

Analisis Segmentasi Wilayah Pemetaan Distribusi Air Menggunakan Metode K-Means Clustering Pada Data Distribusi PDAM Tirtanadi

Kurnia Simbolon¹, Muhammad Syahril², Rita Hamdani³

^{1,2,3} Sistem Informasi, STMIK Triguna Dharma

Email: ¹kurniasimbolon52@gmail.com, ²muhammadsyahril.tgd@gmail.com, ³r1t4.hamdani@gmail.com

Email Penulis Korespondensi: kurniasimbolon52@gmail.com

Article History:

Received Jul 20th, 2026

Revised Aug 24th, 2026

Accepted Aug 31th, 2026

Abstrak

Air bersih merupakan kebutuhan dasar manusia yang sangat penting bagi kesehatan dan kualitas hidup. Di Indonesia, PDAM memiliki peran krusial dalam menyediakan layanan air bersih, termasuk PDAM Tirtanadi cabang Diski di Sumatera Utara. Namun, dalam pelaksanaannya, masih ditemukan berbagai permasalahan terkait ketidakseimbangan distribusi air antar wilayah pelanggan, yang berdampak pada ketidakpuasan masyarakat serta inefisiensi operasional. Penelitian ini bertujuan untuk menganalisis segmentasi wilayah distribusi air bersih dengan menerapkan metode K-Means Clustering, salah satu algoritma dalam Data Mining yang efektif dalam mengelompokkan data berdasarkan karakteristik tertentu. Metode ini dipilih karena kemampuannya dalam mengelompokkan wilayah dengan efisien dan otomatis, sehingga dapat membantu PDAM mengidentifikasi wilayah prioritas yang membutuhkan perhatian lebih dalam aspek distribusi. Melalui proses pengelompokan wilayah berdasarkan data distribusi air, penelitian ini menghasilkan pemetaan yang lebih akurat dan strategis. Hasil analisis diharapkan dapat menjadi dasar pengambilan keputusan dalam meningkatkan pemerataan distribusi air serta efisiensi operasional PDAM Tirtanadi cabang Diski. Dengan demikian, pendekatan berbasis data ini berpotensi memberikan kontribusi nyata dalam pengelolaan layanan publik yang lebih baik dan berkelanjutan.

Kata Kunci: Distribusi air bersih, PDAM Tirtanadi, K-Means Clustering, Data Mining, Segmentasi Wilayah.

Abstract

Clean water is a fundamental human need that is crucial for health and quality of life. In Indonesia, Regional Drinking Water Companies (PDAM) play a vital role in providing clean water services, including PDAM Tirtanadi, Diski branch, located in North Sumatra. However, in practice, various issues still arise, particularly the imbalance in water distribution across customer areas, which leads to public dissatisfaction and operational inefficiencies. This study aims to analyze the segmentation of clean water distribution areas using the K-Means Clustering method, one of the data mining algorithms known for its effectiveness in grouping data based on specific characteristics. This method is chosen for its ability to efficiently and automatically cluster regions, helping PDAM identify priority areas that require more attention in terms of distribution. By grouping regions based on water distribution data, this research produces more accurate and strategic mapping. The analysis results are expected to serve as a basis for decision-making to improve the equity of water distribution and operational efficiency at PDAM Tirtanadi, Diski branch. Thus, this data-driven approach has the potential to make a real contribution to better and more sustainable public service management.

Keywords: Clean water distribution, PDAM Tirtanadi, K-Means Clustering, Data Mining, Area Segmentation.

1. PENDAHULUAN

Air adalah sumber daya alam yang esensial bagi kehidupan manusia. Namun, di era modern ini, dunia dihadapkan pada krisis air yang semakin akut. Kekurangan sumber daya air bersih menjadi salah satu tantangan terbesar yang dihadapi umat manusia saat ini [1]. Ketersediaan air bersih merupakan salah satu indikator utama dalam meningkatkan kualitas hidup masyarakat serta menunjang pembangunan daerah secara berkelanjutan. Di tengah pertumbuhan penduduk yang pesat dan urbanisasi yang semakin meningkat, kebutuhan akan pasokan air bersih menjadi semakin kompleks. Kondisi ini mendorong PDAM Tirtanadi, sebagai penyedia utama layanan air bersih di wilayah Sumatera Utara, untuk terus berinovasi dalam pengelolaan distribusi air agar tercapai pemerataan pelayanan.

Air bersih merupakan kebutuhan dasar bagi kehidupan dan merupakan indikator utama dalam peningkatan kualitas hidup masyarakat. Di Indonesia, PDAM berperan penting sebagai penyedia layanan air bersih bagi masyarakat. Namun, meskipun keberadaan PDAM sudah cukup luas, masih terdapat permasalahan terkait distribusi air bersih yang dirasakan oleh setiap pelanggan. Distribusi ini tidak hanya berdampak pada kenyamanan dan kesehatan konsumen, tetapi juga memengaruhi efisiensi operasional dan keberlanjutan pengelolaan sumber daya air oleh PDAM. Permasalahan distribusi air bersih yang dialami oleh pelanggan PDAM perlu mendapatkan perhatian serius mengingat dampaknya yang luas terhadap kesejahteraan masyarakat. Berdasarkan pengamatan awal, terdapat indikasi bahwa meskipun PDAM memiliki kapasitas air yang cukup besar, distribusi air tidak selalu merata sehingga menyebabkan ketidakpuasan pelanggan. Analisis tentang distribusi jaringan air bersih dapat membuat suatu sistem distribusi air bersih yang sampai ke rumah-rumah masyarakat sekitar [2].

Kondisi geografis wilayah juga berperan signifikan dalam distribusi air bersih. Jarak yang jauh antara pelanggan dengan sumber air dapat menyebabkan penurunan volume air yang diterima, sekaligus meningkatkan biaya operasional PDAM. Selain faktor teknis dan geografis, aspek ekonomi juga turut mempengaruhi distribusi air bersih. Rata-rata tagihan air yang tinggi dapat menjadi refleksi dari biaya operasional yang meningkat akibat jarak distribusi yang jauh atau inefisiensi dalam sistem distribusi. Dengan demikian, integrasi antara data operasional PDAM dan metode analitik seperti K-Means Clustering memiliki potensi besar untuk meningkatkan efisiensi pelayanan dan mengurangi distribusi air bersih bagi pelanggan.

Pada penelitian ini, fokus diarahkan kepada analisis segmentasi wilayah distribusi air bersih di PDAM Tirtanadi cabang Diski melalui penerapan metode K-Means Clustering. Metode ini dipilih karena kemampuannya dalam mengelompokkan data berdasarkan kesamaan karakteristik secara otomatis dan efisien [3]. Dengan pendekatan tersebut, diharapkan dapat diperoleh pemetaan yang jelas mengenai kondisi distribusi air serta identifikasi wilayah-wilayah prioritas yang membutuhkan peningkatan pasokan. Sehingga, solusi yang ditawarkan tidak hanya bersifat deskriptif, melainkan juga memberikan rekomendasi strategis bagi pihak PDAM dalam mengoptimalkan operasional distribusi air.

Data mining merupakan proses dengan kecerdasan buatan, teknik statistik, matematika, dan machine learning untuk mengidentifikasi dan mengekstraksi suatu informasi yang bermanfaat dan pengetahuan terkait dari berbagai database yang besar. Dalam konteks segmentasi wilayah, Data Mining membantu perusahaan untuk menggali informasi yang lebih mendalam tentang perilaku pelanggan, preferensi, dan kebiasaan pembelian mereka, yang tidak selalu terlihat pada data permukaan [4].

Metode K-Means Clustering adalah salah satu algoritma yang populer digunakan dalam segmentasi wilayah. K-Means Clustering adalah teknik pengelompokan data non-hirarki yang memisahkan data ke dalam cluster, mengelompokkan data dengan fitur yang sama bersama-sama dan mengelompokkan data dengan karakteristik yang berbeda ke dalam kelompok yang berbeda. K-Means bekerja dengan cara mengelompokkan data ke dalam beberapa cluster berdasarkan kedekatan karakteristik antar data. Metode ini memiliki kelebihan dalam hal kecepatan dan kemudahan implementasi, sehingga cocok untuk digunakan dalam analisis segmentasi wilayah distribusi air bersih untuk lingkup PDAM Tirtanadi cabang Diski [5]. Dengan Data Mining metode K-Means Clustering diharapkan nantinya mendapatkan beberapa hal yang berguna untuk segmentasi pasar pada PDAM Tirtanadi cabang Diski.

2. METODOLOGI PENELITIAN

Metode penelitian dalam studi ini disusun sebagai serangkaian langkah ilmiah yang bertujuan untuk mengumpulkan dan menganalisis data dalam upaya menjawab permasalahan distribusi air bersih di PDAM Tirtanadi cabang Diski. Penelitian dilakukan secara langsung dengan menghimpun data distribusi air bersih pelanggan, yang kemudian dianalisis menggunakan metode K-Means Clustering. Rangkaian proses ini dimulai dari identifikasi variabel yang relevan, pengumpulan data, hingga analisis hasil pengelompokan wilayah. Melalui pendekatan ini, diharapkan dapat diperoleh pemetaan wilayah yang lebih akurat sebagai dasar penyusunan rekomendasi strategis untuk meningkatkan efektivitas distribusi air bersih di masa mendatang. Adapun tahapan yang dilakukan dalam penelitian guna untuk menyelesaikan permasalahan yang telah dijelaskan pada Bab sebelumnya termasuk pada bagian latar belakang permasalahan, mencakup pada:

1. Teknik Pengumpulan Data (*Data Collecting*)

Teknik pengumpulan data dalam penelitian ini merupakan proses untuk memperoleh informasi mengenai karakteristik, kondisi, aktivitas, atau fenomena yang berhubungan dengan topik penelitian. Pengumpulan data dilakukan secara langsung di PDAM Tirtanadi cabang Diski untuk mendukung analisis segmentasi wilayah distribusi air bersih menggunakan metode K-Means Clustering. Dalam pelaksanaannya, pengumpulan data dilakukan melalui dua pendekatan berikut:

a. Observasi

Metode pengumpulan data ini digunakan untuk memperoleh informasi yang relevan melalui peninjauan langsung ke PDAM Tirtanadi cabang Diski yang berlokasi di Jalan Medan - Binjai KM 10,5, Diski, Kabupaten Deli Serdang, Sumatera Utara. Observasi langsung dilakukan untuk memahami lebih dalam kondisi distribusi air bersih, pola distribusi yang terjadi di berbagai wilayah, serta tantangan yang dihadapi dalam pelayanan pelanggan. Pendekatan ini juga bertujuan untuk mengumpulkan data kualitatif yang tidak dapat sepenuhnya diperoleh hanya melalui dokumentasi internal, sehingga hasil penelitian menjadi lebih komprehensif.

Dalam penelitian ini, terdapat beberapa variabel yang digunakan untuk melakukan segmentasi wilayah distribusi air bersih di PDAM Tirtanadi cabang Diski, yaitu sebagai berikut:

1. Kapasitas Air
Jumlah maksimal air bersih yang tersedia untuk didistribusikan ke pelanggan dalam satu periode tertentu.
2. Volume Air Terdistribusi
Jumlah aktual air bersih yang berhasil disalurkan kepada pelanggan, yang mencerminkan tingkat efektivitas distribusi air.
3. Jarak dari Sumber Air
Jarak antara titik sumber air dengan lokasi pelanggan, yang berpotensi mempengaruhi kualitas dan kuantitas air yang diterima.
4. Rata-Rata Tagihan Air
Rata-rata biaya yang dibebankan kepada pelanggan dalam satu periode, yang dapat menjadi indikator terhadap volume pemakaian air dan efisiensi distribusi.

2.1 PDAM

Perusahaan Daerah Air Minum (PDAM) merupakan salah satu perusahaan daerah yang bergerak dalam pengolahan air baku menjadi air yang dapat dikonsumsi atau diminum. Kualitas air minum yang baik tentunya dapat kita lihat dari nilai persyaratan yang diberikan atau ditunjukkan dalam tabel persyaratan untuk air minum. Jika nilai parameter air berada dibawah baku mutu persyaratan Menkes berarti air tersebut memenuhi standar untuk dikonsumsi. [6].

PDAM adalah layanan penggunaan air bersih dengan sistem berlangganan dan membayar setiap bulan untuk sejumlah pemakaian air yang digunakan, dengan melihat meteran air yang biasa diletakkan di depan rumah pelanggan. Meteran air ini menghubungkan pipa instalasi PDAM lingkungan menuju pipa instalasi air di dalam rumah [7].

Distribusi air merupakan proses penyaluran air dari sumber (mata air, sungai, danau atau sumur dalam) menuju konsumen melalui jaringan perpipaan. Tantangan utama dalam distribusi air antara lain adalah kehilangan air (non-revenue water), kebocoran jaringan, keterbatasan kapasitas produksi, serta jarak dan kondisi topografi. Ketersediaan dan kelancaran distribusi air sangat tergantung pada faktor teknis, geografis, dan manajemen. Oleh karena itu, analisis terhadap wilayah-wilayah yang mengalami distribusi air menjadi sangat penting dalam perencanaan distribusi yang lebih efektif dan efisien.

2.2 Distribusi air

Air adalah sumber daya alam yang esensial bagi kehidupan manusia. Namun, di era modern ini, dunia dihadapkan pada krisis air yang semakin akut. Kekurangan sumber daya air bersih menjadi salah satu tantangan terbesar yang dihadapi umat manusia saat ini. Faktor Krisis air bisa disebabkan oleh yang pertama adalah adanya perubahan iklim. terjadinya kekeringan yang parah bisa menyebabkan kurangnya pasokan air bersih. lalu kedua adalah pencemaran air. air-air yang tercemar oleh limbah industri, pertanian dan rumah tangga ini menyebabkan air menjadi tidak layak untuk di konsumsi. dan yang ketiga adalah pertumbuhan populasi yang pesat meningkatkan peminatan air bersih tapi pasokan air tetap saja terbatas. Distribusi air bersih adalah kondisi di mana ketersediaan air tidak mencukupi untuk memenuhi kebutuhan masyarakat baik dalam kuantitas maupun kualitas [8]. Menurut *Standardized Precipitation And Evapotranspiration Index* (SPEI), distribusi air dapat disebabkan oleh beberapa faktor, antara lain [9]:

1. Peningkatan Permintaan Air (*Over-consumption of Water*). Konsumsi air yang berlebihan menjadi salah satu penyebab utama distribusi air saat ini.
2. Pertumbuhan Penduduk yang Terus-Menerus. Bertambahnya populasi dunia meningkatkan kebutuhan terhadap sumber daya air.
3. Pemanasan Global (*Global Warming*). Peningkatan evapotranspirasi, yaitu penguapan air dari tanah dan tumbuhan ke atmosfer. Meningkatnya kapasitas atmosfer untuk menahan air, menyebabkan redistribusi air dari permukaan bumi ke atmosfer.

4. Perubahan Iklim (*Climate Change*). Mengakibatkan cuaca ekstrem lebih sering, lebih lama, dan lebih intens (misalnya kekeringan parah). Menyebabkan siklus hidrologi menjadi lebih intens, mempercepat kehilangan massa air dari daratan.
5. Manajemen Air yang Buruk (*Poor Water Management*). Contohnya di Yaman, manajemen air yang tidak efektif memperparah krisis air selama musim kemarau.
6. Dampak *Drought*/Kekeringan. Kekeringan yang semakin parah mengurangi ketersediaan air di daratan, baik permukaan maupun bawah tanah. Wilayah dengan permintaan air tinggi dan pasokan rendah (seperti area pertanian, pertambangan, atau perkotaan) lebih rentan terdampak.
7. Ketergantungan pada Air Tanah. Ketergantungan tinggi pada air tanah, sementara recharge air tanah tidak mencukupi, menyebabkan kekurangan air jangka panjang.
8. Perubahan Aliran dan Cadangan Air. Penurunan aliran air permukaan dan perubahan cadangan air bawah tanah akibat aktivitas manusia dan perubahan iklim.

2.3 Data Mining

Data Mining adalah suatu proses penambangan atau penemuan informasi baru yang dilakukan dengan cara mencari sebuah pola atau aturan tertentu dari sejumlah data yang menumpuk dan dikatakan data besar. *Data mining* adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam *database*. *Data mining* merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan di dalam *database* besar [10].

Data Mining merupakan serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. *Data Mining* mulai ada sejak 1990-an sebagai cara yang benar dan tepat untuk mengambil pola dan informasi yang digunakan untuk menemukan hubungan antara data untuk melakukan pengelompokan ke dalam satu atau lebih *cluster* sehingga objek-objek yang berada dalam satu *cluster* akan mempunyai kesamaan yang tinggi antara satu dengan lainnya. *Data Mining* merupakan bagian dari proses penemuan pengetahuan dari basis data *Knowledge Discovery in Databases* [11].

Menurut penulis, *Data Mining* adalah proses untuk menggali informasi tersembunyi yang bernilai dari sekumpulan data besar dengan menggunakan teknik analisis statistik, pemrograman, dan kecerdasan buatan. Tujuan utama dari *Data Mining* bukan sekedar mengumpulkan data, tetapi untuk memahami pola, tren, atau hubungan yang bisa dijadikan dasar pengambilan keputusan yang lebih tepat dan strategis.

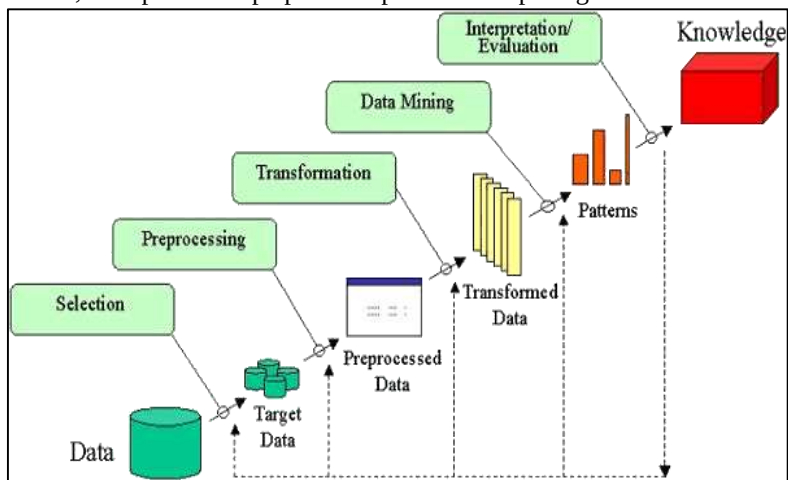
Data Mining dapat dianggap sebagai proses penyaringan informasi yang berharga dari “tambang data” yang sangat luas. Proses ini melibatkan beberapa tahapan penting seperti *data cleaning*, *data integration*, *data selection*, *data transformation*, *data mining*, dan *evaluation/presentation*. Setiap tahapan tersebut berkontribusi dalam membentuk hasil akhir berupa pengetahuan atau informasi yang dapat digunakan untuk mendukung proses pengambilan keputusan, baik dalam bidang bisnis, kesehatan, pendidikan, maupun sektor lainnya. *Data Mining* bukan hanya alat teknis, tetapi juga pendekatan strategis yang memungkinkan organisasi untuk memahami perilaku, memprediksi tren masa depan, dan mengoptimalkan proses operasional.

2.3.1 Teknik Data Mining

Secara umum, teknik data mining dapat dikelompokkan menjadi beberapa kategori utama berdasarkan jenis analisis yang dilakukan dan tujuan pengolahan datanya. Berikut adalah beberapa kelompok utama dalam data mining:

1. Deskripsi
Deskripsi merupakan cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data yang dimiliki.
2. Estimasi
Estimasi hampir sama dengan klasifikasi, hanya saja nilai peubah atau variable target estimasi lebih ke arah data angka atau numeric.
3. Prediksi
Prediksi adalah suatu cara dalam menerka/menebak sebuah nilai yang belum diketahui sebelumnya.
4. Klasifikasi
Dalam klasifikasi terdapat target variable bertipe kategori, contohnya adalah penggolongan pendapatan yang dapat dipisahkan kedalam tiga kategori, yaitu tinggi, sedang, dan rendah.
5. Pengklasteran
Pengklasteran adalah pengelompokan data *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek atau titik-titik yang memiliki kemiripan satu dengan yang lainnya.
6. Asosiasi
Asosiasi bertugas menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja. Suatu sifat yang menjadi sebuah ciri- ciri dari suatu objek disebut dengan karakteristik.

Pada proses *Data Mining* yang biasanya disebut *knowledge discovery database (KDD)*. *Knowledge Discovery Databases (KDD)* adalah penerapan metode saintifik pada *Data Mining*. Dalam penjelasan ini *Data Mining* merupakan satu langkah dari proses KDD, terdapat beberapa proses seperti terlihat pada gambar dibawah 2.1 [12].



Gambar 1 Proses *Knowledge Discovery Database*

Penjelasan pada gambar proses *Knowledge Discovery Database (KDD)* terdapat beberapa proses dari tahap menjalankan KDD yaitu sebagai berikut:

1. Seleksi Data (*Selection*)
Selection berarti proses seleksi/pemilihan dari data yang dilakukan sebelum menuju pada tahap penelusuran dalam *Knowledge Discovery Database (KDD)* dimulai dengan ketentuan data dipilih berdasarkan tujuan. Data yang di kumpulkan akan digunakan dalam pemrosesan yang dilakukan oleh *Data Mining*.
2. Pemilihan Data (*Preprocessing/Cleaning*)
Proses *preprocessing* adalah suatu proses yang meliputi antara lain menghapus atau membuang data ganda yang tidak digunakan, memeriksa data yang dianggap tidak konsisten, dan memperbaiki kesalahan pada data, seperti kesalahan penulisan (*tipografi*).
3. Transformasi (*Transformation*)
Pada fase ini yang dilakukan adalah mengubah suatu bentuk data yang belum mempunyai beberapa entitas yang jelas ke dalam data yang siap untuk dilakukan proses *Data Mining*.
4. *Data Mining*
Pada proses ini, yang dilakukan adalah melakukan penerapan algoritma atau metode pencarian pengetahuan dari data yang dihasilkan pada proses transformasi.
5. Interpretasi/Evaluasi (*Interpretation/Evaluation*)
Pada fase ini yang paling terakhir ini, proses yang dilakukan adalah proses membentuk sebuah output atau hasil yang mudah dimengerti dan berbentuk sebuah informasi yang bermanfaat [13].

2.4 Metode K-Means

K-Means merupakan salah satu algoritma dalam *Data Mining* yang dapat digunakan untuk melakukan pengelompokan/*Clustering* suatu data. Ada banyak pendekatan untuk membuat *cluster*, diantaranya adalah membuat aturan yang mendikte keanggotaan dalam group yang sama berdasarkan tingkat persamaan diantara anggota-anggotanya. *K-Means* adalah suatu teknik pengelompokan data yang mana keberadaan tiap-tiap titik data dalam suatu *cluster* ditentukan oleh derajat keanggotaan [15]. Pendekatan lainnya adalah dengan membuat sekumpulan fungsi yang mengukur beberapa properti dari pengelompokan tersebut sebagai fungsi dari beberapa parameter dari sebuah *Clustering*. Metode *K-Means* adalah metode yang termasuk dalam algoritma *Clustering* berbasis jarak yang membagi data ke dalam sejumlah *cluster* dan algoritma ini hanya bekerja pada atribut numerik.

Metode *K-Means* termasuk dalam *partitioning Clustering* yang memisahkan data ke daerah bagian yang terpisah. Dalam metode *K-Means* setiap data harus termasuk ke *cluster* tertentu pada suatu tahapan proses, pada tahapan proses berikutnya dapat berpindah ke *cluster* yang lain [16].

Pengelompokan data dengan metode *K-Means* dilakukan dengan algoritma [17]:

1. Tentukan jumlah kelompok.
2. Alokasikan data ke dalam kelompok secara acak
3. Hitung pusat kelompok (*centroid*/rata-rata) dari data yang ada di masing-masing kelompok. Lokasi *centroid* setiap kelompok diambil dari rata-rata (*mean*) semua nilai data pada setiap fiturnya. Jika M menyatakan jumlah data dalam sebuah kelompok, i menyatakan fitur ke-i dalam sebuah kelompok, dan p menyatakan dimensi data, maka persamaan untuk menghitung *centroid* fitur ke-i digunakan persamaan 1.

$$C_i = \frac{1}{M} \sum_{j=1}^m X_j \dots\dots\dots (1)$$

persamaan 1 dilakukan sebanyak p dimensi dari i=1 sampai dengan i=p.

4. Alokasikan masing-masing data ke *centroid*/rata-rata terdekat. Ada beberapa cara yang dapat dilakukan untuk mengukur jarak data ke pusat kelompok, diantaranya adalah *Euclidean* . Pengukuran jarak pada ruang jarak (*distance space*) *Euclidean* dapat dicari menggunakan persamaan 2.

$$d = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} \dots\dots\dots (2)$$

Pengalokasian kembali data ke dalam masing-masing kelompok dalam metode *K-Means* didasarkan pada perbandingan jarak antara data dengan *centroid* setiap kelompok yang ada . Data dialokasikan ulang secara tegas ke kelompok yang mempunyai *centroid* dengan jarak terdekat dari data tersebut. Pengalokasian data ini menurut *MacQueen* (1967) dapat ditentukan menggunakan persamaan 3.

$$a = \begin{cases} 1 & d = \min\{D(x, c)\} \\ 0 & \text{Lainnya} \end{cases} \dots\dots\dots (3)$$

a adalah nilai keanggotaan titik xi ke pusat kelompok c1, d adalah jarak terpendek dari data xi ke K kelompok setelah dibandingkan, dan c1 adalah *centroid* (pusat kelompok) ke-1.

$$J = \sum_{i=1}^n \sum_{k=1}^K a_{ik} D(x, c)^2 \dots\dots\dots (4)$$

n adalah jumlah data, k adalah jumlah kelompok, ai1 adalah nilai keanggotaan titik data xi ke kelompok cl yang diikuti. a mempunyai nilai 0 atau 1. Apabila data merupakan anggota suatu kelompok, nilai ai1 = 1. Jika tidak, nilai ai1 = 0

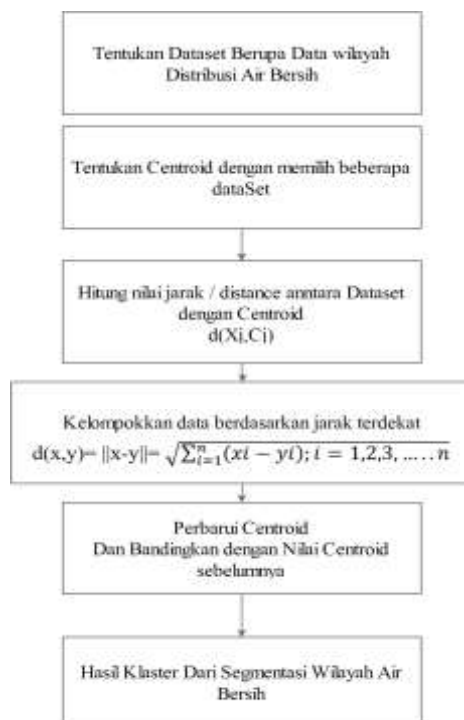
5. Menghitung nilai rasio antara *Between-Cluster Variation* (BCV) dan *Within-Cluster Variation* (WCV)

$$BCV = d(m_1 - m_2) + d(m_1 - m_3) + d(m_2 - m_3) \dots\dots\dots (5)$$

Kembali ke langkah 3, apabila masih ada data yang berpindah kelompok atau apabila ada perubahan nilai *centroid* di atas nilai ambang yang ditentukan, atau apabila perubahan nilai pada fungsi objektif yang digunakan masih di atas nilai ambang yang ditentukan [17].

3. HASIL DAN PEMBAHASAN

Kerangka kerja metode, atau yang biasa disebut sebagai *algoritma* sistem, adalah langkah-langkah yang dilakukan oleh sistem dalam memproses dan menyelesaikan permasalahan yang dihadapi. Berikut ini adalah alur kerja atau langkah-langkah pemecahan masalah dengan menggunakan metode *K-Means Clustering*, yang akan diterapkan untuk analisis segmentasi wilayah distribusi air bersih pada PDAM Tirtanadi cabang Diski. Kerangka kerja dari metode *K-Means* adalah sebagai berikut:



Gambar 2 Kerangka Kerja metode *K-Means*

Tahap persiapan ini adalah langkah awal yang penting dalam mempersiapkan data yang dibutuhkan untuk proses Data Mining. Tujuan dari tahap ini adalah untuk memastikan bahwa data yang digunakan relevan dengan permasalahan yang akan dipecahkan, terjamin kebenarannya, dan tersedia dalam format yang sesuai. Persiapan ini melibatkan serangkaian kegiatan yang harus dilakukan sebelum proses pengumpulan dan pengolahan data dimulai. Pada tahap ini, langkah-langkah penting disusun untuk memastikan efisiensi waktu dan pekerjaan. Kegiatan yang termasuk dalam tahap persiapan ini antara lain:

1. Perumusan dan Identifikasi Masalah
Langkah pertama adalah merumuskan dan mengidentifikasi masalah yang akan dipecahkan dalam penelitian, sehingga tujuan dan fokus analisis menjadi jelas.
2. Observasi dan Peninjauan Langsung di Lokasi Masalah
Dilakukan pengamatan langsung di lapangan untuk memahami kondisi nyata dan menggali informasi terkait masalah yang dihadapi.
3. Penentuan Kebutuhan Data, Sumber Data, dan Pengadaan Administrasi Perencanaan Data
Tahap ini melibatkan penentuan jenis data yang diperlukan, sumber data yang akan digunakan, serta perencanaan administrasi
4. Perencanaan Jadwal dan Desain Perencanaan
Menyusun rencana jadwal dan desain yang jelas untuk pengumpulan data, analisis, dan implementasi solusi dalam penelitian ini, agar semua kegiatan dapat berjalan secara terstruktur dan efisien.

Persiapan yang teliti sangat penting untuk mencegah pekerjaan yang harus dilakukan berulang kali, sehingga proses pengumpulan data dapat berjalan secara maksimal. Berikut ini adalah contoh sampel data yang diperoleh dari PDAM Tirtanadi cabang Diski yang digunakan dalam penelitian ini:

Tabel 1 Tabel Normalisasi Data

No	Wilayah	X ₁	X ₂	X ₃	X ₄
1	Jalan Asoka Komplek BTN Suka Maju	10,000	9,500	6	42,000
2	Gaperta VIII	9,000	8,500	7	43,000
3	Kelambir	8,000	7,200	8	44,000
4	Gatot Subroto	12,000	11,000	5	40,000
5	Amal Luhur	11,000	10,000	6	41,000
6	Jalan Budi Luhur	10,500	10,000	7	42,000
7	Kapten Muslim	11,500	10,500	6	40,000
8	Jalan Binjai KM 12	12,500	12,000	10	43,000
9	Setia Luhur	10,000	9,000	7	42,000
10	Bakti Kompleks	9,500	8,500	8	44,000
11	Akasia Raya	11,000	10,500	6	41,000
12	Komp. Padang Hijau	10,800	10,000	7	42,000
13	Bakti Luhur	11,200	10,700	6	41,000
14	Brawijaya Komplek Kodam	12,000	11,000	5	40,000
15	Jalan Merdeka	11,000	10,000	6	41,000
16	Komplek Sentosa	10,500	9,800	7	42,000
17	Jalan Merpati	9,800	9,200	7	43,000
18	Gatot Subroto II	12,200	11,200	5	40,000
..	..	..	..	..	..
50	Komplek Sentra Agul (Sei Agul)	9,300	9,000	5	42,000

3.1 Proses Perhitungan K-Means

Tahap ini dilakukan penerapan algoritma *K-Means* dengan rumus :

$$d(x,y)= \|x-y\|=\sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1,2,3, \dots n$$

Penerapan jumlah *cluster* (K) yaitu 3 *cluster*. Setelah menetapkan jumlah *cluster*, Tentukan titik pusat awal *cluster* (*centroid*). Berikut ini titik *centroid* yang telah dipilih:

Tabel 2 Tabel Data *Centroid* Awal

Centroid	Wilayah	X ₁	X ₂	X ₃	X ₄	Keterangan
Centroid 1	Gatot Subroto	12,000	11,000	5	40,000	Area dengan kinerja baik (rendah distribusi)
Centroid 2	Jalan Budi Luhur	10,500	10,000	7	42,000	Area dengan kinerja sedang (distribusi moderat):

Centroid 3	Babura (Padang Bulan)	7,500	7,000	12	48,000	Area dengan kinerja rendah (tinggi distribusi)
------------	-----------------------------	-------	-------	----	--------	---

Hitung jarak data ke *centroid* menggunakan rumus *Euclidean*, data tersebut akan ditetapkan sebagai anggota dari *cluster* terdekatnya. Menghitung *Distance* (jarak) antara *variable* dari setiap sampel data dengan *centroid* yaitu :

Tabel 3 Jarak Dengan Jalan Asoka Komplek BTN Suka Maju

No.	Nama Wilayah	X ₁	X ₂	X ₃	X ₄
1	Jalan Asoka Komplek BTN Suka Maju	10,000	9,500	6	42,000

a. Dengan *Centroid* 1

- Jarak antara data Jalan Asoka Komplek BTN Suka Maju dengan titik C1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(10000 - 12000)^2 + (9500 - 11000)^2 + (6 - 5)^2 + (42000 - 40000)^2}$$

$$= 3201.56$$

b. Dengan *Centroid* 2

- Jarak antara data Jalan Asoka Komplek BTN Suka Maju dengan titik C2

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(10000 - 10500)^2 + (9500 - 10000)^2 + (6 - 7)^2 + (42000 - 42000)^2}$$

$$= 707.11$$

c. Dengan *Centroid* 3

- Jarak antara Data Jalan Asoka Komplek BTN Suka Maju dengan titik C3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(10000 - 7500)^2 + (9500 - 7000)^2 + (6 - 12)^2 + (42000 - 48000)^2}$$

$$= 6964.20$$

Lakukan proses perhitungan yang sama sampai dengan objek ke 50. Adapun hasil dari perhitungan iterasi 1 dapat dilihat pada tabel dibawah ini. Dimana jarak terdekat dilihat dari perhitungan yang paling dekat ke pusat *cluster*. Untuk lebih lengkapnya jarak pada setiap baris data, hasilnya seperti pada tabel berikut :

Tabel 4 Iterasi 1

No	Nama Wilayah	X1	X2	X3	Min	Cluster
1	Jalan Asoka Komplek BTN Suka Maju	3201.56	707.11	6964.20	707.11	Cluster 2
2	Gaperta VIII	4924.43	2345.21	5431.39	2345.21	Cluster 2
3	Kelambir	6814.69	4253.23	4036.09	4036.09	Cluster 3
4	Gatot Subroto	0.00	2692.58	10012.49	0.00	Cluster 1
5	Amal Luhur	1732.05	1118.03	8381.53	1118.03	Cluster 2
6	Jalan Budi Luhur	2692.58	0.00	7348.47	0.00	Cluster 2
7	Kapten Muslim	707.11	2291.29	9604.69	707.11	Cluster 1
8	Jalan Binjai KM 12	3201.57	3000.00	8660.25	3000.00	Cluster 2
9	Setia Luhur	3464.10	1118.03	6800.74	1118.03	Cluster 2
10	Bakti Kompleks	5338.54	2692.58	4716.99	2692.58	Cluster 2
11	Akasia Raya	1500.00	1224.75	8573.22	1224.75	Cluster 2
12	Komp. Padang Hijau	2537.72	300.00	7475.96	300.00	Cluster 2
13	Bakti Luhur	1315.30	1407.13	8739.57	1315.30	Cluster 1

14	Brawijaya Komplek Kodam	0.00	2692.58	10012.49	0.00	Cluster 1
15	Jalan Merdeka	1732.05	1118.03	8381.53	1118.03	Cluster 2
16	Komplek Sentosa	2773.09	200.00	7269.11	200.00	Cluster 2
...	...	...	...	...	...	...
50	Komplek Sentra Agul (Sei Agul)	3910.24	1562.05	6575.72	1562.05	Cluster 2

Dari Tabel 3.5 didapat keanggotaan sebagai berikut :

1. *Cluster 1* = { Gatot Subroto, Kapten Muslim, Bakti Luhur, Brawijaya Komplek Kodam, Gatot Subroto II }
2. *Cluster 2* = { Jalan Asoka Komplek BTN Suka Maju, Gaperta VIII, Amal Luhur, Jalan Budi Luhur, Jalan Binjai KM 12, Setia Luhur, Bakti Kompleks, Akasia Raya, Komp. Padang Hijau, Jalan Merdeka, Komplek Sentosa, Jalan Merpati, Komplek Sukses, Jalan Pahlawan, Telaga Sari (Sunggal), Sei Mencirim (Sunggal), Sunggal Kanan (Sunggal), Medan Krio (Sunggal), Sumber Melati Diski (Sunggal), Mulioarjo (Sunggal), Kesawan (Sei Agul), Silalas (Sei Agul), Glugur Kota (Sei Agul), Pulo Brayan Kota (Sei Agul), Karang Berombak (Sei Agul), Sei Agul, Komplek Sentra Agul (Sei Agul)}
3. *Cluster 3* = { Kelambir, Suka Maju (Sunggal), Sei Beras Sekata (Sunggal), Tanjung Selamat (Sunggal), Paya Geli (Sunggal), Puji Mulyo (Sunggal), Sei Semayang (Sunggal), Serba Jadi (Sunggal), Kampung Lalang (Sunggal), Purwodadi (Sunggal), Babura (Padang Bulan), Darat (Padang Bulan), Merdeka (Padang Bulan), Padang Bulan, Petisah Hulu (Padang Bulan), Titi Rantai (Padang Bulan), Komplek Merpati (Padang Bulan), Komplek Bunga (Padang Bulan)}

Maka dari hasil *cluster* diatas kemudian dicari pembaharuan *centroid* dari rata-rata *cluster* yang dibentuk.

Pembaruan *centroid* dari hasil *cluster* dilakukan seperti berikut :

- a. *Cluster 1*
= rata-rata (*Cluster 1*)
= (11780.00; 10880.00; 5.40; 40200.00)
- b. *Cluster 2*
= rata-rata (*Cluster 2*)
= (9892.59; 9374.07; 6.85; 42333.33)
- c. *Cluster 3*
= rata-rata (*Cluster 3*)
= (8088.89; 7588.89; 11.56; 46500.00)

Ada beberapa iterasi lagi yang tidak dicantumkan pada perhitungan ini, untuk itu dibawah ini adalah deskripsi dari hasil iterasi ke 4. Karena pada iterasi ke-4 nilai *centroid* tidak berubah dari nilai *centroid* sebelumnya, maka algoritma dihentikan. Maka dari itu sesuai dengan perhitungan algoritma *K-Means* maka diperoleh hasil akhir sebagai berikut.

1. Hasil *Cluster 1* yang menunjukkan segmentasi dengan kinerja baik (minim distribusi)

Tabel 5 Hasil *Cluster 1*

No	Wilayah	X ₁	X ₂	X ₃	X ₄	Keterangan
1	Gatot Subroto	12,000	11,000	5	40,000	
2	Amal Luhur	11,000	10,000	6	41,000	
3	Kapten Muslim	11,500	10,500	6	40,000	Minim
4	Jalan Binjai KM 12	12,500	12,000	10	43,000	Distribusi
5	Akasia Raya	11,000	10,500	6	41,000	
6	Komp. Padang Hijau	10,800	10,000	7	42,000	
7	Bakti Luhur	11,200	10,700	6	41,000	
8	Brawijaya Komplek Kodam	12,000	11,000	5	40,000	Minim
9	Jalan Merdeka	11,000	10,000	6	41,000	Distribusi
10	Gatot Subroto II	12,200	11,200	5	40,000	
11	Jalan Pahlawan	10,800	10,500	7	41,000	

2. Hasil *Cluster 2* yang menunjukkan segmentasi area dengan kinerja sedang (distribusi moderat) yang berarti terkadang sering mengalami distribusi.

Tabel 6 Hasil *Cluster 2*

No	Wilayah	X ₁	X ₂	X ₃	X ₄	Keterangan
1	Jalan Asoka Komplek BTN Suka Maju	10,000	9,500	6	42,000	Distribusi Moderat
2	Gaperta VIII	9,000	8,500	7	43,000	

3	Jalan Budi Luhur	10,500	10,000	7	42,000
4	Setia Luhur	10,000	9,000	7	42,000
5	Bakti Kompleks	9,500	8,500	8	44,000
6	Komplek Sentosa	10,500	9,800	7	42,000
7	Jalan Merpati	9,800	9,200	7	43,000
8	Komplek Sukses	10,200	10,000	6	42,000
9	Telaga Sari (Sunggal)	9,000	8,500	9	43,000
10	Sei Mencirim (Sunggal) Sunggal Kanan	9,200	8,700	10	44,000
11	(Sunggal)	9,000	8,600	9	44,000
12	Medan Krio (Sunggal) Sumber Melati Diski	9,500	9,000	8	42,000
13	(Sunggal)	8,800	8,300	11	44,000
14	Mulioorejo (Sunggal)	9,000	8,600	10	44,000
15	Kesawan (Sei Agul)	9,500	9,200	4	42,000
16	Silalas (Sei Agul)	9,300	9,000	5	42,000
17	Glugur Kota (Sei Agul) Pulo Brayan Kota (Sei	9,400	9,100	4	41,000
18	Agul)	9,600	9,300	5	43,000
19	Karang Berombak (Sei Agul)	9,500	9,200	4	42,000
20	Sei Agul	9,400	9,100	4	41,000
21	Komplek Sentra Agul (Sei Agul)	9,300	9,000	5	42,000

3. Hasil *Cluster* 3 yang menunjukkan segmentasi area dengan kinerja rendah (Tinggi Distribusi)

Tabel 7 Hasil *Cluster* 3

No	Wilayah	X ₁	X ₂	X ₃	X ₄	Keterangan
1	Kelambir	8,000	7,200	8	44,000	Tinggi Distribusi
2	Suka Maju (Sunggal) Sei Beras Sekata	8,800	8,200	10	45,000	
3	(Sunggal)	8,500	8,000	11	46,000	
4	Tanjung Selamat (Sunggal)	8,700	8,300	10	45,000	
5	Paya Geli (Sunggal)	8,800	8,400	11	46,000	
6	Puji Mulyo (Sunggal)	8,600	8,200	10	45,000	
7	Sei Semayang (Sunggal)	8,400	8,000	12	47,000	
8	Serba Jadi (Sunggal) Kampung Lalang	8,200	7,800	12	46,000	
9	(Sunggal)	8,500	8,000	11	45,000	
10	Purwodadi (Sunggal)	8,700	8,300	11	46,000	
11	Babura (Padang Bulan)	7,500	7,000	12	48,000	
12	Darat (Padang Bulan)	7,700	7,200	13	48,000	
13	Merdeka (Padang Bulan)	7,300	6,800	12	47,000	
14	Padang Bulan Petisah Hulu (Padang	7,600	7,000	13	48,000	
15	Bulan)	7,800	7,200	14	49,000	
16	Titi Rantai (Padang Bulan)	7,500	7,000	13	47,000	
17	Komplek Merpati (Padang Bulan)	7,600	7,100	12	48,000	

Komplek Bunga (Padang					
18	Bulan)	7,400	6,900	13	47,000

4. KESIMPULAN

Berdasarkan hasil penelitian dan pengembangan sistem segmentasi wilayah distribusi air bersih di PDAM Tirtanadi Cabang Diski menggunakan metode K-Means Clustering, diperoleh beberapa kesimpulan penting. Metode K-Means berhasil diterapkan secara efektif untuk menganalisis segmentasi wilayah distribusi air bersih dengan mengelompokkan wilayah ke dalam tiga segmen berdasarkan kesamaan karakteristik data, sehingga memudahkan dalam memahami dan mengukur distribusi air secara objektif. Faktor-faktor utama yang mempengaruhi segmentasi ini meliputi kapasitas air yang tersedia, volume air yang terdistribusi, jarak wilayah dari sumber air, serta rata-rata tagihan air. Keempat variabel tersebut terbukti berkontribusi signifikan dalam membentuk pola segmentasi wilayah. Pola distribusi yang ditemukan menunjukkan perbedaan yang jelas antar segmen, di mana wilayah yang dekat dengan sumber air dan memiliki kapasitas tinggi cenderung tergolong dalam kelompok yang tidak mengalami gangguan distribusi, sementara wilayah yang jauh, berkapasitas rendah, dan volume distribusi kecil cenderung mengalami masalah distribusi. Aplikasi berbasis web yang dikembangkan menggunakan Laravel 11 mampu menjalankan proses K-Means secara otomatis dan cepat, dengan antarmuka yang sederhana serta fitur input data, perhitungan, dan laporan hasil clustering yang sistematis. Implementasi aplikasi ini di lingkungan nyata menunjukkan kinerja optimal dalam membantu PDAM Tirtanadi melakukan segmentasi wilayah, sehingga hasil klasterisasi dapat dijadikan dasar dalam pengambilan keputusan operasional dan perencanaan distribusi air ke depan.

UCAPAN TERIMAKASIH

Terima Kasih diucapkan kepada kedua orang tua serta keluarga yang selalu memberi motivasi, Doa dan dukungan moral maupun materi, serta pihak-pihak yang telah mendukung dalam proses pembuatan jurnal ini yang tidak dapat disebutkan satu persatu. Kiranya jurnal ini bisa memberi manfaat bagi pembaca dan dapat meningkatkan kualitas jurnal.

DAFTAR PUSTAKA

- [1] G. F. A. Sari, "Krisis Air Menangani Penyediaan Air Bersih Di Dunia Yang Semakin Kekurangan Sumber Daya," *Jurnal Ilmiah Research Student*, vol. 1, no. 5, pp. 334-341, 2024.
- [2] Khavidiana Fahmi Addien, Ratna Septi Hendrasari, "Analisis Sistem Distribusi Air Dari Mata Air Di Dusun Kawedan Dengan Menggunakan Software Epanet 2.2," *Jurnal Ilmiah Teknik dan Ilmu Komputer*, vol. 3, no. 3, pp. 167-176, 2024.
- [3] Kartika Zahretta Wijaya, Arif Djunaidy, dan Faizal Mahananto, "Segmentasi Pelanggan Menggunakan Algoritma K-Means dan Analisis RFM di Ova Gaming E-Sports Arena Kediri," *JURNAL TEKNIK ITS*, vol. 10, no. 2, pp. 230-237, 2021.
- [4] Gustientiedina, M.Hasmil Adiyaa, Yenny Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru," *Jurnal Nasional Teknologi Dan Sistem Informasi*, vol. 5, no. 1, pp. 17-24, 2019.
- [5] Tifani Amalina, Danendra Bima Adhi Pramana, Betha Nurina Sari, "Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food," *Jurnal Ilmiah Wahana Pendidikan*, vol. 8, no. 15, pp. 574-583, 2022.
- [6] Ika Suhaila, Abdul Kadir, & Armansyah Matondang, "Analisis Kualitas Pelayanan Air Minum pada Perusahaan Daerah Air Minum (PDAM) Tirtanadi Cabang Tuasan," *Jurnal Ilmiah Administrasi Publik dan Pemerintahan*, vol. 1, no. 2, pp. 71-80,, 2022.
- [7] Perkim, "PDAM atau Sumur Gali?," *perkim.id*, 7 6 2021. [Online]. Available: <https://perkim.id/permukiman/pdam-atau-sumur-gali/#:~>. [Diakses 24 3 2025].
- [8] Gadis Fransiska Apriliana Sari, Devi Yolanda, "Krisis Air Menangani Penyediaan Air Bersih Di Dunia Yang Semakin Kekurangan Sumber Daya," *Jurnal Ilmiah Research Student*, vol. 1, no. 5, pp. 334-341, 2024.
- [9] K. Unfried, "Water scarcity and social conflict," *Journal of Environmental Economics and Management*, vol. 113, 2022.
- [10] Agung Nugraha, Odi Nurdiawan, Gifthera Dwilestari, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport," (*Jurnal Mahasiswa Teknik Informatika*, vol. 6, no. 2, pp. 849-855, 2022.
- [11] Putri Mai Sarah Tarigan, Jaya Tata Hardinata, Hendry Qurniawan, M.Safii, Riki Winanjaya5, "Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang (Studi Kasus : Toko Sinar Harahap)," *Jurnal UMJ*, vol. 12, no. 2, p. 51 – 61, 2022.

- [12] K. Erwansyah, "Implementasi Data Mining Untuk Menganalisa Hubungan Data Penjualan Produk Bahan Kimia Terhadap Persediaan Stok Barang Menggunakan Algoritma FP (Frequent Pattern) Growth Pada PT. Grand Multi Chemicals," *J-SISKO TECH*, vol. 2, no. 2, pp. 30-40, 2019.
- [13] Aulia Rizqi Wahidah, Yogi Bachtiar, Rayung Wulan, "Sistem Pendukung Analisa Key Performance Indicator (Kpi) Menggunakan Metode Data Mining Berbasis Web Python Programming," *JRKT (Jurnal Rekayasa Komputasi Terapan)*, vol. 2, no. 3, pp. 151-158, 2022.
- [14] Inas Ajeng Nur Afifah, Heri Nurdianto, "Data Mining Clustering Dalam Pengelompokan Buku Perpustakaan Menggunakan Algoritma K-Means," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 8, no. 3, pp. 802-814, 2023.
- [15] Normah, Siti Nurajizah, Arinda Salbinda, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten," *Jurnal Teknik Komputer AMIK BSI*, vol. 7, no. 2, pp. 158-163, 2021.
- [16] Safira Azis, Widya; Dedy Atmajaya, dan, "IMPLEMENTASI DATA MINING DENGAN ALGORITMA K-MEANS CLUSTERING UNTUK ANALISIS BISNIS PADA PERUSAHAAN ASURANSI," *Djtechno : Jurnal Teknologi Informasi*, vol. 5, no. 1, pp. 139-152, 2024.
- 17 E. Pepayosa, B. Surbakti, A. F. Boy, S. Yakub, S. Informasi, and S. T. Dharma, "Pemanfaatan Data Mining Untuk Segmentasi Pelanggan Dalam Meningkatkan Efektivitas Penawaran Di Showroom Menggunakan Metode K-Means," vol. 5, pp. 544–555, 2026.