



## Segmentasi Ketercapaian Vaksinasi Covid-19 di Kabupaten Bandung Dengan Menggunakan K-Medoids Clustering

Soviyanti Rosmini<sup>1</sup>, Yulison Herry Chrisnanto<sup>2</sup>, Wina Witanti<sup>3</sup>

<sup>1</sup>Teknik Informatika, Universitas Jenderal Achmad Yani, Cimahi

<sup>2</sup>Teknik Informatika, Universitas Jenderal Achmad Yani, Cimahi

<sup>3</sup>Teknik Informatika, Universitas Jenderal Achmad Yani, Cimahi

e-mail: soviyantir18@if.unjani.ac.id<sup>1</sup>, yhc@if.unjani.ac.id<sup>2</sup>, wni@if.unjani.ac.id<sup>3</sup>

---

### ABSTRAK

#### **Kata Kunci:**

Vaksinasi  
Pengelompokan  
K-Medoids

Pandemi COVID-19 yang melanda China pada tanggal 31 Desember 2019 menyebar ke seluruh negara. Indonesia menjadi salah satu negara yang terjangkit virus COVID-19, masuknya virus COVID-19 ini di Indonesia pada bulan Maret 2020. Jawa Barat menjadi salah satu provinsi yang terjangkit pandemi COVID-19 termasuk Kabupaten Bandung. Untuk mencegah penularan yang semakin meningkat, maka pemerintah Kabupaten Bandung mewajibkan masyarakatnya untuk melaksanakan vaksin. Target pelaksanaan vaksinasi di Kabupaten Bandung belum mencapai 90%, khususnya pada vaksin dosis lengkap. Tujuan pengelompokan vaksinasi berdasarkan capaian targetnya yaitu untuk membantu pemerintah kabupaten Bandung dalam menjangkau masyarakat yang belum menerima vaksin dosis lengkap, maka perlu adanya data berupa pengelompokan dari penerima vaksin. Pengelompokan capaian target vaksinasi akan menggunakan metode *Clustering K-Medoid*. Hasil dari pengelompokan tersebut diperoleh 4 capaian target vaksinasi diantaranya *cluster 0* sebanyak 3040 kategori, *cluster 1* sebanyak 63 kategori, *cluster 2* sebanyak 195 kategori, dan *cluster 3* sebanyak 217 kategori. Pengujian *silhouette coefficient* menunjukkan hasil 0.7595 yang menunjukkan hasil interpretasinya kuat.

### ABSTRACT

#### **Keyword:**

Vaccination  
Grouping  
K-Medoids

*The covid-19 pandemic that swept across China on December 31, 2019 spread across the country. Indonesia became one of the countries with the covid-19 virus, the entry of the covid-19 virus in Indonesia in March 2020. West Java became one of the provinces affected by the covid-19 pandemic, including the municipality of bandung. To prevent increased infections, the bandung government mandated its people to carry out vaccines. The target for vaccination in bandung district has not reached 90%, especially with a full dose vaccine. The goal of grouping vaccinations based on the target's reach is to help the bandung government reach those who have not received a full dose of the vaccine, and then there must be data for grouping the vaccine from those who receive it. Clustering the target measurement of vaccination will use clustering k-medoid methods. The result of the grouping is from cluster 4 produces 3040 categories, cluster 1 produces 63 categories, cluster 2 produces 195, and cluster 3 produces 217 categories. Silhouette coefficient tests show 0.7595 with strong interpretive results.*



## PENDAHULUAN

Pandemi COVID-19 yang melanda China pada tanggal 31 Desember 2019 menyebar ke seluruh negara [1], [2]. Indonesia menjadi salah satu negara yang terjangkit virus COVID-19, masuknya virus COVID-19 ini di Indonesia pada bulan Maret 2020 yang menyebabkan terhambatnya berbagai aktifitas [3]. Upaya pemerintah dalam menangani kasus pandemi ini yaitu dengan menerapkan protokol kesehatan berupa pembatasan sosial berskala besar, mencuci tangan dengan menggunakan sabun, memakai masker, menjaga jarak, membatasi kegiatan diluar ruangan serta menjauhi kerumunan [3]. Adapun upaya lainnya yang mulai diterapkan yaitu program vaksinasi COVID-19 yang wajib diikuti oleh masyarakat Indonesia guna menangani melonjaknya kasus COVID-19 [2], [4], [5].

Jawa Barat menjadi salah satu provinsi yang terjangkit pandemi COVID-19 termasuk kabupaten Bandung. Target pelaksanaan vaksinasi di Kabupaten Bandung belum mencapai 90%, khususnya pada vaksin dosis ke 2 dan 3 [5]. Berdasarkan capaian target vaksinasi tersebut, maka diperlukan pengelompokan vaksinasi berdasarkan capaian target penyuntikkan vaksin lengkap. Adapun faktor-faktor yang mempengaruhi belum tercapainya target vaksinasi lengkap yaitu beredarnya berita *hoax* mengenai efek samping vaksin sehingga masyarakat ikut terbawa berita tersebut, adapun keraguan terhadap keamanan dan keselamatan vaksin, dan banyaknya opini negatif yang menggiring sesama masyarakat untuk tidak melaksanakan vaksinasi [6].

Penelitian ini menggunakan data vaksinasi harian yang diperoleh dari portal resmi Kementerian Kesehatan dari bulan Januari 2021 hingga Juli 2022. Metode penelitian yang digunakan pada penelitian ini yaitu dengan menggunakan algoritma *K-Medoids clustering*. Penelitian ini dilakukan untuk mengetahui bagaimana algoritma *K-Medoids clustering* dalam menentukan pencapaian vaksinasi pada dosis 1, dosis 2, dosis 3 di kabupaten Bandung berdasarkan kategori penerimanya dan bagaimana algoritma *K-Medoids clustering* dalam menentukan persebaran penerima vaksinasi yang lebih mendominasi berdasarkan pencapaiannya. Adapun tujuan dari penelitian ini yaitu menunjukkan hasil segmentasi pada data vaksinasi COVID-19 berdasarkan ketercapaian target vaksinasi pada dosis ke-1, dosis ke-2 dan dosis ke-3 yang akan dikelompokkan kedalam segmentasi ketercapaiannya dan menunjukkan hasil persebaran penerima vaksinasi yang lebih mendominasi pada setiap pencapaiannya.

Penelitian terdahulu pada algoritma *K-Medoids clustering* yaitu pada penelitian [4] mengelompokkan data COVID-19 sebanyak 3 *cluster* dari 34 *record* data, 1 *record* pada *cluster* pertama, 2 *record* pada *cluster* kedua dan 31 *record* pada *cluster* ketiga. Adapun penelitian [7] mengelompokkan kasus penyebaran COVID-19 kedalam 3 *cluster* yaitu pada *cluster* 1 perlu meningkatkan lagi protokol kesehatan menjadi lebih ketat, sedangkan pada *cluster* 2 dan *cluster* 3



perlu mempertahankan kondisi penyebaran COVID-19 yang masih termasuk kedalam golongan rendah.

## METODE

Metode penelitian yang digunakan yaitu menggunakan algoritma *K-Medoids clustering*. Adapun tahapan yang dilakukan untuk melakukan pengelompokan data vaksinasi yaitu pengambilan dataset disesuaikan dengan penelitian yang dilakukan yaitu dataset vaksinasi harian di kabupaten Bandung. Dataset diperoleh dari portal resmi Vaksinasi Kementerian Kesehatan. Dataset diambil pada tanggal 10 Juli 2022 dengan jumlah data sebanyak 10.331 data.

Langkah-langkah dalam melakukan *pre-processing* data yaitu dengan melakukan analisis dataset, identifikasi dataset, dan *cleaning* dataset. Analisis dataset dilakukan untuk mengetahui jumlah atribut, jumlah dataset, jumlah atribut yang tidak digunakan, jumlah atribut yang digunakan, jumlah atribut bersifat numerik dan nominal serta jumlah pada baris data yang kosong. Identifikasi dataset dilakukan untuk menentukan jumlah atribut dan kategori penerima vaksin yang akan digunakan. Atribut yang digunakan pada penelitian ini sebanyak 5 atribut yaitu kategori, tanggal, divaksin 1, divaksin 2 dan divaksin 3. Kategori penerima vaksin yang digunakan sebanyak 30 kategori. *Cleaning* dataset dilakukan untuk menghapus data yang duplikasi, dan data yang kosong. Tahapan ini akan menghapus pada 2 kategori yang tidak akan digunakan seperti menunggu konfirmasi dan WNA. Adapun pengambilan sampel pada setiap kategori penerima vaksinnya yaitu sebanyak 120 sampel.

*K-Medoids clustering* yaitu suatu metode pada *data mining* pada metode *clustering* non hirarki yang akan mengelompokkan sejumlah data [8]. Pengelompokan data dilakukan jika nilai K sebagai jumlah *cluster* sudah ditentukan. Adapun langkah-langkah yang digunakan untuk pengelompokan diantaranya [4], [9], [10] :

1. Melakukan inisialisasi terhadap pusat *cluster*, ditentukan sebanyak K.
2. Melakukan alokasi pada setiap objek data pada jarak cluster terdekat dengan menggunakan rumus *Euclidean Distance*.

$$d(i,j) = \sqrt{(x_i - x_j)^2 - (y_i - y_j)^2} \quad (1)$$

Keterangan :

$x_i$  : objek x ke-i

$y_i$  : objek y ke-i

n : banyaknya objek



3. Melakukan pemilihan objek secara acak pada masing-masing *cluster* untuk kandidat medoid baru.
4. Menghitung jarak pada setiap objek yang berada pada masing-masing *cluster* dengan kandidat medoid yang baru.
5. Melakukan perhitungan pada total simpangan yang dilakukan pada nilai total *distance* baru dan nilai total *distance* lama. Saat melakukan perhitungan total simpangan, perlu diperhatikan bahwa  $S < 0$  maka perlu menukar objek dengan data cluster, sehingga membentuk kumpulan K objek baru sebagai medoid.
6. Ulangi proses 3 sampai 5 hingga tidak ada perubahan pada medoid.

Metode *Elbow* digunakan untuk menentukan jumlah *cluster* [11]. Jumlah *cluster* yang ditentukan diperoleh dari hasil persentase yang membandingkan jumlah *cluster* yang membentuk siku pada suatu titik tertentu. Untuk mendapatkan nilai pembandingnya dapat dihitung dengan menggunakan rumus SSE [11].

$$SSE = \sum_{k=1}^K \sum_{X_i \in S_k} \| X_i - C_k \|^2 \quad (2)$$

Keterangan :

K : jumlah *cluster*

$X_i$  : data ke-i

$C_k$  : centroid *cluster*

*Silhouette Coefficient* adalah proses evaluasi untuk melakukan pengujian terhadap kualitas *cluster* [12]. *Silhouette Coefficient* memiliki rentang dari -1 sampai dengan 1. Jika nilai mendekati angka 1 maka nilai tersebut baik, jika mendekati -1 tidak baik. Pengujian dengan *Silhouette Coefficient* yaitu dimulai dengan mencari jarak dari rata-rata *cluster* yang sama yaitu:

$$a(i) = \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i, j) \quad (3)$$

Keterangan :

A : banyaknya data yang berada di *cluster* A

Setelah mendapatkan nilai a(i), nilai b(i) dicari untuk memilih jarak paling minimum yang diperoleh dari jarak rata-rata data i pada *cluster* yang berbeda. Untuk menghitung nilai b(i) dengan menggunakan persamaan 4.

$$b(i) = \min_{C \neq A} \frac{1}{|C|} \sum_{j \in C} d(i, j) \quad (4)$$

Mencari nilai d(I,C) untuk menghitung jarak rata rata data i, pada semua data di *cluster* C dengan menggunakan persamaan 5.

$$d(i, C) = \frac{1}{|C|} \sum_{j \in C} d(i, j) \quad (5)$$



Keterangan :

C : banyaknya data pada *cluster* C

Perhitungan *Silhouette Coefficient* dilakukan setelah nilai  $a(i)$  dan  $b(i)$  diketahui. Proses perhitungan *Silhouette Coefficient* yaitu dengan menggunakan persamaan 6.

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

Nilai  $s(i)$  dapat dikatakan berada pada rentang -1 dan 1 pada setiap nilainya, dilakukan interpretasi dengan cara sebagai berikut [13] :

**Tabel 1.** Kriteria *Silhouette Coefficient*

<b>Silhouette Coefficient</b>	<b>Hasil Interpretasi</b>
0.71 – 1.00	Kuat
0.51 – 0.70	Baik
0.26 – 0.50	Lemah
$\leq 0.25$	Buruk

## HASIL DAN PEMBAHASAN

Nilai  $K=4$  diperoleh berdasarkan perhitungan dengan menggunakan rumus SSE yang berada pada persamaan (2). *Cluster* yang digunakan pada penelitian ini yaitu *cluster* 0 (C0), *cluster* 1 (C1), *cluster* 2 (C2) dan *cluster* 3 (C3). Penentuan *cluster* C0, C1,C2,C3 dimulai dari 0 berdasarkan array yang terbentuk.

Adapun indikator yang menentukan keterangan pada setiap target vaksinasi yaitu :

**Tabel 2.** Indikator Ketercapaian Vaksinasi

<b>Indikator</b>	<b>Target Vaksin</b>	<b>Keterangan</b>
Sudah Tercapai	95%	3420
Cukup Tercapai	79%	2884
Belum Tercapai	19%	684
Tidak Tercapai	< 19%	< 684

*Sumber : artikel berita harian Kementerian Kesehatan RI*

Bagian keterangan menunjukkan hasil perhitungan dari :

- 95% dari 3600 = 3420
- 79% dari 3600 = 2884
- 19% dari 3600 = 684



Perhitungan *cluster* dilakukan dengan menggunakan metode *K-Medoids clustering*. Untuk menghitung jarak yang paling dekat dengan *medoids* dilakukan dengan menggunakan rumus *Euclidean Distance* yaitu pada persamaan (1). Pada perhitungan ini, menggunakan sampel data sebanyak 120 data disetiap kategorinya dari 3600 populasi. Hasil perhitungan dengan nilai  $K=4$  diperoleh hasil seperti yang ditampilkan pada Tabel (3)

**Tabel 3.** Hasil Perhitungan Algoritma K-Medoids Clustering

No	Kategori	Divaksin1	Divaksin2	Divaksin3	Jarak	Cluster
1	ANAK-ANAK, 23/10/2021	12	1	0	3,162	C0
2	ANAK-ANAK, 24/10/2021	12	0	0	3	C0
3	ANAK-ANAK, 25/10/2021	25	1	0	16,031	C2
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
3600	WARTAWAN DAN PEKERJA MEDIA, 30/06/2022	0	0	1	9,055	C0

Hasil ketercapaian pada setiap cluster dijelaskan pada tabel (4) :

**Tabel 4.** Hasil Cluster

Cluster	Jumlah Vaksinasi	Keterangan Target Vaksin
C0	3040	Cukup Tercapai
C1	63	Tidak Tercapai
C2	195	Tidak tercapai
C3	217	Tidak Tercapai

Berdasarkan hasil pada tabel (4), *cluster* 0 menunjukkan hasil yang menunjukkan cukup tercapai, sedangkan pada *cluster* 1, *cluster* 2 dan *cluster* 3 menunjukkan hasil tidak tercapai berdasarkan indikatornya.

Adapun mayoritas pada setiap *cluster* yang lebih mendominasi bahwa penerima vaksinasi terbanyak pada setiap clusternya terdapat pada beberapa kategori penerima vaksinasi. Mayoritas vaksinasi pada setiap clusternya dijelaskan pada tabel (5).



**Tabel 5.** Mayoritas Penerima Vaksinasi

Cluster	Mayoritas	Jumlah
C0	BUMD, Disabilitas Lainnya, Pedagang Pasar, dan TNI	120
C1	Pendidik	18
C2	Anak-anak	32
C3	Pelayan Publik	37

Pengujian *cluster* pada nilai K yang akan dijelaskan pada Tabel (6).

**Tabel 6.** Hasil Pengujian Cluster

Nilai Cluster	Silhouette Coefficient	Hasil Interpretasi
2	0.7462	Kuat
3	0.7404	Kuat
4	0.7595	Kuat
5	0.6933	Kuat

*Silhouette coefficient* akan melakukan perhitungan dengan 3 tahap.

- Tahap pertama yaitu menghitung jarak rata-rata pada semua objek dalam clusternya dengan menggunakan persamaan (3).
- Tahap kedua yaitu menghitung jarak rata-rata pada semua objek lain di *cluster* lain, dengan menggunakan persamaan (4) dan (5).
- Tahap ketiga yaitu menghitung angka *silhouette coefficient* nya dengan menggunakan persamaan (6).

Hasil perhitungan dengan menggunakan *silhouette coefficient* pada seluruh nilai K dinyatakan kuat, karena berdasarkan tabel kriteria *silhouette coefficient* jika berada pada rentang 0.71 hingga 1.00 maka hasil interpretasi-nya dinyatakan kuat. Nilai K=4 menunjukkan hasil 0.7595 dimana hasil perhitungan pengujian clusternya dinyatakan kuat.

## SIMPULAN

Dalam penelitian Segmentasi Ketercapaian Vaksinasi COVID-19 di Kabupaten Bandung Dengan Menggunakan K-Medoids Clustering memberikan hasil pengelompokkan sebanyak 4 *cluster* dengan target ketercapaian vaksinasinya tidak tercapai. Adapun mayoritas penerima vaksinasi pada setiap clusternya yaitu pada *cluster* 0 mayoritas penerima vaksinasinya yaitu BUMD, disabilitas lainnya, pedagang pasar dan TNI sebanyak 120 penerima. *Cluster* 1 mayoritas penerima vaksinnnya yaitu pendidik sebanyak 18 penerima. *Cluster* 2 mayoritas penerima vaksinnnya yaitu anak-anak sebanyak 32 penerima. *Cluster* 3 mayoritas penerima vaksinnnya yaitu pelayan publik sebanyak 37



penerima. Banyaknya dari kategori lain yang belum mencapai target vaksinasi COVID-19 sehingga status ketercapaian vaksinasi di kabupaten Bandung masih tidak tercapai.

#### DAFTAR RUJUKAN

- [1] N. P. E. D. Yanti, I. M. A. D. P. Nugraha, G. A. Wisnawa, N. P. D. Agustina, and N. P. A. Diantari, “Gambaran Pengetahuan Masyarakat Tentang COVID-19 dan Perilaku Masyarakat di Masa Pandemi COVID-19,” *Jurnal Keperawatan Jiwa*, vol. 8, no. 4, pp. 491–504, 2020, doi: 10.26714/jkj.8.4.2020.491-504.
- [2] P. P. Pratama, H. Witriyono, R. Toyib, and A. Sonita, “Analisis dan Penerapan Algoritma C4.5 Untuk Klasifikasi Penerima Vaksinasi COVID-19 di Kota Bengkulu,” *Jurnal Innovation Informatics*, vol. 1, no. 2, pp. 105–112, 2022.
- [3] Triyo Rachmadi, Titi Pudji Rahayu, Ari Waluyo, and Wakhid Yuliyanto, “Pemberian Vaksinasi COVID-19 Bagi Masyarakat Kelompok Petugas Pelayanan Publik di Kecamatan Buluspesantren,” *JURPIKAT (Jurnal Pengabdian Kepada Masyarakat)*, vol. 2, no. 2, pp. 104–119, 2021, doi: 10.37339/jurpikat.v2i2.643.
- [4] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. Ilmi R.H.Zer, and D. Hartama, “Analisis Algoritma K-Medoids Clustering dalam Pengelompokan Penyebaran COVID-19 di Indonesia,” *Jti (Jurnal Teknologi Informasi)*, vol. 4, no. 1, pp. 166–173, 2020.
- [5] I. M. Apriliani, N. P. Purba, L. P. Dewanti, H. Herawati, and I. Faizal, “Pemetaan Capaian Vaksinasi COVID-19 di Provinsi Jawa Barat Tahun 2021,” *Media Publikasi Promosi Kesehatan Indonesia*, vol. 5, no. 1, pp. 545–551, 2021, doi: <https://doi.org/10.56338/mppki.v5i1.2226>.
- [6] N. P. Astuti, E. G. Z. Nugroho, J. C. Lattu, I. R. Potempu, and D. A. Swandana, “Persepsi Masyarakat Terhadap Penerimaan Vaksinasi COVID-19: Literature Review,” *Jurnal Keperawatan*, vol. 13, no. 3, pp. 569–580, 2021, doi: 10.32583/keperawatan.v13i3.1363.
- [7] Husdi and M. Nanja, “Clustering Daerah Penyebaran Covid-19 di Indonesia Menggunakan Algoritma K-Medoids,” *Jurnal JTIC (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 6, no. 4, pp. 609–615, 2022, doi: <https://doi.org/10.35870/jti.k.v6i4.608> Keywords:
- [8] D. I. Komputer, F. Matematika, D. A. N. Ilmu, and P. Alam, “Aplikasi Berbasis Web Untuk Deteksi Pencilan Titik Panas Menggunakan Algoritma Clustering K-Means dan Framework Shiny,” 2015.
- [9] T. I. Hermanto and Y. Muhyidin, “Analisis Sebaran Titik Rawan Bencana dengan K-Means Clustering dalam Penanganan Bencana,” *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, vol. 5, no. 1, pp. 406–416, 2021, doi: <http://dx.doi.org/10.30645/j-sakti.v5i1.332>.
- [10] F. Yazid and M. Affandes, “Clustering Data Polutan Udara Kota Pekanbaru dengan Menggunakan Metode K-Means Clustering,” *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu*





- Komputer dan Teknologi Informasi*, vol. 3, no. 2, pp. 76–81, 2018, doi: 10.24014/coreit.v3i2.4419.
- [11] S. A. Mousa, A. A. Elbiomy, and A. E. Abo-Hussien, “A Comparison Study of K-Means and K-Medoids Algorithms With an Application on COVID-19 Data as an Example,” *The 32nd Annual International Conference on Statistics and Modeling in Human and Social Sciences*, no. February, 2022.
- [12] D. A. I. C. Dewi and D. A. K. Pramita, “Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali,” *Matrix : Jurnal Manajemen Teknologi dan Informatika*, vol. 9, no. 3, pp. 102–109, 2019, doi: 10.31940/matrix.v9i3.1662.
- [13] L. Kaufman and P. J. Rousseeuw, *Finding Groups In Data An Introduction to Cluster Analysis*, vol. 1999, no. December. USA, 1990.