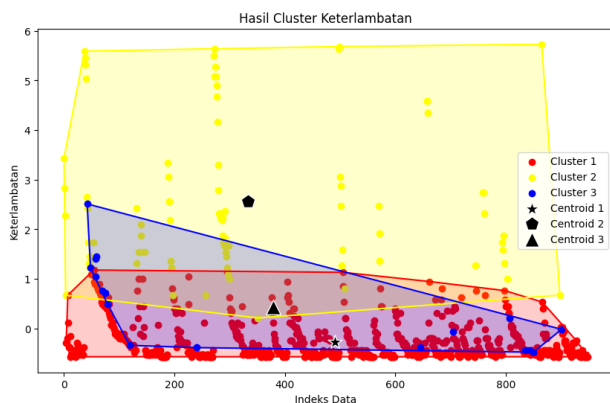
Hasil visualisasi tahun masa pajak menggambarkan pola yang sangat jelas terkait

perkembangan data dan pengelompokannya dalam *cluster* dari tahun ke tahun. Pada periode awal antara tahun 2013 hingga 2021, hampir seluruh objek pada dikelompokkan ke dalam *cluster* 2 (warna kuning) dengan sedikit kontribusi dari *cluster* 3 (warna biru) dimulai dari tahun 2019. Mulai tahun 2022, sebagian besar data objek pajak sudah mendominasi di *cluster* 1 hingga mencapai puncaknya pada tahun 2024 dan 2025 dengan jumlah sebanyak 278 dan 464 objek pajak. Pada rentang waktu 2021 hingga 2025 *cluster* 2 dan *cluster* 3 masih tetap muncul, namun kuantitasnya jauh lebih kecil dibandingkan dengan *cluster* 1. Kondisi tersebut menggambarkan sebelum tahun 2021, data objek pajak homogen ke dalam *cluster* dengan jumlah yang sedikit. Setelahnya data objek pajak homogen ke *cluster* yang lebih besar.

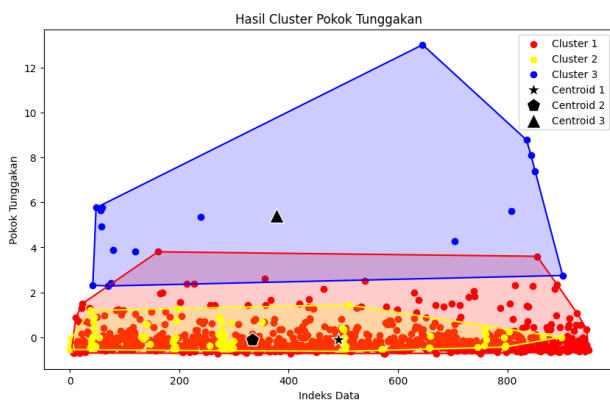


Gambar 3. Visualisasi Hasil *Clustering* Variabel X_3

Berdasarkan diagram hasil visualisasi *clustering* di atas, distribusi masa pajak tiap *cluster* menunjukkan adanya pola musiman yang cukup kuat. Hasil *clustering* didominasi oleh data objek pajak yang termasuk dalam *cluster* 1 dengan puncak tertinggi pada bulan Februari dan Maret dengan jumlah data sebanyak 127 dan 201 objek pajak. Selain itu pada bulan Januari dan April, *cluster* 1 juga memiliki jumlah data yang cukup tinggi sebanyak 97 dan 113 data. Dari hasil *clustering* tersebut, triwulan 1 memiliki intensitas data yang lebih tinggi dibandingkan dengan bulan-bulan lainnya. Pada periode triwulan 2 dan triwulan 3, data yang termasuk pada *cluster* 1 masih mendominasi di tiap bulannya. Namun jumlah data yang masuk ke dalam *cluster* 2 dan *cluster* 3 memiliki perbandingan yang lebih banyak dengan data *cluster* 1 per bulannya.

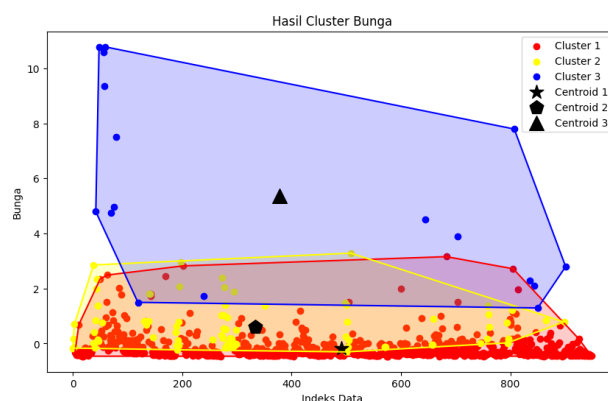
Gambar 4. Visualisasi Hasil Clustering Variabel X_4

Hasil segmentasi menggunakan K-Prototypes Clustering menunjukkan tiga karakteristik utama pada objek pajak reklame. Cluster 1 ditandai dengan keterlambatan rendah, dengan rata-rata 7,425 bulan, median 4 bulan, rentang 1–39 bulan, serta standar deviasi kecil (7,893) yang menggambarkan sebaran data homogen dan pola pembayaran yang konsisten, terlihat pada scatter-plot melalui persebaran data berwarna merah yang berada di bagian bawah. Cluster 2 memiliki karakteristik keterlambatan sangat tinggi dengan rata-rata 69,046 bulan, median 62 bulan, rentang 20–138 bulan, serta standar deviasi besar (32,299) yang menunjukkan variasi ekstrem dan adanya objek pajak dengan tunggakan sangat lama, tercermin dari sebaran data berwarna kuning yang luas dan jauh dari centroid pada visualisasi. Sementara itu, Cluster 3 menggambarkan keterlambatan sedang dengan rata-rata 22,706 bulan, median 18 bulan, rentang 3–68 bulan, dan standar deviasi 19,035 yang menunjukkan variasi tidak terlalu besar, sesuai dengan scatter-plot berwarna biru yang memiliki persebaran lebih kecil dibandingkan cluster 2 dan mencirikan pola pembayaran yang relatif konsisten.

Gambar 5. Visualisasi Hasil Clustering Variabel X_5

Hasil segmentasi pada variabel pokok tunggakan membentuk tiga cluster dengan karakteristik yang berbeda. **Cluster 1** (wilayah merah) merepresentasikan objek pajak reklame dengan pokok tunggakan relatif rendah, ditunjukkan oleh

rata-rata Rp3.795.243,89 dan median Rp2.661.110,00, dengan jangkauan luas hingga Rp27.486.866 akibat adanya nilai sangat kecil (Rp4.138,00) serta beberapa pencilan bernilai lebih dari Rp20 juta, sehingga cluster ini mencakup objek pajak dengan tunggakan rendah hingga menengah. **Cluster 2** memiliki karakteristik serupa namun dengan sebaran yang lebih homogen, dengan rata-rata Rp3.798.345,966, median Rp2.503.913,00, dan standar deviasi lebih kecil (Rp2.886.724,063), sehingga cluster ini menggambarkan objek pajak dengan pokok tunggakan menengah yang lebih seragam dibandingkan cluster 1. Sementara itu, **Cluster 3** (wilayah biru) menunjukkan segmentasi objek pajak dengan pokok tunggakan sangat tinggi, dengan rata-rata Rp37.322.025,235, median Rp36.956.250,00, jangkauan mencapai Rp65.264.008,00, serta standar deviasi besar (Rp16.828.427,106) yang mengindikasikan distribusi yang timpang dan pola sebaran data yang tidak stabil.

Gambar 6. Visualisasi Hasil Clustering Variabel X_6

Hasil clustering pada variabel bunga tunggakan membentuk tiga segmentasi utama. **Cluster 1** (wilayah merah) terdiri dari objek pajak reklame dengan bunga tunggakan rendah, ditunjukkan oleh rata-rata Rp490.387,81, median Rp219.000,00, serta distribusi yang rapat di sekitar nilai kecil dengan centroid dekat angka 0. **Cluster 2** menggambarkan objek pajak dengan bunga tunggakan menengah, memiliki rata-rata Rp1.811.370,38, median Rp1.201.878,24, dan variabilitas lebih tinggi (SD Rp1.370.101,76), sehingga cluster ini berperan sebagai segmen transisi antara bunga tunggakan sedang dan tinggi dengan letak centroid di antara dua cluster lainnya. Sementara itu, **Cluster 3** berisi objek pajak dengan bunga tunggakan sangat tinggi, dengan rata-

rata Rp9.825.461,85, median Rp8.700.000,00, jangkauan Rp15.965.100,00, dan standar deviasi besar (Rp5.764.579,15), yang tercermin dari persebaran titik data yang jauh di nilai tinggi dan tidak stabil pada scatter-plot.

PENUTUP

SIMPULAN

Sebanyak 949 objek pajak reklame berhasil dikelompokkan menjadi tiga segmentasi menggunakan algoritma K-Prototypes dengan kualitas pengelompokan yang cukup baik berdasarkan nilai ASW. **Cluster 1** merupakan kelompok terbesar dan didominasi oleh UPTD Enggal, Kedaton, dan Kedamaian, dengan masa pajak yang umumnya berada pada empat tahun terakhir serta banyak muncul pada awal tahun. Objek pajak dalam cluster ini memiliki keterlambatan pembayaran yang rendah, pokok tunggakan rendah hingga menengah, serta bunga tunggakan yang relatif kecil. **Cluster 2** didominasi objek pajak dari Kedaton, Enggal, dan Kedamaian, dengan masa pajak yang lebih lama dan sebaran bulan yang merata. Kelompok ini dicirikan oleh keterlambatan pembayaran yang tinggi, pokok tunggakan pada kategori menengah dengan sebaran data lebih seragam, serta bunga tunggakan yang berada pada tingkat sedang. Sementara itu, **Cluster 3** yang beranggotakan paling sedikit didominasi UPTD Enggal dan Tanjung Karang Pusat, memiliki masa pajak dalam rentang beberapa tahun terakhir dengan pola bulanan yang tidak khas. Objek pajak dalam cluster ini memiliki keterlambatan menengah, namun pokok dan bunga tunggakan tergolong sangat tinggi, sehingga menjadi segmen dengan tingkat risiko tunggakan terbesar.

SARAN

Saran untuk penelitian selanjutnya dapat membandingkan hasil segmentasi menggunakan metode evaluasi lainnya seperti metode *Dunn Index* atau metode lainnya.

DAFTAR PUSTAKA.

- Annas, S., Irwan, I., Safei, R. H., & Rais, Z. (2022). K-Prototypes Algorithm for Clustering The Tectonic Earthquake in Sulawesi Island. *Jurnal Varian*, 5(2), 191–198. <https://doi.org/10.30812/varian.v5i2.1908>
- Davies Raihannabil, S. (2024). Penerapan Metode Hierarchical Clustering untuk Klasterisasi Provinsi di Indonesia berdasarkan Indikator Status Gizi Anak Baduta (Bawah Dua Tahun) Tahun 2023. *Emerging Statistics and Data Science Journal*, 2(3), 424–436.
- Pradhana, O. J., Saepudin, D., & Indwiarti. (2022). Penerapan Clustering Pada Tipe Data Campuran Menggunakan K-Prototype Pada Perusahaan Multifinance. *e-Proceeding of Engineering*, 9(Juni), 1196–2007.
- Primadeni, F., Ichsan Pradana, A., & Purwanto, E. (2024). Analisis Clustering Untuk Segmentasi Wilayah Berdasarkan Karakteristik PBB di Kabupaten Sragen. *Jurnal Pendidikan Dan Teknologi Indonesia*, 4(11), 405–417. <https://doi.org/10.52436/1.jpti.500>
- Purba, A. F., Mustafid, M., & Kartikasari, P. (2025). Penerapan Algoritma K-Prototype untuk Pengelompokan Desa di Kabupaten Bekasi Berdasarkan Infrastruktur Digital. *Jurnal Gaussian*, 13(2), 479–489. <https://doi.org/10.14710/j.gauss.13.2.479-489>
- Rais, Z., Annas, S., & Refaldy, M. (2024). Algoritma K-Prototype dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Indikator Kesejahteraan Rakyat Tahun 2020. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 6(3), 144–151. <https://doi.org/10.35580/variansiunm20>
- Undang-Undang Republik Indonesia Nomor 1 Tahun 2022 Tentang Hubungan Keuangan Antara Pemerintah Pusat dan Pemerintahan Daerah, 4 (2022).
- Widodo, W., Widodo, N. M., & Prihadyatama, A. (2024). Rasio Efektivitas Dan Kontribusi Pajak Daerah Terhadap Pendapatan Asli Daerah (Studi Kasus Kota Madiun Tahun 2018-2022). *Jurnal Maneksi*, 13, 9.