

Penerapan Algoritma *K-Means Clustering* untuk Mengelompokkan Kecamatan di Kabupaten Grobogan Menurut Tingkat Kesejahteraan Keluarga Tahun 2020

Asa Nugrahaini Itsnal Muna^{1*}, Rahmadi Yotenka¹

¹Universitas Islam Indonesia, Jl. Kaliurang KM 14.5, DI. Yogyakarta, Indonesia 55584

*Corresponding Author: 19611139@students.uii.ac.id



P-ISSN: 2986-4178

E-ISSN: 2988-4004

Riwayat Artikel

Dikirim: 03 Januari 2023

Direvisi: 21 Agustus 2023

Diterima: 29 September 2023

ABSTRAK

Kesejahteraan keluarga adalah keadaan di mana kebutuhan dasar, sosial, dan perkembangan keluarga terpenuhi secara maksimal. Kabupaten Grobogan memiliki tingkat kesejahteraan keluarga yang berbeda-beda di setiap kecamatannya. Berbagai upaya dilakukan oleh pemerintah di Kabupaten Grobogan untuk mengatasi masalah-masalah yang berkaitan dengan kesejahteraan keluarga, mulai dari penyaluran bantuan dalam bentuk bantuan konsumtif hingga bantuan yang lebih produktif, namun seringkali terkendala masalah distribusi. Dengan menggunakan analisis *k-means clustering*, 19 kecamatan di Kabupaten Grobogan diklasifikasi dan ditentukan karakteristiknya guna membantu pemerintah dalam melaksanakan kebijakan yang berkaitan dengan masalah kesejahteraan keluarga di setiap kecamatan. Ada dua kelompok berbeda yang diidentifikasi dalam penelitian ini. Kecamatan pada kelompok 1 memiliki tingkat kesejahteraan keluarga yang tinggi untuk keluarga tahap I, II, III, dan III+. Sementara itu kelompok 2 memiliki tingkat kesejahteraan keluarga rendah untuk keluarga tahap I, II, III, dan III+.

Kata Kunci: Kesejahteraan Keluarga, K-Means Clustering, Keluarga Sejahtera

ABSTRACT

Family welfare is a circumstance wherein the own circle of relatives's basic, social, and developmental desires are met to the fullest. Grobogan Regency has one of a kind stages of own circle of relatives welfare in every sub-district. The authorities has made numerous efforts withinside the Grobogan Regency to triumph over troubles associated with own circle of relatives welfare, starting from the distribution of help withinside the shape of consumptive help to extra efficient help, however distribution troubles are regularly hampered. Using k-manner clustering analysis, 19 sub-districts in Grobogan District had been categorized and their traits decided to help the authorities in enforcing rules associated with own circle of relatives welfare problems in every sub-district. There are wonderful companies recognized on this study. The sub-districts in institution 1 have a excessive degree of own circle of relatives welfare for

degrees I, II, III, and III plus families. Meanwhile, institution 2 has a low degree of own circle of relatives welfare for level I, II, III, and III+ families.

Keywords: Family Welfare, K-Means Clustering, Properous Family

1. Pendahuluan

Keluarga sejahtera adalah keluarga yang semua anggotanya menikah secara sah, yang mampu memenuhi kebutuhan dasar material dan spiritualnya, yang bertakwa kepada Tuhan Yang Maha Esa, dan yang mengabdikan diri untuk membina niat baik antara mereka sendiri dan dengan masyarakat yang lebih besar. Keluarga sejahtera tahap I hingga II, keluarga sejahtera tahap III dan III plus, serta keluarga sejahtera tahap III dan IV merupakan keadaan keluarga sejahtera dalam penilaian BKKBN [1]. Untuk mengatasi masalah kesejahteraan keluarga, Pemerintah Kabupaten Grobogan telah melakukan beberapa program, mulai dari bantuan konsumsi hingga dalam bentuk dukungan produktif, namun pendistribusiannya seringkali bermasalah serta penyaluran bantuan tidak merata. Kabupaten Grobogan telah melaksanakan Program Keluarga Harapan (PKH), namun informasi penyaluran tersebut belum optimal karena masih disampaikan dari mulut ke mulut, seperti pada kasus PKH. Teknik analisis klusterisasi dapat digunakan oleh pemerintah untuk mengidentifikasi daerah mana yang memiliki tingkat kesejahteraan keluarga rendah, sehingga dapat memprioritaskan kesejahteraan keluarga di tempat-tempat tersebut.

Clustering merupakan cara terbaik untuk mengidentifikasi karakteristik tingkat kesejahteraan keluarga pada suatu komunitas tertentu. Pemerintah Kabupaten Grobogan akan dapat lebih menyesuaikan inisiatif kesejahteraan keluarga di Grobogan dengan membentuk kelompok dengan karakteristik yang sama. “Penerapan. Algoritma K-Means dalam Pengelompokan Kesejahteraan Rakyat di Kabupaten Karawang” oleh (Fitriani, Padilah, & Sari, 2021) merupakan penelitian yang sebanding dengan penelitian ini. Menurut kesejahteraan penduduknya, Kecamatan Karawang dikelompokkan menjadi tiga. Menggunakan data dari 30 kecamatan, hasil penelitian menunjukkan tujuh kecamatan memiliki tingkat kesejahteraan masyarakat tinggi, delapan kecamatan memiliki tingkat kesejahteraan sedang, dan lima belas kecamatan memiliki tingkat kesejahteraan masyarakat rendah [2].

Akhyar (Akhyar, 2017) membuat model *clustering* untuk kabupaten dan kota di Jawa Timur berdasarkan indikator pembangunan ekonomi. Kabupaten dan kota di Jawa Timur akan diklasifikasikan menurut tingkat perkembangannya dalam penelitian ini. 26 kota dan kabupaten memiliki indikator ekonomi tinggi, sedangkan 12 kota dan kabupaten memiliki indikator ekonomi rendah, menurut kesimpulan penelitian ini, yang meneliti 38 kota dan kabupaten [3]. (Ramdhani, Hoyyi, & Mukid 2015) dalam artikelnya yang berjudul “Pengelompokan Provinsi di Indonesia Berdasarkan Karakteristik Kesejahteraan Rakyat Menggunakan Metode Klaster K-Means”. Tujuannya adalah untuk mengkategorikan provinsi-provinsi di Indonesia berdasarkan tingkat kesejahteraan penduduknya. Berdasarkan penelusuran terhadap 33 provinsi, enam provinsi memiliki tingkat kesejahteraan masyarakat rendah, delapan provinsi memiliki kesejahteraan masyarakat sedang, dan 19 provinsi memiliki tingkat kesejahteraan masyarakat tinggi [4].

Berdasarkan uraian penelitian di atas, belum terdapat penelitian yang menggunakan algoritma *k-means* untuk mengelompokkan kesejahteraan keluarga di Kabupaten Grobogan berdasarkan tingkatannya. Sehingga penelitian ini difokuskan pada

penerapan dari algoritma *k-means* berdasarkan tingkat kesejahteraan keluarga di Kabupaten Grobogan.

2. Metodologi Penelitian

2.1 Data dan Sumber Data

Angka 2021, Badan Pusat Statistik Kabupaten Grobogan, memberikan data sekunder yang digunakan yaitu tingkat keluarga sejahtera (pra sejahtera, KS 1, KS II, KS III, dan KS III+).

2.2 Landasan Teori

Tinjauan pustaka adalah metode penelitian yang digunakan dalam proyek ini. Informasi dapat diperoleh melalui berbagai literatur, diantaranya buku, jurnal, serta internet. Dengan menggunakan analisis *k-means clustering* untuk mengklasifikasikan kecamatan di Kabupaten Grobogan menurut kesejahteraan keluarganya. Berikut ini adalah berbagai tahapannya.

1. Meng-*input* data.
2. Analisis deskriptif.
3. Pengujian asumsi dengan memastikan sampel sudah representatif, pengujian KMO dan pengujian multikolinearitas [5].

Berikut rumus KMO :

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2} \quad i, j = 1, 2, 3, \dots, p \quad (1)$$

dengan :

r_{ij} : koefisien korelasi antara variabel i dan j

a_{ij} : koefisien korelasi parsial variabel i dan j

Untuk mengetahui adanya hubungan linear antara dua variabel atau lebih dapat digunakan dengan uji multikolinearitas. Adanya multikolinearitas dapat dideteksi melalui beberapa cara [6]:

- a. Menghitung koefisien korelasi antar variabel bebas.

$$r = \frac{n \sum XY - \sum X \sum Y}{\sqrt{(n \sum X^2 - (\sum X)^2)(n \sum Y^2 - (\sum Y)^2)}} \quad (2)$$

dengan :

X : variabel independen

Y : variabel dependen

n : banyaknya sampel

- b. Menghitung nilai *tolerance* atau VIF.

$$VIF = \frac{1}{1 - R_i^2} \quad (3)$$

dengan R_i^2 adalah koefisien determinasi

4. Menentukan jumlah *cluster* dengan menggunakan pendekatan metode *within sum square* atau *elbow* dan *silhouette*.
 - a. Metode *Elbow*

$$SSE = \sum_{k=1}^k \sum_{x_i \in S_k} ||x_1 - C_k|| \quad (4)$$

dimana :

k : jumlah *cluster*

x_i : data ke- i

C_k : *centeroid* dari *cluster*

b. Metode *Silhouette*

- 1) Menghitung rata-rata jarak objek dengan semua objek lain yang berada di dalam satu *cluster*.

$$a(i) = \frac{1}{[A] - 1} \sum_{j \in A, j \neq i} d(i, j) \quad (5)$$

- 2) Menghitung rata-rata jarak objek dengan semua objek lain yang berada pada *cluster* lain, kemudian ambil nilai paling minimum.

$$a(i, C) = \frac{1}{[A]} \sum_{j \in C} d(i, j) \quad (6)$$

- 3) Menghitung nilai *silhouette coefficient*.

$$b(i) = \min_{C \neq A} d(i, C) \quad (7)$$

Nilai hasil akhir berada pada kisaran angka -1 hingga 1.

5. Melakukan analisis *k-means*.

Adapun urutan dalam melakukan analisis *cluster* adalah sebagai berikut :

- a. Menetapkan ukuran jarak antar data.

Jarak *Euclidean* digunakan jika tidak terjadi korelasi dengan rumus :

$$d(y, x) = \sqrt{\sum_{k=1}^l (y_k - x_k)^2} ; l = 1, 2, 3, \dots, n \quad (8)$$

dengan :

$d(y, x)$: kuadrat jarak *Euclid* antar objek y dengan objek pada x

y_k : nilai dari objek y pada variabel ke- k

x_k : nilai dari objek x pada variabel ke- k

- b. Melakukan proses standarisasi data jika diperlukan.

Berikut rumus standarisasi data :

$$z = \frac{x - \mu}{\sigma} \quad (9)$$

dimana :

x : nilai data

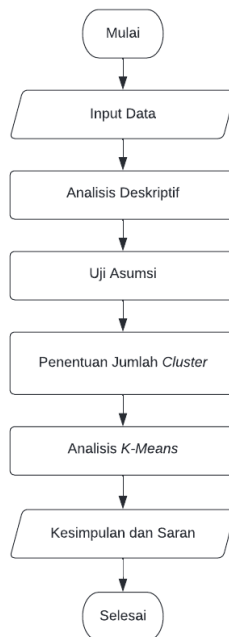
μ : nilai rata-rata

σ : standar deviasi

- c. Proses *clustering* dan penamaan *cluster* yang terbentuk.

6. Menarik kesimpulan dan saran.

Dari penjelasan di atas peneliti menguraikan dalam bentuk diagram alir untuk memudahkan. **Gambar 1** merupakan diagram alir penelitian :



Gambar 1 Diagram Alir Penelitian

3. Hasil dan Pembahasan

3.1. Analisis Deskriptif

Sebelum melakukan pengujian asumsi, terlebih dahulu dilakukan analisis deskriptif. Berikut ini merupakan hasil analisis deskriptif yang disajikan dalam **Tabel 1**.

Tabel 1 Analisis Deskriptif

Statistik Deskriptif	Variabel				
	Pra Sejahtera	Keluarga Sejahtera Tahap I	Keluarga Sejahtera Tahap II	Keluarga Sejahtera Tahap III	Keluarga Sejahtera Tahap III+
<i>Min</i>	4359,0	334,0	547,0	667,0	0,0
Kuartil 1	11379,0	1308,0	1977,0	1280,0	103,0
Median	15466,0	2900,0	3496,0	2131,0	192,0
<i>Mean</i>	15935,0	3077,0	3609,0	3188,0	354,1
Kuartil 3	20795,0	3365,0	4692,0	5134,0	467,0
<i>Max</i>	31868,0	9447,0	8176,0	6999,0	1395,0

Berdasarkan **Tabel 1** diketahui bahwa terdapat 5 variabel yang digunakan, yaitu pra sejahtera, keluarga sejahtera tahap I, keluarga sejahtera tahap II, keluarga sejahtera tahap III, dan keluarga sejahtera tahap III plus. Mayoritas masyarakat di Kabupaten Grobogan memiliki tingkat kesejahteraan keluarga pada tingkatan pra sejahtera yang memiliki nilai analisis deskriptif yang lebih tinggi.



Gambar 2 Bar Chart Tingkat Kesejahteraan Keluarga Tahun 2020

Berdasarkan **Gambar 2** dapat diketahui bahwa kecamatan Purwodadi memiliki keluarga sejahtera tahap I, II, III, dan, III+ yang tinggi dibandingkan dengan kecamatan lain yang mana mengindikasikan tingkat kesejahteraan keluarga di kecamatan Purwodadi lebih tinggi dibandingkan yang lain. Sementara itu, kecamatan Toroh memiliki keluarga pra sejahtera yang tinggi. Namun secara keseluruhan jumlah tingkat kesejahteraan keluarga pada setiap kecamatan di Kabupaten Grobogan memiliki jumlah yang berbeda, sehingga terlihat kecamatan mana yang lebih tinggi dari kecamatan lain pada variabel tertentu

3.2. Pengujian Asumsi Pada Analisis Cluster

3.2.1. Uji Kecukupan dan Uji Kelayakan Model

Dilakukan uji kecukupan data dengan menggunakan metode *Kaiser-Meyer-Olkin* (KMO) dan uji kelayakan model dengan menggunakan kriteria *Measure of Sampling Adequacy* (MSA). Uji KMO merupakan indeks perbandingan jarak antara nilai koefisien korelasi terhadap korelasi parsial.

Tabel 2 Uji Kecukupan dan Uji Kelayakan Model

Variabel	KMO	MSA
Keluarga Pra Sejahtera		0,56
Keluarga Sejahtera Tahap I		0,68
Keluarga Sejahtera Tahap II	0,57	0,85
Keluarga Sejahtera Tahap III		0,57
Keluarga Sejahtera Tahap III+		0,54

Berdasarkan nilai KMO atau *overall* MSA sebesar $0,57 > 0,5$ dan nilai *Measure of Sampling Adequacy* (MSA) atau *MSA for each item* diperoleh kesimpulan bahwa untuk

setiap kelompok kecamatan, sampel mewakili populasi sehingga layak untuk dilakukan analisis *cluster*.

3.2.2. Uji Multikolinearitas

Berikut merupakan hasil uji multikolinearitas :

Tabel 3 Uji Multikolinearitas Menggunakan Korelasi

	<i>Correlations</i>				
	Keluarga Pra Sejahtera	Keluarga Sejahtera Tahap I	Keluarga Sejahtera Tahap II	Keluarga Sejahtera Tahap III	Keluarga Sejahtera Tahap III+
Keluarga Pra Sejahtera	1	-0,2987	0,1982	0,4130	0,0603
Keluarga Sejahtera Tahap I	-0,2987	1	0,3917	0,1660	0,7705
Keluarga Sejahtera Tahap II	0,1982	0,3917	1	0,4465	0,6143
Keluarga Sejahtera Tahap III	0,4130	0,1660	0,4465	1	0,5834
Keluarga Sejahtera Tahap III+	0,0603	0,7705	0,6143	0,5834	1

Berdasarkan **Tabel 3** diperoleh bahwa nilai korelasi variabel $< 0,8$ yang mana mengindikasikan tidak terjadi gejala multikolinearitas serta uji asumsi terpenuhi. Pengujian multikolinearitas juga dapat menggunakan nilai VIF pada masing-masing variabel yang disajikan pada **Tabel 4**.

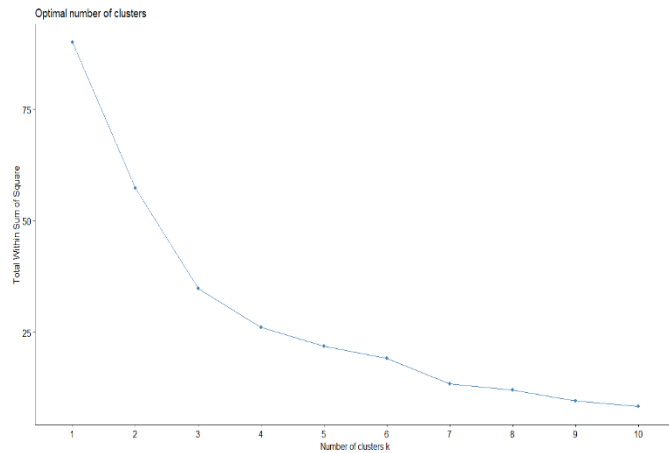
Tabel 4 Uji Multikolinearitas Menggunakan VIF

Variabel	Nilai VIF
Keluarga Pra Sejahtera	1,5095
Keluarga Sejahtera Tahap I	3,5539
Keluarga Sejahtera Tahap II	1,6551
Keluarga Sejahtera Tahap III	2,1725
Keluarga Sejahtera Tahap III+	5,9769

Berdasarkan **Tabel 4** dapat dilihat bahwa nilai VIF pada masing-masing variabel memiliki nilai < 10 . Hal ini mengindikasikan bahwa pada setiap variabel tidak terjadi multikolinearitas.

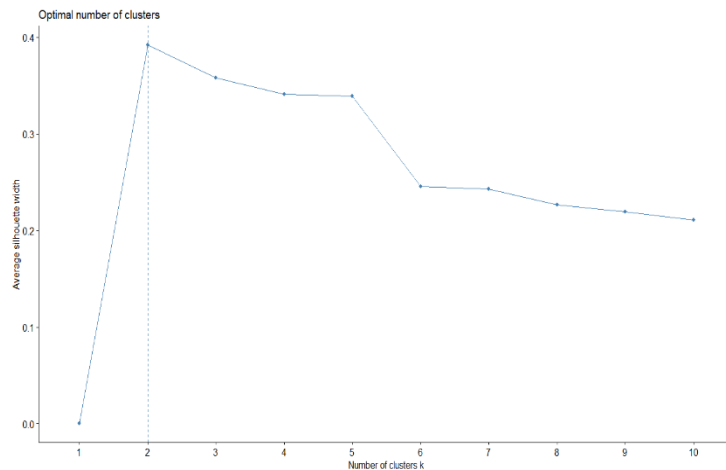
3.3. Penentuan Jumlah Cluster

Penentuan jumlah *cluster* atau banyaknya nilai k dapat ditentukan langsung oleh peneliti atau dapat berdasarkan beberapa pendekatan untuk mendapatkan nilai k yang *optimal*. Pendekatan tersebut antara lain dengan metode *within sum square* atau *elbow* dan *silhouette*.



Gambar 3 Penentuan Jumlah Cluster dengan Metode *Within Sum Square*

Berdasarkan **Gambar 3** grafik tersebut landai setelah banyaknya $k = 2$. Sehingga dapat diduga bahwa k optimal menurut metode *elbow* atau *within sum square* adalah 2.



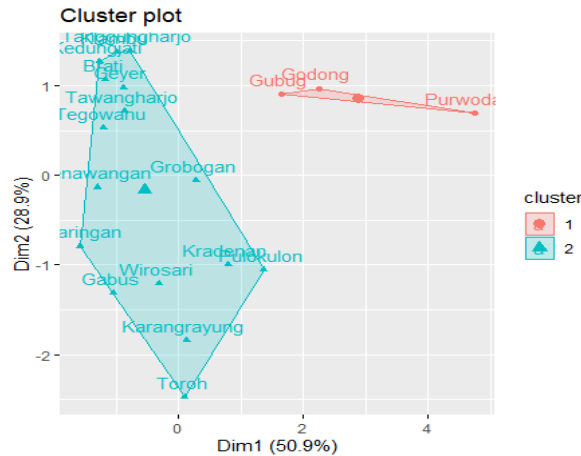
Gambar 4 Penentuan Jumlah Cluster dengan Metode *Silhouette*

Dengan pendekatan *silhouette*, menunjukkan k optimal yaitu saat nilai $k = 2$, ditandai dengan garis vertikal yang berada pada sumbu x saat $k = 2$.

Berdasarkan nilai k yang diperoleh menggunakan metode *within sum square* atau *elbow* dan *silhouette*, didapatkan nilai k yang optimum adalah 2. Sehingga akan digunakan $k = 2$ untuk mengelompokkan kecamatan di Kabupaten Grobogan berdasarkan tahapan perkembangan keluarga sejahtera.

3.4. Analisis *K-Means Clustering*

Metode *k-means* digunakan untuk mengklasifikasikan kecamatan di Kabupaten Grobogan berdasarkan tingkat kesejahteraan keluarga:



Gambar 5 Plot Pembentukan Cluster K-Means

Berdasarkan hasil *plot* pembentukan *cluster* pada Gambar 4 didapatkan hasil pengelompokan kecamatan dengan metode *k-means* seperti pada Tabel 5.

Tabel 5 Hasil Pengelompokan Cluster K-Mean

Cluster	Jumlah Anggota	Anggota Cluster
1	3	Purwodadi, Godong, Gubug.
2	16	Kedungjati, Karangrayung, Penawangan, Toroh, Geyer, Pulokulon, Kradenan, Gabus, Ngaringan, Wirosari, Tawangharjo, Grobogan, Brati, Klambu, Tegowanu, Tanggunharjo.

Terdapat 3 anggota kecamatan pada *cluster* 1 dan 16 anggota kecamatan pada *cluster* 2. Hasil *clustering* tersebut sesuai dengan statistika deskriptif yang ada dimana kecamatan Purwodadi masuk dalam *cluster* 1 dan kecamatan Toroh masuk dalam *cluster* 2.

Setiap *cluster* diprofilkan untuk menemukan rata-rata dari setiap variabel di setiap *cluster*, untuk menentukan fitur dari setiap *cluster*.

Tabel 6 Profilisasi Cluster

Cluster	Rata-rata					Keterangan
	Keluarga Pra Sejahtera	Keluarga Sejahtera Tahap I	Keluarga Sejahtera Tahap II	Keluarga Sejahtera Tahap III	Keluarga Sejahtera Tahap III Plus	
1	14136,667	7150,667	58808,000	5470,333	1081,000	Tinggi
2	16271,688	2313,438	3196,813	2759,813	217,813	Rendah

Tabel 6 menunjukkan nilai rata-rata setiap variabel di setiap *cluster*, dan interpretasi berikut dapat dibuat berdasarkan informasi ini:

- Jika melihat pada klaster 1, akan menemukan keluarga yang berkecukupan dalam keempat tingkatan, yang berarti klaster 1 dapat digabungkan menjadi kelompok dengan tingkat kesejahteraan keluarga yang tinggi. Hal ini dikarenakan cluster 1 memiliki rata-rata tingkat kesejahteraan keluarga paling tinggi.
- Klaster 2 yaitu klaster kecamatan dengan tingkat kesejahteraan keluarga rata-rata, dengan keluarga sejahtera tahap I, II, III, dan III+ rendah, sehingga kelompok kecamatan pada klaster 2 dapat dikelompokkan menjadi kelompok kecamatan. dengan tingkat kesejahteraan keluarga yang rendah.

4. Kesimpulan

Berdasarkan hasil pendekatan *k-means clustering* yang digunakan untuk mengklasifikasikan kecamatan di Kabupaten Grobogan menurut kesejahteraan keluarga tahun 2020, maka dapat ditarik kesimpulan sebagai berikut:

1. Ada tiga kecamatan di *cluster* 1 dan enam belas di *cluster* 2 di Kabupaten Grobogan, menurut analisis *clustering* lima faktor yang berhubungan dengan tingkat kesejahteraan keluarga.
2. Sedangkan klaster 2 merupakan kelompok kecamatan dengan tingkat perkembangan rendah, sedangkan klaster 1 merupakan kelompok kecamatan dengan tingkat perkembangan tinggi untuk rumah tangga sejahtera pada tahap I, II, III, dan III+.
3. Kecamatan Purwodadi, Godong, dan Gubug merupakan klaster 1 atau klaster tinggi. Serta yang lainnya di Kedungjati, Karangrayung dan Penawangan, merupakan bagian dari Klaster 2, yang juga mencakup Toroh, Geyer, Pulokulon, Kradenan, Gabus, Wirosari, Tawangharjo, Grobogan, Brati, Klambu, Tegowanu, Tanggunharjo dan Ngaringan.

Daftar Pustaka

- [1] BKKBN, “Pemutakhiran Data Keluarga,” 2011. <http://aplikasi.bkkbn.go.id/mdk/BatasanMDK.aspx> (diakses 12 Mei 2022).
- [2] D. Fitriani, T. N. Padilah, dan B. N. Sari, “Penerapan Algoritma K-Means dalam Pengelompokan Kesejahteraan Rakyat di Kabupaten Karawang,” *J. Ilm. Komput.*, vol. 17, hal. 73–82, 2021.
- [3] S. Akhyar, “Pengelompokan Kabupaten / Kota di Jawa Timur Berdasarkan Indikator Pembangunan Ekonomi menggunakan Model-Based Clustering,” Institut Teknologi Sepuluh Nopember, 2017.
- [4] F. Ramdhani, A. Hoyyi, dan M. A. Mukid, “Pengelompokan Provinsi di Indonesia Berdasarkan Karakteristik Kesejahteraan Rakyat Menggunakan Metode K-Means Cluster,” *J. Gaussian*, vol. 4, hal. 875–884, 2015.
- [5] J. F. J. Hair, R. . Anderson, R. . Tatham, dan W. C. Black, *Multivariate Data Analysis Sixth Edition*. New Jersey: Pearson Education, Inc, 2006.
- [6] L. Rahmawati, “Analisis Kelompok dengan Menggunakan Metode Hierarki untuk Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Kesehatan,” Universitas Negeri Malang, 2013.