

METODE K-MEANS UNTUK CLUSTERING DATA PELANGGAN BERDASARKAN PEMAKAIAN AIR PADA PDAM TIRTA GIRI NATA KOTA CIREBON

Muhammad Abdul Jabbar¹, Ardhi Akmaludin Jadhira²

¹Politeknik Mardira Indonesia, Majalengka, Indonesia

² Politeknik Negeri Subang, Subang, Indonesia

Penulis Korespondensi: Jabbarmuhammad20@gmail.com

ABSTRAK

Peningkatan jumlah pelanggan dan kebutuhan air bersih menyebabkan Perusahaan Umum Daerah (Perumda) Air Minum Tirta Giri Nata Kota Cirebon memerlukan pengelolaan data pelanggan yang lebih efektif untuk mendukung perencanaan distribusi air dan pengambilan keputusan. Proses identifikasi pola pemakaian air yang masih dilakukan secara konvensional menyulitkan perusahaan dalam mengetahui karakteristik konsumsi pelanggan. Penelitian ini bertujuan menerapkan algoritma K-Means Clustering untuk mengelompokkan pelanggan berdasarkan tingkat pemakaian air sehingga diperoleh informasi mengenai kelompok pelanggan dengan konsumsi rendah, sedang, dan tinggi. Penelitian menggunakan pendekatan kuantitatif dengan tahapan pengumpulan data, *data preprocessing*, penentuan jumlah cluster ($k = 3$), perhitungan jarak menggunakan Euclidean Distance, pembaruan centroid secara iteratif hingga konvergen, serta evaluasi hasil menggunakan Davies–Bouldin Index (DBI). Data yang digunakan berjumlah 42 pelanggan berdasarkan volume pemakaian air (m^3). Hasil penelitian menunjukkan bahwa algoritma K-Means berhasil membentuk tiga cluster, yaitu Cluster 1 sebanyak 20 pelanggan dengan rata-rata pemakaian $8,5 m^3$ (kategori rendah), Cluster 2 sebanyak 15 pelanggan dengan rata-rata $18,3 m^3$ (kategori sedang), dan Cluster 3 sebanyak 7 pelanggan dengan rata-rata $36,4 m^3$ (kategori tinggi). Hasil validasi menghasilkan nilai DBI sebesar 0,231, yang menunjukkan kualitas cluster yang baik. Penelitian ini membuktikan bahwa algoritma K-Means efektif digunakan untuk melakukan segmentasi pelanggan berdasarkan pola konsumsi air sehingga dapat mendukung pengelolaan distribusi air, evaluasi penggunaan, dan pengambilan keputusan pada Perumda Air Minum Tirta Giri Nata Kota Cirebon.

Kata Kunci: *K-Means Clustering*, pemakaian air, PDAM, Davies–Bouldin Index.

Riwayat Artikel :

Tanggal diterima : 27-06-2026

Tanggal terbit : 23-07-2026

Kutipan :

Jabbar, M. A., & Jadhira, A. A. (2026). METODE K-MEANS UNTUK CLUSTERING DATA PELANGGAN BERDASARKAN PEMAKAIAN AIR PADA PDAM TIRTA GIRI NATA KOTA CIREBON. *INFOTECH Journal*, 12(2), 182–187. <https://doi.org/10.31949/infotech.v12i2.19034>

1. PENDAHULUAN

Menurut (Herawati et al., 2023) Air bersih merupakan salah satu kebutuhan dasar manusia yang memiliki peran penting dalam menunjang kehidupan, kesehatan, dan aktivitas sehari-hari. Air digunakan untuk berbagai kebutuhan domestik seperti minum, memasak, mandi, mencuci, serta menjaga kebersihan lingkungan. Seiring dengan meningkatnya jumlah penduduk, perkembangan wilayah, dan aktivitas ekonomi, kebutuhan terhadap air bersih terus mengalami peningkatan. Namun, di sisi lain, ketersediaan air bersih yang memenuhi persyaratan kualitas, kuantitas, dan kontinuitas masih menjadi tantangan di berbagai daerah di Indonesia. (Nurfiah et al., 2025) Di kawasan pedesaan, pengelolaan layanan air bersih seringkali menghadapi berbagai tantangan, seperti keterbatasan infrastruktur, tingkat kebocoran yang tinggi, dan kesulitan dalam memahami pola konsumsi pelanggan.

Tingkat konsumsi air di wilayah Kabupaten Cirebon yang menjadi daerah layanan PDAM merupakan salah satu indikator penting yang dapat dimanfaatkan sebagai dasar dalam perencanaan produksi air bersih pada periode berikutnya. Berdasarkan data penggunaan air bersih pelanggan, total volume konsumsi pada tahun 2021 tercatat sebesar 6.850.831 m³, kemudian mengalami peningkatan pada tahun 2022 menjadi 6.920.429 m³. Peningkatan tersebut menunjukkan adanya kecenderungan bertambahnya kebutuhan air bersih masyarakat sehingga diperlukan perencanaan dan pengelolaan distribusi air yang lebih efektif (Nurdin et al., 2025)

Perusahaan Daerah Air Minum (PDAM) Tirta Giri Nata merupakan salah satu Badan Usaha Milik Daerah (BUMD) yang mempunyai tugas memberikan pelayanan air bersih untuk masyarakat Kota Cirebon. Pada saat ini PDAM Giri Nata Kabupaten Nata terbagi dalam 5 cabang di Kabupaten Cirebon yaitu terdiri dari daerah kantor PDAM Tirtagiri Nata yaitu yang ada di Kota Cirebon, PDAM Tirta Jati yang berada di Kabupaten Cirebon, PDAM Tirtajati Unite Desa Babadan, PDAM Gegesik dan PDAM Arjawinangun. Saat ini PDAM kota Cirebon mampu melayani kurang lebih 60% dari seluruh penduduk kota Cirebon, sedangkan target pelayanan air bersih untuk skala kota besar adalah 80%. Hal ini terjadi karena kebutuhan air bersih dari tahun ke tahun selalu meningkat akibat dari penambahan jumlah penduduk, kemajuan teknologi serta peningkatan ekonomi masyarakat sedangkan debit air baku yang diolah PDAM kurang dari 80%.

Menurut (Yusa et al., 2018) menjelaskan bahwa segmentasi atau pengelompokan pelanggan pada suatu perusahaan dapat membantu perusahaan tersebut dalam membuat keputusan terhadap kelompok pelanggan tertentu sesuai kebijakan suatu perusahaan. pelanggan berdasarkan pola konsumsi pemakaian air terhadap kebutuhan air perkotaan

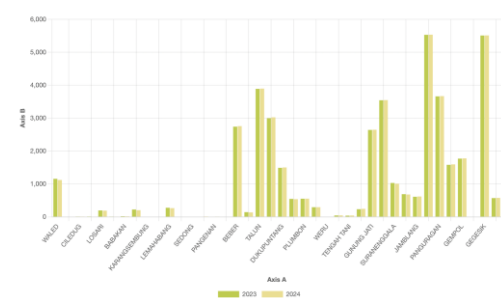
menggunakan metode *Clustering Kohonen atau Self Organized Maps (SOM)* serta menurut (Mikael Kia Mado, 2025) dalam penelitian algoritma clustering K-Means ini sangat relevan untuk mengcluster pelanggan.

Berdasarkan data yang telah didapatkan ketika melakukan observasi ditempat penelitian tentang menentukan data pelanggan pada pemakaian air berdasarkan pada perusahaan umum daerah air minum (PDAM) yang telah dilakukan, berikut contoh data yang tertulis tangan dalam buku yang disajikan dalam bentuk tabel.

Tabel 1. informasi umum pelanggan

No	Clustering Pelanggan	Tahun			
		0-10 m3 Rp	11-20 m3 Rp	21-30 m3 Rp	>30 m3 Rp
1.	Sosial Umum	2.300	2.300	2.300	2.300
2	Sosial Khusus	2.300	2.900	3.200	3.300
3	Rumah tangga 1	2.300	3.500	3.600	3.800
4	Rumah tangga 2	2.900	4.900	6.400	6.800
5	Rumah tangga 3	3.200	5.400	6.800	7.300
6	Rumah tangga 4	3.500	6.400	7.800	8.800
7	Instansi Pemerintah	3.800	6.800	9.300	13.000
8	Niaga Kecil	4.900	6.400	7.800	12.300
9	Industri Kecil	4.900	6.400	7.800	12.300
10	Niaga Besar	8.300	9.800	13.300	18.700
11	Industri Besar	9.800	11.300	14.700	22.100
12	Khusus	Berdasarkan Kesepakatan			

Berdasarkan tabel.1 penulis simpulkan bahwa dari pengelompokan data dapat diketahui kelompok nama pelanggan PDAM Kabupaten Cirebon yaitu di PDAM Tirta Giri Nata yang pemakaian M3. Yang artinya pada data pelanggan paling terendah pada sosial umum terdapat tarif yang paling rendah tagihannya sedangkan pada industri besar memakan tarif tagihan air PDAM paling tinggi sedangkan untuk yang khusus berdasarkan kesepakatan dari pihak tertentu berikut gambar 1.



Gambar 1. Statistik Pelanggan PDAM Kabupaten Cirebon (*Open Data Kabupaten Cirebon, 2025*)

Menurut (Annisa et al., 2022) Metode *K-Means* merupakan salah satu algoritma *clustering* yang digunakan untuk mengelompokkan sejumlah objek atau data ke dalam beberapa kelompok (*cluster*) berdasarkan tingkat kemiripan karakteristiknya. Setiap objek akan ditempatkan pada cluster yang memiliki nilai rata-rata (*centroid*) paling dekat dengan objek tersebut. Algoritma ini bekerja secara iteratif dengan memperbarui posisi centroid hingga diperoleh kelompok yang stabil atau tidak terjadi lagi perpindahan anggota antarcluster. Konsep kerja *K-Means* memiliki kemiripan dengan algoritma *Expectation-Maximization* (EM) pada *Gaussian Mixture Model*, karena keduanya sama-sama bertujuan menentukan pusat kelompok secara optimal melalui proses iterasi untuk memperoleh hasil pengelompokan yang terbaik dan berdasarkan penelitian yang dilakukan oleh (Sri Tria Siska, 2016), algoritma *K-Means* digunakan untuk mengelompokkan pelanggan PDAM berdasarkan kubikasi air menjadi kelompok hemat, sedang, dan boros. Penelitian tersebut berhasil menghasilkan informasi mengenai pola penggunaan air pelanggan, namun masih berfokus pada proses analisis menggunakan perangkat *RapidMiner* dan belum mengembangkan aplikasi yang dapat digunakan secara langsung dalam operasional PDAM.

Selanjutnya, penelitian (Lestari, 2023) menerapkan algoritma *K-Means* untuk mengelompokkan pelanggan SPAM Akidah berdasarkan tingkat penggunaan air dengan bantuan *RapidMiner* serta melakukan evaluasi menggunakan Davies-Bouldin Index (DBI). Penelitian ini menunjukkan kualitas hasil clustering yang baik, namun belum mengintegrasikan proses clustering ke dalam aplikasi pengelolaan pelanggan.

Penelitian (Mutammimul Ula et al., 2023) mengombinasikan metode *K-Nearest Neighbor* (KNN) dan *K-Means* untuk menentukan pelanggan baru dan mengelompokkan pelanggan berdasarkan wilayah. Fokus penelitian tersebut lebih diarahkan pada proses klasifikasi pelanggan dan pengelompokan wilayah sehingga belum membahas pengelolaan data pemakaian air pelanggan dalam bentuk aplikasi terintegrasi.

Sementara itu, penelitian (Sinaga et al., 2019) menggunakan algoritma *K-Means* untuk mengelompokkan jumlah pelanggan perusahaan air bersih berdasarkan provinsi di Indonesia. Penelitian tersebut memberikan gambaran persebaran jumlah pelanggan pada tingkat nasional, namun objek penelitian masih berupa data agregat sehingga belum dapat dimanfaatkan untuk pengelolaan data pelanggan pada tingkat operasional PDAM.

Penelitian yang dilakukan oleh (Hairani et al., 2026) menerapkan algoritma *K-Means* yang dikombinasikan dengan model *RFM* (*Recency, Frequency, Monetary*) untuk melakukan segmentasi pelanggan PDAM berdasarkan tingkat loyalitas. Hasil penelitian menunjukkan bahwa metode tersebut mampu menghasilkan kelompok pelanggan

yang memiliki karakteristik serupa sehingga dapat membantu perusahaan dalam menyusun strategi pelayanan, meningkatkan kualitas layanan, serta mendukung proses pengambilan keputusan secara lebih efektif.

Penelitian lain oleh (Gunawan Zain et al., 2025) memanfaatkan algoritma *K-Means* untuk mengelompokkan stok material di PDAM Kota Makassar Wilayah Pelayanan VI. Penelitian tersebut menggunakan metode *Elbow* dan *Silhouette Coefficient* untuk menentukan jumlah cluster optimal. Hasilnya diperoleh dua *cluster*, yaitu material dengan tingkat penggunaan tinggi dan rendah, dengan nilai *silhouette* sebesar 0,937, sehingga mampu membantu pengelolaan persediaan material menjadi lebih efisien.

Selanjutnya, Penelitian yang dilakukan oleh (Akbar et al., 2025) menerapkan algoritma *K-Means Clustering* untuk melakukan segmentasi tunggakan pelanggan pada Perumdam Tirta Anai. Penelitian tersebut menggunakan dua variabel utama, yaitu jumlah bulan tunggakan dan total tunggakan pelanggan, dengan tujuan mengelompokkan pelanggan ke dalam tiga kategori, yaitu tunggakan rendah, sedang, dan tinggi. Hasil pengelompokan dimanfaatkan sebagai dasar dalam menentukan strategi penagihan, monitoring pelanggan, serta penertiban sambungan guna meningkatkan pendapatan Perusahaan dan (Rudi Cahyono et al., n.d.) penelitian ini menerapkan algoritma *K-Means Clustering* untuk mengelompokkan pelanggan berdasarkan pola konsumsi air bersih. Penelitian ini bertujuan untuk menerapkan algoritma *K-Means Clustering* dalam mengelompokkan pelanggan berdasarkan tingkat penggunaan atau debit air. Hasil pengelompokan diharapkan dapat mengidentifikasi pelanggan dengan tingkat konsumsi air rendah, sedang, dan tinggi sehingga dapat menjadi dasar dalam pengelolaan data pelanggan, evaluasi pola penggunaan air, serta mendukung pengambilan keputusan yang lebih efektif dalam pengelolaan layanan air bersih.

Pembaruan penelitian ini terletak pada pemanfaatan algoritma *K-Means Clustering* untuk mengelompokkan pelanggan berdasarkan besarnya debit pemakaian air yang tercatat pada data penggunaan pelanggan. Berbeda dengan penelitian sebelumnya yang berfokus pada klasifikasi kualitas air dan penentuan pelanggan baru, penelitian ini menghasilkan segmentasi pelanggan berdasarkan pola konsumsi air. Informasi tersebut dapat dimanfaatkan oleh PDAM untuk mengidentifikasi pelanggan dengan konsumsi air tinggi, sedang, dan rendah sebagai dasar dalam pengelolaan distribusi air, evaluasi penggunaan, serta perencanaan layanan yang lebih efektif.

2. METODE

2.1. Jenis Penelitian

Jenis penelitian yang digunakan untuk studi kasus Algoritma metode *K-Means* yang berhasil peneliti buat dengan metode terapan dan peneliti

menghasilkan sesuatu yang dapat memecahkan masalah yang dihadapi.

2.2. Jenis dan Sumber Data

Jenis data yang digunakan adalah data kuantitatif. Dimana data kuantitatif adalah jenis informasi yang dapat dihitung, berupa angka atau nominal.

2.3. Metode Pengumpulan Data

berikut, Berdasarkan sumber information yang digunakan pada jurnal ini, maka metode pengumpulan data yang penulis gunakan adalah sebagai berikut,

- Wawancara, adalah metode pengumpulan data melalui kegiatan tanya jawab langsung dengan individual yang ada pada objek penelitian. Wawancara penulis lakukan dengan karyawan maupun pimpinan PDAM.
- Dokumentasi adalah metode pengumpulan data dengan mengumpulkan catatan - catatan atau document – dokumen penulis dapatkan dari metode wawancara, penulis kumpulkan menjadi satu sehingga menjadi sebuah dokumen yang siap digunakan untuk kepentingan penelitian.
- Studi Pustaka, adalah mempelajari karya ilmiah, buku ilmiah, dan sumber ilmiah lainnya yang sesuai dengan penelitian dan memiliki hubungan dengan masalah yang diteliti. Referensi ilmiah yang penulis gunakan adalah sumber-sumber yang terdapat dalam daftar kepustakaan.

2.4. Penerapan implementasi

k-means untuk clustering data pelanggan berdasarkan pemakaian air pada Perusahaan Umum Daerah Air Minum (PDAM) Perancangan Algoritma *K-Means Clustering* Penerapan implementasi antara lain :

- Menentukan nilai k sebagai jumlah dalam cluster yang ingin dibentuk.
- Bangkitkan k *centroid* (titik pusat cluster) awal secara random
- Hitung jarak setiap data ke masing-masing centroid
- Setiap data memilih *centroid* yang terdekat.
- Tentukan posisi *centroid* yang baru dengan cara menghitung nilai rata-rata dari data-data yang terletak pada *centroid* yang sama. 6. Kembali ke langkah-3 jika posisi *centroid* baru dengan *centroid* yang lama tidak sama.

3. PEMBAHASAN

3.1. Sistem yang sedang berjalan

Proses pengelompokan pelanggan berdasarkan pemakaian air pada Perumda Air Minum Tirta Giri Nata Kota Cirebon masih dilakukan secara manual dengan melihat data pemakaian pelanggan yang tersimpan pada arsip dan laporan bulanan. Kondisi tersebut menyebabkan proses identifikasi pelanggan dengan tingkat konsumsi air tinggi, sedang, dan rendah membutuhkan waktu yang lama serta berpotensi menimbulkan kesalahan dalam pengambilan keputusan. Selain itu, belum adanya sistem pengelompokan pelanggan secara otomatis menyebabkan pihak PDAM mengalami kesulitan

dalam melakukan evaluasi pola penggunaan air dan menentukan strategi pelayanan yang tepat.

3.2. Penerapan Metode K-Means

Penelitian ini menerapkan metode *K-Means Clustering* untuk mengelompokkan data pelanggan berdasarkan besarnya debit pemakaian air. Data yang digunakan terdiri dari 42 data pelanggan dengan atribut utama berupa jumlah pemakaian air (m3) diambil dari satu desa di Kabupaten Cirebon.

Tabel 2. Sampel data

No	ID Pelanggan	Nama Pelanggan	Pemakaian Air (m ³)
1	P001	Andi	6
2	P002	Budi	7
3	P003	Candra	8
4	P004	Dedi	9
5	P005	Eko	8
6	P006	Farhan	7
7	P007	Gilang	10
8	P008	Hadi	9
9	P009	Indra	8
10	P010	Joko	7
11	P011	Kurnia	8
12	P012	Lukman	9
13	P013	Maman	10
14	P014	Nanda	8
15	P015	Oki	7
16	P016	Putra	9
17	P017	Rudi	8
18	P018	Sandi	10
19	P019	Tono	7
20	P020	Yusuf	9
21	P021	Agus	15
22	P022	Bagus	16
23	P023	Dian	17
24	P024	Erwin	18
25	P025	Fajar	19
26	P026	Hendra	20
27	P027	Iwan	21
28	P028	Jamal	18
29	P029	Kiki	17
30	P030	Lina	19
31	P031	Mila	20
32	P032	Nina	22
33	P033	Oman	18
34	P034	Rina	17
35	P035	Sinta	18
36	P036	Teguh	34
37	P037	Ujang	35
38	P038	Wawan	36
39	P039	Yudi	37
40	P040	Zaki	38
41	P041	Asep	36
42	P042	Bayu	39

Tahapan penelitian dimulai dari pengumpulan data, pembersihan data (*data preprocessing*), penentuan jumlah *cluster*, proses perhitungan *centroid*, pengukuran jarak menggunakan *Euclidean Distance*, hingga proses iterasi dilakukan sampai tidak terjadi perpindahan anggota *cluster*.

Pada penelitian ini ditentukan tiga kelompok ($k=3$), yaitu:

- Cluster 1 (C1): pelanggan dengan tingkat pemakaian air rendah.
- Cluster 2 (C2): pelanggan dengan tingkat pemakaian air sedang.
- Cluster 3 (C3): pelanggan dengan tingkat pemakaian air tinggi.

Menentukan centroid awal *Centroid* awal dipilih secara acak dari data pemakaian pelanggan.

Tabel 3. Menentukan centroid awal

Cluster	Centroid Awal (m ³)
C1	8
C2	18
C3	36

Perhitungan menggunakan rumus *Euclidean Distance*:

$$d(x, c) = \sqrt{(x - c)^2} \quad (1)$$

Karena hanya menggunakan satu atribut (pemakaian air), maka perhitungannya menjadi:

Pelanggan P002 memiliki pemakaian 12 m³

Jarak ke C1

$$d = \sqrt{(12 - 8)^2} = 4$$

Jarak ke C2

$$d = \sqrt{(12 - 18)^2} = 6$$

Jarak ke C3

$$d = \sqrt{(12 - 36)^2} = 24$$

Karena jarak terkecil adalah 4, maka pelanggan P002 masuk ke *Cluster* 1.

Tabel 4. Hasil Iterasi ke 1

Cluster	Jumlah Pelanggan
C1	18
C2	17
C3	7

Centroid baru diperoleh dari rata-rata seluruh anggota *cluster*.

Cluster 1

$$\frac{8 + 7 + 9 + 10 + 8 + \dots}{18} = 8.5$$

Cluster 2

$$\frac{17 + 19 + 20 + 18 + \dots}{17} = 18.3$$

Cluster 3

$$\frac{35 + 36 + 40 + 34 + 37 + \dots}{7} = 36.4$$

Hasil pengelompokan menunjukkan bahwa sebagian besar pelanggan berada pada kategori pemakaian rendah dan sedang, sedangkan pelanggan dengan pemakaian tinggi hanya terdiri dari beberapa pelanggan yang memiliki konsumsi air di atas rata-rata. Pengelompokan tersebut memberikan informasi yang lebih jelas mengenai karakteristik pelanggan sehingga memudahkan pihak PDAM dalam melakukan pengawasan terhadap penggunaan air, evaluasi tarif, dan penyusunan kebijakan distribusi air yang lebih efektif.

3.3. Analisis Sistem

Sistem dibangun berbasis web menggunakan bahasa pemrograman PHP dan *database MySQL* dengan pendekatan metode pengembangan *Waterfall*. Sistem memiliki beberapa fitur utama, yaitu login pengguna, pengelolaan data pelanggan, pengelolaan data debit air, transaksi pembayaran, serta proses *clustering* menggunakan algoritma *K-Means*.

Implementasi sistem menghasilkan sebuah aplikasi yang mampu menampilkan hasil pengelompokan pelanggan secara otomatis. Admin dapat melihat data pelanggan berdasarkan kategori pemakaian air, sedangkan pelanggan dapat memperoleh informasi penggunaan air dan tagihan secara lebih cepat dan akurat.

3.4. Analisis Pembahasan dan Hasil

Hasil penelitian menunjukkan bahwa algoritma *K-Means* mampu melakukan segmentasi pelanggan berdasarkan pola konsumsi air dengan baik dengan hasil DBI 0.231 dan Hasil ini sejalan dengan penelitian (Hairani et al., 2026) yang menyatakan bahwa metode *K-Means* efektif digunakan dalam segmentasi pelanggan karena mampu menghasilkan kelompok pelanggan yang memiliki karakteristik serupa.

Perbedaan penelitian ini dengan penelitian sebelumnya terletak pada fokus pengelompokan yang dilakukan secara spesifik terhadap besarnya debit pemakaian air pelanggan dan implementasinya pada aplikasi berbasis web. Dengan adanya sistem ini, proses identifikasi pelanggan menjadi lebih cepat, akurat, dan dapat digunakan sebagai dasar pengambilan keputusan dalam meningkatkan kualitas pelayanan PDAM.

Pengelompokan pelanggan PDAM Tirta Giri Nata Kota Cirebon berdasarkan pemakaian air dapat dilihat pada tabel 7.

Tabel 5. Pengelompokan Pelanggan

cluster	Jumlah Pelanggan	Rata - Rata	Kategori
C1	20	8.5	Rendah
C2	15	18.3	Sedang
C3	7	36.4	Tinggi

Hal ini terjadi karena sebagian besar pelanggan merupakan pelanggan rumah tangga dengan kebutuhan air yang relatif rendah, sedangkan pelanggan dengan konsumsi tinggi berasal dari sektor usaha atau rumah tangga dengan jumlah

penghuni yang lebih banyak. Oleh karena itu, algoritma K-Means membentuk Cluster 1 sebagai kelompok dengan anggota terbanyak.

3.5. Perhitungan DBI

Tabel 6. Perhitungan DBI

cluster	Jumlah Pelanggan	Rata - Rata	Scatter
C1	20	8.5	0.85
C2	15	18.3	0.92
C3	7	36.4	0.76

a. Menghitung Antar Jarak Centroid (Euclied Distance)

Tabel 7. Menghitung Jarak Centroid

Pasangan Cluster	Jarak centroid
C1-C2	7.20
C1-C3	12.10
C3-C2	8.40

b. Menghitung Rij

$$(C1 - C2) R12 = \frac{0,85 + 0,92}{7,20} = \frac{1,77}{7,20} = 0,246$$

$$(C1 - C3) R12 = \frac{0,85 + 0,76}{12,10} = \frac{1,61}{12,10} = 0,133$$

$$(C2 - C3) R12 = \frac{0,92 + 0,76}{8,40} = \frac{1,68}{8,40} = 0,200$$

c. Menghitung Nilai Maksimum setiap Cluster

$$R1 = \max (0,246; 0,133)$$

$$R2 = \max (0,246; 0,200)$$

$$R3 = \max (0,133; 0,200)$$

d. Menghitung Davies-Bouldin Index

$$DBI = \frac{0,246 + 0,246 + 0,200}{3} = \frac{0,692}{3} = \mathbf{0,231}$$

4. KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa penerapan metode K-Means Clustering berhasil mengelompokkan pelanggan PDAM Tirta Giri Nata Kota Cirebon berdasarkan tingkat pemakaian air ke dalam tiga kategori, yaitu pelanggan dengan konsumsi air rendah, sedang, dan tinggi. Implementasi metode K-Means pada aplikasi berbasis web mampu membantu proses pengolahan data pelanggan secara lebih cepat, efektif, dan terkomputerisasi sehingga memudahkan pihak PDAM dalam melakukan evaluasi penggunaan air, pengelolaan distribusi, dan pengambilan keputusan. Untuk penelitian selanjutnya, metode K-Means dapat dikombinasikan dengan metode lain atau menggunakan jumlah data yang lebih besar agar hasil segmentasi pelanggan menjadi lebih optimal.

Selain berhasil mengelompokkan pelanggan berdasarkan tingkat pemakaian air, hasil validasi menggunakan Davies-Bouldin Index (DBI) menghasilkan nilai sebesar 0,231, yang menunjukkan bahwa kualitas cluster yang terbentuk sudah baik. Hal ini mengindikasikan bahwa algoritma K-Means efektif digunakan untuk melakukan segmentasi pelanggan berdasarkan pola konsumsi air.

PUSTAKA

- Akbar, S. C. D., Defit, S., & Hendrik, B. (2025). Segmentasi Tunggalan Pelanggan Menggunakan Algoritma K-Means Cluster pada Perusahaan Air Minum Daerah. *Jurnal Pustaka AI (Pusat Akses Kajian Teknologi Artificial Intelligence)*, 5(2), 349–355. <https://doi.org/10.55382/jurnalpustakaai.v5i2.1215>
- Annisa, K., Serasi Ginting, B., & Syari, M. A. (2022). PENERAPAN DATA MINING PENGELOMPOKAN DATA PENGGUNA AIR BERSIH BERDASARKAN KELUHANNYA MENGGUNAKAN METODE CLUSTERING PADA PDAM LANGKAT. *Jurnal Sistem Informasi Kaputama (JSIK)*, 6(2). www.kaputama.ac.id
- Gunawan Zain, S., Amalia Tahir, R., Akram Nur Risal, A., & Korespondensi, P. (2025). CLUSTERING STOK MATERIAL DI PDAM KOTA MAKASSAR WILAYAH PELAYANAN VI MENGGUNAKAN ALGORITMA K-MEANS. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, 12(4), 2355–2699.
- Hairani, H., Wahyudi, R., Gede Yogi Pratama, & M.Khaerul Ihsan. (2026). Application of RFM Model and K-Means Algorithm in Customer Loyalty Segmentation of Indonesian Regional Water Utility Company (PDAM). *International Journal of Engineering and Computer Science Applications (IJECSA)*, 5(1), 1–8. <https://doi.org/10.30812/ijecea.v5i1.6087>
- Herawati, V., Indah Sari, K., Kunci, K., & Bersih, A. (2023). ANALISIS KEBUTUHAN AIR BERSIH DI DESA LIMAU MANIS KECAMATAN TANJUNG MORAWA. *JTSIP*, 2(2), 2023. <https://jurnal.uisu.ac.id/index.php/JTSIP>
- Lestari, P. D. (2023). DATAMINING PADA PENJUALAN AIR BERSIH DI SPAM AKIDAH MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING MENGGUNAKAN RAPIDMINER. Dalam *Jurnal Mahasiswa Teknik Informatika* (Vol. 7, Nomor 1).
- Mikael Kia Mado, P. (2025). Implementasi Algoritma Clustering K-Means untuk Segmentasi Pelanggan di E-Commerce. Dalam *Jurnal Indonesia : Manajemen Informatika dan Komunikasi (JIMIK)* (Vol. 6, Nomor 3). <https://journal.stmiki.ac.id>
- Mutammimul Ula, Muthmainnah, Ridha Maulana, & Veri Ilhadi. (2023). Penerapan KNN Penentuan Pelanggan Baru PDAM dan Clustering K-Means Berdasarkan Wilayah. *Jurnal Informatika dan Teknologi Komputer*. <https://ejurnalunsam.id/index.php/jicom/>
- Nurdin, Suarna, N., & Prihartono, W. (2025). ALGORITMA REGRESI LINIER

- SEDERHANA UNTUK PREDIKSI
PENGUNAAN VOLUME AIR
BERDASARKAN JENIS PELANGGAN
PDAM. *Jurnal Kecerdasan Buatan dan
Teknologi Informasi*, 4(1), 43–52.
<https://doi.org/10.69916/jkbt.v4i1.187>
- Nurfiah, N., Satria, B., & Al Hafiz, M. (2025).
SEGMENTASI PELANGGAN LAYANAN
AIR BERSIH DI KAWASAN PEDESAAN
MENGUNAKAN K-MEANS
CLUSTERING BERDASARKAN METRIK
DURASI, KONSUMSI RATA-RATA, DAN
TOTAL VOLUME (DCV). *JUPI (Jurnal
Ilmiah Penelitian dan Pembelajaran
Informatika)*, 10(4), 3896–3906.
<https://doi.org/10.29100/jupi.v10i4.9371>
- Open Data Kabupaten Cirebon. (2025).
[https://opendata.cirebonkab.go.id/dataset/jum
lah-pelanggan-pdam-menurut-kecamatan-di-
kabupaten-cirebon](https://opendata.cirebonkab.go.id/dataset/jumlah-pelanggan-pdam-menurut-kecamatan-di-kabupaten-cirebon).
- Rudi Cahyono, G., Negeri Banjarmasin, P., &
Lambung Mangkurat, U. (t.t.). *KLASIFIKASI
KUALITAS AIR PDAM MENGGUNAKAN
ALGORITMA KNN DAN K-MEANS*.
- Sinaga, L., Ahmad, A., Safii, M., Tunas, S.,
Pematangsiantar, B., Supratman, J. W. R.,
Siantar, K., & Siantar, P. (2019).
PENERAPAN DATA MINING PADA
JUMLAH PELANGGAN PERUSAHAAN
AIR BERSIH MENURUT PROVINSI
MENGUNAKAN METODE K-MEANS
CLUSTERING. Dalam *119 JURNAL
RESISTOR* (Vol. 2, Nomor 2). Online.
[http://jurnal.stiki-
indonesia.ac.id/index.php/jurnalresistor](http://jurnal.stiki-indonesia.ac.id/index.php/jurnalresistor)
- Sri Tria Siska. (2016). ANALISA DAN
PENERAPAN DATA MINING UNTUK
MENENTUKAN KUBIKASI AIR
TERJUAL BERDASARKAN
PENGELOMPOKAN PELANGGAN
MENGUNAKAN ALGORITMA K-
MEANS CLUSTERING. *JURNAL
TEKNOLOGI INFORMASI & PENDIDIKAN*,
9.
- Yusa, H., Sudarma, M., & Pramaita, N. (2018). HR
Potensi Pelanggan Tunggal PDAM
Menggunakan Metode K-Medoids dengan
Optimasi Ant Colony Optimization (ACO).
Majalah Ilmiah Teknologi Elektro, 17(3),
353.
[https://doi.org/10.24843/mite.2018.v17i03.p0
8](https://doi.org/10.24843/mite.2018.v17i03.p08)