SPEKTRUM AKADEMIKA

Volume. 1, Nomor. 2, 2026

KLASTERISASI KOMODITAS PANGAN BERDASARKAN STRUKTUR PENYEDIAAN DAN KETERGANTUNGAN IMPOR PADA NERACA BAHAN MAKANAN PROVINSI JAWA TIMUR TAHUN 2025 MENGGUNAKAN K-MEANS

Adinda Putri Rachmawati[1], Selvy Dwi Yulita Sari[2], Prismahardi Aji Riyantoko[3], Alfan Rizaldy[4]

[1,2,3,4] **Program Studi Sains Data, Universitas Pembangunan Nasional "Veteran" Jawa Timur**

24083010005@student.upnjatim.ac.id, 24083010095@student.upnjatim.ac.id

Abstrak

Ketersediaan pangan merupakan aspek penting dalam ketahanan pangan daerah yang dipengaruhi oleh struktur penyediaan dan penggunaan pangan. Penelitian ini bertujuan mengelompokkan komoditas pangan berdasarkan struktur penyediaan dan ketergantungan impor menggunakan data Neraca Bahan Makanan (NBM) Provinsi Jawa Timur Tahun 2025. Metode yang digunakan adalah K-Means dengan lima indikator, yaitu rasio impor, rasio ekspor, rasio penggunaan untuk pakan, rasio penggunaan untuk pengolahan, dan ketersediaan pangan per kapita. Tahapan penelitian meliputi pembersihan data, pembentukan variabel, standardisasi, klasterisasi, dan evaluasi model. Sebanyak 134 komoditas menghasilkan tiga klaster. Klaster 1 memiliki rasio impor dan ekspor tertinggi, Klaster 2 memiliki rasio pengolahan dan ketersediaan per kapita tertinggi, sedangkan Klaster 3 memiliki anggota terbanyak dengan rasio impor relatif rendah. Nilai Silhouette Score sebesar 0,3735 menunjukkan pemisahan antarklaster belum optimal. Hasil penelitian memberikan gambaran karakteristik komoditas pangan sebagai informasi pendukung evaluasi ketersediaan pangan daerah. Penelitian selanjutnya disarankan menggunakan data beberapa tahun dan membandingkan metode klasterisasi lain.

Kata kunci: *k-means; ketergantungan impor; ketersediaan pangan; klasterisasi; neraca bahan makanan*

Abstract

Food availability is an important aspect of regional food security influenced by supply structure and food utilization. This study aims to classify food commodities based on supply structure and import dependency using the 2025 Food Balance Sheet of East Java Province. The K-Means algorithm was applied using five indicators: import dependency ratio, export ratio, feed utilization ratio, processing utilization ratio, and food availability per capita. The analysis included data cleaning, feature formation, standardization, cluster determination, clustering, and model evaluation. A total of 134 food commodities were grouped into three clusters. Cluster 1 had the highest average import and export ratios, Cluster 2 had the highest processing utilization ratio and per capita availability, while Cluster 3 contained the largest number of commodities with a relatively low import ratio. The Silhouette Score of 0.3735 indicated that cluster separation was not yet optimal. Nevertheless, the results provide information on food commodity characteristics that can support regional food availability evaluation. Future research may use multi-year data and compare other clustering methods to improve clustering performance.

Keywords: *clustering; food availability; food balance sheet; import dependency; k-means*

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

PENDAHULUAN

Ketahanan pangan menjadi aspek yang krusial dalam pembangunan, sebab menyangkut kemampuan suatu wilayah untuk mencukupi kebutuhan pangan masyarakat. Ketersediaan pangan dipengaruhi oleh produksi dalam negeri, impor, ekspor, perubahan stok, dan penggunaan pangan. Wahida et al. (2025) serta Hakim et al. (2026) menekankan pentingnya kemandirian dan ketergantungan impor dalam menilai kondisi ketersediaan pangan.

Neraca Bahan Makanan (NBM) merupakan salah satu instrumen untuk menggambarkan kondisi penyediaan dan penggunaan pangan pada suatu wilayah dalam periode tertentu. NBM menyajikan informasi mengenai produksi, perubahan stok, impor, ekspor, serta pemanfaatan pangan untuk berbagai kebutuhan. Melalui informasi tersebut, kondisi ketersediaan pangan dapat ditinjau berdasarkan jumlah dan struktur penyediaannya. Penyediaan pangan domestik menggambarkan jumlah pangan yang tersedia di suatu wilayah setelah memperhitungkan komponen produksi, perdagangan, dan perubahan stok. Penelitian Atasa dan Nugroho (2021), Atasa et al. (2022), serta Renaldi dan Pardian (2024) telah memanfaatkan NBM untuk menganalisis ketersediaan pangan serta melihat keterkaitannya dengan kebutuhan pangan wilayah.

Ketergantungan impor merupakan salah satu aspek yang dapat digunakan untuk melihat sejauh mana ketersediaan suatu komoditas pangan didukung oleh pasokan dari luar wilayah. Semakin besar kontribusi impor terhadap penyediaan domestik, semakin besar pula peran pasokan dari luar wilayah dalam memenuhi ketersediaan komoditas tersebut. Selain rasio impor, struktur penyediaan dan penggunaan pangan dapat digambarkan melalui beberapa indikator lain, yaitu rasio ekspor, rasio penggunaan untuk pakan, dan rasio penggunaan untuk bahan baku industri atau pengolahan. Penelitian mengenai ketergantungan pangan menunjukkan bahwa kontribusi impor dapat menjadi salah satu indikator dalam menilai kerentanan ketersediaan pangan. Penggabungan beberapa indikator penyediaan dan penggunaan pangan dapat memberikan gambaran karakteristik komoditas yang lebih beragam dibandingkan hanya menggunakan jumlah produksi atau ketersediaan.

Untuk mengelompokkan objek berdasarkan kemiripan karakteristiknya, teknik *data mining* seperti *clustering* banyak digunakan. Salah satu algoritma yang umum diterapkan adalah K-Means, yaitu metode pengelompokan yang membagi data ke dalam sejumlah klaster berdasarkan jarak antara data dan pusat klaster (*centroid*) menggunakan *Euclidean Distance*. Sebelum proses klasterisasi, data perlu melalui standardisasi (misalnya menggunakan *StandardScaler*) agar perbedaan skala antarvariabel tidak mendominasi perhitungan jarak. Kualitas hasil pengelompokan selanjutnya dapat dievaluasi menggunakan metrik seperti *Silhouette Score*, *Davies-Bouldin Index* (DBI), dan *Calinski-Harabasz Index* (CHI).

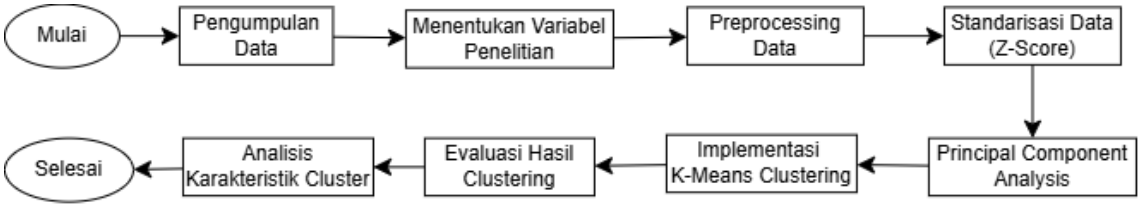
Beberapa penelitian terdahulu seperti Arpipi (2020), Dahwanu et al. (2025), serta Sumargo dan Handhayani (2025) telah menerapkan algoritma K-Means dalam analisis keragaman konsumsi dan harga pangan berdasarkan karakteristik wilayah. Namun, pengelompokan komoditas berdasarkan struktur penyediaan dan pemanfaatan menggunakan data NBM masih perlu dikembangkan. Oleh karena itu, penelitian ini bertujuan mengelompokkan komoditas pangan berdasarkan data NBM Provinsi Jawa Timur Tahun 2025 menggunakan algoritma K-Means dengan variabel rasio ketergantungan impor, rasio ekspor,

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

rasio penggunaan untuk pakan, rasio pengolahan, dan ketersediaan pangan per kapita. Hasil penelitian diharapkan dapat menggambarkan karakteristik komoditas sebagai bahan pendukung evaluasi ketersediaan pangan dan kebijakan ketahanan pangan daerah.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *data mining* melalui algoritma *K-Means Clustering*. Penelitian bertujuan mengelompokkan komoditas pangan berdasarkan struktur penyediaan dan ketergantungan impor pada Neraca Bahan Makanan (NBM) Provinsi Jawa Timur Tahun 2025. Tahapan penelitian meliputi pengumpulan data, *preprocessing*, pembentukan variabel, standardisasi, proses *clustering*, evaluasi, dan interpretasi hasil sebagaimana ditunjukkan pada Gambar 1.



Gambar 1. Diagram Alir Penelitian

Pengumpulan Data

Data yang digunakan merupakan data sekunder yang bersumber dari dokumen hasil kajian Neraca Bahan Makanan Provinsi Jawa Timur Tahun 2025. Data diperoleh dalam format Microsoft Excel pada lembar kerja Tabel NBM dengan objek penelitian berupa komoditas pangan. Setiap komoditas diperlakukan sebagai satu objek pengamatan yang akan dikelompokkan berdasarkan karakteristik penyediaan dan penggunaannya.

Variabel Penelitian

Variabel penelitian dibentuk dari komponen penyediaan dan penggunaan pangan dalam NBM. Seluruh variabel penelitian merupakan data numerik berskala rasio yang digunakan sebagai dasar pengelompokan komoditas pangan menggunakan algoritma K-Means sebagaimana disajikan pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Keterangan
X_1	Rasio ketergantungan impor
X_2	Rasio ekspor
X_3	Rasio penggunaan untuk pakan
X_4	Rasio penggunaan untuk pengolahan
X_5	Ketersediaan pangan per kapita

Preprocessing Data

Tahap *preprocessing* dilakukan untuk memastikan data siap digunakan dalam proses pengelompokan. Tahapan ini meliputi penghapusan baris yang bukan merupakan data komoditas, pemeriksaan nilai kosong,

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

serta seleksi data dengan nilai total penyediaan dan bahan makanan lebih dari nol. Nilai nol pada komponen produksi, impor, dan ekspor tetap dipertahankan karena merupakan bagian dari karakteristik komoditas. Selanjutnya, dilakukan pembentukan lima variabel penelitian dan standardisasi menggunakan metode Z-Score melalui persamaan:

$$Z_{ij} = (X_{ij} - \mu_j) \,/\, \sigma_j \quad (1)$$

Keterangan: Z_{ij} = nilai hasil standardisasi variabel ke-j pada objek ke-i; X_{ij} = nilai awal variabel ke-j pada objek ke-i; μ_j = rata-rata variabel ke-j; dan σ_j = standar deviasi variabel ke-j.

Principal Component Analysis (PCA)

Principal Component Analysis (PCA) digunakan untuk mereduksi lima variabel penelitian menjadi dua komponen utama, yaitu PC1 dan PC2, sehingga pola persebaran data dapat divisualisasikan dalam bentuk dua dimensi. Proses PCA menggunakan data yang telah distandardisasi dengan persamaan:

$$PC_{2D} = ZV_2 \quad (2)$$

Keterangan: PC_{2D} = hasil transformasi menjadi dua komponen utama; Z = matriks data hasil standardisasi; dan V_2 = dua vektor eigen terpilih.

K-Means Clustering

Pengelompokan komoditas pangan dilakukan menggunakan algoritma K-Means berdasarkan kedekatan karakteristik variabel menggunakan jarak *Euclidean*. Perhitungan jarak antara objek dan pusat klaster dinyatakan dengan persamaan:

$$d(x_i, c_j) = \sqrt{[\,\textstyle\sum_{k=1..p} (x_{ik} - c_{jk})^2\,]} \quad (3)$$

Keterangan: $d(x_i, c_j)$ = jarak objek terhadap *centroid*; x_{ik} = nilai variabel ke-k pada objek ke-i; c_{jk} = nilai variabel ke-k pada *centroid* klaster ke-j; dan p = jumlah variabel.

Evaluasi dan Interpretasi Hasil

Evaluasi hasil *clustering* dilakukan untuk menentukan jumlah klaster terbaik dan mengukur kualitas pengelompokan. Pengujian dilakukan pada jumlah klaster $k = 2$ hingga $k = 8$ menggunakan empat metode, yaitu *Elbow Method*, *Silhouette Score*, *Davies-Bouldin Index* (DBI), dan *Calinski-Harabasz Index* (CHI).

Elbow Method digunakan untuk menentukan jumlah klaster berdasarkan nilai *Within-Cluster Sum of Squares* (WCSS), dengan persamaan:

$$WCSS = \textstyle\sum_{j=1..K} \sum_{xi \in Cj} \|x_i - \mu_j\|^2 \quad (4)$$

Keterangan: K = jumlah klaster; C_j = anggota klaster ke-j; x_i = objek data ke-i; dan μ_j = *centroid* klaster ke-j.

Silhouette Score digunakan untuk mengukur kesesuaian objek terhadap klasternya berdasarkan jarak dalam klaster dan klaster terdekat lainnya, dengan persamaan:

$$S(i) = (b(i) - a(i)) \,/\, \max(a(i), b(i)) \quad (5)$$

Keterangan: $a(i)$ = rata-rata jarak objek ke anggota klaster yang sama; $b(i)$ = rata-rata jarak objek ke klaster terdekat lainnya; dan $S(i)$ = nilai *Silhouette* objek ke-i.

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Davies-Bouldin Index (DBI) digunakan untuk mengukur kekompakan dan pemisahan antarklaster, dengan persamaan:

$$DBI = (1/K) \sum_{i=1..K} \max_{j \neq i} ((S_i + S_j) / M_{ij}) \quad (6)$$

Keterangan: K = jumlah klaster; S_i dan S_j = rata-rata jarak anggota terhadap *centroid* masing-masing klaster; serta M_{ij} = jarak antara *centroid* klaster ke-*i* dan ke-*j*.

Calinski-Harabasz Index (CHI) digunakan untuk mengukur perbandingan variasi antarklaster dan variasi dalam klaster, dengan persamaan:

$$CHI = [B_K / (K - 1)] / [W_K / (n - K)] \quad (7)$$

Keterangan: B_K = jumlah kuadrat jarak antarklaster; W_K = jumlah kuadrat jarak dalam klaster; K = jumlah klaster; dan n = jumlah seluruh objek data.

HASIL DAN PEMBAHASAN

Gambaran Umum dan Hasil Pengolahan Data

Data yang digunakan dalam penelitian ini bersumber dari Neraca Bahan Makanan (NBM) Provinsi Jawa Timur Tahun 2025, khususnya pada lembar Tabel NBM. Berdasarkan hasil pembacaan data, diperoleh 150 baris komoditas pangan. Selanjutnya, dilakukan proses pembersihan data dengan mengeluarkan komoditas yang memiliki nilai penyediaan domestik kurang dari atau sama dengan nol serta komoditas dengan nilai bahan makanan kurang dari atau sama dengan nol. Hasilnya, sebanyak 16 komoditas dikeluarkan sehingga diperoleh 134 komoditas yang digunakan dalam proses analisis sebagaimana disajikan pada Tabel 2.

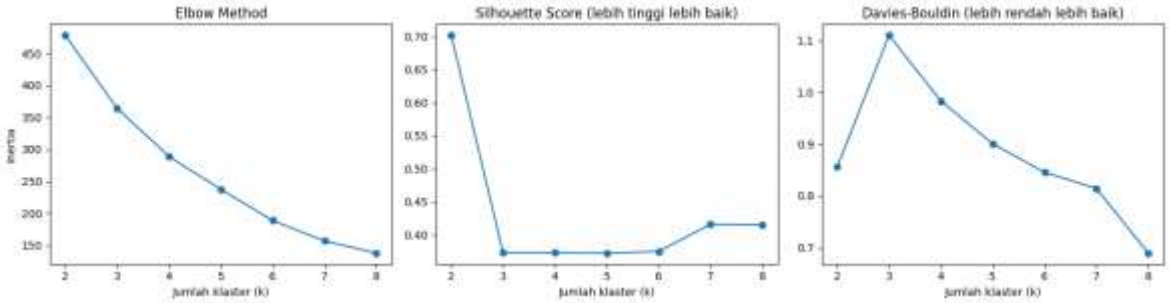
Tabel 2. Hasil Pembersihan Data

Keterangan	Jumlah
Jumlah data awal	150
Data dengan penyediaan ≤ 0	15
Data dengan bahan makanan ≤ 0	1
Jumlah data yang dikeluarkan	16
Jumlah data akhir	134

Penentuan Jumlah Klaster

Penentuan jumlah klaster dilakukan dengan membandingkan nilai *inertia* melalui *Elbow Method*, *Silhouette Score*, dan *Davies-Bouldin Index* pada jumlah klaster (*k*) antara 2 hingga 8. Hasil evaluasi tersebut disajikan pada Gambar 2.

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)



Gambar 2. Hasil Evaluasi Penentuan Jumlah Klaster

Berdasarkan Gambar 2, nilai *Silhouette Score* tertinggi diperoleh pada $k = 2$ sebesar 0,7016. Namun, jumlah anggota pada kedua klaster yang terbentuk sangat tidak seimbang, yaitu 129 dan 5 komoditas. Kondisi tersebut menunjukkan bahwa nilai *Silhouette Score* yang tinggi tidak selalu menghasilkan pembagian kelompok yang proporsional. Oleh karena itu, evaluasi dilanjutkan dengan mempertimbangkan ukuran anggota klaster dan metrik lainnya.

Pada kandidat $k = 3$ dan $k = 4$, nilai *Silhouette Score* yang diperoleh relatif berdekatan, masing-masing sebesar 0,3735 dan 0,3734. Keduanya juga memenuhi batas minimum lima anggota pada setiap klaster. Berdasarkan rekomendasi hasil evaluasi dan pertimbangan keterinterpretasian hasil, penelitian ini menggunakan tiga klaster sebagai jumlah pengelompokan. Pemilihan tersebut merupakan keputusan berdasarkan pertimbangan beberapa indikator, bukan karena $k = 3$ menghasilkan nilai terbaik pada seluruh metrik.

Hasil Klasterisasi Komoditas Pangan

Proses klasterisasi dilakukan menggunakan algoritma K-Means dengan inisialisasi *k-means*++ dan jumlah klaster sebanyak tiga. Berdasarkan hasil pengelompokan terhadap 134 komoditas, diperoleh distribusi anggota pada masing-masing klaster sebagaimana disajikan pada Tabel 3.

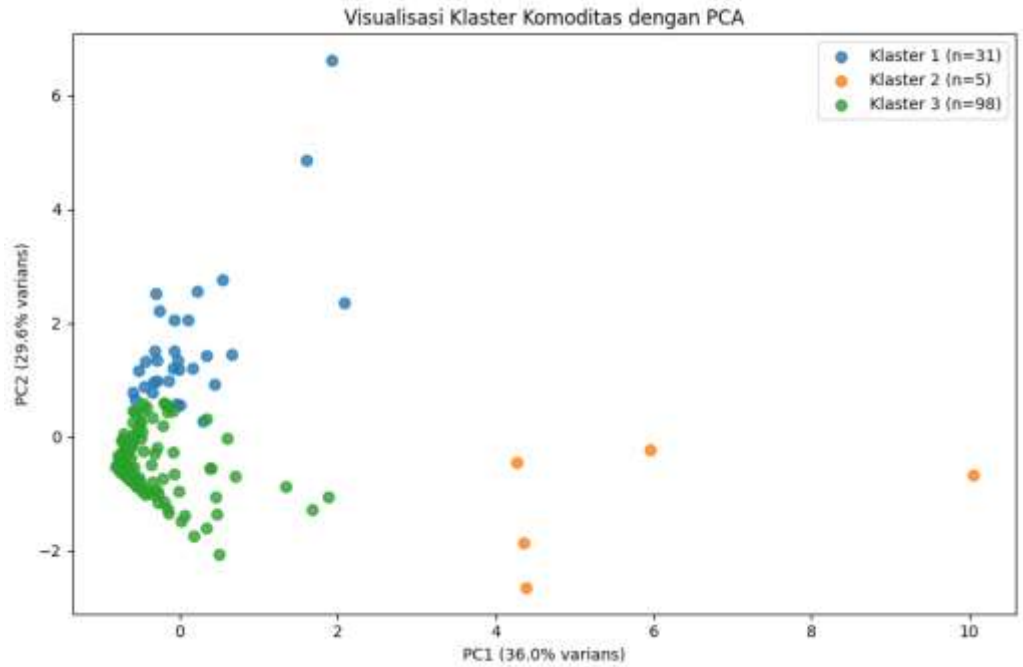
Tabel 3. Jumlah Komoditas pada Setiap Klaster

Klaster	Jumlah Komoditas	Persentase
1	31	23,13%
2	5	3,73%
3	98	73,13%
Total	**134**	**100%**

Hasil tersebut menunjukkan bahwa Klaster 3 memiliki jumlah anggota paling banyak, yaitu 98 komoditas atau 73,13% dari keseluruhan data. Sementara itu, Klaster 1 terdiri atas 31 komoditas dan Klaster 2 terdiri atas lima komoditas. Perbedaan jumlah anggota tersebut menunjukkan adanya variasi karakteristik penyediaan dan penggunaan pangan di antara komoditas yang dianalisis.

Untuk melihat persebaran hasil pengelompokan, dilakukan reduksi dimensi menggunakan *Principal Component Analysis* (PCA) ke dalam dua dimensi, sedangkan proses pembentukan klaster tetap menggunakan lima fitur penelitian. Hasil visualisasi ditampilkan pada Gambar 3.

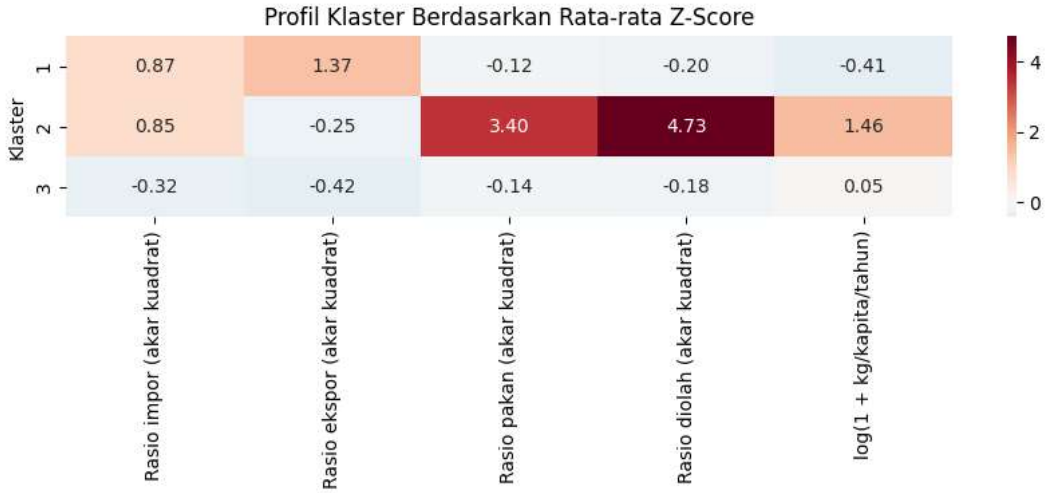
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)



Gambar 3. Visualisasi Hasil Klasterisasi Komoditas Pangan Menggunakan PCA

Persebaran komoditas menunjukkan terbentuknya tiga kelompok dengan pola penyebaran yang berbeda. Dua komponen utama pada visualisasi tersebut mampu menjelaskan 65,63% variasi data. Dengan demikian, visualisasi PCA memberikan gambaran mengenai pola pengelompokan, meskipun tidak seluruh variasi data dapat ditampilkan dalam dua dimensi.

Selanjutnya, karakteristik masing-masing klaster dianalisis berdasarkan nilai rata-rata fitur yang telah distandardisasi dan disajikan dalam bentuk *heatmap* pada Gambar 4.



Gambar 4. Profil Klaster Berdasarkan Rata-Rata Z-Score

Terlihat perbedaan karakteristik relatif setiap klaster terhadap rata-rata keseluruhan fitur. Nilai positif menunjukkan bahwa rata-rata fitur pada klaster tersebut berada di atas rata-rata data, sedangkan nilai negatif menunjukkan posisi di bawah rata-rata data. Untuk memperjelas karakteristik setiap kelompok, dilakukan perhitungan rata-rata fitur menggunakan skala data asli sebagaimana disajikan pada Tabel 4.

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Tabel 4. Rata-Rata Indikator pada Setiap Klaster (Skala Asli)

Indikator	Klaster 1	Klaster 2	Klaster 3
Jumlah komoditas	31	5	98
Rasio impor	1,1545	0,7801	0,0816
Rasio ekspor	0,3033	0,0334	0,0219
Rasio penggunaan pakan	0,0006	0,0534	0,0004
Rasio penggunaan pengolahan	0,0000	0,4221	0,0004
Ketersediaan (kg/kapita/tahun)	1,0994	31,1977	5,0219

Klaster 1 memiliki rata-rata rasio impor sebesar 1,1545 dan rasio ekspor sebesar 0,3033. Kedua nilai tersebut merupakan yang tertinggi dibandingkan klaster lainnya. Karakteristik ini menunjukkan bahwa komoditas dalam Klaster 1 memiliki kontribusi perdagangan luar wilayah yang relatif lebih menonjol dalam struktur penyediaannya. Selain itu, rata-rata ketersediaan per kapita pada kelompok ini merupakan yang terendah, yaitu 1,0994 kg/kapita/tahun.

Klaster 2 memiliki karakteristik yang berbeda dengan rata-rata rasio penggunaan untuk pengolahan sebesar 0,4221 dan ketersediaan per kapita sebesar 31,1977 kg/kapita/tahun. Kedua nilai tersebut merupakan yang tertinggi di antara ketiga klaster. Klaster ini terdiri atas lima komoditas, yaitu kedelai, sagu, kacang tanah, beras, dan ubi kayu. Hasil tersebut memperlihatkan bahwa kelompok ini memiliki karakteristik penggunaan untuk pengolahan dan ketersediaan per kapita yang lebih menonjol dibandingkan kelompok lainnya.

Sementara itu, Klaster 3 merupakan kelompok terbesar dengan 98 komoditas. Rata-rata rasio impornya sebesar 0,0816, lebih rendah dibandingkan Klaster 1 dan Klaster 2. Rasio ekspor, penggunaan pakan, dan penggunaan pengolahan pada kelompok ini juga relatif rendah. Berdasarkan karakteristik tersebut, Klaster 3 menggambarkan kelompok komoditas dengan kontribusi impor dan penggunaan untuk pakan maupun pengolahan yang relatif lebih rendah berdasarkan indikator penelitian. Secara keseluruhan, hasil pengelompokan memperlihatkan bahwa komoditas pangan di Provinsi Jawa Timur memiliki karakteristik penyediaan dan penggunaan yang beragam.

Validasi dan Interpretasi Hasil Klaster

Evaluasi hasil klasterisasi dilakukan menggunakan *Silhouette Score*, *Davies-Bouldin Index* (DBI), dan *Calinski-Harabasz Index* (CHI). Selain itu, dilakukan perbandingan dengan metode Ward serta pengujian stabilitas melalui 100 kali pengulangan sebagaimana disajikan pada Tabel 5.

Tabel 5. Hasil Evaluasi Model K-Means

Indikator Evaluasi	Nilai
Silhouette Score keseluruhan	0,3735
Davies-Bouldin Index	1,1100
Calinski-Harabasz Index	55,0527

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Indikator Evaluasi	Nilai
ARI K-Means dengan Ward	0,3885
Rata-rata ARI stabilitas (100 pengulangan)	0,7413
Standar deviasi ARI stabilitas	0,2091
Silhouette rata-rata Klaster 1	0,1363
Silhouette rata-rata Klaster 2	0,1403
Silhouette rata-rata Klaster 3	0,4605

Nilai *Silhouette Score* keseluruhan sebesar 0,3735 menunjukkan bahwa pemisahan antarklaster masih memiliki ruang untuk perbaikan. Nilai DBI sebesar 1,1100 dan CHI sebesar 55,0527 digunakan sebagai indikator tambahan untuk menilai kekompakan serta pemisahan kelompok. Jika dilihat berdasarkan nilai *Silhouette* rata-rata setiap klaster, Klaster 3 memiliki nilai tertinggi sebesar 0,4605, sedangkan Klaster 1 dan Klaster 2 masing-masing sebesar 0,1363 dan 0,1403. Hal ini menunjukkan bahwa kualitas pemisahan pada masing-masing kelompok belum seragam.

Perbandingan hasil K-Means dengan metode Ward menghasilkan nilai *Adjusted Rand Index* (ARI) sebesar 0,3885, yang menunjukkan bahwa susunan keanggotaan klaster dari kedua metode memiliki tingkat kesesuaian yang terbatas. Sementara itu, pengujian stabilitas melalui 100 pengulangan menghasilkan rata-rata ARI sebesar 0,7413 dengan standar deviasi 0,2091, yang menggambarkan bahwa struktur pengelompokan memiliki tingkat kestabilan tertentu meskipun masih bervariasi pada setiap pengulangan.

Dengan demikian, hasil klasterisasi dapat digunakan untuk menggambarkan perbedaan karakteristik komoditas pangan berdasarkan lima indikator yang digunakan. Namun, interpretasi hasil perlu mempertimbangkan nilai evaluasi yang diperoleh, ketidakseimbangan jumlah anggota klaster, serta keterbatasan data yang hanya mencakup satu periode, yaitu tahun 2025. Hasil pengelompokan ini tidak secara langsung menunjukkan tingkat kerawanan pangan, tetapi dapat menjadi informasi pendukung dalam mengidentifikasi komoditas berdasarkan struktur penyediaan dan ketergantungan impornya.

PENUTUP

Kesimpulan

Penelitian ini menghasilkan pengelompokan 134 komoditas pangan NBM Provinsi Jawa Timur Tahun 2025 menjadi tiga klaster menggunakan algoritma K-Means berdasarkan lima indikator penyediaan dan penggunaan pangan. Klaster 1 memiliki rasio impor dan ekspor tertinggi, Klaster 2 memiliki rasio pengolahan dan ketersediaan per kapita tertinggi, sedangkan Klaster 3 merupakan kelompok terbesar dengan rasio impor relatif rendah. Nilai *Silhouette Score* sebesar 0,3735 menunjukkan bahwa pemisahan antarklaster masih belum optimal. Meskipun demikian, hasil pengelompokan dapat memberikan gambaran perbedaan karakteristik komoditas pangan sebagai informasi pendukung evaluasi ketersediaan pangan daerah.

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Saran

Penelitian ini terbatas pada data satu periode (tahun 2025) dan ketidakseimbangan jumlah anggota antarklaster. Penelitian selanjutnya disarankan menggunakan data deret waktu beberapa tahun, menambahkan variabel lain yang relevan, serta membandingkan algoritma klasterisasi dan metode penentuan jumlah klaster lainnya untuk memperoleh hasil pengelompokan yang lebih optimal.

DAFTAR PUSTAKA

Arpipi, M. Y. R. (2020). Eksplorasi penggunaan algoritma K-Means clustering untuk analisis tren harga komoditas wilayah Indonesia Timur. *Jurnal Komputer dan Informatika*, 15(2), 251–257.

Arthur, D., & Vassilvitskii, S. (2007). k-means++: The advantages of careful seeding. *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, 1027–1035.

Atasa, D., & Nugroho, T. W. (2021). Analisis ketersediaan pangan Kota Malang. *Mimbar Agribisnis: Jurnal Pemikiran Masyarakat Ilmiah Berwawasan Agribisnis*, 7(2), 1185–1194. https://doi.org/10.25157/ma.v7i2.5191

Atasa, D., Laily, D. W., & Wijayanti, P. D. (2022). Dinamika ketersediaan pangan dan alih fungsi lahan pertanian Kota Malang. *Agrinika: Jurnal Agroteknologi dan Agribisnis*, 6(1), 10–22. https://doi.org/10.30737/agrinika.v6i1.2171

Badan Pangan Nasional. (2024). *Neraca bahan makanan Indonesia 2022–2024*. https://data.badanpangan.go.id/statisticpublications/56w

Dahwanu, O., Abdillah, N., Akbar, N., & Alghifari, H. (2025). Klasterisasi keragaman jenis konsumsi pangan menggunakan algoritma K-Means dan implikasinya terhadap kebijakan daerah. *J-ENSITEC (Journal of Engineering and Sustainable Technology)*, 12(1), 10293–10298. https://doi.org/10.31949/j-ensitec.v12i01.16079

Danasari, I. F., Sari, N. M. W., Septiadi, D., & Supartiningsih, N. L. S. (2023). Dependency on beef import to support food security in Indonesia. *IOP Conference Series: Earth and Environmental Science*, 1253(1), Article 012088. https://doi.org/10.1088/1755-1315/1253/1/012088

Hakim, D. L., Machalett, B., Suhendra, A., Arifin, S., Kurniawan, A., & Putra, M. G. (2026). Developing a comprehensive food availability vulnerability index (FAVI) to address food security challenges in the ASEAN region. *Agriculture & Food Security*, 15, Article 55. https://doi.org/10.1186/s40066-026-00627-7

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666. https://doi.org/10.1016/j.patrec.2009.09.011

Kristianto, A., Kastaman, R., Pujianto, T., & Setiasih, I. (2011). Kajian neraca bahan makanan dalam rangka penyusunan strategi dan kebijakan ketahanan pangan Kota Bandung hingga tahun 2020. *Teknotan: Jurnal Industri Teknologi Pertanian*, 5(2).

Mardianto, M. F. F., Siregar, N. R. A. A., Soewignjo, S., Putri, F. F. R., Prayogi, H., Imama, C., Amelia, D., Sediono, & Dewi, D. A. (2024). Time series clustering analysis for increases food commodity prices in Indonesia based on K-Means method. *Journal of Human, Earth, and Future*, 5(3), 319–329. https://doi.org/10.28991/HEF-2024-05-03-02

Muhammad, A. M., Napitupulu, D. M. T., & Suandi. (2022). Keterkaitan lahan pangan dengan neraca bahan makanan dan pola pangan harapan Kota Jambi. *Jurnal Pembangunan Berkelanjutan*, 5(1), 30–36. https://doi.org/10.22437/jpb.v5i1.18626

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., VanderPlas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(85), 2825–2830.

Priambodo, N. D. (2019). Analisis neraca bahan makanan Kota Kediri tahun 2018. *Jurnal Agriovet*, 1(2), 225–236.

Renaldi, E., & Pardian, P. (2024). Analisis neraca bahan makanan Kabupaten Garut tahun 2022. *Mimbar Agribisnis: Jurnal Pemikiran Masyarakat Ilmiah Berwawasan Agribisnis*, 10(2), 3292–3296. https://doi.org/10.25157/ma.v10i2.14746

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Sifriyani, Budiantara, I. N., Mardianto, M. F. F., Febriyani, E. R., Chairunnisa, N. R., & Putri, A. C. (2024). Implementation of data mining and spatial mapping in determining national food security clusterization. *Jurnal Teori dan Aplikasi Matematika*, 8(3), 800–810. https://doi.org/10.31764/jtam.v8i3.19912

Sumargo, E. G., & Handhayani, T. (2025). Analisis clustering harga pangan di pasar tradisional wilayah Kalimantan menggunakan algoritma K-Means dan Fuzzy C-Means. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(5). https://doi.org/10.36040/jati.v9i5.15135

Suryadi, J. (2024). Clustering ketidakcukupan konsumsi pangan per kabupaten/kota menggunakan algoritma K-Means dan Fuzzy C-Means. *Jurnal Komputer dan Informatika*, 19(1), 30–35.

Wahida, Perdana, R. P., Septanti, K. S., Suryana, E. A., Ulpah, A., Raharjo, A. S. S., Suharyono, S., & Yuliani, F. (2025). Dinamika tingkat kemandirian pangan: Implikasinya terhadap kebijakan ketahanan pangan nasional. *Analisis Kebijakan Pertanian*, 23(1), 37–55. https://doi.org/10.21082/akp.v23n1.2025.37-55

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)