



## Visualisasi Data Penerimaan PKB dan Identifikasi Dominasi Kategori Kendaraan menggunakan Pendekatan *Clustering*

Winda Yunia Purnama<sup>1\*</sup>, Dedek Nurlina<sup>2</sup>, Armansyah<sup>3</sup>

<sup>1-3</sup>Universitas Islam Negeri Sumatera Utara, Indonesia

Email: [winda0701233169@uinsu.ac.id](mailto:winda0701233169@uinsu.ac.id)<sup>1</sup>, [dedek0701233170@uinsu.ac.id](mailto:dedek0701233170@uinsu.ac.id)<sup>2</sup>, [armansyah@uinsu.ac.id](mailto:armansyah@uinsu.ac.id)<sup>3</sup>

\*Penulis Korespondensi: [winda0701233169@uinsu.ac.id](mailto:winda0701233169@uinsu.ac.id)

**Abstract.** Motor Vehicle Tax (PKB) as the primary source of Regional Own Revenue (PAD) in North Sumatra has not been utilized optimally due to the lack of systematic analysis to identify contribution patterns across vehicle categories. This study aims to cluster vehicle categories based on PKB contributions using the K-Means Clustering algorithm, in order to provide clear visual insights for decision-making. The data used were derived from PKB realization across five vehicle categories (Passenger Cars, Buses, Cargo Vehicles, Motorcycles, and Special Vehicles) at 34 SAMSAT Technical Implementation Units (UPT) in North Sumatra for the period January–August 2025. The methodology includes data preprocessing, category aggregation, Min-Max Scaling normalization, and optimal cluster determination using the Elbow method. The results indicate an optimal cluster of  $K=3$  (Silhouette Coefficient=0.62; Davies–Bouldin Index=0.78), with Passenger Cars as the dominant category (71.8%) and Motorcycles at an intermediate level (18.2%). Data visualization through monthly trend charts, proportion heatmaps, and bar charts of the top 10 UPTs reveals seasonal fluctuation patterns in PKB revenue. These findings can serve as a basis for priority collection strategies and data-driven transportation infrastructure planning for regional institutions and SAMSAT North Sumatra

**Keywords:** Data Visualization; K-Means Clustering; Motor Vehicle Tax; Regional Own Revenue; Vehicle Contribution Analysis.

**Abstrak.** Pajak Kendaraan Bermotor sebagai sumber utama Pendapatan Asli Daerah di Provinsi ini belum dimanfaatkan secara optimal karena minimnya analisis sistematis untuk mengidentifikasi pola kontribusi kategori kendaraan. Penelitian ini bertujuan mengelompokkan kategori kendaraan berdasarkan kontribusi PKB menggunakan algoritma K-Means Clustering, guna memberikan wawasan visual yang jelas bagi pengambilan keputusan. Data yang digunakan berasal dari realisasi PKB lima kategori kendaraan (Mobil Penumpang, Bus, Barang, Sepeda Motor, Kendaraan Khusus) di 34 UPT SAMSAT Provinsi ini dari periode Januari–Agustus 2025. Metode meliputi pra-pemrosesan data, agregasi kategori, normalisasi Min-Max Scaling, dan penentuan cluster optimal menggunakan metode Elbow. Hasil menunjukkan cluster optimal  $K=3$  (Silhouette Coefficient=0,62; Davies–Bouldin Index=0,78), dengan Mobil Penumpang sebagai dominan (71,8%) dan Sepeda Motor menengah (18,2%). Visualisasi data melalui diagram tren bulanan, heatmap proporsi, dan diagram batang 10 UPT teratas mengungkap pola fluktuasi musiman PKB. Temuan ini dapat menjadi dasar strategi penagihan prioritas dan perencanaan infrastruktur transportasi berbasis data bagi Lembaga dan SAMSAT Provinsi.

**Kata Kunci:** Analisis Kontribusi Kendaraan; K-Means Clustering; Pajak Kendaraan Bermotor; Pendapatan Asli Daerah; Visualisasi Data.

### 1. PENDAHULUAN

Pajak Kendaraan Bermotor (PKB) menjadi salah satu sumber penting Pendapatan Asli Daerah (PAD) yang berperan dalam pembiayaan pembangunan dan fasilitas publik. Seiring meningkatnya jumlah kendaraan bermotor di Provinsi Sumatera Utara, data penerimaan PKB juga semakin besar dan beragam. Berdasarkan data statistik tahun 2025, jumlah kendaraan bermotor di Sumatera Utara terdiri atas 872.015 unit mobil penumpang, 10.541 unit bus, 333.530 unit truk, dan 7.176.912 unit sepeda motor. Besarnya jumlah kendaraan tersebut menunjukkan potensi penerimaan PKB yang tinggi dengan kontribusi berbeda pada setiap kategori kendaraan dan wilayah UPT SAMSAT.

Selain itu, perubahan nilai penerimaan pada setiap periode menunjukkan adanya pola tertentu yang perlu dipahami lebih mendalam. Namun, data PKB masih banyak disajikan dalam bentuk tabel dan laporan angka sehingga pola kontribusi dan dominasi kategori kendaraan belum terlihat secara jelas. Oleh karena itu, diperlukan pendekatan visualisasi data agar pola penerimaan PKB dapat dipahami dengan lebih terstruktur dan mendukung penyusunan kebijakan perpajakan daerah yang lebih tepat (Hanifah *et al.*, 2023).

Sejumlah studi terdahulu membuktikan bahwa *clustering* berbasis K-Means efektif dalam menyajikan pola perpajakan daerah secara terstruktur dan mudah diinterpretasikan. Nursaniah, Rahaningsih, Ali, dan Dienwati (2024) berhasil mengelompokkan penerimaan PBB di 69 kelurahan Kota Tasikmalaya ke dalam enam klaster sehingga distribusi kontribusi antarwilayah lebih mudah diidentifikasi, sementara Nur, Bahtiar, Jannah, dan No (2025) memetakan kepatuhan 376 wajib pajak PBB di Kecamatan Sakra ke dalam tiga kelompok menggunakan K-Means dengan optimasi *elbow* sebagai dasar strategi intervensi yang lebih tepat sasaran. Pada konteks kendaraan bermotor, Wiradiansya, Elfianty, dan Fredricka (2024) mengklasifikasikan tunggakan PKB per wilayah di Kabupaten Bengkulu Selatan ke dalam klaster tinggi dan rendah sebagai bahan evaluasi Kantor SAMSAT, sedangkan Rakhmawati, Aletha, Widiastuti, Sidum, dan Kaki (2020) membuktikan bahwa klasterisasi K-Means terhadap 1.849 peraturan daerah di Pulau Jawa mampu mengungkap konsentrasi kebijakan antarwilayah secara lebih jelas. Temuan-temuan ini menegaskan efektivitas pendekatan *clustering* dalam menyajikan pola perpajakan daerah secara sistematis. Berdasarkan hal tersebut, penelitian ini bertujuan mengklasterisasi dan memvisualisasikan data penerimaan Pajak Kendaraan Bermotor (PKB) guna mengidentifikasi dominasi kontribusi tiap kategori kendaraan.

Tujuan klasterisasi adalah mengelompokkan kategori kendaraan berdasarkan pola penerimaannya, sedangkan tujuan visualisasi adalah menyajikan hasil pengelompokan tersebut secara grafis agar mudah dipahami oleh pemangku kebijakan. Untuk mencapai kedua tujuan tersebut, penelitian ini menggunakan metode K-Means sebagai pendekatan pengelompokan dan visualisasi grafis sebagai media penyajian hasilnya. Pendekatan *clustering* dipilih karena mampu mengelompokkan data berdasarkan kemiripan karakteristik tanpa label sebelumnya sehingga pola kontribusi tiap kategori kendaraan dapat terlihat secara alami. Jumlah klaster ditentukan melalui *elbow method*, yang telah terbukti efektif pada berbagai studi sebelumnya (Tammam *et al.*, 2025). Untuk memastikan kualitas pengelompokan yang dihasilkan, penelitian ini juga menggunakan metrik evaluasi tambahan berupa *Silhouette Score* dan *Davies-Bouldin Index* (DBI).

Penggunaan dua metrik ini secara bersamaan telah terbukti memberikan panduan yang efektif dalam memilih algoritma serta menentukan jumlah kluster optimal tanpa memerlukan label data awal (Hasan, 2024). Hasil penelitian ini diharapkan memberikan dasar yang lebih konkret bagi pemerintah daerah dalam menyusun strategi peningkatan Pendapatan Asli Daerah (PAD) secara tepat sasaran (Nursaniah *et al.*, 2024).

Namun demikian, studi-studi tersebut sebagian besar masih berfokus pada pengelompokan wilayah atau wajib pajak dan belum menyentuh visualisasi dominasi kontribusi per kategori kendaraan secara spesifik di Sumatera Utara. Data per kategori kendaraan merupakan cerminan langsung dari struktur penerimaan PKB yang belum banyak divisualisasikan secara sistematis. Oleh karena itu, diperlukan suatu pendekatan yang mampu memvisualisasikan data penerimaan PKB secara sistematis berdasarkan kategori kendaraan di 34 UPT SAMSAT Sumatera Utara periode Januari–Agustus 2025 dengan memanfaatkan *K-Means Clustering* sebagai metode pengelompokannya. Visualisasi yang dihasilkan diharapkan dapat menjadi acuan yang lebih konkret bagi Badan Pendapatan Daerah dan SAMSAT Sumatera Utara dalam merancang strategi penagihan prioritas, pemberian insentif, serta perencanaan infrastruktur yang lebih terarah. Makalah ini disusun sebagai berikut: Bagian 2 membahas tinjauan literatur, Bagian 3 menjelaskan metode penelitian, Bagian 4 menyajikan hasil dan pembahasan, dan Bagian 5 berisi kesimpulan (Asti *et al.*, 2023).

## 2. TINJAUAN TEORITIS

### **Pajak Kendaraan Bermotor (PKB)**

Pajak Kendaraan Bermotor (PKB) menjadi salah satu sumber Pendapatan Asli Daerah (PAD) yang penting, berasal dari sektor transportasi, dan berkontribusi signifikan terhadap penerimaan keuangan daerah (Asti *et al.*, 2023). PKB terbukti berpengaruh signifikan terhadap penerimaan PAD, dan besarnya kontribusi bergantung pada jumlah kepemilikan kendaraan di suatu wilayah (Hanifah *et al.*, 2023). Objek PKB mencakup berbagai jenis kendaraan bermotor, mulai dari sepeda motor, mobil penumpang, hingga kendaraan pengangkut barang. Oleh karena itu, memahami kategori kendaraan penting untuk mengetahui pola kontribusi pajak dan mengetahui jenis kendaraan yang paling banyak berkontribusi terhadap pendapatan PKB (Rahmawati & Effendi, 2025). Selain itu, analisis data kendaraan bermotor juga dapat membantu pemerintah daerah dalam mengetahui kelompok wajib pajak yang paling berpotensi memberikan kontribusi besar terhadap Pendapatan Asli Daerah (PAD) (Budiarti *et al.*, 2023).

## Visualisasi Data

Dalam menganalisis data perpajakan, visualisasi data sangat penting karena mampu menampilkan pola dan tren penerimaan pajak daerah secara lebih cepat, jelas, dan terstruktur. Visualisasi hasil pengelompokan data dapat membantu pihak perpajakan dalam menentukan prioritas penagihan, memahami pola tunggakan pajak, serta mengenali aktivitas pembayaran wajib pajak (Kaharu & Panyili). Penelitian lain juga menunjukkan bahwa visualisasi klasifikasi wajib pajak dapat membantu dalam penyusunan strategi *Customer Relationship Management* (CRM) serta memudahkan identifikasi dominasi kategori kendaraan secara lebih objektif (Anwar *et al.*, 2025). Lebih lanjut, efektivitas visualisasi semakin meningkat apabila didahului oleh proses klasifikasi berbasis *data mining*. Penerapan teknik klasifikasi berbasis *data mining* pada data wajib pajak kendaraan bermotor terbukti mampu mengungkap potensi kepatuhan wajib pajak yang sulit teridentifikasi hanya dari laporan tabular (Compliance, 2024). Dengan demikian, visualisasi data bukan hanya berperan sebagai media penyampaian informasi, tetapi juga menjadi sarana pendukung pengambilan keputusan dalam pengelolaan penerimaan pajak daerah (Kholijah & Syafmen, 2025).

### *K-Means Clustering*

Pendekatan yang sering digunakan dalam menganalisis data adalah *clustering*, yaitu metode pembelajaran tidak terawasi (*unsupervised learning*) yang mengelompokkan data berdasarkan tingkat kemiripan karakteristiknya tanpa memerlukan label data sebelumnya (Wiradiansya *et al.*, 2024). Metode *clustering* dinilai efektif dalam mengelompokkan wajib pajak, menentukan prioritas penagihan, serta mengidentifikasi kendaraan yang dominan pada data perpajakan daerah (Tammam *et al.*, 2025). Selain itu, metode *clustering* juga mampu mengungkap pola data yang sulit terlihat pada laporan konvensional dan dapat diterapkan secara fleksibel pada data transaksi yang terus berkembang (Budiarti *et al.*, 2023). Salah satu metode pengelompokan yang paling banyak digunakan adalah *K-Means Clustering*. Metode ini bekerja secara iteratif dengan meminimalkan jarak antara data dan pusat kluster (*centroid*) sehingga terbentuk kelompok data yang memiliki karakteristik serupa (Asti *et al.*, 2023). Berbagai penelitian menunjukkan bahwa metode *K-Means* sangat efektif untuk mengelompokkan penerimaan PKB, mengklasifikasikan tunggakan pajak, menentukan jumlah kluster yang optimal, serta mengidentifikasi kategori kendaraan yang paling dominan (Kendaraan & Barat, 2025). Selain mudah diterapkan, metode ini juga efektif dalam mengolah data numerik, seperti besaran pajak yang dibayarkan dan tahun pembuatan kendaraan (Kaharu & Panyili).

Oleh karena itu, metode *K-Means Clustering* dinilai sesuai untuk membantu menganalisis dan memvisualisasikan data penerimaan PKB secara lebih sistematis sehingga menghasilkan informasi yang mudah dipahami (Rahmawati & Effendi, 2025). Hal tersebut sejalan dengan penelitian Apuilino, Seno, Achmadi, dan Industri (2021) yang membuktikan bahwa penerapan *K-Means* pada data retribusi PAD menghasilkan tingkat akurasi sebesar 99,68% dan mampu mengidentifikasi restoran serta kafe sebagai sektor dengan kontribusi terbesar di Kabupaten Gresik (Apuilino *et al.*, 2021).

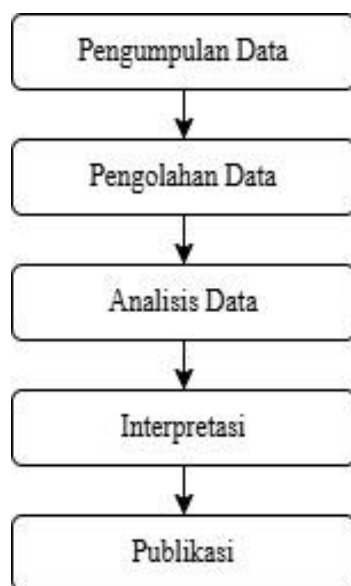
### 3. METODE

Penelitian ini menggunakan pendekatan kuantitatif deskriptif eksploratori untuk mengidentifikasi pola kontribusi penerimaan PKB berdasarkan kategori kendaraan tanpa hipotesis awal. Pendekatan ini dipilih karena seluruh data yang dianalisis bersifat numerik dan bertujuan mengeksplorasi pola pengelompokan alami berdasarkan karakteristik penerimaan pajak antar UPT. Data yang digunakan adalah data realisasi penerimaan PKB periode Januari–Agustus 2025 yang diperoleh dari Badan Pendapatan Daerah (Bapenda) Provinsi Sumatera Utara melalui izin resmi dari instansi terkait. Data mencakup 34 UPT dan 11 jenis kendaraan yang kemudian diagregasi menjadi 5 kategori utama yaitu Mobil Penumpang, Mobil Bus, Mobil Barang, Sepeda Motor, dan Kendaraan Khusus.

Analisis dilakukan melalui serangkaian tahapan yang saling berkesinambungan mengacu pada pendekatan knowledge Discovery in Database (KDD). Pendekatan KDD mencakup serangkaian tahapan sistematis mulai dari seleksi data, preprocessing, transformasi, penerapan algoritma, hingga interpretasi hasil. Data mentah terlebih dahulu melalui proses pembersihan dengan mengisi nilai kosong dan mengkonversi format angka dari string ke numerik, dilanjutkan agregasi 11 jenis kendaraan menjadi 5 kategori utama. Data kemudian dinormalisasi menggunakan Min-Max Scaling agar seluruh kategori memiliki bobot yang setara dalam proses pengelompokan, mengingat skala penerimaan antar kategori sangat bervariasi dari jutaan hingga ratusan miliar rupiah.

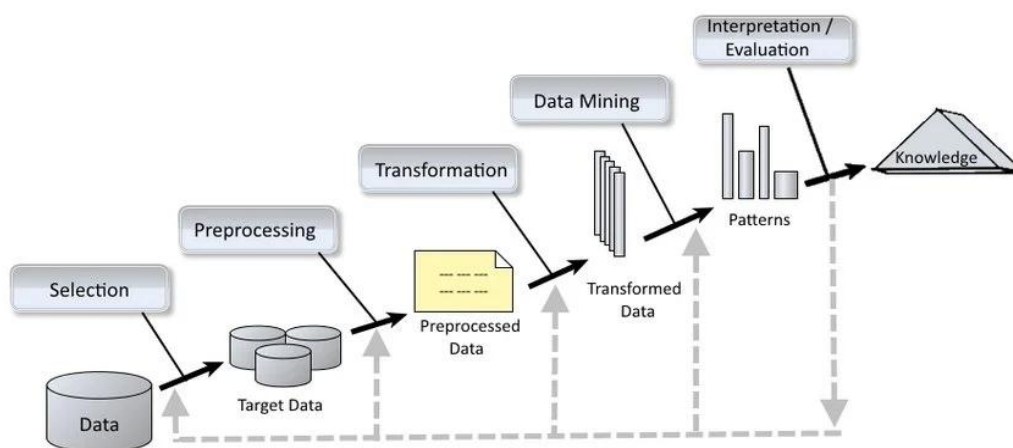
Selanjutnya, pengelompokan 34 UPT dilakukan menggunakan algoritma K-Means Clustering dengan penentuan jumlah cluster optimal melalui metode Elbow. Hasil pengelompokan dievaluasi menggunakan Silhouette Coefficient dan Davies-Bouldin Index untuk mengukur kualitas cluster yang terbentuk, serta divisualisasikan dalam bentuk grafik tren bulanan, heatmap proporsi, diagram batang, dan scatter plot PCA untuk memudahkan interpretasi pola penerimaan antar UPT.

Kerangka alur penelitian mulai dari pengumpulan data hingga penarikan kesimpulan disajikan pada Gambar 1. sedangkan alur metode analisis data menggunakan pendekatan KDD mulai dari selection data hingga knowledge disajikan pada Gambar 2.



**Gambar 1.** Kerangka Penelitian.

Data yang digunakan adalah data sekunder yang diambil dengan observasi langsung ke lapangan dan pengambilan data dari Badan Pendapatan Daerah (Bapenda) Provinsi Sumatera Utara. Selanjutnya, data diolah menggunakan bahasa pemrograman Python melalui Google Colab. Tahap analisis data dilakukan menggunakan pendekatan knowledge Discovery in Databases (KDD) mulai dari seleksi data hingga evaluasi hasil clustering. Hasil penelitian kemudian diinterpretasikan untuk memperoleh informasi yang sesuai dengan tujuan penelitian dan selanjutnya didokumentasikan dalam bentuk artikel.



**Gambar 2.** Diagram Metode Penelitian (KDD).

#### 4. HASIL DAN PEMBAHASAN

##### Deskripsi Data dan Pra-Pemrosesan

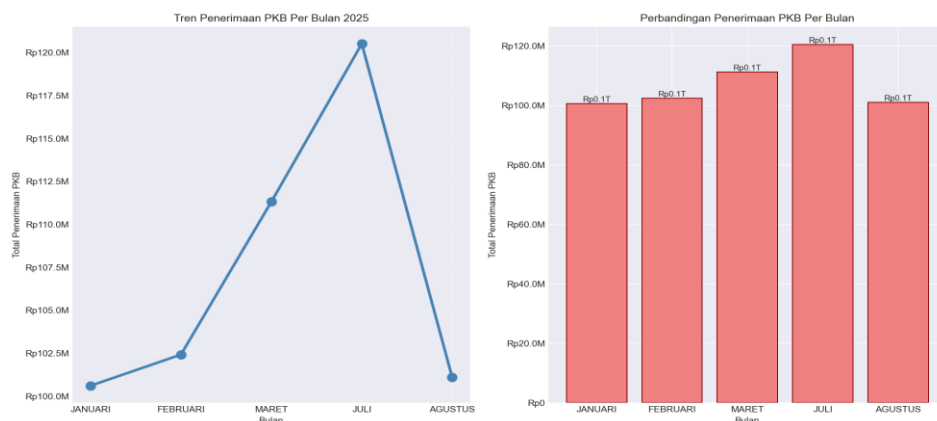
Data realisasi penerimaan PKB periode Januari–Agustus 2025 diperoleh dari sistem e-keuangan Badan Pendapatan Daerah (Bapenda) Provinsi Sumatera Utara, mencakup 34 UPT SAMSAT dan 11 jenis kendaraan yang setelah melalui proses pembersihan, agregasi, dan normalisasi Min-Max Scaling menjadi 5 kategori utama. Statistik deskriptif hasil pra-pemrosesan disajikan pada Tabel 1

**Tabel 1.** Statistik Deskriptif Penerimaan PKB per Kategori (dalam Miliar Rupiah).

| Kategori         | Total  | Rata-rata per UPT | Minimum      | Maksimum                | Standar Deviasi |
|------------------|--------|-------------------|--------------|-------------------------|-----------------|
| Mobil Penumpang  | 571,74 | 16,82             | 0,44 (Salak) | 207,36<br>(Medan Utara) | 48,23           |
| Sepeda Motor     | 143,06 | 4,21              | 0,20 (Salak) | 50,61<br>(Medan Utara)  | 13,15           |
| Mobil Barang     | 83,56  | 2,46              | 0,10 (Salak) | 25,96<br>(Medan Utara)  | 5,98            |
| Mobil Bus        | 4,50   | 0,13              | 0,00         | 1,97<br>(Medan Utara)   | 0,41            |
| Kendaraan Khusus | 0,031  | 0,0009            | 0,00         | 0,023<br>(Binjai)       | 0,005           |

Tabel 1. menunjukkan bahwa Mobil Penumpang mencatat total penerimaan tertinggi sebesar Rp571,74 miliar dengan rata-rata per UPT Rp16,82 miliar, diikuti Sepeda Motor (Rp143,06 miliar), Mobil Barang (Rp83,56 miliar), Mobil Bus (Rp4,50 miliar), dan Kendaraan Khusus (Rp0,031 miliar). UPT Salak secara konsisten mencatat nilai minimum di hampir semua kategori, sementara Medan Utara mendominasi nilai maksimum di empat dari lima kategori. Standar deviasi Mobil Penumpang yang tinggi (48,23) mengindikasikan persebaran penerimaan yang sangat tidak merata antar UPT, sehingga kesenjangan skala yang besar antar kategori ini menjadi alasan utama diterapkannya normalisasi Min-Max Scaling sebelum proses clustering agar kategori dengan nilai besar tidak mendominasi pengelompokan secara tidak proporsional.

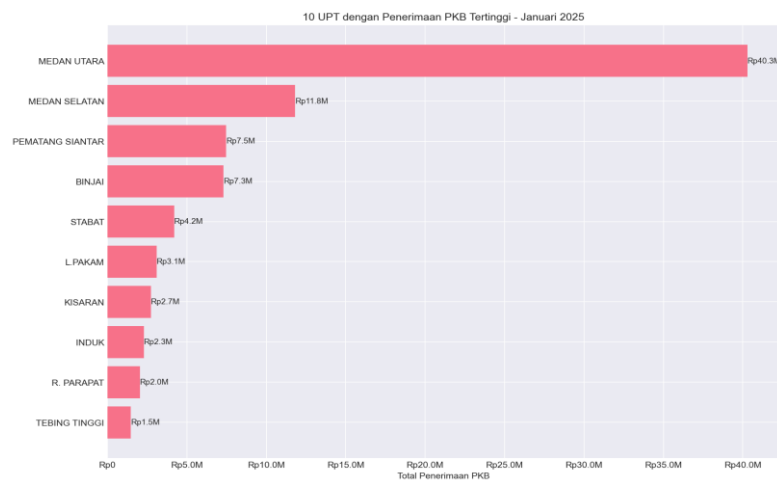
##### Analisis Tren Penerimaan PKB



**Gambar 3.** Analisis Tren Penerimaan Bulanan Total (Januari–Agustus 2025).

Gambar 3 menunjukkan fluktuasi penerimaan PKB Januari–Agustus 2025 dengan puncak pada Juli (Rp120,52 miliar) dan titik terendah pada April (Rp99,57 miliar), mencerminkan selisih sebesar Rp20,95 miliar. Penurunan April dipicu pergeseran pembayaran akibat libur Lebaran, sedangkan lonjakan Juli didorong berakhirnya masa pemutihan dan meningkatnya kesadaran pajak menjelang akhir semester. Pola musiman yang sistemik dan dapat diprediksi ini, sebagaimana dikonfirmasi Kholijah & Syafmen (2025), memberi peluang bagi pemerintah daerah untuk mengintensifkan layanan pada bulan puncak sekaligus merancang intervensi khusus pada periode penurunan.

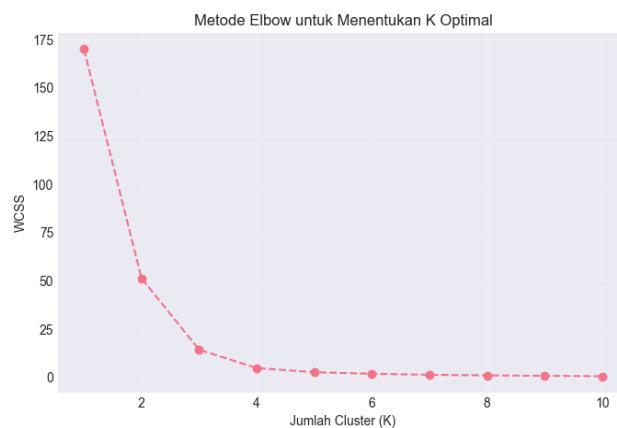
### Konsentrasi Penerimaan per UPT



**Gambar 4.** 10 UPT dengan Penerimaan PKB Tertinggi (Bulan Januari 2025).

Gambar 4. menunjukkan bahwa Medan Utara menempati posisi pertama dengan penerimaan yang jauh melampaui UPT lainnya, diikuti Binjai di posisi kedua. Pola konsentrasi ini juga konsisten pada total periode Januari–Agustus (tidak disajikan), di mana Medan Utara dan Medan Selatan menjadi dua tertinggi.

### Penentuan Cluster Optimal dan Hasil Clustering



**Gambar 5.** Metode Elbow untuk Menentukan Jumlah Cluster Optimal.

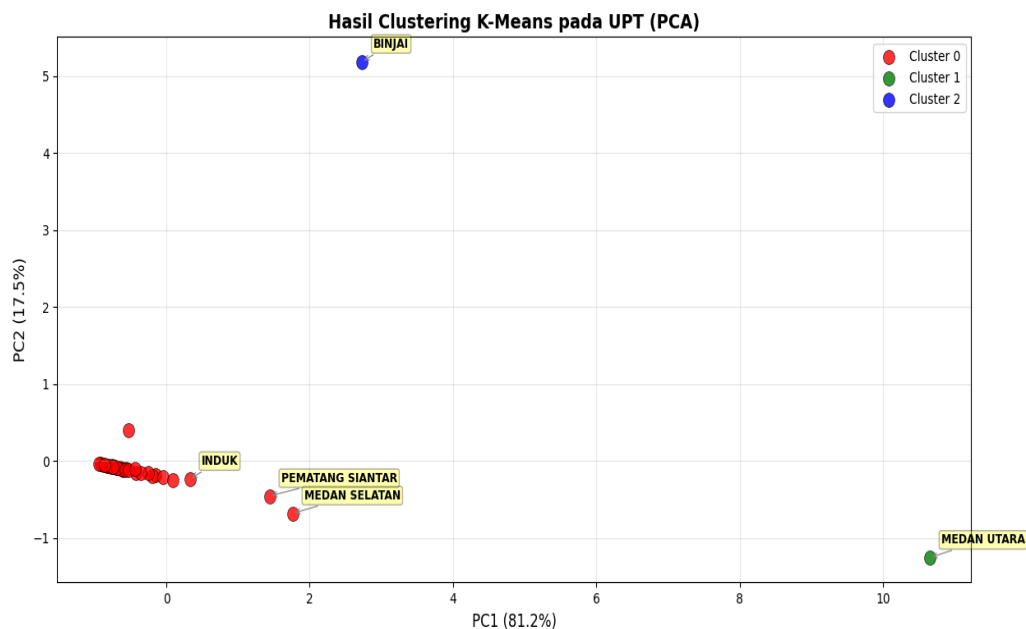
Gambar 5. menunjukkan bahwa penurunan nilai WCSS (Within Cluster Sum of Squares) mulai melandai secara signifikan pada  $K=3$ , yang menandakan titik siku (elbow) berada di posisi tersebut. Nilai Silhouette Coefficient = 0,62 mengkonfirmasi kualitas cluster yang baik karena nilai di atas 0,5 mengindikasikan pemisahan antar cluster yang cukup jelas, sedangkan Davies–Bouldin Index = 0,78 yang relatif rendah semakin memperkuat validitas pengelompokan. Dengan demikian, jumlah cluster optimal ditetapkan  $K=3$ .

Hasil pengelompokan 34 UPT menggunakan K-Means  $K=3$  disajikan pada Tabel 2.

**Tabel 2.** Profil Rata-rata Penerimaan per Cluster (dalam Miliar Rupiah).

| Cluster   | Jumlah UPT      | Mobil Penumpang | Mobil Bus | Mobil Barang | Sepeda Motor | Kendaraan Khusus |
|-----------|-----------------|-----------------|-----------|--------------|--------------|------------------|
| Cluster 0 | 32              | 10,05           | 0,07      | 1,54         | 2,67         | 0,001            |
| Cluster 1 | 1 (Binjai)      | 45,65           | 0,41      | 8,46         | 7,08         | 0,023            |
| Cluster 2 | 1 (Medan Utara) | 207,36          | 1,97      | 25,95        | 50,61        | 0,005            |

Tabel 2. menunjukkan bahwa 32 dari 34 UPT masuk ke Cluster 0 dengan rata-rata penerimaan Mobil Penumpang Rp10,05 miliar, sementara Binjai (Cluster 1) dan Medan Utara (Cluster 2) berdiri sebagai outlier positif dengan penerimaan masing-masing Rp45,65 miliar dan Rp207,36 miliar. Pemisahan ketiga cluster divisualisasikan menggunakan reduksi dimensi PCA pada Gambar 6.



**Gambar 6.** Visualisasi Clustering UPT Menggunakan PCA.

Gambar 6. mempertegas pemisahan cluster secara visual melalui reduksi dimensi PCA, di mana Medan Utara dan Binjai tampak sebagai titik yang terisolasi jauh dari kelompok utama, sementara 32 UPT Cluster 0 membentuk kelompok yang rapat. PC1 yang menjelaskan 81,2% variansi data mengindikasikan bahwa p

Perbedaan kinerja antar UPT bersifat struktural dan bukan sekadar fluktuasi acak. Kedua UPT unggulan yang berlokasi di kawasan perkotaan padat kendaraan ini dapat dijadikan benchmark kinerja bagi UPT Cluster 0, di mana faktor keberhasilan seperti kepadatan kendaraan, efektivitas penagihan, dan aksesibilitas layanan perlu diteliti lebih lanjut.

### **Dominasi Kategori Kendaraan dan Implikasi Kebijakan**

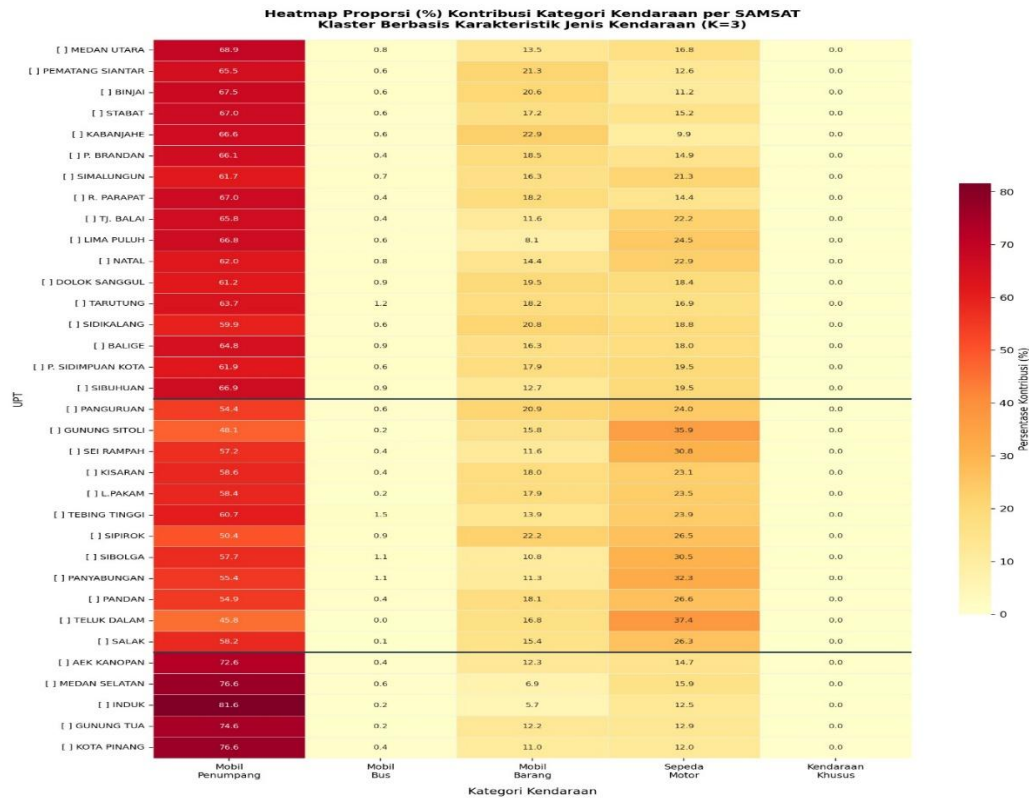
Kontribusi dan status dominasi setiap kategori kendaraan terhadap total penerimaan PKB disajikan pada Tabel 3. dengan visualisasi distribusi proporsi pada Gambar 7, dan Gambar 8

**Tabel 3.** Hasil Identifikasi Dominasi Kategori Kendaraan.

| <b>Kategori</b>  | <b>Kontribusi Rata-rata (%)</b> | <b>Status Dominasi</b> |
|------------------|---------------------------------|------------------------|
| Mobil Penumpang  | 71,2                            | DOMINAN                |
| Sepeda Motor     | 17,8                            | MENENGAH               |
| Mobil Barang     | 10,4                            | NON-DOMINAN            |
| Mobil Bus        | 0,6                             | NON-DOMINAN            |
| Kendaraan Khusus | 0,004                           | NON-DOMINAN            |
| Mobil Penumpang  | 71,2                            | DOMINAN                |

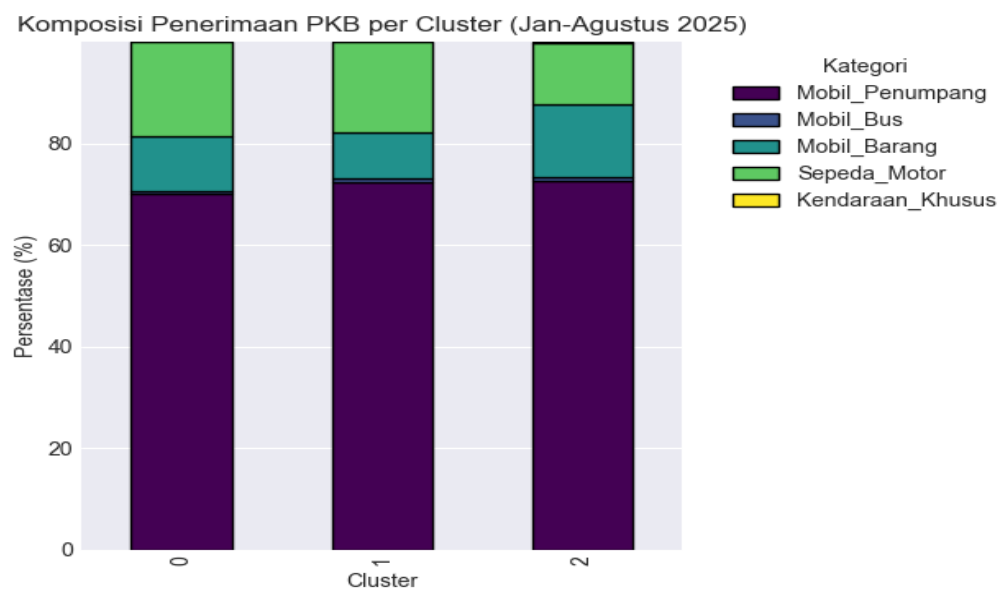
Tabel 3. menunjukkan bahwa Mobil Penumpang merupakan kategori dominan dengan kontribusi rata-rata 71,2%, diikuti Sepeda Motor (17,8%), Mobil Barang (10,4%), Mobil Bus (0,6%), dan Kendaraan Khusus (kurang dari 0,01%).

Gambar 7 memperlihatkan variasi intensitas dominasi tersebut antar UPT. Pada Klaster 0, seluruh 17 UPT mencatat proporsi Mobil Penumpang dalam rentang 59,9–68,9%, dengan Sepeda Motor berkisar 9,9–25,5% dan Mobil Barang 8,1–22,9%. Klaster 1 menunjukkan pergeseran yang cukup mencolok — proporsi Mobil Penumpang turun ke kisaran 45,8–60,7%, sementara Sepeda Motor meningkat hingga 23,9–37,4%, paling tinggi tercatat di Teluk Dalam (37,4%) dan Gunung Sitoli (35,9%). Klaster 2 justru mencatat proporsi Mobil Penumpang tertinggi di antara ketiga klaster, yakni 72,6–81,6%, dengan Sepeda Motor yang relatif rendah di kisaran 12,0–26,3%.



**Gambar 7.** Heatmap Proporsi Penerimaan per Cluster (%).

Pola ini semakin dikonfirmasi oleh Gbr. 7 yang menampilkan komposisi penerimaan PKB per klaster periode Januari–Agustus 2025. Ketiga klaster memperlihatkan struktur batang yang hampir identik secara proporsional — Mobil Penumpang mendominasi sekitar 70% di seluruh klaster, dengan kontribusi Sepeda Motor dan Mobil Barang mengisi porsi berikutnya. Konsistensi ini menegaskan bahwa dominasi Mobil Penumpang bersifat sistemik dan tidak bergantung pada skala penerimaan absolut suatu wilayah..



**Gambar 8.** Komposisi Penerimaan PKB per Cluster.

Dari perspektif kebijakan, temuan ini mengisyaratkan dua hal. Pertama, intensifikasi PKB pada segmen Mobil Penumpang akan memberikan dampak fiskal paling besar secara keseluruhan. Kedua, wilayah Klaster 1 — khususnya UPT dengan proporsi Sepeda Motor tinggi seperti Teluk Dalam dan Gunung Sitoli — menyimpan potensi penerimaan dari segmen roda dua yang masih dapat dioptimalkan melalui pendekatan yang lebih kontekstual.

## **5. KESIMPULAN**

Penelitian ini berhasil mencapai dua tujuan utamanya: memvisualisasikan pola penerimaan PKB secara sistematis dan mengidentifikasi dominasi kontribusi per kategori kendaraan di 34 UPT SAMSAT Sumatera Utara periode Januari–Agustus 2025.

Dari sisi dominasi kategori, Mobil Penumpang terbukti menjadi tulang punggung penerimaan PKB dengan kontribusi rata-rata **71,2%** (Rp571,74 miliar), konsisten di seluruh cluster. Sepeda Motor berada di posisi menengah (18,2%), sementara Mobil Barang, Mobil Bus, dan Kendaraan Khusus secara kolektif hanya menyumbang 13,3%. Pola dominasi ini berlaku merata di semua UPT, menandakan bahwa struktur penerimaan PKB Sumatera Utara sangat bergantung pada satu kategori kendaraan.

Dari sisi pengelompokan, K-Means dengan  $K=3$  (Silhouette Coefficient=0,62; Davies–Bouldin Index=0,78) berhasil memisahkan 34 UPT menjadi tiga kelompok yang berbeda secara signifikan: Medan Utara sebagai UPT unggulan tertinggi, Binjai sebagai UPT menengah, dan 32 UPT lainnya dalam kelompok penerimaan rendah-seragam. Visualisasi PCA yang menjelaskan 81,2% variansi data mempertegas bahwa kesenjangan kinerja antar UPT bersifat struktural, bukan sekadar fluktuasi acak.

Dari sisi temporal, pola musiman dengan puncak Juli dan lembah April mengindikasikan bahwa faktor waktu (periode jatuh tempo, libur nasional) turut memengaruhi perilaku wajib pajak secara sistemik.

Secara implikatif, temuan ini memberikan tiga rekomendasi kebijakan: 1) memprioritaskan intensifikasi pada segmen Mobil Penumpang melalui tarif progresif atau insentif kendaraan ramah lingkungan; 2) meningkatkan kepatuhan segmen Sepeda Motor melalui kemudahan pembayaran digital dan program pemutihan; serta 3) menjadikan Medan Utara dan Binjai sebagai benchmark kinerja bagi UPT Cluster 0. Secara metodologis, penelitian ini menunjukkan bahwa K-Means Clustering merupakan pendekatan yang efektif untuk analisis distribusi PKB berbasis data, dan dapat diadopsi oleh Bapenda provinsi lain.

Untuk penelitian selanjutnya, disarankan memperluas periode data (minimal 3 tahun) untuk menangkap pola musiman yang lebih robust, mengintegrasikan variabel eksternal seperti PDRB dan kepadatan penduduk per wilayah UPT, serta membandingkan performa K-Means dengan algoritma alternatif seperti DBSCAN atau Hierarchical Clustering guna menghasilkan segmentasi yang lebih adaptif terhadap data outlier.

## DAFTAR REFERENSI

- Anwar, M., Firmansyah, E., Sutara, B., & K. K. Clustering. (2025). Penentuan Prioritas Penagihan PBB Berbasis K-Means Clustering di Desa Gunasari. *Vol. 5*(2), 101–110.
- Apuilino, F., Seno, I., Achmadi, S., & F. T. Industri. (2021). Penerapan Metode Clustering pada Analisis Realisasi Pendapatan Asli Daerah dengan Algoritma K-Means. *Vol. 5*(2).
- Asti, D., *et al.* (2023). Penerapan Algoritma K-Means untuk Mengetahui Tingkat Kepatuhan Wajib Pajak Kendaraan Bermotor pada UPT Samsat Medan Selatan. *Vol. 2*(4), 190–198.
- Budiarti, M. G., Rahaningsih, N., Dana, R. D., & T. Informatika. (2023). Analisis Cluster Data Daftar Kendaraan Bermotor Menggunakan Algoritma K-Means. *Vol. 7*(6), 3286–3292.
- Hanifah, N., *et al.* (2023). Pengaruh Pajak Kendaraan Bermotor (PKB), Bea Balik Nama Kendaraan Bermotor (BBNKB), dan Pajak Air Permukaan (PAP) terhadap Pendapatan Asli Daerah (PAD) di UPTD Pusat Pengelolaan Pendapatan Daerah Wilayah (PPPDW) Kota Sukabumi. *Vol. 2*, 1–13.
- Hasan, Y., & K. Dbscan. (2024). Pengukuran Silhouette Score dan Davies-Bouldin Index pada Hasil Cluster. *Vol. 6*(1), 60–74.
- Hasan, Y., U. B. Darma, K. Medan, & S. Utara. (2024). Pengukuran Silhouette Score dan Davies-Bouldin Index pada Hasil Cluster K-Means dan. *Vol. 12*(3).
- Ilmiah, J., & W. Pendidikan. (2023). DOI: <https://doi.org/10.5281/zenodo.10081332> p-ISSN: e-ISSN: Accredited by Directorate General of Strengthening for Research and Development Available online at <https://jurnal.peneliti.net/index.php/JIWP>. *Vol. 9*, 547–558.
- Kaharu, N. A., & Panyili, A. S. P. Penerapan Algoritma K-Means untuk Menentukan Kelas Besar Tunggakan Pajak Kendaraan Bermotor. *Vol. XIV*(1), 70–78.
- Kasus, S., Citra, P. T., Medica, T., Mubarak, H., Irawan, B., & Suprapti, T. (2024). Clustering Data Penjualan Produk Kesehatan Menggunakan Algoritma K-Means di PT. Citra Teknik Medica. *Vol. 8*(2), 2076–2082.
- Kendaraan, P., & J. Barat. (2025). *Al-Kharaj: Jurnal Ekonomi, Keuangan & Bisnis Syariah*. *Vol. 7*, 334–344. <https://doi.org/10.47467/alkharaj.v7i1.5084>
- Kholijah, G., & Syafmen, W. (2025). Analisis Pengelompokan Bulan Berdasarkan Jenis Surat Tahun 2023 Menggunakan Metode K-Means di UPTD/SAMSAT Kabupaten Bungo. *Vol. 10*, 38–45.
- Nur, A. M., Bahtiar, H., Jannah, M. A., & No, V. (2025). Implementasi Algoritma K-Means Clustering dalam Mengelompokkan Kepatuhan Wajib Pajak Bumi dan Bangunan dengan Optimasi Elbow. *Infotek: Jurnal Informatika dan Teknologi*, 8(1).

- Nursaniah, R., Rahaningsih, N., Ali, I., & Dienwati, N. (2024). Pengelompokan Data Penerimaan Pajak Bumi dan Bangunan Berdasarkan Kelurahan di Kota Tasikmalaya Menggunakan Algoritma K-Means. *Vol. 8(2)*, 1477–1483.
- Rahmawati, E., & Effendi, P. M. (2025). Pemodelan Keputusan Konsumen untuk Pembelian Mobil Bekas Menggunakan K-Means Clustering: Consumer Decision K-Means Clustering Modeling for Used Car Purchases Using. *Vol. 6*, 22–33. <https://doi.org/10.37373/infotech.v6i1.1587>
- Rakhmawati, N. A., Aletha, D. B., Widiastuti, L. P. G., Sidum, A., & Kaki, L. (2020). Klasterisasi Peraturan Daerah di Seluruh Wilayah Jawa dengan Menggunakan Algoritma K-Means. *Vol. 9*, 242–246.
- T. Compliance, V. (2024). Penerapan Data Mining Menggunakan Metode Teknik Klasifikasi untuk Melihat Potensi Kepatuhan Wajib Pajak Kendaraan. *Vol. 3(1)*, 1–14.
- Tammam, B. F., Ibrahim, A., Indah, D. R., & Oklilas, A. F. (2025). Taxpayer Classification Using K-Means Clustering to Support CRM Strategy Development: Case Study of Prabumulih City Samsat. *Vol. 7(4)*, 3454–3482. <https://doi.org/10.63158/journalisi.v7i4.1365>
- Tania, N. E., Saputri, S. N., Yusniawaty, A. P., & Nurpriyanti, A. W. (2025). Kepuasan Mahasiswa dalam Berbelanja Online di Aplikasi *e-commerce*: Pendekatan Deskriptif Kuantitatif. *(1)*, 1–14.
- Wiradiansya, E., Elfianty, L., & Fredricka, J. (2024). Klasterisasi Data Kendaraan Bermotor Berdasarkan Tunggakan Pajak pada Kantor Samsat Kabupaten Bengkulu Selatan Menggunakan Metode K-Means Clustering. *Vol. 6*, 164–173.