

Analisis Pengelompokan Tingkat Ketahanan Pangan Provinsi di Indonesia Tahun 2019–2023 Menggunakan Algoritma K-Means

Yunisha Erin Prestisya¹, Nadila Nur Rahma², Dhea Aningthia³

^{1,2,3}STIKOM Tunas Bangsa, Pematangsiantar

e-mail: yunishaerinprestisya@gmail.com, nurrahmanadila2@gmail.com, dheaaningthia@gmail.com

Abstrak

Penelitian ini menganalisis pengelompokan tingkat ketahanan pangan provinsi di Indonesia periode 2019–2023 menggunakan algoritma K-Means. Data berupa skor indeks ketahanan pangan provinsi dianalisis dengan perangkat lunak RapidMiner. Tiga cluster terbentuk, masing-masing mewakili kategori rendah, sedang, dan tinggi. Pengukuran kinerja dengan Davies Bouldin Index sebesar $-0,614$ menandakan hasil clustering yang baik dan pemisahan cluster yang jelas. Hasil menunjukkan disparitas ketahanan pangan antarprovinsi yang signifikan. Provinsi dengan cluster rendah memiliki ketersediaan dan akses pangan terbatas, sementara cluster tinggi menunjukkan ketahanan pangan kuat dan distribusi merata. Studi ini memberikan informasi penting untuk mendukung kebijakan pembangunan ketahanan pangan yang lebih tepat sasaran di Indonesia.

Kata kunci: Ketahanan pangan, K-Means, Clustering

Abstract

This study analyzes the grouping of provincial food security levels in Indonesia for the period 2019–2023 using the K-Means algorithm. Data in the form of provincial food security index scores were analyzed using RapidMiner software. Three clusters were formed, each representing the low, medium, and high categories. Performance measurement using the Davies Bouldin Index of -0.614 indicates good clustering results and clear cluster separation. The results show significant disparities in food security between provinces. Provinces in the low cluster have limited food availability and access, while the high cluster shows strong food security and equitable distribution. This study provides important information to support more targeted food security development policies in Indonesia

Keywords: Food security, K-Means, Clustering

1. Pendahuluan

Ketahanan pangan menjadi salah satu isu penting yang mendapat perhatian utama dalam upaya pembangunan nasional[1][2][3]. Pangan selain fungsi gizi, juga menjaga stabilitas sosial, ekonomi, dan politik. Berdasarkan Undang-Undang Nomor 18 Tahun 2012 tentang Pangan, ketahanan pangan adalah kondisi ketersediaan pangan cukup, berkualitas, aman, merata, dan mudah diakses oleh masyarakat[4].

Indonesia sebagai negara kepulauan memiliki karakteristik geografis, sosial, dan ekonomi yang beragam. Keragaman ini berdampak pada perbedaan kemampuan tiap provinsi dalam menjaga ketersediaan, akses, dan pemanfaatan pangan. Data Badan Pangan Nasional (2023) melalui Indeks Ketahanan Pangan (IKP) menunjukkan bahwa terdapat ketimpangan signifikan antarprovinsi dalam pencapaian ketahanan pangan[5][6][7]. Beberapa provinsi telah mencapai kondisi ketahanan pangan yang relatif baik, sementara provinsi lain masih menghadapi masalah akses pangan, distribusi, maupun kualitas konsumsi masyarakat. Oleh karena itu, diperlukan metode analisis yang mampu mengelompokkan provinsi

berdasarkan tingkat ketahanan pangannya untuk membantu pemerintah dalam perumusan kebijakan yang lebih tepat sasaran[8][9].

Salah satu metode yang dapat diterapkan dalam proses pengelompokan data adalah algoritma *K-Means*. *K-Means* merupakan salah satu metode *unsupervised learning* yang banyak digunakan dalam analisis kluster karena mampu mengelompokkan data berdasarkan kesamaan karakteristik[10][11][12][13].

Penelitian ini memiliki ruang lingkup yang jelas, yaitu menganalisis tingkat ketahanan pangan di 34 provinsi di Indonesia selama periode 2019–2023 dengan fokus pada pengelompokan menggunakan algoritma *K-Means*. Batasan penelitian meliputi penggunaan data indeks ketahanan pangan dari Badan Pangan Nasional dan metode clustering tanpa memproses faktor eksternal yang lebih luas. Asumsi utama adalah data indeks yang digunakan valid dan representatif kondisi ketahanan pangan tiap provinsi. Hasil pengelompokan ini diharapkan dapat memberikan gambaran yang lebih jelas mengenai pola distribusi ketahanan pangan di Indonesia serta menjadi dasar dalam pengambilan kebijakan yang berbasis data[14][15][16][17][18]. Tujuan penelitian ini adalah mengelompokkan tingkat ketahanan pangan provinsi guna memberikan gambaran pola distribusi ketahanan pangan, serta hasilnya diharapkan dapat mendukung perumusan kebijakan yang lebih tepat sasaran oleh pemerintah pusat dan daerah. Melalui proses pengelompokan ini, diharapkan dapat diperoleh informasi mengenai kondisi ketahanan pangan antar daerah, sehingga pemerintah pusat maupun daerah dapat menyusun strategi yang lebih efektif dalam meningkatkan ketahanan pangan nasional[19][20].

2. Metode Penelitian

2.1. Data Mining

Penelitian ini menggunakan metode data mining berbasis algoritma *K-Means* untuk mengelompokkan provinsi menurut tingkat ketahanan pangan[21]. Pengolahan data dilakukan dengan bantuan perangkat lunak *RapidMiner* versi 8.1 yang menyediakan antarmuka visual untuk pemodelan *clustering*.

2.. Algoritma *K-Means*

Algoritma *K-Means* merupakan metode klusterisasi yang sesuai untuk data dengan atribut bertipe numerik. Algoritma ini bekerja dengan cara mengelompokkan data berdasarkan jarak terdekat ke pusat kluster (*centroid*). rumus jarak Euclidean (*Euclidean Distance*):

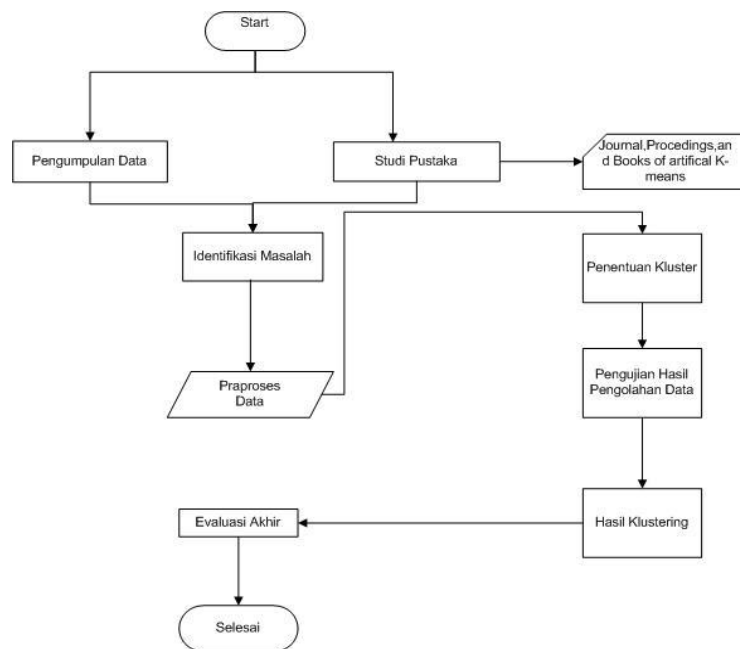
$$D_e = \sqrt{\sum_{i=1}^n (x_i - s_i)^2} \quad (1)$$

Dimana:

- D_e adalah jarak *Euclidean*
- n adalah jumlah atribut
- x_i adalah koordinat data
- s_i adalah koordinat *centroid*

2.3 Tahapan Penelitian

Penelitian ini dilakukan melalui sejumlah tahap inti, yang mencakup proses pengumpulan data, tahap praproses, implementasi algoritma *K-Means*, evaluasi kinerja, serta analisis hasil. Rangkaian tahapan secara keseluruhan ditampilkan pada Gambar 1:



Gambar 1. Tahapan Penelitian

3. Hasil dan Pembahasan

3.1. Pengumpulan Data

Data yang digunakan merupakan data tingkat ketahanan pangan provinsi di Indonesia dari tahun 2019 hingga 2023. Data ini terdiri dari skor indeks ketahanan pangan per provinsi selama lima tahun berturut-turut. Total data yang dianalisis sebanyak 34 provinsi dengan beberapa atribut terkait ketahanan pangan per tahun.

Tabel 1. Dataset Ketahanan Pangan Nasional

Provinsi	2019	2020	2021	2022	2023
Aceh	66,22	70,92	71,63	70,16	72,96
Sumatera Utara	69,81	71,84	72,25	71,22	75,97
Sumatera Barat	75,43	78,64	79,55	79,45	83,22
Riau	62,37	64,12	66,84	67,59	68,68
Jambi	68,23	70	74,18	69,5	72,17
Sumatera Selatan	69,3	68,67	69,55	69,64	73,82
Bengkulu	61,78	70,28	70,32	67,99	72,27
Lampung	71,36	77,43	77,96	78,61	81,56
Kepulauan Bangka Belitung	56,03	71,21	73,22	71,71	71,14
Kepulauan Riau	59,26	62,7	63,26	63,83	65,1
DKI Jakarta	66,87	77,97	78,01	78,25	83,8
Jawa Barat	76,44	76,78	77,79	77,55	82,19
Jawa Tengah	78,85	82,31	82,73	82,95	84,8
DI Yogyakarta	83,63	80,67	81,43	80,88	83,17
Jawa Timur	73,71	79,9	79,7	79,85	82,46
Banten	74,47	73,48	74,38	73,78	78,71
Bali	85,15	84,54	83,82	85,19	87,65
Nusa Tenggara Barat	62,43	75,6	75,67	76,58	76,51
Nusa Tenggara Timur	50,69	66,92	67,35	68,42	71,25

Kalimantan Barat	55,17	71,13	71,32	70,81	72,2
Kalimantan Tengah	71,57	72,58	73,68	69,96	68,9
Kalimantan Selatan	74,71	80,04	80,29	81,05	81,26
Kalimantan Timur	76,9	78,24	77,46	77,65	72,29
Kalimantan Utara	73,12	71,9	73,02	71,04	74,59
Sulawesi Utara	81,44	77,79	78,3	74,3	77,32
Sulawesi Tengah	68,17	75,1	75,73	75,92	75,83
Sulawesi Selatan	78,69	81,81	80,82	81,38	83,36
Sulawesi Utara	76,99	77,06	76,64	75,04	74,96
Gorontalo	69,06	80,4	80,52	80,35	81,63
Sulawesi Barat	60,37	76,36	75,49	74,04	73,03
Maluku	52,35	58,15	58,7	60,2	64,37
Maluku Utara	66,58	63,12	59,58	58,39	62,34
Papua Barat	30,12	49,4	46,05	45,92	47,95
Papua	25,13	34,79	35,48	37,8	42,27

3.2. Praproses Data

Tahap praproses meliputi pembersihan data dan pemilihan atribut nilai ketahanan pangan tahun sebagai variabel input. Data diperiksa untuk menghindari duplikasi atau nilai kosong agar sesuai untuk analisis *clustering*.

3.3. Penentuan Cluster

Pengelompokan dilakukan dengan algoritma *K-Means* mengelompokkan provinsi berdasarkan indeks ketahanan pangan selama periode 2019–2023. Pengukuran kinerja menggunakan nilai "Davies Bouldin Index". Semakin mendekati nol nilai indeks ini, semakin baik hasil clustering.

3.4 Pengujian Hasil Pengolahan Data

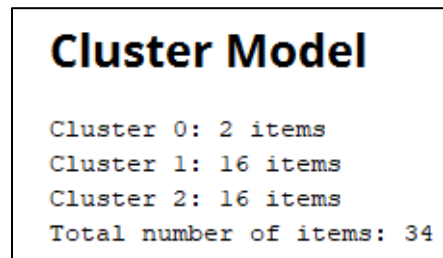
Hasil evaluasi kinerja model clustering ditunjukkan melalui PerformanceVector. Nilai Average within centroid distance sebesar -97.184 menunjukkan jarak rata-rata antara data dan pusat clusternya. Nilai Davies Bouldin Index sebesar -0.614 menunjukkan tingkat pemisahan antar cluster yang baik.

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: -97.184
Avg. within centroid distance_cluster_0: -115.000
Avg. within centroid distance_cluster_1: -128.785
Avg. within centroid distance_cluster_2: -63.355
Davies Bouldin: -0.614
```

Gambar 2. Performance Vector

Sebagai hasil dari penentuan jumlah cluster sebanyak 3 dan nilai *Davies Bouldin Index* sebesar -0.614, didapatkan jumlah anggota *cluster* 0 sebanyak 2 provinsi, *cluster* 1 sebanyak 16 provinsi, dan *cluster* 2 sebanyak 16 provinsi.



Gambar 3. Model Hasil *Clustering* Ketahanan Pangan

Penelitian ini menggunakan data indeks ketahanan pangan yang telah tervalidasi, namun terdapat keterbatasan pada cakupan variabel yang hanya mencakup aspek ketahanan pangan tanpa memperhitungkan faktor sosial-ekonomi atau kebijakan lokal yang dapat memengaruhi kondisi riil.

3.5 Hasil *Clustering*

Untuk memahami karakteristik tiap kelompok hasil pengelompokan, diperlukan analisis terhadap nilai centroid dari masing-masing *cluster*. Nilai *centroid* merepresentasikan rata-rata posisi setiap atribut dalam kelompok tersebut, sehingga dapat menggambarkan profil umum dari tiap *cluster*. Berikut disajikan tabel hasil perhitungan centroid dari algoritma K-Means yang menunjukkan nilai rata-rata setiap variabel pembentuk ketahanan pangan pada masing-masing *cluster*.

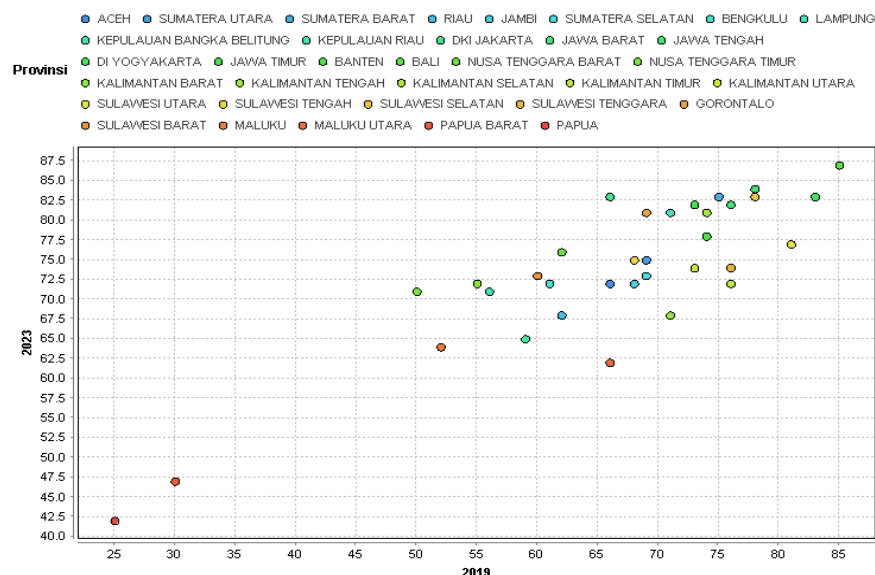
Attribute ↑	cluster_0	cluster_1	cluster_2
2019	27.500	62.438	75.188
2020	41.500	68.625	78.375
2021	40.500	69.312	78.500
2022	41	68.312	78.375
2023	44.500	70.500	80.375

Gambar 4. Titik Pusat Klaster Ketahanan Pangan Provinsi di Indonesia Periode 2019–2023

Berdasarkan gambar di atas, *cluster* 0 memiliki rata-rata nilai ketahanan pangan terendah (27,50–44,50) yang menggambarkan daerah dengan tingkat ketersediaan pangan yang masih terbatas, akses distribusi yang belum merata, serta kerentanan ekonomi yang cukup tinggi. Wilayah dalam kelompok ini perlu mendapatkan perhatian lebih dalam hal peningkatan produksi dan pemerataan distribusi pangan. *Cluster* 1 menunjukkan nilai ketahanan pangan pada kategori sedang (62,43–70,50). Provinsi dalam kelompok ini cenderung memiliki kondisi ketahanan pangan yang cukup stabil, dengan ketersediaan dan akses pangan yang relatif baik namun belum sepenuhnya merata. Upaya peningkatan efisiensi distribusi dan penguatan cadangan pangan daerah masih diperlukan agar dapat mencapai kategori tinggi. Sementara itu, *Cluster* 2 memiliki nilai rata-rata tertinggi (75,18–80,37), yang menunjukkan tingkat ketahanan pangan yang baik dan stabil. Provinsi dalam kelompok ini umumnya memiliki ketersediaan pangan yang melimpah, infrastruktur distribusi yang mendukung, serta tingkat kesejahteraan masyarakat yang lebih tinggi.

3.6 Visualisasi Hasil *Clustering*

Untuk memahami pola distribusi antarprovinsi dalam setiap kelompok hasil klasterisasi, dilakukan visualisasi menggunakan *scatter plot*. Grafik ini menggambarkan persebaran nilai ketahanan pangan tiap provinsi berdasarkan hasil pengelompokan dengan algoritma *K-Means* pada periode 2019–2023.



Gambar 5. Scatter Plot Hasil Clustering

Sumbu X merepresentasikan nilai ketahanan pangan tahun 2019, sedangkan sumbu Y menunjukkan nilai tahun 2023. Warna merah mewakili cluster rendah, yaitu provinsi seperti Papua dan Papua Barat yang memiliki nilai ketahanan pangan terendah dan cenderung tidak stabil. Warna hijau menunjukkan cluster sedang, meliputi provinsi seperti Sumatera Utara, Sulawesi Tengah, dan Kalimantan Selatan dengan kondisi ketahanan pangan cukup baik namun belum merata. Sementara itu, warna biru menggambarkan cluster tinggi, seperti Bali, DI Yogyakarta, dan Jawa Tengah, yang memiliki ketahanan pangan tinggi dan distribusi pangan yang merata. Pola persebaran titik memperlihatkan gradasi horizontal dari kiri ke kanan, di mana semakin ke kanan, tingkat ketahanan pangan semakin baik. Hal ini mengindikasikan peningkatan kinerja ketahanan pangan secara umum pada sebagian besar provinsi Indonesia selama periode pengamatan.

4. Kesimpulan

Analisis yang dilakukan dengan metode *K-Means* menghasilkan pengelompokan tingkat ketahanan pangan provinsi di Indonesia selama periode 2019–2023 ke dalam tiga kategori utama: rendah, sedang, dan tinggi. Skor *Davies–Bouldin Index* sebesar -0.614 mengindikasikan bahwa hasil klusterisasi memiliki pemisahan antar kelompok yang berkualitas baik. Cluster rendah menggambarkan provinsi dengan keterbatasan ketersediaan dan akses pangan, cluster sedang menunjukkan kondisi cukup stabil namun masih memerlukan peningkatan distribusi, sedangkan cluster tinggi mencerminkan provinsi dengan ketahanan pangan yang kuat dan distribusi yang merata. Hasil ini secara langsung menjawab tujuan penelitian, yaitu menghasilkan model pengelompokan yang mampu memetakan ketimpangan ketahanan pangan antarprovinsi di Indonesia. Penelitian ini dapat dijadikan dasar bagi pemerintah dalam menyusun kebijakan berbasis data untuk memperkuat daerah dengan kategori rendah. Penelitian selanjutnya disarankan untuk menambahkan variabel sosial ekonomi dan menggunakan metode *clustering* agar hasil analisis lebih komprehensif dan akurat.

Daftar Pustaka

- [1] T. K. Wijayanti, K. I. Hafsari, M. T. Al Hijrah, dan R. Kurniawan, “Implementasi Algoritma K-Means pada Pengelompokan Ketahanan Pangan di Indonesia Menurut Kabupaten/Kota,” in *PROSIDING SEMINAR NASIONAL SAINS DATA*, 2024, vol. 4, no. 1, hal. 52–63.
- [2] A. Ramadhan, K. Prawita, M. A. Izzudin, dan G. Amandha, “Analisis strategi dan klusterisasi ketahanan pangan nasional dalam menghadapi pandemi covid-19,” *Teknol. Pangan Media Inf. Dan*

- Komun. Ilm. Teknol. Pertan.*, vol. 12, no. 1, hal. 110–122, 2021.
- [3] A. Idris dan M. H. ZM, “Ketahanan Pangan, Air, Energi Dan Pertanian: Analisis Sebaran Spasial Dan Clustering Provinsi Di Indonesia,” *The Journalish Soc. Gov.*, vol. 4, no. 5, hal. 27–40, 2023.
 - [4] R. Farismana, “Penerapan K-Means Clustering Untuk Pemetaan Produktivitas Padi Dan Prediksi Panen Di Kabupaten Indramayu,” *J. Inf. Syst. Applied, Manag. Account. Res.*, vol. 8, no. 3, hal. 589–605, 2024.
 - [5] E. R. Arini, “Penerapan K-Means Cluster di Provinsi Jawa Timur Berdasarkan Ketahanan Pangan,” *J. Sci. Nusantara*, vol. 3, no. 1, hal. 32–36, 2023.
 - [6] F. N. Arifin dan O. Juwita, “Klasterisasi Wilayah Kabupaten Jember Berdasarkan Tingkat Ketahanan Pangan,” *INFORMAL Informatics J.*, vol. 7, no. 2, hal. 95–99, 2022.
 - [7] L. M. A. Fadila dan N. A. Putri, “Analisis Perkembangan Ketahanan Pangan di Indonesia: Pendekatan Menggunakan Big Data dan Data Mining,” in *Seminar Nasional Official Statistics*, 2023, vol. 2023, no. 1, hal. 247–256.
 - [8] M. Diarty dan A. W. Wijayanto, “Analisis Aspek Ketahanan Pangan Indonesia dengan Hard dan Soft Clustering, 2022,” *Rekayasa*, vol. 17, no. 1, hal. 108–123, 2024.
 - [9] F. Nasution, “Penerapan Metode K-Means Clustering untuk Mengelompokkan Ketahanan Tanaman Pangan Kabupaten/Kota di Provinsi Sumatera Utara.” Universitas Sumatera Utara, 2019.
 - [10] K. Z. Fuddiani, “Pengelompokan Desa di Kabupaten Tuban Berdasarkan Kondisi Ketahanan Pangan Menggunakan Metode Integrasi K-Means dan Self Organizing Maps.” Institut Teknologi Sepuluh Nopember, 2024.
 - [11] W. W. Prastanika dan A. W. Wijayanto, “Analisis Hard dan Soft Clustering Untuk Pengelompokan Indikator Ketahanan Pangan Indonesia 2021,” *JUSTIN (Jurnal Sist. dan Teknol. Informasi)*, vol. 11, no. 4, hal. 596–604, 2023.
 - [12] A. S. S. Khan, M. Fatekurohman, dan Y. S. Dewi, “Perbandingan Algoritma K-Medoids Dan K-Means Dalam Pengelompokan Kecamatan Berdasarkan Produksi Padi Dan Palawija Di Jember,” *J. Stat. Dan Komputasi*, vol. 2, no. 2, hal. 67–75, 2023.
 - [13] D. S. Perdana, “Penerapan Metode K-Means Clustering Untuk Mengelompokkan Ketahanan Pangan,” *J. Sist. Inf.*, vol. 3, no. 2, hal. 67–72, 2022.
 - [14] F. M. Izzadin, “Optimasi Jumlah Cluster K-Means dengan Metode Elbow dan Silhouette Pada Produktivitas Tanaman Pangan di Provinsi Jawa Tengah Tahun 2018,” 2020.
 - [15] A. Z. Umarni, “Analisi Cluster untuk Mengelompokkan Kabupaten/Kota di Sulawesi Selatan Berdasarkan Produktivitas Tanaman Pangan Tahun 2023.” Institut Teknologi Kalimantan, 2025.
 - [16] A. L. Zahra, S. T. R. Ada, dan A. F. Ardini, “Implementasi Clustering Algoritma K-Means Pada Produksi Beras di Provinsi Jawa Timur Tahun 2022,” *J. Comput. Inf. Syst. Ampera*, vol. 5, no. 3, hal. 191–202, 2024.
 - [17] A. I. Safe’i, “Aplikasi K-Means untuk Pengelompokan Kabupaten dan Kota Berdasarkan Produktivitas Tanaman Pangan di Provinsi Jawa Tengah Tahun 2015,” 2018.
 - [18] M. N. Maulina, “Optimalisasi Cluster pada Cluster Hierarki Menggunakan Pseudo F-Statistic Calinski Harabasz Untuk Ketahanan Pangan,” *J. Gaussian*, vol. 14, no. 1, hal. 1–12, 2025.
 - [19] S. Fitri, “Penentuan cluster terbaik pada pengelompokan ketahanan pangan menggunakan metode Self-Organizing Maps (SOM) dan metode Silhouette Coefficient (sc): Studi kasus Provinsi Jawa Timur.” Universitas Islam Negeri Maulana Malik Ibrahim, 2022.
 - [20] W. Budiaji dan J. Pancawati, “Medoid-based Clustering pada Kecamatan di Kabupaten Lebak dan Pandeglang Provinsi Banten Berdasarkan Trilogi Ketahanan Pangan,” *J Stat. J. Ilm. Teor. dan Apl. Stat.*, vol. 15, no. 1, 2022.
 - [21] A. S. Yaumi, Z. Zulfiqar, dan A. Nugroho, “Klasterisasi Karakter Konsumen Terhadap Kecenderungan Pemilihan Produk Menggunakan K-Means,” *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 5, no. 3, hal. 195, 2020, doi: 10.31328/jointecs.v5i3.1523.