JITET (Jurnal Informatika dan Teknik Elektro Terapan)
Vol. 13 No. 1, pISSN: 2303-0577 eISSN: 2830-7062 http://dx.doi.org/10.23960/jitet.v13i1.5694

OPTIMASI PENGELOLAAN SAMPAH MELALUI MODEL PENGELOMPOKAN DENGAN ALGORITMA K-MEANS

Rifqi Nugraha[1*], Nana Suarna[2], Irfan Ali[3], Dede Rohman[4]

[1,2,3,4]STMIK IKMI Cirebon; Jl. Perjuangan No. 10B, Majasem, Cirebon, Jawa Barat 45135, Telp. (0231)490480

Received: 14 Desember 2024
Accepted: 14 Januari 2025
Published: 20 Januari 2025

Keywords:
Clustering, K-Means, Sampah, KDD, strategi pengelolaan.

Corespondent Email:
nugraharifqi0@gmail.com

Abstrak. Pengelolaan sampah di Jakarta menghadapi tantangan besar akibat volume dan karakteristik lokasi yang bervariasi. Penelitian ini dimaksudkan untuk menginvestigasi dampak dari beberapa faktor tersebut dengan menerapkan algoritma K-Means Clustering, untuk mengidentifikasi karakteristik setiap cluster dan mengusulkan strategi pengelolaan sampah yang lebih efektif. Metode yang digunakan adalah Proses *Knowledge Discovery in Database (KDD)*, yang mencakup tahap pengumpulan data, preprocessing, transformasi, dan analisis *clustering*. Hasil eksperimen menunjukkan bahwa jumlah klaster yang optimal adalah empat, menunjukan nilai DBI sebesar 0,220, menandakan kualitas klaster yang baik. *Cluster* 0, yang memiliki volume sampah tinggi, mendominasi Jakarta Barat dan memerlukan perhatian khusus dalam pengelolaan. *Cluster* 1, dengan volume sampah sedang, tersebar merata di seluruh daerah, menunjukkan potensi untuk meningkatkan efisiensi pengelolaan. Sementara itu, *Cluster* 2 dan Cluster 3 dengan volume rendah, berada di Jakarta Pusat dan Jakarta Utara, memerlukan pendekatan khusus seperti pengembangan fasilitas pengolahan sampah dan sistem daur ulang yang terintegrasi. Penelitian ini mengindikasikan bahwa pengelompokan menggunakan K-Means dapat memberikan wawasan berharga untuk mendukung strategi pengelolaan sampah yang berkelanjutan di Jakarta, serta menjadi dasar bagi pengambilan keputusan yang lebih terarah dalam memperbaiki kualitas lingkungan kota.

Abstract. *Waste management in Jakarta faces major challenges due to varying volume and location characteristics. This research is intended to investigate the impact of some of these factors by applying the K-Means Clustering algorithm, to identify the characteristics of each cluster and propose more effective waste management strategies. The method used is the Knowledge Discovery in Database (KDD) process, which includes the stages of data collection, preprocessing, transformation, and clustering analysis. The experimental results show that the optimal number of clusters is four, showing a DBI value of 0.220, indicating good cluster quality. Cluster 0, which has a high waste volume, dominates West Jakarta and requires special attention in management. Cluster 1, with moderate waste volume, is evenly distributed across the region, showing potential to improve management efficiency. Meanwhile, Cluster 2 and Cluster 3 with low volumes, located in Central Jakarta and North Jakarta, require special approaches such as the development of waste treatment facilities and integrated recycling systems. This research indicates that clustering using K-Means can provide valuable insights to support sustainable waste management strategies in Jakarta, as well as a basis for more targeted decision-making in improving the city's environmental quality.*

1 PENDAHULUAN

Masalah sampah di Indonesia telah menjadi isu nasional yang penting, karena dampaknya meliputi hilangnya nilai estetika lingkungan, pencemaran tanah, air, dan udara, munculnya berbagai sumber penyakit, serta potensi bencana alam jangka panjang. Meskipun permasalahan ini telah ada selama bertahun-tahun, penyelesaiannya masih belum tuntas. Situasi ini disebabkan oleh kurangnya pemahaman menyeluruh di antara para pemangku kepentingan mengenai cara pengelolaan sampah yang efektif, sehingga tantangan ini terus berlanjut tanpa solusi yang signifikan [1].

Di Indonesia, pengelolaan sampah merupakan tantangan besar yang semakin rumit akibat tingginya jumlah penduduk [2]. Menurut sumber data dari KLHK tahun 2023 melaporkan, total timbulan sampah di Indonesia mencapai 38,4 juta ton, dengan 61,58% sampah berhasil dikelola, namun 38,42% sisanya tidak terkelola dengan baik. Meskipun pemerintah telah menetapkan Undang-Undang Nomor 18 tentang pengelolaan sampah sejak 2008, penerapannya belum sepenuhnya efektif di semua kota [3]. Pertumbuhan jumlah penduduk yang pesat telah menyebabkan daerah pemukiman menjadi lebih luas dan padat, sekaligus meningkatkan produksi sampah secara signifikan. Masalah ini juga dialami oleh hampir semua kota besar di Indonesia [4]

Salah satu contohnya adalah Kota Jakarta. Berdasarkan data BPS (Badan Pusat Statistik) DKI Jakarta pada tahun 2023, jumlah penduduk di Jakarta mencapai 10.672.100 jiwa. Dengan populasi yang besar dan berbagai aktivitas masyarakat sehari-hari, masalah sampah menjadi tak terhindarkan. Ini terlihat dari meningkatnya volume sampah yang dihasilkan di suatu daerah seiring dengan pertumbuhan jumlah penduduk dan peningkatan pendapatan per kapita [5].

Dari masalah yang telah dijelaskan, diperlukan solusi yang efektif untuk penanganan, termasuk metode yang dapat mengidentifikasi wilayah di Kota Jakarta yang menghasilkan sampah terbanyak. Sehubungan dengan itu, diperlukan pengelompokan data salah satunya menggunakan algoritma *K-means clustering* yang merupakan bagian dalam data mining. *K-means clustering* adalah algoritma yang digunakan untuk mengelompokkan variabel-variabel penelitian ke dalam kluster atau polinomial yang tersedia dalam pengelompokan [6], dengan menerapkan proses clustering menggunakan metode k-means, kita akan dapat mengidentifikasi area di Jakarta yang membutuhkan perhatian lebih, sehingga pengelolaan sampah dapat dioptimalkan.

Beberapa penelitian sebelumnya relevan dengan analisis pengelompokan wilayah berdasarkan volume sampah di Kabupaten Karawang. Data mencakup volume sampah di setiap kecamatan, menghasilkan dua kluster: Cluster 0 dengan 6 kecamatan bervolume tinggi dan Cluster 1 dengan 24 kecamatan bervolume rendah. Evaluasi menggunakan indeks Davies-Bouldin (DBI) menunjukkan nilai 0,869, sedangkan nilai Silhouette mencapai 0,591, menandakan kualitas kluster yang baik [7].

Penelitian yang dilakukan oleh Magriaty dkk. di Kabupaten Tapin menerapkan metode PCA dan K-Means clustering untuk mengidentifikasi tiga zona pengelolaan sampah: Zona Tipe 1 (25 desa perkotaan dengan kerentanan tinggi), Zona Tipe 2 (36 desa pedesaan-perkotaan dengan pertumbuhan penduduk tinggi), dan Zona Tipe 3 (36 desa terpencil dengan kerentanan rendah). Pengelolaan difokuskan pada Zona Tipe 1 dan Tipe 2, yang mencakup 73,33% wilayah, sementara Zona Tipe 3, yang mencakup 26,67% wilayah, dikelola secara mandiri. [8].

Penelitian oleh Gusta Nugraha, dkk di Kota Magelang menggunakan K-Means clustering mengelompokkan wilayah penghasil sampah ke dalam tiga cluster berdasarkan volume dan area penyumbang. Evaluasi dengan Silhouette Score menghasilkan nilai 0,79, menunjukkan kualitas cluster yang baik. Hasil penelitian ini memberikan dasar bagi strategi pengelolaan sampah yang lebih efisien, termasuk optimalisasi peran bank sampah dan pengurangan dampak negatif sampah terhadap lingkungan serta kesehatan masyarakat [9].

2 TINJAUAN PUSTAKA

2.1 Clustering

Data mining merupakan teknik yang digunakan untuk menggali informasi dari suatu basis data. Dalam praktiknya, data mining mengaplikasikan berbagai metode, seperti kecerdasan buatan, pembelajaran mesin, dan teknik statistik. Metode ini berfungsi untuk mengekstrak dan mengidentifikasi data yang relevan dari basis data yang besar [10]. Clustering adalah salah satu metode dalam data mining yang menerapkan teknik *Unsupervised Learning* [11]. Clustering adalah proses pengelompokan titik data ke dalam dua atau lebih kelompok, di mana titik-titik dalam kelompok yang sama memiliki kesamaan yang lebih tinggi dibandingkan dengan titik-titik di kelompok lainnya, berdasarkan informasi yang ada dalam data tersebut. [12].

2.2 Algoritma *K-Means Clustering*

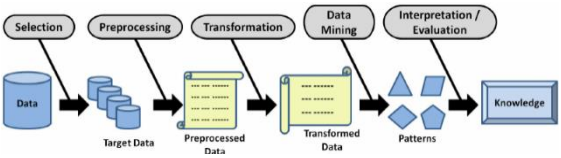
Algoritma K-means adalah metode pengelompokan data yang bersifat non-hierarkis, dengan tujuan untuk membagi data ke dalam satu atau beberapa kelompok atau kluster [13]. Dalam penerapan algoritma k-means, terdapat beberapa langkah yang harus diikuti, yaitu:

a. Menentukan jumlah klaster yang dibutuhkan.
b. Hanya atribut numerik yang digunakan

Algoritma K-means dimulai dengan secara acak menentukan pusat kluster awal, kemudian setiap data dikelompokkan ke dalam kluster terdekat berdasarkan jarak terkecil. Proses ini diikuti dengan pembaruan posisi pusat kluster secara bertahap hingga semua data terkelompokkan secara optimal ke dalam kluster yang terbentuk [14].

3 METODE PENELITIAN

KDD (Knowledge Discovery in Databases) adalah metodologi yang digunakan dalam penelitian ini untuk mendapatkan informasi dalam dataset. Metodologi *KDD* mencakup lima langkah sistematis[15] sebagai berikut.



Gambar 1. Tahapan Proses KDD

3.1 Data Selection

Pada tahap awal ini, fokus utamanya adalah memilih dataset yang sesuai dan relevan dengan tujuan penelitian.

3.2 Preprocessing Data

Tahap kedua dalam proses pra-pemrosesan data melibatkan pembersihan dan penghapusan data yang tidak relevan, menghilangkan nilai yang tidak sesuai, serta melakukan normalisasi data jika diperlukan.

3.3 Transformation Data

Pada tahap ini, data diubah menjadi format yang sesuai untuk digunakan dalam proses data mining, dengan tujuan memudahkan pengolahan data oleh alat dan algoritma yang akan digunakan.

3.4 Data Mining

Data mining adalah proses untuk mengidentifikasi pola atau informasi dalam kumpulan data selama tahap transformasi data. Dalam hal ini, metode K-Means digunakan sesuai dengan tujuan penelitian.

3.5 Evaluation

Pada tahap ini, dilakukan pengecekan kesesuaian data tersebut dianalisis berdasarkan fakta dan asumsi awal. Dalam penelitian ini, metode evaluasi yang digunakan adalah metode DBI *(Davies Bouldin Index)*.

3.6 Knowledge

Pola, hubungan dan tren data dianalisis secara mendalam pada tahap ini, untuk menghasilkan kesimpulan yang relevan dan mendukung pengambilan keputusan berdasarkan informasi dari hasil analisis.

4 HASIL DAN PEMBAHASAN

Tahapan penelitian ini mengacu pada metodologi yang diterapkan, yaitu metode *KDD (Knowledge Discovery in Databases).*

4.1 Data Selection

Data yang digunakan pada penelitian ini adalah data volume sampah Jakarta pada tahun 2024. Data ini adalah data sekunder yang diperoleh melalui situs resmi Kaggle, yang memiliki 54343 record dan 8 Atribut. Hasil *selection* dari dataset yang tidak digunakan adalah 4 atribut sehingga han ya akan menggunakan 4 atribut yaitu, kecamatan, wilayah, Panjang/luas, dan volume sampah perhari. Berikut hasil akhir data selection ditampilkan pada **Gambar 2**.

NO	kecamatan	wilayah	panjang/luas	volume_sampah_perhari(m3)
1	Johar Baru	Jakarta Pusat	464	4
2	Johar Baru	Jakarta Pusat	464	4
3	Johar Baru	Jakarta Pusat	464	4
4	Johar Baru	Jakarta Pusat	606	4
5	Johar Baru	Jakarta Pusat	310	7
6	Johar Baru	Jakarta Pusat	620	1
7	Johar Baru	Jakarta Pusat	300	12
8	Johar Baru	Jakarta Pusat	500	1
9	Johar Baru	Jakarta Pusat	500	1
54333	...	...	...	...
54334	Cakung	Jakarta Timur	700	0
54335	Cakung	Jakarta Timur	600	0
54336	Cakung	Jakarta Timur	800	0
54337	Cakung	Jakarta Timur	1000	0
54338	Cakung	Jakarta Timur	800	0
54339	Cakung	Jakarta Timur	800	0
54340	Cakung	Jakarta Timur	700	0
54341	Cakung	Jakarta Timur	1000	0
54342	Cakung	Jakarta Timur	1000	0
54343	Cakung	Jakarta Timur	400	2

Gambar 2. Atribut yang digunakan

4.2 *Preprocessing Data*

Pada tahap ini, dilakukan pemeriksaan terhadap nilai yang hilang dan duplikasi data. Sesuai dengan tujuannya, data yang tidak normal atau tidak sesuai akan ditangani, baik dengan cara dihapus maupun diperbaiki.

4.3 *Transformation Data*

Proses transformasi data, dilakukan perubahan format tipe data agar sesuai dengan tujuan penelitian sehingga dataset dapat diolah dengan baik.

```
#merubah tipe data menjadi menjadi numeric

df[ 'Luas(m2)' ] = pd.to_numeric(df[ 'Luas(m2)' ], errors='coerce')

df[ 'Volume_Sampah_Perhari(m3)' ] = pd.to_numeric(df[ 'Volume_Sampah_Perhari(m3)' ]

le=LabelEncoder()
df['Wilayah']=le.fit_transform(df['Wilayah'])
df['Kecamatan']=le.fit_transform(df['Kecamatan'])

df.dtypes
```

***Gambar 3**. Srcipt code change data type*

Dalam penelitian ini, data non-numerik akan dikonversi menjadi tipe data numerik. Hasilnya dapat dilihat pada **Gambar 4.**

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 54343 entries, 0 to 54342
Data columns (total 4 columns):
 #   Column                    Non-Null Count  Dtype
---  ------                    --------------  -----
 0   Kecamatan                 54343 non-null  int64
 1   Wilayah                   54343 non-null  int64
 2   Luas(m2)                  41044 non-null  float64
 3   Volume_Sampah_Perhari(m3) 36640 non-null  float64
dtypes: float64(2), int64(2)
```

Gambar 4. Hasil transform data

4.4 *Data Mining*

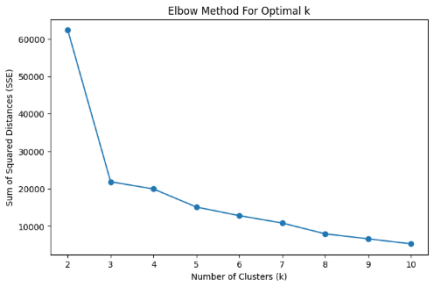
Pada tahap ini, algoritma *K-Means Clustering* diterapkan untuk mengategorikan wilayah dan kecamatan berdasarkan volume penyebaran sampah. Dengan memulai Langkah awal yaitu menentukan jumlah klaster terlebih dahulu. Penentuan jumlah kluster ini dapat dilakukan dengan berbagai metode, salah satunya yang paling umum digunakan adalah Metode Elbow.

Metode Elbow melakukan pemetaan grafik SSE untuk menemukan titik optimal jumlah kluster. Grafik kurva siku menunjukkan perubahan signifikan pada jumlah kluster, menandai titik optimal [16].

```
# Menggunakan metode Elbow untuk menentukan jumlah cluster
sse = []
silhouette_scores = []
k_range = range(2, 11)
for k in k_range:
    kmeans = KMeans(n_clusters=k, random_state=42)
    kmeans.fit(data_scaled)
    sse.append(kmeans.inertia_)
plt.figure(figsize=(8, 5))
plt.plot(k_range, sse, marker='o')
plt.xlabel('Number of Clusters (k)')
plt.ylabel('Sum of Squared Distances (SSE)')
plt.title('Elbow Method For Optimal k')
plt.show()
```

***Gambar 5**. Script Code Elbow Methode*

Setelah menentukan jumlah cluster, selanjutnya dibuat grafik Elbow untuk memvisualisasikan hubungan antara jumlah klaster yang berbeda dengan nilai SSE.



Gambar 6. Hasil Visual Elbow Method

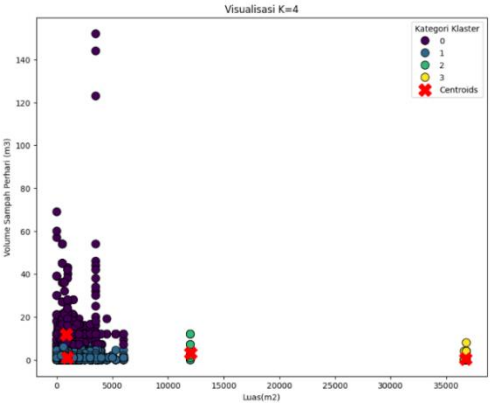
Pada **Gambar 6.** menunjukkan hasil pencapaian nilai K terbaik, yang menunjukkan jumlah kluster yang ideal untuk proses

klasterisasi data menggunakan algoritma *K-Means*. Dengan menggunakan 2 metrik evaluasi klaster: *inertia / Sum of Square Error* dan *Davies-Bouldin Indek* (DBI).

4.5 Evaluation

Hasil penelitian yang didasarkan pada perbandingan nilai Davies Bouldin Index (DBI) dengan metode K-Means Clustering. Pada **Gambar 6.** ditemukan bahwa cluster dengan nilai yang paling mendekati 0 berada pada K=4 dengan nilai DBI 0.220.

Pada **Gambar 7.** merupakan hasil cluster terbaik menggunakan K=4 dengan nilai *Davies Bouldin Index* (DBI)=0.220.



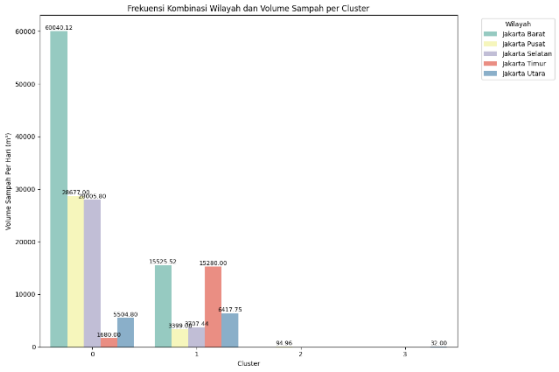
Gambar 7. Visualisasi *Cluster K=4*

Kemudian pada **Tabel 1.** menunjukan hasil clustering beserta jumlah anggota pada masing masing cluster.

Tabel 1. Hasil *Cluster*

No	Cluster	Anggota
1	Cluster 0	10447
2	Cluster 1	43803
3	Cluster 2	31
4	Cluster 3	62
Total number of items		54343

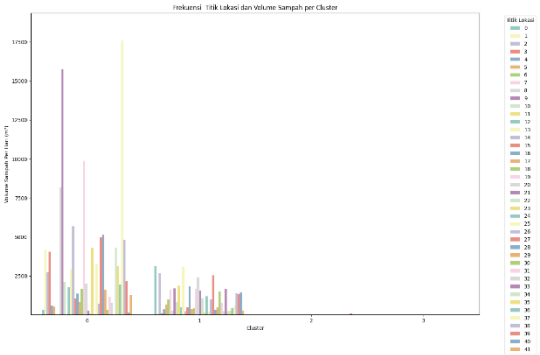
Setelah dilakukan pengelompokan awal kemudian dilanjutkan dengan melakukan pengelompokan berdasarkan Wilayah sebaran sampah. Ditampilkan pada gambar berikut.



Gambar 8. Hasil Visual Cluster Per Wilayah

Berdasarkan hasil pada **Gambar 8.** terdapat empat *Cluster* volume sampah yang diidentifikasi. Pada *Clusster* 0, volume sampah tertinggi mencapai 60.040,12 m³, volume sampah terendah sebesar 1.680,00 m³, dan volume sampah dengan intensitas sedang berada pada angka 28.005,80 m³. Selanjutnya, pada *Cluster* 1, volume sampah tertinggi tercatat sebesar 15.525,52 m³, volume sampah terendah sebesar 3.399,08 m³, dan volume sampah dengan intensitas sedang mencapai 6.417,75 m³. Berbeda dengan *Cluster* lainnya, pada *Cluster* 2 hanya terdapat volume sampah tertinggi yang bernilai 94,96 m³, sedangkan pada *Cluster* 3, volume sampah tertinggi tercatat sebesar 32,00 m³.

Selanjutnya pada **Gambar 9.** ditampilkan hasil cluster sebaran sampah berdasarkan Kecamatan.



Gambar 9. Hasil Cluster Per Kecamatan

Berdasarkan hasil pada **Gambar 9.** terdapat empat cluster dengan karakteristik yang berbeda.

Pada **Cluster 0**, volume sampah tertinggi tercatat sebesar 17.592,00 m³ yang berasal dari Kecamatan Tanjung Priok, sementara volume sampah terendah sebesar 8,00 m³ ditemukan di Kecamatan Ciracas. Volume sampah dengan intensitas sedang pada

cluster ini berada di angka 4.974,92 m³, seperti yang tercatat di Kecamatan Pasar Minggu.

Pada **Cluster 1**, volume sampah tertinggi mencapai 3.108,00 m³ di Kecamatan Cakung, sedangkan volume terendah sebesar 122,08 m³ ditemukan di Kecamatan Cilandak. Volume sampah dengan intensitas sedang di cluster ini berada pada kisaran 1.653,00 m³, seperti yang terlihat di Kecamatan Kramat Jati. Berbeda dengan cluster lainnya.

Pada **Cluster 2** hanya memiliki satu data, yaitu Kecamatan Duren Sawit, dengan volume sampah sebesar 94,96 m³. Kemudian pada **Cluster 3** hanya mencakup satu data, yakni Kecamatan Cilincing, dengan volume sampah tertinggi sebesar 32,00 m³.

4.6 Knowledge

Tahap terakhir ini adalah tahap dimana informasi diperoleh dari data yang sudah diekstrak sebelumnya. Telah diketahui sebelumnya, bahwa jumlah kluster yang optimal adalah 4 cluster. Dimana, cluster 0 merupakan wilayah dengan jumlah penyebaran sampah tinggi, sedangkan cluster 1 merupakan wilayah dengan jumlah penyebaran sampah sedang kemudian cluster 2 dan cluster 3 adalah wilayah dengan penyebaran sampah rendah.

Jumlah anggota cluster 0 adalah 10447 anggota, kemudian jumlah anggota cluster 1 adalah 43803 anggota, jumlah cluster 2 adalah 31 anggota dan jumlah cluster 3 adalah 62 anggota. Dari hasil evaluasi cluster, didapatkan nilai DBI 0.220.

Berdasarkan hasil tersebut, nilai evaluasi kluster yang diperoleh dalam penelitian ini menunjukkan bahwa kualitas kluster yang dihasilkan cukup baik. Hal ini berarti bahwa nilai-nilai dalam dataset memiliki kecenderungan untuk mirip satu sama lain dan berbeda secara signifikan dari objek dalam kluster lainnya.

5 KESIMPULAN

Berdasarkan hasil penelitian klasterisasi pada volume sampah di Jakarta menggunakan algoritma *K-Means Clustering*, dapat disimpulkan beberapa hal sebagai berikut:

1. Analisis clustering data volume sampah Jakarta dilakukan dengan metode *K-Means Clustering* yang membagi data ke dalam beberapa kelompok berdasarkan kemiripan karakteristik. Data ini berfokus pada atribut utama seperti *volume sampah*, *luas area*, *wilayah*, dan *kecamatan*. Pengujian dimulai dengan nilai K dari 2 hingga 10, di mana setiap nilai K diuji dengan menghitung nilai DBI untuk mengetahui performa dari tiap pembagian cluster.
2. Berdasarkan hasil pengujian, nilai DBI paling mendekati 0, yang menunjukkan kualitas klasterisasi yang baik, diperoleh pada saat K=4 dengan nilai DBI sebesar 0.220. Ini berarti bahwa K=4 adalah nilai optimal untuk pembagian cluster pada data volume sampah Jakarta, menghasilkan cluster yang lebih kompak dan terpisah dengan baik dibandingkan jumlah cluster lainnya.
3. Hasil pengelompokan memberikan gambaran mengenai komposisi volume sampah berdasarkan tingkat tinggi, sedang dan rendahnya sampah yang dominan di tiap cluster. Dengan memahami karakteristik tiap kelompok, dengan ini perencanaan kebijakan bisa lebih efektif dan terarah. Dijelaskan sebagai berikut.
    1) **Cluster 0 Volume Sampah Tinggi**: Cluster ini didominasi oleh wilayah Jakarta barat dengan jumlah volume 60040.12. Kelompok ini mungkin membutuhkan perhatian lebih terkait pengelolaan sampah atau yang menjadi prioritas pembersihan yang efektif untuk meningkatkan produktivitas sektor lingkungan.
    2) **Cluster 1 Volume Sampah Sedang**: Terdiri dari semua wilayah dengan volume sampah yang rata rata tidak jauh berbeda yaitu 3.00(m3). Kelompok ini memiliki volume sampah yang relatif stabil di semua wilayahnya, yang menandakan adanya pola yang cukup merata dalam pengelolaan sampah. Namun, pengelolaan sampah masih perlu difokuskan untuk mengoptimalkan efisiensi dan mengurangi dampak lingkungan.
    3) **Cluster 2 Volume Sampah Rendah**: Hanya terdiri dari wilayah Jakarta Pusat dengan jumlah volume 94.96 (m3). Wilayah Jakarta Pusat memiliki volume sampah yang lebih tinggi karena kepadatan aktivitas urban dan komersial.

Usaha yang dapat dilakukan meliputi pengembangan fasilitas pengolahan sampah skala besar, penerapan sistem pengelolaan sampah efisien di sektor bisnis, dan kampanye pengurangan sampah plastik. Selain itu, pembangunan fasilitas daur ulang terintegrasi di area publik juga perlu diperkuat.

4) **Cluster 3 Volume Sampah Rendah**: Hanya terdiri dari wilayah Jakarta Utara dengan jumlah volume 32.00(m3).

6 UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada semua pihak yang telah memberikan dukungan dalam penelitian ini.

7 DAFTAR PUSTAKA

[1] A. S. Kanda and C. Puspita Sari, "Analisis Permasalahan Dan Kebijakan Penanggulangan Sampah di Daerah Pajajaran Kota Bandung," vol. 2, no. 1, pp. 61–69, 2024.

[2] D. Ruwandara, M. Jajuli, and A. Rizal, "Analisis Algoritma K-Means Clustering Untuk Daerah Penyebaran Sampah di Kota Bekasi," *JOINS (Journal Inf. Syst.*, vol. 6, no. 1, pp. 56–63, 2021, doi: 10.33633/joins.v6i1.4085.

[3] sipsn.menlhk.go.id, "Capaian Kinerja Pengelolaan Sampah," sipsn.menlhk.go.id. Accessed: Nov. 28, 2024. [Online]. Available: https://sipsn.menlhk.go.id/sipsn/

[4] M. Chaerul and S. A. Rahayu, "Cost Benefit Analysis dalam Pengembangan Fasilitas Pengolahan Sampah: Studi Kasus Kota Pekanbaru," *J. Pengelolaan Sumberd. Alam dan Lingkung. (Journal Nat. Resour. Environ. Manag.*, vol. 9, no. 3, pp. 710–722, 2019, doi: 10.29244/jpsl.9.3.710-722.

[5] Jakarta.bps.go.id, "Jumlah Penduduk Menurut Kabupaten/Kota di Provinsi DKI Jakarta," jakarta.bps.go.id. Accessed: Nov. 28, 2024. [Online]. Available: https://jakarta.bps.go.id/id/statistics-table/2/MTI3MCMy/jumlah-penduduk-menurut-kabupaten-kota-di-provinsi-dki-jakarta-html

[6] T. F. Ardika Putri, "Clustering Sampah Yang Dihasilkan Oleh Masyarakat Kota Surabaya Dengan Menggunakan Metode K-Means," *J. SENOPATI Sustain. Ergon. Optim. Appl. Ind. Eng.*, vol. 5, no. 1, pp. 38–46, 2023, doi: 10.31284/j.senopati.2023.v5i1.4549.

[7] F. Salsabila, T. Ridwan, and H. H, "Analisa Volume Penyebaran Sampah Di Karawang Menggunakan Algoritma K-Means Clustering," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, 2024, doi: 10.23960/jitet.v12i2.4226.

[8] R. Magriaty, K. Murtilaksono, and S. Anwar, "Analisis K-Means Cluster untuk Identifikasi Kawasan Pengelolaan Sampah di Kabupaten Tapin Provinsi Kalimantan Selatan," *J. Reg. Rural Dev. Plan.*, vol. 7, no. 1, pp. 79–90, 2023, doi: 10.29244/jp2wd.2023.7.1.79-90.

[9] Y. Gusta Nugraha, M. Maimunah, and P. Sukmasetya, "Penerapan Algoritma K-Means Clustering untuk Daerah Penyebaran Sampah Kelurahan," 2023. doi: 10.47065/bits.v5i2.4158.

[10] Dodi Nofri Yoliad, "Data mining Dalam Analisis Tingkat Penjualan Barang ElektronikMenggunakan Algoritma K-means," *Insearch (Information Syst. Res. J.*, vol. 3, no. 1, 2023.

[11] T. M. M. Tyas and A. I. Purnamasari, "Penerapan Algoritma K-means dalam Mengelompokkan Demam Berdarah Dengue Berdasarkan Kabupaten," *Blend Sains J. Tek.*, vol. 1, no. 4, pp. 277–283, 2023, doi: 10.56211/blendsains.v1i4.231.

[12] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan," *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.

[13] M. F. Al Halik and L. Septiana, "Analisa Data Untuk Prediksi Daerah Rawan Bencana Alam Di Jawa Barat Menggunakan Algoritma K-Means Clustering," *J. Inf. Syst. Applied, Manag. Account. Res.*, vol. 6, no. 4, pp. 856–870, 2022, doi: 10.52362/jisamar.v6i4.939.

[14] F. Khoirunnisa and Y. Rahmawati, "Komparasi 2 Metode Cluster Dalam Pengelompokan Intensitas Bencana Alam Di Indonesia," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 1, pp. 68–79, 2024, doi: 10.23960/jitet.v12i1.3619.

[15] F. Alghifari and D. Juardi, "Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes," *J. Ilm. Inform.*, vol. 9, no. 02, pp. 75–81, 2021, doi: 10.33884/jif.v9i02.3755.

[16] N. A. Maori and E. Evanita, "Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 14, no. 2, pp. 277–288, 2023, doi: 10.24176/simet.v14i2.9630.