

Penerapan Metode *Non-Hierarchical Clustering K-Means* untuk Pengelompokan Kabupaten/Kota di Provinsi Jawa Barat Berdasarkan Faktor Produksi Padi Tahun 2021

Syahla Anisah*, Lisnur Wachidah

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

*syahla.anisah111@gmail.com, wachidah.lisnur07@gmail.com

Abstract. Rice production in districts / cities in West Java often differences between districts / cities. The West Java regional government needs to conduct a study to determine the rice production groups in West Java Districts / Cities in order to meet the food needs of the people of West Java through policies related to increasing rice production by grouping districts / cities based on rice production. Rice production factors include planting area, harvest area, and productivity. Rice production groups are determined into 3 categories, namely low production, medium production, and high production. The grouping of districts/cities based on rice production uses non-hierarchical cluster analysis using the K-Means method. This study uses West Java rice production data in 2021. Based on the results of the study, three clusters were obtained, namely cluster 1 showing areas with low rice production with the characteristics of low planting area, low harvest area, and low productivity containing 11 districts/cities. Cluster 2 shows areas with medium rice production with the characteristics of low planting area, low harvest area, and high productivity containing 7 districts/cities. Cluster 3 shows areas with high rice production with high planting area, high harvest area, and high productivity containing 9 districts/cities.

Keywords: *Rice Production, Non-Hierarchical Cluster Analysis, K-Means.*

Abstrak. Produksi padi pada kabupaten/kota di Jawa Barat sering terjadi perbedaan antar kabupaten/kota. Pemerintah daerah Jawa Barat perlu melakukan sebuah kajian untuk mengetahui kelompok produksi padi di Kabupaten/Kota Jawa Barat agar dapat terpenuhinya kebutuhan pangan masyarakat Jawa Barat melalui kebijakan yang berhubungan dengan peningkatan produksi padi dengan cara mengelompokkan kabupaten/kota berdasarkan produksi padi. Faktor produksi padi diantaranya adalah luas tanam, luas panen, dan produktivitas. Kelompok produksi padi ditentukan menjadi 3 kategori yaitu produksi rendah, produksi sedang, dan produksi tinggi. Pengelompokan kabupaten/kota berdasarkan produksi padi tersebut menggunakan analisis klaster non hierarki dengan metode *K-Means*. Penelitian ini menggunakan data produksi padi Jawa Barat Tahun 2021. Berdasarkan hasil penelitian diperoleh tiga klaster, yaitu klaster 1 menunjukkan daerah dengan produksi padi rendah dengan karakteristik luas tanam rendah, luas panen rendah, dan produktivitas rendah yang berisi 11 kabupaten/kota. Klaster 2 menunjukkan daerah dengan produksi padi sedang dengan karakteristik luas tanam rendah, luas panen rendah, dan produktivitas tinggi yang berisi 7 kabupetn/kota. Klaster 3 menunjukkan daerah dengan produksi padi tinggi dengan luas tanam tinggi, luas panen tinggi, dan produktivitas tinggi yang berisi 9 kabupeten/kota.

Kata Kunci: *Produksi Padi, Analisis Klaster Non Hierarki, K-Means.*

A. Pendahuluan

Menurut BPS Provinsi Jawa Barat, total produksi padi di Provinsi Jawa Barat selama tahun 2021 sekitar 9,11 juta ton Gabah Kering Giling (GKG) mengalami kenaikan sebanyak 96,80 ribu ton atau 1,07 persen dibandingkan tahun 2020 yang sebesar 9,02 juta ton [1]. Produksi padi Provinsi Jawa Barat secara provinsi memang mengalami peningkatan, tetapi secara nasional telah terjadi pergeseran peringkat tertinggi provinsi penghasil produksi padi tahun 2021. Jawa Timur berada di peringkat pertama. Jawa Tengah peringkat kedua, dan Jawa Barat berada di peringkat ketiga. Hal ini terjadi karena adanya kesenjangan antara produksi padi pada kabupaten/kota di Jawa Barat. Dinas Tanaman Pangan dan Hortikultura Jawa Barat (2022) mendata terjadi penurunan luas lahan pertanian padi di Jawa Barat pada tahun 2021. Penurunan ini terkait dengan alih fungsi lahan pada kabupaten/kota di Jawa Barat akibat adanya kegiatan pembangunan. Pemerintah daerah Jawa Barat perlu melakukan sebuah kajian untuk mengetahui kelompok produksi padi di kabupaten/kota agar terpenuhi kebutuhan pangan masyarakat Jawa Barat.

Berdasarkan karakteristik setiap kabupaten/kota dilakukan pengelompokkan melalui analisis kluster [2]. Kelompok produksi padi ditentukan menjadi 3 kategori yaitu produksi rendah, produksi sedang, dan produksi tinggi [3]. Berdasarkan pengelompokkan tersebut, digunakan analisis kluster non hierarki dengan metode *K-Means*. Pengelompokkan ini bermanfaat bagi pemerintah khususnya pemerintah di Provinsi Jawa Barat dalam menentukan daerah mana yang dapat diprioritaskan untuk meningkatkan produksi padi melalui kebijakan yang berhubungan dengan peningkatan produksi padi, penyediaan sarana dan prasarana yang dapat menunjang produksi padi agar terpenuhinya kebutuhan pangan masyarakat Jawa Barat.

Berdasarkan uraian dari latar belakang yang telah dipaparkan, maka perumusan masalah dalam penelitian ini adalah “Bagaimana hasil pengelompokkan wilayah kabupaten/kota di Provinsi Jawa Barat berdasarkan faktor produksi padi dengan menggunakan metode Analisis Kluster *K-Means*”. Selanjutnya, tujuan dalam penelitian ini adalah untuk mengetahui hasil pengelompokkan wilayah kabupaten/kota di Provinsi Jawa Barat berdasarkan faktor produksi padi dengan menggunakan metode Analisis Kluster *K-Means*.

B. Metodologi Penelitian

Standarisasi Data

Data yang digunakan dalam analisis kluster jika terdapat perbedaan variabel dengan satuan nilai yang besar antar variabel, maka perlu dilakukan proses standarisasi data dengan transformasi data asli sebelum dianalisis lebih lanjut [4]. Perbedaan satuan data yang besar menyebabkan perhitungan jarak menjadi tidak valid. Transformasi data dapat dilakukan terhadap variabel ke dalam bentuk *z-score* [5], yaitu transformasi data dalam bentuk normal baku $N(0,1)$ dirumuskan sebagai berikut:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (1)$$

Keterangan:

Dimana $i = 1, 2, \dots, n$ dan $j = 1, 2, \dots, p$

n : banyaknya observasi

p : banyaknya variabel

z_{ij} : data hasil standarisasi observasi ke- i pada variabel ke- j

x_{ij} : observasi ke- i pada variabel ke- j

\bar{x}_j : rata-rata variabel ke- j

s_j : simpangan baku variabel ke- j

Uji Kaiser Meyer Olkin (KMO)

Uji KMO digunakan untuk melihat kecukupan sampel dalam mewakili populasi sehingga sampel yang digunakan sudah cukup untuk dianalisis lebih lanjut. Keputusan sampel yang digunakan dapat dikatakan mewakili populasi (sampel yang representatif) yaitu jika nilai KMO berkisar 0,5 sampai 1 [6] dengan rumus sebagai berikut [7]:

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1; j \neq i}^p r_{x_i x_j}^2}{\sum_{i=1}^p \sum_{j \neq i}^p r_{x_i x_j}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{x_i x_j}^2} \quad (2)$$

dengan:

$$r_{x_i x_j} = \frac{n \sum_{i=1}^p x_i \sum_{j=1}^p x_j - (\sum_{i=1}^p x_i)(\sum_{j=1}^p x_j)}{\sqrt{\{n \sum_{i=1}^p x_i^2 - (\sum_{i=1}^p x_i)^2\} \{n \sum_{j=1}^p x_j^2 - (\sum_{j=1}^p x_j)^2\}}}$$

$$a_{x_i x_j} = \frac{r_{x_i x_j}^{-1}}{\sqrt{r_{x_i x_j}^{-1}}}$$

Keterangan:

Dimana $i = 1, 2, \dots, n$ dan $j = 1, 2, \dots, p$

n : banyaknya observasi

p : banyaknya variabel

$r_{x_i x_j}$: koefisien korelasi terobservasi antara variabel i dan variabel j

$a_{x_i x_j}$: koefisien korelasi parsial antara variabel i dan variabel j

Pemeriksaan Multikolinearitas

Pada pengujian analisis kluster, salah satu asumsi yang harus dipenuhi yaitu tidak terjadi multikolinieritas atau tidak terdapat multikolinieritas diantara variabel-variabel independen. Cara indentifikasi adanya multikolinieritas adalah melalui nilai *Tolerance* dan menghitung nilai *Varians Inflation Factor* (VIF) dengan keputusan apabila nilai *Tolerance* lebih dari 0,1 dan nilai VIF kurang dari 10 maka tidak terdapat multikolinieritas antar variabel independen. Rumus untuk mencari nilai *Tolerance* dan nilai VIF sebagai berikut [8]:

$$Tolerance = 1 - R_j^2 \quad (3)$$

$$VIF = \frac{1}{1 - R_j^2} \quad (4)$$

dengan:

$$R_j^2 = (r_{x_i x_j})^2 \times 100\%$$

Keterangan:

Dimana $j = 1, 2, \dots, p$

p : banyaknya variabel

R_j^2 : nilai koefisien determinasi variabel ke- j

$r_{x_i x_j}$: koefisien korelasi terobservasi antara variabel i dan variabel j

Pembentukan Kluster

Metode non hierarki analisis kluster dimulai dengan menentukan banyaknya jumlah kluster yang diinginkan (misalnya akan dibentuk dua kluster, tiga kluster, empat kluster atau lebih) [9], setelah ditentukan banyaknya jumlah kluster maka proses dilanjutkan tanpa mengikuti proses hirarki, metode ini sering disebut sebagai metode *K-Means Clustering*. Prosedur *K-Means Cluster* ini biasanya digunakan apabila banyaknya kluster (k) sudah ditentukan atau dispesifikasikan terlebih dahulu, kemudian membagi n objek kedalam k kluster sesuai karakteristik tertentu. Walaupun dalam beberapa metode, yang tergabung dalam prosedur ini, banyaknya kluster bisa bervariasi selama analisis berlangsung [10]. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu kluster yang sama dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan kluster yang lain [11].

Secara umum metode *K-Means* menggunakan algoritma dasar sebagai berikut:

1. Menentukan jumlah kluster k
2. Menentukan k *centroid cluster* (titik pusat kluster) awal secara acak dari objek-objek yang tersedia sebanyak k kluster. *centroid cluster* ditulis dengan c_j (titik pusat kluster ke- j)
3. Menghitung jarak setiap data yang ada terhadap setiap pusat kluster di masing-masing kluster menggunakan rumus *Euclidian Distance*. Perolehan nilai jarak, didapat dari nilai minimum setiap objek terhadap *centroid cluster*. Dimana hasil jarak minimum ini menunjukkan suatu objek semakin mirip dengan objek lainnya yang berada dalam satu kluster yang sama. Jarak *euclidean* merupakan akar dari jumlah kuadrat perbedaan nilai tiap variabel. Rumus jarak *euclidean* sebagai berikut:

$$D(x_i, y_j) = \sqrt{\sum_{l=1}^p (x_{il} - y_{jl})^2} \quad (5)$$

Keterangan:

Dimana $l = 1, 2, \dots, p$; $i = 1, 2, \dots, n$; dan $j = 1, 2, \dots, n$

$D(x_i, y_j)$: banyaknya variabel

p : banyaknya variabel

n : banyaknya data

x_{il} : data ke- i pada variabel ke- l

y_{jl} : titik pusat kluster ke- j pada variabel ke- l

4. Kelompokkan setiap data ke jarak terdekat pusatnya
5. Melakukan iterasi, kemudian menentukan posisi pusat kluster baru c_{kj} dengan cara mengitung nilai rata-rata dari data-data yang ada pada kluster yang sama dengan rumus sebagai berikut [12]:

$$c_{kj} = \left(\frac{1}{n_k}\right) \sum d_{kl} \quad (6)$$

Keterangan:

c_{kj} : pusat kluster baru ke- j pada variabel ke- l dalam kluster ke- k

n_k : banyaknya anggota objek dalam kluster ke- k

d_{kl} : data pada variabel ke- l dalam kluster ke- k

6. Kembali ke langkah 3 jika masih ada data yang berpindah kluster, sedangkan jika tidak data yang berpindah kluster maka proses kluster selesai. Hal ini disebut juga dengan proses iterasi.

Interpretasi Kluster

Interpretasi kluster merupakan proses menamai dan menandai dengan suatu nama dari masing-masing kluster yang telah terbentuk. Nama yang diberikan kepada masing-masing kluster sesuai dengan keperluan peneliti, dengan syarat bahwa nama yang diberikan kepada kluster tersebut dapat mewakili karakteristik atau ciri dari masing-masing kelompok dengan dilakukan pengkajian mengenai *centroid* dimana *centroid* ini merupakan rata-rata nilai objek yang terdapat dalam kluster pada setiap variabel.

C. Hasil Penelitian dan Pembahasan

Standarisasi Data

Standarisasi data dilakukan karena data pada penelitian ini memiliki satuan variabel yang berbeda yaitu variabel Luas Tanam dengan satuannya Hektar, variabel Luas Panen dengan satuannya Hektar, dan variabel Produktivitas dengan satuannya Kuintal/Hektar. Berdasarkan persamaan (1) dengan menggunakan *software* SPSS data standarisasi disajikan dalam tabel berikut ini:

Tabel 1. Hasil Standarisasi Data pada Variabel Luas Tanam, Luas Panen, dan Produktivitas

No	Wilayah Jawa Barat	ZLuas Tanam	ZLuas Panen	ZProduktivitas
1	Kabupaten Bogor	0,0747	-0,0321	-1,0395
2	Kabupaten Sukabumi	1,4350	0,4870	0,0503
3	Kabupaten Cianjur	1,2769	0,8757	-0,7893
4	Kabupaten Bandung	0,3775	-0,1433	0,4536
5	Kabupaten Garut	0,7654	0,2678	0,3150
6	Kabupaten Tasikmalaya	0,6470	0,3806	0,3895
7	Kabupaten Ciamis	-0,0582	-0,0712	0,5777
8	Kabupaten Kuningan	-0,2400	-0,1970	0,3915
9	Kabupaten Cirebon	0,2286	0,4131	-0,0469
10	Kabupaten Majalengka	0,5002	0,6165	-0,2309
11	Kabupaten Sumedang	0,0933	-0,0982	-1,4841
12	Kabupaten Indramayu	2,4345	2,7122	0,7472
13	Kabupaten Subang	1,6033	1,6913	0,6087
14	Kabupaten Purwakarta	-0,4614	-0,4284	-1,0085
15	Kabupaten Karawang	1,5935	2,2409	1,1195
16	Kabupaten Bekasi	-0,3822	0,6622	0,8568
17	Kabupaten Bandung Barat	-0,3577	-0,5138	-0,3633
18	Kabupaten Pangandaran	-0,5785	-0,5134	-1,0602
19	Kota Bogor	-1,0250	-0,9606	-2,0301
20	Kota Sukabumi	-0,9846	-0,9215	0,8527
21	Kota Bandung	-1,0101	-0,9446	2,5774
22	Kota Cirebon	-1,0302	-0,9575	-1,0126
23	Kota Bekasi	-1,0264	-0,9535	-1,3373
24	Kota Depok	-1,0320	-0,9605	0,0400
25	Kota Cimahi	-1,0302	-0,9560	0,0875
26	Kota Tasikmalaya	-0,8729	-0,8231	0,0772
27	Kota Banjar	-0,9403	-0,8686	1,2580

Uji Kaiser Meyer Olkin (KMO)

Berdasarkan persamaan (2) dengan menggunakan *software* SPSS diperoleh hasil *output* pada Uji KMO sebagai berikut:

Tabel 2. Uji Kaiser Meyer Olkin (KMO)

Kaiser-Meyer-Olkin Test		
Uji	Nilai	Keterangan
KMO	0,500	Mewakili Populasi

Pada tabel diatas, diperoleh nilai KMO sebesar 0,500. Karena 0,500 berkisar antara 0,5 sampai 1 maka diperoleh bahwa sampel yang digunakan sudah cukup mewakili populasi serta sampel yang digunakan sudah cukup untuk dianalisis lebih lanjut.

Pemeriksaan Multikolinearitas

Berdasarkan persamaan (3) dan persamaan (4) dengan menggunakan *software* SPSS diperoleh hasil *output* sebagai berikut:

Tabel 3. Uji Multikolinearitas

Model	Tolerance	VIF	Keterangan
Luas Tanam	0,123	8,130	Tidak Terjadi Multikolinearitas
Luas Panen	0,119	8,406	Tidak Terjadi Multikolinearitas
Produktivitas	0,910	1,098	Tidak Terjadi Multikolinearitas

Pada tabel diatas, menunjukkan bahwa semua variabel memiliki nilai *tolerance* > 0,1 dan nilai VIF < 10 maka dapat disimpulkan tidak terdapat multikolinearitas di antara semua variabel.

Pembentukan klaster

Menentukan *Centroid Cluster* (Titik Pusat Klaster)

Jumlah klaster yang akan dibentuk (*k*) pada proses clustering dengan metode K-Means sebanyak tiga *centroid* yaitu c_1 (*centroid* pada klaster 1), c_2 (*centroid* pada klaster 2), dan c_3 (*centroid* pada klaster 3). Penentuan nilai *centroid* diperoleh secara acak dari objek dalam penelitian. Nilai *centroid* pada masing-masing klaster disajikan sebagai berikut:

Tabel 4. Centroid Cluster

	Klaster		
	1	2	3
Luas Tanam	-1,0250	-1,0101	2,4345
Luas Panen	-0,9606	-0,9446	2,7122
Produktivitas	-2,0301	2,5774	0,7472

Pada tabel diatas, nilai *centroid* pada klaster 1 terletak pada data ke-19 yaitu Kota Bogor atau $c_1(-1,0250; -0,9606; -2,0301)$, nilai *centroid* pada klaster 2 terletak pada data ke-21 yaitu Kota Bandung atau $c_2(-1,0101; -0,9446; 2,5774)$, dan nilai *centroid* pada klaster 3 terletak pada data ke-12 yaitu Kabupaten Indramayu atau $c_3(2,4345; 2,7122; 0,7472)$.

Menghitung Jarak dan Mengelompokkan Setiap Data ke Jarak Terdekat Pusatnya

Pada langkah ini, dilakukan perhitungan untuk menentukan jarak. Dari hasil tersebut, dapat digunakan untuk mengelompokkan setiap data ke jarak terdekat pusatnya. Selanjutnya, menentukan pusat klaster yang baru dan dilakukan kembali perhitungan untuk menentukan jarak. Sehingga proses ini dapat dikatakan dengan proses iterasi. Berdasarkan persamaan (5) untuk menghitung jarak dan berdasarkan persamaan (6) untuk menentukan pusat klaster baru dengan menggunakan software SPSS diperoleh hasil output sebagai berikut:

Tabel 5. Hasil Iterasi Klaster

Iterasi	Klaster		
	1	2	3
1	1,489	1,775	1,849
2	0,130	0,000	0,154
3	0,133	0,000	0,148
4	0,000	0,000	0,000

Pada tabel diatas, dalam mengelompokkan produksi padi berdasarkan kabupaten/kota di Provinsi Jawa Barat menjadi tiga klaster dilakukan iterasi sebanyak 4 kali. Angka-angka pada tabel tersebut menunjukkan nilai signifikansi hasil dari iterasi. Proses iterasi berhenti pada iterasi ke-4 karena nilai *centroid* sudah signifikan artinya tidak ada lagi perpindahan pengelompokkan kabupaten/kota antar klaster. Proses ini dilakukan untuk mendapatkan klaster yang tepat. Selanjutnya, disajikan tabel jarak setiap data dengan pusat klaster sebagai berikut:

Tabel 6. Jarak Setiap Data dengan Pusat Klaster

No	Wilayah Jawa Barat	Jarak Kab/Kota ke c_1	Jarak Kab/Kota ke c_2	Jarak Kab/Kota ke c_3	Jarak Minimum	Klaster
1	Kabupaten Bogor	0,985	2,129	2,014	0,985	1
2	Kabupaten Sukabumi	2,542	2,281	0,675	0,675	3
3	Kabupaten Cianjur	2,468	2,779	1,055	1,055	3
4	Kabupaten Bandung	1,727	1,022	1,467	1,022	2
5	Kabupaten Garut	2,047	1,536	0,905	0,905	3
7	Kabupaten Ciamis	1,638	0,647	1,711	0,647	2
8	Kabupaten Kuningan	1,370	0,663	1,902	0,663	2
9	Kabupaten Cirebon	1,594	1,468	1,183	1,183	3
10	Kabupaten Majalengka	1,822	1,837	0,936	0,936	3
11	Kabupaten Sumedang	1,141	2,554	2,345	1,141	1
12	Kabupaten Indramayu	4,837	4,226	2,132	2,132	3
13	Kabupaten Subang	3,562	2,933	0,840	0,840	3
14	Kabupaten Purwakarta	0,349	2,005	2,543	0,349	1
15	Kabupaten Karawang	4,154	3,314	1,521	1,521	3
16	Kabupaten Bekasi	2,158	1,030	1,716	1,030	2
17	Kabupaten Bandung Barat	0,573	1,372	2,283	0,573	1
18	Kabupaten Pangandaran	0,282	2,065	2,694	0,282	1
19	Kota Bogor	1,291	3,136	3,755	1,291	1
20	Kota Sukabumi	1,735	0,784	2,998	0,784	2
21	Kota Bandung	3,438	1,775	3,778	1,775	2
22	Kota Cirebon	0,513	2,172	3,244	0,513	1
23	Kota Bekasi	0,694	2,473	3,378	0,694	1
24	Kota Depok	0,995	1,266	3,002	0,995	1
25	Kota Cimahi	1,036	1,230	2,998	1,036	1
26	Kota Tasikmalaya	0,947	1,109	2,790	0,947	1
27	Kota Banjar	2,118	0,749	3,041	0,749	2

Pada tabel diatas, merupakan hasil akhir dari proses iterasi ke-4 berdasarkan Tabel 5. Dalam mengelompokkan 27 Kabupaten/Kota Provinsi Jawa Barat kedalam tiga klaster produksi padi dilihat dari jarak minimum kabupaten/kota ke pusat klaster. Misalkan berdasarkan Tabel 6 untuk Kabupaten Bogor, memiliki nilai minimumnya yaitu 0,985 yang merupakan nilai jarak objek dengan titik pusat klaster 1. Maka Kabupaten Bogor merupakan anggota dari klaster 1.

Menginterpretasikan Hasil Klaster yang Diperoleh

1. Menentukan Anggota Pada Setiap Klaster
2. Kabupaten/kota yang merupakan anggota setiap klaster disajikan pada tabel berikut ini:

Tabel 7. Keanggotaan Kabupaten/Kota Pada Setiap Kluster

Clustering		
1	2	3
Kab. Bogor	Kab. Bandung	Kab. Sukabumi
Kab. Sumedang	Kab. Ciamis	Kab. Cianjur
Kab. Purwakarta	Kab. Kuningan	Kab. Garut
Kab. Bandung Barat	Kab. Bekasi	Kab. Tasikmalaya
Kab. Pangandaran	Kota Sukabumi	Kab. Cirebon
Kota Bogor	Kota Bandung	Kab. Majalengka
Kota Cirebon	Kota Banjar	Kab. Indramayu
Kota Bekasi		Kab. Subang
Kota Depok		Kab. Karawang
Kota Cimahi		
Kota Tasikmalaya		

Pada tabel diatas, dari ketiga kluster yang terbentuk, kluster 1 merupakan kluster dengan jumlah anggota terbanyak yaitu sebanyak 11 kabupaten/kota, kluster 2 sebanyak 7 kabupaten/kota, dan kluster 3 sebanyak 9 kabupaten/kota.

3. Visualisasi Kluster

Visualisasi dilakukan menggunakan diagram peta (kartogram) untuk menggambarkan kabupaten/kota di Provinsi Jawa Barat sesuai dengan klusternya masing-masing yang disajikan pada Gambar 1.

**Gambar 1.** Daerah Hasil Kluster dengan Metode *K-Means*

Berdasarkan Gambar diatas, menggambarkan Kabupaten/Kota di Provinsi Jawa Barat. Dari visualisasi gambar diagram peta tersebut dapat diketahui bahwa kabupaten/kota di Provinsi Jawa Barat terbagi menjadi 3 kluster yang ditandai dengan adanya tiga warna. Kluster 1 ditandai dengan warna hijau muda, kluster 2 ditandai dengan warna hijau toska, dan kluster 3 ditandai dengan warna hijau tua.

4. Interpretasi Kluster Pada Setiap Variabel

Interpretasi kluster merupakan proses akhir dalam membentuk kluster. Dalam menginterpretasikan hasil kluster, dapat dilihat melalui *centroid* pada setiap variabel yang disajikan pada tabel berikut ini:

Tabel 8. Nilai *Centroid* Pada Setiap Variabel

	Klaster		
	1	2	3
Luas Tanam	-0,6588	-0,4626	1,1649
Luas Panen	-0,6546	-0,3549	1,0761
Produktivitas	-0,8301	0,9954	0,2404

Variabel pada Tabel diatas, memiliki nilai negatif dan nilai positif. Nilai negatif memiliki arti bahwa data berada di bawah rata-rata total dan nilai positif memiliki arti bahwa data berada di atas rata-rata total. Tabel diatas menunjukkan karakteristik pada masing-masing klaster. Berikut merupakan penjelasan untuk masing-masing klaster. Klaster 1 memiliki nilai rata-rata untuk variabel luas tanam, luas panen, dan produktivitas berada di bawah rata-rata populasi. Hal ini terbukti dari nilai negatif (-) untuk semua variabel. Klaster 2 memiliki nilai rata-rata untuk variabel luas tanam dan luas panen berada di bawah rata-rata populasi, dan variabel produktivitas berada di atas rata-rata populasi. Hal ini terbukti dari nilai negatif (-) untuk variabel luas tanam dan luas panen serta nilai positif (+) untuk variabel produktivitas. Klaster 3 memiliki nilai rata-rata untuk variabel luas tanam, luas panen, dan produktivitas berada di atas rata-rata populasi. Hal ini terbukti dari nilai positif (+) untuk semua variabel.

D. Kesimpulan

Berdasarkan hasil dan pembahasan dapat disimpulkan bahwa metode *K-Means* dalam pembentukan kelompok berdasarkan faktor produksi padi mengelompokkan 27 Kabupaten/Kota Provinsi Jawa Barat menjadi 3 klaster. Klaster 1 memiliki karakteristik luas tanam yang rendah, luas panen yang rendah, dan produktivitas yang rendah. Berdasarkan ciri-ciri yang dimiliki klaster 1 dapat diberi nama dengan produksi padi rendah. Anggota pada klaster 1 merupakan kabupaten/kota memiliki lahan yang sudah banyak digunakan aktivitas pembangunan. Sehingga produksi padi di klaster 1 dapat ditingkatkan melalui optimalisasi lahan yang ada seperti perbaikan akses irigasi, penggunaan pupuk kimia yang berimbang, dan peningkatan produktivitas dengan penggunaan bibit unggul. Klaster 2 memiliki karakteristik luas tanam yang rendah, luas panen yang rendah, dan produktivitas yang tinggi. Berdasarkan ciri-ciri yang dimiliki klaster 2 dapat diberi nama dengan produksi padi sedang. Peningkatan produksi padi di klaster 2 dapat ditingkatkan dengan cara memanfaatkan lahan-lahan pertanian yang selama ini tidak digunakan untuk bercocok tanam padi dan menciptakan pertanian yang berkelanjutan atau pertanian ramah lingkungan dengan menggunakan cara hayati (pembasmi alami yang sesuai) dalam membasmi hama. Klaster 3 memiliki karakteristik luas tanam yang tinggi, luas panen yang tinggi, dan produktivitas yang tinggi. Berdasarkan ciri-ciri yang dimiliki klaster 3 dapat diberi nama dengan produksi padi tinggi. Meskipun klaster 3 merupakan produksi padi tinggi, tetapi peningkatan produksi tetap harus dilakukan dengan cara meningkatkan Indeks Pertanaman (IP) melalui optimalisasi pengelolaan sumber daya air, meningkatkan produktivitas padi melalui penggunaan bibit unggul serta menerapkan sistem jajar legowo dalam menanam padi.

Acknowledge

Penelitian ini dapat terlaksana dengan baik atas bantuan, bimbingan, saran dan dukungan dari berbagai pihak. Oleh karena itu, peneliti mengucapkan terima kasih kepada Allah SWT atas rahmat dan hidayah-Nya, kedua orang tua serta keluarga yang senantiasa mendo'akan dan memberi dukungan baik moral maupun materi, Bapak Dr. Aceng Komarudin Mutaqin, S.Si., M.T., M.Si. selaku dosen pembimbing yang telah memberikan motivasi, saran, menyumbangkan ilmu, waktu, kesempatan, kesabaran, serta kemudahan bagi penulis untuk dapat menyelesaikan skripsi ini dengan baik, para dosen Statistika Unisba yang telah memberikan ilmu pengetahuannya, dan teman-teman yang telah memberikan kritikan, bantuan, dan dorongannya.

Daftar Pustaka

- [1] Badan Pusat Statistik Provinsi Jawa Barat, “Pada 2021, luas panen padi mencapai sekitar 1,60 juta hektar dengan produksi sebesar 9,11 juta ton GKG,” 2022. <https://jabar.bps.go.id/pressrelease/2022/03/01/1007/pada-2021--luas-panen-padi-mencapai-sekitar-1-60-juta-hektar-dengan-produksi-sebesar-9-11-juta-ton-gkg.html> (accessed Jul. 25, 2023).
- [2] N. Azizah, “Pemodelan Spatial Autoregressive (SAR-X) pada Perkawinan Usia Anak di Indonesia,” *Jurnal Riset Statistika*, pp. 1–10, Jul. 2023, doi: 10.29313/jrs.v3i1.1643.
- [3] R. H. B. Bangun, “Analisis Klaster Non Heirarki Dalam Pengelompokan Kabupaten/Kota di Sumatera Utara Berdasarkan Faktor Produksi Padi,” *Jurnal Agribisnis Sumatera Utara*, vol. 4, no. 1, pp. 54–61, 2016.
- [4] Sophia Annisa Faisal and N. A. K. Rifai, “Penerapan Metode Hierarchical Clustering Multiscale Bootstrap untuk Pengelompokan Indikator Indeks Pembangunan Manusia Tahun 2021 di Jawa Barat,” *Bandung Conference Series: Statistics*, vol. 3, no. 1, pp. 100–106, 2023, doi: 10.29313/bcss.v3i1.6327.
- [5] S. Nugroho, *Statistika Mutivariat Terapan*. 2008.
- [6] A. N. Fathia, R. Rahmawati, and Tarno, “Analisis Klaster Kecamatan Di Kabupaten Semarang Berdasarkan Potensi Desa Menggunakan Metode Ward Dan Single Linkage,” *Gaussian*, vol. 5, no. 4, pp. 801–810, 2016.
- [7] J. Gaussian, “136924-ID-pengelo-mpokan-kualitas-udara-ambien-men,” vol. 4, no. 2, pp. 393–402, 2015.
- [8] K. Dan, M. Terhadap, K. Manajerial, and P. Jambi, “1) 2) 3),” pp. 1–12.
- [9] A. Khaerunnisa, “Analisis Tingkat Kelulusan Mahasiswa di Unisba dengan menggunakan Algoritma K-Means Clustering,” *Jurnal Riset Matematika*, pp. 67–76, Jul. 2022, doi: 10.29313/jrm.v2i1.1018.
- [10] N. Hajarisman, “Statistika Multivariat: Analisis Klaster,” 2019.
- [11] Y. Agusta, “K-Means-Penerapan, Permasalahan dan Metode Terkait,” *Jurnal Sistem dan Informatika*, vol. 3, no. Pebruari, pp. 47–60, 2007.
- [12] E. Rivani, “Aplikasi K-Means Cluster untuk Pengelompokkan Provinsi Berdasarkan Produksi Padi, Jagung, Kedelai, dan Kacang Hijau Tahun 2019,” *Jurnal Mat Stat*, vol. 10, no. 2, pp. 122–134, 2010.
- [13] R. Umami, W. E. Putra, & R. Friyani, *Pengaruh Partisipasi Penyusunan Anggaran, Gaya Kepemimpinan dan Motivasi Terhadap Kinerja Manajerial (studi kasus pada Kantor Wilayah Kementerian Agama Provinsi Jambi)*, 2020.