



Perbandingan Algoritma K-Means dan K-Medoids Clustering Penjualan Menu Restoran Bebek Cak Selem Gresik

Maratus Solikhah Tiara Wulan Sari^{1*}, Dhian Satria Yudha Kartika², Agung Brastama Putra³

¹⁻³Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur, Indonesia

E-mail: tiarawulansari999@gmail.com¹, dhian.satria@upnjatim.ac.id², agungbp@upnjatim.ac.id³

*Penulis Korespondensi: 19082010103@student.upnjatim.ac.id

Abstract. Restoran Bebek Cak Selem Gresik in Gresik, East Java, recorded 11,339 transactions across 39 menu items from January to May 2025, yet this data had never been used as a basis for sales strategy. This study applies and compares K-Means and K-Medoids to group menu items by units sold and price, using the CRISP-DM framework. The optimal number of clusters was determined via the Elbow Method and evaluated using Silhouette Score, Davies-Bouldin Index, and Adjusted Rand Index (ARI). K=3 was found optimal, with both algorithms producing an identical Silhouette Score of 0.5368, Davies-Bouldin Index of 0.5861, and ARI of 1.0, proving both generate exactly the same cluster partition. Three menu segments were formed: 8 high-selling, 19 medium-selling, and 12 low-selling menus, each with a distinct sales-strategy recommendation, translated into an interactive HTML dashboard in owner-friendly language. K-Means is the more practical choice for restaurant-scale data, delivering equal clustering quality at substantially lower computation cost.

Keywords: Clustering; CRISP-DM; K-Means; K-Medoids; Restaurant.

Abstrak. Restoran Bebek Cak Selem Gresik mencatat 11.339 transaksi dari 39 menu selama Januari-Mei 2025, namun data tersebut belum pernah dimanfaatkan sebagai dasar strategi penjualan. Penelitian ini menerapkan dan membandingkan algoritma K-Means dan K-Medoids untuk mengelompokkan menu berdasarkan jumlah terjual dan harga, menggunakan kerangka kerja CRISP-DM. Jumlah cluster optimal ditentukan melalui Elbow Method dan dievaluasi menggunakan Silhouette Score, Davies-Bouldin Index, serta Adjusted Rand Index (ARI). Hasil penelitian menunjukkan K=3 sebagai jumlah cluster optimal, dengan kedua algoritma menghasilkan Silhouette Score 0,5368, Davies-Bouldin Index 0,5861, dan ARI sebesar 1,0, membuktikan kedua algoritma menghasilkan partisi cluster yang identik. Terbentuk tiga segmen menu, yaitu 8 menu penjualan tinggi, 19 menu penjualan sedang, dan 12 menu penjualan rendah, masing-masing disertai rekomendasi strategi penjualan berbeda, dan diterjemahkan ke *dashboard* interaktif berbasis HTML dengan bahasa yang mudah dipahami pemilik restoran. Algoritma K-Means menjadi pilihan lebih praktis karena menghasilkan kualitas klaster setara K-Medoids dengan waktu komputasi jauh lebih singkat.

Kata Kunci: Clustering; CRISP-DM; K-Means; K-Medoids; Restoran.

1. LATAR BELAKANG

Sektor kuliner Indonesia tumbuh pesat, dengan industri makanan dan minuman menjadi salah satu penyumbang PDB terbesar dan tulang punggung UMKM di seluruh Indonesia (Badan Pusat Statistik, 2023). Persaingan yang ketat menuntut pelaku usaha kuliner memahami pola permintaan pelanggan dan merespons dinamika pasar secara cepat dan berbasis data (Sibarani et al., 2024). Banyak restoran skala menengah ke bawah mencatat data transaksi rutin melalui sistem kasir, namun data tersebut hanya disimpan sebagai arsip administratif tanpa diolah menjadi informasi bernilai strategis, sehingga strategi penjualan yang disusun cenderung intuitif dan subjektif (Wisudawati et al., 2024).

Salah satu teknik data mining yang dapat digunakan untuk menganalisis pola penjualan secara sistematis adalah clustering, yaitu pengelompokan data ke dalam beberapa kelompok

berdasarkan kemiripan karakteristiknya. Triyandana et al. (2022) menerapkan K-Means untuk mengelompokkan menu di DPom *Coffee* berdasarkan tingkat penjualan sebagai dasar kebijakan stok dan strategi penjualan. Ana et al. (2024) membuktikan klasterisasi menu di Restoran Sumoboo menggunakan K-Means menghasilkan keputusan lebih akurat mengenai menu prioritas dan manajemen stok. Gantara & Ali (2023) juga menerapkan K-Means untuk mengelompokkan barang ritel berdasarkan tingkat penjualan guna mendukung keputusan stok.

Dua algoritma clustering yang paling banyak digunakan dan sering diperbandingkan performanya adalah K-Means dan K-Medoids. K-Means populer karena unggul dalam kecepatan komputasi dan kemudahan implementasi, namun sensitif terhadap *outlier* karena menggunakan nilai rata-rata sebagai pusat cluster. K-Medoids merupakan pengembangan yang lebih robust terhadap outlier karena menggunakan objek aktual dataset sebagai medoid, bukan nilai rata-rata (Fathurrahman et al., 2023). Perbandingan kedua algoritma telah banyak dilakukan pada berbagai domain, di antaranya data obat di puskesmas (Farissa et al., 2021), pemetaan sebaran Covid-19 (Fathurrahman et al., 2023), data transaksi bongkar muat (Kamila et al., 2019), sebaran cacat pada anak (Marlina et al., 2018), strategi promosi pendidikan (Anggraeni & Tresnawati, 2020), serta penerapan CRISP-DM pada klasterisasi UMKM (Iswavigra et al., 2023). Namun penerapannya pada studi kasus penjualan menu restoran UMKM masih relatif terbatas, menjadi celah penelitian (*research gap*) yang coba dijawab pada penelitian ini.

Selain proses pengelompokan data, penyajian hasil analisis dalam format yang mudah dipahami pemilik restoran juga merupakan aspek penting. Visualisasi data melalui *dashboard* interaktif terbukti mampu menjembatani kesenjangan antara hasil analisis data kompleks dengan kebutuhan pengguna bisnis yang memerlukan informasi mudah dipahami dan dapat langsung ditindaklanjuti (Akbar & Octaviany, 2021).

Restoran Bebek Cak Selem Gresik merupakan restoran kuliner khas bebek di Kabupaten Gresik, Jawa Timur, dengan 39 jenis menu berharga Rp4.000-Rp40.000. Berdasarkan data transaksi Januari-Mei 2025, tercatat 11.339 entri transaksi (11.083 setelah pembersihan duplikat). Meskipun volume transaksinya tinggi, restoran ini belum memiliki sistem analisis data penjualan yang terstruktur, sehingga pemilik tidak mengetahui secara pasti menu yang paling berkontribusi terhadap pendapatan, menu yang berpotensi ditingkatkan melalui promosi, maupun menu yang perlu dievaluasi karena penjualannya rendah. Kondisi ini juga menyulitkan pengendalian stok bahan baku, sehingga sering terjadi kelebihan atau kekurangan persediaan.

Berdasarkan uraian tersebut, penelitian ini bertujuan menerapkan dan membandingkan algoritma K-Means dan K-Medoids dengan kerangka CRISP-DM untuk mengelompokkan 39 menu Bebek Cak Seleem Gresik berdasarkan jumlah terjual dan harga ke dalam kategori penjualan tinggi, sedang, dan rendah, serta menentukan algoritma paling optimal melalui Silhouette Score, Davies-Bouldin Index, dan Adjusted Rand Index. Hasil pengelompokan diterjemahkan menjadi rekomendasi strategi penjualan untuk setiap kelompok menu dan divisualisasikan dalam *dashboard* interaktif berbasis HTML.

2. KAJIAN TEORITIS

Penelitian ini menggunakan kerangka kerja CRISP-DM (*Cross Industry Standard Process for Data Mining*), yang terdiri atas enam tahap: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment* (Ansori & Wulandari, 2022; Singgalen, 2023). Kerangka ini banyak digunakan karena memberikan alur kerja data mining yang sistematis, mulai dari pemahaman kebutuhan bisnis hingga penerapan hasil analisis ke dalam pengambilan keputusan.

K-Means bekerja dengan menentukan centroid awal secara acak, menghitung jarak setiap data ke centroid, mengelompokkan data ke cluster terdekat, lalu memperbarui centroid berdasarkan rata-rata anggota cluster secara iteratif hingga konvergen (Wakhidah, 2010). K-Medoids bekerja dengan prinsip serupa, namun pusat cluster (medoid) merupakan objek data aktual yang paling representatif, sehingga lebih tahan terhadap outlier dibandingkan K-Means (Marlina et al., 2018). Jumlah cluster optimal umumnya ditentukan melalui Elbow Method, yaitu dengan mengamati titik belokan pada grafik Within-Cluster Sum of Squares (WCSS) terhadap jumlah cluster (Maori & Evanita, 2023).

Kualitas hasil clustering dapat dievaluasi menggunakan Silhouette Score, yang mengukur seberapa baik suatu data cocok dengan cluster-nya sendiri dibandingkan cluster terdekat lainnya, dengan nilai mendekati 1 menunjukkan struktur cluster yang semakin baik (Mulyani et al., 2023). Davies-Bouldin Index (DBI) mengukur rasio antara sebaran dalam cluster dan jarak antar cluster, dengan nilai yang semakin kecil menunjukkan pemisahan antar cluster yang semakin baik (Hasan, 2024). Adjusted Rand Index (ARI) digunakan untuk mengukur kesamaan keanggotaan cluster antara dua hasil pengelompokan yang berbeda; nilai 1 menunjukkan kedua hasil identik, sedangkan nilai mendekati 0 setara dengan pengelompokan acak. Pendekatan evaluasi serupa turut digunakan Yolandari et al. (2025) dalam membandingkan performa K-Means dan DBSCAN pada data ulasan wisatawan menggunakan Silhouette Index dan Davies-Bouldin Index.

Implementasi algoritma clustering pada penelitian ini menggunakan bahasa pemrograman Python dengan pustaka scikit-learn dan scikit-learn-extra, yang umum digunakan dalam *machine learning* karena kemudahan penggunaan dan ketersediaan pustaka analisis data yang lengkap (Alfarizi et al., 2023). Hasil clustering selanjutnya divisualisasikan melalui *dashboard* interaktif berbasis web, yang terbukti efektif menyampaikan hasil analisis data penjualan kepada pengguna non-teknis dalam berbagai studi kasus UMKM (Tsani et al., 2025).

3. METODE PENELITIAN

Penelitian ini merupakan penelitian terapan (*applied research*) yang menggunakan pendekatan kuantitatif dengan kerangka kerja CRISP-DM. Data yang digunakan adalah data transaksi penjualan menu Restoran Bebek Cak Selem Gresik periode Januari-Mei 2025 yang bersumber dari sistem kasir internal restoran, berjumlah 11.339 entri transaksi dengan atribut ID Transaksi, Tanggal, Nama Menu, Kategori, Quantity, Harga Satuan, dan Total Harga. Populasi penelitian adalah seluruh transaksi penjualan pada periode tersebut, dan proses clustering difokuskan pada dua variabel, yaitu Jumlah Terjual dan Harga, dari 39 menu yang tersedia.

Analisis data diawali dengan pemilihan atribut Jumlah Terjual dan Harga sebagai variabel clustering, dilanjutkan dengan pemeriksaan missing value dan data duplikat. Data transaksi yang telah bersih kemudian diagregasi dari level transaksi menjadi level menu, dengan total *Quantity* dihitung sebagai Jumlah Terjual dan rata-rata Harga Satuan sebagai Harga per menu. Karena kedua variabel memiliki rentang nilai yang jauh berbeda, normalisasi Min-Max Scaling diterapkan agar keduanya memiliki bobot setara dalam perhitungan jarak pada kedua algoritma, dengan persamaan:

$$X' = (X - X_{\min}) / (X_{\max} - X_{\min}) \dots\dots\dots (i)$$

Data yang telah dinormalisasi selanjutnya diproses menggunakan algoritma K-Means dan K-Medoids yang diimplementasikan dengan Python (scikit-learn dan scikit-learn-extra) menggunakan jarak Euclidean dan random state tetap agar hasil dapat direplikasi (Alfarizi et al., 2023). Jumlah cluster optimal ditentukan melalui Elbow Method dengan menjalankan kedua algoritma pada nilai K dari 1 hingga 10 dan mengamati nilai WCSS yang dihasilkan, sementara label tinggi, sedang, dan rendah pada setiap cluster ditentukan berdasarkan rata-rata Jumlah Terjual anggotanya.

Kualitas dan kesamaan hasil clustering dievaluasi menggunakan tiga metrik. Silhouette Score dihitung dengan persamaan:

$$s(i) = (b(i) - a(i)) / \max \{a(i), b(i)\} \dots\dots\dots (ii)$$

Dengan $a(i)$ adalah rata-rata jarak data ke- i terhadap data lain dalam cluster yang sama, dan $b(i)$ adalah rata-rata jarak data ke- i terhadap cluster terdekat lainnya (Mulyani et al., 2023).

Davies-Bouldin Index dihitung dengan persamaan:

$$DBI = (1/k) \sum \max [(S_i + S_j) / d_{ij}] \dots\dots\dots (iii)$$

dengan S_i dan S_j adalah rata-rata jarak antar anggota dalam cluster i dan j , dan d_{ij} adalah jarak antar centroid/medoid cluster i dan j (Hasan, 2024). Adjusted Rand Index digunakan untuk mengukur kesamaan keanggotaan cluster antara hasil K-Means dan K-Medoids secara langsung. Tahap akhir penelitian adalah penerjemahan hasil clustering menjadi rekomendasi strategi penjualan dan visualisasinya ke dalam *dashboard* interaktif berbasis Python untuk pengolahan data serta HTML, CSS, dan JavaScript untuk tampilan dan interaktivitas visual (Tsani et al., 2025).

4. HASIL DAN PEMBAHASAN

Karakteristik Data Menu

Statistik deskriptif dihasilkan terhadap 39 data menu hasil agregasi, sebagaimana ditampilkan pada Tabel 1.

Tabel 1. Statistik Deskriptif Data Menu.

Statistik	Jumlah Terjual	Harga (Rp)
n	39	39
Rata-rata	410,95	16.333,33
Std. Deviasi	407,77	9.540,31
Minimum	102	4.000
Median	253	15.000
Maksimum	1.978	40.000

Rata-rata Jumlah Terjual sebesar 410,95 porsi dengan standar deviasi yang hampir setara mengindikasikan variasi penjualan yang cukup tinggi antar menu, sehingga proses clustering relevan dilakukan untuk mengidentifikasi pola segmentasi menu berdasarkan performa penjualannya. Seluruh 11.339 baris transaksi dinyatakan lengkap tanpa missing value, sementara 256 baris (2,3%) teridentifikasi sebagai duplikat dan dikeluarkan dari analisis untuk menjaga validitas data sebelum diagregasi menjadi 39 baris data menu.

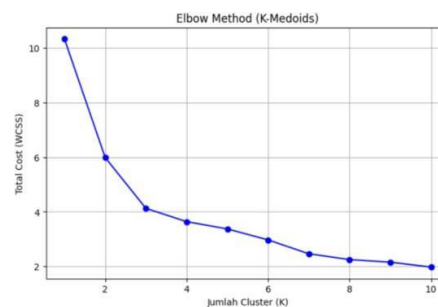
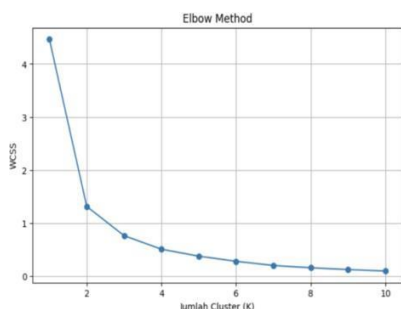
Penentuan Cluster dan Hasil Pengelompokan

Elbow Method dijalankan untuk kedua algoritma pada nilai K dari 1 hingga 10. Nilai WCSS menurun tajam dari $K=1$ hingga $K=3$, kemudian melandai signifikan mulai $K=4$, sebagaimana dirangkum pada Tabel 2. Pola ini mengonfirmasi titik elbow berada pada $K=3$

untuk kedua algoritma, sehingga jumlah cluster optimal yang digunakan adalah K=3 (Maori & Evanita, 2023), sebagaimana ditampilkan pada Gambar 1 dan Gambar 2.

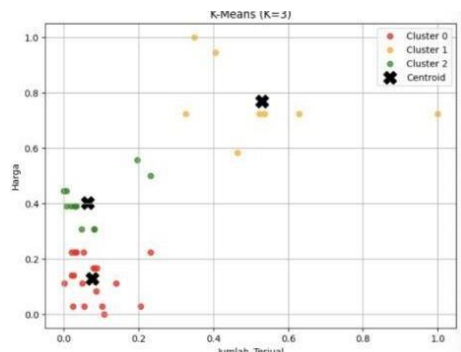
Tabel 2. Nilai WCSS dan Selisih Antar Nilai K.

K	WCSS	Selisih terhadap K sebelumnya
1	$\pm 4,5$	-
2	1,311993	3,152117
3	0,761726	0,550267
4	0,505920	0,255806
5-10	menurun landai	0,027223-0,130415

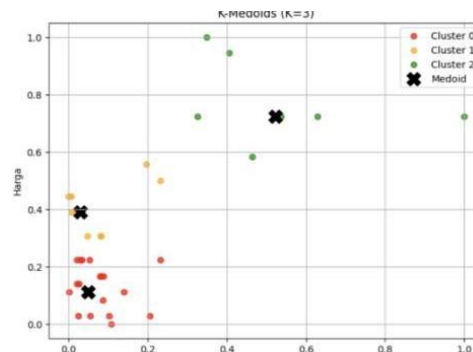


Gambar 1. Hasil Elbow Method K-Means. **Gambar 2.** Hasil Elbow Method K-Medoids.

Perlu ditegaskan bahwa penamaan kelompok (tinggi, sedang, rendah) didasarkan pada rata-rata jumlah terjual, bukan pada tingkat harga. Dengan K=3, kedua algoritma membagi 39 menu ke dalam tiga cluster dengan karakteristik yang identik: kelompok penjualan tinggi berisi 8 menu dengan harga Rp25.000-Rp40.000 (didominasi olahan bebek dan ayam, rata-rata harga Rp31.625), kelompok penjualan sedang berisi 19 menu dengan harga terjangkau Rp4.000-Rp12.000 (rata-rata harga Rp8.579), dan kelompok penjualan rendah berisi 12 menu dengan harga menengah Rp15.000-Rp24.000 (rata-rata harga Rp18.417). Total penjualan masing-masing kelompok berturut-turut adalah 8.756, 4.620, dan 2.651 porsi menariknya, kelompok dengan volume penjualan tertinggi justru berada pada rentang harga tertinggi, karena menu olahan bebek dan ayam merupakan identitas utama restoran sehingga permintaannya tetap tinggi meskipun harganya lebih mahal dibanding menu lain.



Gambar 3. Hasil Clustering K-Means (K=3).



Gambar 4. Hasil Clustering K-Medoids (K=3).

Sumbu horizontal pada Gambar 3 dan Gambar 4 merepresentasikan Jumlah Terjual dan sumbu vertikal merepresentasikan Harga, keduanya dalam skala hasil normalisasi *Min-Max* (0-1). Pola sebaran titik pada kedua gambar terlihat serupa, yang secara visual memperkuat temuan nilai Adjusted Rand Index sebesar 1,0 bahwa K-Means dan K-Medoids menghasilkan partisi cluster yang identik pada dataset ini.

Evaluasi dan Perbandingan Algoritma

Kualitas dan kesamaan hasil clustering kedua algoritma dievaluasi menggunakan Silhouette Score, Davies-Bouldin Index, dan Adjusted Rand Index, sebagaimana ditampilkan pada Tabel 3.

Tabel 3. Hasil Evaluasi dan Perbandingan K-Means dan K-Medoids.

Metrik Evaluasi	K-Means	K-Medoids
Silhouette Score	0,5368	0,5368
Davies-Bouldin Index	0,5861	0,5861
Adjusted Rand Index (ARI)	1,0	1,0

Silhouette Score sebesar 0,5368 mengindikasikan struktur cluster yang cukup baik, karena nilai yang mendekati 1 menunjukkan pemisahan antar cluster yang semakin baik (Mulyani et al., 2023). Davies-Bouldin Index sebesar 0,5861 yang mendekati 0 juga mengindikasikan kualitas clustering yang cukup baik (Hasan, 2024). Yang lebih penting, nilai ARI sebesar 1,0 membuktikan bahwa K-Means dan K-Medoids menghasilkan partisi cluster yang benar-benar identik pada dataset ini setiap menu dialokasikan ke kelompok yang merepresentasikan karakteristik penjualan yang sama pada kedua algoritma. Karena kedua algoritma menghasilkan kualitas clustering yang identik, pertimbangan pemilihan algoritma dapat dilihat dari sisi efisiensi komputasi: K-Means cenderung lebih ringan karena pembaruan pusat cluster hanya menghitung rata-rata anggota cluster, sedangkan K-Medoids memerlukan proses pertukaran (*swap*) kandidat medoid yang lebih kompleks dan mahal secara komputasi, khususnya pada dataset yang lebih besar (Marlina et al., 2018).

Pembahasan

Temuan bahwa K-Means dan K-Medoids menghasilkan partisi cluster yang identik pada penelitian ini menunjukkan pola yang tidak selalu konsisten dengan penelitian terdahulu. Farissa et al. (2021) yang membandingkan kedua algoritma pada data obat di Puskesmas Karangsambung melaporkan kualitas cluster K-Means hanya masuk kategori medium structure berdasarkan nilai Silhouette, sementara Fathurrahman et al. (2023) pada data sebaran Covid-19 turut menemukan kemiripan hasil yang sejalan dengan temuan penelitian ini. Kesamaan hasil pada studi kasus penjualan menu Bebek Cak Selem Gresik diduga terkait dengan penggunaan dua variabel numerik (Jumlah Terjual dan Harga) yang memiliki pola distribusi

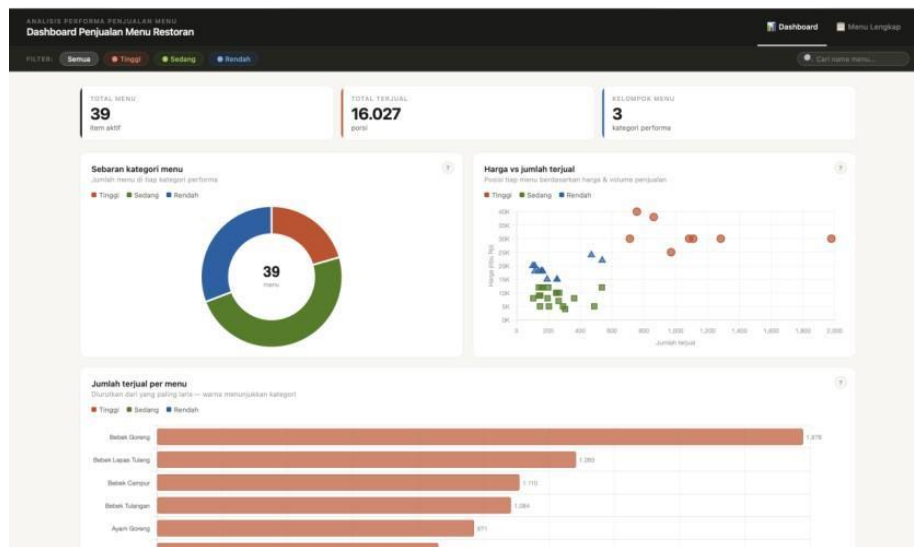
relatif terpisah secara alami, sehingga ruang pencarian solusi clustering menjadi lebih sederhana dibandingkan penelitian dengan variabel yang lebih banyak atau heterogen, seperti pada Anggraeni & Tresnawati (2020).

Dibandingkan dengan Triyandana et al. (2022) yang menerapkan K-Means pada klasterisasi menu DPom *Coffee* dan memperoleh Davies-Bouldin Index sebesar 0,457, nilai DBI pada penelitian ini (0,5861) sedikit lebih tinggi, mengindikasikan pemisahan antar cluster yang sedikit kurang tegas. Perbedaan ini kemungkinan dipengaruhi rentang harga menu Bebek Cak Selem Gresik yang jauh lebih lebar (Rp4.000-Rp40.000) dibandingkan data kafe pada penelitian acuan, sehingga variasi harga turut memengaruhi kepadatan (*density*) masing-masing cluster. Pendekatan evaluasi kombinasi Silhouette Index dan Davies-Bouldin Index yang digunakan pada penelitian ini juga sejalan dengan Yolandari et al. (2025). Meskipun demikian, ketiga segmen menu yang terbentuk tetap menunjukkan karakteristik bisnis yang jelas dan dapat ditindaklanjuti, sejalan dengan penekanan Ana et al. (2024) dan Gantara dan Ali (2023) bahwa hasil clustering perlu bersifat aplikatif bagi pengambilan keputusan operasional, bukan semata-mata optimal secara statistik.

Secara keseluruhan, kontribusi utama penelitian ini terletak pada penerapan sekaligus perbandingan langsung K-Means dan K-Medoids pada studi kasus penjualan menu restoran skala UMKM, domain yang belum banyak dieksplorasi pada penelitian acuan yang lebih berfokus pada sektor kesehatan, pendidikan, dan ritel. Integrasi hasil clustering ke dalam *dashboard* interaktif turut menjawab kebutuhan praktis pemilik restoran yang umumnya tidak memiliki latar belakang statistik, sehingga hasil analisis data mining dapat langsung ditindaklanjuti dalam pengambilan keputusan operasional sehari-hari.

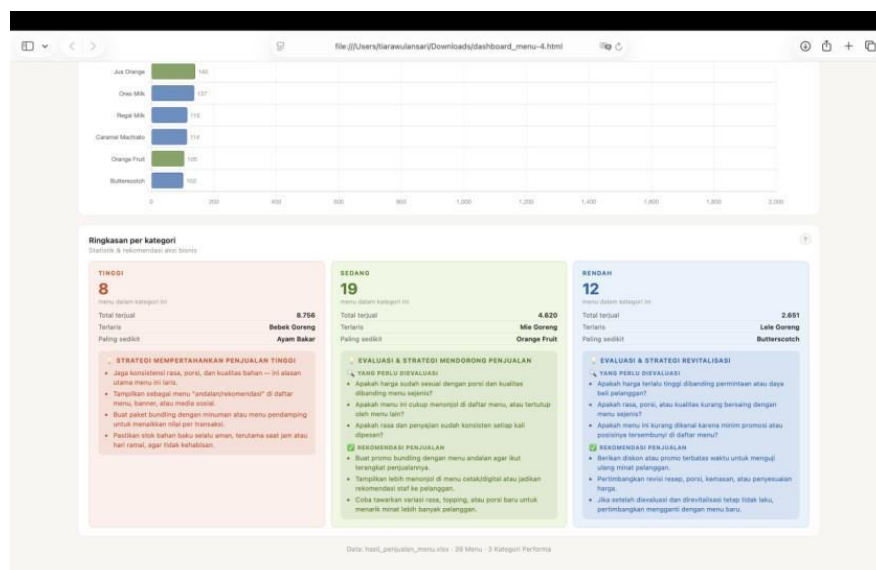
Implikasi Praktis: *Dashboard* Interaktif bagi Pemilik Restoran

Hasil clustering dan rekomendasi strategi penjualan divisualisasikan melalui *dashboard* interaktif berbasis HTML yang dirancang untuk menjembatani kesenjangan antara hasil analisis statistik dan kebutuhan pengambilan keputusan sehari-hari pemilik restoran; *dashboard* ini memuat filter kategori, ringkasan indikator, grafik distribusi dan sebaran cluster, ranking menu, serta tabel data yang saling melengkapi dalam satu tampilan. *Dashboard* ini menerjemahkan istilah teknis clustering menjadi kategori yang mudah dipahami pemilik restoran, sehingga hasil penelitian dapat langsung ditindaklanjuti dalam pengambilan keputusan bisnis sehari-hari tanpa memerlukan pemahaman teknis mengenai algoritma yang digunakan (Akbar & Octaviany, 2021), sebagaimana ditampilkan pada Gambar 5.



Gambar 5. *Dashboard* Interaktif Hasil Clustering Penjualan Menu.

Berdasarkan hasil evaluasi tersebut, dirumuskan rekomendasi strategi penjualan yang berbeda untuk masing-masing kelompok menu. Kelompok penjualan tinggi (8 menu) menyumbang penjualan tertinggi, sehingga strategi yang disarankan adalah mempertahankan kualitas dan konsistensi rasa, optimalisasi promosi di media sosial dan *platform* pesan-antar, serta program loyalitas pelanggan. Kelompok penjualan sedang (19 menu) menunjukkan performa stabil dengan harga terjangkau, sehingga disarankan program *bundling* atau paket hemat untuk meningkatkan nilai transaksi per pelanggan. Kelompok penjualan rendah (12 menu) memerlukan peningkatan visibilitas melalui *promosi happy hour*, rekomendasi aktif oleh pelayan, dan pemanfaatan media sosial, disertai evaluasi harga dan kualitas produk, sebagaimana dirangkum pada Gambar 6.



Gambar 6. Ringkasan Rekomendasi Strategi Penjualan per Kategori pada *Dashboard* Interaktif.

5. KESIMPULAN DAN SARAN

Penerapan algoritma K-Means dan K-Medoids dengan kerangka kerja CRISP-DM berhasil mengelompokkan 39 menu Restoran Bebek Cak Seleem Gresik ke dalam tiga cluster optimal ($K=3$) berdasarkan variabel Jumlah Terjual dan Harga, yaitu kelompok penjualan tinggi (8 menu, Rp25.000-Rp40.000), kelompok penjualan sedang (19 menu, Rp4.000-Rp12.000), dan kelompok penjualan rendah (12 menu, Rp15.000-Rp24.000). Kedua algoritma menghasilkan kualitas clustering yang identik, dengan Silhouette Score 0,5368, Davies-Bouldin Index 0,5861, dan Adjusted Rand Index 1,0, yang membuktikan bahwa keduanya menghasilkan partisi cluster yang sama persis pada data ini. Berdasarkan hasil perbandingan performa tersebut, algoritma K-Means direkomendasikan sebagai algoritma yang lebih praktis untuk diterapkan pada data penjualan menu Bebek Cak Seleem Gresik, bukan karena perbedaan hasil pengelompokan, melainkan karena pertimbangan teoretis mengenai efisiensi komputasi K-Means yang tidak memerlukan proses evaluasi pertukaran medoid.

Hasil pengelompokan berhasil diterjemahkan menjadi rekomendasi strategi penjualan yang berbeda untuk setiap kelompok menu dan divisualisasikan dalam *dashboard* interaktif berbasis HTML yang dapat langsung digunakan oleh pemilik restoran. Penelitian ini memiliki keterbatasan, yaitu periode data yang hanya mencakup lima bulan (Januari-Mei 2025) sehingga belum menangkap pola musiman sepanjang tahun, serta pengujian waktu komputasi (runtime) kedua algoritma yang tidak dilakukan secara langsung, sehingga klaim efisiensi komputasi K-Means masih bersifat teoretis. Penelitian selanjutnya disarankan untuk memperluas periode data menjadi minimal satu tahun penuh, menambahkan variabel lain seperti frekuensi pemesanan dan margin keuntungan, mengukur dan membandingkan waktu komputasi kedua algoritma secara langsung, serta mengintegrasikan *dashboard* dengan sistem *Point of Sale* (POS) restoran secara *real-time*.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pemilik dan pengelola Restoran Bebek Cak Seleem Gresik atas izin akses data transaksi penjualan yang digunakan dalam penelitian ini, serta kepada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur atas dukungan yang diberikan. Artikel ini merupakan bagian dari skripsi penulis pertama.

DAFTAR REFERENSI

- Akbar, R., & Octaviany, M. (2021). Perancangan visualisasi dashboard dan *clustering* dengan menerapkan *business intelligence* pada Dinas DPMPTSP Kabupaten Dharmasraya. *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, 7(3), 340-350. <https://doi.org/10.26418/jp.v7i3.49719>
- Alfarizi, M. R. S., Al-Farish, M. Z., Taufiqurrahman, M., Ardiansah, G., & Elgar, M. (2023). Penggunaan Python sebagai bahasa pemrograman untuk *machine learning* dan *deep learning*. *KARIMAH TAUHID: Karya Ilmiah Mahasiswa Bertauhid*, 2(1), 1-6.
- Ana, E. S., Anwar, S., & Fuad, M. H. (2024). Analisa tingkat penjualan makanan dan minuman dengan klasterisasi menggunakan algoritma K-Means. *Jurnal INSAN: Journal of Information System Management and Innovation*, 4(1), 45-54. <https://doi.org/10.31294/jinsan.v4i1.3763>
- Anggraeni, N. L., & Tresnawati, S. (2020). Komparasi algoritma K-Means dan K-Medoids untuk menangani strategi promosi di Politeknik TEDC Bandung. *Jurnal TEDC*, 14(2), 337-342.
- Ansori, Y., & Wulandari, C. (2022). CRISP-DM method on Indonesian micro industries (UMKM) using K-Means clustering algorithm. *MATICS: Jurnal Ilmu Komputer dan Teknologi Informasi*, 14(2), 35-40. <https://doi.org/10.18860/mat.v14i2.13760>
- Badan Pusat Statistik. (2023). *Statistik industri mikro dan kecil 2023*. <https://www.bps.go.id/id/publication/2023/11/30/statistik-industri-mikro-dan-kecil-2023.html>
- Farissa, R. A., Mayasari, R., & Umaidah, Y. (2021). Perbandingan algoritma K-Means dan K-Medoids untuk mengelompokkan data obat dengan *Silhouette Coefficient* di Puskesmas Karangsambung. *Journal of Applied Informatics and Computing*, 5(2), 119-125. <https://doi.org/10.30871/jaic.v5i2.3237>
- Fathurrahman, F., Harini, S., & Kusumawati, R. (2023). Evaluasi *clustering* K-Means dan K-Medoids pada persebaran COVID-19 di Indonesia dengan metode *Davies-Bouldin Index* (DBI). *Jurnal Mnemonic*, 6(2), 117-128. <https://doi.org/10.36040/mnemonic.v6i2.6642>
- Gantara, N. P., & Ali, I. (2023). Penerapan metode K-Means clustering pada penjualan barang di Sports Station. *E-Link: Jurnal Teknik Elektro dan Informatika*, 18(1), 28-39. <https://doi.org/10.30587/e-link.v18i1.5339>
- Hasan, Y. (2024). Pengukuran *Silhouette Score* dan *Davies-Bouldin Index* pada hasil *cluster* K-Means dan DBSCAN. *Jurnal Informatika dan Teknik Elektro Terapan*, 12(3S1). <https://doi.org/10.23960/jitet.v12i3S1.5001>
- Iswavigra, D. U., Zen, L. E., Okfalisa, & Hanim, H. (2023). *Marketing strategy* UMKM dengan CRISP-DM *clustering* dan *promotion mix* menggunakan metode K-Medoids. *Jurnal Informasi dan Teknologi*, 5(1), 45-54.
- Kamila, I., Khairunnisa, U., & Mustakim, M. (2019). Perbandingan algoritma K-Means dan K-Medoids untuk pengelompokan data transaksi bongkar muat di Provinsi Riau. *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, 5(1), 119-125. <https://doi.org/10.24014/rmsi.v5i1.7381>

- Maori, N. A., & Evanita, E. (2023). Metode *elbow* dalam optimasi jumlah *cluster* pada K-Means clustering. *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, 14(2), 277-288. <https://doi.org/10.24176/simet.v14i2.9630>
- Marlina, D., Lina, N., Fernando, A., & Ramadhan, A. (2018). Implementasi algoritma K-Medoids dan K-Means untuk pengelompokan wilayah sebaran cacat pada anak. *Jurnal CoreIT*, 4(2), 64-70. <https://doi.org/10.24014/coreit.v4i2.4498>
- Mulyani, H., Setiawan, R. A., & Fathi, H. (2023). Optimization of K value in clustering using *Silhouette Score* (Case study: Mall customers data). *Journal of Information Technology and Its Utilization*, 6(2), 45-50. <https://doi.org/10.56873/jitu.6.2.5243>
- Sibarani, B. E., Setiawan, S., Hadi, T., Williams, T., & Mkhize, T. (2024). Penggunaan data analitik dalam strategi pemasaran untuk mempertahankan loyalitas pelanggan. *Jurnal MENTARI: Manajemen, Pendidikan dan Teknologi Informasi*, 3(1), 30-39. <https://doi.org/10.33050/mentari.v3i1.592>
- Singgalen, Y. A. (2023). Penerapan metode CRISP-DM dalam klasifikasi data ulasan pengunjung destinasi Danau Toba menggunakan algoritma *Naïve Bayes Classifier* (NBC) dan *Decision Tree* (DT). *JURIKOM (Jurnal Riset Komputer)*, 10(3). <https://doi.org/10.30865/mib.v7i3.6461>
- Triyandana, G., Putri, L. A., & Umaidah, Y. (2022). Penerapan *data mining* pengelompokan menu makanan dan minuman berdasarkan tingkat penjualan menggunakan metode K-Means. *Journal of Applied Informatics and Computing*, 6(1), 40-46. <https://doi.org/10.30871/jaic.v6i1.3824>
- Tsani, A. J. B., Mahardika, F., & Santika, D. (2025). Pembuatan *web dashboard* interaktif untuk analisis data penjualan *vending machine*. *Modem: Jurnal Informatika dan Sains Teknologi*, 3(2), 51-61. <https://doi.org/10.62951/modem.v3i2.406>
- Wakhidah, N. (2010). *Clustering* menggunakan algoritma K-Means. *Jurnal Transformatika*, 8(1), 33-36. <https://doi.org/10.26623/transformatika.v8i1.45>
- Wisudawati, N., Rahmawati, D., & Amelia, S. (2024). Strategi pengambilan keputusan berbasis *data analytics* pada UMKM untuk meningkatkan penjualan. *Data Sciences Indonesia (DSI)*, 4(2), 88-97.
- Yolandari, N. A., Butarbutar, L. E., Rajagukguk, G. C. H., Zulfi, M. F., Arnita, & Ramadhani, F. (2025). Analisis perbandingan K-Means dan DBSCAN dalam pengelompokan *travel review ratings* menggunakan evaluasi *Silhouette Index* dan *Davies-Bouldin Index*. *Jurnal Informatika dan Teknik Elektro Terapan*, 13(3), 470-481. <https://doi.org/10.23960/jitet.v13i3.6884>