

# ANALISIS SEGMENTASI PELANGGAN UNTUK MENYUSUN STRATEGI PROMOSI MENGGUNAKAN K-MEANS: STUDI KASUS DI PT XYZ

Nur Anindya Putri Haryandini<sup>1\*</sup>, Willy Prihartono<sup>2</sup>, Fathurrohman<sup>3</sup>

<sup>1,2,3</sup> Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon; Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135; Telp. (0231) 490480

Received: 14 Desember 2024  
Accepted: 14 Januari 2025  
Published: 20 Januari 2025

## Keywords:

Data Mining, K-Means,  
Segmentasi Pelanggan,  
Strategi Promosi.

## Correspondent Email:

putrianindya1902@gmail.com

**Abstrak.** Perkembangan teknologi informasi telah membawa dampak signifikan dalam strategi pemasaran, termasuk di industri otomotif. Penelitian ini bertujuan untuk menyusun strategi promosi berbasis segmentasi pelanggan dengan menggunakan algoritma data mining *K-Means* di PT XYZ. Segmentasi ini memungkinkan perusahaan untuk menargetkan kelompok pelanggan tertentu dengan strategi yang relevan dan lebih efektif. Data yang digunakan meliputi atribut usia, jenis kelamin, alamat, dan preferensi kendaraan dari faktur transaksi periode Desember 2023 hingga Mei 2024. Metode penelitian ini mengadopsi tahapan *Knowledge Discovery in Databases* (KDD) yang mencakup seleksi data, *preprocessing*, transformasi, klusterisasi dengan *K-Means*, dan evaluasi hasil. Penentuan jumlah cluster optimal dilakukan menggunakan indeks *Davies-Bouldin* (DBI). Penelitian ini menemukan bahwa segmentasi terbaik diperoleh pada jumlah cluster (*k*) sebesar 2 dengan nilai DBI sebesar 0,802. Kluster ini mengidentifikasi karakteristik utama pelanggan, seperti preferensi kendaraan, usia, dan distribusi wilayah tempat tinggal. Hasil segmentasi digunakan untuk merancang strategi promosi yang lebih terarah. Strategi berbasis data ini membantu PT XYZ dalam mengoptimalkan sumber daya pemasaran, meningkatkan loyalitas pelanggan, dan memaksimalkan potensi pasar. Penelitian ini memberikan kontribusi signifikan terhadap penerapan *K-Means* di industri otomotif dan menyediakan dasar bagi inovasi strategi pemasaran berbasis data.

**Abstract.** The development of information technology has had a significant impact on marketing strategies, including in the automotive industry. This study aims to develop a promotion strategy based on customer segmentation using the *K-Means* data mining algorithm at PT XYZ. This segmentation enables the company to target specific customer groups with more relevant and effective strategies. The data used includes attributes such as age, gender, address, and vehicle preferences from transaction invoices for the period of December 2023 to May 2024. The research method adopts the stages of *Knowledge Discovery in Databases* (KDD), which includes data selection, *preprocessing*, transformation, clustering with *K-Means*, and result evaluation. The optimal number of clusters is determined using the *Davies-Bouldin Index* (DBI). The study found that the best segmentation was achieved with 2 clusters, resulting in a DBI value of 0.802. These clusters identify key customer characteristics such as vehicle preferences, age, and the geographic distribution of residences. The segmentation results are used to design more

*targeted promotion strategies. This data-driven strategy helps PT XYZ optimize marketing resources, enhance customer loyalty, and maximize market potential. This study makes a significant contribution to the application of K-Means in the automotive industry and provides a foundation for data-driven marketing strategy innovations.*

---

## 1. PENDAHULUAN

Persaingan antara bisnis adalah salah satu karakteristik dunia usaha. Agar bisnis dapat mengambil manfaat dari persaingan, mereka harus mampu mengembangkan dan memanfaatkan sumber daya yang tersedia dalam berbagai industri dan kebutuhan bisnis [1]. Selain itu, dalam lingkungan bisnis yang lebih kompetitif, bisnis harus menjaga hubungan positif dengan pelanggan mereka. Salah satu strategi yang efektif untuk mengelola hubungan ini adalah melalui segmentasi pelanggan, yang mendorong pelanggan untuk bekerja sama berdasarkan kesamaan yang relevan [2].

PT XYZ merupakan salah satu perusahaan yang bergerak di industri kendaraan bermotor. Meskipun PT XYZ memiliki jangkauan pelanggan yang luas, perusahaan belum sepenuhnya mengoptimalkan data untuk strategi pemasaran. Tanpa segmentasi yang tepat, perusahaan mengalami kesulitan dalam menargetkan kelompok pelanggan yang sesuai. Akibatnya, kampanye pemasaran sering kali gagal mencapai sasaran. Pemahaman yang lebih baik tentang perilaku pelanggan membantu PT XYZ menganalisis segmen pasar yang paling menguntungkan. Hal ini menyebabkan sumber daya pemasaran tidak dimanfaatkan secara maksimal. Dalam situasi seperti ini, PT XYZ perlu segera meningkatkan strategi pemasarannya agar mampu bersaing di pasar. Oleh karena itu, segmentasi pelanggan menjadi langkah penting dalam memahami perilaku konsumen dan pola pembelian.

Segmentasi adalah proses membagi pasar menjadi segmen-segmen yang lebih kecil dan mengelompokkan pelanggan berdasarkan karakteristik tertentu untuk mengembangkan strategi bisnis yang sesuai [3].

Untuk melaksanakan promosi yang efektif dan efisien, analisis data dilakukan agar dapat diperoleh kumpulan informasi menggunakan metode data *mining*. Salah satu metode analisis data *mining* yang digunakan adalah *K-Means Clustering*. Dengan metode ini, data dapat

dikelompokkan ke dalam kluster berdasarkan korelasi antara data yang dimaksud [4]. *Clustering* adalah proses pengelompokan dan analisis data ke dalam beberapa set data sehingga titik data yang serupa dikelompokkan dalam satu *cluster*, sedangkan titik data yang berbeda dikelompokkan dalam *cluster* yang berbeda. [5].

Berdasarkan beberapa penelitian sebelumnya, metode *K-Means* telah terbukti berhasil dalam segmentasi pelanggan. Menurut penelitian yang dilakukan oleh [6] algoritma *K-Means* menunjukkan kinerja yang baik dalam segmentasi data di pasar otomotif. Hal ini disebabkan oleh nilai *Silhouette Coefficient* yang cukup tinggi, yaitu 0,716, yang berarti lebih baik dibandingkan dengan metode *DBSCAN* dengan nilai 0,298 (*MinPts* 5, *Eps* 0,529) dan *Hierarchical*, yang juga cukup tinggi. Dibandingkan dengan metode *DBSCAN* dan *Hierarchical* yang lebih mapan, model KNIME menunjukkan eksekusi yang cukup unik.

Peneliti lain, yaitu [7] menyebutkan segmentasi pelanggan menggunakan metode *clustering K-Means* yang bekerja dengan baik dalam mengelompokkan pelanggan. Tiga kriteria digunakan untuk menentukan pengelompokan pelanggan: usia, jenis kelamin, frekuensi transaksi, jenis pembayaran, dan kota tempat pembelian dalam satu bulan. Setelah proses segmentasi data historiografi selesai, terbentuk tiga kluster, yaitu kluster 1, 2, dan 3. Hasil segmentasi pelanggan yang telah divisualisasikan dan diinterpretasikan diharapkan dapat membantu bisnis lebih memahami kebutuhan pelanggan dan juga dapat digunakan sebagai data untuk menentukan strategi pemasaran yang tepat.

Tujuan dari penelitian ini adalah untuk mengidentifikasi segmentasi pelanggan PT XYZ menggunakan algoritma *clustering K-Means*. Hasil segmentasi ini akan digunakan untuk menentukan strategi pemasaran yang lebih relevan dan efektif berdasarkan preferensi pelanggan.

## 2. TINJAUAN PUSTAKA

### 2.1. Segmentasi Pelanggan

Memanfaatkan segmentasi pelanggan adalah salah satu metode yang digunakan untuk memahami karakteristik pelanggan. Salah satu cara yang dapat digunakan untuk memperkuat hubungan antara pelanggan dan bisnis adalah dengan menawarkan tugas yang berbeda berdasarkan karakteristik masing-masing pelanggan. Data *mining* digunakan untuk melakukan segmentasi pelanggan guna memahami karakteristik pelanggan yang tercermin dalam data. Metode untuk memahami segmentasi pelanggan suatu perusahaan adalah melalui analisis cluster. Tujuan dari segmentasi adalah menyesuaikan produk, layanan, dan pemasaran untuk setiap segmen. Proses segmentasi mengidentifikasi pelanggan berdasarkan karakteristik mereka sebagai sebuah kelompok.[4]

### 2.2. Data Mining

Data *mining* sering disebut sebagai teknik yang digunakan untuk memahami pengetahuan yang terkandung dalam basis data. Data *mining* adalah proses di mana teknik statistik, matematis, kecerdasan, dan pembelajaran mesin digunakan untuk mengidentifikasi dan mengklasifikasikan informasi yang berguna serta pengetahuan terkait[8]. Data *mining* adalah prosedur yang digunakan untuk melakukan analisis pada kumpulan data besar yang berguna untuk mengidentifikasi informasi baru yang termasuk dalam kumpulan data yang dapat digunakan dalam pengumpulan data di masa depan [9].

### 2.3. Algoritma K-Means

*K-means clustering* adalah teknik *clustering* non-hierarkis yang mengelompokkan data ke dalam unit-unit pada tingkat tertentu [10]. *K-Means Clustering* adalah teknik klasifikasi di mana kumpulan data dibagi ke dalam beberapa grup. Pertama kali *Mac Queen* menggunakan algoritma *K-Means* pada tahun 1967. Metode *K-Means* melibatkan perbandingan data yang ada dari beberapa grup, di mana data dari satu grup memiliki karakteristik yang mirip dengan grup lainnya dan memiliki karakteristik yang berbeda dari data grup lainnya [11].

### 2.4. Strategi Promosi

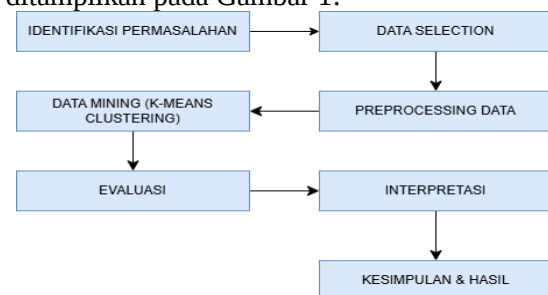
Promosi adalah jenis kegiatan yang digunakan untuk memperkenalkan produk atau layanan kepada masyarakat, seperti makanan, barang dagangan, bisnis, dan lainnya. Ada empat jenis kegiatan promosi, yaitu hubungan masyarakat, penjualan pribadi, penjualan, dan promosi penjualan [10].

### 2.6. DBI

*Davies-Bouldin Index* (DBI) adalah metrik evaluasi internal yang digunakan untuk menilai kualitas *clustering*. Metrik ini mengevaluasi beberapa parameter dari setiap kluster yang dihasilkan dari kluster lain, dan metrik yang lebih akurat menunjukkan *clustering* yang lebih baik [12]

## 3. METODE PENELITIAN

Metodologi penelitian yang digunakan dalam studi ini adalah penelitian kuantitatif, di mana data disajikan dalam format nominal/angka pasti atau data yang dapat diinterpretasikan pada skala numerik [1]. Analisis data dalam penelitian ini menggunakan algoritma *K-Means Clustering*. Fungsi dari algoritma *K-Means Clustering* adalah untuk mengklasifikasikan pelanggan berdasarkan data yang telah dikumpulkan sebelumnya dari biro jasa kantor yang bekerja sama dengan PT XYZ. Langkah-langkah metodologi penelitian ditampilkan pada Gambar 1:



Gambar 1 Tahapan Penelitian

## 4. HASIL DAN PEMBAHASAN

### 4.1. Hasil

#### 4.1.1. Identifikasi Permasalahan

Pada tahap ini, masalah yang dihadapi oleh PT XYZ diidentifikasi, terutama yang berkaitan dengan efektivitas strategi promosi berdasarkan statistik penjualan produk. Selain

itu, tujuan dari penelitian ini adalah untuk mengidentifikasi strategi pemasaran yang lebih efektif melalui segmentasi pelanggan menggunakan metode *K-Means* dan analisis literatur.

#### 4.1.2. Data Selection

Data yang digunakan dalam penelitian ini adalah informasi sekunder yang diperoleh dari biro jasa kantor yang bekerja sama dengan PT XYZ. Data ini berasal dari kertas faktur pembelian kendaraan dan kemudian didigitalisasi ke dalam format *Microsoft Excel* (xlsx) untuk digunakan sebagai bahan penelitian. Data yang digunakan merupakan data pelanggan dari bulan Desember 2023 sampai Mei 2024. Proses pengumpulan data dilakukan tanpa menggunakan survei, kuesioner, atau metode pendahuluan lainnya. Empat atribut Usia, Kecamatan, Kota/Kabupaten, dan Tipe Motor merupakan fokus dari 833 titik data yang digunakan. Berikut adalah dataset-nya:

| 1  | USIA | KECAMATAN    | KOTA/KABUPATEN    | TYPE MOTOR |
|----|------|--------------|-------------------|------------|
| 2  | 20,0 | MUNDU        | KABUPATEN CIREBON | B6H A/T    |
| 3  | 22,0 | PANGENAN     | KABUPATEN CIREBON | BBP A/T    |
| 4  | 34,0 | PALIMANAN    | KABUPATEN CIREBON | BEJ A/T    |
| 5  | 34,0 | TALUN        | KABUPATEN CIREBON | BEJ A/T    |
| 6  | 34,0 | KEJAKSAN     | KOTA CIREBON      | BJM A/T    |
| 7  | 52,0 | PANGENAN     | KABUPATEN CIREBON | BJM A/T    |
| 8  | 26,0 | HARJAMUKTI   | KOTA CIREBON      | B6H A/T    |
| 9  | 21,0 | KEJAKSAN     | KOTA CIREBON      | BJM A/T    |
| 10 | 34,0 | SUSUKAN      | KABUPATEN CIREBON | B6H A/T    |
| 11 | 24,0 | CILEDUG      | KABUPATEN CIREBON | BJM A/T    |
| 12 | 21,0 | PALIMANAN    | KABUPATEN CIREBON | BBP A/T    |
| 13 | 19,0 | PABUARAN     | KABUPATEN CIREBON | BEJ A/T    |
| 14 | 26,0 | DEPOK        | KABUPATEN CIREBON | BEJ A/T    |
| 15 | 21,0 | PALIMANAN    | KABUPATEN CIREBON | BEJ A/T    |
| 16 | 30,0 | HARJAMUKTI   | KOTA CIREBON      | BBP A/T    |
| 17 | 26,0 | KLANGENAN    | KABUPATEN CIREBON | B3W A/T    |
| 18 | 41,0 | WERU         | KABUPATEN CIREBON | BJM A/T    |
| 19 | 34,0 | KESAMB       | KOTA CIREBON      | B6H A/T    |
| 20 | 42,0 | MUNDU        | KABUPATEN CIREBON | B3W-S A/T  |
| 21 | 27,0 | SUSUKAN      | KABUPATEN CIREBON | BBP A/T    |
| 22 | 31,0 | ARJAWINANGUN | KABUPATEN CIREBON | BJM A/T    |
| 23 | 33,0 | KAPETAKAN    | KABUPATEN CIREBON | B6H A/T    |
| 24 | 26,0 | WALD         | KABUPATEN CIREBON | BJM A/T    |
| 25 | 56,0 | PABUARAN     | KABUPATEN CIREBON | B6H A/T    |
| 26 | 44,0 | SEDONG       | KABUPATEN CIREBON | B6H-AI A/T |
| 27 | 21,0 | GEBANG       | KABUPATEN CIREBON | B6H-AI A/T |

Gambar 2 Dataset

#### 4.1.3. Preprocessing

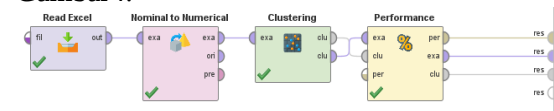
Pada tahap ini, operator *Nominal to Numerical* digunakan untuk mengonversi atribut kategorikal, seperti Kecamatan, Kota/Kabupaten, dan Nama Kendaraan, menjadi nilai numerik. Berikut adalah data setelah dikonversi menjadi numerik :

| Row No. | KECAMATAN | KOTA/KABU... | TYPE MOTOR | USIA |
|---------|-----------|--------------|------------|------|
| 1       | 0         | 0            | 0          | 20   |
| 2       | 1         | 0            | 1          | 22   |
| 3       | 2         | 0            | 2          | 34   |
| 4       | 3         | 0            | 2          | 34   |
| 5       | 4         | 1            | 3          | 34   |
| 6       | 1         | 0            | 3          | 52   |
| 7       | 5         | 1            | 0          | 26   |
| 8       | 4         | 1            | 3          | 21   |
| 9       | 6         | 0            | 0          | 34   |
| 10      | 7         | 0            | 3          | 24   |
| 11      | 2         | 0            | 1          | 21   |
| 12      | 8         | 0            | 2          | 19   |
| 13      | 9         | 0            | 2          | 26   |
| 14      | 2         | 0            | 2          | 21   |
| 15      | 5         | 1            | 1          | 30   |

Gambar 3 Transformasi

#### 4.1.4. Data Mining (K-Means Clustering)

Langkah berikutnya adalah proses data mining, yang dilakukan menggunakan teknik klusterisasi dengan algoritma *K-means* dalam aplikasi *Rapidminer*. Ada beberapa langkah yang terlibat dalam proses pemodelan algoritma *K-means*, seperti menganalisis data, mengklasifikasikan data sesuai dengan kriteria, mengonversi atribut kategorikal ke bentuk numerik, memulai proses klasifikasi, dan menilai hasil dari proses klasifikasi. Proses yang lebih panjang dari model algoritma *K-Means* dapat ditampilkan dalam Gambar4:



Gambar 4 Model Clustering K-Means

#### 4.1.5. Evaluasi

Pada fase evaluasi, dilakukan eksperimen klusterisasi untuk menentukan jumlah kluster dengan kinerja terbaik [4]. Penelitian ini didasarkan pada metode evaluasi seperti *Davies-Bouldin Index* (DBI). Metode DBI digunakan untuk meningkatkan kualitas kluster dengan mengukur jarak antar kluster (jarak intra-kluster) dan antar kluster (jarak inter-kluster), di mana nilai DBI yang lebih kecil menghasilkan hasil kluster yang lebih baik.

Pengujian ini dilakukan dengan berbagai nilai  $k$ , yaitu  $k=2$ ,  $k=3$ ,  $k=4$ , dan  $k=5$ . Pemilihan nilai  $k$  dilakukan secara eksperimental dengan mencoba beberapa kemungkinan jumlah kluster melalui proses iteratif. Proses iteratif ini berarti menjalankan algoritma *K-Means* secara berulang-ulang untuk setiap nilai  $k$ , mulai dari  $k=2$ ,  $k=3$ ,  $k=4$  hingga  $k=5$ . Hasil eksperimen untuk setiap kluster adalah sebagai berikut:

Tabel 1 Hasil Eksperimen DBI dari  $K=2$  sampai  $K=5$

| Parameter K | DBI   |
|-------------|-------|
| 2           | 0.802 |
| 3           | 1.028 |
| 4           | 0.921 |
| 5           | 0.854 |

#### 4.1.6. Interpretasi

Berdasarkan hasil evaluasi kinerja, kelas terbaik ditemukan pada  $k = 2$ , dengan *Davies-Bouldin Index* (DBI) sekitar 0,802. Nilai DBI sebesar 0 menunjukkan korelasi positif antara klaster terbaik dan klaster ideal. Jumlah item di setiap klaster adalah sebagai berikut:

#### Cluster Model

Cluster 0: 512 items  
Cluster 1: 321 items  
Total number of items: 833

Gambar 5 Cluster Model

Centroid dari setiap klaster dapat dilihat pada Gambar6:

| Attribute     | cluster_0 | cluster_1 |
|---------------|-----------|-----------|
| KECAMATAN     | 12.406    | 48.903    |
| KOTAKABUPATEN | 1.148     | 1.062     |
| TYPE MOTOR    | 3.244     | 3.227     |
| USIA          | 33.264    | 32.128    |

Gambar 6 Nilai Centroid

Berdasarkan Gambar 6 di atas, berikut adalah interpretasi untuk setiap atribut:

##### 1. Kecamatan

- Klaster 0 memiliki rata-rata sebesar 12,406.
- Klaster 1 memiliki rata-rata sebesar 48,903.
- Artinya, lebih banyak orang dari lebih banyak kecamatan yang tinggal di Klaster 1 dibandingkan dengan Klaster 0.

##### 2. Kota/Kabupaten

- Klaster 0 memiliki rata-rata sebesar 1,148.
- Klaster 1 memiliki rata-rata sebesar 1,062.
- Ini menunjukkan bahwa distribusi pelanggan di Klaster 0 dan 1 hampir sama di seluruh kota atau wilayah, dengan perbedaan yang sangat kecil.

##### 3. Type Motor

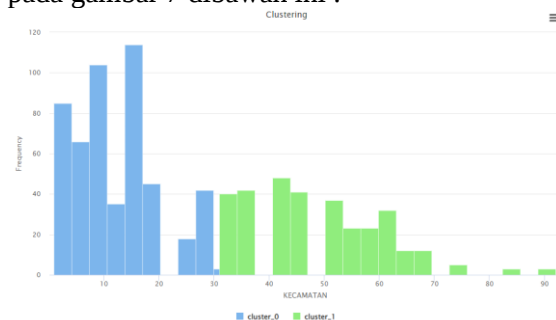
- Klaster 0 memiliki rata-rata sebesar 3,244.
- Klaster 1 memiliki rata-rata sebesar 3,227.
- Klaster ini menunjukkan preferensi yang mirip terkait tipe motor, dengan perbedaan yang sangat kecil.

##### 4. Usia

- Klaster 0 memiliki nilai rata-rata usia sebesar 33,264 tahun.
- Klaster 1 memiliki nilai rata-rata usia sebesar 32,128 tahun.
- Pelanggan di klaster 0 sedikit lebih dominan dibandingkan dengan pelanggan di klaster 1.

#### 4.2. Pembahasan

Berdasarkan hasil pengelompokan klaster diatas, diperoleh hasil persebaran data. Berikut ini visualisasi hasil persebaran data berdasarkan atribut kecamatan dapat dilihat pada gambar 7 dibawah ini :



Gambar 7 Visualisasi Persebaran Atribut Kecamatan

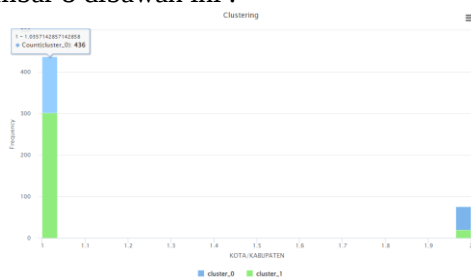
Tabel 2 menampilkan hasil *clustering* berdasarkan hasil pengelompokan di atas

Tabel 2 Hasil Pengelompokan Atribut Kecamatan

| Kecamatan  | C0 | Kecamatan   | C1 |
|------------|----|-------------|----|
| Mundu      | 13 | Plumbon     | 9  |
| Pangenan   | 20 | Pabedilan   | 31 |
| Palimanan  | 32 | Kedawung    | 18 |
| Talun      | 20 | Gempol      | 24 |
| Kejaksan   | 19 | Dukupuntang | 23 |
| Harjamukti | 30 | Panguragan  | 7  |
| Susukan    | 17 | Pasaleman   | 18 |
| Ciledug    | 63 | Pekalipan   | 6  |
| Pabuaran   | 16 | Plered      | 10 |
| Depok      | 25 | Lemahabang  | 10 |
| Klangenan  | 23 | Ciwaringin  | 7  |
| Weru       | 12 | Babakan     | 37 |

|              |    |                   |    |
|--------------|----|-------------------|----|
| Kesambi      | 28 | Lemahwung<br>kuk  | 14 |
| Arjawinangun | 14 | Suranenggala      | 9  |
| Kapetakan    | 28 | Tengah Tani       | 15 |
| Waled        | 63 | Greged            | 8  |
| Sedong       | 6  | Astanajapura      | 8  |
| Gebang       | 39 | Gegesik           | 9  |
| Sumber       | 18 | Karangsemb<br>ung | 15 |
| Karangwareng | 16 | Jamblang          | 12 |
| Gunung Jati  | 14 | Beber             | 12 |
| Losari       | 12 | Susukan<br>Lebak  | 5  |
| Plumbon      | 3  | Kaliwedi          | 3  |
|              |    | Pabuaran          | 3  |

Hasil persebaran data berdasarkan atribut Kota/Kabupaten dapat ditampilkan pada gambar 8 dibawah ini :



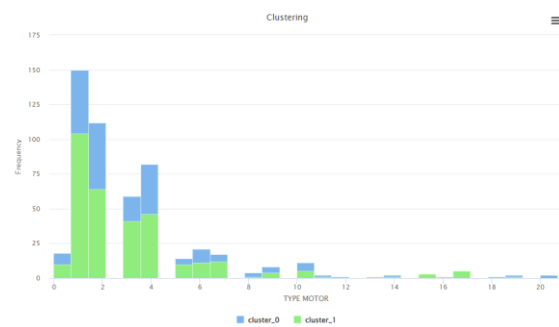
*Gambar 8 Visualisasi Persebaran Atribut Kota/Kabupaten*

Tabel 3 menampilkan hasil *clustering* berdasarkan hasil pengelompokkan di atas :

*Tabel 3 Hasil Pengelompokkan Atribut Kota/Kabupaten*

| Kota/Kab          | C0  | Kota/Kab          | C1  |
|-------------------|-----|-------------------|-----|
| Kota Cirebon      | 76  | Kota Cirebon      | 20  |
| Kabupaten Cirebon | 436 | Kabupaten Cirebon | 301 |

Hasil persebaran data berdasarkan atribut *Type Motor* ditampilkan pada gambar 9 dibawah ini:



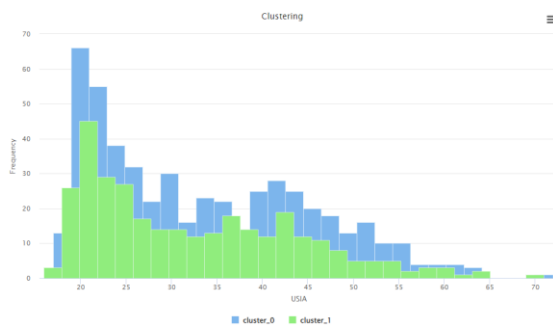
*Gambar 9 Visualisasi Persebaran Atribut Type Motor*

Tabel 4 menampilkan hasil *clustering* berdasarkan hasil pengelompokkan di atas :

*Tabel 4 Hasil Pengelompokkan Atribut Type Motor*

| Type Motor | C0  | Type Motor | C1  |
|------------|-----|------------|-----|
| BBP-A A/T  | 18  | BBP-A A/T  | 10  |
| B6H A/T    | 150 | B6H A/T    | 106 |
| BBP A/T    | 112 | BBP A/T    | 64  |
| BEJ A/T    | 59  | BEJ A/T    | 38  |
| BJM A/T    | 82  | BJM A/T    | 46  |
| B3W A/T    | 14  | B3W A/T    | 12  |
| B3W-S A/T  | 21  | B3W-S A/T  | 11  |
| B6H-AI A/T | 17  | B6H-AI A/T | 11  |
| BDJ M/T    | 4   | BDJ M/T    | 1   |
| BBS-R M/T  | 8   | BBS-R M/T  | 4   |
| BKU A/T    | 11  | BKU A/T    | 16  |
| 2PV R M/T  | 2   | 2PV R M/T  | 1   |
| B9T M/T    | 1   | BPV A/T    | 1   |
| BPV A/T    | 1   | B3M M/T    | 2   |
| B3M M/T    | 4   | B3F-I A/T  | 3   |
| B3F-I A/T  | 1   | BDJ-A M/T  | 1   |
| BDJ-A M/T  | 1   | BKU-S      | 5   |
| BKU-S A/T  | 3   |            |     |
| SE88       | 1   |            |     |
| BJ8 W A/T  | 2   |            |     |

Hasil persebaran data berdasarkan atribut *Usia* ditampilkan pada Gambar 10 di bawah ini.



Gambar 10 Visualisasi Persebaran Atribut Usia

Tabel 5 menampilkan hasil *clustering* berdasarkan hasil pengelompokan di atas :

Tabel 5 Hasil Pengelompokan Atribut Usia

| Usia | C0 | C1 | Usia | C0 | C1 |
|------|----|----|------|----|----|
| 17   | 2  | 2  | 42   | 17 | 13 |
| 18   | 11 | 9  | 43   | 13 | 6  |
| 19   | 31 | 17 | 44   | 12 | 6  |
| 20   | 35 | 23 | 45   | 8  | 6  |
| 21   | 26 | 22 | 46   | 12 | 7  |
| 22   | 29 | 15 | 47   | 12 | 4  |
| 23   | 26 | 14 | 48   | 6  | 4  |
| 24   | 12 | 18 | 49   | 4  | 4  |
| 25   | 13 | 9  | 50   | 9  | 3  |
| 26   | 19 | 5  | 51   | 7  | 2  |
| 27   | 13 | 12 | 52   | 9  | 3  |
| 28   | 9  | 9  | 53   | 3  | 2  |
| 29   | 16 | 5  | 54   | 7  | 3  |
| 30   | 14 | 9  | 55   | 3  | 2  |
| 31   | 5  | 5  | 56   | 7  | 1  |
| 32   | 11 | 5  | 57   | 1  | 1  |
| 33   | 7  | 7  | 58   | 3  | 3  |
| 34   | 16 | 6  | 59   | 1  | 0  |
| 35   | 13 | 7  | 60   | 3  | 3  |
| 36   | 9  | 10 | 61   | 3  | 0  |
| 37   | 6  | 8  | 62   | 1  | 1  |
| 38   | 8  | 6  | 63   | 1  | 0  |

|    |    |   |    |   |   |
|----|----|---|----|---|---|
| 39 | 11 | 8 | 64 | 2 | 2 |
| 40 | 14 | 6 | 70 | 0 | 1 |
| 41 | 11 | 6 | 71 | 1 | 0 |

Berdasarkan tabel hasil *clustering* di atas, setiap kluster memiliki karakteristik unik yang dapat membantu segmentasi pelanggan. Berikut adalah beberapa poin penting yang dapat ditekankan:

- Kluster 0 :
  - Dua kecamatan dengan jumlah penduduk terbanyak adalah **Ciledug** dan **Waled**, sedangkan dua kecamatan dengan jumlah penduduk paling sedikit adalah **Sedong** dan **Plumbon..**
  - Kendaraan yang paling populer adalah motor dengan tipe transmisi otomatis, seperti **Nmax** atau **Aerox**.
  - Kelompok usia yang paling dominan adalah **19–23 tahun**.
- Kluster 1 :
  - Kecamatan dengan jumlah penduduk terbanyak adalah **Babakan** dan **Pabedilan**, sedangkan kecamatan dengan jumlah penduduk paling sedikit adalah **Kaliwedi** dan **Pabuaran**.
  - Kendaraan yang paling populer adalah motor **Nmax**.
  - Kelompok usia yang paling dominan adalah **19–22 tahun**

## 5. KESIMPULAN

Berdasarkan hasil pengujian diatas, dapat disimpulkan hal-hal sebagai berikut :

- Metode *K-Means* berhasil mengidentifikasi pelanggan PT XYZ berdasarkan atribut usia, jenis kelamin, lokasi geografis, dan preferensi kendaraan. Hasil analisis menunjukkan bahwa metode ini berhasil membentuk dua kelas utama dengan karakteristik yang berbeda, sebagaimana ditunjukkan oleh nilai *Davies Bouldin Index* (DBI) terbaik sebesar 0,802 untuk jumlah kluster optimal  $k=2$ .
- Berdasarkan hasil segmentasi pelanggan menggunakan metode *K-Means*, berikut adalah strategi

promosi yang disarankan untuk tim pemasaran PT XYZ:

- a. Promosikan keunggulan kendaraan seperti Nmax dan Aerox, yang menjadi favorit pelanggan. Fokuskan kampanye pada fitur-fitur seperti kenyamanan berkendara, efisiensi bahan bakar, dan desain modern.
- b. Luncurkan kampanye media sosial interaktif melalui platform seperti *Instagram*, *TikTok*, dan *YouTube* untuk menarik minat pelanggan muda. Tawarkan program cicilan ringan dengan tenor fleksibel yang sesuai dengan kebutuhan mereka.
- c. Berikan diskon pembelian berikutnya sebagai insentif loyalitas bagi pelanggan yang telah melakukan pembelian sebelumnya. Tawarkan program tukar tambah kendaraan lama dengan kendaraan baru yang sesuai dengan kebutuhan mereka.
- d. Tawarkan diskon khusus di kecamatan dengan sedikit pembelian, seperti Ciledug, Waled, Kaliwedi, dan Pabuaran. Adakan promosi lokal seperti pameran kendaraan di pusat keramaian wilayah ini untuk meningkatkan visibilitas produk.
- e. Untuk meningkatkan loyalitas pelanggan dapat dilakukan acara komunitas motor, seperti touring bersama atau *workshop* pemeliharaan kendaraan, untuk membangun kedekatan dengan pelanggan. Tawarkan diskon perawatan rutin atau suku cadang untuk pelanggan setia. Jalankan program poin loyalitas di mana pelanggan dapat mengumpulkan poin dari pembelian atau servis yang dapat ditukarkan dengan hadiah menarik.

## UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak terkait yang telah memberi dukungan terhadap penelitian ini.

## DAFTAR PUSTAKA

- [1] W. W. Kristianto and C. Rudianto, "Penerapan Data Mining Pada Penjualan Produk Menggunakan Metode K-Means Clustering (Studi Kasus Toko Sepatu Kakikaki)," *J. Pendidik. Teknol. Inf.*, no. 5, pp. 90–98, 2020.
- [2] C. H. Ardana, A. A. A. A. A. Khoyum, and M. Faisal, "Segmentasi Pelanggan Penjualan Online Menggunakan Metode K-means Clustering," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 9, no. 1, pp. 1–9, 2024, doi: 10.14421/jiska.2024.9.1.1-9.
- [3] E. F. L. Awalina and W. I. Rahayu, "Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail," *J. Teknol. dan Inf.*, vol. 13, no. 2, pp. 122–137, 2023, doi: 10.34010/jati.v13i2.10090.
- [4] K. Anam, D. Sudrajat, and D. A. Kurnia, "Analisis Segmentasi Pelanggan Menggunakan Metode K-Means Clustering," *J. ICT Inf. Commun. Technol.*, vol. 21, no. 2303–3363, pp. 273–278, 2022.
- [5] B. T. Kristanti, A. Junaidi, and E. P. Mandyartha, "Implementasi K-Means Clustering Dalam Segmentasi Pelanggan Berdasarkan Usia, Pendapatan, Dan Model Rfm (Studi Kasus: Lantikya Store Jombang)," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4677.
- [6] S. D. K. Wardani, A. S. Ariyanto, M. Umroh, and D. Rolliawati, "Perbandingan Hasil Metode Clustering K-Means, Db Scanner & Hierarchical Untuk Analisa Segmentasi Pasar," *JIKO (Jurnal Inform. dan Komputer)*, vol. 7, no. 2, p. 191, 2023, doi: 10.26798/jiko.v7i2.796.
- [7] S. A. Perdana, S. F. Florentin, and A. Santoso, "ANALISIS SEGMENTASI PELANGGAN MENGGUNAKAN K-MEANS CLUSTERING STUDI KASUS APLIKASI ALFAGIFT Satria," *Sebatik*, vol. 26, no. 2, pp. 420–427, 2022, doi: 10.46984/sebatik.v26i2.2134.
- [8] S. Syahputra, S. Ramadani, and A. M. H. Pardede, "Menentukan Strategi Promosi Menggunakan Algoritma Clustering K-Means," *JOISIE (Journal Inf. Syst. Informatics Eng.)*, vol. 4, no. 1, p. 7, 2020, doi: 10.35145/joisie.v4i1.510.
- [9] Sharyanto and D. Lestari, "Penerapan Data Mining Untuk Menentukan Segmentasi Pelanggan Dengan Menggunakan Algoritma K-Means dan Model RFM Pada E-Commerce," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 4,

- pp. 866–871, 2022, doi: 10.30865/jurikom.v9i4.4525.
- [10] O. O. Tensao, I. N. Y. A. Wijaya, and K. Q. Fredlina, “Analisa Data Mining dengan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru Pada STMIK Primakara,” *Inf. (Jurnal Inform. dan Sist. Informasi)*, vol. 14, no. 1, pp. 1–17, 2022, doi: 10.37424/informasi.v14i1.135.
- [11] Haris Kurniawan, Sarjon Defit, and Sumijan, “Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal,” *J. Appl. Comput. Sci. Technol.*, vol. 1, no. 2, pp. 80–89, 2020, doi: 10.52158/jacost.v1i2.102.
- [12] K. Dbscan and Y. Hasan, “Pengukuran Silhouette Score dan Davies-Bouldin Index pada Hasil Cluster,” vol. 06, no. 01, pp. 60–74, 2024.