

Pemanfaatan Algoritma K-Means Clustering Pada Sistem Rental Mobil

Sri Wulandari Maesaroh, T.M Diansyah*, Risiko Liza, Yessi Fitri Annisa Lubis

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

Email: ¹wulandarimaesaroh@gmail.com, ^{2,*}dian.10.22@gmail.com, ³risko.liza@gmail.com, ⁴yessy.annisa@gmail.com

Email Penulis Korespondensi: dian.10.22@gmail.com

Abstrak—PT. Station Armada Indonesia merupakan salah satu perusahaan yang bergerak dibidang jasa penyewaan rental mobil. Dengan banyaknya jenis pilihan mobil yang ditawarkan tersebut tidak jarang banyak pelanggan yang merasa bingung dalam memilih jenis mobil apa yang sesuai dengan kebutuhannya. Masalah seperti ini banyak dialami oleh para pelanggan yang bingung dengan banyaknya pilihan jenis mobil yang ada. Dalam penelitian ini menggunakan algoritma *k-means* untuk mengelompokkan mobil berdasarkan beberapa atribut. Algoritma *k-means* dapat digunakan untuk mengelompokkan data jenis mobil untuk membantu memberikan rekomendasi pemilihan jenis mobil. Tujuan penelitian ini adalah untuk mempermudah pelanggan dalam pemilihan jenis mobil yang paling diminati dan sebagai bahan bagi PT. Station Armada Indonesia untuk merespon perubahan pasar dengan lebih baik, dan mencapai hasil yang lebih baik. Mengelompokkan armada rental mobil berdasarkan harga sewa dan jarak tempuh dengan pemanfaatan menggunakan algoritma *k-means* ini dapat membantu PT. Station Armada Indonesia mengelompokkan jenis mobil. Dari hasil pengelompokkan diperoleh dua kelompok *cluster* dengan dengan karakter *cluster* pertama adalah kurang diminati pelanggan dan kelompok *cluster* kedua adalah paling diminati pelanggan. Sehingga perusahaan dapat dengan mudah mempersiapkan jenis armada yang paling diminati. Dalam penerapan metode data *mining* dengan menggunakan *k-means* sangat membantu dan mempermudah PT. Station Armada Indonesia untuk mengembangkan strategi pemasaran dan penawaran yang lebih efektif. Dengan mengelompokkan jenis mobil dengan implementasi *k-means* dapat mempermudah pengetahuan pelanggan memilih jenis mobil berdasarkan kebutuhan pelanggan.

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

Abstract—PT. Station Armada Indonesia is one of the companies engaged in the car rental service sector. With the many types of car choices offered, it is not uncommon for many customers to feel confused in choosing what type of car suits their needs. This problem is often experienced by customers who are confused by the many choices of car types available. In this study, the k-means algorithm was used to group cars based on several attributes. The k-means algorithm can be used to group car type data to help provide recommendations for choosing a car type. The purpose of this study is to make it easier for customers to choose the type of car that is most in demand and as material for PT. Station Armada Indonesia to respond better to market changes and achieve better results. Grouping car rental fleets based on rental prices and mileage by utilizing the k-means algorithm can help PT. Station Armada Indonesia group car types. From the grouping results, two cluster groups were obtained with the character of the first cluster being less in demand by customers and the second cluster group being the most in demand by customers. So that the company can easily prepare the type of fleet that is most in demand. In the application of data mining methods using k-means is very helpful and makes it easier for PT. Station Armada Indonesia to develop more effective marketing and offering strategies. By grouping car types with the implementation of k-means can facilitate customer knowledge in choosing car types based on customer needs.

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

1. PENDAHULUAN

Perkembangan kebutuhan masyarakat terhadap sarana transportasi roda empat mengalami peningkatan yang signifikan seiring dengan mobilitas yang semakin tinggi dalam kehidupan sehari-hari. Mobil bukan hanya sekadar alat transportasi pribadi, tetapi juga digunakan untuk berbagai kepentingan, seperti perjalanan bisnis, wisata keluarga, hingga kebutuhan logistik dalam skala kecil. Namun, tidak semua individu maupun perusahaan memiliki kemampuan untuk membeli mobil secara langsung karena alasan biaya yang relatif tinggi, kebutuhan yang bersifat sementara, maupun pertimbangan efisiensi. Hal inilah yang menjadi dasar berkembangnya layanan rental mobil atau penyewaan mobil di berbagai daerah, termasuk di Indonesia.

Rental mobil merupakan sebuah layanan yang memberikan fleksibilitas kepada pelanggan untuk menggunakan mobil dalam jangka waktu tertentu tanpa harus melakukan pembelian [1]. Hal ini tentunya memberikan keuntungan ganda: bagi pelanggan, layanan rental mobil menjadi solusi yang lebih ekonomis dan efisien; sementara bagi perusahaan penyedia jasa, bisnis ini menjadi peluang yang menjanjikan karena mampu menjangkau pasar yang luas. Bisnis rental mobil dapat dikategorikan sebagai salah satu sektor jasa yang memiliki prospek cerah karena mampu menjawab kebutuhan masyarakat akan mobilitas, baik untuk tujuan pribadi maupun komersial [2].

Pertumbuhan bisnis rental mobil semakin pesat dengan hadirnya berbagai perusahaan yang menawarkan beragam pilihan armada kendaraan, mulai dari mobil keluarga, mobil mewah, hingga kendaraan komersial. Armada yang ditawarkan juga memiliki variasi harga sewa, kapasitas penumpang, serta kualitas layanan yang berbeda-beda. Namun, di tengah perkembangan tersebut, persaingan antar penyedia jasa rental mobil semakin ketat. Hal ini menuntut setiap perusahaan untuk mampu menyusun strategi pemasaran yang efektif agar tetap kompetitif. Perusahaan yang gagal membaca preferensi pelanggan akan kesulitan mempertahankan pelanggan lama dan menarik pelanggan baru.

Dalam konteks persaingan yang ketat ini, salah satu tantangan utama yang dihadapi perusahaan rental mobil adalah bagaimana memahami perilaku, preferensi, dan kebutuhan pelanggan. Setiap pelanggan tentu memiliki pertimbangan yang berbeda saat memilih kendaraan, misalnya berdasarkan harga sewa, kenyamanan, maupun jarak tempuh kendaraan. Perusahaan yang mampu mengelompokkan pelanggan atau produk berdasarkan karakteristik tertentu akan lebih mudah

dalam menyusun strategi layanan yang tepat sasaran. Oleh karena itu, diperlukan pendekatan berbasis data yang mampu membantu perusahaan dalam menganalisis pola penyewaan, preferensi pelanggan, dan karakteristik armada yang dimiliki.

Salah satu solusi yang dapat digunakan untuk mengatasi permasalahan tersebut adalah penerapan data mining. Data mining adalah proses menggali informasi berharga dari kumpulan data yang besar dengan tujuan menemukan pola, hubungan, maupun tren tersembunyi yang dapat mendukung pengambilan keputusan [3]. Dalam konteks bisnis rental mobil, data mining dapat dimanfaatkan untuk menganalisis data penyewaan, mengelompokkan pelanggan, serta memberikan rekomendasi strategi pemasaran yang lebih tepat.

Salah satu teknik dalam data mining yang sering digunakan adalah clustering. Clustering merupakan metode pengelompokan data ke dalam kelompok (*cluster*) berdasarkan kesamaan karakteristik tertentu. Objek yang berada dalam satu cluster memiliki kemiripan yang tinggi, sedangkan antar cluster memiliki perbedaan yang signifikan [4]. Dengan demikian, clustering dapat membantu perusahaan rental mobil untuk mengelompokkan data penyewaan berdasarkan atribut penting seperti harga sewa dan jarak tempuh kendaraan.

Dari berbagai algoritma clustering, K-Means menjadi salah satu metode yang paling populer dan banyak digunakan. K-Means bekerja dengan cara membagi data ke dalam sejumlah cluster yang ditentukan sebelumnya, sehingga setiap data dengan karakteristik yang mirip akan berada pada cluster yang sama [5][6]. Keunggulan utama algoritma K-Means adalah kesederhanaannya, kecepatan dalam pemrosesan, serta kemampuannya untuk menangani data dalam jumlah besar [7]. Karena alasan inilah, K-Means banyak digunakan dalam berbagai bidang, mulai dari pengelompokan dokumen, analisis perilaku pelanggan, hingga sistem rekomendasi produk.

Beberapa penelitian sebelumnya yang relevan dengan penerapan algoritma K-Means clustering dalam berbagai bidang dapat menjadi acuan dalam penelitian ini. Penelitian yang dilakukan oleh Zunaida dkk. (2024) membahas Penerapan Data Mining untuk Clustering Penduduk Miskin di Kota Tanjungbalai. Hasil penelitian menunjukkan bahwa penerapan data mining dengan algoritma K-Means mampu menghasilkan informasi baru berupa pengelompokan data penduduk miskin yang dapat dijadikan bahan pertimbangan bagi pihak terkait dalam mengambil keputusan. Proses clustering yang dilakukan menghasilkan keluaran sesuai dengan yang diharapkan sehingga dapat mendukung pengambilan kebijakan yang lebih tepat sasaran [8]. Selanjutnya, penelitian oleh Nirwana (2024) menerapkan algoritma K-Means clustering dalam Evaluasi Hasil Pembelajaran Siswa. Berdasarkan pengujian yang dilakukan, diperoleh nilai silhouette score tertinggi sebesar 0,9168 pada jumlah cluster 2. Hasil ini menunjukkan bahwa pengelompokan data ke dalam kelompok-kelompok tersebut memiliki kualitas yang baik. Penelitian ini juga menegaskan bahwa penggunaan silhouette score sebagai metrik evaluasi dapat menjadi panduan yang efektif dalam menentukan jumlah cluster optimal dalam analisis clustering [9]. Sementara itu, Amri dkk. (2024) melakukan penelitian mengenai Penerapan Algoritma K-Means Clustering dalam Mengelompokkan Smartphone yang Rekomendasi Berdasarkan Spesifikasi. Hasil penelitian menunjukkan bahwa smartphone dapat dikelompokkan ke dalam tiga kategori, yaitu kelompok Rekomendasi (225 item), Sangat Direkomendasikan (98 item), dan Kurang Direkomendasikan (27 item). Temuan ini memberikan manfaat praktis bagi penjual untuk menyesuaikan stok smartphone sesuai kebutuhan konsumen berdasarkan harga dan spesifikasi yang paling diminati [10]. Penelitian lain dilakukan oleh Ananda dkk. (2024) yang berjudul Implementasi Data Mining dengan Algoritma K-Means Clustering untuk Analisis Bisnis pada Perusahaan Asuransi. Penelitian ini berhasil mengelompokkan data pelanggan berdasarkan karakteristik seperti umur, jenis kelamin, dan riwayat klaim. Hasil pengelompokan memberikan wawasan berharga bagi perusahaan dalam menyusun strategi pemasaran yang lebih tepat serta meningkatkan efisiensi operasional. Dengan demikian, penelitian ini memberikan kontribusi signifikan dalam meningkatkan daya saing perusahaan asuransi melalui penerapan teknologi data mining [11].

Berdasarkan penelitian-penelitian tersebut, dapat disimpulkan bahwa algoritma K-Means clustering telah banyak digunakan di berbagai bidang, mulai dari pengelompokan data penduduk, evaluasi pembelajaran, rekomendasi produk, hingga analisis bisnis perusahaan. Namun, penerapan algoritma ini dalam konteks rental mobil, khususnya untuk mengelompokkan kendaraan berdasarkan harga sewa dan jarak tempuh, masih jarang dilakukan. Hal inilah yang menjadi gap dan sekaligus kontribusi dari penelitian ini.

Dengan demikian, penelitian ini tidak hanya memiliki kontribusi teoritis dalam pengembangan penerapan K-Means clustering pada sektor jasa, tetapi juga kontribusi praktis bagi perusahaan rental mobil dalam pengelolaan data dan strategi bisnis. Melalui hasil pengelompokan yang akurat, perusahaan diharapkan dapat lebih memahami kebutuhan pelanggan, mengoptimalkan penggunaan armada, serta meningkatkan daya saing di tengah kompetisi industri rental mobil yang semakin kompetitif.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining merupakan suatu proses untuk menggali dan menemukan pola, pengetahuan, serta informasi yang bermanfaat dari sekumpulan data yang besar. Proses ini dilakukan dengan menggunakan berbagai teknik analisis, seperti klasifikasi, regresi, clustering, asosiasi, dan deteksi anomali [12]. Dalam konteks penelitian ini, data mining digunakan sebagai metode utama untuk mengolah data yang diperoleh dari sistem rental mobil. Dengan memanfaatkan data mining, peneliti dapat mengidentifikasi pola perilaku pelanggan, tingkat penggunaan kendaraan, serta kecenderungan penyewaan berdasarkan periode tertentu.

Tahapan dalam data mining secara umum mengikuti model CRISP-DM (Cross Industry Standard Process for Data Mining) yang meliputi: (1) *Business Understanding*, yaitu memahami tujuan bisnis yang ingin dicapai melalui data mining; (2) *Data Understanding*, yaitu tahap pengumpulan dan eksplorasi data; (3) *Data Preparation*, yaitu membersihkan, memilih, dan mengubah data agar siap untuk diproses; (4) *Modeling*, yaitu penerapan algoritma tertentu untuk menemukan pola dalam data; (5) *Evaluation*, yaitu tahap mengevaluasi hasil model; dan (6) *Deployment*, yaitu mengimplementasikan hasil analisis ke dalam sistem yang berjalan[13][14]. Dengan pendekatan ini, data yang diperoleh dari sistem rental mobil dapat diolah menjadi informasi yang lebih terstruktur dan bermanfaat untuk pengambilan keputusan.

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 [15]. Algoritma *k-means* berguna untuk melakukan *clustering* secara *unsupervised learning* pada sebuah data berdasarkan parameter tertentu. 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][17].

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[18].

K-means clustering banyak digunakan dalam berbagai aplikasi seperti segmentasi pelanggan, pengelompokkan data pasar, analisis citra, pengenalan pola dan masih banyak lagi. Algoritma *k-means clustering* membantu dalam memahami struktur data yang rumit, mengidentifikasi kelompok-kelompok yang sama, memberikan wawasan tentang pola dan hubungan dalam data[19]. Langkah-langkah melakukan *clustering* atau pengelompokkan dengan metode *k-means clustering* sebagai berikut[20]:

- Pilih jumlah *cluster* k.
- Inisialisasi k sebagai *centroid*, yang dilakukan dengan cara acak atau random.
- Menghitung jarak objek ke *centroid* dengan menggunakan rumus *Euclidean Distance* yang dirumuskan sebagai berikut:

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

Keterangan:

De = *Euclidean Distance*

i = Banyaknya objek

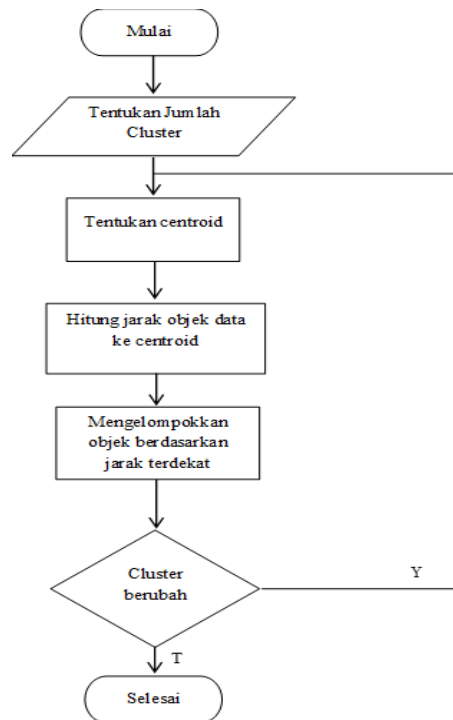
(x, y) = Koordinat objek

(s, t) = Koordinat *centroid*

- Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan centroidnya.
- 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.

2.2 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. Berikut ini gambar *flowchart* dari metode *k-means* seperti yang terlihat pada Gambar 1:



Gambar 1. Flowchart Metode K-Means

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:

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

- 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
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.

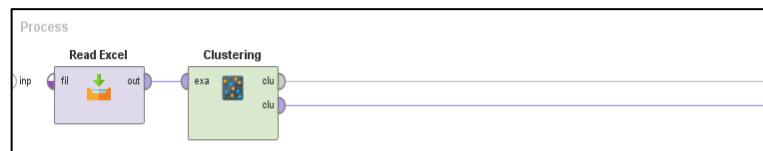
Tabel 4. Hasil Iterasi Ke-1

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:

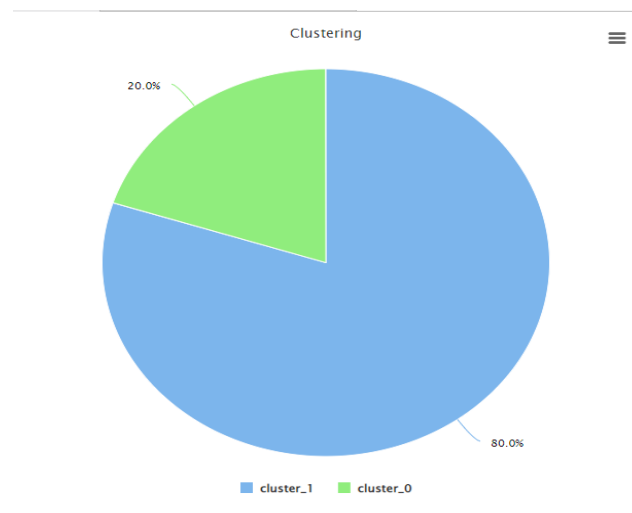
**Gambar 2.** Tampilan Menghubungkan *Read Excel* Dengan *Clustering*

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

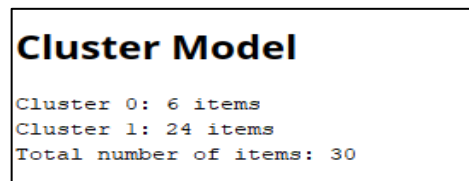
Row No.	Merek Mobil	cluster	Jumlah	Harga sewa	Jarak Temp...
1	Avanza ADJ (...)	cluster_1	25	400000	2376
2	Avanza ADR (...)	cluster_1	20	400000	1893
3	Avanza MAB (...)	cluster_1	23	400000	2826
4	Avanza ADQ (...)	cluster_1	27	400000	1790
5	Xenia AED (M...	cluster_1	18	400000	1160
6	Xenia AEO (M...	cluster_1	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. Tampilah *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:

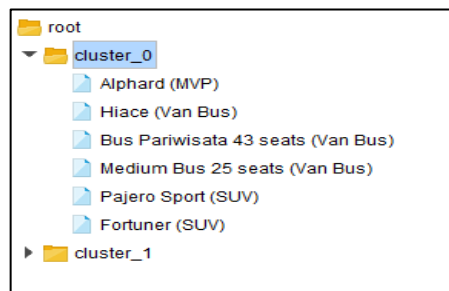
**Gambar 4.** Tampilan *Visualization*

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:



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 6.



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

Mengelompokkan armada rental mobil berdasarkan harga sewa dan jarak tempuh dengan pemanfaatan menggunakan algoritma *k-means* ini dapat membantu PT. Station Armada Indonesia mengelompokkan jenis mobil. Dari hasil pengelompokkan diperoleh dua kelompok *cluster* dengan dengan karakter *cluster* pertama adalah kurang diminati pelanggan dan kelompok *cluster* kedua adalah paling diminati pelanggan. Sehingga perusahaan dapat dengan mudah mempersiapkan jenis armada yang paling diminati. Dalam penerapan metode data *mining* dengan menggunakan *k-means* sangat membantu dan mempermudah PT. Station Armada Indonesia untuk mengembangkan strategi pemasaran dan penawaran yang lebih efektif. Menerapkan metode *k-means* pada RapidMiner sangat optimal dalam pengelolaan armada mobil yang dimana dilakukan dengan memasukkan sampel data yang akan menjadi *database* pada *Microsoft Excel*. Data tersebut kemudian dihubungkan ke dalam *tools* RapidMiner dan akan diolah dalam bentuk *k-means*. Didapatkan hasil bahwa jenis mobil pada bulan agustus yang dibagi menjadi 2 *cluster*. *Cluster* 0 adalah kelompok jenis mobil yang kurang diminati pelanggan yang memiliki 6 jenis mobil dengan rata-rata jarak tempuh pelanggan 673km - 2050km. pada *cluster* 1 kelompok jenis mobil yang paling diminati pelanggan yang memiliki 24 jenis mobil dengan rata-rata jarak tempuh pelanggan 270km-3388km. Dengan mengelompokkan jenis mobil dengan implementasi *k-means* dapat mempermudah pengetahuan pelanggan memilih jenis mobil berdasarkan kebutuhan pelanggan.

REFERENCES

- [1] S. W. Maesaroh, T. M. Diansyah, R. Liza, and Y. F. A. Lubis, "Pemanfaatan Algoritma K-Means Clustering Pada Sistem Rental Mobil," *Bulletin of Computer Science Research*, vol. 5, no. 3, pp. 176–181, 2025, doi: 10.47065/bulletincsr.v5i3.494.
- [2] M. R. Tetlageni and A. Solichin, "Klasterisasi Penyewaan Kendaraan Menggunakan Metode K-Means Pada PT. Mardika Daya Tribuana," *Bit (Fakultas Teknologi Informasi Universitas Budi Luhur)*, vol. 20, no. 2, pp. 141–148, 2023, doi: 10.36080/bit.v20i2.2496.
- [3] C. Hardjono and S. M. Isa, "Implementation of Data Mining for Churn Prediction in Music Streaming Company Using 2020 Dataset," *Journal on Education*, vol. 5, no. 1, pp. 1189–1197, 2022, doi: 10.31004/joe.v5i1.740.
- [4] A. R. Savira, "Optimizing Clustering Models Using Principle Component Analysis for Car Customers," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 18, no. 2, pp. 1–5, 2024, doi: 10.22146/ijccs.94744.
- [5] A. Yudistira and R. Andika, "Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering," *Journal of Artificial Intelligence and Technology Information*, vol. 1, no. 1, pp. 20–28, 2023, doi: 10.58602/jaiti.v1i1.22.
- [6] P. Apriyani, A. R. Dikananda, and I. Ali, "Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi," *Hello World Jurnal Ilmu Komputer*, vol. 2, no. 1, pp. 20–33, 2023, doi: 10.56211/helloworld.v2i1.230.

- [7] S. N. B. Sembiring, H. Winata, and S. Kusnasari, "Pengelompokan Prestasi Siswa Menggunakan Algoritma K-Means," *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, vol. 1, no. 1, pp. 31–40, 2022, doi: 10.53513/jursi.v1i1.4784.
- [8] Z. Sitorus and S. Suhartika, "Penerapan Data Mining Untuk Clustering Penduduk Miskin Di Kota Tanjungbalai Menggunakan Metode Algoritma K-Means," *Journal of Science and Social Research*, vol. 7, no. 1, pp. 212–218, 2024, doi: 10.54314/jssr.v7i1.1732.
- [9] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM)*, vol. 3, no. 1, pp. 46–56, 2024, doi: 10.58602/jima-ilkom.v3i1.26.
- [10] A. M. Nur, M. Saiful, H. Bahtiar, and M. T. Hidayat, "Penerapan Algoritma K-Means Clustering Dalam Mengelompokkan Smartphone Yang Rekomendasi Berdasarkan Spesifikasi," *Infotek: Jurnal Informatika dan Teknologi*, vol. 7, no. 2, pp. 478–488, 2024, doi: 10.29408/jit.v7i2.26283.
- [11] A. A. A. Putri and S. A. Rahmah, "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, doi: 10.46576/djtechno.v5i1.4537.
- [12] D. Marlina and M. Bakri, "Penerapan Data Mining Untuk Memprediksi Transaksi Nasabah Dengan Algoritma C4. 5," *Jurnal Teknologi Dan Sistem Informasi*, vol. 2, no. 1, pp. 23–28, 2021, doi: 10.33365/jtsi.v2i1.627.
- [13] U. Suriani, "Penerapan Data Mining untuk Memprediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma Decision Tree C4. 5," *Journal of Computer and Information Systems Ampera*, vol. 4, no. 2, pp. 55–65, 2023, doi: 10.51519/journalcisa.v4i2.393.
- [14] F. Alghifari and D. Juardi, "Penerapan Data Mining Pada Penjualan Makanan dan Minuman Menggunakan Metode Algoritma Naïve Bayes: Studi Kasus: Makan Barbeque Sepuasnya," *Jurnal Ilmiah Informatika*, vol. 9, no. 02, pp. 75–81, 2021, doi: 10.33884/jif.v9i02.3755.
- [15] S. W. Harjono, N. W. Utami, and I. G. A. P. D. Putri, "Klasterisasi tingkat penjualan pada startup Panak. id dengan algoritma K-Means," *Jurnal Ilmiah Teknologi Informasi Asia*, vol. 17, no. 1, pp. 55–66, 2023, doi: 10.32815/jitika.v17i1.888.
- [16] M. Rosadi, D. A. Nurhasanah, and M. S. Hasibuan, "Clustering Panjang Ruas Jalan di BBPJN Sumut Menggunakan Algoritma K-Means," *Journal of Computer Science and Informatics Engineering*, vol. 2, no. 1, pp. 29–38, 2023, doi: 10.55537/cosie.v2i1.567.
- [17] Y. Sopyan, A. D. Lesmana, and C. Juliane, "Analisis Algoritma K-Means dan Davies Bouldin Index dalam Mencari Cluster Terbaik Kasus Perceraian di Kabupaten Kuningan," *Build. Informatics, Technol. Sci*, vol. 4, no. 3, pp. 1464–1470, 2022, doi: 10.47065/bits.v4i3.2697.
- [18] M. Qusyairi, Z. Hidayatullah, and A. Sandi, "Penerapan K-Means Clustering Dalam Pengelompokan Prestasi Siswa Dengan Optimasi Metode Elbow," *Infotek: Jurnal Informatika dan Teknologi*, vol. 7, no. 2, pp. 500–510, 2024, doi: 10.29408/jit.v7i2.26375.
- [19] M. R. Kusnaldi, T. Gulo, and S. Aripin, "Penerapan Normalisasi Data Dalam Mengelompokkan Data Mahasiswa Dengan Menggunakan Metode K-Means Untuk Menentukan Prioritas Bantuan Uang Kuliah Tunggal," *Journal of Computer System and Informatics (JoSYC)*, vol. 3, no. 4, pp. 330–338, 2022, doi: 10.47065/josyc.v3i4.2112.
- [20] I. W. S. A. Nugraha, "Clustering Pemetaan Tingkat Kemiskinan di Provinsi Jawa Barat Menggunakan Algoritma K-Means," *Jurnal Ilmiah Wahana Pendidikan*, vol. 9, no. 2, pp. 234–244, 2023, doi: 10.5281/zenodo.7567622.