



## *Implementation of K-Means Algorithm for Clustering Sales Data Based on Sales Patterns*

### **Implementasi Algoritma K-Means untuk Pengelompokan Data Penjualan Berdasarkan Pola Penjualan**

**Arfigo Yahya<sup>1\*</sup>, Rakhmat Kurniawan<sup>2</sup>**

<sup>1,2,3,4</sup>Program Studi Sistem Informasi, Fakultas Sains dan Teknologi,  
Universitas Islam Negeri Sumatera Utara, Indonesia

E-Mail: <sup>1</sup>arfigoyahya06@gmail.com, <sup>2</sup>rakhmat.kr@uinsu.ac.id

Received Oct 17th 2024; Revised Dec 25th 2024; Accepted Jan 11th 2025; Available Online Jan 13th 2025, Published Jan 26th 2025

Corresponding Author: Arfigo Yahya

Copyright © 2025 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

#### **Abstract**

*Toko Sembako Aceh Mulia has a distribution center. This distribution center stores many different products to be sold, looking at the store's distribution center still involves bookkeeping for various transaction information, and still reviewing the products to be purchased to meet stock at the distribution center, there is still no estimate for products that are generally sought after by customers, To produce the development of products that are currently popular and minimize products that are not often purchased which cause losses. In the existing problem using the K-Means method, by clustering or grouping the goods sold into 3 parts, namely very popular, selling well and less popular, with this division it is expected to make it easier for Toko Sembako Aceh Mulia in developing strategies in managing stocking of goods by grouping goods sold with very popular criteria. The clustering results from 30 datasets that have been analyzed show that C1 contains 21 types of products which can be categorized as the best-selling with a value of 181.23 in the type of Kapal Api coffee product and C2 contains 9 types of products which can be categorized as the least popular with a value of 228.03 in the type of Kapal Api coffee product.*

*Keyword: Clustering, Grocery Store, K-Means Algorithm, Products.*

#### **Abstrak**

Toko Sembako Aceh Mulia memiliki sebuah pusat distribusi. Pusat distribusi ini menyimpan banyak produk berbeda yang akan dijual, melihat pada pusat distribusi toko masih melibatkan pembukuan untuk berbagai informasi transaksi, dan masih mengkaji produk yang akan dibeli untuk memenuhi stok di pusat distribusi, masih belum ada estimasi untuk produk yang umumnya dicari oleh pelanggan, untuk menghasilkan perkembangan produk yang sedang populer dan meminimalisir produk tidak sering dibeli yang menyebabkan kerugian. Dalam permasalahan yang ada menggunakan metode *K-Means*, dengan mengklaster atau mengelompokkan barang-barang yang terjual menjadi 3 bagian yaitu sangat laku, laku dan kurang laku, dengan pembagian seperti ini diharapkan memudahkan Toko Sembako Aceh Mulia dalam menyusun strategi dalam management penyetokan barang dengan pengelompokan barang yang terjual dengan kriteria sangat laku. Hasil *clustering* dari 30 dataset yang sudah di analisis menunjukkan bahwa C1 terdapat 21 jenis produk dimana dapat dikategorikan paling laris dengan nilai 181,23 pada jenis produk kopi kapal api dan C2 terdapat 9 jenis produk dimana dapat dikategorikan paling tidak laris dengan nilai 228,03 pada jenis produk kopi kapal api.

Kata Kunci: *Algoritma K-Means, Pengelompokan, Produk, Toko Sembako*

#### **1. PENDAHULUAN**

Sembilah bahan pokok (Sembako) merupakan sembilan kebutuhan pokok masyarakat Indonesia yang terdiri dari makanan atau minuman yang digunakan untuk memenuhi kehidupan sehari-hari [1]. Barang-barang yang dijual di toko sembako Aceh Mulia berpasaran dalam bidang sembako kebutuhan sehari-hari, antara lain adalah beras, kecap, gula, minuman kemasan dan makanan kemasan, telur, indomie, bumbu dapur kemasan, dan lain-lain [2]. Toko Sembako Aceh Mulia memiliki sebuah pusat distribusi. Pusat distribusi ini menyimpan banyak produk berbeda yang akan dijual, melihat pada pusat distribusi toko masih melibatkan pembukuan

untuk berbagai informasi transaksi, dan masih mengkaji produk yang akan dibeli untuk memenuhi stok di pusat distribusi, masih belum ada estimasi untuk produk yang umumnya dicari oleh pelanggan, untuk menghasilkan perkembangan produk yang sedang populer dan meminimalisir produk tidak sering dibeli yang menyebabkan kerugian.

*Clustering* merupakan salah satu teknik dari salah satu fungsionalitas *data mining* [3], *algoritma K-Means* merupakan algoritma pengelompokan sejumlah data menjadi kelompok – kelompok data tertentu (*cluster*) [4]. Kelebihan dari metode yang dikembangkan oleh McQueen ini yaitu, mampu mengelompokkan dokumen dalam jumlah yang besar dengan waktu komputasi yang cepat [5]. Prinsip dari metode ini adalah mempartisi suatu koleksi dokumen menjadi beberapa *cluster* dan menentukan *centroid* (titik pusat awal) secara acak [6]. Tetapi, metode *K-Means* ini juga memiliki kelemahan yaitu sensitif terhadap *outlier* [7]. *Complete Linkage Clustering* digunakan untuk mengukur jarak yang terbesar [8]. Dapat juga diartikan cara mengukur jarak data dari yang terjauh [9]. *K-Means* membagi sekumpulan sampel ke dalam  $k$  kelompok atau *cluster* yang terpisah menggunakan nilai rata-rata anggota sebagai indikator utama. Titik ini biasanya disebut *Centroid*, mengacu pada entitas aritmatika dengan nama yang sama, dan direpresentasikan sebagai vektor dalam ruang dimensi yang berubah-ubah [10], dan *K-medoids* atau sering dikenal sebagai PAM (*Partitioning Around Medoids*), merupakan algoritma yang menerapkan objek sebagai perwakilan di setiap *cluster* yaitu *k-medoid*. *K-medoids* dapat meminimalkan jumlah perbedaan antara titik data dalam *cluster* dengan titik data terpilih di setiap *cluster* sebagai pusat (*k-medoid*) [11]. *Algoritma K-Means* hanya bisa menggunakan data bersifat numerik atau angka dalam mengolah atau pengaplikasiannya pada data. Ketidakmiripan suatu data dengan kelompok dapat direpresentasikan dengan jarak antara data dengan *centroid* (titik pusat) *cluster*. Jumlah terbaik dari kelompok tergantung pada jarak yang disebut *apriori* [12].

Melihat dari permasalahan-permasalahan yang sudah dipaparkan sebelumnya, dan mengusulkan teknik *data mining* untuk menyelesaikan masalah yang ada dengan mengolah data historis transaksi penjualan yang sudah ada. Dalam permasalahan yang ada menggunakan metode *K-Means*, dengan mengklaster atau mengelompokkan barang-barang yang terjual menjadi 3 bagian yaitu sangat laku, laku dan kurang laku, dengan pembagian seperti ini diharapkan memudahkan Toko Sembako Aceh Mulia dalam menyusun strategi dalam management penyetokan barang dengan pengelompokan barang yang terjual dengan kriteria sangat laku, maka menyetok barang dalam jumlah 3 banyak, dengan kriteria laku, maka menyetok barang dengan jumlah sedang dan kriteria kurang laku, maka menyetok barang dengan jumlah sedikit.

Terdapat beberapa penelitian sebelumnya dengan *algoritma K-Means* seperti produk laris dan kurang laris pada toko alfamart, dengan hasil penelitian dari proses *mining* membentuk *cluster* yang digunakan untuk memberi saran pertimbangan dalam menentukan strategi penjualan yakni mengeliminasi produk dengan posisi *cluster* terbawah dan memberikan *reward* untuk pelanggan dengan posisi *cluster* teratas [13]. Kemudian pengelompokan jenis merk laptop, pada penelitian ini telah berhasil mengelompokkan data laptop menggunakan *algoritma K-Means clustering* menjadi 3 kelompok yaitu untuk kelompok 1 berjumlah 11 data, kelompok 2 berjumlah 6 data, dan kelompok 3 berjumlah 8 data [14]. Kemudian pengelompokan penjualan kartu paket internet pada penelitian tersebut dalam menentukan kartu paket internet yang paling banyak diminati konsumen masih menggunakan sistem manual yaitu Telkomsel [15]. Kemudian pengelompokan penjualan parfume, hasil klaster yang diperoleh pada algoritma *K-Means* menunjukkan bahwa kelompok C0 produk terlaris meliputi 9 produk, kelompok C1 produk terlaris meliputi 3 produk, dan kelompok C2 mencakup 13 produk [16]. Kemudian pengelompokan untuk segmentasi konsumen, penelitian yang telah dilakukan dengan mengelompokkan konsumen yang paling berpotensi di PT. Super Sukses Niaga pada tahun 2018 ada pada *cluster* 2 yaitu nilai rata-rata total belanja setahun paling tinggi yaitu 4198.6656 [17].

## 2. METODOLOGI PENELITIAN

Metode penelitian yang digunakan pada penelitian ini, yaitu metode kuantitatif. Metode kuantitatif yang digunakan pada suatu penelitian, dimana data yang akan di analisis berupa angka (numerik). Penelitian ini mengambil objek dari data historis penjualan Januari – Oktober 2024. Variable penelitian yang digunakan berupa nama dan jumlah produk yang terjual setiap bulannya. Langkah-langkah penelitian ini dapat dilihat pada Gambar 1.

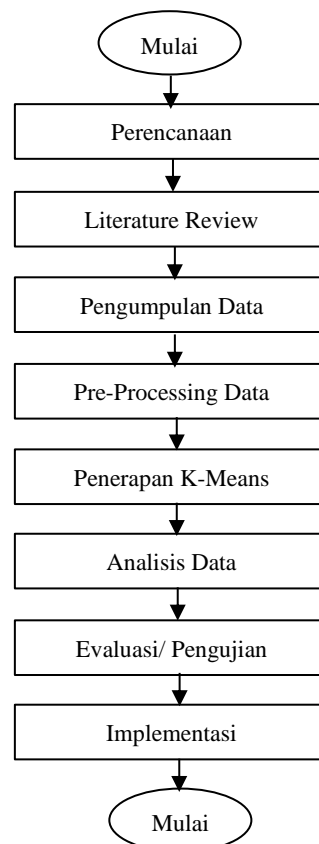
### 2.1. Perencanaan Penelitian

Secara umum perencanaan diartikan sebagai suatu aktivitas yang menelaah apa yang akan dilakukan, bagaimana melakukannya dan kapan dilakukan. Perencanaan penelitian atau dalam berbagai buku teks penelitian disebut dengan desain penelitian yang berarti semua proses yang diperlukan dalam perencanaan dan pelaksanaan penelitian. Penelitian ini dilakukan dengan menerapkan metode *K-Means* dalam proses pengelompokan data penjualan sembako.

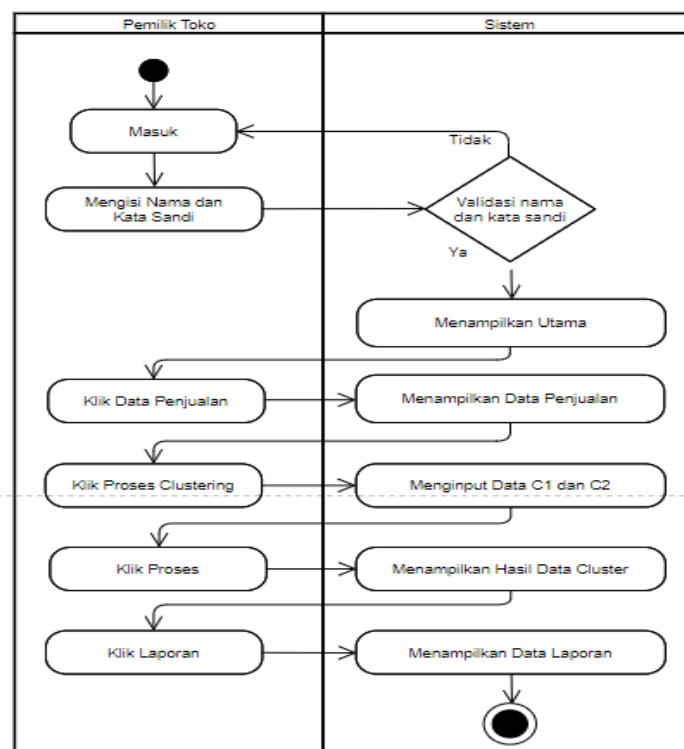
### 2.2. Perancangan Sistem

Sistem yang berjalan saat ini pada Toko Sembako Aceh Mulia masih manual dengan menggunakan buku catatan pada pendataan transaksi penjualan pelanggan, lalu tidak adanya analisis pola penjualan produk

yang diminati pada Toko Sembako Aceh Mulia. Dari analisis permasalahan sistem berjalan, dan dapat menyimpulkan bahwa dibutuhkannya sebuah sistem yang dapat membantu proses pencatatan transaksi penjualan dan analisis pola penjualan produk yang diminati di Toko Sembako Aceh Mulia. Berikut usulan kebutuhan sistem yang dapat dilihat pada Gambar 2.



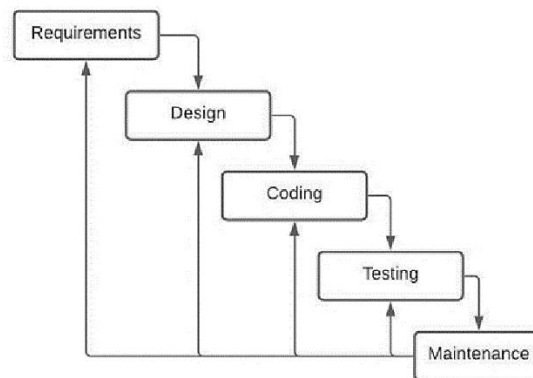
**Gambar 1.** Tahap Penelitian



**Gambar 2.** Activity Diagram Perancangan Sistem

### 2.3. Metode Pengembangan Sistem

Metode yang digunakan untuk mengembangkan sistem analisis pola penjualan produk terlaris pada Toko Sembako Aceh Mulia menggunakan metode *Waterfall* yang terlihat pada gambar 3.



**Gambar 3.** Tahapan *Waterfall*

Berikut langkah-langkah dalam proses penelitian:

1. *Requirement Analysis*

Pada tahap pertama, atau analisis kebutuhan, sebelum melakukan pengembangan perangkat lunak, pengembang perlu mengetahui dan memahami informasi tentang kebutuhan pengguna terhadap perangkat lunak. Metode pengumpulan data diperoleh dari penelitian literature dan dokumentasi mengenai implementasi *algoritma K-Means* sebagai pemilihan produk yang laris dan tidak laris pada penjualan produk. Data yang dipakai ialah history penjualan pada bulan Januari – Oktober 2024 pada Toko Sembako Aceh Mulia.

2. *Design*

Pada tahap ini menggunakan *Unified Modelling Language (UML)*, perancangan aplikasi dan perancangan *user interface*.

3. Implementasi

Selanjutnya pada tahap implementasi dilakukan menggunakan tools *Microsoft Excel 2019* sebagai perhitungan *dataset*.

4. Pengujian

Tahap selanjutnya, pengujian, menggunakan pengujian *White Box Testing* dan pengujian *Black Box Testing* untuk menyesuaikan aplikasi dengan kebutuhan pengguna.

### 2.4. Algoritma K-Means

*Algoritma K-Means* merupakan salah satu metode pengelompokan data non-hierarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih 13 kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diset dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok [18].

*Clustering* merupakan salah satu teknik dari salah satu fungsionalitas *data mining*, *algoritma K-Means* merupakan *algoritma* Pengelompokan sejumlah data menjadi kelompok – kelompok data tertentu (*cluster*). Kelebihan dari metode yang dikembangkan oleh McQueen ini yaitu, mampu mengelompokkan dokumen dalam jumlah yang besar dengan waktu komputasi yang cepat. Prinsip dari metode ini adalah suatu koleksi dokumen menjadi beberapa *cluster* dan menentukan *centroid* (titik pusat awal) secara acak. Tetapi, metode *K-Means* ini juga memiliki kelemahan yaitu sensitif terhadap *outlier* [19]. Secara umum, cara kerja dari *Algoritma K-Means clustering* adalah sebagai berikut [20]:

1. Menentukan  $k$  sebagai jumlah kluster yang di inginkan.
2. Menentukan nilai random untuk pusat *centroid* (*cluster* awal) sebanyak  $k$ .
3. Menghitung jarak setiap data input terhadap masing – masing *centroid* menggunakan rumus jarak *Euclidean (Euclidean Distance)* sampai menemukan jarak terdekat dari setiap data dengan *centroid*, berikut adalah persamaan pada rumus *Euclidean Distance*, yang ditunjukkan pada persamaan 1.

$$d(a_i, b_j) = \sqrt{\sum (a_i - b_j)^2} \quad (1)$$

4. Mengklasifikasikan data berdasarkan kedekatan-nya dengan *centroid* (jarak terkecil).
5. Memperbaharui nilai *centroid*, nilai *centroid* baru diperoleh dari rata – rata *cluster* yang bersangkutan dengan menggunakan persamaan 2.

$$b_j(t + 1) = \frac{1}{NS_j} \sum_{jesj} a_j \quad (2)$$

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Pengumpulan Data

Sumber data yang digunakan pada penelitian ini berupa data history penjualan dari buku transaksi toko. Data tersebut diperoleh dari melakukan observasi ataupun pengamatan langsung di Toko Sembako Aceh Mulia saat sebelum penerapan aktivitas pengumpulan informasi dengan melangsungkan identifikasi kepada objek penelitian, kegiatan kerja serta materi amatan serta melakukan wawancara secara langsung dengan pemilik Toko Sembako Aceh Mulia terkait pola penjualan produk yang digunakan untuk menarik konsumen.

**Tabel 1.** Data Transaksi

No	Produk	2023	2024
1	Dove	203	214
2	Minyak Goreng 1 Liter	170	180
3	Sunligh	220	240
4	Kopi Kapal Api	350	320
5	Detergen Baju Daia	250	234
6	Kecap Bango	137	159
7	Saus Sambal ABC	223	234
8	Roti	168	158
9	Garam	170	236
10	Beras	180	211

#### 3.2. Pre-Processing Data

Tahapan sebelum pemrosesan data dalam *data mining* lebih dikenal dengan tahap *preprocessing* data, adapun manfaat dari proses ini adalah untuk meningkatkan kualitas hasil pengelompokan data menggunakan *algoritma K-Means*. Penjelasan tahapan *pre-processing* data sebagai berikut:

1. Seleksi Atribut, pada tahap ini data dilakukan pengumpulan data dari buku catatan transaksi penjualan yang ada pada Toko Sembako Aceh Mulia, kemudian pemilihan data yang sesuai dengan variable/parameter yang digunakan dalam penelitian ini yakni data produk dan tahun penjualan. Pada tahap ini data yang dipakai adalah data Januari – Oktober 2024. Data yang dipakai adalah 10 jenis barang terlaris dari 50 jenis barang pada Januari – Oktober 2024 untuk melakukan analisis menggunakan *algoritma K-Means*.
2. Pembersihan Data, pada tahap ini dilakukan pengambilan data-data yang sesuai dengan variabel yang akan digunakan, tahap ini dilakukan proses pemisahan beberapa data seperti data produk penjualan yang terdapat pada tahun yang sama agar data tersebut dapat data produk yang sebelumnya di cantumkan bulan pada tahap ini hanya tersisa data tahun saja sesuai variabel yang dipakai dalam penelitian ini yakni tahun Januari – Oktober 2024 dikarenakan dari persetujuan pihak toko dan yang ingin memakai data transaksi per 2 tahun saja. Kemudian dilakukan perbaikan terhadap kesalahan penulisan pada nama produk dan kesalahan penulisan lain yang terdapat dalam data. Setelah melalui proses *cleaning* data yang sebelumnya sejumlah 203 pada tahun Januari 2024 menjadi 214 pada Februari 2024 pada produk dove, data yang akan dilakukan proses *clustering* untuk mendapatkan pola produk terlaris pada Toko Sembako Aceh Mulia.
3. Transformasi Data Pada tahap ini dilakukan perubahan data yang memiliki tipe data yang awalnya tidak bisa diolah secara matematis menjadi data yang bisa diolah. Tujuan dari transformasi data untuk menghindari data yang rusak dan tidak valid. Pada proses *K-Means* dilakukan transformasi berupa data alpha numerik (teks) menjadi numerik.

**3.3. Penerapan K-Means**

Data yang digunakan pada penelitian ini adalah data produk penjualan pada tahun Januari – Oktober 2024 pada Toko Sembako Aceh Mulia. Data dibawah ini adalah atribut data yang akan diolah, berupa nama produk dan produk yang terjual setiap tahunnya.

1. Perhitungan Jarak Pusat Cluster

Hitung jarak pusat cluster data terhadap setiap centroid produk Dove (Centroid 1 = 203, 214), Berikut hasil perhitungan jarak pusat cluster, ditunjukkan pada tabel 1.

**Tabel 2.** Perhitungan Jarak Pusat Cluster

No	Produk	C1	C2	Jarak Terpendek
1	Dove	0,000000000	47,38143096	0,000000000
2	Minyak Goreng 1 Liter	47,38143096	0,000000000	0,000000000
3	Sunligh	31,06444913	78,10249676	31,06444913
4	Kopi Kapal Api	181,2318956	228,035085	181,2318956
5	Detergen Baju Daia	51,07837116	96,5194281	51,07837116
6	Kecap Bango	85,91274643	39,11521443	39,11521443
7	Saus Sambal ABC	28,28427125	75,66372975	28,28427125
8	Roti	66,03786792	22,09072203	22,09072203
9	Garam	39,66106403	56,00000000	39,66106403
10	Beras	23,19482701	32,57299495	23,19482701

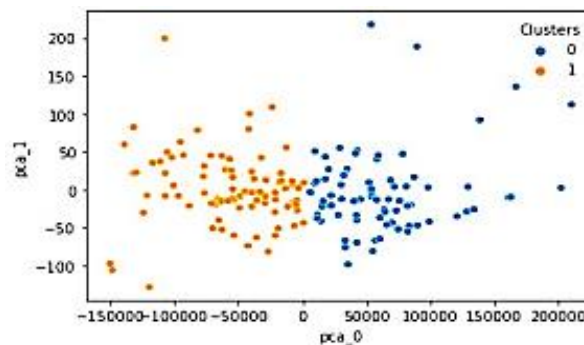
2. Pengelompokan Data

Selanjutnya pencarian jarak terpendek dari angka nilai perhitungan C1 dan C2, yaitu jarak terpendek dari Dove dengan nilai terpendek 0 yang artinya Dove dengan nilai C1 adalah produk sangat laris. Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat cluster, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat cluster terdekat, ditunjukkan pada tabel 3.

**Tabel 3.** Pengelompokan Data

No	Produk	C1	C2
1	Dove	1	
2	Minyak Goreng 1 Liter		1
3	Sunlight	1	
4	Kopi Kapal Api	1	
5	Detergen Baju Daia	1	
6	Kecap Bango		1
7	Saos Sambal ABC	1	
8	Roti		1
9	Garam	1	
10	Beras	1	

3. Untuk menentukan jumlah cluster terbaik, digunakan *Silhouette Coefficient*. Hasil perhitungan *Silhouette Coefficient* untuk beberapa jumlah cluster ditunjukkan pada Gambar 4. nilai *Silhouette Coefficient*, yaitu 0.5793 dengan jumlah cluster terbaik adalah k = 2.



**Gambar 4.** Visualisasi Hasil Clustering

**3.4. Analisis Data**

Suatu proses memecahkan masalah pada sistem yang utuh menjadi berbagai komponen sistem yang dapat dipahami tentang cara kerjanya dan bagaimana pengaruhnya antar komponen yang satu dengan

komponen lainnya. Tujuannya untuk mengevaluasi dan mengidentifikasi segala kebutuhan dan hambatan dalam membuat sebuah sistem agar sistem dapat berjalan dengan baik.

#### 1. Proses Pengambilan Data

Pada proses ini data yang diambil dari satu tabel dengan atribut total penjualan Januari – Oktober 2024 serta rata – rata penjualan dalam 2 tahun terakhir. Tabel ini yang akan dipakai untuk proses *clustering*.

#### 2. Proses *Clustering*

Pada penelitian *clustering* peneliti mengelompokkan data penjualan menjadi 2 *cluster*, dimana *cluster* 1 (C1) adalah mewakili data sangat laris, *cluster* 2 (C2) adalah mewakili data tidak laris. Pada langkah awal sebelum proses *clustering* dilakukan, terlebih dahulu sistem akan memilih produk sebanyak 2 angka yang akan digunakan sebagai nilai *centroid* pertama. Dan pada proses ini apabila ada angka yang sama, maka akan diulang sampai ketiga angka yang terpilih adalah berbeda. Setelah diperoleh 2 angka total penjualan, akan dilakukan proses *clustering* dengan *Algoritma K-Means*.

#### 3. Pada menu hasil *clustering* proses mencari jenis produk pengelompokan data penjualan juga menjadi 2 *cluster*, dimana *cluster* 1 (C1) adalah mewakili data produk sangat laris, *cluster* 2 (C2) adalah mewakili data produk tidak laris.

### 3.5. Evaluasi/Pengujian

Analisa metode *silhouette* dapat digunakan untuk validasi *algoritma K-Means*. Hal ini dilakukan dengan melihat besar nilai  $s$  dari hasil perhitungan dengan menggunakan bantuan *software MatLab*. Hasil perhitungan nilai *silhouette coefficient* dapat bervariasi antara -1 hingga 1. Jika  $si = 1$  berarti objek  $i$  sudah berada dalam cluster yang tepat. Jika nilai  $si = 0$  maka objek  $i$  berada di antara dua *cluster* sehingga objek tersebut tidak jelas harus dimasukkan ke dalam *cluster* A atau *cluster* B. Akan tetapi, jika  $si = -1$  artinya struktur *cluster* yang dihasilkan *overlapping*, sehingga objek  $i$  lebih tepat dimasukkan ke dalam *cluster* yang lain. *Silhouette Coefficient* merupakan salah satu teknik evaluasi yang digunakan untuk menentukan kualitas sebuah *cluster*. Dalam proses perhitungannya, teknik ini menggunakan rumus *Euclidean distance*.

Setelah dilakukan *clustering* pada data produk dengan menggunakan metode *K-Means* dan *tools Matlab*, maka didapatkan hasil *cluster* 1 berjumlah 75 produk, *cluster* 2 berjumlah 67 produk, *cluster* 3 berjumlah 55 produk, *cluster* 4 berjumlah 61 produk. Berikut hasil *clustering* menggunakan validitas *silhouette coefficient* (SI) dapat dilihat pada tabel 4.

**Tabel 4.** Hasil *Clustering*

No	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Hasil Cluster
1	0.6872	1.0951	1.0063	1.1838	1
2	0.7822	0.7920	0.5235	1.0162	3
3	0.4437	0.8345	0.6643	0.9008	1
4	0.3719	0.8375	0.7025	0.6502	1
5	0.6525	1.1022	0.9657	0.7574	1
...	...	...	...	...	...
25	1.1054	0.8629	0.9424	0.6903	4

### 3.6. Implementasi Sistem

Halaman ini menampilkan proses dan perhitungan dari *algoritma K-Means*. Pengguna harus mencari terlebih dahulu data transaksi berdasarkan nama produk, setelah itu masukan C1 dan C2 yang ingin dihitung dapat dilihat pada Gambar 5.

### 3.7. Diskusi

Berdasarkan hasil diskusi, bahwa metode *K-Means clustering* dapat menghasilkan produk yang laris dan kurang laris. Selain itu, pengujian dengan menggunakan sistem berbasis *website* dan *excel* sebagai perhitungan manual dengan dapat memunculkan hasil yang serupa, yaitu masing-masing memiliki anggota kelompok *cluster* 1 pada sistem memiliki *cluster* 1 berjumlah 75 produk yang laris yaitu dengan nilai 181,2318956 pada jenis produk kopi kapal api, *cluster* 2 berjumlah 67 produk yang biasa dengan nilai 51,07837116 pada jenis produk detergen baju daia, *cluster* 3 berjumlah 55 produk dan *cluster* 4 berjumlah 61 produk yang kurang laris nilai 0 pada jenis produk sabun dove dan minyak goreng 1 liter. Dengan bantuan *algoritma K-Means*, *data mining* menawarkan banyak keuntungan dan mempermudah proses pembuatan rencana manajemen stok produk penjualan pada data transaksi Toko Sembako Aceh Mulia yang ada.

Pada penelitian yang dilakukan oleh (Aceng Supriyadi, Agung Triayudi, dan Ira Diana Sholihati, 2021) [21]. Dengan hasil penelitian *clustering K-Means* melalui aplikasi web diperoleh kesesuaian jumlah dan susunan anggota *cluster* dengan perhitungan secara manual dan *tool Rapidminer* sebesar 100% untuk *cluster* 1 yaitu sebagai *cluster* dengan produktivitas rendah dengan jumlah anggota sebanyak 11 armada, pada *cluster*

ke-2 sebagai *cluster* dengan produktivitas sedang sebesar 96% dengan jumlah anggota pada perhitungan melalui aplikasi web sebanyak 21 armada, sedangkan secara manual dan juga *tool Rapidminer* sebanyak 20 armada, untuk *cluster* ke-3 sebagai *cluster* dengan produktivitas tinggi. Diperoleh persentase kesesuaian sebesar 97% dengan jumlah anggota 34 armada pada perhitungan aplikasi *web* dan sebanyak 35 armada pada perhitungan secara manual dan *tool Rapidminer*. Sedangkan pada penelitian yang sudah diteliti maka perbandingan pengujian hanyalah dengan menggunakan sistem berbasis *website* dan *excel* sebagai perhitungan manual dengan dapat memunculkan hasil yang serupa, yaitu masing-masing memiliki anggota kelompok *cluster* 1 pada sistem memiliki *cluster* 1 berjumlah 75 produk yang laris yaitu dengan nilai 181,2318956 pada jenis produk kopi kapal api, *cluster* 2 berjumlah 67 produk yang biasa dengan nilai 51,07837116 pada jenis produk detergen baju daia, *cluster* 3 berjumlah 55 produk dan *cluster* 4 berjumlah 61 produk yang kurang laris nilai 0 pada jenis produk sabun dove dan minyak goreng 1 liter pada Toko Sembako Aceh Mulia.

Hasil penelitian ini, akan memberikan kelebihan wawasan mendalam tentang preferensi konsumen dalam industri produk kebutuhan sehari-hari di Toko Sembako Aceh Mulia. Metode *K-Means clustering* diharapkan dapat mengidentifikasi pola yang mungkin terlewatkan, memberikan landasan bagi penyedia layanan untuk merancang strategi pemasaran yang lebih tepat sasaran. Implikasi praktisnya mencakup pengembangan produk yang lebih sesuai dengan preferensi konsumen, meningkatkan daya saing bisnis, dan meningkatkan kepuasan pelanggan. Bagi peneliti, temuan ini dapat memotivasi penelitian lanjutan di bidang analisis data dan perilaku konsumen. Keseluruhannya, penelitian ini diharapkan dapat memberikan kontribusi signifikan pada pemahaman dan pengembangan produk kebutuhan sehari-hari di Toko Sembako Aceh Mulia. Dengan keterbatasan ini menghambat pemahaman tentang perilaku pembelian, pengelolaan persediaan produk yang efektif serta pengetahuan strategi penjualan yang dimiliki pemilik toko dan saran untuk pengembang sistem untuk penelitian selanjutnya diharapkan menggunakan aplikasi berbasis android yang dapat digunakan melakukan handphone saja serta dapat menambahkan sistem transaksi penjualan secara *online*.

K-MEANS CLUSTERING APLIKASI BERBASIS WEBSITE DENGAN ALGORITMA K-MEANS METODE CLUSTERING | 01:32:44 am , 17-Jun-2023 |

Home Data Penjualan Proses Clustering Hasil Clustering Laporan Alur Sistem Logout

Proses Clustering

C1x	<input type="text" value="45"/>	C1y	<input type="text" value="25"/>
C2x	<input type="text" value="30"/>	C2y	<input type="text" value="20"/>

No	Nama Produk	Total Stok Produk	2021 - 2022			Total	Aksi	
			2021	2022	Rata-rata Penjualan		Editt	Delete
1	Minyak Goreng 1 Liter	525	170	180	175	525	Editt	Delete
2	Roti	489	168	158	163	489	Editt	Delete
3	Garam	609	170	236	203	609	Editt	Delete
4	Beras	586	180	211	195	586	Editt	Delete
5	Sunligh	690	220	240	230	690	Editt	Delete
6	Kopi Kapal Api	1005	350	320	335	1005	Editt	Delete

Gambar 5. Halaman Proses *Clustering*

#### 4. KESIMPULAN

Sistem ini dapat dipakai untuk membantu pihak toko dalam mencari dan menentukan produk terlaris dan tidak terlaris terutama pada penyediaan stok produk yang mulai menipis digudang toko. Hasil *clustering* dari 30 *dataset* yang sudah di analisis menunjukkan bahwa C1 terdapat 21 jenis produk dimana dapat dikategorikan paling laris dengan nilai 181,23 pada jenis produk kopi kapal api dan C2 terdapat 9 jenis produk dimana dapat dikategorikan paling tidak laris dengan nilai 228,03 pada jenis produk kopi kapal api. Pola penjualan produk pada Toko Sembako Aceh Mulia dari hasil implementasi dengan menggunakan *Algoritma K-Means* adalah dilakukan dengan menganalisis data mentah penjualan sembako dari Toko Sembako Aceh Mulia kemudian data tersebut diinisialisasikan kedalam bentuk angka lalu dilakukan proses perhitungan *distance score* sehingga terbentuk hasil *cluster* (hasil pengelompokkan).

#### REFERENSI

[1] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *JURIKOM (Jurnal Ris. Komputer)*, vol. 8, no. 6, p. 219, 2021, doi: 10.30865/jurikom.v8i6.3655.



- [2] F. A. Bramasta and R. Halilintar, "Penerapan Data Mining Untuk Menentukan Strategi Penjualan Toko Sepatu," *Pros. SEMNAS INOTEK ...*, pp. 236–241, 2021, [Online]. Available: <https://proceeding.unpkediri.ac.id/index.php/inotek/article/view/1135%0Ahttps://proceeding.unpkediri.ac.id/index.php/inotek/article/download/1135/736>
- [3] A. Ikhwan and N. Aslami, "Implementasi Data Mining untuk Manajemen Bantuan Sosial Menggunakan Algoritma K-Means," *J. Teknol. Inf.*, vol. 4, no. 2, pp. 208–217, 2020, doi: 10.36294/jurti.v4i2.2103.
- [4] M. H. Fakhriza and K. Umam, "Analisis Produk Terlaris Menggunakan Metode K-Means Clustering Pada "Pt.Sukanda Djaya," *JIKA (Jurnal Inform.*, vol. 5, no. 1, p. 8, 2021, doi: 10.31000/jika.v5i1.3236.
- [5] M. I. Jambak and R. Efendi, "Pengaruh Reduksi Dimensi Terhadap Metode Pengklasteran Berbasis Centroid Dan Metode Pengklasteran Berbasis Density Dalam Pengklasteran Dokumen Teks," *Indones. J. Bus. Intell.*, vol. 4, no. 2, p. 53, 2021, doi: 10.21927/ijubi.v4i2.1918.
- [6] I. Syafrinal and E. L. Febrianti, "Penerapan Algoritma K-Means Pada Aplikasi Data Mining Untuk Menentukan Pola Penjualan (Studi Kasus: Zahra Mart)," *J. Digit.*, vol. 13, no. 1, p. 31, 2023, doi: 10.51920/jd.v13i1.320.
- [7] A. Yani, Z. Azmi, and D. Suherdi, "Implementasi Data Mining Menganalisa Data Penjualan Menggunakan Algoritma K-Means Clustering," *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 2, no. 2, p. 315, 2023, doi: 10.53513/jursi.v2i2.6357.
- [8] A. Ayu, D. Sulistyawati, and M. Sadikin, "SISTEMASI: Jurnal Sistem Informasi Penerapan Algoritma K-Medoids untuk Menentukan Segmentasi Pelanggan," vol. 10, pp. 516–526, 2021, [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [9] S. Rahmatullah, M. Mukrim, and M. N. Pramitha, "Data mining untuk menentukan produk terlaris menggunakan metode naive bayes," *J. Inf. Dan Komput.*, vol. 7, pp. 57–64, 2019.
- [10] A. Rozaq, "Implementation of K-Means and Agglomerative Hierarchical Methods to House Clusterization," *J. Media Inform. Budidarma*, vol. 6, no. 2, p. 933, 2022, doi: 10.30865/mib.v6i2.3573.
- [11] Z. Muttaqin, "Implementasi Unsupervised Learning Pada Nilai Jasmani Kesamaptaaan Sekolah Polisi Negara Dengan Metode Clustering Analysis," *PROSISKO J. Pengemb. Ris. dan Obs. Sist. Komput.*, vol. 10, no. 1, pp. 18–23, 2023, doi: 10.30656/prosisko.v10i1.6269.
- [12] R. S. Ramadhan and S. Utara, "Clustering Daerah Rawan Angin Puting Beliung Pada Kabupaten Di Sumatera Utara Dengan Algoritma K-Means," *J. Ilm. Komputasi*, vol. 23, no. 2, pp. 247–252, 2024, doi: 10.32409/jikstik.23.2.3578.
- [13] I. Nawangsih, R. Puspita, and Suherman, "Implementasi Algoritma K-Means Dalam Mengkategorikan Produk Terlaris Dan Kurang Laris Pada Toko Alfamart Cikarang," *Pelita Teknol.*, vol. 16, no. 1, pp. 79–87, 2021, [Online]. Available: <https://jurnal.pelitabangsa.ac.id/index.php/pelitatekno/article/view/674>
- [14] P. Ulil, F. Aulia, and S. Saepudin, "Penerapan Data Mining K-Means Clustering Untuk Mengelompokkan Berbagai Jenis Merk Laptop," 2021.
- [15] H. Nicodemus Turnip and H. Fahmi, "Hendra Turnip, 2 Hasanul Fahmi (Penerapan Data Mining Pada Penjualan Kartu Paket Internet Yang Banyak Diminati Konsumen Dengan Metode K-Means) Penerapan Data Mining Pada Penjualan Kartu Paket Internet Yang Banyak Diminati Konsumen Dengan Metode K-Means," *JIKOMSI J. Ilmu Komput. dan Sist. Inf.*, vol. 4, no. 2, pp. 36–41, 2021.
- [16] 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/jieeee.v2i4.1181.
- [17] K. Auliasari and M. Kertaningtyas, "Penerapan Algoritma K-Means untuk Segmentasi Konsumen Menggunakan R," *J. Teknol. dan Manaj. Inform.*, vol. 5, no. 2, 2019, doi: 10.26905/jtmi.v5i2.3644.
- [18] Desy Julika Sari, W. Handoko, and Parini, "JUTSI ( Jurnal Teknologi dan Sistem Informasi ) Bangunan Terlaris Dengan Menggunakan Metode K-Means Di Ud Maju Bersama Mahasiswa Prodi Sistem Informasi , STMIK Royal Dosen Prodi Sistem Informasi , STMIK Royal JUTSI ( Jurnal Teknologi dan Sistem Informasi," vol. 2, no. 2, pp. 93–102, 2022.
- [19] A. R. Danurisa and J. Heikal, "Customer Clustering Using the K-Means Clustering Algorithm in the Top 5 Online Marketplaces in Indonesia," *BudapestInternational Res. Critics Institute-Journal*, vol. 5, no. 3, pp. 24287–24301, 2022, [Online]. Available: <https://doi.org/10.33258/birci.v5i3.6450>
- [20] M. M. Ridzki, I. Hadijah, M. Mukidin, A. Azzahra, and A. Nurjanah, "K-Means Algorithm Method for Clustering Best-Selling Product Data at XYZ Grocery Stores," *Int. J. Soc. Serv. Res.*, vol. 3, no. 12, pp. 3354–3367, 2023, doi: 10.46799/ijssr.v3i12.652.
- [21] A. Supriyadi, A. Triayudi, and I. D. Sholihati, "Perbandingan Algoritma K-Means Dengan K-Medoids Pada Pengelompokan Armada Kendaraan Truk Berdasarkan Produktivitas," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 6, no. 2, pp. 229–240, 2021, doi: 10.29100/jupi.v6i2.2008.