

Analisis Segmentasi Wilayah Penjualan menggunakan Algoritma K-Means Clustering

Erdin Gari¹, Retno Sari²

^{1,2}Universitas Nusa Mandiri / Informatika
e-mail: ¹erdingari599@gmail.com, ²retno.rnr@nusamandiri.ac.id

Diterima	Direvisi	Disetujui
09-10-2025	11-11-2025	22-12-2025

Abstrak - Persaingan bisnis pada era saat ini merupakan tantangan besar pada suatu perusahaan dalam meningkatkan jumlah penjualan suatu produk, sehingga perusahaan dituntut merancang strategi penjualan yang efektif dan tepat sasaran. Salah satu tantangan yang dihadapi oleh PT Masada Abadi, Tangerang, adalah bagaimana menentukan wilayah penjualan kabel listrik yang optimal guna meningkatkan daya saing dan volume penjualan di Tengah persaingan yang cukup kompetitif pada saat ini. Permasalahan yang dihadapi PT Masada Abadi, Tangerang, terdapat pada analisa pola pembelian pelanggan berdasarkan wilayah yang tidak optimal sehingga menghambat efektivitas strategi pemasaran dan distribusi produk. Mengatasi permasalahan tersebut, penelitian ini menerapkan algoritma K-Means clustering sebagai solusi dalam mengoptimalkan wilayah penjualan kabel listrik berdasarkan data. Pemilihan algoritma K-Means clustering bertujuan agar pola penjualan di kelompokkan berdasarkan jumlah penjualan dan total penjualan pada setiap jenis kabel di semua wilayah penjualan sehingga menghasilkan visualisasi kluster yang dapat diinterpretasikan. Hasil pengujian dari penerapan algoritma K-Means clustering diperoleh 3 cluster wilayah penjualan dengan wilayah penjualan paling rendah terdapat 3 wilayah yaitu Tangerang, Surabaya dan Jakarta Selatan, wilayah penjualan sedang terdapat 4 wilayah yaitu Jakarta pusat, Bandung, Jakarta Timur dan Jogjakarta, dan wilayah penjualan paling tinggi terdapat 3 wilayah yaitu Kalimantan, Jakarta Barat dan Bekasi.

Kata Kunci: K-means, Clustering, Penjualan Kabel Listrik.

Abstract - Business competition in the current era is a big challenge for a company in increasing the number of sales of a product, so that the company is required to design an effective and targeted sales strategy. One of the challenges faced by PT Masada Abadi, Tangerang, is how to determine the optimal sales area for electrical cables to increase competitiveness and sales volume in the midst of quite competitive competition at this time. The problem faced by PT Masada Abadi, Tangerang, lies in the analysis of customer purchasing patterns based on non-optimal regions, thus hampering the effectiveness of product marketing and distribution strategies. To overcome this problem, this study applies the K-Means clustering algorithm as a solution in optimizing the sales area for electrical cables based on data. The selection of the K-Means clustering algorithm aims to group sales patterns based on the number of sales and total sales for each type of cable in all sales areas so as to produce interpretable cluster visualization. The test results from the K-Means clustering algorithm yielded three clusters of sales regions. The three regions with the lowest sales were Tangerang, Surabaya, and South Jakarta; the four regions with the medium sales were Central Jakarta, Bandung, East Jakarta, and Yogyakarta; and the three regions with the highest sales were Kalimantan, West Jakarta, and Bekasi.

Keywords: K-means, Clustering, Electrical Cable Sales

PENDAHULUAN

Pemanfaatan data saat ini merupakan kebutuhan yang tidak dapat dihindari untuk mendukung pengambilan keputusan. Data yang dimiliki dianalisis untuk mendapatkan informasi yang bermanfaat. Perusahaan harus mampu memahami karakteristik pelanggan ataupun wilayah penjualannya, karena dua hal ini merupakan hal yang penting.

Pengetahuan mengenai wilayah penjualan, sangat diperlukan perusahaan dalam menentukan keputusan, karena setiap wilayah memiliki karakteristik yang berbeda sehingga teknik dalam pemasarannya pun berbeda. Untuk mengetahui tantangan apa saja yang dihadapi dalam pemasaran produk di wilayah-wilayah diperlukan analisis data. Pendekatan yang dapat dilakukan untuk melakukan analisis data dengan menggunakan data mining.

PT Benteng Masada Abadi memiliki tantangan sendiri dalam mengelola distribusi serta

pemasaran produknya. PT Benteng Masada Abadi belum melakukan analisis segmentasi wilayah sehingga belum ada pengelompokan wilayah. Data penjualan yang ada belum dimanfaatkan dengan maksimal sehingga pengambilan keputusan seperti penentuan prioritas pemasaran dan distribusi menjadi kurang optimal.

Analisis segmentasi dinilai dapat memberikan solusi kepada perusahaan untuk meningkatkan efisiensi strategi pemasaran (Perdana et al., 2022). Segmentasi memungkinkan perusahaan mendapatkan manfaat karena dapat memasarkan produk berdasarkan kebutuhan pelanggan individu (Rohman & Wibowo, 2024). Menganalisis segmentasi pasar menjadi sangat penting untuk memahami siapa yang akan menjadi pangsa pasar dan mengetahui bagaimana kebiasaan untuk mengetahui segmentasi (Annisa & Rusvinasari, 2024).

Data mining merupakan salah satu solusi dalam segmentasi (Ahsina et al., 2022). Metode *clustering* digunakan untuk menemukan struktur klaster dalam suatu data yang dicirikan oleh kesamaan terbesar dalam klaster yang sama dan perbedaan terbesar antara klaster yang berbeda (Sinaga & Yang, 2020). Metode clustering dianggap sangat penting karena mengkategorikan kumpulan data yang besar ke dalam kelompok-kelompok yang mudah dipahami dan bermakna, mengungkap pola dan struktur tersembunyi (Miraftabzadeh et al., 2023).

Metode *clustering* yang banyak digunakan yaitu metode k-means, karena metode yang populer untuk mengelompokkan karena kesederhanaan, efisiensi, skalabilitas dan kemudahan implementasinya (Miraftabzadeh et al., 2023), serta populer untuk identifikasi kesamaan antar objek berdasarkan vektor jarak yang cocok untuk kumpulan data kecil (Abdullah et al., 2022) dan dinilai mudah dimengerti dan efektif dalam mengelompokkan data (Syaharani et al., 2025).

Telah dilakukan penelitian sebelumnya yaitu telah dianalisis proses penerimaan mahasiswa baru belum dioptimalkan. Tidak ada pengelompokan calon mahasiswa berdasarkan karakteristik tertentu yang dapat dijadikan dasar pengambilan keputusan strategis dalam pemilihan program studi. Penelitian ini bertujuan untuk melakukan pengelompokan data mahasiswa baru dengan menerapkan pendekatan K-Means *Clustering* dan bantuan aplikasi RapidMiner, serta mencari jumlah kluster (K) optimal berdasarkan evaluasi Davies Bouldin Index. Hasil penelitian menunjukkan bahwa penerapan algoritma K-Means menghasilkan 2 kluster optimal dengan nilai Davies Bouldin Index = -0.278. Penerapan metode K-Means *Clustering* pada penelitian ini akan memberikan dampak positif terhadap peningkatan dan efisiensi di masa mendatang (Ilmiyah & Bahtiar, 2024).

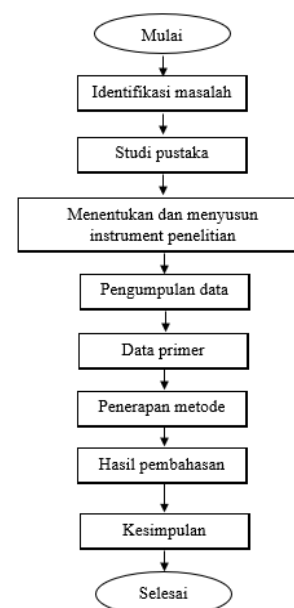
Kampus STMIK Primakara membutuhkan strategi pemasaran yang baik dalam melakukan promosi untuk mendapatkan target mahasiswa agar

lebih efektif dan efisien. Pendekatan yang belum digunakan adalah pola pendaftaran mahasiswa berdasarkan wilayah, program studi dan gelombang. Penelitian ini menerapkan metode K-Means *Clustering* yang bertujuan mengelompokkan data mahasiswa baru dengan menerapkan *clustering* guna mengetahui wilayah asal, program studi pilihan dan Waktu pendaftaran paling diminati, sehingga hasil pengelompokan dapat dijadikan dasar untuk mengatur strategi proses promosi yang lebih efektif dan tepat sasaran. Hasil penelitian yang telah dilakukan menghasilkan 3 cluster yaitu cluster 0 sebanyak 906 mahasiswa, cluster 1 sebanyak 28 mahasiswa dan cluster 2 sebanyak 77 mahasiswa. Berdasarkan hasil pengelompokan ini, disimpulkan bahwa mahasiswa terbanyak berasal dari Denpasar Timur dan Gianyar dengan program studi sistem informasi dan teknik informatika serta mendaftar di gelombang pertama, mahasiswa paling sedikit berasal dari Kupang dan Larantuka dengan program studi sistem informasi dan teknik informatika serta mendaftar di gelombang pertama dan mahasiswa dengan jumlah rata-rata berasal dari Bekasi dengan program studi sistem informasi dan teknik informatika serta mendaftar di gelombang pertama (Tensao et al., 2022).

Penelitian ini bertujuan untuk mengetahui segmentasi wilayah penjualan agar mendapatkan solusi yang tepat untuk mengatasi wilayah penjualan yang rendah dan memaksimalkan penjualan pada wilayah penjualan tertinggi.

METODE PENELITIAN

Tahapan penelitian dapat dilihat pada gambar berikut:



Sumber: (hasil Penelitian, 2025)

Gambar 1 Tahapan Penelitian

Berdasarkan gambar tahapan penelitian, berikut adalah uraian mengenai tahapan penelitian dari setiap proses yang dilakukan :

1. Identifikasi masalah
Identifikasi masalah dilakukan melalui analisis terhadap kondisi perusahaan, di mana ditemukan ketidaksesuaian antara potensi data transaksi wilayah penjualan yang dimiliki dengan pemanfaatannya saat ini. Identifikasi masalah dilakukan dengan mengetahui PT Benteng Masada Abadi memiliki data transaksi penjualan yang tersebar di berbagai wilayah, namun belum dilakukan pengelompokan atau segmentasi wilayah berdasarkan potensi penjualan dan lokasi geografis. Akibatnya, strategi pemasaran belum optimal karena tidak diperolehnya data yang akurat. Kebutuhan pengelompokan wilayah menggunakan metode analisis seperti K-Means *Clustering* untuk membantu mengelompokkan wilayah berdasarkan kesamaan pola transaksi dan lokasi, sehingga perusahaan dapat fokus pada wilayah dengan potensi penjualan tertinggi.
2. Studi Pustaka
Tahapan yang dimana peneliti melakukan pencarian berbagai literatur atau sumber-sumber informasi yang berkaitan dengan topik penelitian. Sumber-sumber dapat berupa buku, artikel jurnal, disertasi, tesis, laporan penelitian, dan sumber-sumber terpercaya lainnya yang terkait dengan topik penelitian.
3. Menentukan dan menyusun instrumen penelitian
Pada tahapan ini, instrumen penelitian ditentukan berdasarkan wilayah penjualan, jumlah transaksi, volume penjualan, tipe kabel listrik, merek kabel listrik dan tipe konsumen berdasarkan data yang ada pada PT Benteng Masada Abadi dan data tersebut diperoleh peneliti dengan melakukan pengambilan data laporan penjualan langsung di PT Benteng Masada Abadi.
4. Pengumpulan data
Data dikumpulkan berdasarkan kebutuhan pada instrumen penelitian yang telah ditentukan sebelumnya dimana data yang dikumpulkan adalah data primer atau data yang diperoleh langsung dari PT Benteng Masada Abadi. Teknik pengumpulan data yang digunakan penulis merupakan *purposive sampling* dimana data tersebut dikumpulkan dalam periode Januari 2024 hingga Desember 2024. Proses pengumpulan dilakukan secara sistematis sehingga hasil analisis dapat mencerminkan kondisi secara akurat dalam periode waktu tersebut.
5. Data Primer
Data primer adalah data yang diperoleh peneliti dari sumbernya aslinya secara

langsung berdasarkan hasil wawancara, observasi, survey atau eksperimen. Teknik pengumpulan data yang digunakan penulis merupakan *purposive sampling* dimana data tersebut dikumpulkan dalam periode Januari 2024 hingga Desember 2024. Proses pengumpulan dilakukan secara sistematis sehingga hasil analisis dapat mencerminkan kondisi secara akurat dalam periode waktu tersebut.

6. Penerapan Metode
Metode K-Means *clustering* di terapkan secara manual oleh peneliti dan kemudian dilakukan pengujian menggunakan aplikasi *RapidMiner* versi 10.3.1. Tahapan manual dan menggunakan *RapidMiner* dilakukan penulis untuk mengetahui kebenaran hasil perhitungan manual dengan hasil pengujian sehingga dapat diperoleh kesimpulan yang dapat dibuktikan terkait hasil yang diperoleh memiliki hasil yang sama atau memiliki perbedaan.
7. Hasil pembahasan
Hasil pembahasan merupakan analisa peneliti berdasarkan penerapan metode dan pengujian model. Pada tahapan ini penulis menguraikan seluruh tahapan dari proses perhitungan dan pengujian hingga memperoleh sebuah keputusan.
8. Kesimpulan
Kesimpulan merupakan tahap akhir yang diambil peneliti dari hasil penelitian dimana pada tahapan ini penulis memberikan poin penting dari hasil penelitian yang dilakukan sehingga dapat diketahui secara ringkas terkait hasil penelitian.

HASIL DAN PEMBAHASAN

A. Sample Data

Sampel data yang digunakan adalah data penjualan kabel listrik dari bulan Januari sampai dengan Desember 2024 dengan total jenis kabel adalah 11 dan total wilayah penjualan adalah 10. Sampel data dapat dilihat pada tabel 1 Sampel Data

B. Normalisasi Sampel Data

Sebelum menerapkan metode K-Means sampel data dilakukan normalisasi dengan menggunakan *decimal scaling method* yang mana data dibagi dengan 1.000.000 atau (10^6) , sehingga nilai yang awalnya 10.000.000 menjadi 10. Normalisasi ini bertujuan agar data menjadi sederhana dan mudah dipahami serta sederhana. Hasil normalisasi dapat dilihat pada tabel 2 sampel data hasil normalisasi

C. Penerapan Algoritma K-Means

1) Menentukan jumlah *cluster*

Jumlah *cluster* yang digunakan sebanyak 3 yaitu:

- C1 = tinggi
- C2 = sedang
- C3 = rendah

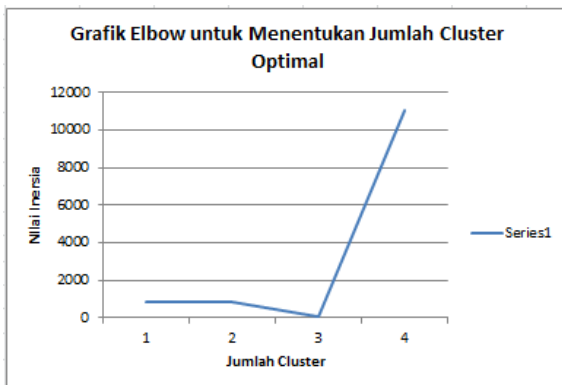
Keterangan :

Cluster dibagi 3 dilakukan untuk menentukan dan melakukan perbandingan setiap wilayah tingkat penjualan rendah, sedang dan tinggi. Pembagian cluster ditentukan berdasarkan total penjualan dari setiap wilayah penjualan dimana total penjualan pada setiap wilayah bergantung pada jumlah kebutuhan yang diminta dan total transaksi pada setiap wilayah penjualan.

Dalam menentukan berapa banyak cluster yang digunakan dengan menggunakan metode Elbow, yang dapat dilihat pada tabel berikut:

Tabel 3 Hasil nilai perhitungan dengan metode Elbow pada Setiap Cluster

Jumlah Cluster	Nilai Inersia
1	873,33
2	884,6666667
3	31,77777778
4	11033,98378



Gambar 1 Grafik Elbow untuk menentukan jumlah cluster optimal

Dari hasil grafik dapat dilihat titik elbow berada pada K=3, menunjukkan bahwa jumlah cluster optimal adalah 3.

- 2) Inisialisasi centroid secara acak
Centroid awal :

Tabel 3 Centroid awal

C	1	1	2	5	1	0	5	5	1	3	7
1	0	5	5	0	5	0	5	5	0	3	7
C	2	1	2	3	1	2	1	1	2	1	4
2	6	8	8	6	3	2	7	4	0	9	4
C	3	2	3	4	1	3	2	2	2	1	1
3	5	5	5	5	8	0	5	0	6	2	8

Sumber : (Hasil Penelitian, 2025)

- 3) Menghitung jarak

Jarak dihitung menggunakan rumus jarak Euclidean (Euclidean Distance). Berikut adalah persamaan Euclidian Distance :

$$d(x_i, u_j) = \sqrt{\sum (x_i - u_j)^2}$$

keterangan :

x_i : data kriteria

u_j : centroid pada cluster ke-j

Perhitungan iterasi 1 dapat dilihat pada tabel dibawah ini :

Tabel 4 Iterasi 1

Data	C1	C2	C3	MIN	CLUSTER
W1	0	38.31	60.84	0	1
W2	49.98	14.56	14.97	14.56	2
W3	32.51	5.92	28.44	5.92	2
W4	38.31	0	22.58	0	2
W5	53.22	15.03	7.75	7.75	3
W6	7.75	31.18	53.70	7.75	1
W7	21.10	17.64	40.11	17.64	2
W8	26.76	11.83	34.32	11.83	2
W9	60.84	22.58	0.00	0	3
W10	14	24.86	47.35	14	1

Sumber: (Hasil Penelitian, 2025)

Langkah selanjutnya adalah melakukan pembaharuan nilai centroid, berikut rumus yang digunakan untuk memperbaharui nilai centroid :

$$u_j(t+1) = \frac{1}{n_{sj}} \sum_{j \in s_j} x_j$$

Keterangan :

$u_j(t+1)$: centroid baru pada iterasi ke (t+1)

n_{sj} : banyak data pada cluster sj.

Diketahui :

C1 = (W1, W6, W10)

C2 = (W2, W3, W4, W7, W8)

C3 = (W5, W9)

Maka :

C1 = (W1+W6+W10)/3

C1, K1 = (10+12+15) / 3
= 12

C2 = (W2+W3+W4+W7+W8) / 5

C2, K1 = (30+24+26+20+22) / 5
= 24

C3 = (W5, W9) / 2

C3, K1 = (32+35) / 2
= 34

Berikut nilai hasil perhitungan centroid baru dapat dilihat pada tabel dibawah ini :

Tabel 5 Nilai centroid baru

C	1	1	2	1	2	1	1	1	1	1	1
1	2	8	8	4	6	2	7	6	1	4	8
C	2	1	2	3	1	2	1	1	1	1	1
2	4	6	6	4	2	0	5	2	8	8	3
C	3	2	3	4	1	2	2	1	2	1	1
3	4	4	4	4	8	9	4	9	5	2	8

Perhitungan iterasi 2 dapat dilihat pada tabel dibawah ini :

Tabel 6 Iterasi 2

Data	C1	C2	C3	MIN	CLUSTER	KETERANGAN
W1	7.23	33.52	57.03	7.23	1	TETAP
W2	43.28	17.68	12.25	12.25	3	BERUBAH
W3	25.58	2.31	24.64	2.31	2	TETAP
W4	31.40	5.17	18.79	5.17	2	TETAP
W5	46.33	19.87	3.87	3.87	3	TETAP
W6	0.94	26.43	49.90	0.94	1	TETAP
W7	14.16	12.87	36.30	12.87	2	TETAP
W8	19.82	7.34	30.51	7.34	2	TETAP
W9	53.94	27.44	3.87	3.87	3	TETAP
W10	6.80	20.08	43.55	6.80	1	TETAP

Sumber: (Hasil Penelitian, 2025)

Pembaharuan centroid iterasi 2

Tabel 7 Nilai centroid baru

C	1	2	3	4	5	6	7	8	9	10	11	12
C1	2	8	8	4	6	2	7	6	1	4	8	
C2	2	1	2	3	1	1	1	1	1	1	7	1
C3	3	5	5	3	2	9	4	1	7	5	2	
C4	3	2	3	4	1	2	2	2	2		1	
C5	2	2	2	2	7	8	2	1	4	11	7	

Sumber: (Hasil Penelitian, 2025)

Iterasi 3

Perhitungan iterasi 3 dapat dilihat pada tabel dibawah ini :

Tabel 8 Iterasi 3

Data	C1	C2	C3	MIN	CLUSTER	KETERANGAN
W1	7	30	54	7	1	TETAP
W2	43	22	8	8	3	TETAP
W3	26	3	22	3	2	TETAP
W4	31	9	17	9	2	TETAP
W5	46	24	3	3	3	TETAP
W6	1	22	47	1	1	TETAP
W7	14	9	34	9	2	TETAP
W8	20	3	28	3	2	TETAP
W9	54	31	7	7	3	TETAP
W10	7	16	41	7	1	TETAP

Sumber: (Hasil Penelitian, 2025)

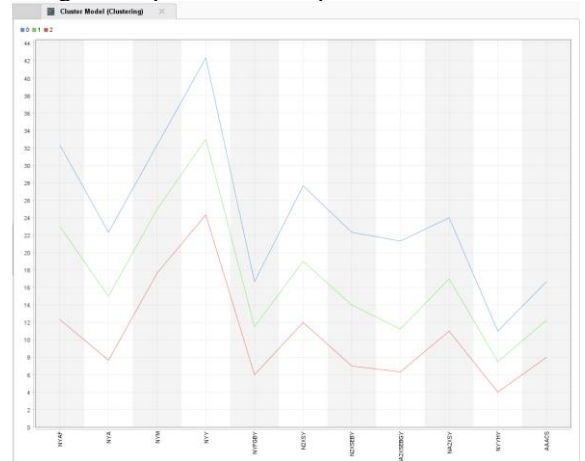
Pada iterasi 3 tidak terjadi perubahan dan proses perhitungan jarak selesai.

- 4) Pengujian Pada Aplikasi Rapidminer
 - a. Centroid tabel yang dihasilkan

Centroid tabel yang dihasilkan merupakan baris-baris tabel yang menunjukan nilai rata-rata atribut pada setiap cluster 0, cluster 1 dan cluster 2. Gambar centroid tabel dapat dilihat pada gambar 2.

- b. Centroid plot yang dihasilkan

Centroid plot merupakan visualisasi dari setiap atribut jenis kabel yang telah dilakukan pengujian pada data uji, sehingga pada centroid plot dapat memberikan pendekatan visualisasi dalam menganalisa perbedaan setiap cluster.

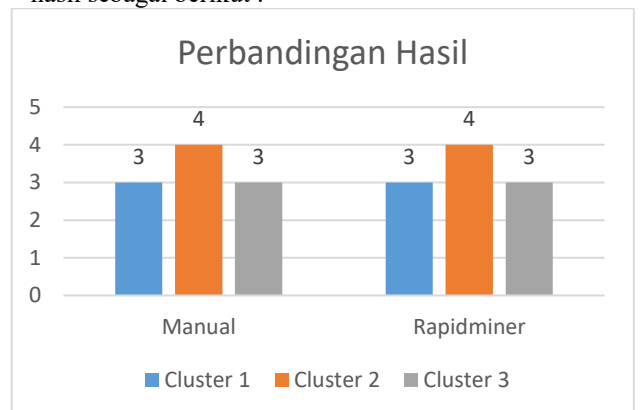


Sumber: (Hasil Penelitian, 2025)

Gambar 3 Centroid plot yang dihasilkan

Potensi Hasil

Berdasarkan kedua tahapan yang dilakukan yaitu perhitungan manual dan pengujian menggunakan RapidMiner diperoleh perbandingan hasil sebagai berikut :



Sumber: (Hasil Penelitian, 2025)

Gambar 4 Grafik perbandingan hasil

Penerapan K-means dan perbandingan hasil yang diperoleh dari gambar diatas dapat diperoleh potensi hasil sebagai berikut:

1. Hasil penerapan metode K-Means clustering dan pengujian menggunakan aplikasi RapidMiner memiliki hasil output yang sama yaitu cluster 1 sebanyak 3 , cluster 2 sebanyak 4 dan cluster 3 sebanyak 3.

2. Cluster C1 sebanyak 3 cluster yaitu cluster rendah merupakan presentasi penjualan paling rendah, cluster sedang merupakan rata-rata penjualan dan cluster tinggi merupakan presentasi penjualan tertinggi kabel listrik. Untuk penjualan tertinggi yaitu W2 (Kalimantan), W5 (Jakarta Barat), W9 (Bekasi). Untuk penjualan kabel dengan penjualan sedang yaitu cluster C2 sebanyak 4 cluster yaitu W3 (Jakarta Pusat), W4 (Bandung), W7(Jakarta Timur), W8 (Jogjakarta), yang merupakan cluster terendah atau wilayah penjualan kabel sedikit yaitu cluster C3 sebanyak 3 cluster yaitu. diperoleh W1 (Tangerang), W6 (Surabaya), W10 (Jakarta Selatan)
3. Pada *centroid plot view* hasil pengujian *RapidMiner* diperoleh bahwa kabel penjualan terendah adalah kabel jenis AAACS dan kabel penjualan tertinggi adalah NYY.

KESIMPULAN

Pengelompokan pada segmentasi wilayah dilakukan dengan mengelompokan wilayah penjualan dan total penjualan untuk setiap jenis kabel sehingga hasil pengelompokan memperoleh data wilayah dan data total penjualan pada setiap jenis kabel serta total penjualan kabel secara keseluruhan. Pada penerapan menggunakan algoritma K-Means *clustering* data penjualan dilakukan normalisasi terlebih dahulu kemudian dibagi menjadi 3 cluster yaitu rendah, sedang dan tinggi. Hasil pengujian dari penerapan algoritma K-Means *clustering* diperoleh 3 cluster dimana wilayah penjualan dengan penjualan paling rendah C3 terdapat 3 wilayah yaitu Tangerang, Surabaya dan Jakarta Selatan, wilayah penjualan sedang C2 terdapat 4 wilayah yaitu Jakarta Pusat, Bandung, Jakarta Timur dan Jogjakarta, dan wilayah penjualan paling tinggi C1 terdapat 3 wilayah yaitu Kalimantan, Jakarta Barat dan Bekasi.

REFERENSI

Abdullah, D., Susilo, S., Ahmar, A. S., Rusli, R., & Hidayat, R. (2022). The application of K-means clustering for province clustering in Indonesia of the risk of the COVID-19 pandemic based on COVID-19 data. *Quality and Quantity*, 56(3), 1283–1291. <https://doi.org/10.1007/s11135-021-01176-w>

Ahsina, N., Fatimah, F., & Rachmawati, F. (2022). Analisis Segmentasi Pelanggan Bank Berdasarkan Pengambilan Kredit Dengan

Menggunakan Metode K-Means Clustering. *Jurnal Ilmiah Teknologi Infomasi Terapan*, 8(3). <https://doi.org/10.33197/jitter.vol8.iss3.2022.883>

Annisa, L. H., & Rusvinasari, D. (2024). Segmentasi Pembelian Produk Menggunakan Algoritma K-Means Berdasarkan Clusterisasi pada pemilihan menu yang ada diUMKM Kuliner. *Jurnal Informatika: Jurnal Pengembangan IT*, 9(3), 203–212. <https://doi.org/10.30591/jpit.v9i3.6556>

Ilmiyah, E., & Bahtiar, A. (2024). Penerapan Algoritma K-Means Clustering Untuk Mengelompokkan Data Mahasiswa Baru. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 1543–1547. <https://doi.org/10.36040/jati.v8i2.9005>

Miraftabzadeh, S. M., Colombo, C. G., Longo, M., & Foiadelli, F. (2023). K-Means and Alternative Clustering Methods in Modern Power Systems. *IEEE Access*, 11(September), 119596–119633. <https://doi.org/10.1109/ACCESS.2023.3327640>

Perdana, S. A., Florentin, S. F., & Santoso, A. (2022). Analisis Segmentasi Pelanggan Menggunakan K-Means Clustering Studi Kasus Aplikasi Alfagift. *Sebatik*, 26(2), 446–457. <https://doi.org/10.46984/sebatik.v26i2.2134>

Rohman, N., & Wibowo, A. (2024). Perbandingan Metode K-Medoids dan Metode K-Means Dalam Analisis Segmentasi Pelanggan Mall. *SINTECH (Science and Information Technology) Journal*, 7(1), 49–58. <https://doi.org/10.31598/sintechjournal.v7i1.1507>

Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access*, 8, 80716–80727. <https://doi.org/10.1109/ACCESS.2020.2988796>

Syahrani, A., Ilhadi, V., & Karima, A. (2025). Analisis Clustering Segmentasi Pelanggan Pada Cv Prima Jaya Utama Menggunakan Metode K-Means. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(5), 8847–8854. <https://doi.org/10.36040/jati.v9i5.15238>

Tensao, O. O., Wijaya, I. N. Y. A., & Fredlina, K. Q. (2022). Analisa Data Mining dengan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru Pada STMIK Primakara. *INFORMASI (Jurnal Informatika Dan Sistem Informasi)*, 14(1), 1–17.

Tabel 1 Sampel Data

No	Wilayah	NYAF	NYA	NYM	NYN	NYFGBY	N2XSY	N2XSEBY	NA2XSEBGY	NA2XSY	NYHYH	AAACS	Total Penjualan
1	Tangerang	10.jt	5.jt	15.jt	20.jt	5.jt	10.jt	5.jt	5.jt	10.jt	3.jt	7.jt	95.jt

No	Wilayah	NYAF	NYA	NYM	NYY	NYFGBY	N2XSY	N2XSEBY	NA2XSEBGY	NA2XSY	NYHYH	AAACS	Total Penjualan
2	Kalimantan	30.jt	20.jt	30.jt	40.jt	15.jt	25.jt	20.jt	26.jt	22.jt	10.jt	15.jt	253.jt
3	Jakarta Pusat	24.jt	16.jt	26.jt	34.jt	12.jt	20.jt	15.jt	12.jt	18.jt	8.jt	13.jt	198.jt
4	Bandung	26.jt	18.jt	28.jt	36.jt	13.jt	22.jt	17.jt	14.jt	20.jt	9.jt	14.jt	217.jt
5	Jakarta Barat	32.jt	22.jt	32.jt	42.jt	17.jt	28.jt	22.jt	18.jt	24.jt	11.jt	17.jt	265.jt
6	Surabaya	12.jt	8.jt	18.jt	25.jt	6.jt	12.jt	7.jt	6.jt	11.jt	4.jt	8.jt	117.jt
7	Jakarta Timur	20.jt	12.jt	22.jt	30.jt	10.jt	16.jt	11.jt	9.jt	14.jt	6.jt	10.jt	160.jt
8	Jogjakarta	22.jt	14.jt	24.jt	32.jt	11.jt	18.jt	13.jt	10.jt	16.jt	7.jt	12.jt	179.jt
9	Bekasi	35.jt	25.jt	35.jt	45.jt	18.jt	30.jt	25.jt	20.jt	26.jt	12.jt	18.jt	289.jt
10	Jakarta Selatan	15.jt	10.jt	20.jt	28.jt	7.jt	14.jt	9.jt	8.jt	12.jt	5.jt	9.jt	137.jt

Tabel 2 Sampel data hasil normalisasi

DATA	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11
W1	10	5	15	20	5	10	5	5	10	3	7
W2	30	20	30	40	15	25	20	26	22	10	15
W3	24	16	26	34	12	20	15	12	18	8	13
W4	26	18	28	36	13	22	17	14	20	9	14
W5	32	22	32	42	17	28	22	18	24	11	17
W6	12	8	18	25	6	12	7	6	11	4	8
W7	20	12	22	30	10	16	11	9	14	6	10
W8	22	14	24	32	11	18	13	10	16	7	12
W9	35	25	35	45	18	30	25	20	26	12	18
W10	15	10	20	28	7	14	9	8	12	5	9

Sumber : (Hasil Penelitian, 2025)

Attribute	cluster_0	cluster_1	cluster_2
NYAF	32.333	23	12.333
NYA	22.333	15	7.667
NYM	32.333	25	17.667
NYY	42.333	33	24.333
NYFGBY	16.667	11.500	6
N2XSY	27.667	19	12
N2XSEBY	22.333	14	7
NA2XSEBGY	21.333	11.250	6.333
NA2XSY	24	17	11
NYHYH	11	7.500	4
AAACS	16.667	12.250	8

Sumber: (Hasil Penelitian, 2025)

Gambar 2 Centroid tabel yang dihasilkan