

# Klasterisasi Peserta KB Aktif di Desa Kalirejo Lawang Menggunakan Metode *K-Means*

Afifah Vera Ferencia Fitria Ningrum<sup>1\*</sup>, Mochammad Anshori<sup>2</sup>, Risqy Siwi Pradini<sup>3</sup>

<sup>1\*,2,3</sup> Program Studi Informatika, Sains dan Teknologi, Institut Teknologi, Sains, dan Kesehatan RS. Dr. Soepraoen Kesdam V/BRW, Kota Malang, Provinsi Jawa Timur, Indonesia.

*Email:* afifahveraferencia@gmail.com<sup>1\*</sup>, moanshori@itsk-soepraoen.ac.id<sup>2</sup>, risqypradini@itsk-soepraoen.ac.id<sup>3</sup>

## Histori Artikel:

*Dikirim* 15 Desember 2024; *Diterima dalam bentuk revisi* 25 Desember 2024; *Diterima* 5 Januari 2025; *Diterbitkan* 10 Januari 2025. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

## Abstrak

Program keluarga berencana (KB) di Desa Kalirejo Lawang menghadapi tantangan dalam proses pengklasterisasian peserta KB aktif di desa tersebut. Dalam pengklasterisasian itu memakan waktu dan rentan terhadap kesalahan. Dengan dasar tersebut diusulkan solusi klasterisasi dengan menggunakan algoritma K-Means. Eksperimen yang dilakukan dengan menguji jumlah klaster 2 hingga 8 dan di evaluasi menggunakan silhouette score. Hasil penelitian menunjukkan bahwa jumlah klaster optimal adalah dua, dibuktikan dengan nilai silhouette score sebesar 0.447. Nilai tersebut menunjukkan kualitas klasterisasi terbaik dibandingkan dengan jumlah klaster lainnya, dimana skor untuk klaster 3 hingga 8 tidak ada yang melebihi nilai tersebut. Hal ini membuktikan bahwa klasterisasi dengan dua klaster memberikan hasil yang paling optimal. Skor yang lebih rendah pada klaster 3 hingga 8 mengindikasikan bahwa pembagian klaster menjadi lebih banyak tidak menciptakan pemisahan yang jelas atau memperburuk keseragaman dalam klaster. Kesimpulan dari penelitian ini menunjukkan bahwa metode K-Means dapat diterapkan dan handal untuk di aplikasikan sebagai metode dalam klasterisasi peserta KB aktif bagi desa Kalirejo, Lawang. Dengan kecepatan dan keakuratan K-means memberikan solusi signifikan untuk meningkatkan efisiensi program KB di tingkat desa. Implikasi praktis dari penelitian ini adalah memberikan dasar yang lebih terstruktur untuk perencanaan dan pengambilan keputusan dalam program KB di tingkat desa. Temuan ini menjadi langkah penting dalam optimalisasi pengelolaan data program KB, membuka peluang untuk implementasi lebih luas di wilayah lainnya.

**Kata Kunci:** Keluarga Berencana(KB); Klasterisasi; K-Means.

## Abstract

The Family Planning (KB) program in Kalirejo Lawang Village faces challenges in the process of clustering active participants, which is time-consuming and prone to errors. Based on these challenges, a clustering solution using the K-Means algorithm was proposed. Experiments were conducted by testing the number of clusters from 2 to 8 and evaluating them using the silhouette score. The results of the study showed that the optimal number of clusters is two, as indicated by a silhouette score of 0.447. This value represents the best clustering quality compared to other cluster numbers, where the scores for clusters 3 to 8 did not exceed this value. This demonstrates that clustering into two groups provides the most optimal results. Lower scores for clusters 3 to 8 indicate that dividing the clusters into more groups did not create clear separations or worsened the cohesion within clusters. The conclusion of this study shows that the K-Means method can be applied and is reliable for clustering active KB participants in Kalirejo Lawang Village. With its speed and accuracy, K-Means offers a significant solution to improving the efficiency of the KB program at the village level. The practical implication of this research is to provide a more structured basis for planning and decision-making in the KB program at the village level. These findings mark an important step in optimizing the management of KB program data, opening opportunities for broader implementation in other areas.

**Keyword:** Family Planning (KB); Clustering; K-Means.

## 1. Pendahuluan

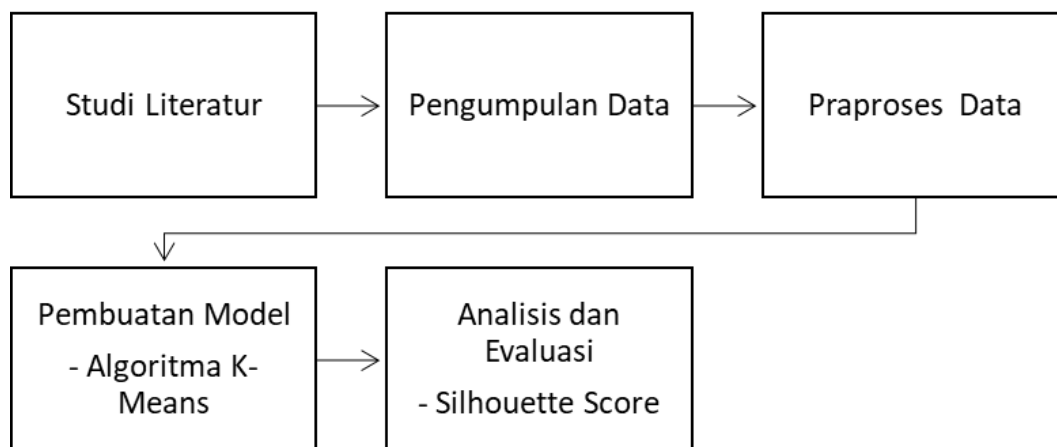
Program keluarga berencana (KB) merupakan inisiatif pemerintah untuk membantu pasangan suami istri dalam merencanakan kehamilan, dengan mempertimbangkan jarak dan kesiapan, serta menentukan jumlah anak yang diinginkan (Swastika *et al.*, 2024). Selain itu, program ini bertujuan untuk meningkatkan kesejahteraan dengan membatasi jumlah anak dalam keluarga (Bella Dwi Saputri & Diah Indriani, 2022). Partisipasi masyarakat dalam program KB diharapkan dapat menurunkan angka kelahiran di wilayah tertentu, yang pada gilirannya berkontribusi pada kesejahteraan penduduk secara keseluruhan. Melalui penyediaan layanan kesehatan reproduksi yang berkualitas, program KB diharapkan dapat meningkatkan kualitas hidup masyarakat (Tiffani *et al.*, 2020). Upaya pengendalian angka kelahiran melalui program KB tidak hanya bertujuan untuk meningkatkan kesejahteraan keluarga secara individu, tetapi juga untuk mengelola jumlah penduduk guna mencegah masalah sosial dan ekonomi yang timbul akibat pertumbuhan penduduk yang tidak terkendali. Pertumbuhan penduduk yang tidak terkendali dapat memicu berbagai masalah sosial dan ekonomi, seperti keterbatasan lapangan kerja, ketimpangan kesejahteraan, dan peningkatan angka kriminalitas (Maharani & Yotenka, 2024). Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN) mengelola program KB dengan tujuan mengendalikan laju pertumbuhan penduduk dan menurunkan angka kelahiran. Keberhasilan program KB diharapkan dapat menghasilkan pertumbuhan penduduk yang lebih seimbang serta meningkatkan kualitas hidup masyarakat (Lino *et al.*, 2021). Penelitian ini berfokus pada Desa Kalirejo Lawang sebagai lokasi untuk melakukan pengkasterisasian peserta KB aktif. Pemilihan desa ini didasarkan pada kebutuhan untuk meningkatkan efektivitas program KB di wilayah tersebut. Program KB di Desa Kalirejo Lawang bertujuan untuk menurunkan laju pertumbuhan penduduk, mengurangi kepadatan penduduk, dan menekan angka kelahiran yang masih tinggi. Namun, tingkat penggunaan alat kontrasepsi di setiap RW masih rendah, yang mengakibatkan pencapaian target pengendalian penduduk belum optimal. Perbedaan tingkat partisipasi KB antar RW menciptakan ketidakseimbangan dalam pencapaian target pengendalian penduduk. Saat ini, petugas BKKBN di Desa Kalirejo Lawang masih mengelompokkan peserta KB secara manual berdasarkan jenis alat kontrasepsi yang digunakan, yang mengarah pada proses yang kurang efisien dan rentan terhadap kesalahan.

Kemajuan teknologi menawarkan solusi melalui penggunaan metode *klasterisasi* untuk mengelompokkan peserta KB aktif secara lebih efisien. *Klasterisasi* adalah proses pengelompokan data tanpa label ke dalam beberapa *cluster* berdasarkan kesamaan tertentu (Mauladi & Susilo, 2021; Fitri *et al.*, 2023). Salah satu metode yang banyak digunakan adalah algoritma *K-Means*, yang telah diaplikasikan dalam berbagai bidang, seperti pendidikan, kesehatan, sosial, dan ilmu komputer (Pipit Mulyah *et al.*, 2020). Algoritma *K-Means* unggul karena sifatnya yang linear, sehingga dapat mencapai akurasi tinggi, serta cepat dan fleksibel dalam penggunaannya (Rahman Sahputra & Kurniawan, 2024). Beberapa penelitian terdahulu menunjukkan keberhasilan *K-Means* sebagai metode *klasterisasi*, dibandingkan dengan *Agglomerative Nesting* (Kasus & Indonesia, 2023) dan *K-Medoids* (Nafilah *et al.*, 2024). Dalam bidang kesehatan, *K-Means* telah berhasil diterapkan untuk *klasterisasi* layanan kesehatan (Nur, 2024), deteksi penyakit ginjal (Nithya *et al.*, 2020), kesehatan mental (Gao *et al.*, 2023), dan pencarian pola dalam data kesehatan (Haraty *et al.*, 2015). Beberapa penelitian sebelumnya telah menggunakan *K-Means* untuk *klasterisasi* dalam bidang kesehatan, dengan fokus pada pengelompokan peserta aktif program KB. Penelitian yang dilakukan oleh Swastika *et al.* (2024) menggunakan metode *K-Means* untuk data yang bersumber dari Dinas P2KBP3A Asahan tahun 2021. Penelitian lain oleh Sitepu *et al.* (2022) juga menggunakan metode yang sama, dengan data yang bersumber dari peserta KB Kota Binjai. Selain itu, Amelia *et al.* (2022) melakukan *klasterisasi* data yang bersumber dari website resmi BPS Kabupaten Cirebon. Ketiga penelitian ini menggunakan metode *K-Means* untuk *klasterisasi* peserta KB, namun terdapat perbedaan signifikan di antara ketiganya. Penelitian pertama hanya berfokus pada proses *klasterisasi* menggunakan data Dinas P2KBP3A Asahan. Penelitian kedua juga mencakup visualisasi data dan pembuatan *antarmuka pengguna* untuk menampilkan hasil analisis.

Penelitian ketiga hanya berfokus pada *klasterisasi*. Ketiga penelitian ini belum diterapkan pada tingkat desa seperti Desa Kalirejo Lawang. Penelitian ini bertujuan untuk mengisi kekosongan tersebut dengan menggunakan metode *K-Means* untuk pengelolaan data peserta KB aktif di tingkat desa, khususnya Desa Kalirejo Lawang. Penelitian ini bertujuan untuk mengklasterisasi data peserta KB di Desa Kalirejo Lawang menggunakan metode *K-Means*. Mengingat banyaknya data terkait penggunaan alat kontrasepsi di desa ini, pengklasteran ini diharapkan dapat mempermudah petugas BKKBN dalam mengidentifikasi peserta KB aktif dengan akurat. Dengan menerapkan metode *klasterisasi K-Means*, penelitian ini diharapkan dapat mendukung pengambilan keputusan dalam pengklasteran peserta KB aktif yang lebih tepat dalam pelaksanaan program KB di Desa Kalirejo Lawang.

## 2. Metode Penelitian

Metode yang digunakan pendekatan kuantitatif dengan metode klasterisasi untuk menganalisis data peserta KB di Desa Kalirejo Lawang. Pendekatan ini dipilih untuk memastikan pengelompokan peserta berdasarkan kesamaan karakteristik yang relevan, sehingga memudahkan identifikasi dan pengelolaan data peserta KB secara efisien. Kerangka berpikir pada penelitian ini ditunjukkan oleh gambar 1 berikut:



Gambar 1. Tahapan Penelitian

Kerangka berpikir penelitian ini dirancang untuk menyelesaikan permasalahan yang ada. Secara umum, terdapat lima tahap yang dilakukan dalam penelitian ini, yaitu studi literatur, pengumpulan data, praproses data, pembuatan model, serta analisis dan evaluasi. Berikut adalah uraian lebih lanjut mengenai tahap-tahap tersebut:

### 1) Studi Literatur

Langkah pertama dalam penelitian ini adalah melakukan studi literatur, yaitu meninjau penelitian-penelitian sebelumnya yang berkaitan dengan program KB serta penerapan metode *klasterisasi K-Means*. Tujuan dari studi literatur ini adalah untuk memperoleh pemahaman terkait hasil-hasil, kelemahan, dan parameter yang digunakan dalam penelitian-penelitian sebelumnya mengenai pengelompokan peserta KB aktif. Dengan memahami penelitian terdahulu, peneliti dapat memperoleh panduan mengenai penerapan metode *K-Means* dalam pengelompokan peserta KB aktif. Studi literatur ini juga bertujuan untuk memperdalam pemahaman mengenai mekanisme kerja *K-Means*, yang merupakan persiapan penting untuk penerapan metode ini di Desa Kalirejo Lawang. Diharapkan, penelitian ini dapat memaksimalkan hasil dengan memanfaatkan kelebihan dari penelitian sebelumnya, serta mengatasi kekurangannya (Nurahman & Susanto, 2023).

2) Pengumpulan Data

Tahap kedua dalam penelitian ini adalah pengumpulan data. Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari BKKBN di Desa Kalirejo Lawang. Data yang dikumpulkan mencakup periode waktu dari bulan Januari hingga September, dan disajikan secara rinci dalam tabel 1.

Tabel 1. Data Peserta KB Aktif Desa Kalirejo Lawang

No.	Tanggal Ikut KB	NIK Istri	Nama Umur				Jumlah Anak		Umur Anak Terakhir	RT/RW	Tempat Pelayanan	Jenis Kontrasepsi
			Istri	Suami	Istri	Suami	L	P				
1	06/02/2023	54378179	maya	agus	37	38	1	1	2.5 tahun	01/01	Bidan Titik	Suntik
2	02/03/2023	35678	wati	arif	30	40	1	2	18 bulan	01/01	RS Siti Mariam	MOW
3	01/01/2023	35821	fatim	budi	45	47	0	2	16 tahun	01/01	Bidan Titik	Suntik
...	...	...	...	...	...	...	...	...	...	...	...	...
1203	26/06/2024	35128	laila	gunawan	22	24	0	1	12 minggu	03/16		pil
1204	20/02/2023	35497	ani	nanang	46	48	0	2	13 tahun	03/16	puskesmas Lawang	Implant
1205	21/09/2021	35217	fitri	nuril	36	40	1	1	9 tahun	03/16		Pil

Berdasarkan tabel 1, data yang digunakan dalam penelitian ini berisi informasi mengenai peserta KB di Desa Kalirejo Lawang. Data tersebut mencakup tanggal keikutsertaan dalam program KB, Nomor Induk Keluarga (NIK) istri, nama istri dan suami, umur istri dan suami, jumlah anak, umur anak terakhir, RT/RW, tempat pelayanan, dan jenis kontrasepsi. Tabel 1 hanya menunjukkan sebagian kecil dari dataset yang digunakan, dengan beberapa baris data yang memberikan gambaran mengenai struktur dan atribut yang ada. Pada atribut NIK serta nama istri dan suami, data tersebut disamarkan untuk menjaga kerahasiaan, karena merupakan data pribadi yang bersifat sensitif dan tidak dapat dipublikasikan. Dataset ini digunakan sebagai dasar untuk pembuatan model *klasterisasi* dalam rangka mengelompokkan peserta KB aktif di Desa Kalirejo Lawang.

3) Praproses Data

Tahap ketiga dalam penelitian ini adalah praproses data, yang merupakan proses pembersihan data guna memastikan kualitas data yang tinggi dan mendukung hasil penelitian yang lebih akurat dan valid (Madani *et al.*, 2024). Pada tahap ini, informasi pribadi seperti nama dan NIK dihapus untuk menjaga kerahasiaan data. Selain itu, atribut yang dianggap tidak relevan, seperti tanggal keikutsertaan KB dan tempat layanan, dihilangkan karena tidak memberikan kontribusi langsung terhadap analisis *klasterisasi*. Langkah ini bertujuan untuk mengurangi kompleksitas data sehingga algoritma dapat memproses informasi yang relevan dengan lebih efisien dan akurat.

4) Pembuatan Model

Langkah berikutnya adalah pembuatan model klasterisasi menggunakan algoritma K-Means. Klasterisasi adalah teknik yang membagi data ke dalam beberapa kelompok berdasarkan karakteristik tertentu. Algoritma K-Means mengelompokkan data ke dalam beberapa klaster di mana data dalam satu klaster memiliki kesamaan (homogenitas) karakteristik. Metode ini dikenal cepat dan mampu menangani data dalam jumlah besar, meskipun kurang akurat dalam menangani noise dan data yang terisolasi (Hutagalung & Sonata, 2021)(Pratistha & Kristianto, 2024) (Hutagalung, 2022). Algoritma K-Means menghitung jarak antar data dengan jarak kuadrat Euclidean. Dengan kata lain jarak antara dua titik data  $x_i, x_j \in x$  (Cordeiro de Amorim & Makarenkov, 2023), ditunjukkan pada persamaan berikut:

$$d(x_i, x_j) = \sum_{v=1}^m (x_{iv} - x_{jv})^2$$

Dimana  $x_{iv}$  dan  $x_{jv}$  masing-masing merupakan nilai dari  $x_i$  dan  $x_j$  yang berkaitan dengan fitur atau variabel ( $v$ ), serta  $m$  adalah jumlah total fitur yang ada. Klasterisasi K-Means yang menghasilkan nilai minimum pada fungsi objektif dianggap sebagai solusi yang paling optimal. Fungsi objektif ini dirumuskan sebagai persamaan berikut ini:

$$W = \sum_{i=1}^k \sum_{x \in S_i} d(x_i, z_l)$$

Dimana  $z_l$  merupakan pusat kluster  $S_l \in S$ , yang dihitung sebagai rata-rata dari setiap komponen semua  $x_i$  yang termasuk  $S_l$ . Dalam penelitian ini, peneliti menggunakan tool Orange untuk pembuatan model. Orange menawarkan banyak keuntungan karena memungkinkan pengguna untuk bereksperimen dengan berbagai alur kerja dan menerapkan berbagai widget yang tersedia hingga mencapai hasil yang diinginkan (Ricardo *et al.*, 2021). Dalam penerapan algoritma K-Means peneliti menetapkan parameter seperti nilai awal centroid yang dipilih secara acak, jumlah iterasi maksimum, dan toleransi untuk konvergensi. Parameter ini digunakan untuk memastikan algoritma dapat menemukan pembagian kluster yang optimal dengan mempertimbangkan distribusi data.

#### 5) Analisis dan Evaluasi

Tahap terakhir adalah analisis dan evaluasi hasil klasterisasi menggunakan *silhouette score*. Fungsi *silhouette score* ini menghitung rata-rata koefisien siluet dengan mempertimbangkan jarak *intra-kluster* serta jarak ke kluster terdekat untuk setiap titik data. Koefisien siluet sampel dihitung dengan rumus (Shahapure & Nicholas, 2020).

$$S = \frac{(b - a)}{\max(a, b)}$$

Dimana  $b$  adalah rata-rata antara sampel dengan semua sampel lain dalam kluster yang sama. Sedangkan  $a$  adalah rata-rata jarak antara sampel dan semua sampel dalam kluster terdekat, serta  $\max(a, b)$  adalah nilai maksimum antara  $a$  dan  $b$ . Analisis penelitian ini menghasilkan skor siluet yang memberikan gambaran tentang pengelompokan data. Skor siluet dengan nilai mendekati +1 menunjukkan bahwa titik data berada dalam tinggi. Sebaliknya, skor mendekati 0 menunjukkan bahwa titik data seharusnya mungkin berada di kluster lain. Sementara itu, skor mendekati -1 menunjukkan bahwa titik data lebih cocok berada di kluster rendah. Setelah mendapatkan analisis tersebut maka akan ditarik kesimpulan dan akan dilakukan evaluasi dari hasil skor *siluet* yang didapatkan.

## 3. Hasil dan Pembahasan

### 3.1 Hasil

Berdasarkan eksperimen yang telah dilakukan, hasil dan pembahasan diperoleh dengan mengamati performa klasterisasi menggunakan metode *K-Means*. Tahap pertama yang dilakukan adalah praproses data yang telah dikumpulkan sebelumnya. Pada tahap ini, dilakukan pembersihan data untuk memastikan kualitas data yang tinggi dan kesesuaian dengan tujuan penelitian. Atribut yang bersifat pribadi dan tidak relevan dihapus untuk mempermudah proses *klasterisasi*. Beberapa atribut yang dihapus antara lain tanggal keikutsertaan dalam program KB, NIK istri, nama istri dan suami, umur istri dan suami, umur anak terakhir, serta tempat pelayanan. Tabel 2 berikut menunjukkan data peserta KB di Desa Kalirejo Lawang setelah dilakukan pembersihan data.

Tabel 2. Data Setelah Dilakukan Praproses

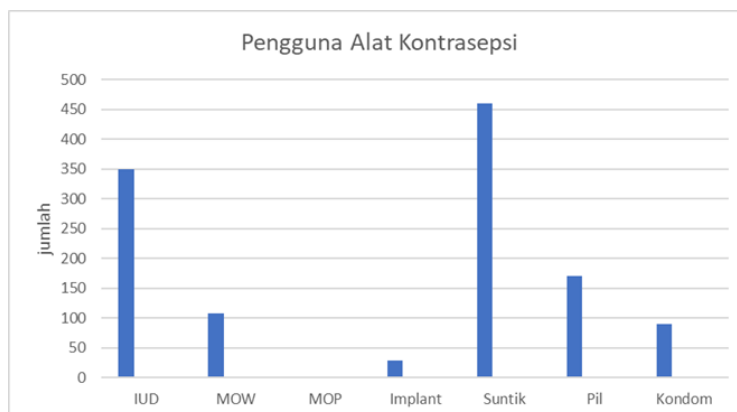
RT	RW	IUD	MOW	MOP	Implant	Suntik	Pil	Kondom
RT 01	RW 01	1	1	0	3	10	0	0
RT 02	RW 01	1	1	0	0	5	1	0
RT 03	RW 01	1	1	0	0	5	0	0
...	...	...	...	...	...	...	...	...
RT 01	RW 16	1	1	0	2	0	2	0
RT 02	RW 16	1	1	0	0	0	4	1
RT 03	RW 16	0	0	0	1	0	2	0

Pada tabel 2, terdapat sembilan atribut dengan jumlah baris data sebanyak 75. Dari sembilan atribut tersebut, meliputi RT (Rukun Tetangga), RW (Rukun Warga), IUD (*Intrauterine Device*), MOW (Medis Operatif Wanita), MOP (Medis Operatif Pria), Implant, Suntik, Pil, dan Kondom. Kesembilan atribut ini dianggap relevan dan akan digunakan dalam penelitian ini.

Tabel 3. Rentang data

Fitur	Data range
RT	1 – 7
RW	1 – 16
IUD	0 – 13
MOW	0 – 5
MOP	0 – 1
Implant	0 – 3
Suntik	0 – 28
PIL	0 – 13
Kondom	0 – 8

Pada tabel 3 menunjukkan nama fitur dan rentang nilainya. Diketahui terdapat 9 fitur, yaitu RT, RW, IUD, MOW, MOP, Implant, Suntik, PIL, dan Kondom. Fitur RT memiliki rentang data 1-7, sedangkan RW memiliki rentang 1-16. Pengguna alat kontrasepsi IUD berkisar antara 0-13, sementara MOW berkisar antara 0-5, dan MOP berkisar antara 0 dan 1. Pengguna alat kontrasepsi implant berkisar antara 0-3, sedangkan pengguna suntik berkisar antara 0-28. Penggunaan pil berkisar antara 0-13, dan penggunaan kondom berkisar antara 0-8. Rentang ini memberikan gambaran tentang penyebaran penggunaan alat kontrasepsi di Desa Kalirejo Lawang, untuk membantu memahami seberapa besar variasi penggunaan di setiap RT/RW.



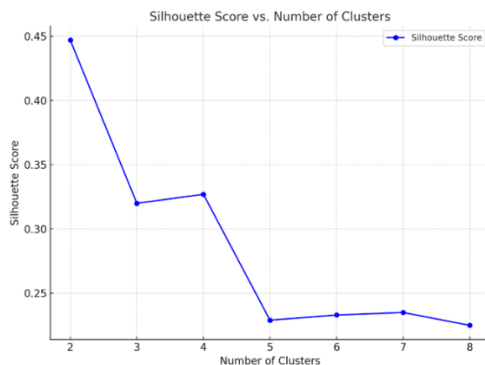
Gambar 2. Pengguna Alat Kontrasepsi Desa Kalirejo Lawang

Berdasarkan gambar 2 menunjukkan jumlah pengguna alat kontrasepsi di Desa Kalirejo Lawang. Terlihat bahwa suntik menjadi jenis alat kontrasepsi yang paling banyak digunakan dengan total pengguna 460 orang. Di urutan kedua, IUD digunakan oleh 349 orang, diikuti oleh pil kontrasepsi dengan jumlah 171 pengguna. Sementara itu, penggunaan kondom sekitar 89 orang. Jenis kontrasepsi lainnya, seperti MOW memiliki 107 pengguna, sedangkan implant 28 orang dan MOP 1 orang. Dari data tersebut menunjukkan bahwa jenis kontrasepsi yang paling dominan adalah suntik dan IUD. Sedangkan alat kontrasepsi yang paling jarang digunakan adalah MOP. Setelah dilakukan praproses data, dilakukan proses pembuatan model K-Means. Penelitian ini menggunakan *tool orange* (Demšar *et al.*, 2013). Skenario eksperimen klusterisasi K-Means dengan menguji nilai k secara bertingkat. Pengujian nilai k dibatasi dengan rentang 2 hingga 8. Diberikan batasan nilai k untuk mengetahui seberapa baik model saat dilakukan perubahan banyaknya kluster. Selain itu, dengan penambahan jumlah kluster, tidak selalu menghasilkan kluster yang optimal. Peneliti melakukan evaluasi berdasarkan nilai *silhouette score* dari tiap uji nilai k. Sehingga didapatkan hasil yang ditunjukkan pada tabel 4 berikut.

Tabel 4. Hasil *Silhouette Score*

Jumlah Kluster	<i>Silhouette Score</i>
2	0.447
3	0.320
4	0.327
5	0.229
6	0.233
7	0.235
8	0.225

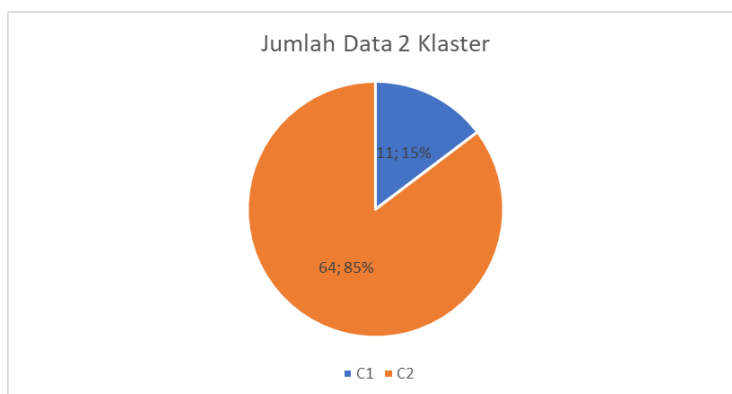
Berdasarkan Tabel 4 di atas, *Silhouette Score* digunakan untuk mengevaluasi kualitas klusterisasi berdasarkan jumlah kluster yang berbeda. Nilai ini mengukur seberapa mirip suatu data dalam kluster tertentu dibandingkan dengan data dalam kluster lainnya. *Silhouette Score* berkisar antara -1 hingga 1, di mana nilai yang lebih tinggi menunjukkan klusterisasi yang lebih baik. Dari tabel tersebut, terlihat bahwa jumlah kluster 2 menghasilkan nilai *Silhouette Score* tertinggi, yaitu 0.447, dibandingkan dengan jumlah kluster lainnya. Hal ini menunjukkan bahwa dengan dua kluster, distribusi data dalam masing-masing kluster lebih baik. Kondisi ini tercermin dari kedekatan antar data dalam kluster dan keterpisahan antar kluster yang lebih optimal dibandingkan dengan jumlah kluster lainnya. Hasil ini memberikan indikasi bahwa dua kluster memberikan kualitas klusterisasi yang lebih baik dibandingkan jumlah kluster lainnya (3, 4, 5, 6, 7, atau 8). Pada jumlah kluster tiga, nilai *Silhouette Score* menurun menjadi 0.320, menunjukkan sedikit penurunan dalam kualitas klusterisasi. Untuk empat kluster, skor meningkat menjadi 0.327, meskipun kualitasnya masih berada di bawah dua kluster. Selanjutnya, dengan lima kluster, skor turun drastis menjadi 0.229, yang menandakan penurunan signifikan dalam kualitas klusterisasi. Skor terendah tercatat pada enam kluster dengan nilai 0.233, diikuti oleh tujuh kluster dengan nilai 0.235, dan delapan kluster dengan nilai 0.225. Hal ini menunjukkan bahwa penambahan jumlah kluster tidak meningkatkan kualitas klusterisasi, seperti yang tercermin pada nilai *Silhouette Score* yang semakin menurun dengan bertambahnya jumlah kluster.



Gambar 3. Grafik Hasil *Silhouette Score*

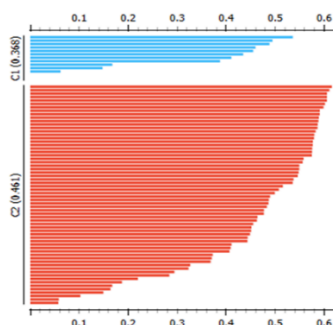
Grafik pada Gambar 3 menunjukkan hubungan antara jumlah kluster dan nilai *Silhouette Score*. Pada sumbu x (horizontal) merepresentasikan jumlah kluster, sementara pada sumbu y (vertikal) merepresentasikan nilai *Silhouette Score*. Berdasarkan grafik, nilai *Silhouette Score* tertinggi tercatat ketika jumlah kluster adalah 2, yang mengindikasikan bahwa pemilihan dua kluster lebih optimal dibandingkan dengan jumlah kluster lainnya. Pada kluster 3, nilai *Silhouette Score* mengalami penurunan dibandingkan dengan kluster 2, yang mengindikasikan distribusi data yang kurang baik. Pada kluster 4, nilai *Silhouette Score* sedikit meningkat dibandingkan dengan kluster 3, namun tetap lebih rendah dari kluster 2. Penurunan signifikan terlihat pada kluster 5, di mana nilai *Silhouette Score* mendekati 0, yang menunjukkan bahwa klusterisasi menjadi sangat tidak optimal. Pada kluster 6, nilai *Silhouette Score* tetap rendah dibandingkan dengan kluster 2. Hal serupa juga terjadi pada kluster 7, di mana nilai *Silhouette Score* masih rendah. Kluster 8 memiliki nilai terendah, yang menunjukkan pembagian data yang kurang optimal, dengan nilai *Silhouette Score* sebesar 0.225, yang mendekati nilai 0. Nilai *Silhouette Score* yang kurang dari 1 dapat dikatakan kurang optimal.

Berdasarkan hasil ini, jumlah kluster 2 adalah yang paling tepat untuk digunakan dalam klusterisasi peserta KB aktif di Desa Kalirejo Lawang. Oleh karena itu, dalam penelitian ini, peneliti akan menggunakan dua kluster untuk mengelompokkan peserta KB aktif di Desa Kalirejo Lawang, yaitu kluster "tinggi" dan "rendah". Implikasi praktis dari distribusi kluster ini adalah bahwa pengelompokan peserta KB berdasarkan tingkat partisipasi atau penggunaan alat kontrasepsi yang lebih tinggi atau lebih rendah dapat membantu merancang kebijakan atau intervensi yang lebih tepat sasaran. Sebagai contoh, kelompok "tinggi" mungkin memerlukan pendekatan yang lebih intensif untuk meningkatkan penggunaan alat kontrasepsi, sementara kelompok "rendah" mungkin membutuhkan edukasi atau fasilitas yang lebih mudah diakses untuk meningkatkan partisipasi dalam program KB. Klusterisasi ini juga dapat digunakan untuk menilai efektivitas program KB yang ada dengan memantau perubahan dalam kluster dari waktu ke waktu.



Gambar 4. Data 2 Kluster

Gambar 4 diatas adalah sebaran klaster dari klasterisasi menggunakan *K-Means*. Diketahui terdapat 2 klaster, yaitu C1 dan C2. Dari 75 data, terdapat 11 data yang dikelompokkan ke dalam klaster C1. Sedangkan untuk klaster C2, terdapat 65 data didalamnya. Hal ini menunjukkan bahwa klasterisasi data didominasi oleh klaster C2 dengan 85% data didalamnya, diikuti klaster C1 15%.



Gambar 5. Hasil *Silhouette Plot*

Gambar 5 menampilkan diagram batang horizontal yang terbagi menjadi dua kelompok utama: kelompok "C1" di bagian atas (berwarna biru) dan "C2" di bagian bawah (berwarna merah). Setiap kelompok memiliki bilah-bilah yang mewakili nilai individual, yang digambarkan pada sumbu horizontal (nilai dari 0 hingga 0,6). Kelompok "C1" diberi label dengan angka 0,368, yang tampaknya menunjukkan rata-rata atau nilai agregat lainnya dari data dalam kelompok tersebut. Sementara itu, kelompok "C2" memiliki label 0,461, yang mungkin juga mewakili metrik agregat serupa. Perbedaan panjang bilah rata-rata menunjukkan bahwa data pada kelompok "C2" secara umum memiliki nilai yang lebih besar dibandingkan dengan "C1". Distribusi data di kedua kelompok cukup teratur, tanpa fluktuasi ekstrem, yang menunjukkan keteraturan data. Sumbu horizontal menunjukkan nilai metrik tertentu, sementara sumbu vertikal mengindikasikan urutan atau kategori yang terkait dengan bilah tersebut. Perbedaan warna mempertegas pembagian antar kelompok. Analisis semacam ini relevan untuk mengelompokkan data atau menampilkan perbandingan metrik antara dua kelompok yang berbeda.

### 3.2 Pembahasan

Hasil klasterisasi menggunakan metode *K-Means* pada data peserta KB di Desa Kalirejo Lawang menunjukkan bahwa jumlah klaster dua (2) memberikan hasil yang lebih optimal dibandingkan jumlah klaster lainnya. Berdasarkan nilai *Silhouette Score*, klasterisasi dengan dua klaster menghasilkan nilai tertinggi, yaitu 0.447, yang menunjukkan bahwa pembagian data menjadi dua kelompok lebih terorganisir dan terpisah dengan jelas. Temuan ini sesuai dengan penelitian oleh Amelia *et al.* (2022), yang juga menunjukkan efektivitas *K-Means* dalam mengelompokkan peserta KB dengan dua klaster, memudahkan analisis yang lebih tepat. Penurunan nilai *Silhouette Score* pada jumlah klaster lebih dari dua mengindikasikan bahwa menambah jumlah klaster justru mengurangi kualitas klasterisasi, yang sejalan dengan penelitian Swastika *et al.* (2024), yang menemukan bahwa lebih banyak klaster tidak selalu meningkatkan ketepatan pengelompokan dalam aplikasi *K-Means*. Pada klaster 2, pembagian data menjadi dua kelompok utama kelompok "tinggi" dan "rendah" dapat menunjukkan perbedaan jelas antara tingkat partisipasi atau penggunaan alat kontrasepsi. Kelompok "C1" (tinggi) dan "C2" (rendah) yang terlihat pada Gambar 4 memperlihatkan bahwa kelompok "C2" memiliki nilai lebih tinggi dibandingkan dengan kelompok "C1", mengindikasikan bahwa data pada kelompok "C2" lebih homogen dalam hal penggunaan alat kontrasepsi.

Pembagian ini dapat mempermudah desain kebijakan yang lebih sesuai dengan kebutuhan masing-masing kelompok. Lino *et al.* (2021) juga menunjukkan pentingnya pengelompokan dalam program KB untuk meningkatkan hasil dan menjangkau kelompok yang memerlukan perhatian lebih.

Perbedaan nilai *Silhouette Score* untuk setiap jumlah kluster menunjukkan bahwa menambah kluster lebih dari dua menyebabkan distribusi data yang lebih buruk, yang berarti informasi relevan mulai terpecah. Hal ini sesuai dengan temuan Haraty *et al.* (2015), yang mengemukakan bahwa lebih banyak kluster tidak selalu memberikan manfaat lebih dalam hal pemisahan data yang lebih baik. Pada kluster 5 dan seterusnya, penurunan nilai *Silhouette Score* semakin tajam, yang menunjukkan bahwa klusterisasi dengan jumlah kluster lebih banyak memperburuk hasil. Implikasi praktis temuan ini adalah bahwa mengelompokkan peserta KB menjadi dua kluster "tinggi" dan "rendah" dapat membantu merancang kebijakan yang lebih terarah. Kelompok "tinggi" mungkin memerlukan pendekatan yang lebih intensif dalam mempertahankan atau meningkatkan penggunaan alat kontrasepsi, sementara kelompok "rendah" membutuhkan perhatian lebih pada edukasi dan peningkatan aksesibilitas layanan KB. Sebagai contoh, kelompok dengan partisipasi rendah mungkin memerlukan edukasi lebih banyak atau fasilitas yang lebih mudah dijangkau. Klusterisasi ini juga memungkinkan penilaian program KB dari waktu ke waktu dengan memantau perubahan dalam kluster, sebagaimana diungkapkan oleh Fitri *et al.* (2023), yang menekankan bahwa pengelompokan data secara dinamis meningkatkan evaluasi program. Hasil ini mendukung penggunaan dua kluster dalam mengelompokkan peserta KB aktif di Desa Kalirejo Lawang dan memberikan dasar yang kuat untuk merancang kebijakan lebih tepat sasaran sesuai dengan kebutuhan masing-masing kelompok.

#### 4. Kesimpulan

Dalam penelitian ini, peneliti melakukan klusterisasi peserta KB aktif di Desa Kalirejo Lawang menggunakan metode *K-Means*. Data yang digunakan mencakup informasi peserta KB selama periode Januari hingga September 2024 yang diperoleh dari BKKBN Desa Kalirejo Lawang. Eksperimen dilakukan dengan menguji nilai  $k$  pada rentang 2 hingga 8. Hasil analisis menunjukkan bahwa jumlah kluster yang paling optimal adalah dua, yaitu kluster C1 dan C2. Hal ini dibuktikan dengan nilai *Silhouette Score* tertinggi sebesar 0.447, yang mendekati 1 dan menunjukkan hasil yang lebih baik. Meskipun nilai 0.447 tergolong rendah, klusterisasi dengan dua kluster ini menghasilkan distribusi data yang lebih baik dibandingkan dengan jumlah kluster lainnya. Nilai *Silhouette Score* untuk kluster lainnya, seperti pada kluster 3 (0.320) dan kluster 4 (0.327), lebih rendah, yang mengindikasikan bahwa distribusi data pada kluster-kluster tersebut kurang optimal. Dari total 75 sampel data, 11 data termasuk dalam kluster C1, sementara 65 data masuk ke dalam kluster C2. Hasil klusterisasi ini dapat mendukung pengambilan keputusan dalam perencanaan program KB yang lebih terfokus dan sesuai sasaran. Selain itu, penelitian ini berkontribusi pada peningkatan pengelolaan data di tingkat desa melalui klusterisasi. Penelitian selanjutnya dapat mengeksplorasi kombinasi metode klusterisasi lain, seperti *Hierarchical Clustering* dan *DBSCAN*, untuk meningkatkan akurasi. Hal ini diharapkan dapat lebih meningkatkan efektivitas program KB di wilayah yang lebih luas.

#### 5. Ucapan Terima Kasih

Penulis menyampaikan terima kasih yang sebesar besarnya kepada orang tua dan pasangan atas dukungan moral dan dorongan yang tiada henti selama proses penelitian ini. Terima kasih juga disampaikan kepada Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN) Desa Kalirejo Lawang atas kerja sama dan penyediaan data yang menjadi dasar penelitian ini. Penulis juga menghargai bantuan teknis dan informasi yang telah diberikan oleh petugas BKKBN di Desa Kalirejo Lawang. Penulis juga ingin menyampaikan terima kasih kepada Institut Teknologi, Sains, dan Kesehatan RS. Dr. Soepraoen Kesdam V/BRW sebagai institusi tempat penulis menimba ilmu. Terima kasih atas fasilitas, bimbingan, dan dukungan yang telah diberikan sehingga penelitian ini dapat terselesaikan dengan baik.

## 6. Daftar Pustaka

- Amelia, R., Martanto, M., & Bahtiar, A. (2022). Penerapan Algoritma K-Means Clustering Pada Pengelompokan Pasangan Usia Subur Peserta Kb Di Kabupaten Cirebon. *JURNAL TEKNOLOGI TECHNOSCIENTIA*, 9-14. <https://doi.org/10.34151/technosciantia.v15i1.3849>.
- de Amorim, R. C., & Makarenkov, V. (2023). On k-means iterations and Gaussian clusters. *Neurocomputing*, 553, 126547. <https://doi.org/10.1016/j.neucom.2023.126547>.
- Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., ... & Zupan, B. (2013). Orange: data mining toolbox in Python. *the Journal of machine Learning research*, 14(1), 2349-2353.
- Estupiñán Ricardo, J., Domínguez Menéndez, J. J., Barcos Arias, I. F., Macías Bermúdez, J. M., & Moreno Lemus, N. (2021). Neutrosophic K-means for the analysis of earthquake data in Ecuador. *Neutrosophic Sets and Systems*, 44(1), 29.
- Fitri, E. M., Suryono, R. R., & Wantoro, A. (2023). Klasterisasi Data Penjualan Berdasarkan Wilayah Menggunakan Metode K-Means Pada Pt Xyz. *Jurnal Komputasi*, 11(2), 157-168. <https://doi.org/10.23960/komputasi.v11i2.12582>.
- Gao, C. X., Dwyer, D., Zhu, Y., Smith, C. L., Du, L., Filia, K. M., ... & Cotton, S. M. (2023). An overview of clustering methods with guidelines for application in mental health research. *Psychiatry Research*, 327, 115265. <https://doi.org/10.1016/j.psychres.2023.115265>.
- Ginting, B. S., & Fatmaira, Z. (2022). Clustering Peserta Kb Aktif Di Kota Binjai Menggunakan Metode K-Means (Study Kasus BKKBN Kota Binjai). *Jurnal Sistem Informasi Kaputama (JSIK)*, 6(2), 208-220.
- Hakim, M. L., Ulandari, S., & Hidayati, N. Implementasi Pembelajaran Berbasis Masalah dalam Meningkatkan Kompetensi Pedagogik Guru Sejarah Kebudayaan Islam. *Jurnal Pendidikan Sultan Agung*, 4(3), 223-234.
- Haraty, R. A., Dimishkieh, M., & Masud, M. (2015). An enhanced k-means clustering algorithm for pattern discovery in healthcare data. *International Journal of distributed sensor networks*, 11(6), 615740. <https://doi.org/10.1155/2015/615740>.
- Hutagalung, J. (2022). Pemetaan Siswa Kelas Unggulan Menggunakan Algoritma K-Means Clustering. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 9(1), 606-620. <https://doi.org/10.35957/jatisi.v9i1.1516>.
- Hutagalung, J., & Sonata, F. (2021). Penerapan Metode K-Means Untuk Menganalisis Minat Nasabah. *Jurnal Media Informatika Budidarma*, 5(3), 1187-1194.
- Lino, M., Jedo, A., & Adam, C. V. (2021). Identifikasi faktor-faktor yang mempengaruhi pengambilan keputusan pasangan usia subur dalam mengikuti program kb (studi kasus di desa leraboleng kecamatan titehena kabupaten flores timur). *Jurnal Administrasi dan Demokrasi (Administration and Democracy Journal)*, 1(2), 100-123. <https://doi.org/https://doi.org/10.35508/jad.v1i2.5599>.
- Maharani, S., & Yotenka, R. (2024). Pengelompokan Kecamatan di Daerah Istimewa Yogyakarta berdasarkan Jumlah Pengguna Alat Kontrasepsi Tahun 2022 dengan K-Medoids Cluster:

Pengelompokan Kecamatan di Daerah Istimewa Yogyakarta berdasarkan Jumlah Pengguna Alat Kontrasepsi Tahun 2022 dengan K-Medoids Cluster. *Emerging Statistics and Data Science Journal*, 2(2), 222-237. <https://doi.org/https://doi.org/10.20885.16>.

Munawar, F., Utami, A. S. D., & Manurung, S. B. T. (2024). KLASTERISASI DAERAH PESERTA KB AKTIF DI KABUPATEN ASAHAN MENGGUNAKAN METODE K-MEANS. *J-Com (Journal of Computer)*, 4(1), 58-67. <https://doi.org/https://doi.org/10.33330/j-com.v4i1.3047>.

Nafilah, M., Rahaningsih, N., & Dana, R. D. (2024). PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS UNTUK PENGELOMPOKAN HASIL PERTANIAN DI KABUPATEN CIREBON. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 223-229. <https://doi.org/10.36040/jati.v8i1.8326>.

Nithya, A., Appathurai, A., Venkatadri, N., Ramji, D. R., & Palagan, C. A. (2020). Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images. *Measurement*, 149, 106952. <https://doi.org/10.1016/j.measurement.2019.106952>.

Nur, I. M. (2024). K-Means++ Algorithm for Health Services Clustering Based on Districts in West Java Province. *EKSAKTA: Journal of Sciences and Data Analysis*, 96-102.

Nurahman, N., & Susanto, J. (2023). Klasterisasi Data Penerima Bantuan Langsung Tunai Menggunakan Algoritma K-Means. *JURIKOM (Jurnal Riset Komputer)*, 10(2), 461-470. <https://doi.org/10.30865/jurikom.v10i2.5807>.

Pratistha, R. N., & Kristianto, B. (2024). Implementasi Algoritma K-Means dalam Klasterisasi Kasus Stunting pada Balita di Desa Randudongkal. *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, 5(2), 1193-1205. <https://doi.org/10.35870/jimik.v5i2.634>.

Rahmawati, G., Sanmas, S. A., Nudyawati, E., Syaharani, N. D., Annas, M. R., & Fauzi, F. (2024, October). Studi Perbandingan Performa: Prediksi Status Stunting Pada Anak Berdasarkan Data Antropometri Menggunakan Algoritma Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN). In *PROSIDING SEMINAR NASIONAL SAINS DATA* (Vol. 4, No. 1, pp. 782-790). <https://doi.org/10.47065/bits.v6i1.5206>.

Raihannabil, S. D., Ilyas, H. M. A., Shafira, H. N., Riani, M. A., Hastin, N. N., & Siregar, T. K. H. (2024, November). Perbandingan Agglomerative Nesting dan K-Means untuk Klasterisasi Ketimpangan Gender berdasarkan Dimensi Kesehatan Reproduksi. In *Seminar Nasional Official Statistics* (Vol. 2024, No. 1, pp. 459-470). <https://doi.org/https://doi.org/10.34123/semnasoffstat.v2024i1.1977>.

Sahputra, R. R., & Kurniawan, R. (2024). KLASTERISASI DATA BALITA STUNTING DI KECAMATAN WILAYAH KABUPATEN CIAMIS BERDASARKAN PREVALENSI DENGAN MENGGUNAKAN ALGORITMA K-MEANS. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 2451-2461. <https://doi.org/10.36040/jati.v8i2.9027>.

Saputri, B. D., & Indriani, D. (2022). Pemetaan Cakupan Pengguna KB Aktif Dan Unmet Need Di Kabupaten Ponorogo Tahun 2021. *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, 15(1). <https://doi.org/10.36456/jstat.vol15.no1.a5446>.

- Shahapure, K. R., & Nicholas, C. (2020, October). Cluster quality analysis using silhouette score. In *2020 IEEE 7th international conference on data science and advanced analytics (DSAA)* (pp. 747-748). IEEE. <https://doi.org/10.1109/DSAA49011.2020.00096>.
- Susilo, P. H. (2021). Klasterisasi Virus Covid-19 Di Wilayah Kabupaten Lamongan Dengan Metode K-Means Clustering. *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, *6*(2), 325-335. <https://doi.org/10.29100/jipi.v6i2.1999>.
- Tifannii, W. F., Mayasari, M., & Rifai, M. (2020). Implementasi program keluarga berencana (kb) dalam upaya menekan pertumbuhan penduduk di kelurahan sumur batu kecamatan bantar gebang kota bekasi. *Dinamika: Jurnal Ilmiah Ilmu Administrasi Negara*, *7*(3), 525-540. <https://doi.org/http://dx.doi.org/10.25157/dinamika.v7i3.4348>.