

METODE *ELBOW* DALAM OPTIMASI JUMLAH *CLUSTER* PADA *K-MEANS CLUSTERING*

Nadia Annisa Maori

Program Studi Teknik Informatika
Universitas Islam Nahdlatul Ulama Jepara
Email: nadia@unisnu.ac.id

Evanita

Program Studi Teknik Informatika
Universitas Muria Kudus
Email: evanita@umk.ac.id

ABSTRAK

K-Means clustering merupakan salah satu strategi yang digunakan dalam analisis data dan *machine learning* untuk mengelompokkan data menjadi beberapa kelompok (*cluster*) berdasarkan kemiripan fitur atau atributnya. Metode ini bertujuan untuk meminimalkan jarak antara data dalam satu kelompok dan memaksimalkan jarak antara kelompok yang berbeda. Metode *elbow* merupakan sebuah metode yang diterapkan pada *k-means clustering* untuk menentukan jumlah optimal dari *cluster* yang akan dibentuk. Metode *elbow* membantu dalam menentukan jumlah *cluster* yang tepat untuk data tertentu. Data yang digunakan dalam penelitian ini adalah data penerima Program Keluarga Harapan (PKH) di wilayah Kabupaten Jepara tahun 2022. Tujuan penelitian ini adalah untuk mendapatkan jumlah *cluster* optimal dalam melakukan pengelompokkan penerima PKH, sehingga dapat mengetahui desa yang paling banyak dijumpai pada penerima PKH di Kabupaten Jepara, agar penerima PKH yang mendapatkan lebih optimal dan tepat sasaran. Hasil penelitian setelah dilakukan pengujian dari jumlah $k=2$ sampai dengan $k=6$, didapatkan jumlah *cluster* terbaik menggunakan optimasi metode *elbow* terdapat pada jumlah *cluster* sebanyak 3 ($k=3$) dan didukung dengan nilai *Davies Bouldin Index* (DBI) sebesar 0.554. Pada *cluster_0* terdapat 63 desa, *cluster_1* terdapat 127 desa, dan *cluster_2* terdapat 5 desa yang memiliki penerima PKH lebih dari 900 rumah tangga dan lansia menjadi komponen penerima PKH Kabupaten Jepara yang paling dominan. Hasil pengelompokkan ini diharapkan dapat memudahkan dalam menentukan penerima PKH yang tepat sasaran.

Kata kunci: *k-means, clustering, elbow*

ABSTRACT

K-Means clustering is a strategy used in data analysis and machine learning to group data into several groups (*clusters*) based on the similarity of features or attributes. This method aims to minimize the distance between data in one group and maximize the distance between different groups. The elbow method is a method applied to *k-means clustering* to determine the optimal number of clusters to be formed. The Elbow method helps in determining the right number of clusters for certain data. The data used in this research is Program Keluarga Harapan (PKH) beneficiary data in the Jepara Regency area in 2022. The purpose of this study was to obtain the optimal number of clusters for grouping PKH beneficiaries, to find out the villages with the most PKH beneficiaries in Jepara Regency, so that PKH beneficiaries get more optimal and on target. The results after testing the number of $k = 2$ to $k = 6$, the best number of clusters using the optimization of the elbow method is found in the number of clusters of 3 ($k = 3$) and is supported by the Davies Bouldin Index (DBI) value of 0.554. In *cluster_0* there are 63 villages, *cluster_1* there are 127 villages, and *cluster_2* there are 5 villages which have PKH beneficiaries of more than 900 households and the elderly are the most dominant PKH beneficiary component in Jepara Regency. The results of this grouping are expected to make it easier to determine PKH beneficiaries who are on target.

Keywords: *k-means, clustering, elbow*

1. PENDAHULUAN

Indonesia merupakan salah satu negara berkembang. Negara berkembang diartikan sebagai suatu negara dengan kesejahteraan material yang rendah. Kelompok negara berkembang dapat ditandai dengan beberapa hal, seperti rendahnya usia harapan hidup (51 tahun), tingginya kematian bayi dan anak, kemiskinan dimana-mana, pengangguran, distribusi pendapatan tidak merata dan tingkat kecerdasan dan pendidikan yang masih rendah.

Kemiskinan merupakan ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur menurut garis kemiskinan. Garis kemiskinan adalah nilai pengeluaran kebutuhan minimum makanan (setara 2100 kilo kalori per kapita per hari). Oleh sebab itu kemiskinan salah satu permasalahan serius yang selalu dihadapi oleh setiap negara, termasuk di Indonesia. Keberadaan penduduk miskin di Indonesia selalu mengalami naik turun, hal itu dapat dilihat dari data Badan Pusat Statistik (BPS) dimana presentase kemiskinan pada periode September 2015 hingga September 2021 terus mengalami penurunan (dari 11.13% menjadi 9.71%), namun dalam jangka waktu tersebut, tepatnya pada periode September 2019 hingga September 2020 mengalami peningkatan (dari 9.22% menjadi 10.19%) [1]. Salah satu penyebab meningkatnya pada tahun 2020 adalah adanya kebijakan yang diambil Pemerintah untuk mencegah penyebaran COVID-19, yang mana salah satu kebijakan tersebut yaitu Pembatasan Sosial Berskala Besar (PSBB) yang membuat perekonomian di Indonesia mengalami penurunan, karena banyaknya perusahaan yang tidak dapat melakukan kegiatan produksi.

Kabupaten Jepara yang berada di wilayah Provinsi Jawa Tengah dengan 16 Kecamatan, menduduki peringkat ketiga terendah dalam hal penduduk miskin di wilayah Jawa Tengah setelah Kota Semarang dan Kota Salatiga dengan presentase 7.44% per Maret 2021. Presentase penduduk miskin yang berada di Kabupaten Jepara per Maret 2021 mengalami peningkatan 0.27% dari yang sebelumnya sebesar 7.17% [2]. Kabupaten Jepara terdiri dari 16 Kecamatan, 11 Kelurahan dan 184 Desa [3]. Berdasarkan data yang terdapat pada laman Badan Pusat Statistik (BPS) Jepara, jumlah penduduk yang tinggal di Kabupaten Jepara sejumlah 1.188.510 jiwa yang terdiri dari 588.821 jiwa berjenis kelamin laki-laki dan 590.708 jiwa berjenis kelamin perempuan. Beberapa faktor yang mempengaruhi tingkat kemiskinan di Kabupaten Jepara adalah faktor jumlah penduduk dan faktor pendidikan [4]. Berdasarkan rasio banyaknya penduduk per kilometer persegi, Kabupaten Jepara termasuk padat penduduk. Jumlah penduduk yang tinggi dapat menjadi faktor yang mempengaruhi tingkat kemiskinan [5]. Jika jumlah penduduk melebihi kapasitas ekonomi atau lapangan kerja yang tersedia di suatu daerah, dapat terjadi persaingan yang ketat untuk sumber daya yang terbatas, seperti pekerjaan dan penghasilan. Hal ini dapat menyebabkan tingkat pengangguran dan kemiskinan yang tinggi. Selain itu, jika infrastruktur dan sumber daya publik tidak cukup memadai untuk memenuhi kebutuhan semua penduduk, maka kemungkinan tingkat kemiskinan akan meningkat. Tingkat pendidikan juga merupakan faktor penting yang mempengaruhi kemiskinan. Pendidikan yang baik memberikan kesempatan bagi individu untuk memperoleh keterampilan dan pengetahuan yang diperlukan dalam dunia kerja. Jika tingkat pendidikan rendah, peluang untuk mendapatkan pekerjaan yang baik dan berpenghasilan tinggi menjadi terbatas. Selain itu, pendidikan yang rendah juga dapat mempengaruhi produktivitas dan kualitas tenaga kerja, yang pada gilirannya dapat berdampak negatif pada perekonomian lokal.

Fakta tentang hasil diatas menunjukkan bahwa kemiskinan merupakan sebuah masalah yang harus ditangani dengan serius. Dari sisi pemerintah, dalam rangka menanggulangi kemiskinan tersebut adalah adanya berbagai program bantuan sosial salah satu program penanggulangan kemiskinan tersebut adalah Program Keluarga Harapan (PKH). PKH merupakan salah satu program yang dikelola oleh Kementerian Sosial Republik Indonesia dan bekerja sama dengan instansi terkait, seperti Badan Pusat Statistik (BPS), untuk mengidentifikasi dan menyalurkan keluarga miskin yang memenuhi syarat. Keberadaan PKH bertujuan untuk meningkatkan kondisi sosial ekonomi, meningkatkan taraf pendidikan, serta meningkatkan status kesehatan, gizi ibu hamil dan balita bagi Rumah Tangga Sangat Miskin (RTSM). Pada akhirnya, PKH diharapkan tidak hanya sekedar

menurunkan angka kemiskinan dan meningkatkan sumber daya manusia terutama pada kelompok masyarakat miskin, namun ditujukan untuk memutus rantai kemiskinan itu sendiri.

PKH didesain sebagai program perlindungan sosial yang holistik, yang tidak hanya memberikan bantuan keuangan langsung, tetapi juga melibatkan pendampingan dan pemberdayaan. Melalui program ini, keluarga miskin diberikan akses ke bantuan sosial, layanan kesehatan, pendidikan, dan pelatihan keterampilan sesuai dengan Permensos 1 Tahun 2018 tentang Program Keluarga Harapan. Bantuan keuangan PKH diberikan dalam bentuk tunai dan ditujukan untuk memenuhi kebutuhan dasar seperti makanan, pendidikan, kesehatan, dan perawatan. Jumlah bantuan yang diberikan bervariasi tergantung pada kondisi dan kebutuhan keluarga penerima. Selain bantuan keuangan, PKH juga melibatkan program pendampingan dan pemberdayaan. Program pendampingan mencakup pendidikan kesehatan, peningkatan akses ke layanan pendidikan, serta pembentukan perilaku hidup sehat dan produktif. Sementara itu, program pemberdayaan bertujuan untuk meningkatkan kapasitas keluarga miskin melalui pelatihan keterampilan dan pemberian akses terhadap program-program pemberdayaan ekonomi. Penerima PKH adalah 1) keluarga dengan tingkat kemiskinan yang telah ditetapkan oleh pemerintah, 2) keluarga dengan anggota yang berada dalam kondisi rentan, seperti anak-anak, ibu hamil, lansia, dan penyandang disabilitas, dan 3) keluarga yang memenuhi kriteria pilihan wilayah yang ditetapkan oleh pemerintah. Bantuan sosial PKH dicairkan sebanyak 4 kali dalam setahun, pada bulan Januari, April, Juli dan Oktober.

Dalam upaya penanganan kemiskinan sejatinya pemerintah ataupun masyarakat saling memiliki peran penting di dalamnya. Banyak cara dan upaya yang dapat dilakukan dalam menghadapi permasalahan kemiskinan, salah satunya yaitu dengan melakukan pengelompokan rumah tangga penerima PKH sesuai dengan karakteristik ataupun kategori yang dimiliki. Dari pengelompokan rumah tangga penerima PKH, diharapkan dapat diketahui kategori apa yang banyak dijumpai dari penerima PKH tersebut agar pemberian bantuan dapat dioptimalkan sesuai dengan harapan masyarakat yang membutuhkan, karena dalam hal penyaluran bantuan, masih terdapat beberapa tantangan yang perlu diatasi, seperti ketidakakuratan data penerima, terhentinya bantuan, ketidaksesuaian data penerima, serta beberapa masalah teknis dan fundamental lainnya [6].

Teknik *clustering* adalah suatu metode dalam analisa data untuk menyelesaikan persoalan tentang pengelompokan data. Salah satu metode yang digunakan adalah *K-Means*, yang mempunyai kemampuan dalam klasifikasi data yang cukup besar jumlahnya cukup relatif cepat dan efisien [9]. Kekurangan metode ini adalah penentuan jumlah *K cluster* yang baik digunakan. Walaupun mampu menyelesaikan data dengan jumlah banyak, namun masih belum efisien dalam mengelompokkan dokumen secara tepat [10]. Oleh sebab itu, perlu adanya metode untuk menentukan *K cluster* yang baik. Salah satu metode yang dapat digunakan adalah metode *elbow*, kemampuannya dalam menganalisa jumlah *K cluster* dapat digunakan untuk mengoptimalkan kinerja dari *K-Means*.

Berdasarkan penjelasan diatas, peneliti akan melakukan optimasi *K-Means Clustering* menggunakan metode *elbow* pada data dari PKH. Hasil yang akan dicapai pada penelitian ini adalah diketahuinya kategori dan lokasi yang paling banyak dijumpai pada penerima PKH di Kabupaten Jepara, sehingga dapat mempermudah tim penyeleksi memberikan bantuan sesuai dengan kriteria yang telah ditentukan. Harapannya adalah untuk memastikan agar penerima Program Keluarga Harapan (PKH) dapat tepat sasaran dan memberikan bantuan kepada keluarga miskin dan rentan secara efektif.

2. METODOLOGI PENELITIAN

Metode yang digunakan dalam penelitian ini adalah menggunakan algoritma *K-Means Clustering* dan metode *elbow* sebagai metode optimasi. Metode *elbow* digunakan sebagai penentu jumlah *k* pada proses *clustering*. Langkah-langkah pada penelitian yang akan dilakukan dapat dilihat pada Gambar 1.



Gambar 1. Langkah Metode Penelitian

Penjelasan dari langkah diatas adalah sebagai berikut:

1. **Pengumpulan Data**
Data yang digunakan dalam penelitian ini adalah data penerima Program Keluarga Harapan (PKH) termin I Kabupaten Jepara Tahun 2022 sejumlah 65.455 rumah tangga, yang terbagi menjadi 11 kelurahan dan 195 desa.
2. **Pre-Processing Data**
Data yang terkumpul akan dilakukan proses pengolahan data awal untuk dapat menghasilkan data yang lebih akurat. Proses pengolahan data awal yang digunakan pada penelitian ini adalah *Data Cleaning*, *Data Reduction*, *Data Transformation* dan *Atributte Selection*. Pada proses *attribute selection* didapatkan atribut yang digunakan adalah jumlah rumah tangga penerima PKH, jumlah anak usia dini, jumlah anak usia SD, jumlah anak usia SMP, jumlah anak usia SMA, jumlah lansia dan jumlah disabilitas.
3. **Pengujian Metode**
Pada penelitian ini metode yang digunakan adalah metode *clustering* menggunakan algoritma *k-means clustering* dan metode optimasi menggunakan metode *elbow*. Metode *elbow* digunakan untuk menentukan jumlah *k-cluster* terbaik pada proses clustering melalui perhitungan SSE. Pada proses pengujian menggunakan algoritma *k-means* dilakukan pengujian sebanyak 5 kali, yaitu dari nilai $k=2$ sampai dengan $k=6$.
4. **Evaluasi**
Hasil evaluasi *cluster* menggunakan SSE dari metode *elbow* berdasarkan jumlah k uji.
5. **Hasil**
Hasil dari penelitian ini adalah jumlah *cluster* optimal pada proses *clustering*.

2.1. Algoritma K-Means Clustering

Teknik *clustering* adalah suatu metode dalam analisa data untuk menyelesaikan persoalan tentang pengelompokan data. Salah satu metode *clustering* adalah *K-Means*, yang mempunyai kemampuan dalam klasifikasi data yang cukup besar jumlahnya relatif cepat dan efisien. Kekurangannya, hasil *clustering* dengan menerapkan algoritma *K-Means* bergantung pada penentuan awal pusat *cluster*, sehingga hasil perhitungan *clustering* dengan metode *K-Means* akan baik jika penentuan pusat cluster tepat [9].

Algoritma *K-Means* adalah metode klasterisasi yang dapat mempartisi data ke dalam beberapa kelompok (*cluster*). *K-Means* yaitu salah satu dari metode pengelompokan data nonhierarki (sekatan) yang dapat mempartisi data kedalam bentuk dua kelompok ataupun lebih. Metoda tersebut akan mempartisi data kedalam suatu kelompok dimana data yang berkarakteristik sama akan dimasukkan kedalam satu kelompok sama sedangkan data yang memiliki karakteristik yang berbeda akan dikelompokkan kedalam kelompok lainnya [14]. Tujuan dari pengelompokan yaitu untuk meminimalkan dari fungsi objektif yang diset dalam proses pengelompokan, pada umumnya akan berusaha meminimalkan variasi didalam suatu kelompok dan memaksimalkan variasi antar kelompok. Manfaat dari *clustering* yaitu sebagai Identifikasi Objek (*Recognition*) misal dalam bidang *Image Processing*, *Computer Vision* ataupun *robot vision* [15]. Tidak hanya itu manfaat lain yaitu sebagai SPK (Sistem Pendukung Keputusan) serta data mining semacam pemetaan daerah, segmentasi pasar, manajemen marketing dan sebagainya [16]. Algoritma *K-Means* merupakan salah satu dari algoritma yang banyak digunakan dalam pengelompokan karena kesederhanaan dan efisiensi dan diakui sebagai salah satu dari 10 algoritma data mining teratas oleh IEEE [9].

Algoritma *K-means* bersifat sederhana dalam pengimplementasian dan menjalankan, pengerjaan relatif cepat, dapat juga beradaptasi dan sudah sering digunakan. Langkah-langkah yang terdapat pada Algoritma *K-Means* yaitu:

1. Menentukan nilai *K* sebagai jumlah *cluster* yang akan dibuat.
2. Menentukan titik pusat cluster (*K centroid*) awal secara acak.
3. Setelah mendapatkan titik pusat *cluster*, selanjutnya adalah menghitung jarak setiap data ke masing-masing *centroid* dengan menggunakan rumus korelasi antar dua objek, seperti yang dilihat pada rumus *Euclidean Distance*.

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + \dots (n_i - n_j)^2} \quad (1) [8]$$

Keterangan:

D_{ij} = Jarak antara *i* dan *j*

$X_i X_i$ = Koordinat *x* objek

$X_j X_j$ = Koordinat *x* pusat

$Y_i Y_i$ = Koordinat *y* objek

$Y_j Y_j$ = Koordinat *y* pusat

4. Setelah mendapatkan hasil dari perhitungan jarak setiap data, langkah selanjutnya adalah menggabungkan setiap data yang dilihat berdasarkan jarak terdekat antara data dan cluster.
5. Menentukan centroid baru dengan menghitung nilai rata-rata berdasarkan data pada cluster yang serupa. Untuk perhitungannya dapat menggunakan persamaan berikut:

$$C_i = \frac{x_1 + x_2 + x_3 + \dots + x_n}{\sum x} \quad (2) [8]$$

$$C_i = \frac{x_1 + x_2 + x_3 + \dots + x_n}{\sum x}$$

Keterangan:

x_1 x_1 = nilai record data ke 1

x_2 x_2 = nilai record data ke 2

$\sum x$ $\sum x$ = jumlah record data.

Algoritma *k-means* merupakan algoritma yang dilakukan secara berulang-ulang hingga menghasilkan *cluster* sesuai kelompok pada data-data yang sama dan yang berbeda ditempatkan pada kelompok yang berbeda [10].

2.2. Metode Elbow

Metode *elbow* merupakan salah satu metode yang dapat digunakan untuk menentukan jumlah cluster terbaik, yaitu dengan cara melihat presentase setiap cluster yang akan membentuk siku pada

suatu titik tertentu. Metode *elbow* biasa disajikan dalam bentuk grafik untuk mengetahui lebih jelas siku yang terbentuk. Tujuan dari metode *elbow* adalah untuk memilih nilai k yang kecil dan masih memiliki nilai *withinss* yang rendah. Nilai k pada kombinasi siku dengan *k-means* adalah grafik hubungan *cluster* dengan penurunan *error*. Jumlah *cluster* k yang dihasilkan dari pengujian dengan *k-means* dievaluasi dengan teknik SSE. SSE (*Sum of Square Error*) merupakan rumus yang digunakan untuk mengukur perbedaan antara data yang telah dilakukan sebelumnya [11]. Adapun persamaan yang digunakan dalam metode *elbow* yaitu nilai total WSS (*Within Cluster Sum of Squares*) atau biasa disebut dengan SSE (*Sum Square Error*) dengan rumus:

$$SSE = \sum_{k=1}^K \sum_{x_i \in C_k} (x_i - \varphi_k)^2 \quad (3) \text{ [13]}$$

Keterangan:

C_k C_k = K *cluster* yang terbentuk

k k = banyak *cluster*

x_i x_i = data x pada fitur ke- i

φ_k φ_k = rata-rata K *cluster* pada nilai k ($k=1,2,3,...,K$)

3. HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini adalah data penerima PKH yang berada pada Kabupaten Jepara yang terdiri dari 16 Kecamatan. Data yang tersedia sebanyak 65.455 rumah tangga penerima PKH yang terbagi 11 Kecamatan dan 195 Desa. Data awal disajikan pada Tabel 1 dibawah ini.

Tabel 1. Data Awal Penerima PKH Kabupaten Jepara

Kecamatan	Desa	A	B	C	D	E	F	G
Bangsri	Bangsri	619	105	190	135	147	367	11
	Banjar Agung	263	58	130	89	90	94	0
	Banjaragung	105	0	0	1	0	104	2
	Banjaran	532	83	170	128	130	293	15
	Bondo	662	89	200	139	152	374	24
	Guyangan	613	109	205	135	162	395	7
	Jerukwangi	503	88	155	119	125	265	10
	Kedungleper	352	63	142	84	79	175	6
	Kepuk	354	58	109	65	80	188	11
	Papasan	141	28	44	23	40	71	4
	Srikandang	454	83	156	81	99	243	6
	Tengguli	615	117	190	129	152	394	16
	Wedelan	353	56	121	93	85	182	16
...	...	...	...	...	...	...	...	...
Welahan	Welahan	378	56	152	104	102	196	8

Keterangan:

A = Jumlah Anggota Rumah Tangga Penerima PKH

B = Jumlah Anak Usia Dini (AUD)

C = Jumlah Anak Sekolah Dasar (SD)

D = Jumlah Anak Sekolah Menengah Pertama (SMP)

E = Jumlah Anak Sekolah Menengah Atas (SMA)

F = Jumlah Lansia

G = Jumlah Disabilitas

Data yang sudah dikumpulkan kemudian dilakukan pengolahan data awal untuk mendapatkan data yang akurat. Data hasil pengolahan data awal ditunjukkan pada Tabel 2.

Tabel 2. Hasil Pengolahan Data Awal

No	Desa	A	B	C	D	E	F	G
1	Bakalan	120	17	39	30	48	62	10
2	Balong	376	46	82	61	63	294	9
3	Bandengan	369	76	107	91	82	188	11
4	Bandung	107	14	40	28	30	64	4
5	Bandungharjo	259	26	45	32	45	207	15
6	Bandungrejo	251	69	92	67	70	112	8
7	Bangsri	619	105	190	135	147	367	11
8	Banjaragung	368	58	130	90	90	198	2
9	Banjaran	532	83	170	128	130	293	15
10	Bantrung	497	91	196	113	114	240	10
...	...	...	...	...	...	...	...	...
195	Welahan	378	56	152	104	102	196	8

Data hasil pengolahan data awal yang ditunjukkan pada Tabel 2. Selanjutnya dilakukan proses *clustering* menggunakan algoritma *k-means clustering*. Hasil pengelompokan menggunakan algoritma *k-means clustering* ditunjukkan pada Tabel 3.

Tabel 3. Hasil Clustering Menggunakan Algoritma K-Means Clustering

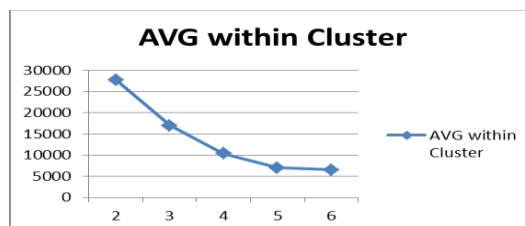
k	2	3	4	5	6
Cluster 1	66 data	127 data	61 data	28 data	28 data
Cluster 2	129 data	63 data	97 data	63 data	63 data
Cluster 3		5 data	32 data	64 data	3 data
Cluster 4			5 data	5 data	64 data
Cluster 5				35 data	35 data
Cluster 6					2 data

Dari hasil Tabel 3. diatas, kemudian dilakukan metode optimasi menggunakan metode *elbow* untuk menentukan jumlah *cluster* terbaik. Berikut adalah hasil perhitungan SSE berdasarkan pengujian nilai $k=2$ sampai dengan $k=6$.

Tabel 4. Hasil Perhitungan SSE tiap Cluster

k	Hasil SSE	Selisih
2	27815,998	0
3	16961,739	10854,26
4	10404,693	6557,046
5	7059,808	3344,885
6	6569,519	490,289

Dari perhitungan SSE metode *Elbow* pada Tabel 4, didapatkan nilai SSE yang memiliki selisih terbesar terdapat pada nilai $k=3$, sehingga jumlah *cluster* optimal yang dapat dibentuk sejumlah 3 *cluster*.

**Gambar 2. Grafik SSE Metode Elbow**

Dari grafik SSE metode *Elbow* pada Gambar 2 terlihat bahwa pada $k=3$ mengalami penurunan paling besar jika dibandingkan jumlah k yang lain. Jumlah *cluster* selain $k=3$ mengalami penurunan yang stabil. Maka dapat disimpulkan jumlah *cluster* optimal yang terbentuk sejumlah 3 *cluster*. Hasil *clustering* dengan nilai $k=3$ dapat ditunjukkan pada Gambar 3 dan Tabel 5.



Gambar 3. Hasil *Clustering*

Tabel 5. Hasil *Clustering* Menggunakan Metode *Elbow* dan Algoritma *K-Means Clustering*

<i>Cluster Model</i>	
<i>Cluster_0</i>	63 items
<i>Cluster_1</i>	127 items
<i>Cluster_2</i>	5 items
Total number of items	195

Pada Tabel 5 menunjukkan jumlah anggota pada setiap *cluster* dengan nilai $k=3$. Dimana, dari jumlah 195 data yang menjadi anggota pada *cluster_0* sejumlah 63 anggota, *cluster_1* sejumlah 127 anggota dan *cluster_2* sejumlah 5 anggota. Hasil untuk setiap *cluster* dari penelitian ini dapat dilihat pada Tabel 6 hingga Tabel 8.

Tabel 6. Hasil Pengelompokkan *Cluster_0*

No	Desa	A	B	C	D	E	F	G
1	Balong	376	46	82	61	63	294	9
2	Bangsri	619	105	190	135	147	367	11
3	Banjaran	532	83	170	128	130	293	15
4	Bantrung	497	91	196	113	114	240	10
5	Banyumanis	514	80	150	112	94	302	42
6	Bategede	406	69	130	132	102	290	21
7	Bawu	661	127	248	146	192	367	11
8	Blingoh	495	54	116	71	94	367	13
9	Bondo	662	89	200	139	152	374	24
10	Bringin	434	57	163	103	107	248	8
...	...	...	...	...	...	...	...	...
63	Kecapi	727	147	289	194	197	375	20

Tabel 6 merupakan hasil pengelompokkan dari *cluster 0* yang mempunyai anggota sebanyak 63 desa. Hasil tersebut menunjukkan bahwa pengelompokkan jumlah rumah tangga penerima PKH pada setiap desa berada diatas 400 rumah tangga dan rata-rata setiap komponen pendidikan (SD, SMP, dan SMA) memiliki jumlah penerima diatas 100, serta pada komponen lansia cukup banyak yang penerima PKH.

Tabel 7. Hasil Pengelompokkan Cluster_1

No	Desa	A	B	C	D	E	F	G
1	Bakalan	120	17	39	30	48	62	10
2	Bandung	107	14	40	28	30	64	4
3	Bandungharjo	259	26	45	32	45	207	15
4	Bandungrejo	251	69	92	67	70	112	8
5	Banyuputih	221	37	87	56	76	100	11
6	Bapangan	133	18	32	34	30	81	1
7	Batealit	312	64	131	74	82	164	7
8	Batukali	156	29	55	41	30	71	5
9	Bendanpete	247	43	81	61	46	141	3
10	Blimbingrejo	302	35	79	56	75	173	58
...	...	...	...	...	...	...	...	...
127	Banjaragung	368	58	130	90	90	198	2

Tabel 7 merupakan hasil pengelompokkan dari *cluster* 1 yang mempunyai sebanyak 127 desa. Hasil tersebut menunjukkan bahwa *cluster_1* mengelompokkan data berdasarkan jumlah rumah tangga penerima PKH (A) berada dibawah 400 rumah tangga dan dari komponen pendidikan, lansia maupun disabilitas masih acak dalam pengelompokkannya.

Tabel 8. Rekapitulasi Hasil Cluster_2

No	Desa	A	B	C	D	E	F	G
1	Cepogo	966	140	301	162	204	626	29
2	Karanggodang	1207	227	480	304	316	562	34
3	Lebak	912	174	339	258	247	498	28
4	Troso	1085	160	315	191	213	753	12
5	Tubanan	1149	118	254	161	216	844	25

Tabel 8 merupakan hasil pengelompokkan dari *cluster* 2 yang mempunyai sebanyak 5 desa. Hasil tersebut menunjukkan bahwa *cluster* ini didominasi oleh desa-desa yang mempunyai penerima PKH diatas angka 900 dan jumlah komponen lansia lebih mendominasi dari komponen pendidikan maupun disabilitas.

4. KESIMPULAN

Penentuan titik pusat *cluster* sangat berpengaruh pada terhadap perhitungan algoritma *k-means*. Dari pengujian menggunakan algoritma *k-means clustering* yang dioptimalkan dengan metode *elbow*, terbentuk jumlah *cluster* yang optimal sebanyak 3 *cluster*. Penentuan $k=3$ menggunakan metode *elbow* ini berdasarkan perhitungan selisih dari rata-rata dari tiap *cluster* yang terbentuk dan nilai *Davies Bouldin Index* (DBI) yang mendekati 0 menunjukkan *cluster* yang terbentuk adalah *cluster* yang optimal. Dalam menentukan jumlah *cluster* menggunakan metode *elbow* mendapatkan hasil yang lebih baik tingkat kemiripan setiap anggotanya dibandingkan dengan menentukan jumlah *cluster* secara acak.

Dari 195 desa dan kelurahan yang berada di Kabupaten Jepara setelah dilakukan *clustering*, bahwa terdapat 5 desa yang memiliki jumlah penerima Program Keluarga Harapan (PKH) yaitu desa Cepogo, Karanggodang, Lebak Troso dan Tubanan, dimana kelima desa tersebut memiliki jumlah rumah tangga penerima PKH diatas 900 rumah tangga dengan komponen yang paling dominan atau banyak dijumpai adalah dari usia lanjut (lansia) pada penerima PKH di masing-masing desa. Hasil pengelompokkan menggunakan algoritma *K-Means Clustering* ini menunjukkan bahwa penerima PKH yang mendapatkan prioritas utama berada pada *cluster* 2. Hasil pengelompokkan ini diharapkan dapat membantu pemerintah dalam menentukan penerima PKH yang akurat dan tepat sasaran sesuai dengan syarat penerima PKH yang telah ditentukan.

DAFTAR PUSTAKA

- [1] B. R. Statistik, “Berita Resmi Statistik,” Bps.Go.Id, 2022, [Online]. Available: <https://jakarta.bps.go.id/pressrelease/2019/11/01/375/tingkat-penghunian-kamar--tpk--hotel--berbintang-dki-jakarta-pada-bulan-september-2019-mencapai-58-97-persen.html>
- [2] Badan Pusat Statistik Daerah Kabupaten Jepara, “Profil Kemiskinan Kabupaten Natuna,” no. 1, pp. 1–8, 2021.
- [3] B. Silvia and F. Azmi, “Analisis Faktor-Faktor Yang Mempengaruhi Persepsi SAK EMKM,” *J. Anal. Bisnis Ekon.*, vol. 17, no. 1, pp. 57–73, 2019.
- [4] E. Syahrani, A. A. Kusumaningdyah, and D. D. Dewa, “Analisis Faktor-Faktor Yang Mempengaruhi Tingkat Kemiskinan Kabupaten/Kota Di Jawa Tengah,” *Geodika J. Kaji. Ilmu dan Pendidik. Geogr.*, vol. 5, no. 2, pp. 247–258, 2021.
- [5] K. U. Albab, “Analisis faktor-faktor yang mempengaruhi tingkat kemiskinan di kabupaten jepara,” *Semin. Nas. KeIndonesiaan III*, pp. 817–825, 2018.
- [6] Zulkifli, Suadi, and Alwi, “Penyalahgunaan Bantuan PKH: Motif dan Dampak (Studi Kasus di Gampong Bluka Teubai Kecamatan Dewantara Kabupaten Aceh Utara),” *J. Sociol. Agama Indones.*, vol. 2, no. 2, pp. 101–110, 2021.
- [7] N. T. Hartanti, “Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 6, no. 2, pp. 82–89, 2020, doi: 10.25077/teknosi.v6i2.2020.82-89.
- [8] Afifi, Wakhid, Nastiti, Dhiya’an Ramadhanty, and Aini, Qurrotul, “Clustering K-Means Pada Data Ekspor (Studi Kasus:PT. Gaikindo),” *Jurnal SIMETRIS.*, vol. 11, no. 1, pp. 45-50, 2020.
- [9] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, “Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [10] M. M. K-means, S. A. Azis, S. Defit, and Y. Yunus, “Jurnal Informatika Ekonomi Bisnis Klasterisasi Dana Bantuan Pada Program Keluarga Harapan (PKH),” vol. 3, pp. 53–59, 2021, doi: 10.37034/infeb.v3i2.66.
- [11] D. Jollyta, S. Efendi, M. Zarlis, and H. Mawengkang, “Optimasi Cluster Pada Data Stunting: Teknik Evaluasi Cluster Sum of Square Error dan Davies Bouldin Index,” *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 918, 2019, doi: 10.30645/senaris.v1i0.100.
- [12] A. T. Rahman, Wiranto, and A. Rini, “Coal Trade Data Clustering Using K-Means (Case Study Pt. Global Bangkit Utama),” *ITSMART J. Teknol. dan Inf.*, vol. 6, no. 1, pp. 24– 31, 2017, doi: 10.20961/ITS.V6I1.11296.
- [13] Dewi, Dewa Ayu Indah Cahya Dewi, and Pramita, Dewa Ayu Kadek, “Analisis Perbandingan Metode Elbow dan Sillhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali,” *Jurnal Matrix*, vol.9, no.3, pp. 102-109, 2019
- [14] Parlaungan, Mustika, Fanisya Alva, Dhika, Harry, “Sistem Rekomendasi Film Menggunakan Metode K-Means Clustering,” *Jurnal SIMETRIS*, vol. 13, no. 2, pp. 1-17

- [15] A. Kroner, M. Senden, K. Driessens, and R. Goebel, "Contextual encoder–decoder network for visual saliency prediction," *Neural Networks*, vol. 129, pp. 261–270, 2020, doi: 10.1016/j.neunet.2020.05.004.
- [16] U. Khalsum, "Implementasi Sistem Pengelompokkan Data Penerima Beasiswa Menggunakan Metode K-MEANS," Universitas Islam Negeri Sultan Syarif Kasim Riau, 2014.