

---

---

## ANALISIS KOMPARATIF ALGORITMA K-MEANS DAN K-MEDOIDS DALAM *CLUSTERING* RASIO DISTRIBUSI ALOKON TERHADAP PUS DI PROVINSI SUMATERA UTARA TAHUN 2025

Panggabean Siahaan<sup>1</sup>, Muhammad Irfan Sarif<sup>2</sup>, Siti Qomariyah<sup>3</sup>, Satria Sinurat<sup>4</sup>,  
Norita Tampubolon<sup>5</sup>

Universitas Pembangunan Panca Budi, Medan

e-mail: <sup>1</sup>[gabesiahaan@gmail.com](mailto:gabesiahaan@gmail.com), <sup>2</sup>[irfanberbagi@gmail.com](mailto:irfanberbagi@gmail.com),

<sup>3</sup>[sitidarma140395@gmail.com](mailto:sitidarma140395@gmail.com), <sup>4</sup>[satria.sinurat@yahoo.co.id](mailto:satria.sinurat@yahoo.co.id), <sup>5</sup>[queen.norita07@gmail.com](mailto:queen.norita07@gmail.com)

**Abstract:** Unequal distribution of contraceptive supplies (Alokon) relative to the target population remains a persistent challenge in family planning programs, particularly in regions with high demographic heterogeneity. Evaluation based on absolute distribution values generates proportional bias, as larger-population areas automatically receive higher volumes without accounting for the proportional needs of the Reproductive Age Couples (*PUS*) population. This study proposes a ratio-based approach — dividing total *Alokon* distributed by the number of *PUS* — as the primary clustering variable to enable proportional comparison and reduce population-scale bias across 33 districts and cities in North Sumatra Province. Two algorithms, K-Means and K-Medoids based on Partitioning Around Medoids (PAM), were comparatively evaluated using Silhouette Score as the evaluation metric. The optimal number of clusters ( $K = 3$ ) was determined through a combination of the Elbow Method — which identified a 75.12% WCSS reduction from  $K = 2$  to  $K = 3$  — and Silhouette Score validation. Results show that both algorithms produced identical cluster compositions: 16 districts in the low-distribution group (48.5%;  $\bar{R} = 0.1687$ ), 13 districts in the moderate group (39.4%;  $\bar{R} = 0.3117$ ), and 4 districts in the high group (12.1%;  $\bar{R} = 0.6077$ ), with equal average Silhouette Scores of  $\bar{s} = 0.6998$  (reasonable structure). Densely populated areas such as Medan City and Deli Serdang — despite receiving the highest absolute distribution volumes — were classified in the low group when measured proportionally, demonstrating the superiority of the ratio-based approach. To the best of the authors' knowledge, this study is the first to apply comparative clustering on *Alokon* distribution using a proportional ratio framework in North Sumatra Province, providing empirical evidence on algorithm performance in normalized health service distribution data.

**Keywords:** *clustering; contraceptive distribution; K-Means; K-Medoids; Silhouette Score*

**Abstrak:** Ketidakteraturan distribusi alat dan obat kontrasepsi (Alokon) terhadap populasi sasaran menjadi tantangan dalam program Keluarga Berencana, khususnya pada wilayah dengan heterogenitas demografis yang tinggi. Evaluasi berbasis nilai absolut distribusi menghasilkan bias proporsional karena wilayah berpenduduk besar secara otomatis menerima volume lebih tinggi tanpa mempertimbangkan proporsi kebutuhan terhadap jumlah Pasangan Usia Subur (PUS). Penelitian ini mengusulkan pendekatan berbasis rasio distribusi Alokon terhadap PUS sebagai variabel *clustering* utama untuk memungkinkan perbandingan yang proporsional dan mengurangi bias skala populasi pada 33 kabupaten/kota di Provinsi Sumatera Utara. Dua algoritma, K-Means dan K-Medoids berbasis *Partitioning Around Medoids* (PAM), dibandingkan menggunakan *Silhouette Score* sebagai metrik evaluasi. Jumlah *cluster* optimal ( $K = 3$ ) ditentukan melalui kombinasi *Elbow Method* — yang mengidentifikasi penurunan WCSS sebesar 75,12% dari  $K = 2$  ke  $K = 3$  — dan validasi *Silhouette Score*. Hasil menunjukkan bahwa kedua algoritma menghasilkan komposisi *cluster* yang identik: 16 wilayah dalam kelompok distribusi rendah (48,5%;  $\bar{R} = 0,1687$ ), 13 wilayah kelompok sedang

(39,4%;  $\bar{R}= 0,3117$ ), dan 4 wilayah kelompok tinggi (12,1%;  $\bar{R}= 0,6077$ ), dengan nilai  $\bar{S}= 0,6998$  (*reasonable structure*) pada keduanya. Wilayah berpenduduk besar seperti Kota Medan dan Deli Serdang — meskipun menerima volume distribusi absolut tertinggi — terklasifikasi dalam kelompok rendah ketika diukur secara proporsional, menunjukkan keunggulan pendekatan berbasis rasio. Sejauh pengetahuan penulis, penelitian ini merupakan studi pertama yang menerapkan *clustering* komparatif pada distribusi Alokon berbasis rasio di Provinsi Sumatera Utara dan memberikan bukti empiris mengenai performa algoritma pada data distribusi layanan kesehatan yang telah dinormalisasi.

**Kata Kunci:** clustering; distribusi kontrasepsi; K-Means; K-Medoids; Silhouette Score

## PENDAHULUAN

Distribusi alat dan obat kontrasepsi (Alokon) merupakan salah satu pilar utama dalam pelaksanaan program Keluarga Berencana (KB) nasional yang dikoordinasikan oleh BKKBN. Ketersediaan Alokon yang memadai dan merata di tingkat wilayah menjadi prasyarat agar akseptor dapat mengakses layanan kontrasepsi secara berkesinambungan. Distribusi yang efektif idealnya bersifat proporsional terhadap kebutuhan populasi sasaran, khususnya jumlah Pasangan Usia Subur (PUS) di masing-masing wilayah. Namun, dalam praktiknya, distribusi Alokon antarkabupaten/kota kerap tidak mencerminkan proporsi kebutuhan yang sesungguhnya, sehingga menimbulkan ketidakseimbangan berupa kelebihan distribusi di sejumlah daerah dan kekurangan di daerah lainnya. Ketidakseimbangan tersebut berpotensi menurunkan efektivitas program KB dan menghambat akselerasi pemerataan layanan kontrasepsi bagi seluruh lapisan masyarakat.

Di Provinsi Sumatera Utara, tantangan distribusi Alokon memiliki dimensi yang khas. Provinsi ini terdiri dari 33 kabupaten/kota dengan karakteristik demografis, geografis, dan sosial-ekonomi yang sangat heterogen — mulai dari wilayah perkotaan padat seperti Kota Medan hingga kabupaten kepulauan seperti Nias Barat yang memiliki keterbatasan aksesibilitas. Keragaman ini menciptakan variasi

kebutuhan Alokon yang signifikan antarwilayah. Secara operasional, mekanisme distribusi Alokon dilaksanakan sebagai tindak lanjut dari surat permintaan yang diajukan oleh masing-masing kabupaten/kota kepada BKKBN Perwakilan Provinsi. BKKBN kemudian memproses seluruh permintaan tersebut dan mendistribusikan Alokon sesuai dengan ketersediaan stok di gudang, yang umumnya berada di bawah angka permintaan karena harus mempertimbangkan kecukupan untuk seluruh 33 kabupaten/kota secara bersamaan. Kesenjangan antara kapasitas stok gudang dan akumulasi permintaan wilayah berimplikasi pada realisasi distribusi yang tidak proporsional terhadap jumlah PUS. Fenomena ini menciptakan variasi tingkat kecukupan layanan yang signifikan antarwilayah.

Pendekatan konvensional dalam mengevaluasi distribusi Alokon umumnya masih menggunakan nilai absolut jumlah Alokon yang didistribusikan tanpa mempertimbangkan kebutuhan relatif tiap wilayah. Hal ini dapat menimbulkan bias dalam interpretasi karena wilayah dengan jumlah penduduk yang besar secara alami akan tercatat menerima distribusi lebih tinggi, meskipun belum tentu proporsional terhadap kebutuhan PUS-nya. Untuk mengatasi keterbatasan tersebut, penelitian ini mengusulkan penggunaan rasio distribusi Alokon terhadap PUS sebagai satuan analisis utama. Rasio ini menghasilkan ukuran yang telah dinormalisasi sehingga mampu mencerminkan tingkat kecukupan

distribusi secara proporsional terhadap populasi sasaran di setiap wilayah, serta memungkinkan perbandingan distribusi yang lebih objektif dan bebas bias skala antarwilayah.

Meskipun studi-studi tersebut menunjukkan relevansi dan potensi *clustering* dalam analisis layanan kesehatan, terdapat beberapa celah yang belum terjawab. Pertama, penelitian sebelumnya umumnya menggunakan nilai absolut sebagai variabel *clustering*, bukan rasio proporsional terhadap populasi sasaran. Kedua, hasil komparasi antara K-Means dan K-Medoids tidak konsisten di berbagai konteks — terdapat studi yang menunjukkan keunggulan K-Means, namun ada pula yang menunjukkan sebaliknya — sehingga evaluasi empiris pada konteks dan karakteristik data yang spesifik tetap diperlukan. Ketiga, penerapan analisis komparatif kedua algoritma secara khusus pada data distribusi Alokon dalam konteks program KB, khususnya di Provinsi Sumatera Utara, belum pernah dilakukan sebelumnya.

Berdasarkan kesenjangan penelitian tersebut, penelitian ini mengusulkan analisis komparatif antara algoritma K-Means dan K-Medoids dalam melakukan *clustering* terhadap rasio distribusi Alokon terhadap PUS pada 33 kabupaten/kota di Provinsi Sumatera Utara tahun 2025. Data yang digunakan bersumber langsung dari sistem pencatatan distribusi BKKBN Perwakilan Provinsi Sumatera Utara, sehingga memiliki tingkat keabsahan dan aktualitas yang tinggi. Penelitian ini berfokus pada analisis perbandingan performa algoritma *clustering* dan pengelompokan wilayah berbasis data, bukan pada perumusan kebijakan distribusi. Tujuan utama penelitian ini adalah: (1) mengidentifikasi pola pengelompokan wilayah berdasarkan tingkat kecukupan rasio distribusi Alokon terhadap PUS, dan (2) mengevaluasi serta membandingkan kualitas *clustering* yang dihasilkan oleh algoritma K-Means dan K-Medoids menggunakan metrik *Silhouette Score*.

## METODE

### Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang bersumber langsung dari sistem pencatatan distribusi BKKBN Perwakilan Provinsi Sumatera Utara tahun 2025 dan Sistem Informasi Keluarga. Penggunaan data institusional ini menjamin keabsahan dan aktualitas informasi yang digunakan dalam model.

**Tabel 1 Spesifikasi Data Penelitian**

Atribut	Keterangan
Sumber data	BKKBN Perwakilan Provinsi Sumatera Utara
Periode	Tahun 2025
Unit analisis	Kabupaten/kota
Jumlah observasi	33 kabupaten/kota
Cakupan jenis Alokon	Pil, Suntik, Implan, IUD, Kondom, MOW
Variabel populasi sasaran	Jumlah PUS per kabupaten/kota
Tools prapemrosesan	Microsoft Excel
Tools analisis	Python 3.10

### Rekayasa Fitur : Perhitungan Rasio Distribusi Alokon terhadap PUS

Sebagaimana diuraikan pada BAB I dan BAB II, penelitian ini menggunakan rasio distribusi Alokon terhadap PUS sebagai variabel analisis utama untuk menghindari bias interpretasi yang ditimbulkan oleh perbedaan skala populasi sasaran yang sangat heterogen antarwilayah di Provinsi Sumatera Utara [6] [7]. Rasio  $R_i$  untuk setiap kabupaten/kota dihitung menggunakan Persamaan (1).

$$R_i = \frac{D_i}{PUS_i} \quad (1)$$

di mana  $R_i$  adalah rasio distribusi Alokon di kabupaten/kota  $i$ ,  $D_i$  adalah jumlah total Alokon yang didistribusikan ke kabupaten/kota  $i$ , dan  $PUS_i$  adalah jumlah Pasangan Usia Subur di kabupaten/kota  $i$ .

Nilai  $R_i$  yang lebih tinggi mengindikasikan tingkat kecukupan distribusi yang lebih tinggi relatif terhadap kebutuhan populasi sasaran di wilayah tersebut, dan sebaliknya. Variabel  $R_i$  merupakan satu-satunya fitur yang digunakan sebagai input dalam seluruh proses *clustering*. Struktur *dataset* setelah tahap ini disajikan pada Tabel 2.

### Normalisasi Data

Guna memastikan stabilitas komputasi dan mencegah dominasi nilai ekstrem dalam perhitungan jarak antarobservasi, dilakukan normalisasi terhadap nilai  $R_i$  menggunakan metode *Min-Max Normalization* [9]. Metode ini membawa seluruh nilai ke rentang [0,1] sesuai Persamaan (2).

$$R'_i = \frac{R_i - R_{\min}}{R_{\max} - R_{\min}} \quad (2)$$

### Penentuan Jumlah Cluster Optimal

Penentuan jumlah *cluster* optimal dilakukan menggunakan dua metode yang diterapkan secara konsisten pada kedua algoritma agar komparasi berjalan pada

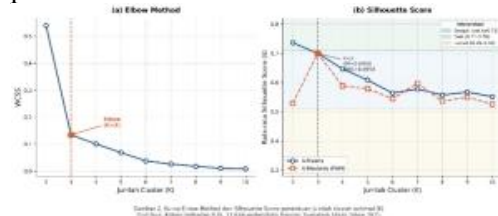
basis yang setara [9] [14]. Nilai yang diuji mencakup rentang  $K = 2$  hingga  $K = 10$ .

*Elbow Method* menghitung nilai *Within-Cluster Sum of Squares* (WCSS) untuk setiap nilai  $K$  menggunakan Persamaan (3). Titik *elbow* — yaitu titik di mana laju penurunan WCSS mulai melambat secara signifikan — diidentifikasi sebagai kandidat jumlah *cluster* optimal.

$$WCSS = \sum_{k=1}^K \sum_{x_i \in C_k} (x_i - \mu_k)^2 \quad (3)$$

di mana  $C_k$  adalah himpunan anggota *cluster* ke- $k$ ,  $x_i$  adalah nilai  $R'_i$  kabupaten/kota yang termasuk dalam *cluster*  $k$ , dan  $\mu_k$  adalah nilai rata-rata  $R'_i$  seluruh anggota *cluster*  $k$ . *Silhouette Score* kemudian digunakan sebagai

validasi silang untuk memastikan nilai  $K$  yang dipilih memiliki rata-rata kualitas pengelompokan tertinggi [14]. Hasil kedua metode divisualisasikan dalam bentuk kurva sebagaimana ditunjukkan pada Gambar 2.



**Gambar 2 Kurva Elbow Method dan Silhouette Score penentuan jumlah cluster optimal (K)**

Hasil *Elbow Method* pada Gambar 2a menunjukkan penurunan WCSS paling tajam terjadi dari  $K = 2$  ke  $K = 3$  sebesar 75,12% (dari 0,5407 menjadi 0,1345), setelah itu laju penurunan melambat secara konsisten. Titik *elbow* teridentifikasi secara objektif pada  $K = 3$ . Hasil *Silhouette Score* pada Gambar 2b menunjukkan nilai  $\bar{s}$  tertinggi K-Means pada  $K = 2$  sebesar 0,7362, sementara K-Medoids mencapai nilai tertinggi pada  $K = 3$  sebesar 0,6998. Pada  $K = 3$ , kedua algoritma menghasilkan nilai  $\bar{s}$  yang identik sebesar 0,6998, keduanya tergolong kategori *reasonable structure* [14].

Berdasarkan kombinasi kedua metode tersebut, penelitian ini menetapkan  $K = 3$  sebagai jumlah *cluster* optimal dengan tiga pertimbangan. Pertama, nilai  $\bar{s} = 0,6998$  pada  $K = 3$  masih tergolong struktur baik untuk kedua algoritma secara bersamaan. Kedua, *Elbow Method* menunjukkan titik siku yang jelas pada  $K = 3$  dengan penurunan WCSS terbesar. Ketiga, tiga kelompok memberikan interpretasi substantif yang lebih informatif — rendah, sedang, dan tinggi — dibandingkan dua kelompok dalam konteks distribusi Alokon.

### Implementasi Algoritma K-Means

Algoritma K-Means diimplementasikan pada *dataset* ternormalisasi ( $R_i$ ) menggunakan Python 3.10 dengan pustaka *scikit-learn* untuk mengelompokkan 33 kabupaten/kota berdasarkan kedekatan jarak terhadap *centroid* [9] [11]. Inisialisasi menggunakan metode *k-means++* dengan ‘*n\_init=50*’ dan ‘*max\_iter=500*’ untuk meningkatkan stabilitas hasil dan meminimalkan risiko konvergensi ke solusi lokal yang suboptimal.

### Implementasi Algoritma K-Medoids (PAM)

Sebagai pembandingan, diimplementasikan algoritma K-Medoids menggunakan metode *Partitioning Around Medoids* (PAM) yang dikembangkan oleh Kaufman dan Rousseeuw [12] melalui pustaka *scikit-learn-extra*. Berbeda dengan K-Means yang menggunakan nilai rata-rata sebagai *centroid*, K-Medoids menggunakan titik data aktual sebagai medoid sehingga lebih *robust* terhadap *outlier* [12]. Pemilihan metode PAM dinilai memadai mengingat ukuran *dataset* relatif kecil ( $n = 33$ ), sehingga perbedaan efisiensi komputasi antara PAM dan variannya tidak berpengaruh secara praktis. Alur implementasi diilustrasikan pada Gambar 4.

### Evaluasi dan Analisis Komparatif

Evaluasi kualitas *clustering* dilakukan menggunakan *Silhouette Score* sebagai metrik utama sesuai Persamaan (6) [14], konsisten dengan kerangka teoritis yang telah dibangun pada subbab 2.6.

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (6)$$

di mana  $a(i)$  adalah rata-rata jarak intra-*cluster* dan  $b(i)$  adalah rata-rata jarak ke *cluster* terdekat lainnya.

## HASIL DAN PEMBAHASAN

### Hasil Clustering

Implementasi algoritma K-Means dan K-Medoids PAM pada 33 kabupaten/kota di Provinsi Sumatera Utara menggunakan  $K = 3$  menghasilkan struktur pengelompokan  $K$

yang stabil dan konsisten. Nilai telah dijustifikasi pada sub-bab 3.5 melalui kombinasi *Elbow Method* — dengan penurunan WCSS terbesar sebesar 75,12% dari  $K = 2$  ke  $K = 3$  — dan *Silhouette Score* yang menghasilkan  $\bar{s} = 0,6998$  untuk kedua algoritma, tergolong kategori *reasonable structure* menurut kriteria Rousseeuw [14]. Statistik deskriptif data rasio distribusi Alokon terhadap PUS ( $R_i$ ) disajikan pada Tabel 4, sebagai konteks sebelum hasil *clustering* diinterpretasikan.

**Tabel 2 Statistik Deskriptif Rasio Distribusi Alokon terhadap PUS**

Statistik	Nilai
Jumlah observasi	33 kabupaten/kota
Rata-rata ( $\bar{R}$ )	0,2782
Standar deviasi	0,1467
Nilai minimum	0,1002 (Kabupaten Dairi)
Nilai maksimum	0,7008 (Kabupaten Mandailing Natal)
Rentang	0,6006

Rentang rasio yang mencapai 0,6006 dengan standar deviasi 0,1467 mengkonfirmasi heterogenitas distribusi yang signifikan antarwilayah, sehingga menjustifikasi penggunaan teknik *clustering* untuk mengidentifikasi pola pengelompokan secara objektif. Kedua algoritma menghasilkan distribusi *cluster* yang identic: 16 wilayah (*Cluster 1* — rendah), 13 wilayah (*Cluster 2* — sedang), dan 4 wilayah (*Cluster 3* — tinggi). Hasil lengkap disajikan pada Tabel 5. Tabel 5. Hasil *Clustering* K-Means dan K-Medoids PAM ( $K = 3$ ) (Lampiran 2)

### Karakteristik Cluster

Ringkasan statistik per*cluster* disajikan pada Tabel 3

### Cluster 1 — Distribusi Rendah ( $\bar{R} = 0,1687$ ; $n = 16$ ; 48,5%)

Cluster ini merupakan kelompok terbesar dengan nilai rasio pada 60,6% dari rata-rata provinsi. Nilai  $\bar{s} = 0,7182$  — tertinggi di antara tiga cluster — menunjukkan kohesi internal yang kuat. Yang menarik, cluster ini tidak hanya mencakup wilayah dengan keterbatasan akses seperti Dairi ( $R = 0,1002$ , terendah seluruh dataset), Samosir, dan seluruh kepulauan Nias, tetapi juga wilayah perkotaan berpenduduk besar seperti Kota Medan dan Deli Serdang. Meskipun keduanya menerima distribusi Alokon absolut tertinggi di provinsi (masing-masing 37.780 dan 58.620 unit), rasio proporsional terhadap PUS-nya tetap rendah karena jumlah PUS yang sangat besar (194.804 dan 265.491). Temuan ini menegaskan bahwa pendekatan berbasis rasio mampu mengungkap ketimpangan yang tidak terlihat pada analisis berbasis nilai absolut [6] [7]. Deli Serdang mencatat *Silhouette Score* terendah dalam seluruh dataset (0,3888), mengindikasikan posisinya di perbatasan antara Cluster 1 dan Cluster 2.

### Cluster 2 — Distribusi Sedang ( $\bar{R} = 0,3117$ ; $n = 13$ ; 39,4%)

Cluster ini mencerminkan kondisi distribusi yang relatif sebanding dengan rata-rata provinsi (112,0% dari  $\bar{R}_{prov}$ ). Nilai  $\bar{s} = 0,6821$  menunjukkan struktur cluster yang baik. Cluster ini didominasi wilayah pantai timur dan Tapanuli, termasuk kelompok Labuhanbatu, Asahan, dan beberapa kota madya. Nias ( $R = 0,2588$ ) memiliki *Silhouette Score* terendah dalam cluster ini (0,3645), mengindikasikan posisinya di perbatasan dengan Cluster 1.

### Cluster 3 — Distribusi Tinggi ( $\bar{R} = 0,6077$ ; $n = 4$ ; 12,1%)

Cluster ini merupakan kelompok paling kecil namun paling distingtif, dengan rasio mencapai 218,4% dari rata-rata provinsi. Keempat anggotanya —

Mandailing Natal, Kota Padang Sidempuan, Humbang Hasundutan, dan Toba — memiliki karakteristik umum berupa jumlah PUS yang relatif kecil (20.575–54.936) namun realisasi distribusi yang tinggi secara proporsional, menghasilkan rasio yang jauh di atas rata-rata provinsi.

### Analisis Komparatif Algoritma

Perbandingan performa kedua algoritma pada  $K = 3$  disajikan pada Tabel 7. Tabel 7. Perbandingan Performa K-Means dan K-Medoids PAM ( $K = 3$ ) (Lampiran 4). Kedua algoritma menghasilkan nilai  $\bar{s}$  yang identik (0,6998) dengan komposisi cluster yang sepenuhnya sama untuk seluruh 33 kabupaten/kota. Secara teoretis, K-Medoids lebih *robust* terhadap *outlier* [12], sedangkan K-Means sensitif terhadap pemilihan centroid awal [9]. Namun, penggunaan k-means++ dengan ‘n\_init=50’ pada penelitian ini terbukti mampu menghasilkan stabilitas yang sebanding dengan K-Medoids PAM. Tidak adanya perbedaan hasil mengindikasikan bahwa data rasio distribusi yang telah dinormalisasi memiliki struktur yang relatif kompak tanpa pengaruh *outlier* ekstrem yang cukup kuat untuk menggeser centroid K-Means secara signifikan. Dengan demikian, perbedaan konseptual antara centroid dan medoid tidak menghasilkan perbedaan praktis pada kasus ini, yang justru memperkuat validitas struktur cluster yang terbentuk — pemartisian yang sama dicapai oleh dua pendekatan yang secara algoritmik berbeda.

### Pembahasan

Penelitian ini menghasilkan dua temuan utama yang saling berkaitan dan menjawab seluruh tujuan yang ditetapkan pada BAB 1.

Temuan pertama adalah bahwa distribusi Alokon terhadap PUS di Provinsi Sumatera Utara membentuk tiga pola yang terdefinisi dengan baik. Pola ini tidak dapat diidentifikasi secara akurat

melalui pendekatan konvensional berbasis nilai absolut, mengingat wilayah dengan distribusi absolut tertinggi justru masuk dalam kelompok distribusi rendah ketika dianalisis secara proporsional. Temuan ini konsisten dengan argumentasi pada BAB 1 dan BAB 2 [6] [7] bahwa pendekatan berbasis rasio lebih relevan untuk analisis distribusi berbasis kebutuhan populasi sasaran yang heterogen. Lebih lanjut, temuan bahwa hampir separuh wilayah Sumatera Utara (48,5%) berada dalam kelompok distribusi rendah konsisten dengan mekanisme distribusi BKKBN yang diuraikan pada BAB 1, di mana realisasi distribusi umumnya berada di bawah angka permintaan akibat keterbatasan stok gudang terhadap akumulasi permintaan 33 wilayah.

Temuan kedua adalah tidak adanya perbedaan performa antara K-Means dan K-Medoids PAM dalam konteks dataset ini. Temuan ini memberikan kontribusi empiris terhadap literatur komparatif yang sebelumnya menunjukkan hasil inkonsisten: Fira et al. [1] menemukan K-Medoids lebih unggul pada data COVID-19 tingkat provinsi, sementara Syamfithriani et al. [2] dan Nirwana et al. [3] menemukan K-Means lebih unggul pada konteks berbeda. Penelitian ini menunjukkan bahwa pada data yang telah dinormalisasi dan tidak mengandung *outlier* ekstrem kondisi yang berbeda dari studi-studi sebelumnya kedua algoritma dapat menghasilkan solusi yang identik. Hal ini selaras dengan Permatasari et al. [5] yang menemukan performa sebanding pada data kesehatan dengan karakteristik serupa. Nilai  $\bar{s} = 0,6998$  menunjukkan bahwa struktur *cluster* valid secara statistik [14] meskipun belum mencapai kategori *strong structure* — yang merupakan karakteristik alami data distribusi kontinu tanpa batas kelompok yang tajam.

## SIMPULAN

### Ringkasan Temuan Penelitian

Penelitian ini telah berhasil memenuhi kedua tujuan yang ditetapkan pada BAB 1: mengidentifikasi pola pengelompokan 33 kabupaten/kota di Provinsi Sumatera Utara berdasarkan rasio distribusi Alokon terhadap PUS, serta mengevaluasi dan membandingkan performa algoritma K-Means dan K-Medoids PAM dalam membentuk pengelompokan tersebut. Melalui penetapan  $K = 3$  yang dijustifikasi secara empiris menggunakan kombinasi *Elbow Method* dan *Silhouette Score*, ketiga kelompok yang terbentuk mencerminkan spektrum tingkat kecukupan distribusi: rendah (16 wilayah; 48,5%), sedang (13 wilayah; 39,4%), dan tinggi (4 wilayah; 12,1%).

### Kesimpulan dan Pendapat Argumentatif

Pertama, pendekatan berbasis rasio distribusi terhadap PUS menunjukkan kecenderungan menghasilkan representasi yang lebih informatif dibandingkan pendekatan berbasis nilai absolut dalam konteks heterogenitas populasi antarwilayah. Fakta bahwa Kota Medan dan Deli Serdang dua wilayah dengan volume distribusi Alokon absolut tertinggi di provinsi justru masuk dalam kelompok distribusi rendah secara proporsional merupakan bukti konkret bahwa evaluasi distribusi berbasis nilai absolut berpotensi menghasilkan bias interpretasi yang dapat memengaruhi pengambilan keputusan. Peneliti berpendapat bahwa temuan ini memiliki implikasi yang lebih luas: di wilayah mana pun dengan disparitas jumlah PUS yang besar antarunit administratif, normalisasi berbasis rasio populasi sasaran seharusnya menjadi standar minimal dalam evaluasi distribusi layanan kesehatan reproduksi, bukan sekadar pilihan metodologis.

Kedua, kesetaraan performa K-Means dan K-Medoids PAM pada dataset ini alih-alih dipandang sebagai hasil yang tidak informatif justru merupakan temuan yang paling signifikan dari perspektif ilmiah. Inkonsistensi hasil komparatif

antara kedua algoritma dalam berbagai studi sebelumnya [1] [2] [3] selama ini menimbulkan pertanyaan terbuka: dalam kondisi data seperti apa salah satu algoritma akan unggul? Penelitian ini memberikan indikasi empiris bahwa pada data distribusi layanan kesehatan yang telah dinormalisasi dan tidak mengandung *outlier* ekstrem, struktur *cluster* cenderung cukup stabil sehingga pilihan antara K-Means dan K-Medoids tidak menghasilkan perbedaan pengelompokan yang berarti. Peneliti berpendapat bahwa kondisi ini — bukan berarti kedua algoritma selalu ekuivalen — melainkan menunjukkan bahwa karakteristik data, khususnya keberadaan *outlier* dan distribusi fitur, merupakan faktor penentu yang lebih kritis daripada pilihan algoritma itu sendiri dalam konteks *clustering* data distribusi layanan kesehatan.

#### DAFTAR PUSTAKA

- [1] A. Fira, C. Rozikin, and Garno, “Komparasi algoritma K-Means dan K-Medoids untuk pengelompokan penyebaran COVID-19 di Indonesia,” *Journal of Applied Informatics and Computing (JAIC)*, vol. 5, no. 2, pp. 133–138, 2021, doi: 10.30871/jaic.v5i2.3286.
- [2] T. S. Syamfithriani, N. Mirantika, and R. Trisudarmo, “Perbandingan algoritma K-Means dan K-Medoids untuk pemetaan daerah penanganan diare pada balita di Kabupaten Kuningan,” *Jurnal Sistem Informasi Bisnis*, vol. 12, no. 2, pp. 132–139, 2023, doi: 10.21456/vol12iss2pp132-139.
- [3] S. D. Nirwana, M. I. Jambak, and A. Bardadi, “Perbandingan algoritma K-Means dan K-Medoids dalam clustering rata-rata penambahan kasus COVID-19 berdasarkan kota/kabupaten di Provinsi Sumatera Selatan,” *Jurnal Sistem Informasi (JSiI)*, vol. 9, no. 2, pp. 126–131, 2022, doi: 10.30656/jsii.v9i2.5127.
- [4] D. K. Sitinjak, B. A. Pangestu, and B. N. Sari, “Clustering jumlah tenaga kesehatan berdasarkan kecamatan di Kabupaten Karawang menggunakan algoritma K-Means,” *Journal of Applied Informatics and Computing (JAIC)*, vol. 6, no. 1, pp. 46–54, 2022, doi: 10.30871/jaic.v6i1.3855.
- [5] R. Permatasari, Sukirman, and F. Elfazza, “Perbandingan optimasi penyuluhan penyakit stunting pada balita integrasi algoritma K-Means dan Partitioning Around Medoids (PAM),” *JIKTIF: Jurnal Ilmu Komputer dan Teknologi Informasi*, vol. 1, no. 2, pp. 90–99, 2024, doi: 10.71466/jiktif.v1i2.43.
- [6] G. J. Oyewole and G. A. Thopil, “Data clustering: application and trends,” *Artif. Intell. Rev.*, vol. 56, no. 7, pp. 6439–6475, 2023, doi: 10.1007/s10462-022-10325-y.
- [7] A. E. Ezugwu *et al.*, “A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects,” *Eng. Appl. Artif. Intell.*, vol. 110, p. 104743, 2022, doi: 10.1016/j.engappai.2022.104743.
- [8] M. Chaudhry, I. Shafi, M. Mahnoor, D. L. R. Vargas, E. B. Thompson, and A. Imran, “A Systematic Literature Review on Identifying Patterns Using Unsupervised Clustering Algorithms: A Data Mining Perspective,” *Symmetry (Basel)*, vol. 15, no. 1679, pp. 1–44, 2023, doi: 10.3390/sym15091679.
- [9] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and H. Jia, “K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data,” *Inf. Sci. (N. Y.)*, vol. 622, pp. 178–210, 2023, doi: 10.1016/j.ins.2022.11.139.
- [10] A. A. Wani, “Comprehensive analysis of clustering algorithms: exploring limitations and innovative solutions,” *PeerJ Comput. Sci.*, vol.

- 10, p. e2286, 2024, doi: 10.7717/peerj-cs.2286.
- [11] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth berkeley symposium on mathematical statistics and probability, volume 1*, L. M. L. Cam and J. Neyman, Eds., University of California Press, 1967, pp. 281–297. Accessed: Apr. 15, 2026. [Online]. Available: <https://projecteuclid.org/euclid.bsmsp/1200512992>
- [12] L. Kaufman and P. J. Rousseeuw, "Clustering by means of medoids," in *Statistical Data Analysis Based on the L1 Norm and Related Methods*, Y. Dodge, Ed., Amsterdam: North-Holland; Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co., 1987, pp. 405–416.
- [13] E. Schubert and P. J. Rousseeuw, "Fast and eager k-medoids clustering: O(k) runtime improvement of the PAM, CLARA, and CLARANS algorithms," *Inf. Syst.*, vol. 101, no. 101804, 2021,
- [14] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, pp. 53–65, 1987, doi: 10.1016/0377-0427(87)90125-7.