

ALGORITMA *K-MEANS* UNTUK MENGELOMPOKKAN KELURAHAN DI PROVINSI BANTEN BERDASARKAN LUAS WILAYAH DAN JUMLAH PENERIMA BANTUAN BERAS BULOG OLEH PT YASA ARTHA TRIMANGUNGGAL

San Prastiwa¹, Arief Wibowo^{2*}

^{1,2}Teknik Informatika, Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia

Email: ¹2011500556@student.budiluhur.ac.id, ^{2*}arief.wibowo@budiluhur.ac.id
(* : corresponding author)

Abstrak- Pendistribusian bantuan beras BULOG kepada masyarakat, khususnya di kelurahan-kelurahan di Provinsi Banten, sering menghadapi tantangan dalam penentuan wilayah distribusi yang tepat sasaran. Tantangan ini terutama terkait dengan ketidakseimbangan antara luas wilayah dan jumlah penerima, yang menyebabkan bantuan tidak selalu sampai kepada yang paling membutuhkan secara efektif. Untuk mengatasi masalah ini, PT Yasa Artha Trimanggunggal menerapkan teknik data mining dengan menggunakan algoritma *K-Means*, sebuah metode *clustering* yang efektif dalam mengelompokkan data berdasarkan karakteristik serupa, yaitu luas wilayah dan jumlah penerima bantuan per kelurahan. Dalam penelitian ini, data yang digunakan mencakup luas wilayah dan jumlah penerima bantuan di setiap kelurahan. Data ini diolah untuk menemukan pola distribusi yang lebih terperinci dan terfokus, guna meningkatkan efisiensi dan efektivitas distribusi bantuan. Metode *Elbow* digunakan untuk menentukan jumlah kluster yang optimal dengan mengevaluasi variasi total dalam setiap jumlah kluster. Hasil analisis menunjukkan bahwa tiga kluster adalah jumlah yang optimal, dengan nilai *SSE* (*Sum of squared errors*) sebesar 5936976,31. Uji coba menunjukkan bahwa penerapan tiga kluster ini meningkatkan efisiensi distribusi bantuan beras dengan keberhasilan mencapai 85%. Pembagian ini menghasilkan pengelompokan data kelurahan yang lebih efisien dan bermakna berdasarkan analisis statistik dan evaluasi kinerja *cluster*. Dengan hasil ini, PT Yasa Artha Trimanggunggal dapat menyusun strategi distribusi bantuan yang lebih tepat sasaran, memastikan bahwa bantuan beras BULOG dapat didistribusikan dengan lebih efektif dan efisien di Provinsi Banten. Teknik data mining yang digunakan dalam penelitian ini diharapkan dapat memberikan kontribusi signifikan bagi optimalisasi distribusi bantuan sosial di masa depan serta menjadi acuan bagi penelitian-penelitian serupa di wilayah lainnya.

Kata Kunci: *K-Means*, Kluster, *Clustering*, Metode *Elbow*, BULOG

K-MEANS ALGORITHM TO GROUP VILLAGES IN BANTEN PROVINCE BASED ON THE AREA AND NUMBER OF RECIPIENTS OF BULOG RICE ASSISTANCE BY PT YASA ARTHA TRIMANGUNGGAL

Abstract- The distribution of BULOG rice aid to communities, particularly in the urban villages of Banten Province, often faces challenges in determining the most targeted distribution areas. These challenges are primarily related to the imbalance between the area size and the number of recipients, which causes the aid not to always reach those most in need effectively. To address this issue, PT Yasa Artha Trimanggunggal has applied data mining techniques using the *K-Means* algorithm, an effective clustering method for grouping data based on similar characteristics, namely the area size and the number of aid recipients per urban village. In this study, the data used includes the area size and the number of aid recipients in each urban village. This data is processed to identify more detailed and focused distribution patterns to improve the efficiency and effectiveness of aid distribution. The *Elbow Method* is employed to determine the optimal number of clusters by evaluating the total variation within each cluster number. The analysis results show that three clusters are optimal, with a *Sum of squared errors* (*SSE*) value of 5936976.31. Trials demonstrate that implementing these three clusters improves the efficiency of rice aid distribution, with a success rate of 85%. This clustering results in a more efficient and meaningful grouping of urban village data based on statistical analysis and cluster performance evaluation. With these results, PT Yasa Artha Trimanggunggal can develop a more targeted aid distribution strategy, ensuring that BULOG rice aid is distributed more effectively and efficiently in Banten Province. The data mining techniques used in this study are expected to make a significant contribution to optimizing social aid distribution in the future and serve as a reference for similar research in other regions.

Keywords: *K-Means*, Cluster, *Clustering*, *Elbow Method*, BULOG

1. PENDAHULUAN

Dalam mengelola program bantuan beras, penting bagi pemerintah dan lembaga terkait untuk memastikan bahwa bantuan tersebut disalurkan secara efisien dan tepat sasaran kepada masyarakat yang membutuhkan. Perusahaan Yasa Artha Trimanunggal yang beralamat di Jalan Indraloka II No. 1850, RT.7/RW.6, Wijaya Kusuma, Kecamatan Grogol Petamburan, Kota Administrasi Jakarta Barat, Daerah Khusus Ibukota Jakarta 11460, terlibat dalam distribusi bantuan beras dari Badan Urusan Logistik (Bulog) di Indonesia, terutama di Provinsi Banten. Perusahaan ini dihadapkan pada gandingan dalam mengelola proses distribusi bantuan tersebut agar berjalan dengan efisien. Masalah utama yang dihadapi dalam proses distribusi bantuan beras meliputi ketidakakuratan data penerima yang dapat mengakibatkan bantuan tidak tepat sasaran, ketidakefisienan distribusi yang dapat menyebabkan penumpukan stok di beberapa lokasi sementara lokasi lain kekurangan, kurangnya pemahaman terhadap profil demografis dan geografis penerima bantuan yang menghambat perancangan kebijakan distribusi yang efektif, keterbatasan sumber daya baik dalam hal logistik maupun tenaga kerja yang menuntut adanya sistem distribusi yang efisien, serta perubahan cepat dalam kondisi sosial dan ekonomi masyarakat yang memerlukan sistem yang adaptif untuk memastikan distribusi tetap relevan dan tepat sasaran.

Pada penelitian sebelumnya yang dilakukan oleh [1] yang berjudul “Implementasi Data Mining untuk Manajemen Bantuan Sosial Menggunakan Algoritma K-Means”, penelitian tersebut untuk mengidentifikasi kelompok prioritas penerima Program Keluarga Harapan (PKH) untuk keluarga berpenghasilan rendah di Kecamatan Medan Tembung, Data yang digunakan yaitu Data Kecamatan Medan Tembung, namun hasil *Clustering* pada penelitian tersebut tidak dievaluasi, sehingga tidak dapat mengetahui kualitas dari hasil *Clustering* yang didapat.

Perbedaan penelitian tersebut dengan penelitian ini yaitu pada proses evaluasi hasil *Clustering*. Pada Penelitian ini, hasil *Clustering* yang didapat dari perhitungan K-Means dievaluasi menggunakan metode Elbow, yang dimana penelitian ini mengevaluasi hasil perhitungan dengan berbagai jumlah *Cluster*, sehingga dapat menentukan kualitas *Clustering* yang baik.

Analisis klaster, yang memungkinkan identifikasi pengelompokan dengan karakteristik yang sama berdasarkan lokasi geografis dan faktor demografis, merupakan salah satu metode yang dapat digunakan untuk memahami pola penerimaan bantuan beras. Metode *K-Means* merupakan pilihan yang tepat dalam analisis klaster untuk distribusi bantuan beras karena metode ini dapat menangani data yang bersifat noise dan kategorik. Dengan menggunakan metode ini, kita dapat membagi penerima bantuan beras menjadi lebih homogen dan mencatat perbedaan yang signifikan secara demografis dan geografis di antara mereka.

Analisis klaster *K-Means* bukan hanya sekedar alat statistik, tetapi juga merupakan instrumen kebijakan yang penting. Dengan pemahaman yang lebih mendalam mengenai profil demografis dan geografis penerima manfaat beras, pemerintah dapat merancang kebijakan yang lebih terfokus, termasuk distribusi sumber daya yang efisien dan program bantuan tambahan yang sesuai dengan kebutuhan setiap kelompok penerima manfaat. Dengan menggunakan metode ini, diharapkan program bantuan beras menjadi lebih inklusif, responsif, dan sesuai dengan kebutuhan semua kelompok penerima manfaat.

2. METODE PENELITIAN

2.1 Data Mining

Data mining adalah proses penggabungan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan menemukan informasi yang relevan dan pengetahuan yang terkait dari *database* atau data warehouse yang sangat besar [2]. Data mining memungkinkan penggunaan beberapa teknik, termasuk pemilihan variabel, regresi, pengelompokan, klasifikasi, dan analisis [3]. Data tentang perusahaan selalu meningkat sebanding dengan tindakan yang terjadi di sana. Tidak diragukan lagi, jumlah dan keragaman barang yang dibeli konsumen akan selalu mempengaruhi jumlah informasi yang tersedia mengenai proses pembuatan produk tertentu.

Akibatnya, *database* bisnis sekarang lebih besar dari sebelumnya dan mencakup lebih banyak data daripada sebelumnya. Ketersediaan basis data yang semakin besar ini jarang dimanfaatkan dalam proses pengambilan keputusan; sebaliknya, firasat para pengambil keputusan lebih sering menjadi pertimbangan. KDD adalah metode komputer yang menggunakan algoritma matematika dan ekstraksi data untuk memperkirakan probabilitas perilaku hipotetis di masa depan [2].

Informasi dapat diekstraksi dari kumpulan data yang sudah ada sebelumnya dengan menggunakan pendekatan KDD [4]. Data mining dapat digunakan untuk memeriksa kumpulan data yang besar dan memberikan informasi dalam bentuk pola yang signifikan yang dapat digunakan untuk mendukung keputusan. Aplikasi data

mining menghasilkan pengetahuan dan informasi yang segar, menarik, dan penting untuk bisnis. Setelah data ini diperiksa, hasilnya ditampilkan untuk membantu pengguna dalam memilih data yang relevan untuk mendukung pengambilan keputusan mereka.

2.2 Clustering

Clustering adalah pendekatan data mining yang menemukan kelompok-kelompok benda berdasarkan kualitas serupa yang memungkinkan mereka dibedakan dari kelompok benda lainnya. Hasilnya, jika dibandingkan dengan item serupa di *cluster* lain, objek di dalam sebuah *cluster* sering kali lebih homogen. *Cluster* adalah kumpulan record yang memiliki kemiripan satu sama lain tetapi berbeda dengan record di *cluster* lain [5].

Clustering adalah proses menemukan pola dalam data dan mengelompokkan data yang serupa [6]. Dalam *Clustering*, item data dikelompokkan, dipecah, atau dibagi menjadi k *cluster*, di mana k kurang dari atau sama dengan data *cluster*. Banyak rumus dan prosedur algoritma jarak - terutama Euclidean Distance - yang dapat digunakan untuk mencapai hal ini. Analisis klaster adalah metode untuk mengorganisir kumpulan data berdasarkan kesamaan yang telah ditentukan [7].

Tujuan utamanya adalah agar sekumpulan item data dapat dikategorikan ke dalam kelompok-kelompok yang dapat dibedakan satu sama lain dan memiliki fitur-fitur tertentu sehingga memudahkan untuk menganalisis dan memahami data lebih lanjut sesuai dengan tujuan proyek. Model yang digunakan diperkirakan dapat menggunakan tipe data seperti data biner, frekuensi, dan interval.

2.3 K-Means

Algoritma *K-Means* termasuk dalam domain teknik pengelompokan pembelajaran tak terawasi yang menggunakan mekanisme partisi untuk membagi data ke dalam beberapa kelompok [8]. Teknik pengelompokan yang paling dasar, *K-Means* membagi data ke dalam k kelompok sesuai dengan *Centroid* dari masing-masing kelompok. Namun, inisialisasi *Centroid* dan parameter k memiliki dampak yang signifikan terhadap hasil *K-Means*. Mengurangi fungsi tujuan yang ditetapkan selama proses pengelompokan adalah tujuan dari teknik pengelompokan data ini [9]. Biasanya, *K-Means* menginisialisasi *centroid* secara acak. Langkah-langkah algoritma *K-Means* seperti yang dinyatakan oleh [2] adalah sebagai berikut:

- Sebagai pusat klaster awal, pilih satu objek secara acak.
- Dengan menggunakan nilai rata-rata objek dalam klaster, kelompokkan setiap objek ke dalam klaster dengan objek terkait.
- Tentukan rata-rata dari setiap item klaster untuk memperbarui nilai rata-rata klaster.
- Lanjutkan langkah kedua dan ketiga hingga tidak ada anggota klaster yang berubah.

2.4 Metode Elbow

Metode *Elbow* adalah metode yang menentukan jumlah *cluster* optimal dengan membandingkan jumlah *cluster* yang membentuk siku pada suatu tempat [10]. Metode *Elbow* menghitung nilai perbandingan dengan menggunakan Sum of Squares (SSE) dari setiap nilai *cluster*, dimana SSE menurun seiring dengan jumlah nilai *cluster* K . Metode *Elbow* menurut [11] merupakan teknik yang melihat persentase hasil perbandingan jumlah *cluster* yang akan membentuk siku di suatu tempat sehingga dapat menghasilkan informasi mengenai jumlah *cluster* yang ideal. Dengan pertama-tama memilih nilai klaster dan kemudian menambahkannya ke model data untuk digunakan dalam mengidentifikasi klaster yang optimal, teknik *Elbow* menghasilkan ide. Selain itu, persentase komputasi yang dihasilkan digunakan untuk membandingkan jumlah *cluster* yang dibuat. Dengan menggunakan grafik sebagai sumber data, nilai persentase yang berbeda dari setiap nilai *cluster* dapat ditampilkan. Nilai *cluster* optimal jika grafik menunjukkan sudut antara nilai *cluster* pertama dan kedua, atau jika nilainya paling banyak berkurang. Rumus SSE adalah persamaan (1) sebagai berikut:

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

Keterangan:

K	=	Jumlah klaster
X _i	=	Data ke i
C _k	=	<i>Centroid</i> klaster ke i

Adapun langkah-langkah dalam pengujian menggunakan *Elbow Method* sebagai berikut:

- Inisialisasi awal nilai K

- b. Naikkan nilai K .
- c. Tentukan hasil *Sum of square error* setiap nilai K .
- d. Lihat hasil *Sum of square error* untuk penurunan tajam pada nilai K .
- e. Tentukan nilai mana yang optimal, yaitu nilai K yang berbentuk siku.

2.5 Tahapan *Knowledge Discover in Database* (KDD)

2.5.1 Data Selection

Menemukan fitur yang relevan dari pengumpulan data adalah tujuan dari seleksi data, yang bertujuan untuk meningkatkan efisiensi analisis data. Pada tahap ini dipilih 8 atribut sebagai berikut:

- a. Provinsi
- b. Kabupaten/Kota
- c. Kecamatan
- d. Kelurahan
- e. Luas Wilayah
- f. Jumlah Penerima
- g. *Latitude*
- h. *Longitude*

2.5.2 Pre-Processing

Ada banyak tahapan yang terlibat dalam tahap pra-pemrosesan data mining untuk studi kasus penerima bantuan pangan. Pertama, informasi dikumpulkan dalam file Excel tentang pengguna bantuan pangan, termasuk kabupaten, kecamatan, desa, wilayah, dan jumlah penerima. Selanjutnya, data dibersihkan dengan menghilangkan entri yang berlebihan, mengisi bagian yang kosong, dan memperbaiki format atau kesalahan pengetikan. Setelah itu, Kelurahan digunakan untuk mengorganisir data dan memformat ulang data yang diperlukan. Setelah itu, informasi tersebut dirangkum dan disusun secara rapi dalam sebuah tabel dengan menghitung jumlah penerima bantuan untuk setiap Kelurahan menggunakan fungsi COUNTIFS di Excel. Pada akhirnya, file Excel dengan data yang telah dirangkum disimpan.

2.5.3 Transformation

Praktik mengubah data yang dipilih agar lebih memenuhi persyaratan prosedur data mining dikenal sebagai transformasi data. Penemuan pengetahuan dalam basis data, atau KDD, menggunakan proses transformasi kreatif yang sangat bergantung pada jenis atau struktur data yang harus diekstrak dari basis data. Data dari PT Yasa Artha Trimanunggal sudah sesuai dan siap untuk data mining, sehingga penulis memilih untuk tidak mengubah data tersebut.

2.5.4 Data Mining

Tahap selanjutnya adalah menganalisa data dengan menggunakan teknik *K-Means Clustering* untuk mengekstrak informasi dari data setelah dibersihkan dan dimodifikasi. Aplikasi yang dibuat dengan memanfaatkan metode dan komputasi algoritma *K-Means Clustering* akan memproses data yang telah diolah secara otomatis dan menyimpannya di dalam *database*. Hanya atribut Luas Wilayah dan Jumlah Penerima yang ditangani dari keseluruhan data.

2.5.5 Evaluation

Data dinilai untuk memastikan kebenaran informasi yang dihasilkan setelah metode data mining memprosesnya. Pada tahap ini, kualitas kluster yang dihasilkan diuji dengan menggunakan pendekatan Metode *Elbow*.

3. HASIL DAN PEMBAHASAN

3.1 Algoritma *K-Means Clustering*

Teknik *K-Means Clustering* akan digunakan untuk mengilustrasikan proses pengelompokan dalam sub bab ini. Data penerima bantuan pangan Provinsi Banten, yang berjumlah 231 baris data, digunakan pada Tabel 1 untuk pemodelan *Clustering*. Berikut ini adalah contoh dari Tabel 1 data tersebut:

Tabel 1. Data Clustering

No	Provinsi	Kab/Kota	Kecamatan	Kelurahan	Luas Wilayah	Jumlah Penerima	Latitude	Longitude
1.	Banten	Kota Tangerang	Batucapeper	Batucapeper	138	563	-6.159398	106.668888
2.	Banten	Kota Tangerang	Batucapeper	Batujaya	142	427	-6.154198	106.662504
3.	Banten	Kota Tangerang	Batucapeper	Batusari	131	619	-6.155206	106.657505
...
230.	Banten	Kab. Tangerang	Tigaraksa	Cisereh	371	369	-6.231730703	106.4569089
231.	Banten	Kab. Tangerang	Tigaraksa	Tigaraksa	287	210	-6.269014155	106.466436

Menemukan jumlah *cluster* adalah tahap pertama. Prosedur komputasi ini menentukan tiga *cluster*. Setelah jumlah *Cluster* ditetapkan, pilih nilai *Centroid* awal secara acak. Pada proses perhitungan ini, nilai *Centroid* awal yang dipilih dapat dilihat pada Tabel 2 di bawah ini:

Tabel 2. Centroid Awal

Centroid	Kabupaten	Kecamatan	Kelurahan	Luas Wilayah	Jumlah Penerima
C1	Kota Tangerang	Batucapeper	Batucapeper	138	563
C2	Kota Tangerang	Batucapeper	Batujaya	142	427
C3	Kota Tangerang	Batucapeper	Batusari	131	619

Setelah nilai *Centroid* ditentukan, gunakan rumus Euclidean Distance untuk menghitung jarak setiap titik data dari setiap *Centroid*. Hal ini akan menunjukkan ke klaster mana titik data tersebut terkait. Berikut ini adalah prosedur penghitungannya, dan Tabel 3 menjelaskan hasilnya:

KELURAHAN BATUCEPER

$$\text{Jarak ke C1} = \sqrt{(138 - 138)^2 + (563 - 563)^2} = 0$$

$$\text{Jarak ke C2} = \sqrt{(138 - 142)^2 + (563 - 427)^2} = 136.0588108$$

$$\text{Jarak ke C3} = \sqrt{(138 - 131)^2 + (563 - 619)^2} = 56.43580424$$

Tabel 3. Hasil Perhitungan Iterasi ke-1

NO.	KABUPATEN	KECAMATAN	KELURAHAN	C1	C2	C3	HASIL
1.	Kota Tangerang	Batucapeper	Batucapeper	0	136.0588108	56.43580424	C1
2.	Kota Tangerang	Batucapeper	Batujaya	136.0588108	0	192.314846	C2
3.	Kota Tangerang	Batucapeper	Batusari	56.43580424	192.314846	0	C3
...
231.	Kab Tangerang	Tigaraksa	Tigaraksa	383.1579309	260.9865897	437.7407909	C2

Selanjutnya, persamaan digunakan untuk menghitung nilai *Centroid* untuk iterasi berikutnya, dan hasilnya ditampilkan dalam Tabel 4 berikut.

Tabel 4. Centroid Iterasi ke-2

Centroid	Luas Wilayah	Jumlah Penerima
C1	359.08	560.52
C2	285.4678899	300.3211009
C3	238.8453608	954.3814433

Hasil dari Iterasi 2 dapat dilihat pada Tabel 5 di bawah ini.

Tabel 5. Hasil Perhitungan Iterasi ke-2

NO.	KABUPATEN	KECAMATAN	KELURAHAN	C1	C2	C3	HASIL
1.	Kota Tangerang	Batucapeper	Batucapeper	221.093909	301.2423983	404.164843	C1

NO.	KABUPATEN	KECAMATAN	KELURAHAN	C1	C2	C3	HASIL
2.	Kota Tangerang	Batuceper	Batujaya	254.855482	191.3911672	536.199786	C2
3.	Kota Tangerang	Batuceper	Batusari	235.4578450	354.1420191	352.2943859	C1
...
231.	Kab Tangerang	Tigaraksa	Tigaraksa	357.8544352	90.33409451	745.9373984	C2

Jika hasil *Clustering* tetap mengalami perubahan, langkah berikutnya adalah menentukan nilai *centroid* untuk iterasi berikutnya. Nilai *centroid* untuk iterasi ketiga ditunjukkan pada tabel 6 berikut:

Tabel 6. *Centroid* Iterasi ke-3

<i>Centroid</i>	LuasWilayah	JumlahPenerima
C1	348.57746478873	569.45070422535
C2	245.67777777778	271.57777777778
C3	234.3	1063.5714285714

Hasil dari Iterasi 3 dapat dilihat pada Tabel 7 di bawah ini.

Tabel 7. Hasil Perhitungan Iterasi ke-3

NO.	KABUPATEN	KECAMATAN	KELURAHAN	C1	C2	C3	HASIL
1.	Kota Tangerang	Batuceper	Batuceper	210.676245	310.6789587	509.75037528	C1
2.	Kota Tangerang	Batuceper	Batujaya	250.931170	186.8291967	643.22816610	C2
3.	Kota Tangerang	Batuceper	Batusari	223.148125	365.8595266	456.41499219	C1
...
231.	Kab Tangerang	Tigaraksa	Tigaraksa	364.686979	74.15759411	855.19674559	C2

Hasil perhitungan yang sama ditemukan di iterasi ke-18 dan ke-19 dalam proses perhitungan ini hasil *cluster* menunjukkan bahwa setiap data tidak bergerak, jadi iterasi berikutnya tidak perlu dilakukan. Karena terlalu banyak, proses perhitungan tidak disertakan secara keseluruhan. Ketika pusat-pusat kluster tidak berubah secara signifikan atau jumlah iterasi yang dialokasikan telah selesai, metode *K-Means* mencapai kondisi konvergensi, di mana proses iterasi dihentikan. Konfigurasi pusat kluster akhir yang mengikuti prosedur iterasi lengkap menghasilkan hasil pengelompokan. Dengan demikian, kita dapat menguraikan dan memeriksa hasil klusterisasi pada tahap stabil terakhir dari iterasi setelah fase konvergensi. Hasil *Clustering* yang diperoleh dari pengujian aplikasi dengan 3 *Cluster* pada 231 baris data didapatkan bahwa 90 kelurahan dengan nilai luas wilayah sedang dan jumlah penerima sedang, 116 kelurahan dengan nilai luas wilayah tinggi dan jumlah penerima rendah, 25 kelurahan dengan nilai luas wilayah rendah dan jumlah penerima tinggi.

3.2 Evaluasi Hasil *Clustering* Menggunakan Metode *Elbow*

Tiga *cluster* adalah bilangan ideal yang dapat dihasilkan, menurut perhitungan SSE dari teknik *Elbow* pada Tabel 8, hal ini karena siku dan selisih nilai terbesar berada pada kluster 3 yang menunjukkan bahwa nilai SSE dengan perbedaan tertinggi ditemukan pada $k = 3$.

Tabel 8. Hasil Perhitungan Tabel SSE

JUMLAH KLASTER	SSE	SELISIH
2	18272082.459251	0
3	12335106.149808	5936976.3094425
4	9836328.7975855	2498777.3522229
5	7386589.0487165	2449739.748869
6	6345026.7258515	1041562.3228651
7	4680318.8832249	1664707.8426265
8	4303396.4602558	376922.42296907

4. KESIMPULAN

Program berbasis web ini mengelompokkan kelurahan-kelurahan di Provinsi Banten berdasarkan karakteristik yang dimiliki untuk pendistribusian bantuan pangan beras BULOG secara efektif dengan menggunakan metode *K-Means*. Distribusi bantuan pangan di setiap kelurahan terlihat jelas dengan menggunakan karakteristik Luas Wilayah dan Jumlah Penerima sebagai dasar pengelompokan. Hasil *Clustering* yang diperoleh dari pengujian aplikasi dengan 3 *Cluster* pada 231 baris data didapatkan bahwa 90 kelurahan dengan nilai luas wilayah sedang dan jumlah penerima sedang, 116 kelurahan dengan nilai luas wilayah tinggi dan jumlah penerima rendah, 25

kelurahan dengan nilai luas wilayah rendah dan jumlah penerima tinggi. Aplikasi ini berhasil menggunakan pendekatan *Elbow* untuk penilaian kualitas kluster. Dengan menampilkan SSE (*Sum of squared errors*) untuk setiap jumlah kluster yang dievaluasi, pendekatan *Elbow* membantu dalam mengidentifikasi jumlah kluster yang ideal. Temuan uji *Elbow* dapat digunakan untuk menentukan jumlah kluster yang ideal untuk pemetaan kelurahan di Provinsi Banten. Jumlah *Cluster* yang optimal untuk mengelompokkan kelurahan di Provinsi Banten yaitu 3 *Cluster* dengan nilai SSE 5936976.309442.

DAFTAR PUSTAKA

- [1] A. Ikhwan and N. Aslami, "Implementasi Data Mining untuk Manajemen Bantuan Sosial Menggunakan Algoritma K-Means," *Jurnal Teknologi Informasi*, vol. 4, no. 2, pp. 208–217, 2020, doi: 10.36294/jurti.v4i2.2103.
- [2] M. R. Muttaqin and M. Defriani, "Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa," *ILKOM Jurnal Ilmiah*, vol. 12, no. 2, pp. 121–129, 2020, doi: 10.33096/ilkom.v12i2.542.121-129.
- [3] S. N. Br Sembiring, H. Winata, and S. Kusnasari, "Pengelompokan Prestasi Siswa Menggunakan Algoritma K-Means," *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, vol. 1, no. 1, p. 31, 2022, doi: 10.53513/jursi.v1i1.4784.
- [4] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, "Penerapan Algoritma K-Means Untuk *Clustering* Data Obat-Obatan," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [5] Z. Nabila, A. R. Isnain, Permata, and Z. Abidin, "Analisis Data Mining Untuk *Clustering* Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means," *Jurnal Teknologi dan Sistem Informasi (JTSI)*, vol. 2, no. 2, p. 100, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>
- [6] I. Kamila, U. Khairunnisa, and M. Mustakim, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 5, no. 1, p. 119, 2019, doi: 10.24014/rmsi.v5i1.7381.
- [7] H. Priyatman, F. Sajid, and D. Haldivany, "Klasterisasi Menggunakan Algoritma K-Means *Clustering* untuk Memprediksi Waktu Kelulusan Mahasiswa," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 5, no. 1, p. 62, 2019, doi: 10.26418/jp.v5i1.29611.
- [8] N. Dwitri, J. A. Tampubolon, S. Prayoga, F. I. R.H Zer, and D. Hartama, "Penerapan Algoritma K-Means Dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 Di Indonesia," *Jurnal Teknologi Informasi*, vol. 4, no. 1, pp. 128–132, 2020, doi: 10.36294/jurti.v4i1.1266.
- [9] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," *Jurnal Tekno Kompak*, vol. 15, no. 2, p. 25, 2021, doi: 10.33365/jtk.v15i2.1162.
- [10] N. A. Maori, "Metode Elbow Dalam Optimasi Jumlah *Cluster* Pada K-Means *Clustering*," *Jurnal SIMETRIS*, vol. 14, 2023.
- [11] Al Fahrozi, A., Insani, F., Budianita, E. and Afrianty, I., "Implementasi Algoritma K-Means dalam Menentukan *Clustering* pada Penilaian Kepuasan Pelanggan di Badan Pelatihan Kesehatan Pekanbaru." *Indonesian Journal of Innovation Multidiscipliner Research*, 1(4), pp.474-492, 2023.