

## Implementasi Data Mining dalam Pengelompokan Data Pembelian Menggunakan Algoritma *K-Means* Pada PT.Otomotif 1

**Denis Dwi Susilo<sup>1\*</sup>, Shofa Shofiah Hilabi<sup>2</sup>, Bayu Priyatna<sup>3</sup>, Elfina Novalia<sup>4</sup>**  
 Sistem Informasi, Universitas Buana Perjuangan, Karawang, Indonesia  
 e-mail *Corresponding Author*: si20.denissusilo@mhs.ubpkarawang.ac.id

### Abstract

*This research is based on the problems of PT. AUTOMOTIVE 1 which is having difficulty monitoring products that consumers are interested in. The research aims to group buyer data using the K-Means algorithm. This research uses transaction data at PT. AUTOMOTIVE 1 during January 2016, with a total of 30 valid transaction records. The analysis process involves using the K-Means algorithm to group transaction data into three clusters, namely: the most popular cluster, the least popular cluster, and the least popular cluster. This research provides valuable insight for PT. AUTOMOTIVE 1 in developing their marketing strategy and inventory management. Applying the clustering method using the K-Means algorithm can help companies optimize their marketing strategy and inventory management.*

**Keywords:** *Clustering Analysis; K-Means; Purchase; Marketing strategy; Transaction Data*

### Abstrak

Penelitian ini didasarkan pada permasalahan PT. OTOMOTIF 1 yang mengalami kesulitan dalam memantau produk yang diminati konsumen. Penelitian bertujuan untuk mengelompokkan data pembelian menggunakan algoritma *K-Means*. Penelitian ini menggunakan data transaksi pada PT. OTOMOTIF 1 selama Januari 2016, dengan total 30 rekaman transaksi yang valid. Proses analisis melibatkan penggunaan algoritma *K-Means* untuk mengelompokkan data transaksi menjadi tiga klaster, yaitu: klaster yang paling diminati, klaster yang sedikit diminati, dan klaster yang kurang diminati. Penelitian ini memberikan wawasan yang berharga bagi PT. OTOMOTIF 1 dalam mengembangkan strategi pemasaran dan manajemen persediaan barang mereka. Penerapan metode *clustering* dengan menggunakan algoritma *K-Means* dapat membantu perusahaan mengoptimalkan strategi pemasaran dan manajemen persediaan barang mereka.

**Kata kunci:** *Analisis Clustering; K-Means; Pembelian; Strategi Pemasaran; Data Transaksi*

### 1. Pendahuluan

Kebutuhan akan transportasi semakin tinggi di masyarakat, yang mengakibatkan orang memiliki kendaraan pribadi untuk kegiatan sehari-hari. Semua pengendara pasti ingin motor mereka beroperasi dengan optimal, sehingga perusahaan yang menjual suku cadang dan layanan perawatan motor perlu meningkatkan efisiensi dan produktivitas dalam usaha mereka. Dengan semakin intensifnya persaingan di dunia bisnis, terutama di sektor penjualan suku cadang motor dan layanan servis, para pengembang harus mencari strategi baru guna mendongkrak promosi dan penjualan produk perusahaannya [1]. Data transaksi dapat kita manfaatkan sebagai solusinya. Salah satu permasalahan eksternal yang dihadapi dunia usaha adalah penggunaan teknologi informasi di pasar yang sangat kompetitif [2].

PT. OTOMOTIF 1 adalah perusahaan di sektor otomotif yang spesialis dalam penjualan sparepart kendaraan bermotor, layanan *service* untuk berbagai merek motor, dan pembelian produk otomotif. Namun, perusahaan ini memiliki permasalahan dalam memantau jenis produk yang diminati oleh konsumen [3]. Oleh karena itu, dibutuhkan sistem yang dapat mendukung pengambilan keputusan yang cepat dan akurat. Solusi yang diusulkan untuk mengatasi masalah tersebut adalah melalui penerapan analisis *clustering* untuk melakukan segmentasi pelanggan dengan menggunakan Algoritma *K-Means* [4].

Membagikan kelompok barang-barang berdasarkan karakteristik khusus merupakan tujuan dari teknik *clustering* [5]. Analisis *cluster* adalah metode multivariat yang berupaya mengorganisasikan item berdasarkan atributnya, dengan tujuan menempatkan objek dalam *cluster* yang sama jika objek tersebut sangat mirip satu sama lain. Penemuan pengetahuan dalam database (KDD) [6], sebagai upaya untuk mengungkap pola atau hubungan berdasarkan data historis [7].



Gambar 1. Proses KDD

Salah satu metode yang banyak digunakan dalam *clustering* adalah algoritma pengelompokan berbasis *K-Means* [8][9]. Metode ini digunakan untuk membantu perusahaan mengidentifikasi barang terlaris dan menghindari penumpukan barang di gudang, memberikan solusi nyata untuk pengelolaan stok. Tujuan utama pengelompokan adalah untuk mengkategorikan semua data ke dalam sub kelompok atau klaster yang berbeda yang relatif seragam di mana kesamaan antar entri dalam kelompok ditingkatkan, sementara kesamaan dengan entri di luar kelompok diminimalkan [10]. Algoritma *K-Means* populer digunakan karena kemudahannya diimplementasikan dan kecepatannya dalam mengelompokkan data besar [11][12].

Penelitian ini bertujuan untuk mengelompokkan data transaksi penjualan atau pembelian *sparepart* kendaraan bermotor, guna mempermudah pihak manajemen PT. OTOMOTIF 1 dalam melakukan perencanaan strategi bisnis perusahaan.

## 2. Tinjauan Pustaka

Penelitian yang dilakukan oleh Niko Suwaryo, Arif Rahman, Dewi Marini Umi Atmaja, dan Amat Basri mengenai *Clustering* Stok Produk Ritel untuk Menentukan Pergerakan Kebutuhan Konsumen Menggunakan Algoritma *K-Means* menunjukkan bahwa menggunakan metode *clustering* pada data produk ritel menghasilkan wawasan dan informasi yang berharga. Dengan menggunakan Data Persyaratan Awal dan Data Persyaratan Akhir serta menerapkan algoritma *K-means*, produk dapat dikelompokkan berdasarkan ketersediaan dan inventaris di gudang. Hasil uji coba menunjukkan terbentuknya tiga klaster: Klaster 1 berisi produk dengan ketersediaan rendah (*Low*), Klaster 2 berisi produk dengan ketersediaan sedang (*Medium*), dan Klaster 3 berisi produk dengan ketersediaan tinggi (*High*). Validasi menggunakan validasi silang *k-fold* semakin memperkuat temuan ini. Uji coba dengan *Rapid Miner* juga memberikan hasil yang serupa, dengan setiap klaster berisi anggota yang sejalan dengan perhitungan manual. Penemuan ini memberikan pemahaman yang lebih mendalam tentang strategi pengelolaan inventaris produk ritel dan berkontribusi secara signifikan terhadap perkembangan sistem manajemen inventaris di sektor ritel [13].

Penelitian yang dilakukan oleh Moh. Azlan Shah Putra, Julhandri, Ihsan Khoir, dan Siti Monalisa merupakan eksplorasi terbaru dalam menggunakan algoritma *Fuzzy C-Means* bersama dengan model RFM untuk mengelompokkan pelanggan di Feandra Cake Shop. Hasil penelitian menunjukkan adanya segmen pelanggan dengan nilai Seumur Hidup Pelanggan (*Customer Lifetime Value/CLV*) tertinggi, yang terdapat dalam cluster 1. Pelanggan dalam klaster ini menunjukkan tingkat loyalitas yang tinggi, ditandai dengan durasi belanja yang singkat, frekuensi pembelian yang sering, dan jumlah transaksi yang signifikan, menunjukkan preferensi untuk transaksi dengan nilai tinggi dengan perusahaan. Temuan ini menyoroti profitabilitas yang tinggi dari segmen-segmen ini, menekankan perlunya strategi khusus untuk menjaga mereka dari pesaing. Selain itu, analisis pola pembelian produk mengungkap berbagai

dinamika penjualan silang di setiap klaster, dengan produk tertentu seperti kue durian dan talam memperoleh popularitas di klaster yang berbeda. Meskipun demikian, produk unggulan perusahaan seperti Suke Ori dan Suke Rasa tidak termasuk dalam rekomendasi produk karena sudah diperoleh oleh pelanggan lama. Oleh karena itu, penelitian ini memberikan wawasan yang berharga untuk meningkatkan manajemen hubungan pelanggan dan merancang strategi pemasaran yang efektif dalam industri kue [14].

Penelitian yang dilakukan oleh Arief Wibowo dan Andy Rio Handoko mengenai Segmentasi Pelanggan Ritel Produk Farmasi Menggunakan Metode *Clustering* Data Mining dengan Analisis *Modified Recency Frekuensi Moneter* (RFM) menunjukkan bahwa inklusi atribut *Quantity* dalam analisis data pelanggan menghasilkan *clustering* yang lebih optimal, yang dikenal sebagai model QRFM (Kuantitas, Kekinian, Frekuensi, dan Moneter). *Cluster* optimal diidentifikasi setelah enam kali percobaan yang melibatkan dua segmen pelanggan, dengan nilai *Davies Bouldin Index* (DBI) sebesar 0,527 untuk atribut QRF yang dianalisis menggunakan algoritma *K-Means*. Ini menunjukkan bahwa atribut *Quantity* memiliki kontribusi yang signifikan dalam pembentukan *cluster* yang lebih baik daripada model RFM asli. Penelitian ini membuka peluang untuk penelitian lebih lanjut dengan mengeksplorasi segmentasi pelanggan berdasarkan metode pembelian, termasuk transaksi *online* dan *offline*. Analisis juga dapat diperluas untuk mengevaluasi pembelian ulang dari jenis produk yang sama pada transaksi berikutnya. Temuan ini menjanjikan potensi untuk strategi pemasaran baru, terutama mengingat peningkatan penetrasi penjualan ritel *online* di Indonesia dalam dekade terakhir, yang diperkirakan akan memengaruhi perilaku dan metodologi pembelian konsumen di masa mendatang [15].

*State of the art* pada artikel ini adalah pengimplementasian data mining pada pengelompokan data pembeli menggunakan algoritma *K-means* pada PT. Otomotif 1 Dalam penelitian ini, tiga klaster dalam perusahaan telah diidentifikasi dengan menggunakan berbagai metode termasuk *RapidMiner* dan penghitungan manual. Klaster tersebut terdiri dari konsumen dengan pola pengeluaran yang berbeda.

### 3. Metodologi

Metode Penelitian ini menggunakan metode yang terstruktur, mendasarkan analisis pada data transaksi dari PT. OTOMOTIF 1 selama Januari 2016. Dari dataset yang dikumpulkan, teridentifikasi 30 rekaman transaksi yang valid dan telah diolah sesuai dengan standar untuk studi tersebut. Data yang akan dianalisis ditetapkan terlebih dahulu. Prosesnya melibatkan pencarian data yang tersedia, pengumpulan tambahan data yang dibutuhkan, dan penggabungan semua informasi ke dalam dataset yang mencakup variabel yang relevan untuk prosedur penelitian ini.

#### 3.1 Algoritma *K-Means Clustering*

*Clustering* adalah proses Mengategorikan data ke dalam beberapa kelompok berdasarkan kesamaan karakteristiknya. Selanjutnya, pada tahap ini, dilakukan proses pengelompokan data Dengan menetapkan bobot-bobot nilai klaster yang kemudian akan dianalisis secara komprehensif untuk mencapai hasil yang optimal.

$$C_i = \frac{1}{M} \sum_{j=1}^M X_j. \quad \dots\dots\dots (1)$$

Keterangan :

- $C_i$  = simbol yang mewakili nilai rata-rata dari sebuah set data.
- $\frac{1}{M}$  = kebalikan dari jumlah total elemen dalam set data dan digunakan untuk menghitung nilai rata-rata.
- $\sum$  = Sigma.
- $M$  = jumlah total elemen dalam set data.
- $J$  = indeks yang digunakan untuk menunjukkan elemen ke-  $J$  dalam set data, yang berkisar dari 1 hingga  $M$ .
- $X_j$  = nilai dari elemen ke-  $J$  dalam set data.

Kemudian Jarak dalam ruang *Euclidean* dapat dihitung dengan menggunakan persamaan untuk mengukur jarak.

$$d = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} \dots\dots\dots (2)$$

Keterangan :

d = jarak antara dua titik dalam ruang dua dimensi.

$\sqrt{\quad}$  = akar kuadrat.

$(X_1 - X_2)^2$  = mewakili kuadrat dari selisih antara koordinat X titik pertama dan kedua.

$(Y_1 - Y_2)^2$  = mewakili kuadrat dari selisih antara koordinat Y titik pertama dan kedua.

### 3.2 Sumber Data

Metode Penelitian ini mengadopsi pendekatan terstruktur, mengandalkan analisis pada data transaksi yang berasal dari PT. OTOMOTIF 1 selama bulan Januari 2016. Dari dataset yang terhimpun, 30 catatan transaksi yang valid telah diidentifikasi dan diproses sesuai dengan standar yang diperlukan untuk penelitian ini. Kumpulan data ini mencakup tiga atribut utama, yaitu kode barang, jumlah, dan nilai pemasukan. Sebagai contoh, kami memiliki data sebagai berikut:

Tabel 1. Data Set Penelitian

No	Kode Barang	Jumlah	Nilai Pemasukan
1.	B6H0A	152	1,903,000
2.	B6H0A	212	13,630,500
3.	B6H0A	272	12,574,200
4.	B6H0A	192	11,486,000
5.	B6H0A	92	1,831,500
6.	B6H0A	182	11,516,000
7.	B5D	262	11,549,600
8.	BK6-01	252	11,135,000
...	...	...	...
30.	BDR0A	292	12,250,700

Kemudian data yang berasal dari PT. OTOMOTIF 1 diproses menggunakan metode clustering, sehingga menghasilkan data yang memiliki nilai dan kegunaan dalam konteks penelitian. *Clustering* adalah proses Mengategorikan data ke dalam beberapa kelompok berdasarkan kesamaan karakteristiknya. Selanjutnya, pada tahap ini, dilakukan proses pengelompokan data Dengan menetapkan bobot-bobot nilai klaster yang kemudian akan dianalisis secara komprehensif untuk mencapai hasil yang optimal.

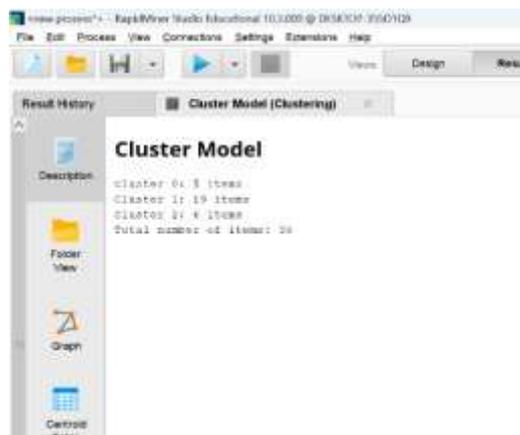
## 4. Hasil dan Pembahasan

### 4.1 Pengolahan Data Dengan *RapidMiner*

Setelah data dimasukkan oleh peneliti dan divalidasi, peneliti melanjutkan untuk memodelkan dataset. Awalnya, data diimpor ke *RapidMiner*, dan kemudian proses dimulai menggunakan metode *K-means* dengan parameter diatur ke  $k = 3$  dan jenis pengukuran didefinisikan sebagai ukuran campuran. Metodologi ini memiliki arti penting karena kumpulan data berisi campuran kode dan nilai numerik, sehingga memerlukan penggunaan Jarak *Euclidean* untuk mengukur jarak antar titik *cluster*.

Gambar 2. Parameter *Cluster*

Kemudian, Data diproses di *RapidMiner*, menghasilkan terciptanya tiga model *cluster*. *Cluster* ini dilambangkan dengan *cluster 0* yang terdiri dari 5 item, *cluster 1* dengan 19 item, dan *cluster 2* dengan 6 item. Perlu dicatat bahwa penomoran *cluster* dimulai dari 0, konsisten dengan konvensi urutan penomoran dalam bahasa pemrograman.

Gambar 3. *Cluster* model

*RapidMiner* menawarkan opsi hasil dalam bentuk struktur pohon, yang menggambarkan besarnya perbandingan antar *cluster* di seluruh kumpulan data. Fitur ini memungkinkan visualisasi tentang bagaimana kumpulan data asli, atau kumpulan akar, dibagi menjadi beberapa *Cluster* anak.

Gambar 4. Perbandingan model *tree*

Tabel 2. Hasil *RapidMiner*

Model <i>cluster</i>	Jumlah
<i>Cluster 0</i>	5
<i>Cluster 1</i>	19
<i>Cluster 2</i>	6
Total <i>cluster</i>	30

#### 4.2 Penghitungan Manual

Kemudian untuk mengkalibrasi hasil dari *RapidMiner* pengujian ini akan dibandingkan dengan cara lain untuk mengevaluasi performa dan validitasnya. Berikut merupakan penamaan awal pada masing-masing data *cluster*.

Tabel 3. Penamaan awal data

Data ke-	<i>Cluster</i> ke	Frek_jual	Total bayar (Rp)
1	$C_1$	152	1,903,000
2	$C_2$	212	13,630,500
3	$C_3$	272	12,574,200
4	$C_1$	192	11,486,000
5	$C_2$	92	1,831,500
6	$C_3$	182	11,516,000
7	$C_1$	262	11,549,600
8	...	...	...
30	$C_3$	292	12,250,700

Setelah menyelesaikan perhitungan awal untuk setiap *cluster*, selanjutnya, kita perlu menghitung jarak antara setiap titik data dan setiap *cluster*.

1. *Centeroid* awal pada jumlah barang

$$C_0 = \frac{\text{data1} + \text{data4} + \text{data7} + \text{data10} + \dots + \text{data28}}{100}$$

$$= \frac{152 + 192 + 262 + 242 + 222 + 172 + 112 + 182 + 252 + 392}{100} = 22$$

$$C_1 = \frac{\text{data2} + \text{data5} + \text{data8} + \text{data11} + \dots + \text{data29}}{100}$$

$$= \frac{212 + 92 + 252 + 142 + 272 + 302 + 112 + 262 + 132 + 262}{100} = 21$$

$$C_2 = \frac{\text{data3} + \text{data6} + \text{data9} + \text{data12} + \dots + \text{data30}}{100}$$

$$= \frac{272 + 182 + 242 + 402 + 322 + 221 + 332 + 212 + 302 + 292}{100} = 28$$

2. *Centeroid* awal pada total bayar

$$C_1 = \frac{\text{data1} + \text{data4} + \text{data7} + \text{data10} + \dots + \text{data28}}{100}$$

$$= \frac{1.903.000 + 11.486.000 + \dots + 12.460.000}{100}$$

$$= 2.258.480$$

$$C_2 = \frac{\text{data2}+\text{data5}+\text{data8}+\text{data11}+\dots+\text{data29}}{100}$$

$$= \frac{13.630.500+1.831.500+\dots+1.821.000}{100}$$

$$= 1.603.330$$

$$C_3 = \frac{\text{data3}+\text{data6}+\text{data9}+\text{data12}+\dots+\text{data30}}{10}$$

$$= \frac{12.574.200+11.516.000+\dots+12.250.700}{100}$$

$$= 2.776.180$$

Langkah berikutnya melibatkan melakukan perhitungan guna menemukan jarak antara Rumus jarak *Euclidean* digunakan untuk membangun setiap titik data awal dan pusat massa. Hasil perhitungan jarak ini akan mempengaruhi bagaimana setiap data ditempatkan ke dalam kluster tertentu.

Jarak antara data awal dan pusat massa *cluster* awal dihitung:

$$d = \sqrt{(22 - 15)^2 + (2.258.480 - 1.903.000)^2}$$

$$= 1.355.480$$

Jarak antara data awal dan pusat massa *cluster* kedua dihitung:

$$d = \sqrt{(21 - 15)^2 + (1.603.480 - 1.903.000)^2}$$

$$= 700.330$$

Jarak antara data awal dan pusat massa *cluster* ketiga dihitung:

$$d = \sqrt{(28 - 15)^2 + (2.776.180 - 1.903.000)^2}$$

$$= 1.873.180$$

Setelah menyelesaikan perhitungan awal untuk setiap *cluster*, selanjutnya, kita perlu menghitung jarak antara setiap titik data dan setiap *cluster*.

Tabel 4. Perhitungan pada hasil jarak data

Data ke-	total	Total bayar (Rp)	C1	C2	C3
1	152	1,903,000	1.455.480	800.330	1.973.180
2	212	13,630,500	1.472.020	2.127.170	954.320
3	272	12,574,200	415.720	980.870	301.980
4	192	11,486,000	872.480	217.330	1.390.180
5	92	1,831,500	1.526.980	871.830	1.954.680
6	182	11,516,000	842.480	97.330	1.360.180
7	262	11,549,600	808.880	63.730	1.326.580
8	...	...	...	...	...
30	292	12,250,700	8.780	657.370	535.480

Kemudian dari hasil perhitungan pada tabel penulis dapat mengetahui pusat cluster yaitu:

1. Untuk Klaster awal, terdapat **lima(5)** data yang terdiri dari nomor 3, 14, 17, 28, dan 30.

No	Klaster awal
1.	3
2.	14
3.	17
4.	28
5.	30

2. Untuk klaster kedua, terdapat **sembilan belas(19)** data yang terdiri dari nomor 1, 4, 5, 6, 7, 8, 10, 11, 12, 13, 18, 19, 20, 23, 24, 25, 26, 27, dan 29.

No	Klaster kedua
1.	1
2.	4
3.	5
4.	6
5.	7
6.	8
7.	10
8.	11
9.	12
10.	13
11.	18
12.	19
13.	20
14.	23
15.	24
16.	25
17.	26
18.	27
19.	29

3. Untuk Klaster ketiga, terdapat **enam(6)** data yang terdiri dari nomor 2, 9, 15, 16, 21, dan 22.

No	Klaster ketiga
1.	2
2.	9
3.	15
4.	16
5.	21
6.	22

#### 4.3 Pelabelan Hasil Klaster

Menginterpretasikan hasil pelabelan klaster sangat krusial untuk memberikan pemahaman yang mendalam tentang signifikansi dari setiap klaster yang terbentuk. Di bawah ini merupakan contoh analisis yang dapat dilakukan terkait dengan hasil pelabelan *cluster*.

##### 1) Klaster C1 (Paling Diminati)

###### a. Ciri – Ciri Anggota Klaster C1

Untuk klaster awal atau C1, terdapat **lima(5)** data yang terdiri dari nomor 3, 14, 17, 28, dan 30.

###### b. Alasan Pemberian Label

Barang-barang ini menjadi favorit pelanggan dengan tingkat pembelian atau permintaan yang jauh lebih tinggi dibandingkan dengan klaster lainnya, berdasarkan data penjualan yang ada.

## 2) Klaster C2 (Sedikit Diminati)

### a. Ciri – Ciri Anggota Klaster C2

Untuk klaster kedua, terdapat **sembilan belas(19)** data yang terdiri dari nomor 1, 4, 5, 6, 7, 8, 10, 11, 12, 13, 18, 19, 20, 23, 24, 25, 26, 27, dan 29.

### b. Alasan Pemberian Label

Barang-barang ini terus memperoleh penjualan, meskipun permintaannya cenderung lebih rendah dibandingkan dengan klaster lainnya. Label "Sedikit Diminati" menunjukkan bahwa meskipun tidak secara keseluruhan diminati, permintaan terhadap barang-barang tersebut memang relatif lebih rendah dibandingkan dengan barang-barang dalam kelompok lain.

## 3) Klaster C3 (Kurang Diminati)

### a. Ciri – Ciri Anggota Klaster C3

Untuk klaster ketiga, terdapat **enam(6)** data yang terdiri dari nomor 2, 9, 15, 16, 21, dan 22.

### b. Alasan Pemberian Label

Label Kurang diminati mengindikasikan bahwa barang tersebut mungkin kurang menarik bagi pelanggan atau memiliki permintaan yang terbatas. Evaluasi ulang terhadap keberadaan item dalam klaster ini mungkin diperlukan, termasuk pertimbangan untuk mengurangi persediaan atau mengalokasikan kembali sumber daya ke produk yang lebih diminati.

## 4.4 Keterbatasan Penelitian

Keterbatasan dari penelitian ini adalah bahwa hanya menggunakan data transaksi dari PT. OTOMOTIF 1 selama Januari 2016 sebagai basis analisis. Oleh karena itu, saran untuk penelitian lebih lanjut adalah melakukan penelitian dengan rentang waktu yang lebih luas, mencakup beberapa bulan atau bahkan beberapa tahun, untuk mendapatkan pemahaman yang lebih menyeluruh tentang tren penjualan dan preferensi konsumen dari waktu ke waktu. Selain itu, penelitian lebih lanjut dapat memperluas cakupan data dengan memperhitungkan variabel tambahan seperti lokasi geografis pelanggan, kondisi ekonomi, dan faktor-faktor lain yang dapat memengaruhi permintaan produk otomotif. Dengan demikian, analisis yang lebih komprehensif dapat dilakukan untuk memberikan wawasan yang lebih mendalam bagi perusahaan dalam mengoptimalkan strategi penjualan dan persediaan barang mereka.

## 5. Simpulan

Setelah penelitian yang dilakukan, penulis memperoleh hasil dengan menggunakan berbagai metode untuk memastikan *cluster* dalam perusahaan. Dengan menggunakan RapidMiner dan penghitungan manual penulis tetap mendapatkan hasil yang sama yaitu *cluster* 1 yang berjumlah 5 data, *cluster* 2 yang berjumlah 19 data, dan *cluster* 3 yang berjumlah 6 data. Dari penelitian ini, dapat disimpulkan bahwa dalam klaster 1 terdapat konsumen yang mengeluarkan kurang dari Rp. 300.000, sedangkan klaster 2 mencakup konsumen dengan pengeluaran antara Rp. 300.000 hingga Rp. 700.000. Sementara pada klaster 3 terdapat konsumen yang mengeluarkan lebih dari Rp. 1.000.000.

Penelitian ini bertujuan untuk memberikan wawasan berharga bagi PT. OTOMOTIF 1 dalam mengembangkan strategi pemasarannya. Saran yang ingin disampaikan oleh peneliti adalah untuk menyesuaikan kembali harga dan keuntungan yang diperoleh dari transaksi yang termasuk dalam klaster 1. Selain itu, peneliti mengusulkan pengembangan layanan online sepanjang waktu melalui *platform* seperti *website* atau bot media sosial, seperti *bot WhatsApp* yang dilengkapi dengan kemampuan transaksi otomatis.

## Daftar Referensi

- [1] S. S. Hilabi, "Rancang Bangun Sistem Inventory Usaha (UMKM) 'Karpel' Desa Kamurang Berbasis Web," *Pros. Konf. Nas. Penelit. dan Pengabd.*, vol. 2 no. 1, pp. 1147–1155, 2022.
- [2] Y. Suhandi, I. Kurniati, and S. Norma, "Penerapan Metode Crisp-DM Dengan Algoritma K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik," *J. Teknol. Inform. dan Komput.*, vol. 6, no. 2, pp. 12–20, 2020, doi: 10.37012/jtik.v6i2.299.
- [3] I. Melani, B. Priyatna, F. Nurapriani, and S. S. Hilabi, "Implementasi Metode K-Means

- Clustering Pada Penilaian Kinerja Karyawan PT Kopetri Citra Abadi," *J. Inf. Interaktif*, vol. 8, no. 1, pp. 24–30, 2023, [Online]. Available: <http://e-journal.janabadra.ac.id/>
- [4] T. Widyanti, S. S. Hilabi, A. Hananto, Tukino, and E. Novalia, "Implementasi K-Means dan K-Nearest Neighbors pada Kategori Siswa Berprestasi," *J. Inf. dan Teknol.*, vol. 5, no. 1, pp. 75–82, 2023, doi: 10.37034/jidt.v5i1.255.
- [5] D. Bahtiar *et al.*, "Pemetaan Penduduk Penerima Bantuan Sosial Desa Waru Jaya Menggunakan Algoritma K-Means Clustering," *Sci. Sacra J. Sains*, vol. 3, no. 2, pp. 29–39, 2023, [Online]. Available: <http://pijarpemikiran.com/index.php/Scientia>
- [6] Y. Prihati, Suwarno, and A. Dharmawan, "Implementasi Algoritma K-Means Untuk Pemetaan Prestasi Akademik Siswa Disekolah Dasar Terang Bagi Bangsa Pati," *Kinabalu*, vol. 11, no. 2, pp. 50–57, 2019.
- [7] F. T. Kesuma, and S. P. Tamba, "Penerapan Data Mining Untuk Menentukan Penjualan Sparepart Toyota Dengan Metode K-Means Clustering," *J. Sist. Inf. dan Ilmu Komput. Prima (JUSIKOM PRIMA)*, vol. 2, no. 2, pp. 67–72, 2020, doi: 10.34012/jusikom.v2i2.376.
- [8] B. Susarianto, & T. Nizami, "Penentuan Penerima Bantuan Sosial Beras untuk Masyarakat Miskin Menggunakan Metode K-Means". *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 12, no. 2, pp. 772-783, 2023.
- [9] A. Wahyudi and R. Utami, "Penggunaan Metode Euclidean Distance Pada Aplikasi Pencarian Lokasi Rumah Sakit di Kota Medan," *Informatics Eng. Electron. Data*, vol. 1, no. 1, pp. 47–58, 2022, doi: 10.59840/ieed.v1i1.193.
- [10] W. Mega, "Clustering Menggunakan Metode K-Means Untuk Menentukan Status Gizi Balita," *J. Inform.*, vol. 15, no. 2, pp. 160–174, 2015.
- [11] T. P. Sari, A. L. Hananto, E. Novalia, T. Tukino, and S. S. Hilabi, "Implementasi Algoritma K-Means dalam Analisis Klasterisasi Penyebaran Penyakit Hiv/Aids," *Infotek J. Inform. dan Teknol.*, vol. 6, no. 1, pp. 104–114, 2023, doi: 10.29408/jit.v6i1.7423.
- [12] N. Afiasari, N. Suarna, and N. Rahaningsi, "Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering dengan Metode K-Means," *J. SAINTEKOM*, vol. 13, no. 1, pp. 100–110, 2023, doi: 10.33020/saintekom.v13i1.402.
- [13] N. S. Niko, A. Rahman, D.M.U Atmaja, and A. Basri, "Klasterisasi Stok Produk Retail Untuk Menentukan Pergerakan Kebutuhan Konsumen Dengan Algoritma K-Means," *Bull. Inf. Technol.*, vol. 4, no. 3, pp. 306–312, 2023, doi: 10.47065/bit.v4i3.736.
- [14] M. A. Shah Putra, S. Monalisa, J. Julhandri, and I. Khoiru, "Penerapan Algoritma Fuzzy C-Means Menggunakan Model Rfm Dalam Klasterisasi Pelanggan Pada Toko Kue Feandra Cake," *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 6, no. 1, pp. 64-69, 2020, doi: 10.24014/rmsi.v6i1.8646.
- [15] A. Wibowo and A. R. Handoko, "Segmentasi Pelanggan Ritel Produk Farmasi Obat Menggunakan Metode Data Mining Klasterisasi Dengan Analisis Recency Frequency Monetary (RFM) Termodifikasi," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 3, pp. 573-580, 2020, doi: 10.25126/jtiik.2020702925.