

# CLUSTERING DATA PENJUALAN TOKO HELM KARTINI MENGUNAKAN METODE K-MEANS

Nicko Kurniawan <sup>1)</sup> Teny Handhayani <sup>2)</sup> Novario Jaya Perdana <sup>3)</sup>

<sup>1) 2) 3)</sup> Teknik Informatika Universitas Tarumanagara

Jl. Letjen Letjen S. Parman No. 1, Jakarta, 11440, Indonesia

email : [nicko.535200065@stu.untar.ac.id](mailto:nicko.535200065@stu.untar.ac.id) <sup>1)</sup>, [tenyh@fti.untar.ac.id](mailto:tenyh@fti.untar.ac.id) <sup>2)</sup>, [novariojp@fti.untar.ac.id](mailto:novariojp@fti.untar.ac.id) <sup>3)</sup>

## ABSTRACT

*This study aims to design a K-Means clustering-based system that can help Kartini Helmet Store in grouping helmet sales data based on product category. With the increasing need for deeper data analysis to support business strategies, this store faces challenges in understanding sales patterns and customer preferences. This system utilizes the K-Means clustering algorithm implemented through the Python programming language to group sales data, facilitate trend analysis, and provide visualization of clustering results. The data used are store sales records in Excel format, which are then processed through stages starting from preprocessing to evaluating clustering results using the silhouette coefficient method to ensure clustering accuracy. The results of this study are expected to help stores in making decisions related to stock management, adjusting marketing strategies, and increasing customer satisfaction.*

## Key words

*Clustering, K-Means, Silhouette, Helmet*

## 1. Pendahuluan

Helm merupakan salah satu sektor dalam pasar otomotif, terutama di negara-negara dengan populasi pengguna sepeda motor yang tinggi seperti Indonesia. Penggunaan helm yang tepat dan berkualitas penting untuk mengurangi risiko cedera kepala pada pengendara sepeda motor. Penelitian menunjukkan bahwa helm dengan standar keamanan yang baik dapat mengurangi risiko cedera kepala hingga tujuh puluh persen [1]. Helm tidak hanya berfungsi sebagai alat pelindung keselamatan, tetapi juga menjadi bagian dari gaya hidup, dengan berbagai desain dan fitur yang terus berkembang mengikuti tren konsumen.

Toko helm Kartini adalah sebuah toko ritel yang berfokus pada penjualan berbagai jenis helm motor untuk memenuhi kebutuhan pengendara di daerah jalan RA Kartini, Kota Bekasi. Untuk mempertahankan posisinya sebagai salah satu toko helm di kawasan tersebut, toko ini menawarkan promosi seperti diskon musiman atau paket produk, yang menarik pelanggan baru dan

mempertahankan pelanggan yang sudah ada. Toko helm Kartini menghadapi tantangan dalam menganalisis data penjualan yang beragam untuk memahami tren penjualan, dan mengenali pola pembelian. Salah satu teknik yang dapat digunakan adalah K-Means Clustering. Strategi penjualan yang efektif di toko helm Kartini dapat berkontribusi terhadap peningkatan dan kepuasan pelanggan. Faktor-faktor seperti variasi produk, kualitas pelayanan, dan lokasi toko mempengaruhi keberhasilan penjualan helm [2].

Clustering adalah suatu proses pengelompokan data atau objek ke dalam suatu kelompok, sehingga setiap data dalam sebuah cluster memiliki keterikatan yang kuat satu sama lain dan memiliki keterikatan yang lemah terhadap data dari cluster lain [3].

K-Means Clustering merupakan salah satu metode data Clustering yang akan mempartisi data yang ada ke dalam bentuk cluster atau kelompok, sehingga data yang memiliki karakteristik yang sama akan dikelompokkan ke dalam clustering dapat digunakan sebagai dasar penentuan strategi pemasaran yang tepat untuk menentukan jumlah objek berdasarkan kriteria-kriteria yang telah ditentukan [4]. Metode algoritma K-Means digunakan karena cukup pengelompokannya cepat dan efektif untuk diimplementasikan terhadap data berjumlah besar. Langkah kerja K-Means adalah membagi data menjadi  $k$  kelompok yang saling berlebihan. Kemudian melalui partisi berulang, K-Means akan mengurangi jarak total antara setiap bagian data dan cluster-nya [5].

Machine learning merupakan cabang dari kecerdasan buatan (Artificial Intelligence) yang memungkinkan sistem untuk belajar secara otomatis dari data tanpa pemrograman eksplisit [6]. Penggunaan machine learning digunakan untuk menemukan pola atau wawasan dalam data tanpa perlu pemrograman manual. Dengan kata lain, sistem machine learning dapat mengembangkan kemampuannya sendiri untuk melakukan tugas-tugas tertentu berdasarkan pengalaman atau data yang diberikan kepadanya.

## 2. Metode Penelitian

### 2.1 Clustering

Clustering adalah tugas pengelompokan kumpulan objek sedemikian rupa sehingga objek berada dalam kelompok yang sama lebih mirip dengan satu sama lain dibandingkan dengan kelompok lain. Ini adalah yang utama dari tugas eksplorasi data mining, dan teknik umum untuk analisis data statistik yang digunakan di banyak bidang, termasuk pembelajaran mesin, pengenalan pola, analisis gambar, pengambilan informasi [7].

### 2.2 K-Means

K-Means clustering adalah salah satu metode paling populer untuk mempartisi data ke dalam kelompok-kelompok berbeda berdasarkan kesamaan fitur, yang banyak digunakan karena kesederhanaan dan efisiensinya [8].

Metode clustering yang bertujuan untuk membagi data menjadi sejumlah cluster K berdasarkan kesamaan antar data. clustering menggunakan metode K-Means dapat membantu toko helm mengelompokkan besar kecilnya data penjualan berdasarkan hasil pendapatan, dan frekuensi pembelian, merk helm, dan produk lainnya yang memungkinkan toko helm untuk mengidentifikasi produk yang laku terjual dan dapat mengembangkan strategi pemasaran dan mengelola stok secara lebih efisien.

Di bawah ini merupakan langkah-langkah dalam menggunakan algoritma clustering K-Means [9]:

1. Tentukan banyaknya cluster yang ingin dibentuk.
2. Tentukan nilai-nilai titik pusat cluster awal (centroid).
3. Hitung jarak setiap data input terhadap semua centroid dengan menggunakan rumus Euclidean Distance di bawah ini hingga mendapatkan hasil dengan nilai jarak paling dekat dengan centroid.

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2} \quad (2.1)$$

Keterangan:

$d$  = jarak  
 $j$  = jumlah atribut  
 $c$  = centroid  
 $x$  = data

4. Masukkan setiap data berdasarkan jarak terdekat dengan centroid (nilai terkecil).
5. Ubah nilai centroid dengan centroid baru yang diperoleh dari perhitungan rata-rata cluster dengan menggunakan rumus di bawah ini:

$$v_{iw} = \frac{1}{N_i} \left( \sum_{k=1}^{N_i} X_{kw} \right) \quad (2.2)$$

Keterangan:

$i, w$  = indeks dari cluster

$w$  = indeks dari variabel

$v_{iw}$  = centroid/ rata-rata cluster ke- $i$  untuk variabel ke- $w$

$x_{kw}$  = nilai data ke- $k$  yang ada di dalam cluster tersebut untuk variabel ke- $w$

$N_i$  = jumlah data yang menjadi anggota cluster ke- $i$

6. Ulang kembali langkah 3 sampai dengan Langkah 5 hingga setiap data tidak berpindah cluster lagi (konvergen).

Berikut data sampel untuk contoh perhitungan data dilihat pada Tabel 1.

Tabel 1. Data Sampel

Data	F1	F2
A	5	10
B	3	7
C	4	9
D	6	14
E	2	5

Keterangan:

F1 = Jumlah terjual produk 1

F2 = Jumlah terjual produk 2

Contoh perhitungan jarak titik ke titik pusat:

A (5, 10)

Jarak A ke C1

$$\sqrt{(5 - 5)^2 + (10 - 10)^2} = 0$$

Jarak A ke C2

$$\begin{aligned} \sqrt{(3 - 5)^2 + (7 - 10)^2} &= \sqrt{(-2)^2 + (-3)^2} = \sqrt{4 + 9} \\ &= \sqrt{13} = 3.61 \end{aligned}$$

Jarak A Lebih dekat ke C1.

B (3, 7)

Jarak B ke C1

$$\begin{aligned} \sqrt{(5 - 3)^2 + (10 - 7)^2} &= \sqrt{2^2 + 3^2} = \sqrt{4 + 9} = \sqrt{13} \\ &= 3.61 \end{aligned}$$

Jarak B ke C2

$$\sqrt{(3 - 3)^2 + (7 - 7)^2} = \sqrt{0 + 0} = 0$$

Jarak B Lebih dekat ke C2.

Perhitungan tersebut dilakukan sampai ke data E

Tabel 2. Hasil Clusteing

Data	Jarak ke C1	Jarak ke C2	Cluster
A	<b>0.00</b>	3.61	1
B	3.61	<b>0.00</b>	2
C	<b>1.41</b>	2.24	1
D	<b>4.12</b>	7.28	1
E	6.40	<b>2.00</b>	2

### 2.3 Silhouette

Algoritma *Silhouette* adalah metode yang banyak digunakan untuk mengevaluasi kualitas kluster yang dihasilkan oleh teknik *clustering*. Metode ini mengukur seberapa mirip suatu objek dengan kluster sendiri dibandingkan dengan kluster lain, sehingga dapat memberikan indikasi seberapa baik objek tersebut ditempatkan [10].

Berikut langkah-langkah dalam melakukan perhitungan *Silhouette* [11]:

1. Menghitung rata-rata jarak objek ke-*i* dengan semua objek yang berada di dalam satu kelompok *A* dengan persamaan 2.3.

$$a(i) = \frac{1}{|A| - 1} \sum_{j \in A, j \neq i} d(i, j), \quad (2.3)$$

dengan *j* adalah objek lain dalam satu kelompok *A* dan  $d(i, j)$  adalah jarak antara objek *i* dan *j*.

2. Menghitung rata-rata jarak objek ke-*i* dengan semua objek yang berada pada kelompok lain dengan persamaan 2.4.

$$d(i, C) = \frac{1}{|C|} \sum_{j \in C} d(i, j), \quad (2.4)$$

dengan  $d(i, C)$  adalah jarak rata-rata objek *i* dengan semua objek pada kelompok lain *C* dimana  $A \neq C$ .

3. Menentukan nilai minimumnya yaitu  $b(i)$  yang menunjukkan perbedaan rata-rata objek *i* untuk kelompok yang terdekat dengan tetangganya dapat dituliskan dengan persamaan 2.5.

$$b(i) = \min_{C \neq A} d(i, C) \quad (2.5)$$

4. Menghitung nilai *Silhouette* dengan persamaan 2.6.

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2.6)$$

Hasil perhitungan  $s(i)$  berada pada kisaran -1 hingga

1. Nilai  $s(i)$  dapat diartikan sebagai:

- a.  $s(i) \approx 1$  artinya objek *i* terletak di kelompok yang tepat (dalam *A*),
- b.  $s(i) \approx 0$  artinya objek *i* terletak di antara 2 kelompok (*A* dan *B*),

- c.  $s(i) \approx -1$  artinya objek *i* terletak di kelompok yang tidak tepat (lebih dekat ke *B* daripada *A*).

5. Menghitung koefisien *Silhouette* yang didefinisikan sebagai rata-rata  $s(i)$  yaitu pada 2.7.

$$SC = \frac{1}{n} \sum_{i=1}^n s(i) \quad (2.7)$$

dengan *n* adalah banyak pengamatan.

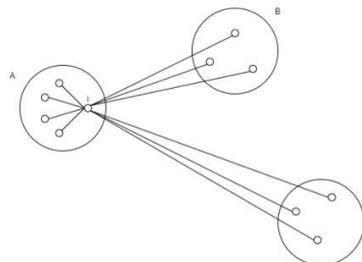
Pengelompokkan terbaik dicapai jika *SC* maksimal artinya meminimalkan jarak dalam kelompok ( $a(i)$ ) sekaligus memaksimalkan jarak antarkelompok ( $b(i)$ ).

Ukuran nilai koefisien *Silhouette* adalah seperti ditunjukkan pada Tabel 3.

Tabel 3. Kategori Nilai Koefisien *Silhouette* dan Interpretasinya [11]

Koefisien <i>Silhouette</i>	Interpretasi
$0.7 < SC \leq 1$	Terdapat ikatan yang sangat baik ( <i>strong structure</i> ) antara objek dan kelompok yang terbentuk.
$0.5 < SC \leq 0.7$	Terdapat ikatan yang cukup baik ( <i>medium structure</i> ) antara objek dan kelompok yang terbentuk.
$0.25 < SC \leq 0.5$	Terdapat ikatan yang lemah ( <i>weak structure</i> ) antara objek dan kelompok yang terbentuk.
$SC \leq 0.25$	Tidak terdapat ikatan antara objek dan kelompok yang terbentuk.

Gambar ilustrasi *Silhouette* dapat dilihat pada Gambar 1 di bawah ini.



Gambar 1. Ilustrasi *Silhouette* [12].

Berikut contoh perhitungannya:

Jarak antar Titik dalam Cluster C1:

A ke C:

$$\sqrt{(5 - 4)^2 + (10 - 9)^2} = 1.41$$

A ke D:

$$\sqrt{(5 - 6)^2 + (10 - 14)^2} = 4.12$$

Rata-rata jarak  $a(i)$  (Rata-rata jarak ke titik lain dalam cluster C1):

$$a(A) = \frac{\text{JarakAC} + \text{JarakAD}}{2} = \frac{1.41 + 4.12}{2} = 2.765 = 2.775$$

Nilai  $b(A)$  ke A ke B dan E di Cluster 2

A ke B:

$$\sqrt{(5 - 3)^2 + (10 - 7)^2} = 3.61$$

A ke E:

$$\sqrt{(5 - 2)^2 + (10 - 5)^2} = 5.83$$

Menghitung *Silhouette Score* untuk A:

$$s(A) = \frac{b(A) - a(A)}{\max(a(A), b(A))} = \frac{4.72 - 2.765}{\max(2.775, 4.6675)} = 0.412$$

Perhitungan tersebut dilakukan sampai ke data E. Hasil perhitungan dapat dilihat pada Tabel 4.

Tabel 4. Hasil Silhouette

Data	Cluster	a(i)	b(i)	Silhouette Score
A	C1	2.765	4.72	0.412
C	C1	3.4	4.495	0.305
D	C1	4.755	7.67	0.38
B	C2	2.24	4.68	0.521
E	C2	2.24	6.57	0.659

Rata-rata silhouette score untuk seluruh data dalam clustering ini dapat dihitung sebagai rata-rata nilai dari semua data:

$$\frac{1.412 + 0.305 + 0.38 + 0.521 + 0.659}{5} \approx 0.455$$

### 2.3 Toko Helm Kartini

Toko Helm Kartini adalah salah satu toko helm yang terletak di Jalan RA Kartini No.22 B-C, RT.006/RW.003, Margahayu, Kec. Bekasi Timur, Kota Bekasi, Jawa Barat 17113. Toko ini menyediakan berbagai jenis helm dari berbagai merk, seperti AGV(Amisano Gino Valenza), NOL(Nolan), INK(Indonesia-Korea, KYT(Koyoto), MDS(Matahari Departement Store), BMC(Bismillah Motor Center), HIU(Shark), NHK(Nippon Hoso Kyokai), GM(Gaul Man), ASCA, ZEUS, CAR(Cargloss), HBC(Helm Butho Cakil), JPX(Jaya Plastik X), NJS(Nusantara Jaya Sentosa, DYR, G2,. Selain helm toko ini juga menjual beberapa barang seperti SRM(Sarung Motor), SRT(Sarung Tangan, GOG(Goggles), Masker, Kaca, Aksesoris helm, dan lainnya seperti minuman.

## 3. Hasil Percobaan

### 3.1 Rancangan Sistem

Pada perancangan ini, sistem yang akan dibangun adalah sistem yang dapat mengelompokkan hasil pendapatan toko helm Kartini. Sistem ini akan menerapkan metode *clustering K-Means* dengan menggunakan bahasa pemrograman Python. Sistem ini akan membaca data yang di-input, melakukan preprocessing data, mengelompokkan data, mengevaluasi hasil *clustering*, dan juga menampilkan visualisasi berupa gambar. Langkah pertama.

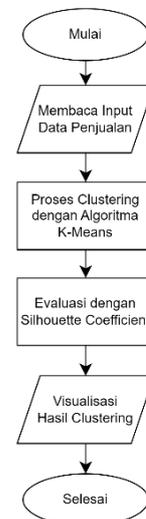
### 3.2 Tahap Perencanaan

Tahap perencanaan ini merupakan tahapan pertama dalam membangun sebuah sistem menurut metodologi, dimana pada tahap perencanaan ini akan ditentukan hal-hal yang berkaitan dengan pembangunan sistem seperti tujuan perancangan, sumber data yang digunakan, dan algoritma yang digunakan sebagai proses. Hal tersebut berfungsi untuk mempermudah memahami kegunaan sistem yang dibangun.

Tujuan dari pembangunan sistem ini adalah untuk membantu memudahkan toko mengatur stok penjualan. Algoritma yang digunakan adalah algoritma *clustering K-Means*. Hasil dari perancangan akan divisualisasikan dalam sebuah website untuk memudahkan pengguna mengakses informasi tersebut. Sumber data yang digunakan dalam sistem ini adalah catatan penjualan toko dalam bentuk file Excel. Algoritma *K-Means* digunakan sebagai metode untuk melakukan proses *clustering*, Hasil *clustering* akan dievaluasi menggunakan metode *Silhouette coefficient* untuk memastikan keakuratan pengelompokan.

### 3.3 Flowchart

Berikut proses alur sistem yang dapat dilihat pada Gambar 2.

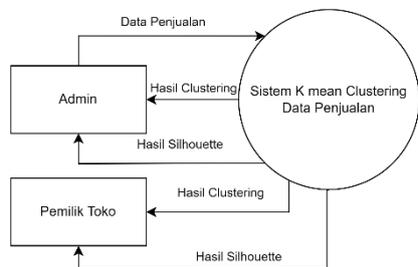


Gambar 2. Flowchart Sistem

Proses ini menggambarkan langkah-langkah dan aliran data yang terjadi dalam proses tersebut. Setiap tahap dalam diagram ini diilustrasikan dengan simbol standar yang menunjukkan jenis aktivitas yang dilakukan, serta panah yang menghubungkan setiap langkah untuk menunjukkan urutan eksekusinya.

### 3.4 Context Diagram

Diagram konteks adalah diagram tingkat tinggi, artinya diagram tersebut tidak membahas seluk-beluk sistem secara terperinci. *Context Diagram* bisa dilihat di Gambar 3.

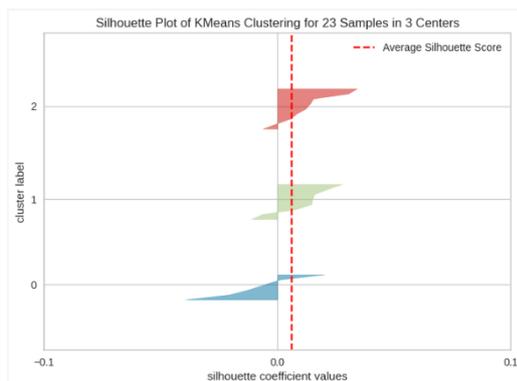


Gambar 3. Context Diagram

### 3.5 Hasil Clustering

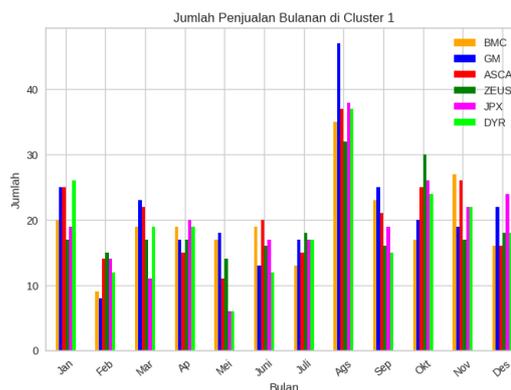
Proses *clustering* dilakukan menggunakan algoritma *K-Means* untuk mengelompokkan data penjualan toko helm Kartini berdasarkan kategori produk. Hasil dari proses *clustering* ini ditampilkan dalam beberapa visualisasi untuk dapat dianalisis lebih lanjut.

Setelah data penjualan helm normalisasi untuk mengurangi perbedaan skala antar-kategori produk, metode *K-Means* diterapkan untuk mengelompokkan data ke dalam 3 cluster. Hasil *clustering* dievaluasi menggunakan *Silhouette Score* untuk mengetahui seberapa baik data terkelompokkan dalam tiap-tiap cluster. Skor *Silhouette* yang dihasilkan adalah 0.0061, menunjukkan bahwa hasil *clustering* cukup optimal. Gambar *Silhouette* dapat dilihat pada Gambar 4.

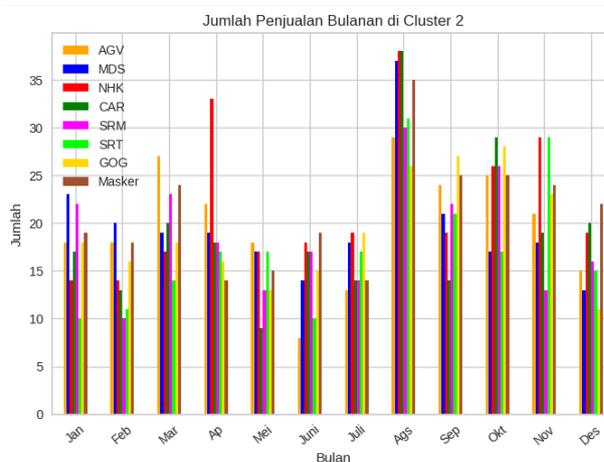


Gambar 4. *Silhouette*

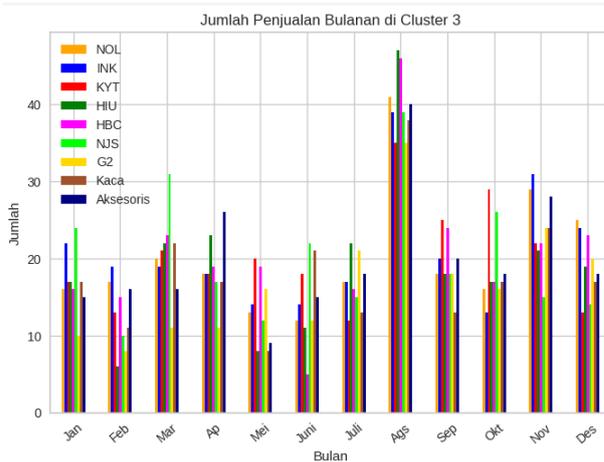
Setelah menentukan cluster untuk tiap kategori produk, data penjualan bulanan dari masing-masing cluster ditampilkan dalam bentuk diagram batang. Setiap cluster menunjukkan pola penjualan yang berbeda, yang dapat digunakan untuk melakukan analisis terhadap performa kategori produk per cluster. Gambar jumlah penjualan bulanan di setiap cluster dapat dilihat pada Gambar 5, Gambar 6, dan Gambar 7.



Gambar 5. Jumlah penjualan bulanan di cluster 1



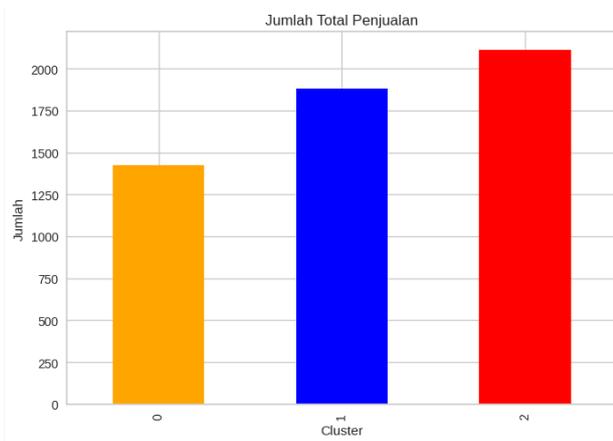
Gambar 6. Jumlah penjualan bulanan di cluster 2



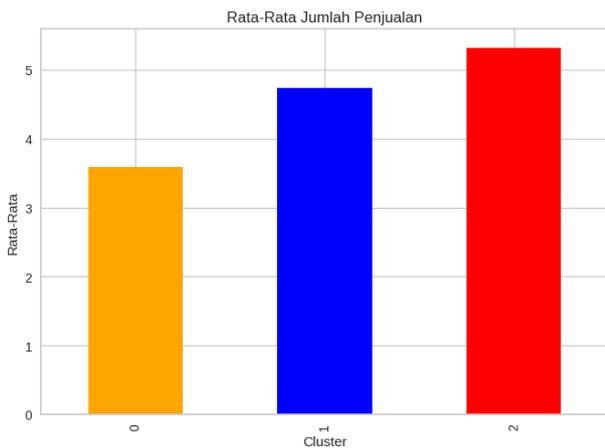
Gambar 7. Jumlah penjualan bulanan di cluster 3

Terlihat masing-masing cluster memiliki distribusi penjualan yang berbeda. Cluster 1 dan Cluster 3 menunjukkan pola penjualan yang cenderung lebih tinggi pada bulan tertentu, sementara Cluster 2 menunjukkan penjualan yang terlihat lebih stabil dibandingkan dengan cluster lainnya.

Pada Gambar 8 dan Gambar 9 terdapat jumlah total dan rata-rata penjualan pada setiap cluster untuk mengetahui seberapa besar jumlah total dan rata-rata setiap masing-masing cluster



Gambar 8. Jumlah total penjualan

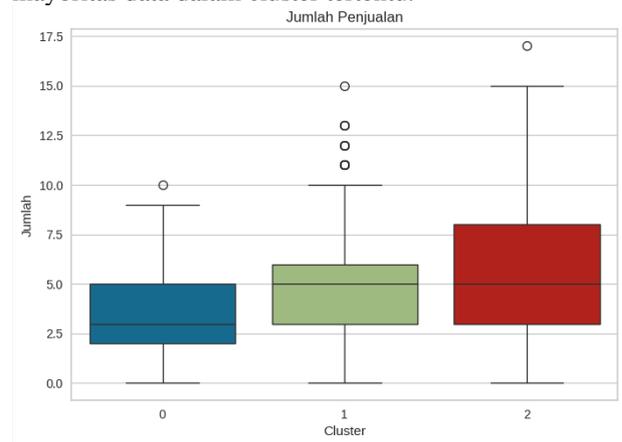


Gambar 9. Rata-rata jumlah penjualan

Hasil menunjukkan bahwa Cluster 3 memiliki jumlah penjualan total tertinggi dibandingkan cluster lainnya, sementara rata-rata penjualan per cluster bervariasi sesuai dengan kategori produk yang dominan dalam masing-masing cluster.

*Boxplot*, atau *box-and-whisker* plot adalah jenis visualisasi data yang digunakan untuk menganalisis dan menggambarkan distribusi nilai numerik dalam satu set data. *Boxplot* berguna untuk menunjukkan gambaran umum tentang rangkuman data, termasuk nilai minimum, maksimum, median, serta kuartil pertama (Q1) dan kuartil ketiga (Q3). *Boxplot* juga membantu mengidentifikasi pencilan (outliers) dalam data. Dalam konteks *clustering* data penjualan, *boxplot* dapat digunakan untuk melihat bagaimana distribusi jumlah penjualan berbeda di setiap cluster. Ini membantu kita memahami variasi penjualan dan mengidentifikasi apakah ada pencilan, seperti data

penjualan yang jauh lebih tinggi atau lebih rendah dari mayoritas data dalam cluster tertentu.



Gambar 10. Boxplot

Pada *boxplot* yang dapat dilihat pada Gambar 10, sumbu horizontal (x-axis) menunjukkan label dari masing-masing cluster, yaitu 0, 1, dan 2, sedangkan sumbu vertikal (y-axis) menunjukkan jumlah penjualan dalam unit yang dijual.

Cluster 0 (Warna Biru): Cluster ini menunjukkan median jumlah penjualan yang lebih rendah dibandingkan dengan cluster lainnya. Mayoritas penjualan dalam cluster ini berada dalam kisaran rendah, dengan beberapa pencilan (outliers) di atas median, yang menandakan adanya data penjualan yang lebih tinggi meskipun secara umum kelompok ini memiliki penjualan yang lebih sedikit.

Cluster 1 (Warna Hijau): Cluster ini memiliki median penjualan yang lebih tinggi dibandingkan dengan Cluster 0, tetapi masih berada dalam kisaran yang lebih rendah dibandingkan Cluster 2. Distribusi data di cluster ini lebih lebar dengan adanya beberapa pencilan di atas, yang menunjukkan adanya penjualan tinggi yang tidak terlalu sering.

Cluster 2 (Warna Merah): Cluster ini memiliki median dan rentang penjualan tertinggi dibandingkan cluster lainnya. Distribusi data pada cluster ini menunjukkan bahwa kelompok ini mencakup data dengan jumlah penjualan yang lebih besar, menandakan bahwa anggota dalam cluster ini memiliki kinerja penjualan yang tinggi.

Perbedaan median dan rentang penjualan di setiap cluster mengindikasikan adanya segmen-segmen penjualan dengan karakteristik yang berbeda. Cluster 0 dapat dikaitkan dengan penjualan yang rendah, Cluster 1 dengan penjualan menengah, dan Cluster 2 dengan penjualan yang tinggi. Hal ini menunjukkan bahwa toko helm Kartini dapat mengambil strategi yang berbeda untuk setiap cluster, seperti fokus pemasaran yang lebih intens pada cluster yang cenderung berpotensi penjualan tinggi.

#### 4. Kesimpulan

Dengan menggunakan *K-Means clustering*, analisis ini berhasil mengelompokkan data penjualan toko helm Kartini menjadi tiga cluster utama. Setiap cluster memiliki karakteristik penjualan yang berbeda, yang dapat membantu dalam pengambilan keputusan terkait strategi bisnis dan pemasaran. Cluster dengan penjualan yang tinggi dapat difokuskan untuk promosi lebih lanjut, sedangkan cluster dengan penjualan rendah dapat dianalisis lebih lanjut untuk memahami faktor yang mempengaruhi penurunan penjualan.

Secara keseluruhan, hasil *clustering* ini memberikan pemahaman untuk toko helm Kartini dalam mengelola stok, mengatur kategori produk, dan meningkatkan efektivitas promosi. Dengan memahami pola penjualan melalui analisis *clustering*, toko dapat membuat keputusan yang lebih tepat guna memenuhi kebutuhan pelanggan dan meningkatkan performa bisnis.

#### REFERENSI

- [1] A. V. Gheorghe, C. Lucaci, and M. G. Gheorghe, "Motorcycle helmets: A review of the effectiveness and standards," *Safety Science*, vol. 118, pp. 93-99, Jan. 2019.
- [2] M. N. Setiawan, S. A. Rachman, and T. Prasetyo, "Analysis of Factors Affecting Motorcycle Helmet Sales in Retail Stores," *Journal of Retail and Consumer Services*, vol. 53, pp. 200-210, Sept. 2020.
- [3] P. Alkhairi and A. P. Windarto, "Penerapan K-Means Cluster Pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara," *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, vol. 1, no. 1, pp. 762-767, 2019, [Online]. Available: <https://seminar-id.com/seminas-sainteks2019.html>
- [4] A. A. Rismayadi, N. N. Fatonah, and E. Junianto, "Algoritma K-Means Clustering Untuk Menentukan Strategi Pemasaran Di Cv. Integreet Konstruksi," *J. Responsif Ris. Sains dan Inform.*, vol. 3, no. 1, pp. 30-36, 2021.
- [5] K. Arai and A. R. Barakbah, "Hierarchical K-means: an algorithm for centroids initialization for K-means," *Reports of the Faculty of Science and Engineering*, vol. 36, no. 1, pp. 25-31, 2007.
- [6] S. Syarif, "Pengenalan Dasar Machine Learning dan Aplikasinya," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 1, pp. 50-56, 2022.
- [7] L. V. Bijuraj, "Clustering and its Applications," in *Proceedings of National Conference on New Horizons in IT - NCNHIT, Coimbatore*, 2013, pp. 169-172.
- [8] A. Kumar and S. V. Bhatia, "A survey on clustering algorithms and K-Means," *Journal of Computing and Information Technology*, vol. 27, no. 4, pp. 229-244, 2019.
- [9] A. Saputra, B. Mulyawan, and T. Sutrisno, "REKOMENDASI LOKASI WISATA KULINER DI JAKARTA MENGGUNAKAN METODE K-MEANS CLUSTERING DAN SIMPLE ADDITIVE WEIGHTING," *Jurnal Ilmu Komputer dan Sistem Informasi*, vol. 7, no. 1, pp. 14-21, 2019, doi: <https://doi.org/10.24912/jiksi.v7i1.5773>.
- [10] Y. T. Wang, J. Liu, dan S. Zhang, "Evaluating Clustering Algorithms with Silhouette Score," *Journal of Data Science and Analytics*, vol. 8, no. 2, pp. 95-104, 2022.

- [11] D. S. Utami, D. Retno, and S. Saputro, "PENGELOMPOKAN DATA YANG MEMUAT PENCILANDENGAN KRITERIA ELBOWDAN KOEFISIEN SILHOUETTE (ALGORITME KMEDOIDS)," in *Prosiding Konferensi Nasional Penelitian Matematika dan Pembelajarannya*, 2018.
- [12] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *J Comput Appl Math*, vol. 20, pp. 53-65, 1987.

**Nicko Kurniawan**, saat ini sebagai mahasiswa Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara..