

# Pemanfaatan Data Mining Untuk Segmentasi Pelanggan Dalam Meningkatkan Efektivitas Penawaran Di Showroom Menggunakan Metode K-Means

Emiya Pepayosa Br Surbakti<sup>1</sup>, Ahmad Fitri Boy<sup>2</sup>, Suardi Yakub<sup>3</sup>

<sup>1,2,3</sup> Sistem Informasi, STMIK Triguna Dharma

Email: <sup>1</sup>emiyasurbakti3@gmail.com, <sup>2</sup>ahmadfitriboy@gmail.com, <sup>3</sup>yakubsuardi@gmail.com

Email Penulis Korespondensi: <sup>1</sup>emiyasurbakti3@gmail.com

## Abstrak

Perkembangan teknologi informasi mendorong perusahaan untuk memanfaatkan data sebagai dasar dalam mengambil keputusan strategis. Salah satu pendekatan yang efektif adalah segmentasi pelanggan menggunakan teknik data mining. Penelitian ini dilakukan pada Showroom Pandu Motor dengan tujuan untuk mengelompokkan pelanggan berdasarkan karakteristik tertentu guna mendukung efektivitas strategi penawaran. Segmentasi ini dianggap penting karena dapat membantu pihak showroom dalam memahami preferensi pelanggan secara lebih spesifik. Metode yang digunakan dalam penelitian ini adalah algoritma K-Means Clustering, yang mampu membagi data ke dalam tiga kelompok pelanggan berdasarkan atribut seperti usia, penghasilan, harga motor, lama kredit, dan status OJK. Proses analisis dilakukan secara bertahap mulai dari pengumpulan data, normalisasi, pemilihan centroid awal, hingga perhitungan jarak Euclidean dan proses iterasi yang berulang sampai hasil pengelompokan stabil (konvergen). Metode ini dinilai efisien karena dapat mengolah data dalam jumlah besar dan menghasilkan segmentasi yang jelas. Hasil dari penelitian ini menunjukkan bahwa algoritma K-Means berhasil mengidentifikasi tiga cluster pelanggan yang berbeda tingkat potensinya, yaitu tinggi, sedang, dan rendah. Dengan informasi tersebut, showroom dapat menyusun strategi pemasaran yang lebih tepat sasaran untuk tiap segmen pelanggan. Penelitian ini diharapkan dapat menjadi referensi bagi pengembangan sistem analisis pelanggan berbasis data mining, khususnya dalam sektor perdagangan otomotif.

**Kata Kunci:** Data Mining, Segmentasi Pelanggan, K-Means Clustering, Showroom, Strategi Pemasaran

## Abstract

*The advancement of information technology encourages companies to utilize data as a foundation for making strategic decisions. One effective approach is customer segmentation using data mining techniques. This study was conducted at Pandu Motor Showroom with the aim of grouping customers based on specific characteristics to support the effectiveness of marketing strategies. This segmentation is considered important because it helps the showroom better understand customer preferences. The method used in this research is the K-Means Clustering algorithm, which is capable of dividing data into three customer groups based on attributes such as age, income, motorcycle price, loan duration, and OJK status. The analysis process was carried out in stages, starting from data collection, normalization, initial centroid selection, Euclidean distance calculation, and iterative processes until the grouping results stabilized (converged). This method is considered efficient as it can process large amounts of data and produce clear segmentation. The results of this study show that the K-Means algorithm successfully identified three customer clusters with different potential levels: high, medium, and low. With this information, the showroom can formulate more targeted marketing strategies for each customer segment. This research is expected to serve as a reference for developing data mining-based customer analysis systems, particularly in the automotive retail sector.* **Keywords:** Data Mining, Customer Segmentation, K-Means Clustering, Showroom, Marketing Strategy.

## 1. PENDAHULUAN

Di era digital dan globalisasi saat ini, keberadaan showroom sebagai salah satu bentuk usaha di bidang perdagangan kendaraan bermotor memiliki peranan yang sangat penting dalam menjembatani interaksi antara produsen dan konsumen [1]. Showroom tidak hanya menjadi tempat penjualan, tetapi juga memberikan pengalaman langsung bagi konsumen untuk menilai produk sebelum membeli, sehingga dapat meningkatkan kepercayaan dan loyalitas pelanggan [2]. Industri otomotif Indonesia menunjukkan pertumbuhan yang signifikan, baik dalam penjualan kendaraan baru maupun bekas [3]. Dalam menghadapi persaingan bisnis yang semakin ketat, showroom dituntut untuk mampu menyusun strategi pemasaran yang lebih tepat sasaran, salah satunya dengan memahami karakteristik pelanggan melalui pendekatan berbasis data [4]. Data telah menjadi aset penting dalam mendukung pengambilan keputusan bisnis. Melalui pendekatan data mining, perusahaan dapat menggali informasi berharga dari kumpulan data besar guna mengidentifikasi pola-pola tertentu dalam perilaku konsumen [5]. Salah satu penerapan data mining yang relevan dalam konteks bisnis adalah segmentasi pelanggan, yaitu proses pengelompokan pelanggan ke dalam beberapa kelompok berdasarkan kemiripan karakteristik atau perilaku [6].

Segmentasi pelanggan membantu perusahaan dalam memahami kebutuhan, preferensi, dan tingkat loyalitas pelanggan, sehingga strategi pemasaran dapat disesuaikan dengan masing-masing segmen secara lebih efektif [7]. Algoritma *K-Means Clustering* merupakan salah satu metode populer dalam segmentasi karena kemampuannya mengelompokkan data tanpa label secara cepat dan efisien [8]. Kelebihan metode ini antara lain mudah diterapkan, skalabel, dan cocok untuk analisis dataset besar, meskipun hasilnya sangat dipengaruhi oleh inisialisasi centroid dan pemilihan jumlah cluster ( $k$ ) [9]. Berbagai penelitian terdahulu telah membuktikan efektivitas metode K-Means dalam bidang pemasaran, termasuk untuk analisis pelanggan di sektor ritel, perbankan, dan e-commerce [10]. Dalam konteks

showroom kendaraan, penerapan metode ini dapat digunakan untuk mengelompokkan pelanggan berdasarkan atribut tertentu, seperti usia, penghasilan, jenis kendaraan yang diminati, hingga frekuensi pembelian [11].

Penelitian ini dilakukan di Showroom Pandu Motor yang berlokasi di Riau, sebuah showroom yang bergerak di bidang jual-beli motor baru dan bekas dengan cakupan wilayah operasional yang luas. Showroom ini menghadapi tantangan dalam menyusun strategi penawaran yang sesuai dengan karakteristik pelanggan yang beragam. Oleh karena itu, diperlukan penerapan metode analisis yang mampu mengelompokkan pelanggan secara otomatis dan sistematis. Berdasarkan hal tersebut, penelitian ini bertujuan untuk menerapkan metode K-Means dalam proses segmentasi pelanggan di Showroom Pandu Motor guna meningkatkan efektivitas strategi penawaran produk. Melalui hasil pengelompokan yang diperoleh, showroom diharapkan dapat merancang pendekatan pemasaran yang lebih terfokus dan relevan dengan kebutuhan setiap segmen pelanggan [12].

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Metodologi penelitian merupakan serangkaian langkah ilmiah yang dirancang secara sistematis untuk memahami dan menyelesaikan permasalahan segmentasi pelanggan di Showroom Pandu Motor. Penelitian ini bertujuan untuk memperoleh data yang mencerminkan kondisi nyata di lapangan sebagai dasar penerapan algoritma K-Means dalam meningkatkan efektivitas penawaran. Adapun tahapan pelaksanaan penelitian ini dijelaskan sebagai berikut:

- Observasi :** Observasi dilakukan dengan cara mengamati secara langsung aktivitas penjualan dan interaksi pelanggan di showroom. Metode ini bersifat non-partisipatif, di mana peneliti hanya bertindak sebagai pengamat. Data yang diperoleh meliputi perilaku dan pola kunjungan pelanggan, yang kemudian dianalisis untuk mendukung proses segmentasi pelanggan menggunakan algoritma K-Means.
- Wawancara :** Metode wawancara digunakan untuk menggali informasi langsung dari pihak-pihak yang terlibat dalam aktivitas penjualan di Showroom Pandu Motor, seperti pemilik atau staf pemasaran. Informasi yang dikumpulkan mencakup faktor-faktor yang memengaruhi segmentasi pelanggan, strategi penawaran yang diterapkan, serta kendala dalam kegiatan pemasaran. Hasil wawancara ini digunakan sebagai dasar untuk memperkaya proses analisis data.

### 2.2 Knowledge Discovery In Database (KDD)

Knowledge Discovery in Database (KDD) adalah proses untuk menemukan informasi yang berguna dan pola-pola menarik dalam kumpulan data besar, yang sering kali kompleks dan tersembunyi. Tujuan KDD adalah mengolah data ini menjadi informasi yang lebih bermakna, sehingga dapat memberikan wawasan baru yang sebelumnya tidak terlihat. Proses ini melibatkan sejumlah tahapan analisis untuk mengidentifikasi pola atau tren yang signifikan, seperti perilaku pelanggan, tren finansial, atau korelasi penting dalam penelitian ilmiah. Salah satu bagian penting dari KDD adalah data mining, yang membantu mengungkap pola-pola tersebut dari data dalam jumlah besar. Informasi yang dihasilkan oleh KDD dapat dimanfaatkan untuk berbagai keperluan, seperti pengambilan keputusan strategis, optimasi bisnis, atau pemahaman yang lebih dalam terhadap suatu fenomena.

- Data Selection :** Pada tahap ini dilakukan pemilihan data yang akan digunakan dalam proses data mining. Data yang digunakan dalam penelitian ini berasal dari database showroom pandu motor yang sudah ada dan terbentuk tetapi belum sesuai dengan format database penelitian.
- Pre procesing :** Tahap ini bertujuan untuk meningkatkan kualitas data, termasuk membersihkan data dari masalah seperti ketidaklengkapan, gangguan, atau outlier. Proses Pre-Processing dan Cleaning Data dilakukan untuk memastikan data bebas dari duplikasi, ketidakkonsistenan, dan kesalahan.
- Data Transformation :** Setelah melalui tahap Pre-Processing langkah berikutnya adalah Transformation, yaitu proses mengubah data dari format awalnya menjadi bentuk yang siap digunakan untuk proses penambangan. Transformasi ini bertujuan untuk mempermudah penggalian data dan membantu menemukan pengetahuan baru. Setelah tahap transformasi selesai, dataset yang telah disiapkan dapat dimanfaatkan untuk data mining.
- Data Mining :** Pada tahap ini, dilakukan pemilihan tugas data mining yang sesuai serta penentuan jenis data mining yang akan diterapkan. Selain itu, algoritma yang akan digunakan dipilih, diimplementasikan, dan dijalankan untuk mendapatkan hasil yang diinginkan.
- Interpretation/Evaluas :** pada tahap ini dilakukan Interpretation/Evaluasi melalui pola yang dihasilkan dari hasil klasifikasi. Pada tahap ini juga penggunaan dan umpan balik secara keseluruhan pada pola dan hasil penemuan diperoleh dengan data mining.

### 2.3 Algoritma K-Means

Algoritma K-Means Clustering adalah salah satu teknik yang sering digunakan dalam analisis data dan data mining. Metode ini berfungsi untuk mengelompokkan data berdasarkan kesamaan tertentu tanpa memerlukan label atau referensi eksternal. Dengan demikian, data akan dibagi ke dalam beberapa grup yang memiliki karakteristik yang mirip. Pendekatan ini sangat berguna untuk menyederhanakan analisis data yang kompleks, membuatnya lebih terstruktur, dan mudah dipahami, terutama ketika menangani dataset yang besar dan beragam.

Metode K-Means bertujuan untuk mengelompokkan data ke dalam beberapa cluster, sehingga data dengan karakteristik serupa dikelompokkan dalam cluster yang sama, sementara data dengan karakteristik yang berbeda ditempatkan dalam cluster yang berbeda. Berikut adalah langkah-langkah melakukan clustering dengan metode k-means:

1. Tentukan jumlah cluster yang diinginkan dengan menetapkan nilai  $k$ .
2. Inisialisasi  $k$  centroid, yang dapat dipilih secara acak pada tahap awal.
3. Setiap data atau objek kemudian diklasifikasikan ke dalam cluster terdekat berdasarkan jarak. Kedekatan ini diukur dari jarak antara objek dengan pusat cluster. Proses ini melibatkan perhitungan jarak dari setiap data terhadap semua centroid, dan data akan dimasukkan ke dalam cluster dengan jarak terkecil. Pengukuran jarak umumnya menggunakan rumus Euclidean, yang dirumuskan sebagai:

$$D(i,j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

**Keterangan:**

- $D(i,j)$  : Jarak data ke  $i$  ke pusat *cluster*  $j$
  - $X_{ki}$  : Data ke  $i$  pada atribut data ke  $k$
  - $X_{kj}$  : Titik pusat ke  $j$  pada atribut ke  $k$
4. Perbarui pusat cluster berdasarkan keanggotaan *cluster* yang ada saat ini. Pusat *cluster* diperoleh dari rata-rata semua data atau objek yang termasuk dalam *cluster* tersebut. Sebagai alternatif, median dari *cluster* tersebut juga dapat digunakan jika diperlukan.
  5. Perbarui penugasan objek sesuai dengan pusat cluster yang baru. Jika pusat cluster tidak mengalami perubahan, proses klusterisasi selesai. Jika belum, ulangi langkah ketiga hingga pusat klaster tidak berubah lagi.

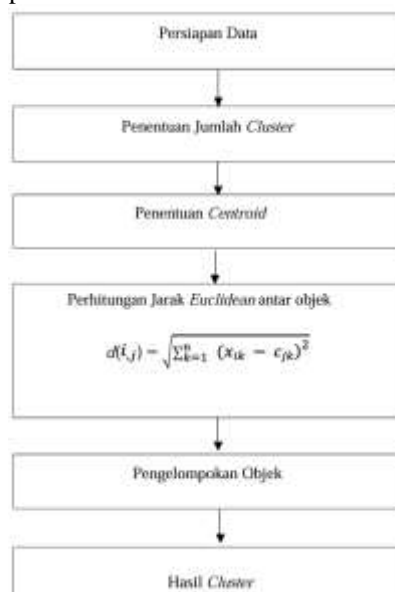
### 3. HASIL DAN PEMBAHASAN

#### 3.1 Penerapan Algoritma K-Means

Metode *K-Means Clustering* diterapkan untuk melakukan segmentasi pelanggan di Showroom Pandu Motor berdasarkan hasil analisis data mining. Segmentasi ini dibagi menjadi tiga *cluster* berdasarkan hasil perhitungan *K-Means Clustering*, yaitu pelanggan dengan tingkat potensi tinggi (C1), pelanggan dengan tingkat potensi sedang (C2), dan pelanggan dengan tingkat potensi rendah (C3). Dengan pembagian ini, *showroom* dapat menyusun strategi penawaran yang lebih efektif sesuai dengan karakteristik setiap *cluster*. Berikut adalah tahapan perhitungan metode *K-Means Clustering* yang digunakan dalam penelitian ini.

##### 3.1.1 Kerangka Kerja Metode K-Means

Kerangka kerja metode K-Means terdiri dari beberapa tahapan utama yang dimulai dengan persiapan data sebagai langkah awal dalam proses analisis. Selanjutnya, dilakukan penentuan jumlah *cluster* yang diinginkan dengan menetapkan nilai  $k$ , diikuti dengan inisialisasi titik pusat (*centroid*) secara acak. Setelah itu, setiap data akan dihitung jaraknya terhadap masing-masing *centroid* menggunakan rumus jarak *Euclidean* untuk menentukan kedekatannya. Data kemudian dikelompokkan ke dalam *cluster* yang memiliki jarak paling dekat dengan *centroid*. Proses ini dilakukan secara iteratif dengan memperbarui posisi *centroid* hingga konvergen, dan hasil akhirnya adalah pembentukan *cluster* yang merepresentasikan segmentasi data secara optimal.



Gambar 1. Kerangka Kerja K-Means

**3.2 Implementasi Data**

Implementasi data dalam penelitian ini mencakup serangkaian proses mulai dari persiapan hingga penerapan algoritma K-Means untuk segmentasi pelanggan. Tahap awal dilakukan dengan menyusun data pelanggan yang terdiri dari atribut usia, penghasilan, harga motor, lama kredit, dan status OJK. Data kemudian dinormalisasi agar berada dalam skala yang seragam dan layak untuk dianalisis menggunakan metode K-Means. Setelah data dinyatakan siap, proses segmentasi dimulai dengan menentukan jumlah *cluster* sebanyak tiga, yang masing-masing merepresentasikan tingkat potensi pelanggan: tinggi (C1), sedang (C2), dan rendah (C3). Inisialisasi centroid awal dipilih secara acak dari data pelanggan. Selanjutnya, dilakukan perhitungan jarak Euclidean antara setiap data dengan masing-masing centroid untuk menentukan kedekatan dan pengelompokan awal. Iterasi kemudian dijalankan dengan memperbarui posisi centroid berdasarkan rata-rata nilai data dalam setiap cluster hingga tercapai kondisi konvergen, yaitu ketika posisi centroid tidak lagi berubah secara signifikan. Hasil akhir dari implementasi ini berupa pengelompokan pelanggan ke dalam tiga cluster yang dapat digunakan sebagai dasar dalam merancang strategi penawaran yang lebih efektif dan sesuai dengan karakteristik tiap segmen. Proses clustering dilakukan dengan menentukan jumlah centroid (k) sebanyak tiga, sesuai dengan kebutuhan segmentasi pelanggan. Penghitungan jarak antar data dan centroid dilakukan menggunakan metode **Euclidean Distance** untuk menentukan kedekatan setiap data terhadap masing-masing cluster. Berikut menyajikan contoh data pelanggan yang digunakan dalam proses segmentasi.

Tabel 1. Tabel Normalisasi Data

No	Nama Pelanggan	Usia (Tahun)	Penghasilan	Harga Motor	Lama Kredit
1	JAMAL	5	55	1994	3
2	KUNDEN	3	45	113	2
3	NARSIM	51	76	2495	3
4	BAYU RIDWAN DANA	27	38	193	4
5	YUDISTIRA	23	3	125	2
6	FERY KURNIAWAN	51	25	1352	3
7	PURYANTO	48	5	108	0
8	SAMUEL	45	7	3148	3
9	SRI LESTARI	3	5	88	0
10	SUDIONO	34	55	83	0
11	BINTARNO	38	68	2236	2
...	.....	.....	.....	...	.....
...	.....	.....	.....	...	.....
...	.....	.....	.....	...	.....
30	AMITA	41	53	1258	1
31	AZUAR	45	6	224	1
32	JEFRIZAL	24	35	118	1
33	RINI ETRI	33	63	1076	0
34	DARIANA	27	32	155	0
35	APRIL YAKUB LASE	44	49	243	3

1. Penentuan Pusat Cluster (Centroid)

Penentuan titik pusat awal *cluster* (*K Centroid*) dalam penelitian ini dilakukan dengan memilih secara acak dari data yang terdapat pada Tabel 3.1. Proses ini menghasilkan titik pusat *cluster* sebagai berikut:

Tabel 2. Data Centroid Awal

Centroid	Kode Nomor	Nama Pelanggan	Usia	Penghasilan	Harga Motor	Lama Kredit
C1	P2	KUNDEN	3	45	113	2
C2	P27	ASMAR	5	5	1356	2
C3	P23	LUKI SETIAWAN	24	33	1154	1

2. Perhitungan Jarak Euclidean Antar Objek

Jarak antara setiap data dan *centroid* dihitung menggunakan rumus *Euclidean*. Data tersebut kemudian akan dikelompokkan ke dalam *cluster* terdekat. Perhitungan jarak (*distance*) dilakukan dengan membandingkan setiap variabel dalam sampel data terhadap *centroid* masing-masing.

Dengan *Centroid* 1 (3,45,113,2)

Jarak antara data P2 dengan titik C1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(3 - 3)^2 + (45 - 45)^2 + (113 - 113)^2 + (2 - 2)^2}$$

$$= 0$$

Dengan *Centroid* 2 (5,5,1356,2)

Jarak antara P2 dengan titik C2

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(3 - 5)^2 + (45 - 5)^2 + (113 - 1356)^2 + (2 - 2)^2}$$

$$= 1243.64$$

Dengan *Centroid* 3 (24,33,1154,1)

Jarak antara P2 dengan titik C3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(3 - 24)^2 + (45 - 33)^2 + (113 - 1154)^2 + (2 - 1)^2}$$

$$= 1041.28$$

Proses perhitungan dilakukan hingga objek ke-35 dengan metode yang sama. Hasil perhitungan pada iterasi pertama dapat dilihat pada tabel di bawah ini, di mana jarak terdekat ditentukan berdasarkan nilai yang paling mendekati pusat *cluster*. Untuk lebih jelasnya, rincian jarak pada setiap baris data disajikan dalam tabel berikut, yang menunjukkan hasil segmentasi pelanggan di *Showroom* Pandu Motor menggunakan metode *K-Means*.

Tabel 3. Iterasi 1

KN	Nama Pelanggan	C1	C2	C3	MIN	Cluster
P1	JAMAL	1881,028	639,957	840,5052	639,957	2
P2	KUNDEN	0	1243,645	1041,281	0	1
P3	NARSIM	2382,685	1142,138	1341,962	1142,138	2
P4	YUDISTIRA	83,83913	1163,678	961,0224	83,83913	1
...	.....	.....	.....	...	.....	...
...	.....	.....	.....	...	.....	...
...	.....	.....	.....	...	.....	...
P33	RINI ETRI	963,6374	287,3186	84,0595	84,0595	3
P34	DARIANA	50,12983	1201,507	999,0055	50,12983	1
P35	APRIL YAKUB LASE	136,3745	1114,552	911,3622	136,3745	1

Dari Tabel 3.3 terdapat keanggotaan sebagai berikut :

- a. *Cluster* 1 = {P2, P4, P5, P7, P9, P10, P12, P16, P17, P18, P19, P20, P25, P26, P29, P31, P32, P34, P35}
- b. *Cluster* 2 = {P1, P3, P6, P8, P11, P13, P14, P15, P22, P24, P27}
- c. *Cluster* 3 = {P21, P23, P28, P30, P33}

Maka dari hasil *Cluster* diatas kemudian dicari pembaruan *centroid* dari rata-rata *cluster* yang dibentuk. Pembaruan *centroid* dari hasil *Cluster* dilakukan seperti berikut :

- a. *Cluster* 1  
= rata-rata (*Cluster* 1)  
= (30,84211;35,26316;130,4737;0,842105)
- b. *Cluster* 2  
=rata-rata (*Cluster* 2)  
=(35,54545;38,54545;2053,727;2,363636)

c. Cluster 3

= rata-rata (Cluster 3)

= (34,4;39,8;1093,4;1)

BCV : *Between Cluster Variation*

WCV : *within Cluster Variation*

Pada langkah ini dihitung pula rasio antara BCV dan WCV karena *centroid*

m1 = (30,84211;35,26316;130,4737;0,842105)      m2 =

(35,54545;38,54545;2053,727;2,363636)      m3 =

(34,4;39,8;1093,4;1)

$$d(m1,m2) \sqrt{\sum_{i=1}^n (xi - yi)^2}$$

$$\sqrt{(30,84211 - 35,54545)^2 + (35,26316 - 38,54545)^2 + (130,4737 - 2053,727)^2 + (0,842105 - 2,363636)^2}$$

$$= 1923,263$$

$$d(m1,m3) \sqrt{\sum_{i=1}^n (xi - yi)^2}$$

$$\sqrt{(30,84211 - 34,4)^2 + (35,26316 - 39,8)^2 + (130,4737 - 1093)^2 + (0,842105 - 1)^2}$$

$$= 962,9436$$

$$d(m2,m3) \sqrt{\sum_{i=1}^n (xi - yi)^2}$$

$$\sqrt{(35,54545 - 34,4)^2 + (38,54545 - 39,8)^2 + (2053,727 - 1093)^2 + (2,363636 - 1)^2}$$

$$= 960,3297$$

BCV = d(m1,m2)+ d(m1,m3)+ d(m2,m3) = **3846.5363**

WCV = yaitu memilih jumlah jarak terkecil dipangkat 2 antara data dengan *centroid* pada masing-masing *cluster*

3. Tentukan *Centroid* Baru Dari Data Yang Ada Di Masing – Masing *Cluster* Dan Perhitungan Iterasi *Cluster* Kembali Selanjutnya hitung *Iterasi* seterusnya sampai nilai *centroid* nya tidak berubah lagi dan Iterasi hentikan. Dibawah ini merupakan perhitungan *Iterasi* ke 4 sebagai *Iterasi* terakhir dari perhitungan sebelumnya.

Setelah didapatkan hasil dari setiap *cluster* kemudian pusat *cluster* baru dihitung berdasarkan data tiap-tiap *cluster* yang sudah didapatkan menggunakan persamaan *Iterasi* 4 yang sesuai dengan pusat *cluster* sebagai berikut :

Tabel 4. Data Centroid Iterasi 4

<i>Centroid</i>	U	P	HM	LK
<i>Centroid</i> 1	30,84211	35,26316	130,4737	0,842105
<i>Centroid</i> 2	35,42857	40,42857	2398,143	2,714286
<i>Centroid</i> 3	35	37,77778	1252,333	1,333333

Hitung jarak data ke *centroid* dengan rumus *Euclidean*, data tersebut akan ditetapkan sebagai anggota dari *cluster* terdekatnya. Menghitung *distance* (jarak) antara *variable* dari setiap sampel data dengan *centroid* yaitu:

a. Dengan *Centroid* 1 (30,84211;35,26316;130,4737;0,842105)

- Jarak antara data P2 dengan titik C1

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

$$= \sqrt{(3 - 30,84211)^2 + (45 - 35,26316)^2 + (113 - 130,473)^2 + (2 - 0,84210)^2}$$

$$= 34.30247$$

b. Dengan *Centroid* 2 (35,42857;40,42857;2398,143;2,714286)

- Jarak antara data P2 dengan titik C2

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

$$= \sqrt{(3 - 35,42857)^2 + (45 - 40,42857)^2 + (113 - 2398,143)^2 + (2 - 2,714286)^2}$$

= **2285.378**

c. Dengan *Centroid* 3 (35;37,77778;1252,333;1,333333)

- Jarak antara data P2 dengan titik C3

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

$$= \sqrt{(3 - 35)^2 + (45 - 37,77778)^2 + (113 - 1252,333)^2 + (1 - 1,333333)^2}$$

= **1139.806**

Lakukan proses perhitungan yang sama sampai dengan objek ke 35. Adapun hasil perhitungan iterasi 2 dapat dilihat pada table dibawah ini. Dimana jarak terdekat dilihat dari perhitungan yang paling dekat ke pusat *cluster*. Untuk lebih lengkapnya jarak pada setiap baris data, hasilnya seperti pada table berikut :

Tabel 5. Iterasi 4

KN	Nama Pelanggan	C1	C2	C3	MIN	Cluster
P1	JAMAL	1863,811	405,5487	742,4748	405,5487	2
P2	KUNDEN	34,30247	2285,378	1139,806	34,30247	1
P3	NARSIM	2364,964	104,3513	1243,358	104,3513	2
P4	BAYU RIDWAN DANA	62,78347	2205,161	1059,367	62,78347	1
...	.....	.....	.....	...	.....	...
...	.....	.....	.....	...	.....	...
...	.....	.....	.....	...	.....	...
P31	AZUAR	99,01502	2174,437	1028,873	99,01502	1
P32	JEFRIZAL	14,2303	2280,179	1134,39	14,2303	1
P33	RINI ETRI	945,9359	1322,341	178,1443	178,1443	3
P34	DARIANA	25,05313	2243,176	1097,379	25,05313	1
P35	APRIL YAKUB LASE	114,1432	2155,177	1009,437	114,1432	1

Dari table 3.13 terdapat keanggotaan sebagai berikut :

3.2.5 *Cluster* 1 = {P2, P4, P5, P7, P7, P9,P10, P12, P16, P17, P18, P19, P20, P25, P26, P29, P31, P32, P34, P35}

3.2.5 *Cluster* 2 = {P1, P3, P8, P11, P13, P15, P22}

3.2.5 *Cluster* 3 = {P6, P14, P21, P23, P24, P27, P28, P30, P33}

4. Pengelompokan Objek

Setelah dilakukan perhitungan jarak antara setiap data pelanggan dengan masing-masing centroid menggunakan rumus Euclidean, maka setiap data dikelompokkan ke dalam cluster dengan jarak terdekat. Proses ini merupakan inti dari metode K-Means, di mana objek-objek data akan terus berpindah dari satu cluster ke cluster lain hingga posisi centroid tidak lagi berubah.

Tabel 6. Pengelompokan Objek

KN	Nama Pelanggan	C1	C2	C3	MIN	Cluster
----	----------------	----	----	----	-----	---------

P2	KUNDEN	34,30247	2285,378	1139,806	34,30247	1
p4	BAYU RIDWAN DANA	62,78347	2205,161	1059,367	62,78347	1
p32	JEFRIZAL	14,2303	2280,179	1134,39	14,2303	1
p34	DARIANA	25,05313	2243,176	1097,379	25,05313	1
p35	APRIL YAKUB LASE	114,1432	2155,177	1009,437	114,1432	1
p1	JAMAL	1863,811	405,5487	742,4748	405,5487	2
p3	NARSIM	2364,964	104,3513	1243,358	104,3513	2
p8	SAMUEL	3017,693	750,663	1895,944	750,663	2
p6	FERY KURNIAWAN	1221,738	1046,373	101,7619	101,7619	3
p14	SUKMA WATI	1443,713	824,3067	322,2274	322,2274	3
p21	SOZANOLO ZALUKHU	993,6187	1274,308	129,6219	129,6219	3
p23	LUKI SETIAWAN	1023,552	1244,219	99,06252	99,06252	3

5. Hasil Proses K-Means

Setelah dilakukan perhitungan dan diperoleh rasio yang tidak mengalami perubahan, proses iterasi dihentikan pada iterasi keempat karena hasil yang didapat sudah konstan. Dengan begitu, posisi *centroid* tetap dan tidak mengalami pergeseran, segmentasi pelanggan berdasarkan metode K-Means dapat ditentukan, yang selanjutnya digunakan untuk meningkatkan efektivitas penawaran di Showroom Pandu Motor.

a. Hasil Cluster 1 yang menunjukkan performa tertinggi

Tabel 7. Hasil Cluster

KN	Nama Pelanggan	C1	C2	C3	MIN	Cluster
p2	KUNDEN	34,30247	2285,378	1139,806	34,30247	1
p4	BAYU RIDWAN DANA	62,78347	2205,161	1059,367	62,78347	1
...	.....	.....	.....	...	.....	...
...	.....	.....	.....	...	.....	...
...	.....	.....	.....	...	.....	...
p34	DARIANA	25,05313	2243,176	1097,379	25,05313	1
p35	APRIL YAKUB LASE	114,1432	2155,177	1009,437	114,1432	1

b. Hasil Cluster 2 menunjukkan volume pembelian dengan performa menengah

Tabel 8. Hasil Cluster

KN	Nama Pelanggan	C1	C2	C3	MIN	Cluster
p1	JAMAL	1863,811	405,5487	742,4748	405,5487	2
p3	NARSIM	2364,964	104,3513	1243,358	104,3513	2
...	.....	.....	.....	...	.....	...
...	.....	.....	.....	...	.....	...
...	.....	.....	.....	...	.....	...
p15	RINTO PRATAMA	2364,744	104,1107	1243,175	104,1107	2

p22	KOKO HARYADI	2353,791	90,58483	1232,07	90,58483	2
-----	--------------	----------	----------	---------	----------	---

c. Hasil Cluster 3 menunjukkan volume pembelian dengan performa rendah

Tabel 9. Hasil Cluster

KN	Nama Pelanggan	C1	C2	C3	MIN	Cluster
p6	FERY KURNIAWAN	1221,738	1046,373	101,7619	101,7619	3
p14	SUKMA WATI	1443,713	824,3067	322,2274	322,2274	3
p21	SOZANOLO ZALUKHU	993,6187	1274,308	129,6219	129,6219	3
p23	LUKI SETIAWAN	1023,552	1244,219	99,06252	99,06252	3
p24	SUTIKNO	1391,743	876,3273	270,4347	270,4347	3
p27	ASMAR	1226,173	1043,189	112,7901	112,7901	3
p28	SUMARNO	724,9088	1543,309	397,9026	397,9026	3
p30	AMITA	1127,712	1140,227	17,31873	17,31873	3

**3.3 Implementasi Sistem Berbasis Web**

Sistem segmentasi pelanggan dalam penelitian ini diimplementasikan dalam bentuk aplikasi berbasis web menggunakan framework Laravel. Tujuan dari pengembangan sistem ini adalah untuk mempermudah pengguna dalam mengelola data pelanggan, melakukan proses clustering dengan algoritma K-Means, serta menampilkan hasil pengelompokan secara interaktif dan informatif.

a. Halaman Login

Halaman login digunakan sebagai pintu akses utama untuk masuk ke dalam sistem. Pengguna harus memasukkan email dan kata sandi yang valid untuk dapat mengakses seluruh fitur sistem.



Gambar 2. Login

b. Halaman Register

Halaman registrasi digunakan untuk membuat akun baru bagi pengguna yang belum memiliki akses.



Gambar 3. Register

c. Halaman Lupa Password

Tampilan lupa *password* dirancang untuk mereset *password* dengan menggunakan email yang terkait.



Gambar 5. Lupa Password

d. Halaman Dashboard

Halaman Dashboard menyajikan ringkasan utama dari sistem segmentasi pelanggan. Berikut ini adalah tampilan dari halaman Dashboard.



Gambar 6. Dashboard

e. Halaman Data Pelanggan

Halaman ini menampilkan seluruh data pelanggan yang dimiliki showroom. Pengguna dapat menambah, mengedit, atau menghapus data pelanggan sesuai kebutuhan sebelum dilakukan proses clustering.



Gambar 7. Data Pelanggan

f. Halaman Data Centroid

Halaman ini menampilkan nilai centroid awal yang digunakan dalam proses clustering. Nilai-nilai ini dapat ditentukan secara manual oleh pengguna atau di-generate secara otomatis oleh sistem.



Gambar 8. Data Centroid

g. Halaman Proses K-Means

Halaman ini merupakan inti dari sistem, di mana pengguna dapat menjalankan algoritma K-Means untuk mengelompokkan data pelanggan. Proses ini akan menghasilkan pembagian data ke dalam tiga cluster berdasarkan kemiripan atribut.



Gambar 9. Proses K-Means

h. Halaman Hasil Proses K-Means

Setelah proses clustering selesai, hasil pengelompokan akan ditampilkan dalam bentuk tabel maupun grafik. Halaman ini memudahkan pengguna dalam menganalisis hasil segmentasi dan memahami karakteristik setiap kelompok pelanggan.



Gambar 10. Hasil Proses K-Means

### 3. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, metode K-Means Clustering terbukti efektif dalam melakukan segmentasi pelanggan di Showroom Pandu Motor. Proses pengelompokan dilakukan secara sistematis melalui tahapan seleksi data, normalisasi, perhitungan jarak Euclidean, hingga iterasi pembaruan centroid hingga mencapai titik konvergen. Segmentasi pelanggan terbagi ke dalam tiga cluster, yaitu pelanggan dengan potensi tinggi, sedang, dan rendah. Hasil segmentasi ini memberikan pemahaman yang lebih mendalam mengenai karakteristik pelanggan dan dapat digunakan sebagai dasar dalam merumuskan strategi pemasaran yang lebih tepat sasaran. Dengan pendekatan ini, showroom dapat mengoptimalkan efektivitas penawaran produk dan meningkatkan peluang penjualan secara signifikan. Oleh karena itu, penerapan algoritma K-Means memiliki potensi besar untuk mendukung pengambilan keputusan berbasis data di sektor penjualan kendaraan bermotor maupun bidang usaha lainnya yang memerlukan segmentasi konsumen.

### UCAPAN TERIMAKASIH

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat-Nya sehingga jurnal ini dapat diselesaikan dengan baik. Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Bapak Ahmad Fitri Boy, S.Kom., M.Kom. selaku dosen pembimbing I dan Bapak Suardi Yakub, S.E., S.Kom., M.M. selaku dosen pembimbing II, yang telah memberikan bimbingan, arahan, serta masukan yang sangat berharga selama proses penyusunan penelitian hingga penulisan jurnal ini. Penulis juga menyampaikan apresiasi kepada semua pihak yang telah membantu, baik secara langsung maupun tidak langsung, dalam mendukung kelancaran penelitian ini. Ucapan terima kasih khusus disampaikan kepada keluarga, sahabat, serta rekan-rekan seperjuangan yang selalu memberikan semangat, doa, dan motivasi yang tidak pernah surut hingga jurnal ini dapat terselesaikan dengan sebaik-baiknya.

### DAFTAR PUSTAKA

- [1] A. A. SHELEMO, "DATA MINING UNTUK SEGMENTASI PELANGGAN DENGAN ALGORITMA K-MEANS STUDI KASUS PADA DATA PELANGGAN DI TOKO RETAIL," *Nucl. Phys.*, vol. 13, no. 1, pp. 104–116, 2023.
- [2] D. A. N. Disiplin, K. Pada, and M. M. Karanganyar, "KINERJA KARYAWAN DITINJAU DARI KEPEMIMPINAN, LINGKUNGAN KERJA PADA SHOWROOM MUHARI MOTOR 651 KARANGANYAR," vol. 04, no. 01, pp. 338–346, 2020.
- [3] R. Riadi and Mesran, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS UNTUK ANALISA PENJUALAN PARFUME," *J. Informatics, Electr. Electron. Eng.*, vol. 2, no. 4, pp. 138–145, 2023, doi: 10.47065/jieee.v2i4.1181.
- [4] N. H. Nabila, A. F. Isnawati, and M. A. Afandi, "ANALISIS PENGGUNAAN TEKNIK K-MEANS CLUSTERING SEBAGAI PENGANTI FUNGSI DEMAPPER PADA SISTEM KOMUNIKASI FBMC OQAM," *Transmisi*, vol. 23, no. 2, pp. 48–57, 2021, doi: 10.14710/transmisi.23.2.48-57.
- [5] N. Hendrastuty, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING DALAM EVALUASI HASIL PEMBELAJARAN SISWA," *J. Ilm. Inform. Dan Ilmu Komput.*, vol. 3, no. 1, pp. 46–56, 2024, [Online]. Available: <https://doi.org/10.58602/jima-ilkom.v3i1.26>
- [6] P. H. Suharti, A. S. Suryandari, and R. N. Amalia, "ANALISIS KINERJA MODUL PENGENDALI TEKANAN UDARA PCT-14 BERBASIS PLC DENGAN BERBAGAI METODA TUNING," *Sebatik*, vol. 26, no. 2, pp. 420–427, 2022, doi: 10.46984/sebatik.v26i2.2134.

- [7] Feronika Dalla and Usman Usman, “ANALISIS STRATEGI PEMASARAN DALAM MENINGKATKAN PENJUALAN OLSHOP PAKAIAN,” *J. Inov. Manajemen, Kewirausahaan, Bisnis dan Digit.*, vol. 1, no. 3, pp. 112–120, 2024, doi: 10.61132/jimakebidi.v1i3.236.
- [8] D. S. Indah and I. Komputer, “PEMANFAATAN TEKNIK CLUSTERING DALAM DATA MINING UNTUK SEGMENTASI PELANGGAN,” vol. 1, no. 6, pp. 1–18, 2024.
- [9] R. Hidayat, “PEMANFAATAN DATA MINING UNTUK MELIHAT MINAT SISWA SETELAH MENYELESAIKAN PENDIDIKAN SEKOLAH MENENGAH ATAS (SMA) DENGAN ALGORITMA K-MEANS CLUSTERING,” *Technol. Informatics Insight J.*, vol. Volume 1, no. 2, 2022, [Online]. Available: <https://jurnal.universitaspurabangsa.ac.id/index.php/tiij>
- [10] C. Zai, “IMPLEMENTASI DATA MINING SEBAGAI PENGOLAHAN DATA,” *J. Portal Data*, vol. 2, no. 3, pp. 1–12, 2022, [Online]. Available: <http://portaldata.org/index.php/portaldata/article/view/107>
- [11] S. Asyuti and A. A. Setyawan, “DATA MINING DALAM PENGGUNAAN PRESENSI KARYAWAN DENGAN CLUSTER MEANS,” *J. Ilm. Sains Teknol. Dan Inf.*, vol. 1, no. 1, pp. 01–10, 2023, [Online]. Available: <https://jurnal.alimspublishing.co.id/index.php/JITI/article/download/6/6>
- [12] M.Hafizh, “PENERAPAN DATA MINING ALGORITMA ASSOCIATION RULE METODE FP-GROWTH UNTUK MENGANALISA TINGKAT KEKERASAN DALAM RUMAH TANGGA,” *Maj. Ilm. UPI YPTK*, vol. 25, no. 1, pp. 99–106, 2018, doi: 10.35134/jmi.v25i1.36.