



Pemanfaatan Algoritma K-Means Clustering Pada Sistem Rental Mobil

Sri Wulandari Maesaroh^{*}, T M Diansyah, Risiko Liza, Yessi Fitri Annisah Lubis

Program Studi Teknik Informatika Fakultas Teknik dan Komputer, Universitas Harapan Medan

Email: ¹*wulandarimaesaroh@gmail.com, ²Dian.10.22@gmail.com, ³risko.liza@gmail.com, ⁴yessi.annisa@gmail.com

Email Penulis Koresponden: wulandarimaesaroh@gmail.com

Abstrak—Penelitian ini menggunakan algoritma K-Means clustering untuk menganalisis data penyewaan mobil dari PT. Station Armada Indonesia, yang bertujuan untuk menyederhanakan pemilihan mobil pelanggan dan meningkatkan daya tanggap pasar perusahaan. Penelitian ini membahas masalah kebingungan pelanggan yang berasal dari berbagai macam jenis mobil yang ditawarkan oleh perusahaan. Dengan menggunakan pengelompokan K-Means pada data penyewaan bulan Agustus 2023, penelitian ini mengelompokkan mobil berdasarkan harga sewa dan jarak tempuh. Dataset, yang awalnya mencakup empat kategori mobil (Minibus MVP, SUV, City Car, dan Van/Bus), kemudian diperinci lebih lanjut untuk memasukkan model mobil individu. Tiga parameter—durasi sewa, harga sewa, dan jarak tempuh—digunakan untuk pengelompokan. Algoritma K-Means, yang dipilih karena kemudahan implementasi dan kecepatannya, diterapkan secara iteratif menggunakan jarak Euclidean untuk menempatkan titik data ke pusat data terdekat. Penelitian ini awalnya mendefinisikan dua cluster. Perhitungan manual, yang dirinci dalam makalah tersebut, menunjukkan proses pengelompokan. Hasil manual ini kemudian dibandingkan dengan hasil yang diperoleh dengan menggunakan RapidMiner Studio versi 10.1, yang menunjukkan efisiensi perangkat lunak dalam menangani proses K-Means. Hasil keluaran RapidMiner mencakup tampilan Data, Statistik, dan Anotasi, yang memberikan analisis komprehensif dari cluster. Pengelompokan akhir, yang dicapai setelah tiga kali iterasi, menunjukkan dua cluster yang berbeda: satu cluster yang mewakili jenis mobil yang kurang populer (Klaster 0), dan klaster lainnya mewakili jenis mobil yang paling populer (Klaster 1). Cluster 0 terdiri dari enam jenis mobil dengan jarak tempuh rata-rata pelanggan mulai dari 673 km hingga 2050 km, sedangkan Cluster 1 terdiri dari 24 jenis mobil dengan jarak tempuh rata-rata antara 270 km dan 3388 km. Temuan ini memungkinkan PT. Station Armada Indonesia untuk mengoptimalkan manajemen armada dan strategi pemasaran dengan berfokus pada jenis mobil yang paling banyak diminati. Studi ini menyimpulkan bahwa pengelompokan K-Means, yang diimplementasikan melalui RapidMiner, menawarkan alat yang berharga untuk meningkatkan pemahaman pelanggan tentang pemilihan mobil dan meningkatkan efisiensi perusahaan secara keseluruhan.

Kata Kunci: Data Mining; Clustering; K-Means; Rental Mobil; Rapidminer

Abstract—This research utilizes the K-Means clustering algorithm to analyze car rental data from PT. Station Armada Indonesia, aiming to simplify customer car selection and improve the company's market responsiveness. The study addresses the problem of customer confusion stemming from the wide variety of car types offered by the company. By employing K-Means clustering on August 2023 rental data, the research groups cars based on rental price and mileage. The dataset, initially encompassing four car categories (Minibus MVP, SUV, City Car, and Van/Bus), was further detailed to include individual car models. Three parameters—rental duration, rental price, and mileage—were used for clustering. The K-Means algorithm, chosen for its ease of implementation and speed, was applied iteratively using Euclidean distance to assign data points to the nearest centroid. The study initially defined two clusters. Manual calculations, detailed in the paper, demonstrate the clustering process. These manual results were then compared against results obtained using RapidMiner Studio version 10.1, showcasing the software's efficiency in handling the K-Means process. The RapidMiner output included Data, Statistics, and Annotations views, providing a comprehensive analysis of the clusters. The final clustering, achieved after three iterations, revealed two distinct clusters: one representing less popular car types (Cluster 0), and the other representing the most popular car types (Cluster 1). Cluster 0 contained six car types with average customer mileage ranging from 673 km to 2050 km, while Cluster 1 included 24 car types with average mileage between 270 km and 3388 km. The findings enable PT. Station Armada Indonesia to optimize fleet management and marketing strategies by focusing on the most in-demand car types. The study concludes that K-Means clustering, implemented via RapidMiner, offers a valuable tool for enhancing customer understanding of car selection and improving the company's overall efficiency.

Keyword: Data Mining; Clustering; K-Means; Car Rental; Rapidminer

1. PENDAHULUAN

Rental mobil merupakan sebuah layanan yang semakin populer di kalangan masyarakat. Layanan ini memungkinkan individu atau perusahaan untuk menyewa mobil dalam jangka waktu tertentu, memberikan fleksibilitas yang tidak dimiliki oleh pemilik mobil. Dengan rental mobil, pelanggan dapat menggunakan kendaraan sesuai kebutuhan tanpa harus mengeluarkan biaya besar untuk membeli mobil. Hal ini sangat menguntungkan bagi mereka yang memerlukan mobil untuk keperluan pribadi, perjalanan bisnis, atau keperluan lainnya. Kebutuhan akan alat transportasi roda empat saat ini menjadi sangat penting dalam memenuhi mobilitas masyarakat. Dalam era modern ini, mobilitas yang cepat dan efisien menjadi salah satu faktor penentu dalam menjalani aktivitas sehari-hari. Rental mobil hadir sebagai solusi yang dapat memenuhi kebutuhan tersebut[1]. Perusahaan penyewaan mobil telah berkembang pesat, menawarkan berbagai merek kendaraan dan fasilitas yang beragam. Persaingan di Pasar Rental Mobil, persaingan antar perusahaan rental mobil sangat ketat. Banyaknya perusahaan yang menawarkan layanan serupa membuat pelanggan memiliki banyak pilihan[2]. Oleh karena itu, dibutuhkan strategi pemasaran yang handal untuk menarik perhatian pelanggan[3]. Rental mobil bukan hanya sekadar bisnis, tetapi juga merupakan jasa yang membantu masyarakat yang memerlukan kendaraan untuk berbagai keperluan.

Variasi Armada dan Kebutuhan Pelanggan, perusahaan rental mobil sering kali memiliki armada yang beragam, dengan variasi harga, ukuran, dan kualitas yang berbeda. Setiap pelanggan memiliki preferensi dan kebutuhan yang unik[4]. Oleh karena itu, penting bagi perusahaan untuk memahami perilaku sewa pelanggan agar dapat menyediakan



layanan yang lebih disesuaikan. Mengelompokkan pelanggan berdasarkan atribut seperti harga sewa dan jarak tempuh dapat membantu perusahaan dalam merencanakan strategi yang tepat[5]. Data Mining sebagai solusi dalam menghadapi tantangan ini, perusahaan rental mobil dapat memanfaatkan data mining sebagai solusi[6]. Data mining adalah proses menggali informasi yang tersembunyi dari dalam database yang besar[7]. Dengan menggunakan data mining, perusahaan dapat menganalisis dan menemukan hubungan yang tidak terduga antar data, yang pada gilirannya dapat membantu dalam pengambilan keputusan[8]. Metode Clustering dalam Data Mining, salah satu metode yang dapat digunakan dalam data mining adalah metode clustering[9]. Clustering adalah teknik yang memisahkan beberapa kelompok data berdasarkan ciri masing-masing. Dalam konteks rental mobil, objek yang dapat dikelompokkan bisa berupa pelanggan, jenis mobil, atau bahkan pola sewa. Dengan memisahkan data ke dalam kelompok yang relevan, perusahaan dapat lebih mudah memahami kebutuhan pelanggan dan meresponsnya dengan lebih baik[10]. Algoritma K-Means dalam Clustering[11], algoritma k-means menjadi salah satu algoritma yang paling penting dalam data mining. Kelebihan dari algoritma ini adalah kemudahan dalam implementasi dan kecepatan dalam proses analisis[12]. K-means merupakan metode pengelompokan yang paling populer dan telah digunakan secara luas untuk menyelesaikan berbagai persoalan[13]. Dalam konteks rental mobil, algoritma k-means dapat digunakan untuk mempartisi data berdasarkan harga sewa dan jarak tempuh, sehingga data dengan karakteristik yang sama akan dikelompokkan ke dalam satu cluster.

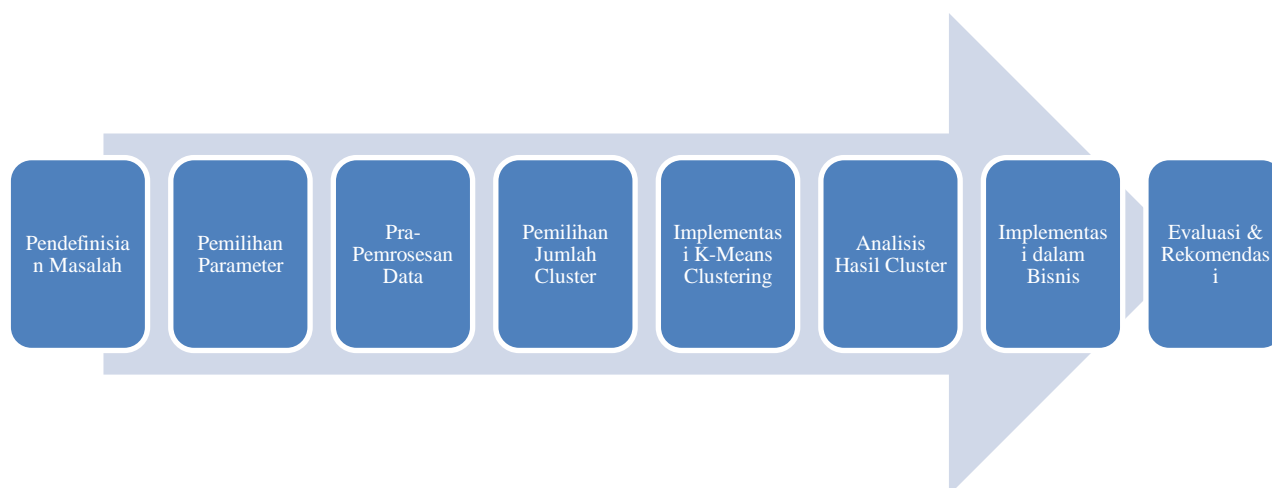
Penerapan K-Means dalam beberapa penelitian telah menggunakan algoritma k-means untuk berbagai tujuan. Misalnya, penelitian oleh [14] menggunakan metode k-means untuk mengelompokkan buku di perpustakaan guna mengetahui jumlah buku yang dipinjam. Penelitian lain oleh [15] menggunakan algoritma k-means clustering untuk mengelompokkan data penjualan dalam strategi penjualan sembako. Penelitian [16] juga menerapkan algoritma k-means clustering untuk mengelompokkan mobil bekas agar konsumen dapat dengan mudah menemukan mobil sesuai kebutuhan mereka. Penelitian Terkait Rental Mobil Dalam penelitian ini, kami akan menggunakan algoritma k-means clustering untuk mengelompokkan jenis mobil berdasarkan harga sewa dan jarak tempuh. Tujuannya adalah untuk mempermudah pelanggan dalam pemilihan jenis mobil yang paling diminati. Selain itu, hasil dari penelitian ini diharapkan dapat menjadi bahan bagi PT. Station Armada Indonesia untuk merespons perubahan pasar dengan lebih baik dan mencapai hasil yang lebih optimal.

2. METODOLOGI PENELITIAN

Metodologi Penelitian Metodologi yang akan digunakan dalam penelitian ini meliputi pengumpulan data, analisis data, dan interpretasi hasil[17]. Data yang akan dikumpulkan mencakup informasi mengenai jenis mobil, harga sewa, dan jarak tempuh. Setelah data terkumpul, langkah selanjutnya adalah melakukan analisis menggunakan algoritma k-means untuk mengelompokkan data tersebut.

2.1 Kerangka Kerja Penelitian

Kerangka kerja penelitian yang menjadi panduan sistematis dalam menjalankan studi[18]. Kerangka ini mencakup langkah-langkah penting mulai dari pendefinisian masalah hingga evaluasi dan rekomendasi, yang bertujuan untuk memastikan penelitian dilakukan secara terstruktur dan efektif dapat dilihat pada Gambar 1 berikut:



Gambar 1. Kerangka Kerja Penelitian

Diagram ini terdiri dari beberapa langkah penting yang harus diikuti untuk mencapai tujuan penelitian.

- Pendefinisian Masalah: Langkah awal untuk mengidentifikasi isu yang akan diteliti.
- Pemilihan Parameter: Menentukan variabel kunci yang relevan dengan masalah.
- Pra-Pemrosesan Data: Mengolah data agar siap digunakan dalam analisis.
- Pemilihan Jumlah Cluster: Menetapkan jumlah cluster optimal untuk analisis K-Means.
- Implementasi K-Means Clustering: Melakukan pengelompokan data berdasarkan algoritma K-Means.



- f. Analisis Hasil Cluster: Mengevaluasi karakteristik setiap cluster yang terbentuk.
- g. Implementasi dalam Bisnis: Menerapkan hasil analisis ke strategi bisnis nyata.
- h. Evaluasi & Rekomendasi: Mengkaji efektivitas dan memberikan saran perbaikan.

Diagram ini memberikan panduan sistematis bagi peneliti dalam menjalankan studi mereka secara terstruktur dan efisien.

2.2 Algoritma K-means Clustering

K-means clustering adalah metode untuk pengelompokan data non-hierarki yang mempartisis data ke dalam bentuk dua atau lebih kelompok, sehingga data berkarakteristik sama dikelompokkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain[19]. Algoritma *k-means* berguna untuk melakukan *clustering* secara *unsupervised learning* pada sebuah data berdasarkan parameter tertentu[20]. Dengan kata lain, algoritma ini tidak memerlukan label atau informasi kelas yang diketahui sebelumnya untuk melakukan *clustering* pada data. Algoritma *k-means clustering* dapat membantu dalam mengelompokkan data menjadi beberapa *cluster* yang homogeny berdasarkan parameter tertentu dan memberikan wawasan yang bermakna dalam analisis data[16]. Fungsi *k-mean clustering* adalah menentukan kelompok-kelompok set data. Caranya adalah dengan melihat jumlah grup yang diwakili variabel K. untuk bisa mengelompokkan data, algoritma ini bekerja secara berulang (iteratif) dari setiap titik data ke grup K yang telah dibuat. Pusat *cluster* adalah rata-rata dari semua data atau objek dalam *cluster* tertentu. Setiap objek kemudian ditugaskan kembali memakai pusat *cluster* yang baru. Jika pusat *cluster* tidak berubah lagi maka proses *clustering* selesai atau kembali ke langkah menentukan centroid samapai pusat *cluster* tidak berubah lagi[21]. *K-means clustering* banyak digunakan dalam berbagai aplikasi seperti segmentasi pelanggan, pengelompokan data pasar, analisis citra, pengenalan pola dan masih banyak lagi[22]. Algoritma *k-means clustering* membantu dalam memahami struktur data yang rumit, mengidentifikasi kelompok-kelompok yang sama, memberikan wawasan tentang pola dan hubungan dalam data. Langkah-langkah melakukan *clustering* atau pengelompokan dengan metode *k-means clustering* sebagai berikut:

- a. Pilih jumlah *cluster* k.
- b. Inisialisasi k sebagai *centroid*, yang dilakukan dengan cara acak atau random.
- c. Menghitung jarak objek ke *centroid* dengan menggunakan rumus *Euclidean Distance* yang dirumuskan. Rumus Jarak Euclidean digunakan untuk mengukur seberapa jauh dua titik dalam ruang dua dimensi, yang sering digunakan dalam analisis data, rumusnya ditulis sebagai:

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (1)$$

Keterangan:

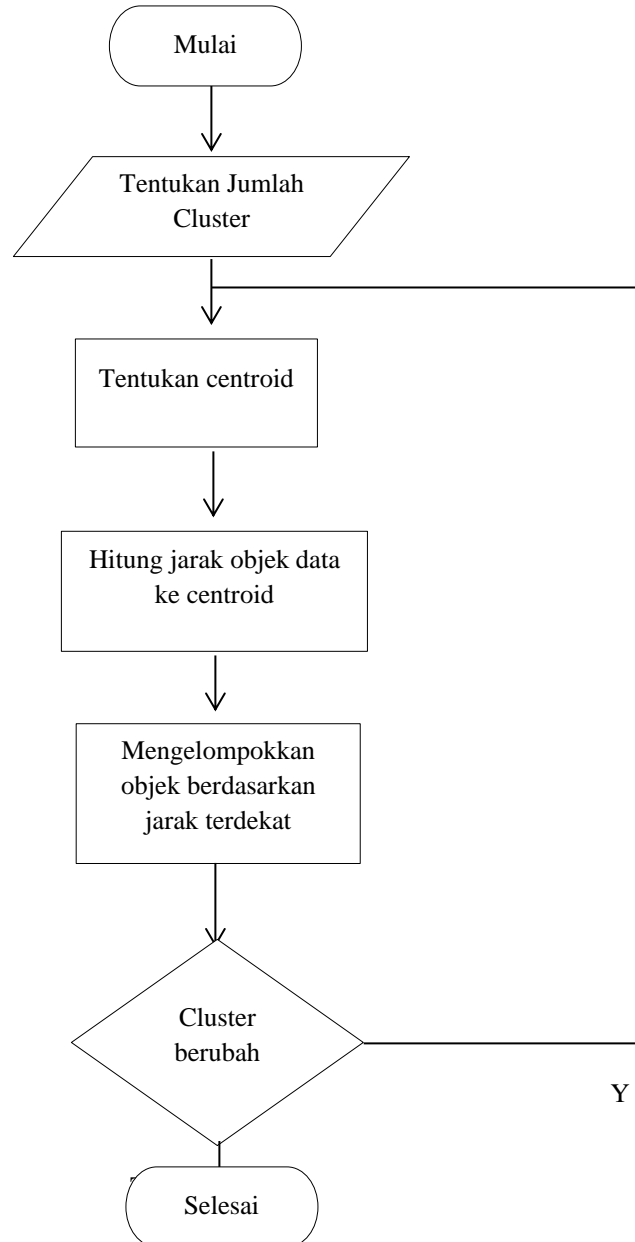
- De = *Euclidean Distance*
- i = Banyaknya objek
- (x, y) = Koordinat objek
- (s, t) = Koordinat *centroid*.

Rumus Jarak Euclidean digunakan untuk mengukur seberapa jauh dua titik dalam ruang dua dimensi, yang sering digunakan dalam analisis data. Dalam rumus tersebut, **De** merujuk pada jarak antara dua titik, yang dihitung dengan mengukur selisih antara koordinat objek yang satu dengan objek yang lain. **i** menggambarkan indeks objek yang sedang dihitung jaraknya, sedangkan **(x, y)** adalah koordinat objek tersebut. Sementara itu, **(s, t)** adalah koordinat dari pusat atau centroid yang berfungsi sebagai titik referensi dalam perhitungan jarak. Dengan demikian, rumus ini bertujuan untuk menentukan kedekatan antara objek dengan titik pusat atau centroid dalam dua dimensi, yang sangat relevan dalam proses analisis seperti pengelompokan data atau analisis klaster, di mana pengukuran jarak antar objek memegang peranan penting.

- d. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan centroidnya.
- e. Tentukan posisi *centroid* baru. Jika pusat *cluster* tidak berubah lagi maka proses *clustering* selesai. Atau kembali ke langkah 3 sampai pusat *cluster* tidak berubah lagi.

2.3 Flowchart

Flowchart atau sering disebut dengan diagram alir merupakan suatu jenis diagram yang merepresentasikan algoritma langkah-langkah instruksi yang berurutan dalam sistem. Seorang analis sistem menggunakan *flowchart* sebagai bukti dokumentasi untuk menjelaskan gambaran logis sebuah sistem yang akan dibangun kepada programmer. *Flowchart* dapat membantu untuk memberikan solusi terhadap masalah yang bisa saja terjadi dalam membangun sistem. *Flowchart* digambarkan dengan menggunakan simbol-simbol. Setiap simbol mewakili suatu proses tertentu. Sedangkan untuk menghubungkan satu proses ke proses selanjutnya digambarkan dengan menggunakan garis penghubung[10]. Berikut ini gambar *flowchart* dari metode *k-means* seperti yang terlihat pada gambar 1:



Gambar 2. Flowchart Metode K-means

Gambar 2. flowchart ini menggambarkan proses algoritma K-Means Clustering. Proses dimulai dengan menentukan jumlah cluster yang diinginkan, diikuti dengan pemilihan centroid awal. Selanjutnya, jarak objek data dihitung ke centroid, dan objek dikelompokkan berdasarkan jarak terdekat; jika cluster berubah, proses akan diulang hingga konvergensi tercapai dan akhirnya menghasilkan hasil clustering yang stabil.

3. HASIL DAN PEMBAHASAN

Pada penelitian ini data yang digunakan adalah data rental mobil PT. Station Armada Indonesia yang dihasilkan dari rekap penyewaan mobil perbulan. Dalam penelitian data yang diambil sebagai sampel yaitu data penyewaan mobil pada bulan agustus 2023. Pada PT. Station Armada Indonesia mempunyai banyak merek mobil yang dikelompokkan menjadi 4 jenis mobil sesuai fungsi dan kegunaannya.

Tabel 1. Data Jenis Mobil PT. Station Armada Indonesia

Jenis Mobil	Harga Sewa
Minibus MVP	400.000 – 2.350.000
SUV	500.000 – 2.200.000
City Car	300.000 – 400.000
Van/Bus	1.600.000 – 3.500.000



Setiap jenis mobil memiliki beberapa jenis merek mobil, agar data dapat diolah maka perlu dilakukan penjabaran tiap-tiap jenis mobil. Pada tiap-tiap jenis mobil penulis mengambil beberapa parameter yaitu banyaknya jumlah sewa, harga sewa, dan jarak tempuh yang akan dijadikan untuk proses pengelompokkan jenis mobil pada penerapan *k-means*. Data yang sudah dijadikan sampel akan dilakukan pengolahan data dengan proses *clustering* dengan menggunakan metode algoritma *k-means clustering* sehingga didapatkan hasil pengelompokkan data yang diinginkan. Pada proses perhitungan dalam *cluster* dengan menggunakan metode algoritma *k-means clustering* memiliki beberapa langkah-langkah sebagai berikut:

a. Menentukan jumlah *cluster*

Menentukan jumlah *cluster* yang digunakan pada data penyewaan mobil sebanyak 2 *cluster*.

b. Menentukan centroid

Menentukan pusat awal *cluster* (centroid). Pusat awal *cluster* diambil dari data ke-1 dan ke-2 sebagai perhitungan pertama.

Tabel 2. Centroid Awal

Centroid	X	Y	Z
C1	25	400.000	2376
C2	20	400.000	1893

Pada Tabel 2 dapat dilihat jumlah *cluster* yang sudah ditentukan sebanyak 2 *cluster* dengan setiap centroidnya memiliki 3 variabel yaitu variabel x, variable y, variabel z. variabel diambil dari data jenis mobil yaitu jangka waktu lama disewakan, harga sewa, jarak tempuh yang berfungsi sebagai parameter dari hasil *clustering*.

c. Menghitung jarak dari centroid

Menghitung jarak antar titik centroid dengan titik tiap objek. Setiap data akan dihitung jarak terdekat dengan setiap centroid yang sudah ditentukan pada tahap sebelumnya dengan menggunakan rumus *Euclidian Distance*. Perhitungannya adalah sebagai berikut :

$$D_{(1,1)} = \sqrt{(25 - 25)^2 + (400.000 - 400.000)^2 + (2376 - 2376)^2}$$

$$= 0$$

$$D_{(1,2)} = \sqrt{(25 - 20)^2 + (400.000 - 400.000)^2 + (2376 - 1893)^2}$$

$$= 483,0258792$$

$$D_{(2,1)} = \sqrt{(20 - 25)^2 + (400.000 - 400.000)^2 + (1893 - 2376)^2}$$

$$= 483,0258792$$

$$D_{(2,2)} = \sqrt{(20 - 20)^2 + (400.000 - 400.000)^2 + (1893 - 1893)^2}$$

$$= 0$$

$$D_{(3,1)} = \sqrt{(23 - 25)^2 + (400.000 - 400.000)^2 + (2826 - 2376)^2}$$

$$= 450,0044444$$

$$D_{(3,2)} = \sqrt{(23 - 20)^2 + (400.000 - 400.000)^2 + (2826 - 1893)^2}$$

$$= 933,0048231$$

...

$$D_{(30,1)} = \sqrt{(25 - 25)^2 + (500.000 - 400.000)^2 + (2150 - 2376)^2}$$

$$= 100000,2554$$

$$D_{(30,2)} = \sqrt{(25 - 20)^2 + (500.000 - 400.000)^2 + (2150 - 1893)^2}$$

$$= 100000,3304$$

d. Pengelompokkan jarak setiap data ke centroid

Pengelompokkan data jarak terdekat ke centroid pada iterasi ke-1

Tabel 3. Pengelompokkan Jarak Setiap Data Ke Centroid Iterasi Ke-1

Data ke	Jenis Mobil	Jarak Ke Centroid		Terdekat	Cluster
		C1	C2		
1	Avanza ADJ (MVP)	0	483,0258792	0	Cluster 1
2	Avanza ADR (MVP)	483,0258792	0	0	Cluster 2



Data ke	Jenis Mobil	Jarak Ke Centroid		Terdekat	Cluster
		C1	C2		
3	Avanza MAB (MVP)	450,0044444	933,0048231	450,0044444	Cluster 1
4	Avanza ADQ (MVP)	586,003413	103,2375901	103,2375901	Cluster 2
5	Xenia AED (MVP)	1216,020148	733,0027285	733,0027285	Cluster 2
6	Xenia AEO (MVP)	1657,004828	1174,000426	1174,000426	Cluster 2
7	Xenia ALT (MVP)	1890,016931	1407,003198	1407,003198	Cluster 2
8	Calya ACL (MVP)	50044,33307	50026,33437	50026,33437	Cluster 2
9	Xpander ADA (MVP)	50034,39867	50018,8213	50018,8213	Cluster 2
10	Xpander ADA (MVP)	50003,67224	50000,15165	50000,15165	Cluster 2
11	Xpander ARM (MVP)	50000,90003	50006,131	50000,90003	Cluster 1
12	Innova Zenix MAE (MVP)	500000	500000,2295	500000	Cluster 1
13	Innova Zenix AEM (MVP)	500001,7583	500000,7107	500000,7107	Cluster 2
14	Innova Zenix AEB (MVP)	500001,0241	500002,235	500001,0241	Cluster 1
15	Innova Reborn (MVP)	450000,4594	450000,0284	450000,0284	Cluster 2
16	XL-7 (MVP)	50036,54508	50020,41649	50020,41649	Cluster 2
17	Alphard (MVP)	1950000,744	1950000,382	1950000,382	Cluster 2
18	Rocky MAE (City Car)	1261,009913	778,0642647	778,0642647	Cluster 2
19	Rocky ADS (CityCar)	1783,010095	1300,000385	1300,000385	Cluster 2
20	Brio (City Car)	50035,3701	50019,54047	50019,54047	Cluster 2
21	Agya (City Car)	100020,5627	100011,9347	100011,9347	Cluster 2
22	Hiace (Van Bus)	3100000,028	3100000,13	3100000,028	Cluster 1
23	Bus Pariwisata 43 Seats (Van Bus)	2600000,193	2600000,052	2600000,052	Cluster 2
24	Medium Bus 25 Seats (Van Bus)	1200001,143	1200000,573	1200000,573	Cluster 2
25	Pajero Sport (SUV)	1800000,044	1800000,002	1800000,002	Cluster 2
26	Fortuner (SUV)	1100000,048	1100000,011	1100000,011	Cluster 2
27	Terios AEE (SUV)	100011,1149	100005,0802	100005,0802	Cluster 2
28	Terios AEM (SUV)	100010,6722	100004,7823	100004,7823	Cluster 2
29	Terios ALT (SUV)	100009,2613	100003,8544	100003,8544	Cluster 2
30	Rush MAB (SUV)	100000,2554	100000,3304	100000,2554	Cluster 1

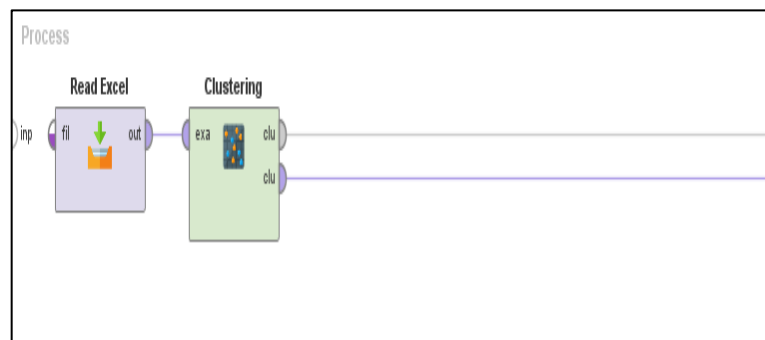
e. Menghitung Centroid Baru

Centroid yang baru dapat ditentukan berdasarkan pengelompokkan anggota masing-masing cluster. Berdasarkan tabel 3 C1 untuk parameter jumlah sewa, harga sewa, dan jarak tempuh memiliki 7 data, sementara C2 untuk parameter jumlah sewa, harga sewa dan jarak tempuh memiliki 23 data.

Tabel 4. Hasil Iterasi Ke-1

C1	1,3,11,12,14,22,30
C2	2,4,5,6,7,8,9,10,13,15,16,17,18,19,20,21,23,24,25,26,27,28,29

Selanjutnya centroid yang baru dapat ditentukan berdasarkan pengelompokkan anggota masing-masing cluster. Berdasarkan hasil pengelompokkan dari seluruh data menggunakan metode *k-means clustering* penulis mendapatkan hasil akhir pengelompokkan hingga iterasi ke-3. Dimana titik centroid tidak berubah lagi dan tidak ada data yang berpindah antar cluster. Pada implementasi dan pengujian sistem penelitian ini, penulis menggunakan sebuah *software* RapidMiner studio versi 10.1. dengan pengujian data dengan *software* RapidMiner penulis akan membandingkan bagaimana hasil pengolahan data secara manual dengan hasil pengolahan data menggunakan *software*. Selanjutnya untuk mengimplementasikan menggunakan Rapidminer diawali mengkoneksikan *Read Excel* dengan *clustering* dengan cara hubungkan *out* pada *Read Excel* dengan *exa* pada *clustering* Setelah itu jalankan dengan klik *start* seperti yang terlihat pada Gambar 2 berikut:



Gambar 2 Tampilan Menghubungkan Read Excel Dengan Clustering

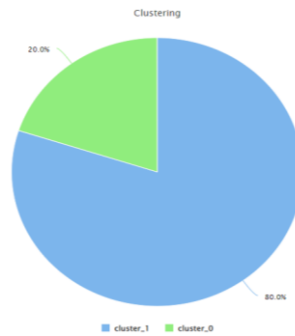


Setelah klik *start* dan program dijalankan akan muncul tampilan seperti pada Gambar 3 berikut:

Row No.	Merek Mobil	cluster	Jumlah	Harga sewa	Jarak Temp...
1	Avanza ADJ (...)	cluster_1	25	400000	2376
2	Avanza ADR (...)	cluster_1	20	400000	1893
3	Avanza MAB (...)	cluster_1	23	400000	2826
4	Avanza ADQ (...)	cluster_1	27	400000	1790
5	Xenia AED (M...)	cluster_1	18	400000	1160
6	Xenia AEO (M...)	cluster_1	21	400000	719
7	Xenia ALT (M...)	cluster_1	17	400000	486
8	Calya ACL (M...)	cluster_1	19	350000	270
9	Xpander ADA ...	cluster_1	30	450000	521
10	Xpander ADA ...	cluster_1	26	450000	1770
11	Xpander ARM...	cluster_1	27	450000	2676
12	Innova Zenix ...	cluster_1	24	900000	2372
13	Innova Zenix ...	cluster_1	30	900000	1050
14	Innova Zenix ...	cluster_1	25	900000	3388
15	Inowa Rebor...	cluster_1	22	850000	1733

Gambar 3 Tampilan *ExampleSet*

Pada bagian *ExampleSet* terdapat beberapa bentuk hasil *cluster* yaitu *Data*, *Statistics*, dan *Annotations*. Pada *ExampleSet* dapat dilihat beberapa tampilan hasil *cluster* yaitu *Data View*. *Data View* merupakan tampilan hasil *cluster* data secara keseluruhan sesuai dengan data yang telah di inputkan seperti terlihat pada gambar 4 berikut :



Gambar 4 Tampilan *Visualization*

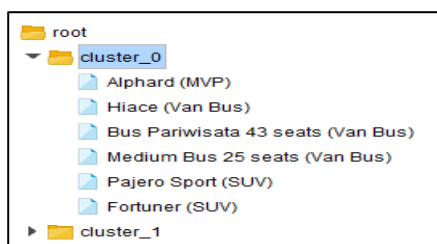
Visualization merupakan tampilan grafik hasil pengelompokan atau *cluster* sampel data jenis mobil dengan 2 *cluster* yang dapat dilihat pada Gambar 4. Sementara untuk *cluster* model terdapat *Description*, *Folder View*, *Graph*, *Centroid Table*, *Plot View* dan *Annotations* seperti dilihat pada Gambar 5 berikut:

```

Cluster Model
Cluster 0: 6 items
Cluster 1: 24 items
Total number of items: 30
    
```

Gambar 5 Tampilan Hasil *Cluster Model*

Pada *Cluster Model (Clustering)* dapat dilihat beberapa tampilan hasil *cluster* yaitu *Description* yang merupakan tampilan hasil pengelompokan berdasarkan *cluster* dan jumlah anggotanya, tampilan *Description* terlihat pada Gambar 5.



Gambar 6 Tampilan *Folder View cluster 0*



Pada Folder *View* menampilkan data bagian-bagian *cluster* secara keseluruhan, dimana masing-masing anggota *cluster* menampilkan jenis mobil. Pada cluster 0 menampilkan jenis mobil dengan kategori kurang diminati yang terlihat pada Gambar 6.

4. KESIMPULAN

Penelitian yang telah dilakukan di PT. Station Armada Indonesia, penulis menarik beberapa kesimpulan penting mengenai pengelompokan armada rental mobil menggunakan algoritma K-Means Clustering. Metode ini terbukti efektif dalam mengelompokkan jenis mobil berdasarkan dua parameter utama: harga sewa dan jarak tempuh. Hasil pengelompokan menunjukkan adanya dua cluster yang berbeda, yaitu Cluster 0, yang terdiri dari enam jenis mobil kurang diminati pelanggan dengan rata-rata jarak tempuh antara 673 km hingga 2050 km, dan Cluster 1, yang mencakup 24 jenis mobil paling diminati dengan rata-rata jarak tempuh antara 270 km hingga 3388 km. Dengan informasi ini, PT. Station Armada Indonesia dapat lebih mudah mempersiapkan armada sesuai dengan preferensi pelanggan dan meningkatkan efisiensi operasional. Penerapan metode data mining melalui K-Means tidak hanya membantu perusahaan dalam memahami karakteristik pasar tetapi juga memungkinkan pengembangan strategi pemasaran dan penawaran yang lebih tepat sasaran. Implementasi algoritma K-Means pada platform RapidMiner menunjukkan hasil optimal dalam pengelolaan armada mobil. Data dari Microsoft Excel dihubungkan ke RapidMiner untuk diolah menjadi model clustering yang informatif. Dengan demikian, penggunaan metode ini memberikan wawasan berharga bagi perusahaan untuk menyesuaikan layanan mereka sesuai kebutuhan pelanggan. Secara keseluruhan, penelitian ini menegaskan bahwa penerapan K-Means Clustering merupakan langkah strategis bagi PT. Station Armada Indonesia untuk meningkatkan kepuasan pelanggan serta merespons dinamika pasar secara lebih efektif dan efisien.

REFERENCES

- [1] S. Aminah and F. Febriansyah, "Sistem Prediksi Kelayakan Rental Mobil Menggunakan Algoritma C4. 5," *JURIKOM (Jurnal Ris. Komputer)*, vol. 8, no. 6, p. 367, 2021.
- [2] N. Yunita and R. Rosmawati, "Sistem Informasi Rental Mobil Berbasis Web Pada PT Karya Mobil," *Simpatik J. Sist. Inf. Dan Inform.*, vol. 1, no. 1, pp. 53–62, 2021.
- [3] H. Kurniawan and S. Defit, "Data mining menggunakan metode k-means clustering untuk menentukan besaran uang kuliah tunggal," *J. Appl. Comput. Sci. Technol.*, vol. 1, no. 2, pp. 80–89, 2020.
- [4] M. K. Putri, J. S. F. M. Rahman, F. A. Nursyifa, S. Alfarisi, T. G. S. Putro, and R. Agustin, "Analisis Segmentasi Pasar Dalam Penggunaan Produk Viefresh Di Wilayah Sekitar Kampus Universitas Muhammadiyah Surabaya," *Balanc. Econ. Business, Manag. Account. J.*, vol. 16, no. 2, 2019.
- [5] J. Nasir, "Penerapan Data Mining Clustering Dalam Mengelompokkan Buku Dengan Metode K-Means," *J. Simetris*, vol. 11, no. 2, pp. 1–13, 2020.
- [6] D. Triyansyah and D. Fitriana, "Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing," *InComTech J. Telekomun. dan Komput.*, vol. 8, no. 3, pp. 163–182, 2018.
- [7] D. Rusdianto and L. Zaelani, "Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Mengetahui Pola Peminjaman Buku di Perpustakaan Universitas Bale Bandung," *J-SIKA/ J. Sist. Inf. Karya Anak Bangsa*, vol. 2, no. 02, pp. 1–10, 2020.
- [8] E. A. Pratiwi, B. Irawan, A. Bahtiar, and N. Rahaningsih, "PENERAPAN FP-GROWTH DALAM MENGANALISIS DATA PENJUALAN DI TOKO X," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 6, pp. 3549–3556, 2023.
- [9] N. F. Adani, A. F. Boy, and R. Syahputra, "Implementasi data mining untuk pengelompokan data penjualan berdasarkan pola pembelian menggunakan algoritma K-Means clustering pada Toko Syihan," *J. Cyber Tech*, vol. 2, no. 5, 2019.
- [10] R. Rahim *et al.*, "TOPSIS Method Application for Decision Support System in Internal Control for Selecting Best Employees," *J. Phys. Conf. Ser.*, vol. 1028, no. 1, p. 012052, Jun. 2018, doi: 10.1088/1742-6596/1028/1/012052.
- [11] A. Pangestu and T. Ridwan, "Penerapan Data Mining Menggunakan Algoritma K-Means Pengelompokan Pelanggan Berdasarkan Kubikasi Air Terjual Menggunakan Weka," *Just IT J. Sist. Informasi, Teknol. Inf. dan Komput.*, vol. 12, no. 3, pp. 67–71, 2022.
- [12] Z. Arif and A. Nurokhan, "Analisis Perbandingan Algoritma Kriptografi Simetris Dan Asimetris Dalam Meningkatkan Keamanan Sistem Informasi," *J. Teknol. Sist. Inf.*, vol. 4, no. 2, pp. 394–405, 2023.
- [13] D. Gopagani and P. V Lakshmi, "Automated machine learning tool: The first stop for data science and statistical model building," *Int. J. Adv. Comput. Sci. Appl.*, no. 2, pp. 410–418, 2020, doi: 10.14569/ijacsa.2020.0110253.
- [14] W. S. Azis and D. Atmajaya, "Pengelompokan Minat Baca Mahasiswa menggunakan Metode K-Means," *Ilk. J. Ilm.*, vol. 8, no. 2, pp. 89–94, 2016.
- [15] M. H. Adlun, R. Astuti, and F. M. Basysyar, "PENGELOMPOKAN DATA PENJUALAN SEMBAKO BERDASARKAN PERILAKU PEMBELI MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 4, pp. 7699–7703, 2024.
- [16] D. Yusuf, E. Sestri, and F. Razi, "Implementasi Teknik Clustering Untuk Pengelompokan Mobil Bekas Berdasarkan Grade Pada Mobi Auto," *J. Teknol. Sist. Inf. dan Sist. Komput. TGD*, vol. 6, no. 2, pp. 307–313, 2023.
- [17] Y. Rifa'i, "Analisis Metodologi Penelitian Kualitatif dalam Pengumpulan Data di Penelitian Ilmiah pada Penyusunan Mini Riset," *Cendekia Inov. Dan Berbudaya*, vol. 1, no. 1, pp. 31–37, 2023.
- [18] M. S. Jailani, "Teknik pengumpulan data dan instrumen penelitian ilmiah pendidikan pada pendekatan kualitatif dan kuantitatif," *IHSAN J. Pendidik. Islam*, vol. 1, no. 2, pp. 1–9, 2023.
- [19] D. Ramdhan, G. Dwilestari, R. D. Dana, and A. Ajiz, "Clustering data persediaan barang dengan menggunakan metode K-Means," *MEANS (Media Inf. Anal. dan Sist.)*, pp. 1–9, 2022.
- [20] H. Priyatman, F. Sajid, and D. Haldivany, "Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu



- Kelulusan Mahasiswa,” *J. Edukasi Dan Penelit. Inform.*, vol. 5, no. 1, p. 62, 2019.
- [21] S. Aulia, “Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-Means Clustering (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja),” *Djtechno J. Teknol. Inf.*, vol. 1, no. 1, pp. 1–5, 2020.
- [22] S. Arifah, E. R. Swedia, and M. R. D. Septian, “ANALISIS PERBANDINGAN ALGORITMA CLUSTERING DALAM MELAKUKAN SEGMENTASI WARNA PADA CITRA JAJAN TRADISIONAL,” *Sebatik*, vol. 27, no. 1, pp. 70–76, 2023.