

# Evaluasi Kinerja *Clustering* K-Means Menggunakan Elbow Method dan Silhouette Coefficient pada Data Penimbangan Sampah

Anis Athifah<sup>1</sup>, Herlawati Herlawati<sup>1,\*</sup>, Fata Nidaul Khasanah<sup>1</sup>

\* Korespondensi: e-mail: [herlawati@ubharajaya.ac.id](mailto:herlawati@ubharajaya.ac.id)

<sup>1</sup> Informatika; Fakultas Ilmu Komputer; Universitas Bhayangkara Jakarta Raya, Jl. Perjuangan No. 81, Marga Mulya, Bekasi Utara, Bekasi, Jawa Barat, 17143, telp/fax: (021) 88955882; e-mail: [202210715265@mhs.ubharajaya.ac.id](mailto:202210715265@mhs.ubharajaya.ac.id), [herlawati@ubharajaya.ac.id](mailto:herlawati@ubharajaya.ac.id), [fata.nidaul.khasanah@dsn.ubharajaya.ac.id](mailto:fata.nidaul.khasanah@dsn.ubharajaya.ac.id)

Submitted : 16 April 2026  
Revised : 29 April 2026  
Accepted : 13 Mei 2026  
Published : 30 Mei 2026

## Abstract

*Bank Sampah Perumahan Sapta Pesona RW 08 Bekasi City has waste weighing data that is still limited to administrative records and has not been optimally used to identify waste characteristic patterns. This study aims to Cluster waste data based on quantity/tonnage and unit price using the K-Means algorithm. The research method used Cross Industry Standard Process for Data Mining (CRISP-DM). The dataset consisted of 810 waste weighing records from January 2024 to December 2025 with item type, quantity/tonnage, and unit price attributes. The number of Clusters was determined using the Elbow Method based on the Sum of Squared Error (SSE), while the Clustering results were evaluated using the Silhouette Coefficient. The results show that the data were grouped into three Clusters with a Silhouette Coefficient value of 0.7844, indicating a strong Cluster structure. Cluster 0 represents waste with low to moderate tonnage and unit price, Cluster 1 represents waste with low tonnage and high unit price, while Cluster 2 represents waste with high tonnage and low to moderate unit price. These results help managers understand waste data characteristics more systematically.*

**Keywords:** *Clustering, CRISP-DM, K-Means, Silhouette Coefficient, Waste Bank*

## Abstrak

Bank Sampah Perumahan Sapta Pesona RW 08 Kota Bekasi memiliki data penimbangan sampah yang masih terbatas sebagai pencatatan administrasi dan belum dimanfaatkan optimal untuk mengetahui pola karakteristik sampah. Penelitian ini bertujuan mengelompokkan data sampah berdasarkan jumlah/tonase dan harga satuan menggunakan algoritma K-Means. Metode penelitian menggunakan Cross Industry Standard Process for Data Mining (CRISP-DM). Data yang digunakan berjumlah 810 data penimbangan periode Januari 2024 sampai Desember 2025 dengan atribut jenis barang, jumlah/tonase, dan harga satuan. Jumlah Cluster ditentukan menggunakan Elbow Method berdasarkan nilai Sum of Squared Error (SSE), sedangkan hasil Clustering dievaluasi menggunakan Silhouette Coefficient. Hasil penelitian menunjukkan data terbagi menjadi tiga Cluster dengan nilai Silhouette Coefficient sebesar 0,7844 yang menunjukkan struktur Cluster kuat. Cluster 0 memiliki tonase dan harga satuan rendah hingga sedang, Cluster 1 memiliki tonase rendah dan harga satuan tinggi, sedangkan Cluster 2 memiliki tonase tinggi dan harga satuan rendah hingga sedang. Hasil ini membantu pengelola memahami karakteristik data sampah secara lebih terstruktur.

**Kata kunci:** *Bank Sampah, Clustering, CRISP-DM, K-Means, Silhouette Coefficient*

**JSRCS is a nationally accredited journal ranked SINTA 5 by the Directorate General of Research Strengthening and Development under Decree No. 225/E/KPT/2022, covering publications from Volume 1, Issue 2 (2020) to Volume 6, Issue 1 (2025). The journal was re-accredited under Decree No. 295/C/C3/KPT/2026 and retains its SINTA 5 status for the period from Volume 6, Issue 2 (2025) through Volume 11, Issue 1 (2030).**

## 1. Pendahuluan

Permasalahan sampah masih menjadi isu lingkungan yang perlu diperhatikan karena dapat berdampak terhadap kebersihan, kesehatan, dan kualitas hidup masyarakat. Sampah yang tidak dikelola dengan baik dapat menyebabkan pencemaran tanah, air, dan udara serta menurunkan nilai estetika lingkungan. Berdasarkan data komposisi sampah, sektor rumah tangga menjadi salah satu penyumbang sampah terbesar dibandingkan sumber lainnya (Nabila et al., 2025). Hal ini menunjukkan bahwa aktivitas masyarakat dalam kehidupan sehari-hari memiliki pengaruh besar terhadap peningkatan jumlah sampah. Oleh karena itu, pengelolaan sampah perlu dilakukan secara berkelanjutan dengan melibatkan peran aktif masyarakat sejak dari sumbernya (Nugraha et al., 2025).

Salah satu upaya pengelolaan sampah berbasis masyarakat adalah melalui program bank sampah. Bank sampah merupakan sistem pengelolaan sampah yang menerapkan kegiatan pemilahan, pengumpulan, penimbangan, dan pencatatan sampah yang masih memiliki nilai ekonomi. Melalui bank sampah, masyarakat tidak hanya berpartisipasi dalam menjaga lingkungan, tetapi juga dapat memperoleh manfaat ekonomi dari sampah yang disetorkan. Selain itu, kegiatan bank sampah menghasilkan data penimbangan yang dapat dimanfaatkan untuk mengetahui pola dan karakteristik sampah secara lebih terstruktur (Kurniawan & Fuaddah, 2024; Retno et al., 2025).

Bank Sampah Perumahan Sapta Pesona RW 08 Kota Bekasi merupakan salah satu bank sampah yang secara rutin melakukan kegiatan pengumpulan dan penimbangan sampah dari warga. Data yang dihasilkan berupa jenis barang, jumlah atau tonase, dan harga satuan. Namun, data penimbangan tersebut masih digunakan sebatas pencatatan administrasi dan laporan, sehingga belum dimanfaatkan secara optimal untuk memperoleh informasi mengenai karakteristik sampah. Kondisi ini menyebabkan pengelola belum dapat mengetahui kelompok sampah yang dominan maupun pola sebaran data penimbangan secara lebih jelas (Safitri et al., 2022).

Pemanfaatan *data mining* dapat digunakan untuk mengolah data penimbangan sampah agar menghasilkan informasi yang lebih bermanfaat. Salah satu teknik dalam *data mining* adalah *Clustering*, yaitu proses pengelompokan data berdasarkan kemiripan karakteristik tertentu (Salsabila et al., 2024). Algoritma *K-Means* merupakan salah satu metode *Clustering* yang banyak digunakan karena memiliki proses perhitungan yang sederhana, efisien, dan sesuai untuk data numerik. Metode ini dapat digunakan untuk mengelompokkan data sampah berdasarkan atribut jumlah atau tonase dan harga satuan sehingga pola karakteristik data dapat terlihat lebih jelas (Yudhistira & Andika, 2023)

Beberapa metode *Clustering* lain seperti *Hierarchical Clustering*, *DBSCAN*, dan *K-Medoids* juga dapat digunakan dalam proses pengelompokan data. Namun, *K-Means* dipilih dalam penelitian ini karena lebih sederhana dan praktis diterapkan pada data numerik dalam jumlah cukup besar. Dalam proses penerapannya, penentuan jumlah *Cluster* menjadi bagian penting agar hasil pengelompokan dapat merepresentasikan kondisi data dengan baik. Oleh

karena itu, penelitian ini menggunakan *Elbow Method* untuk menentukan jumlah *Cluster* optimal berdasarkan nilai *Sum of Squared Error* dan *Silhouette Coefficient* untuk mengevaluasi kualitas hasil *Clustering* (Shodikin & Taniasari, 2025); (Amri et al., 2025).

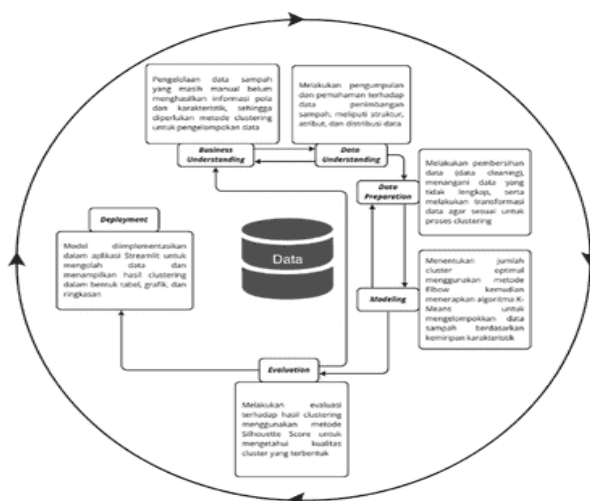
Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengelompokkan data penimbangan sampah pada Bank Sampah Perumahan Sapta Pesona RW 08 Kota Bekasi menggunakan algoritma *K-Means*. Data yang digunakan berjumlah 810 data penimbangan periode Januari 2024 sampai Desember 2025 dengan atribut jenis barang, jumlah atau tonase, dan harga satuan. Penelitian ini diharapkan dapat membantu pengelola bank sampah dalam memahami karakteristik data sampah serta mendukung pengelolaan dan penyaluran sampah ke bank sampah induk secara lebih terstruktur.

## 2. Metode Penelitian

Penelitian ini menggunakan metode *Cross Industry Standard Process for Data Mining* (CRISP-DM) sebagai acuan dalam proses pengolahan data untuk melakukan pengelompokan data sampah dan algoritma *K-Means* sebagai metode pemodelan *Clustering*.

### 2.1. Cross Industry Standard Process for Data Mining (CRISP-DM)

*Cross Industry Standard Process for Data Mining* (CRISP-DM) merupakan metodologi yang digunakan sebagai kerangka kerja dalam proses *data mining* secara sistematis. (Dhewayani et al., 2022). Metode ini terdiri dari enam tahapan utama, yaitu *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*. Dalam penelitian ini, CRISP-DM digunakan untuk mengarahkan proses pengolahan data penimbangan sampah, mulai dari pemahaman permasalahan, pemahaman data, persiapan data, pemodelan menggunakan algoritma *K-Means*, evaluasi hasil *Clustering*, hingga penyajian hasil dalam sistem berbasis *web*.



Sumber: Hasil Penelitian (2026)

Gambar 1. Kerangka Penelitian

### 2.2. K-Means Clustering

*K-Means Clustering* merupakan salah satu metode *Clustering* dalam kategori *unsupervised learning* yang digunakan untuk mengelompokkan data berdasarkan tingkat

kemiripan karakteristik tertentu. Metode ini bekerja dengan membagi data ke dalam sejumlah *Cluster* yang telah ditentukan sebelumnya. Setiap data akan ditempatkan pada *Cluster* yang memiliki jarak paling dekat dengan pusat *Cluster* atau *centroid*. Dengan demikian, data dalam satu *Cluster* memiliki karakteristik yang serupa, sedangkan data antar *Cluster* memiliki perbedaan yang cukup jelas (Rusvinasari & Annisa, 2025).

Perhitungan jarak pada algoritma *K-Means* umumnya menggunakan *Euclidean Distance* sebagaimana ditunjukkan pada persamaan (1).

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Keterangan:

$d(x, y)$  = jarak antara data dan *centroid*

$x_i$  = nilai data pada dimensi ke- $i$

$y_i$  = nilai *centroid* pada dimensi ke- $i$

$n$  = jumlah dimensi data

Proses *K-Means* dilakukan secara iteratif, yaitu menentukan jumlah *Cluster*, menentukan *centroid* awal, menghitung jarak data terhadap *centroid*, mengelompokkan data ke *Cluster* terdekat, menghitung ulang *centroid*, dan mengulangi proses hingga *centroid* tidak mengalami perubahan (Handayani, 2022). Dalam penelitian ini, algoritma *K-Means* digunakan untuk mengelompokkan data penimbangan sampah berdasarkan atribut jumlah/tonase dan harga satuan.

### 2.3. Elbow Method

*Elbow Method* merupakan metode yang digunakan untuk menentukan jumlah *Cluster* optimal pada algoritma *K-Means*. Metode ini dilakukan dengan menghitung nilai *Sum of Squared Errors* (SSE) pada beberapa nilai  $k$ . Nilai SSE menunjukkan tingkat variasi data dalam satu *Cluster*. Semakin kecil nilai SSE, maka data dalam *Cluster* semakin homogen. Rumus SSE ditunjukkan pada persamaan (2).

$$SSE = \sum_{i=1}^n \sum_{j=1}^k (x_i - c_j)^2 \quad (2)$$

Keterangan:

$x_i$  = data ke- $i$

$c_j$  = *centroid Cluster* ke- $j$

$k$  = jumlah *Cluster*

$n$  = jumlah data

Dalam penerapannya, nilai SSE ditampilkan dalam bentuk grafik. Titik perubahan yang membentuk sudut atau *elbow* menunjukkan jumlah *Cluster* yang paling optimal (Farida & Lubis, 2025)

## 2.4. Silhouette Coefficient

*Silhouette Coefficient* merupakan metode evaluasi yang digunakan untuk mengukur kualitas hasil *Clustering*. Metode ini melihat tingkat kedekatan data dalam satu *Cluster* dan perbedaannya dengan *Cluster* lain. Perhitungan metode ini didasarkan pada  $a(i)$ , yaitu rata-rata jarak data dengan anggota *Cluster* yang sama, dan  $b(i)$ , yaitu rata-rata jarak data dengan *Cluster* terdekat lainnya. Rumus *Silhouette Coefficient* ditunjukkan pada persamaan (3).

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (3)$$

Nilai *Silhouette Coefficient* berada pada rentang -1 sampai 1. Semakin tinggi nilai yang dihasilkan, maka kualitas *Cluster* semakin baik karena data berada pada kelompok yang sesuai dan memiliki pemisahan yang jelas dengan *Cluster* lain (Yolandari et al., 2025).

Tabel 1. Kriteria *Silhouette*

Rentang Nilai	Interpretasi
0.71 – 1.00	Struktur kuat
0.51 – 0.70	Struktur baik
0.26 – 0.50	Struktur lemah
$\leq 0.25$	Tidak terstruktur

Sumber: Toto (2025)

## 3. Hasil dan Pembahasan

Pembahasan pada penelitian ini meliputi penerapan algoritma *K-Means* dalam mengelompokkan data penimbangan sampah berdasarkan atribut jumlah/tonase dan harga satuan. Proses pengolahan data dilakukan menggunakan tahapan *Cross Industry Standard Process for Data Mining (CRISP-DM)*, yaitu *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*. Hasil dari setiap tahapan digunakan untuk mengetahui karakteristik kelompok sampah yang terbentuk pada Bank Sampah Perumahan Sapta Pesona RW 08 Kota Bekasi.

### 3.1. Business Understanding

Bank Sampah Perumahan Sapta Pesona RW 08 Kota Bekasi melakukan kegiatan pengumpulan dan penimbangan sampah secara berkala. Data hasil penimbangan tersebut selama ini masih digunakan sebatas untuk pencatatan administrasi dan laporan, sehingga belum dimanfaatkan secara optimal untuk mengetahui pola karakteristik sampah. Kondisi tersebut menyebabkan pengelola belum memperoleh informasi yang jelas mengenai kelompok sampah berdasarkan jumlah/tonase dan harga satuan.

Penelitian bertujuan untuk mengelompokkan data penimbangan sampah menggunakan algoritma *K-Means*. Hasil pengelompokan diharapkan dapat membantu pengelola dalam memahami karakteristik data sampah, seperti kelompok sampah dengan tonase tinggi, harga satuan tinggi, maupun kelompok sampah yang paling umum ditemukan dalam data penimbangan.

### 3.2. Data Understanding

Data yang digunakan dalam penelitian ini merupakan data penimbangan sampah pada Bank Sampah Perumahan Sapta Pesona RW 08 Kota Bekasi periode Januari 2024 sampai Desember 2025. *Dataset* awal berjumlah 810 data dengan atribut jenis barang, jumlah/tonase, dan harga satuan. Atribut jenis barang digunakan sebagai identitas data, sedangkan atribut jumlah/tonase dan harga satuan digunakan sebagai dasar dalam proses *Clustering*.

Adapun contoh data penimbangan sampah yang digunakan dalam penelitian ini dapat dilihat pada Tabel 2.

Tabel 2. Data Penimbangan Sampah

Jenis Barang	Jumlah / Tonase	Harga Satuan
AC / MESIN CUCI	0	45000
ACCU	0	0
ALUMUNIUM POCARI	0	8000
BELING	12.52	300
BESI A	28.3	4000
BESI B / STAL	0	3500
BUKU	44.55	800
CPU	0	40000
DISK / CD	0	7000
DUPLEK	16.12	400
EMBERAN	60.39	1500
GALON	0	4000
GELAS A	1.57	3700
GELAS B	0.7	1500
JERIGEN	0	1500
KALENG	4.1	2400
KARDUS	135.55	1000
KARUNG	0	0
ALUMINIUM	0.25	12000
KIPAS ANGIN	0	0
KORAN A	0	12000
.....	.....	.....
KRESEK / ASOY	13.706	500
KRISTAL / MIKA TEBAL	0	3500
KULKAS	0	0
KUNINGAN	0	80000
MAJALAH	0	1000
MIKA TIPIS / PVC	1.66	300
MINYAK JELANTAH	0	0
MONITOR LCD/TV	0	15000
PARALON	0	1700

Jenis Barang	Jumlah / Tonase	Harga Satuan
PET A	39.71	3200
GALON LE MINERALE	4.36	2000
PET KECAP	0	0
PET WARNA	0	1700
PLASTIK BENING	0	700
PRINTER	0	0
PUTIHAN	15.83	2000
SELOPAN	0.88	1000
TEMBAGA	0	120000
TETRAPACK	1.98	300
TUTUP BOTOL	1.91	3000
TUTUP GALON	0.39	5000
TV MONITOR TABUNG	0	0
YAKULT / IMPEK	0.42	300
CAMPURAN	0	0

Sumber: Hasil Penelitian (2026)

Berdasarkan Tabel 2, data penimbangan sampah memiliki variasi nilai pada atribut jumlah/tonase dan harga satuan. Perbedaan nilai tersebut menunjukkan bahwa setiap jenis sampah memiliki karakteristik yang berbeda, sehingga data dapat digunakan dalam proses pengelompokan menggunakan algoritma *K-Means*.

### 3.3. Data Preparation

Tahap *data preparation* dilakukan untuk menyiapkan data sebelum masuk ke proses pemodelan. Proses ini diawali dengan memilih atribut yang relevan, yaitu jumlah/tonase dan harga satuan sebagai atribut utama dalam proses *Clustering*. Atribut jenis barang tetap digunakan sebagai informasi pendukung untuk membantu membaca hasil pengelompokan.

Hasil pengecekan awal menunjukkan bahwa dataset berjumlah 810 data dengan 3 atribut. Tidak ditemukan *missing value* pada data, namun terdapat 374 data duplikat. Selain itu, atribut jumlah/tonase masih bertipe *object*, sehingga perlu dilakukan transformasi tipe data menjadi numerik. Setelah proses pembersihan data dilakukan, jumlah data menjadi 436 data dan seluruh atribut numerik telah siap digunakan dalam proses normalisasi.

Tabel 3. Hasil Pengecekan Data

Keterangan	Hasil
Jumlah data awal	810
Jumlah atribut	3
<i>Missing value</i>	0
Data duplikat	374
Jumlah data setelah <i>cleaning</i>	436

Sumber: Hasil Penelitian (2026)

Setelah dilakukan pembersihan data, jumlah data menjadi 436 data dengan 3 atribut. Hasil pengecekan menunjukkan bahwa tidak terdapat *missing value* dan data duplikat pada

dataset. Selain itu, atribut jumlah/tonase dan harga satuan telah memiliki tipe data *float64*, sehingga data dapat digunakan pada proses normalisasi dan pemodelan menggunakan algoritma *K-Means*, seperti ditunjukkan pada Gambar 2.

```

Jumlah data setelah cleaning: 436
Jumlah atribut: 3

Missing value setelah cleaning:
JENIS BARANG      0
JUMLAH / TONASE   0
HARGA SATUAN      0
dtype: int64

Jumlah data duplikat setelah cleaning:
0

Tipe data setelah cleaning:
JENIS BARANG      object
JUMLAH / TONASE   float64
HARGA SATUAN      float64
dtype: object
    
```

Sumber: Hasil Penelitian (2026)

Gambar 2. Pengecekan *Dataset* Setelah *Cleaning*

Setelah proses pembersihan data, dilakukan pengecekan distribusi data dan *outlier*. Berdasarkan pengecekan menggunakan histogram dan *boxplot*, atribut jumlah/tonase dan harga satuan memiliki sebaran data yang tidak merata. Sebagian besar data berada pada nilai rendah, sedangkan beberapa data memiliki nilai yang jauh lebih tinggi dibandingkan mayoritas data lainnya. Hasil pengecekan dengan metode *Interquartile Range* menunjukkan terdapat 47 *outlier* pada atribut jumlah/tonase dan 33 *outlier* pada atribut harga satuan. Namun, *outlier* tersebut tidak dihapus karena masih merepresentasikan kondisi nyata, seperti sampah dengan jumlah penimbangan tinggi atau harga satuan tinggi. Selanjutnya dilakukan normalisasi data menggunakan *StandardScaler*. Normalisasi dilakukan karena atribut jumlah/tonase dan harga satuan memiliki skala nilai yang berbeda, sehingga perlu disamakan agar tidak ada atribut yang mendominasi proses perhitungan jarak pada algoritma *K-Means*.

### 3.4. Modeling

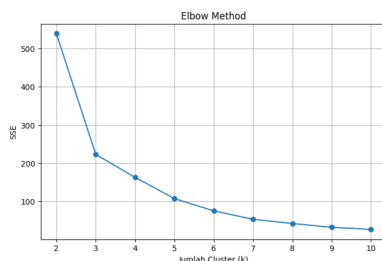
Tahap *modeling* diawali dengan menentukan jumlah *Cluster* optimal menggunakan *Elbow Method*. Nilai *Sum of Squared Error* (SSE) dihitung pada beberapa jumlah *Cluster* untuk melihat titik ketika penurunan nilai mulai melambat. Hasil perhitungan SSE dapat dilihat pada Tabel 4.

Tabel 4. Hasil Perhitungan SSE

Jumlah <i>Cluster</i> (k)	SSE
2	538,9405
3	223,1963
4	163,1220
5	107,6027
6	75,5402
7	53,5027
8	42,4808
9	32,5909
10	27,0921

Sumber: Hasil Penelitian (2026)

Berdasarkan Tabel 4, nilai SSE mengalami penurunan seiring bertambahnya jumlah *Cluster*. Penurunan terbesar terjadi dari  $k = 2$  ke  $k = 3$ , sedangkan setelah  $k = 3$  penurunan nilai SSE mulai melambat. Oleh karena itu, jumlah *Cluster* yang digunakan dalam penelitian ini adalah 3 *Cluster*. Grafik *Elbow Method* dapat dilihat pada Gambar 4.



Sumber: Hasil Penelitian (2026)

Gambar 4. Grafik *Elbow Method*

Setelah jumlah *Cluster* ditentukan, proses *Clustering* dilakukan menggunakan algoritma *K-Means* dengan jumlah *Cluster* sebanyak 3. Hasil dari proses *K-Means* kemudian ditambahkan ke dalam dataset sebagai atribut baru bernama *Cluster*. Contoh hasil penambahan label *Cluster* dapat dilihat pada Tabel 5.

Tabel 5. Contoh Hasil Penambahan Label *Cluster*

Jenis Barang	Jumlah/Tonase	Harga Satuan	<i>Cluster</i>
AC / Mesin Cuci	0,00	45000	1
Accu	0,00	0	0
Aluminium Pocari	0,00	8000	0
Beling	12,52	300	0
Besi A	28,30	4000	0

Sumber: Hasil Penelitian (2026)

Tabel 5 menunjukkan bahwa setiap data telah memperoleh label *Cluster* sesuai hasil pengelompokan. Label tersebut menunjukkan kelompok data berdasarkan kemiripan nilai jumlah/tonase dan harga satuan. Selanjutnya, dilakukan perhitungan jumlah data dan karakteristik pada masing-masing *Cluster*.

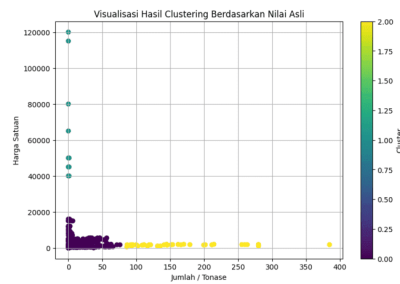
Tabel 6. Jumlah Data dan Karakteristik *Cluster*

<i>Cluster</i>	Jumlah Data	Rata-Rata Tonase	Min Tonase	Max Tonase	Rata-Rata Harga	Min Harga	Max Harga
0	388	11,4196	0,00	76,00	2807,2165	0	16000
1	10	0,3000	0,00	1,00	65000,0000	40000	120000
2	38	158,0930	86,12	384,75	1588,1579	500	2000

Sumber: Hasil Penelitian (2026)

Berdasarkan Tabel 6, *Cluster* 0 memiliki jumlah data terbanyak, yaitu 388 data. Kelompok ini memiliki karakteristik tonase rendah hingga sedang dan harga satuan rendah hingga sedang. *Cluster* 1 terdiri dari 10 data dengan karakteristik tonase rendah dan harga satuan tinggi. Jenis sampah yang termasuk dalam kelompok ini antara lain AC/Mesin Cuci, CPU, Kuningan, dan

Tembaga. Sementara itu, *Cluster 2* terdiri dari 38 data dengan karakteristik tonase tinggi dan harga satuan rendah hingga sedang, seperti Kardus, Emberan, Buku, dan Duplek.



Sumber: Hasil Penelitian (2026)

Gambar 5. Visualisasi Hasil *Clustering*

Hasil visualisasi pada Gambar 5 menunjukkan bahwa data membentuk tiga kelompok berdasarkan karakteristik jumlah/tonase dan harga satuan. Kelompok dengan harga satuan tinggi berada terpisah dari data lain karena memiliki nilai harga yang lebih besar, sedangkan kelompok dengan tonase tinggi membentuk pola tersendiri karena jumlah penimbangannya lebih besar dibandingkan kelompok lain.

### 3.5. Evaluation

Tahap *evaluation* dilakukan untuk menilai kualitas hasil *Clustering* menggunakan *Silhouette Coefficient*. Evaluasi ini digunakan untuk mengetahui seberapa baik data berada dalam *Cluster*-nya masing-masing dan seberapa jauh jaraknya terhadap *Cluster* lain.

```
# Menghitung Silhouette Coefficient
silhouette_final = silhouette_score(X_scaled, df_hasil_cluster['Cluster'])

print("Silhouette Coefficient:", round(silhouette_final, 4))

Silhouette Coefficient: 0.7844
```

Sumber: Hasil Penelitian (2026)

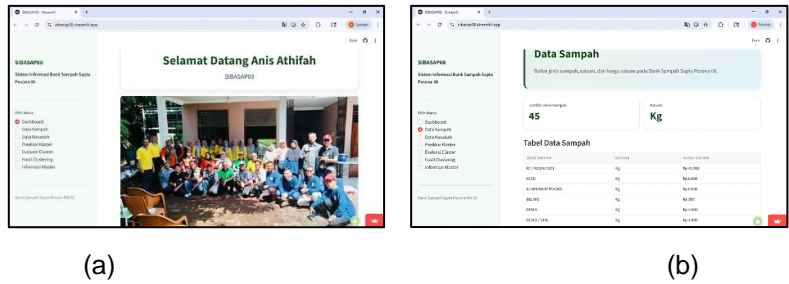
Gambar 6. Perhitungan *Silhouette Coefficient*

Berdasarkan hasil perhitungan, nilai *Silhouette Coefficient* yang diperoleh adalah 0,7844 dengan jumlah *Cluster* sebanyak 3. Nilai tersebut berada pada rentang 0,71–1,00, sehingga termasuk dalam kategori struktur kuat. Hal ini menunjukkan bahwa hasil *Clustering* menggunakan algoritma *K-Means* mampu mengelompokkan data sampah dengan baik berdasarkan kemiripan karakteristik jumlah/tonase dan harga satuan.

### 3.6. Deployment

Pada tahap *deployment*, hasil penelitian diterapkan ke dalam aplikasi berbasis *web* menggunakan *framework Streamlit*. Aplikasi yang dikembangkan bernama SIBASAP08 dan digunakan untuk membantu pengelolaan data Bank Sampah Perumahan Sapta Pesona RW 08 Kota Bekasi. Aplikasi ini terdiri dari beberapa menu utama, yaitu *Dashboard*, Data Sampah, Data Nasabah, Prediksi Klaster, Evaluasi *Cluster*, Hasil *Clustering*, dan Informasi Klaster. Gambar 7 menampilkan menu *Dashboard* dan Data Sampah, Gambar 8 menampilkan menu Data Nasabah dan Prediksi Klaster, Gambar 9 menampilkan menu Evaluasi *Cluster*, sedangkan Gambar 10 menampilkan menu Hasil *Clustering* dan Informasi Klaster.

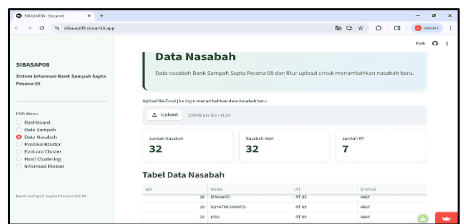
**Evaluasi Kinerja Clustering K-Means Menggunakan Elbow Method dan Silhouette Coefficient pada Data Penimbangan Sampah**



Sumber: Hasil Penelitian (2026)

Gambar 7. (a) Dashboard Aplikasi SIBASAP0, (b) Menu Data Sampah

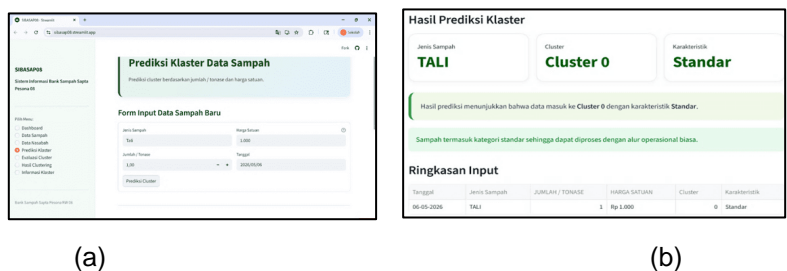
Gambar 7(a) menjelaskan informasi umum sistem dan Gambar 7(b) menjelaskan Menu Data Sampah yang menampilkan data penimbangan sampah.



Sumber: Hasil Penelitian (2026)

Gambar 8. Menu Data Nasabah dan Menu Prediksi Klaster

Menu data nasabah bank sampah dijelaskan dalam Gambar 8 dan juga yang menampilkan jumlah nasabah, nasabah aktif dan jumlah RT.



Sumber: Hasil Penelitian (2026)

Gambar 9. (a) Menu Prediksi Klaster, (b) Hasil Prediksi Klaster

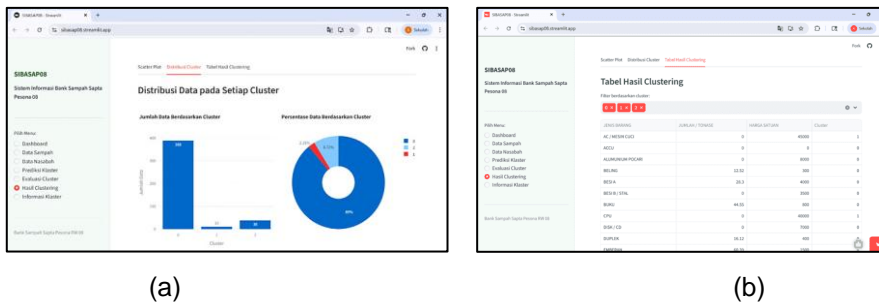
Pada Gambar 9(a) menampilkan Menu Prediksi Klaster yang digunakan untuk memasukkan jumlah/tonase dan harga satuan, dan Gambar 9(b) menampilkan Hasil Prediksi Klaster yang menampilkan hasil kelompok sampah berdasarkan data input, misalnya jenis sampah TALI, Cluster 0 dan Karakteristik Standar.



Sumber: Hasil Penelitian (2026)

Gambar 10. (a) Menu Evaluasi Cluster, (b) Grafik Elbow Method dan Silhouette Coefficient

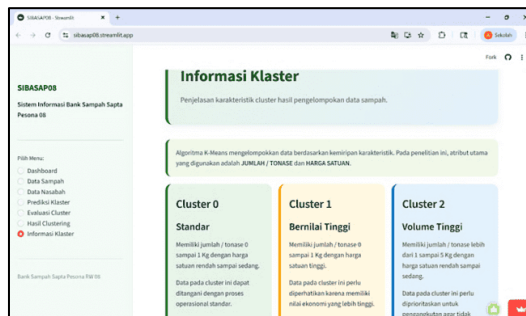
Pada Gambar 10(a) menampilkan hasil evaluasi pengelompokan data, dan Gambar 10(b) menampilkan Grafik *Elbow Method* dan *Silhouette Coefficient* yang digunakan untuk melihat jumlah *Cluster* optimal dan kualitas hasil *Clustering*.



Sumber: Hasil Penelitian (2026)

Gambar 11. (a) Menu Hasil *Clustering*, (b) Tabel Hasil *Clustering*

Pada Gambar 11(a) menjelaskan Tampilan Menu Hasil *Clustering* yang menampilkan hasil pengelompokan data sampah, dan Gambar 11(b) menjelaskan Tampilan Tabel Hasil *Clustering* yang menampilkan data dengan label *Cluster*.



Sumber: Hasil Penelitian (2026)

Gambar 12. Menu Informasi Kluster

Gambar 12 menampilkan halaman Informasi Kluster pada aplikasi pengelolaan data sampah yang menggunakan algoritma K-Means. Halaman ini menjelaskan bahwa data sampah dikelompokkan berdasarkan beberapa atribut seperti nilai, *volume*, dan nama komunitas. Tujuan dari proses *Clustering* ini adalah untuk menemukan pola atau karakteristik data sehingga pengelolaan sampah dapat dilakukan dengan lebih efektif dan terorganisir.

Pada hasil pengelompokan terdapat tiga *Cluster* utama, yaitu *Cluster 0* (Standar) yang berisi data dengan karakteristik rata-rata atau sedang, *Cluster 1* (Bernilai Tinggi) yang berisi sampah dengan nilai ekonomi tinggi, dan *Cluster 2* (*Volume Tinggi*) yang berisi data dengan jumlah atau *volume* sampah besar. Dengan adanya pembagian *Cluster* ini, sistem dapat membantu pengguna dalam menentukan prioritas pengelolaan dan pengambilan keputusan terkait pengolahan sampah. Berdasarkan tampilan tersebut, aplikasi SIBASAP08 dapat menyajikan data sampah, data nasabah, prediksi kluster, evaluasi *Cluster*, hasil *Clustering*, dan informasi kluster secara lebih sederhana sehingga mudah dipahami oleh pengelola bank sampah.

#### **4. Kesimpulan**

Penelitian ini berhasil mengelompokkan data penimbangan sampah pada Bank Sampah Perumahan Sapta Pesona RW 08 Kota Bekasi menggunakan algoritma *K-Means* berdasarkan atribut jumlah/tonase dan harga satuan. Data yang digunakan berjumlah 810 data periode Januari 2024 sampai Desember 2025, kemudian melalui tahap *data preparation* sehingga data siap digunakan dalam proses *Clustering*. Jumlah *Cluster* ditentukan menggunakan *Elbow Method* berdasarkan titik perlambatan penurunan nilai SSE, yaitu sebanyak tiga *Cluster* atau  $k = 3$ . Evaluasi menggunakan *Silhouette Coefficient* menghasilkan nilai 0,7844 yang termasuk kategori struktur kuat. Hasil pengelompokan menunjukkan bahwa *Cluster 0* merupakan kelompok sampah dengan tonase dan harga satuan rendah hingga sedang, *Cluster 1* merupakan kelompok sampah dengan tonase rendah dan harga satuan tinggi, sedangkan *Cluster 2* merupakan kelompok sampah dengan tonase tinggi dan harga satuan rendah hingga sedang. Hasil penelitian ini diterapkan ke dalam aplikasi berbasis *web* menggunakan Streamlit, sehingga dapat membantu pengelola bank sampah dalam memahami karakteristik data sampah dan mendukung pengelolaan sampah secara lebih terstruktur.

#### **Ucapan Terima Kasih**

Peneliti mengucapkan terima kasih kepada para reviewer atas masukan, kritik, dan saran yang membangun. Ucapan terima kasih juga disampaikan kepada Universitas Bhayangkara Jakarta Raya atas dukungan selama proses penelitian hingga artikel ini dapat disusun dengan baik.

#### **Daftar Pustaka**

- Amri, M. A., Fajrin Bakri, M., Akram, A., Risal, N., & Suriyanto, D. F. (2025). Mini-Batch K-Means Clustering Untuk Pengelompokan Kemandirian Daerah Di Sulawesi Selatan. *JIP (Jurnal Informatika Polinema)*, 11, 235.
- Dhewayani, F. N., Amelia, D., Alifah, D. N., Sari, B. N., & Jajuli, M. (2022). Implementasi K-Means Clustering untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM. *Jurnal Teknologi Dan Informasi*, 12(1), 64–77. <https://doi.org/10.34010/jati.v12i1.6674>
- Farida, J. I., & Lubis, A. H. (2025). Grouping Of Tourism Locations In Indonesia Using Distance Variations In The K-Means Algorithm. *Jite*, 8(3), 42–50. <https://doi.org/DOI:10.31289/jite.v8i3SpC.14528>
- Handayani, F. (2022). Aplikasi Aplikasi Data Mining Menggunakan Algoritma K-Means Clustering untuk Mengelompokan Mahasiswa Berdasarkan Gaya Belajar. *Jurnal Teknologi Dan Informasi*, 12(1), 46–63. <https://doi.org/10.34010/jati.v12i1.6733>
- Safitri, N., Myrna, R., & Ismanto, S. U. (2022). Partisipasi masyarakat dalam pengelolaan bank sampah di Kecamatan Jatiasih Kota Bekasi. *JANE (Jurnal Administrasi Negara)*, 14(1), 304–314.

- Kurniawan, A., & Fuaddah, A. (2024). Memberdayakan Rumah Tangga untuk Pengelolaan Sampah Berkelanjutan: Studi Kesadaran Masyarakat di Kota Semarang. *Journal of Urban Sociology*, 1(2), 112. <https://doi.org/10.30742/jus.v1i2.3494>
- Nabila, M. S., Herlawati, H., & Hidayat, A. (2025). Pendeteksian dan Klasifikasi Sampah pada Bank Sampah Berbasis Web Menggunakan YOLOv11. *Journal of Students' Research in Computer Science*, 6(1), 99–112. <https://doi.org/10.31599/r5me0z35>
- Nugraha, R., Suarna, N., Ali, I., & Rohman, D. (2025). Optimasi Pengelolaan Sampah Melalui Model Pengelompokan Dengan Algoritma K-Means. *Jurnal Teknik Informatika dan Teknik Elektro Terapan*, 13(1), 646–652. <https://doi.org/http://dx.doi.org/10.23960/jitet.v13i1.5694>  
OPTIMASI
- Retno, D. P., Putri, H. E., Putri, V. N., Rabbani, A., Angelica, A. M., & Nafi'ah, B. A. (2025). Partisipasi Mahasiswa dalam Sinergi Pemerintah dan Masyarakat untuk Pengelolaan Sampah melalui Bank Sampah di Kelurahan Kenjeran. *Pema*, 5(3), 36–45. <https://doi.org/10.56832/pema.v5i3.1552>
- Rusvinasari, D., & Annisa, L. H. (2025). Klasterisasi Pola Penjualan Menu Makanan pada Rumah Makan menggunakan Metode K-Means Clustering. *Jurnal Informatika: Jurnal Pengembangan IT*, 10(2), 398–409. <https://doi.org/10.30591/jpit.v10i2.8511>
- Salsabila, F., Ridwan, T., Karawang, U. S., & Timur, T. (2024). Analisa Volume Penyebaran Sampah Di Karawang Menggunakan Algoritma K-Means. *Jurnal Teknik Informatika dan Teknik Elektro Terapan*, 12(2). <https://doi.org/http://dx.doi.org/10.23960/jitet.v12i2.4226>
- Shodikin, M., & Taniasari, N. (2025). MALCOM: Indonesian Journal of Machine Learning and Computer Science Comparative Analysis of DBSCAN and K-Means Algorithms in Inpatient Patient Segmentation Using the RFMT Model Approach Analisis Komparasi Algoritma DBSCAN dan K-Means dalam Pemetaan Segmen. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 5, 1231–1241.
- Toto, T. A. (2025). Perbandingan Algoritma K-Means Dan Agglomerative Hierarchical Clustering Untuk Pengelompokan Daerah Penghasil Padi Di Indonesia. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(3). <https://doi.org/10.23960/jitet.v13i3.7251>
- Yolandari, N. A., Butarbutar, L. E., Rajagukguk, G. C. H., Zulfi, M. F., & Armita, F. R. (2025). Analisis Perbandingan K-Means Dan Dbscan Dalam Pengelompokan Data Travel Review Ratings Menggunakan Evaluasi Silhouette Index Dan Davies-Bouldin Index. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(3). <https://doi.org/10.23960/jitet.v13i3.6884>
- Yudhistira, A., & Andika, R. (2023). Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering. *Jurnal Aplikasi Informatika dan Teknologi Informasi*, 1(1), 20–28. <https://doi.org/https://doi.org/10.58602/jaiti.v1i1.22>