

Pengelompokkan Kecamatan di Kabupaten Bima Berdasarkan Jumlah Produksi dan Luas Panen Bawang Merah Tahun 2021 Menggunakan *K-Means Clustering*

Ashabul Akbar Maulana¹, Athallah Widyatama Rafii²,
Yulia Anggi Anjelina³, Edy Widodo⁴

Universitas Islam Indonesia, Fakultas Matematika dan Ilmu Pengetahuan Alam

Jl. Kaliurang No.Km. 14,5, Krawitan, Umbulmartani, Kec. Ngemplak, Kabupaten Sleman,
Daerah Istimewa Yogyakarta 55584

e-mail: 19611101@students.uii.ac.id⁽¹⁾, 19611106@students.uii.ac.id⁽²⁾,
19611104@students.uii.ac.id⁽³⁾, edywidodo@uui.ac.id⁽⁴⁾

ABSTRAK

Bawang merah adalah komoditas pertanian signifikan di Indonesia dengan potensi ekonomi yang menjanjikan. Kabupaten Bima mendominasi produksi bawang merah di Indonesia. Secara geografis, Kabupaten Bima berada pada kisaran ketinggian 0 hingga 477.5 mdpl dengan mayoritas wilayahnya memiliki iklim panas dan kering, menciptakan kondisi yang ideal untuk budidaya bawang merah. Penelitian ini bertujuan menggambarkan serta menganalisis Cluster hasil produksi bawang merah di Kabupaten Bima tahun 2021 menggunakan metode Clustering K-Means. K-Means Clustering merupakan analisis kluster non-hierarkis yang membagi data ke dalam kelompok berdasarkan karakteristik tertentu. Hasil analisis menunjukkan bahwa Kecamatan Lambu mencatat produksi dan luas panen bawang merah tertinggi di Kabupaten Bima, sementara Kecamatan Langgudu dan Donggo tidak menghasilkan bawang merah. Analisis kluster menghasilkan tiga kelompok dengan karakteristik yang berbeda, di mana Cluster 1 memiliki kategori produksi rendah, Cluster 2 memiliki kategori produksi tinggi, dan Cluster 3 memiliki kategori produksi sedang.

Kata kunci : Bawang Merah; *Clustering; K-Means.*

ABSTRACT

Shallots are a significant agricultural commodity in Indonesia with promising economic potential. Bima Regency dominates shallot production in Indonesia. Geographically, Bima Regency spans an elevation range of 0 to 477.5 meters above sea level, with the majority of its area experiencing a hot and dry climate, creating ideal conditions for shallot cultivation. This study aims to depict and analyze the clusters of shallot production outcomes in Bima Regency for the year 2021 using the K-Means Clustering method. K-Means Clustering is a non-hierarchical cluster analysis that groups data based on specific characteristics. The analysis results indicate that the Lambu Sub-district records the highest shallot production and harvest area in Bima Regency, while the Langgudu and Donggo Sub-districts do not yield shallots. The cluster analysis yields three distinct groups with varying characteristics, where Cluster 1 has a low production category, Cluster 2 has a high production category, and Cluster 3 falls into the moderate production category.

Keywords : Red Onion, Clustering, K-Means.

PENDAHULUAN

Bawang merah adalah salah satu komoditas pertanian yang menjanjikan dan memiliki nilai jual yang cukup tinggi. Kabupaten Bima menjadi salah satu daerah penghasil bawang merah terbesar di Indonesia. Dengan letak geografis yang berada pada 0 sampai dengan 477.5 mdpl memiliki iklim yang kering dan suhu udara yang tergolong panas sangat cocok untuk ditanami bawang merah, menjadikan Kabupaten Bima mempunyai potensi yang sangat tinggi dalam menghasilkan bawang merah [1].

Salah satu metode yang dapat digunakan untuk melakukan analisis pada data tersebut adalah analisis clustering. Secara umum, Clustering didefinisikan sebagai pengelompokan data yang menunjukkan kesamaan berdasarkan standar khusus [2]. Sama seperti kelompok profil yang terbagi menjadi beberapa klaster yang saling serupa satu sama lain. Klasterisasi menghasilkan jumlah klaster yang kurang atau setara dengan jumlah data profil yang dimasukkan [3].

Metode k-Means merupakan salah satu metode yang populer digunakan untuk clustering data. *K-Means Clustering* adalah salah satu jenis analisis klaster non-hirarki. Metode ini bertujuan untuk membagi suatu objek data kedalam satu atau lebih jumlah *Cluster* berdasarkan karakteristik tertentu [4]. Terdapat beberapa keunggulan *Clustering K-Means* dibandingkan dengan *Clustering* yang lain diantaranya waktu komputasi yang relatif cepat, dan hasil yang mudah di representasikan [5].

Penelitian yang dilakukan oleh Febianto & Palarasa (2018) [6] yang meneliti tentang pengelompokan daerah di Provinsi Jawa Timur berdasarkan 13 indikator kemiskinan. Penelitian ini menghasilkan 5 cluster dengan pemetaan karakteristik dari setiap kelompok yang terbentuk berdasarkan nilai rata-rata tertinggi dan terendah dari setiap indicator kemiskinan provinsi Jawa Barat tahun 2018.

Penelitian ini bertujuan untuk mengetahui hasil dari pengelompokan daerah di Kabupaten Bima berdasarkan data Jumlah Panen dan Luas Panen Bawang Merah tahun 2021, serta mengetahui karakteristik yang ada dari setiap karakteristik yang terbentuk. Untuk selanjutnya dapat diambil keputusan dalam meningkatkan produksi Bawang Merah di Kabupaten Bima sehingga diharapkan mampu meningkatkan pendapatan petani Bawang Merah di Kabupaten Bima.

Banyak penelitian yang menggunakan metode *clustering K-Means* akan tetapi belum ada penelitian yang mengelompokkan data Jumlah Produksi dan luas panen bawang merah terkhusus pada daerah Kabupaten Bima.

Mengacu pada penjelasan sebelumnya, peneliti menjalankan sebuah penelitian yang memiliki potensi sebagai panduan bagi Pemerintah Kabupaten Bima dalam melaksanakan evaluasi atas hasil kelompok yang terbentuk serta ciri-ciri yang ada melalui judul penelitian "Pengelompokan Kecamatan di Kabupaten Bima Berdasarkan Jumlah Produksi dan Luas Panen Bawang Merah Tahun 2021 Menggunakan *K-Means Clusterin*".

METODE PENELITIAN

Statistika Deskriptif

Statistika merupakan ilmu yang paling berguna untuk mengumpulkan dan mendefinisikan data kuantitatif sedemikian rupa sehingga kemungkinan kesalahan dalam kesimpulan serta estimasi dapat diperkirakan dengan menggunakan penalaran induktif berdasarkan pada probabilitas

matematika. Statistika Deskriptif merupakan kumpulan teknik yang meliputi teknik pengumpulan, penyajian, dan peringkasan data [7].

Statistik deskriptif adalah bidang studi yang berfokus pada cara-cara dan teknik-teknik untuk mengatur dan menyajikan data yang telah terkumpul dalam sebuah penelitian. Misalnya, ini melibatkan penyusunan data dalam bentuk tabel frekuensi atau dalam bentuk grafik, serta menghitung nilai statistik seperti rata-rata, median, variasi, simpangan baku, dan sejenisnya [8].

Statistika deskriptif dapat digunakan untuk melakukan analisa dan menggambarkan sejumlah data hasil penelitian yang tidak digunakan dalam pengambilan kesimpulan yang sifatnya luas. Statistika deskriptif dapat berupa penyajian data dalam bentuk tabel distribusi frekuensi, grafik atau dalam bentuk lainnya yang tidak bertujuan sebagai pengambilan kesimpulan berkenaan dengan generalisasi [9].

Uji Multikolinearitas

Multikolinearitas artinya terdapat hubungan linier yang tinggi antara variabel yang menjelaskan model. Terjadinya Multikolinearitas dapat diketahui dari nilai koefisien korelasi dari masing-masing variabel independen. Tujuan dari Uji multikolinearitas adalah untuk mengetahui dalam suatu model regresi terdapat korelasi antar variabel independen atau variabel bebas yang signifikan. Hal tersebut berarti standar *error* besar, akibatnya ketika koefisien diuji, T-Hitung akan bernilai lebih kecil dari T-Tabel. Hal ini menunjukkan tidak adanya hubungan linier antara variabel independen atau variabel bebas yang dipengaruhi dengan variabel dependen atau variabel terikat [10].

Adanya multikolinearitas dapat mengakibatkan ketidaktepatan penggunaan metode regresi karena perkiraan regresinya menjadi tidak stabil dan koefisien variabel regresinya cenderung besar. Beberapa faktor yang dapat diamati untuk mengidentifikasi keberadaan multikolinearitas [11].

Untuk mengetahui ada atau tidaknya multikolinearitas pada suatu model regresi dapat dilihat dari nilai *tolerance* dan nilai *Variance Inflation Factor* (VIF). Nilai *tolerance* mengukur variabilitas dari variabel bebas yang terpilih dan tidak dapat didefinisikan oleh variabel bebas lainnya, jadi nilai *tolerance* yang rendah sama dengan nilai VIF tinggi, dikarenakan perhitungan $VIF = 1/tolerance$ dan menunjukkan terdapat korelasi yang tinggi. Nilai *cut off* yang digunakan adalah untuk nilai *tolerance* 0,10 atau nilai VIF di atas angka 10. Rumus VIF untuk koefisien dijabarkan sebagai berikut [12].

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

keterangan:

VIF : *Variance Inflation Factor*

R^2 : Koefisien determinasi

j : Jumlah Sampel

Berdasarkan rumus di atas, jika nilai $VIF < 10$ atau nilai *tolerance* $> 0,01$ maka dinyatakan tidak terjadi multikolinearitas. Jika nilai $VIF > 10$ atau nilai *tolerance* $< 0,01$ maka dinyatakan terjadi multikolinearitas.

Ashabul Akbar Maulana¹, Athallah Widyatama Rafii², Yulia Anggi Anjelina³,
Edy Widodo⁴/

Analisis *Clustering*

Analisis *Cluster* adalah salah satu dari metode analisis multivariat yang memiliki tujuan untuk mengelompokkan objek-objek berdasarkan suatu karakteristik yang dimiliki. Analisis *Cluster* meengelompokkan individu atau objek suatu penelitian, sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam satu *Cluster* yang sama. *Cluster-Cluster* yang terbentuk dalam suatu *Cluster* mempunyai ciri yang relatif mirip atau bahkan sama (homogen), sedangkan antar *Cluster* mempunyai ciri yang berbeda (heterogen). Pengelompokan ini dilakukan berdasarkan variabel yang diamati [13]. Analisis kluster adalah salah satu teknik dalam pembelajaran mesin yang memiliki tujuan utama dalam menemukan pola berinformasi atau pola yang bernilai pada kumpulan *big data* [14].

Untuk mendapatkan kelompok yang homogen, maka menggunakan dasar untuk mengelompokkan yaitu kesamaan nilai yang dianalisis. Semakin kecil nilai jarak pada suatu individu terhadap individu lain, maka semakin besar pula kemiripan individu tersebut. Data mengenai ukuran kesamaan tersebut dikelompokkan sehingga dapat ditentukan individu mana yang termasuk ke kelompok yang sesuai [13].

Algoritma *K-Means*

K-Means Clustering adalah salah satu jenis analisis kluster non-hirarki. Metode ini bertujuan untuk membagi suatu objek data kedalam satu atau lebih jumlah *Cluster* berdasarkan karakteristik tertentu [4]. Algoritma K-means merupakan salah satu metode algoritma partisional, karena K-means didasarkan pada pemilihan jumlah kelompok awal dengan menetapkan nilai awal pusatannya. Algoritma K-means mengadopsi langkah-langkah berulang untuk menghasilkan kelompok data yang konsisten [15].

Berikut adalah langkah-langkah dalam melakukan analisis *Clustering K-Means* [16]:

1. Menentukan secara acak jumlah data yang digunakan sebagai pusat *Cluster*
2. Menghitung jarak antara data dan pusat *Cluster*. Untuk menghitung semua jarak antara data dengan pusat *Cluster* dapat menggunakan rumus pada teori jarak euclidian yang dirumuskan sebagai berikut :

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + \dots + (X_{ki} - X_{kj})^2} \quad (2)$$

Keterangan:

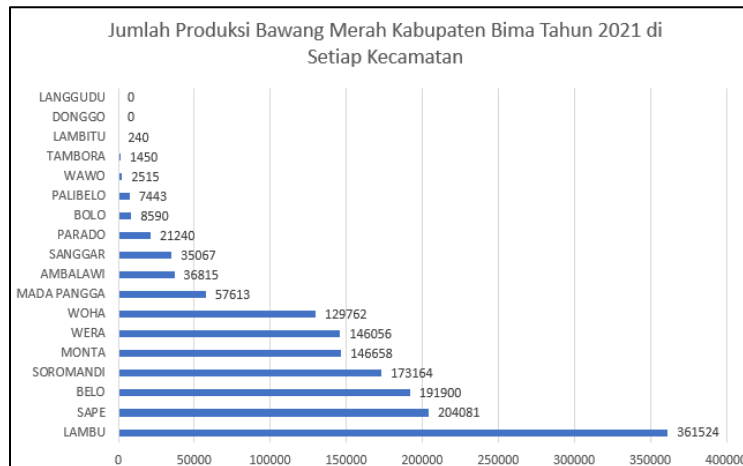
- $D(i, j)$: Jarak data ke-i ke pusat *Cluster* j
 X_{ki} : Data ke-i pada atribut data ke-k
 X_{kj} : Titik pusat ke-j pada atribut data ke-k

3. Data ditempatkan *Cluster* terdekat, dihitung dari pusat *Cluster*.
4. Pusat *Cluster* akan dihitung kembali setelah semua data ditempatkan pada *Cluster*-nya masing-masing.
5. Proses pembagian *Cluster* selesai apabila tidak terdapat perubahan pada nilai *centroid*.

HASIL DAN PEMBAHASAN

Analisis Deskriptif

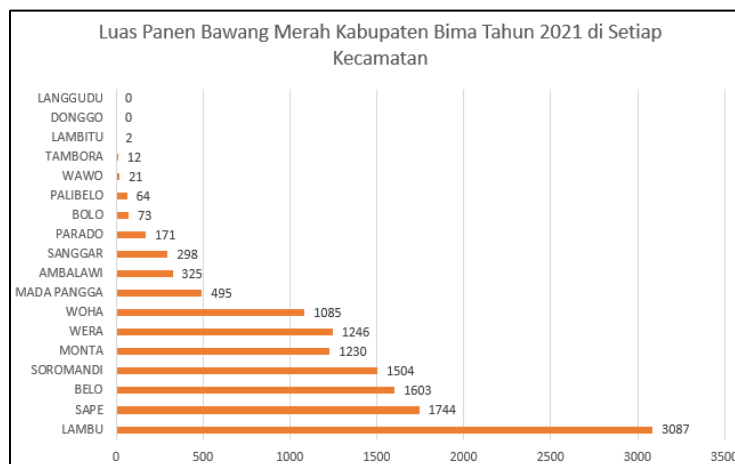
Tujuan dari analisis deskriptif adalah menyajikan data kedalam bentuk yang mudah dipahami dan menarik untuk diambil informasinya oleh pembaca [9]. *Bar chart* digunakan untuk melakukan visualisasi perbandingan jumlah produksi bawang merah untuk setiap Kecamatan yang ada di Kabupaten Bima.



Gambar 1. Jumlah Produksi Bwang Merah Kabupaten Bima Tahun 2021.

Diperoleh bahwa Kecamatan Lambu menjadi daerah dengan hasil produksi bawang merah terbanyak dengan jumlah produksi sebanyak 361524 ton, sedangkan Kecamatan Langgudu dan Donggo adalah daerah yang tidak memproduksi bawang merah pada tahun 2021. Karena daerah tersebut tidak cocok digunakan untuk menanam bawang merah karena keadaan georafis dan memiliki suhu terlalu dingin untuk pertanian bawang merah.

Kemudian peneliti membandingkan luas panen bawang merah kabupaten Bima tahun 2021 untuk setiap Kecamatan akan terlihat seperti pada **Gambar 2**.



Gambar 2. Luas Panen Bwang Merah Kabupaten Bima Tahun 2021.

Dapat dilihat pada **Gambar 2** bahwa Kecamatan Lambu adalah daerah dengan luas panen bawang merah terbesar dengan luas 3087 ha, sedangkan Kecamatan Langgudu dan Donggo menjadi daerah tanpa luas panen bawang Merah. Karena daerah tersebut tidak cocok digunakan untuk menanam bawang merah karena keadaan geografis dan memiliki suhu terlalu dingin untuk pertanian bawang merah.

Analisis Clustering K-Means

Peneliti melakukan proses Klasterisasi dengan tujuan untuk mengidentifikasi ciri khas dari masing-masing kelompok Klasterisasi yang terbentuk. Penelitian ini menggunakan dataset mengenai Jumlah Produksi Bawang Merah di Kabupaten Bima pada Tahun 2021. Sebelum menjalankan analisis Klasterisasi, terlebih dahulu dilakukan pengujian terhadap asumsi multikolinearitas melalui evaluasi nilai Toleransi dan Faktor Inflasi Varians (VIF). Hasil Klasterisasi diharapkan mampu memberikan arahan bagi pemerintah dalam meningkatkan produktifitas serta kesejahteraan petani bawang merah.

Untuk mengidentifikasi adanya multikolinearitas dalam suatu model regresi, dapat dianalisis melalui nilai toleransi dan Variance Inflation Factor (VIF). Toleransi mengukur variasi dari variabel independen yang tidak dapat dijelaskan oleh variabel independen lainnya, sehingga nilai toleransi yang rendah berhubungan dengan nilai VIF yang tinggi. Ini terjadi karena perhitungan $VIF = 1/\text{toleransi}$, yang mengindikasikan adanya korelasi yang kuat. Nilai ambang yang digunakan adalah toleransi 0,10 atau VIF melebihi 10.

Tabel 1. Uji Multikolinearitas.

Variabel	VIF	Tolerance
Luas Panen	1	1

Berdasarkan *output* pada **Tabel 1** didapatkan nilai *VIF* lebih kecil dari 10 dan nilai *Tolerance* lebih besar dari 0,1 dengan demikian menolak H_0 sehingga dinyatakan tidak terjadi Multikolinearitas antara variabel Jumlah Produksi dan Luas Panen.

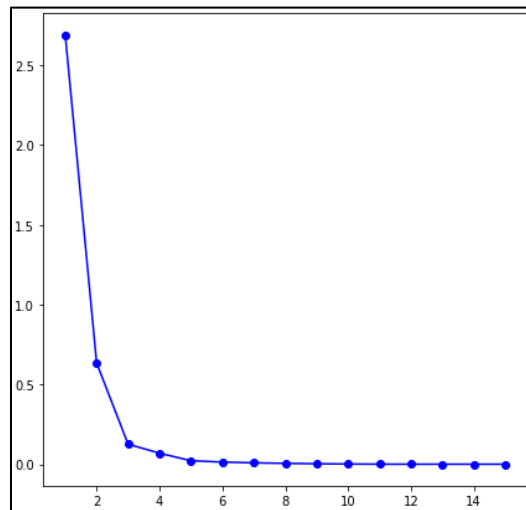
Setelah asumsi multikolinearitas terpenuhi, kemudian peneliti melakukan standarisasi data untuk membuat data memiliki rentang nilai yang sama hal ini disebabkan karena data memiliki satuan yang berbeda serta range data yang berbeda. Selain itu juga, standarisasi data dapat mempermudah dalam melakukan perhitungan. Hasil Standarisasi data dapat dilihat pada Tabel 2 berikut:

Tabel 2. Data Standarisasi.

Kecamatan	Jumlah Produksi	Luas Panen
Ambalawi	0.1018	0.1052
Belo	0.5308	0.5193
Bolo	0.0238	0.0236
Donggo	0.0000	0.0000
Lambitu	0.0007	0.0006
Lambu	1.0000	1.0000
Langgudu	0.0000	0.0000

Kecamatan	Jumlah Produksi	Luas Panen
Mada Pangga	0.1594	0.1603
Monta	0.4057	0.3984
Palibelo	0.0206	0.0207
Parado	0.0588	0.0554
Sanggar	0.0970	0.0965
Sape	0.5645	0.5649
Soromandi	0.4790	0.4872
Tambora	0.0040	0.0039
Wawo	0.0070	0.0068
Wera	0.4040	0.4036
Woha	0.3589	0.3514

Dari data yang telah di standarisasi, kemudian akan ditentukan *Cluster* yang terbentuk dengan menggunakan diagram *Elbow*.



Gambar 3. Diagram *Elbow* Untuk Jumlah *Cluster* Optimal.

Pada **Gambar 3** didapatkan bahwa jumlah *Cluster* optimal yang didapatkan dengan menggunakan metode *elbow* adalah sebanyak 3 *Cluster*. Sehingga, *data* Jumlah Produksi Bawang Merah Kabupaten Bima Tahun 2021 akan dibagi ke dalam 3 *Cluster*.

Setelah jumlah *Cluster* optimal didapatkan, peneliti menentukan pusat *Cluster* pada masing-masing variabel.

Tabel 3. Pusat *Cluster*.

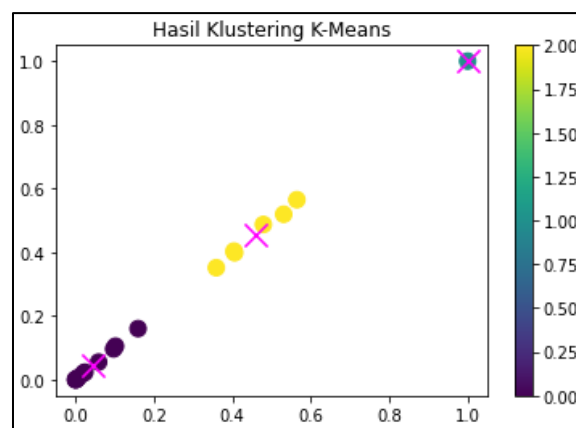
K	Pusat <i>Cluster</i>	
	Jumlah Produksi	Luas Panen
1	0.0430	0.0430
2	1	1
3	0.4571	0.4542

Didapatkan pada **Tabel 3** dengan menggunakan data yang di standarisasi bahwa pada variabel Jumlah Produksi, *Cluster 1* memiliki pusat pada titik 0.0430, *Cluster 2* berpusat pada titik 1, dan *Cluster 3* berpusat pada titik 0.4571. sedangkan, pada variabel Luas Panen, *Cluster 1* memiliki pusat pada titik 0.0430, *Cluster 2* berpusat pada titik 1, dan *Cluster 3* berpusat pada titik 0.4542. Dari hasil yang didapatkan, maka akan dilakukan pelabelan pada data asli sesuai dengan jarak terdekat dari pusat *Cluster*.

Tabel 4. Pelabelan *Data*.

Kecamatan	Jumlah Produksi	Luas Panen	Cluster
Ambalawi	36815	325	1
Bolo	8590	73	1
Donggo	0	0	1
Lambitu	240	2	1
Langgudu	0	0	1
Mada Pangga	57613	495	1
Palibelo	7443	64	1
Parado	21240	171	1
Sanggar	35067	298	1
Tambora	1450	12	1
Wawo	2515	21	1
Lambu	361524	3087	2
Belo	191900	1603	3
Monta	146658	1230	3
Sape	204081	1744	3
Soromandi	173164	1504	3
Wera	146056	1246	3
Woha	129762	1085	3

Dari *data* yang telah diberikan label dapat dilihat bahwa pada *Cluster 1* memiliki 11 anggota, untuk *Cluster 2* terdapat satu anggota, dan untuk *Cluster 3* terdapat 6 anggota. Kemudian peneliti akan menampilkan penyebaran data dari setiap *Cluster* yang terbentuk dalam *scatterplot*.



Gambar 4. Penyebaran *Data Anggota Cluster*.

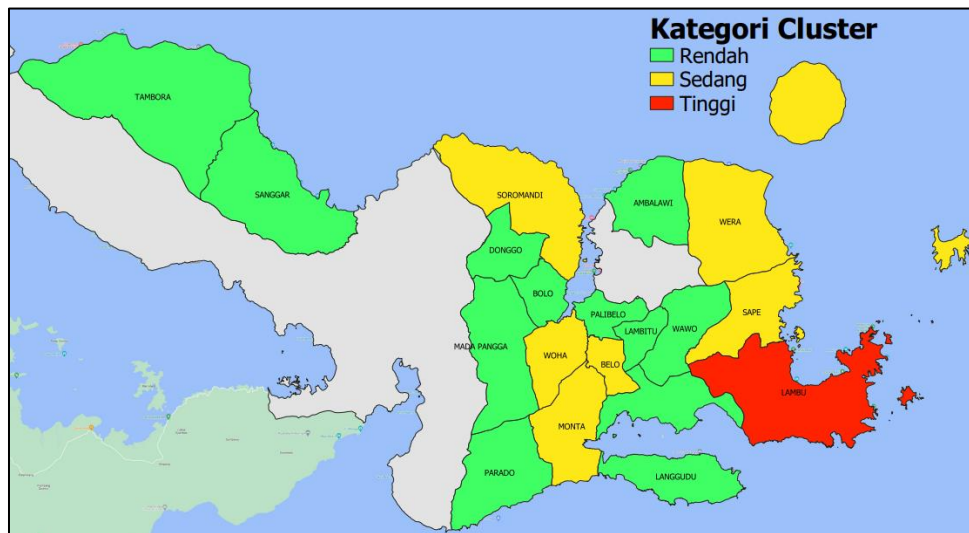
Pada **Gambar 4** dapat dilihat bahwa setiap data yang telah di kelompokkan dalam *Clusternya* masing-masing membentuk hubungan yang linear, dan jarak dari setiap data terhadap pusat *Cluster* yang tidak berjauhan, artinya karakteristik dari setiap anggota cluster memiliki kesamaan. Untuk melihat persamaan atau karakteristik dari setiap anggota *cluster*, yang dapat dilihat berdasarkan rata-rata dari setiap variabel berdasarkan *Clusternya* masing-masing.

Tabel 5. Karakteristik *Cluster*.

<i>Cluster</i>	Rata-rata	
	Jumlah Produksi	Luas Panen
1	15543.00	132.82
2	361524.00	3087.00
3	165270.67	1402.00

Pada **Tabel 5** dapat dilihat bahwa *Cluster 1* adalah *Cluster* yang memiliki anggota dengan rata-rata jumlah produksi dan luas panen bawang merah yang rendah, *Cluster 2* adalah *cluster* yang memiliki anggota dengan rata-rata jumlah produksi dan luas panen yang tinggi, dan *Cluster 3* adalah *Cluster* dengan anggota yang memiliki rata-rata jumlah produksi dan luas panen yang sedang.

Dari hasil tersebut, maka akan dilakukan visualisasi dalam bentuk peta untuk melihat penyebaran wilayah setiap *Cluster* berdasarkan karakteristiknya masing-masing yang dapat dilihat pada **Gambar 5**.



Gambar 5. Visualisasi Penyebaran Anggota Cluster.

KESIMPULAN

Dari hasil statistik deskriptif yang didapatkan, Kecamatan Lambu adalah daerah yang mempunyai Jumlah Produksi dan Luas Panen Bawang merah tertinggi dari seluruh Kecamatan yang ada di Kabupaten Bima. Sedangkan, Kecamatan Langgudu dan Donggo adalah daerah yang tidak memproduksi Bawang Merah karena letak geografis dan suhu di daerah tersebut tidak cocok

untuk proses produksi bawang merah. Berdasarkan hasil Profilisasi *Cluster*, didapatkan hasil bahwa *Cluster* 1 dengan 11 anggota adalah *Cluster* dengan kategori Rendah karena memiliki rata-rata jumlah produksi dan luas panen terendah dari tiga *cluster* yang ada, *Cluster* 2 adalah *Cluster* dengan Kategori Tinggi karena Kecamatan Lambu adalah daerah yang penduduknya didominasi oleh petani bawang merah, dan *Cluster* 3 adalah *Cluster* dengan kategori sedang.

DAFTAR PUSTAKA

- [1] Kementrian Perdagangan RI, 22 Juni 2015. [Online]. Available: http://bppp.kemendag.go.id/media_content/2017/08/Potensi_Bawang_Merah_di_Kabupaten_Bima.pdf.
- [2] R. O. Duda and P. E. Hart, Pattern classification and scene analysis, New York: Wiley, 1973.
- [3] A. Al-Wakeel and J. Wu, "K-means based cluster analysis of residential smart meter measurements," *Energy Procedia*, vol. 88, pp. 754-760, 2016.
- [4] A. N. Khomarudin, "Teknik Data Mining : Algoritma K-Means Clustering," *IlmuKomputer.com*, pp. 7-8, 2016.
- [5] Syafnidawaty, 19 April 2020. [Online]. Available: <https://raharja.ac.id/2020/04/19/k-means-clustering/#:~:text=Ada%20beberapa%20kelebihan%20pada%20algoritma,Mudah%20untuk%20diadaptasi..>
- [6] N. I. Febianto and N. D. Palasara, "Analisis Clustering K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018," *SISFOKOM*, vol. 8, no. 2, pp. 130-140, 2019.
- [7] H. H, Statistika Deskriptif, Banjarmasin: Poliban Press, 2019.
- [8] Maizar, I. Mustika and S. D. Nabella, PENGANTAR STATISTIK 1, Media Sains Indonesia, 2022.
- [9] B. R. Kustitunto B, Statistika 1 (Deskriptif), Depok, 1994.
- [10] Ghozali, Aplikasi Multivariate Dengan Program IBM SPSS 23, Semarang: Badan Penerbit Universitas Diponegoro, 2016.
- [11] V. I. Anggeyeny, Fear of Floating: Studi Empiris Sistem Nilai Tukar secara de facto di Indonesia dalam Periode 1994-2003, Depok: Universitas Indonesia, 2009.
- [12] Roswirman and Elazhari, "Pengaruh Implementasi Manajemen Mutu Terpadu dan Disiplin Kerja Terhadap Kinerja Guru pada Era New Normal di SMK Swasta PAB 2 Helvetia," *All Fields of Science Journal Liaison Academia and Society*, vol. 1, no. 4, pp. 316-333, 2021.
- [13] S. Hussein, "Analisis CLuster," Oktober 2021. [Online].
- [14] L. H. Witten and E. Frank, "Data mining: practical machine learning tools and techniques with Java implementations," *Acm Sigmod Record*, vol. 31, no. 1, pp. 76-77, 2002.
- [15] N. T. Hartanti, "Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional," *Nas. Teknol. dan Sist. Inf*, vol. 6, no. 2, pp. 82-89, 2020.
- [16] A. Muhidin and I. Baragigiratri, "Pemetaan Penduduk Calon Penerima Bantuan Renovasi Rumah Desa Pesangkalan Menggunakan Algoritma Clustering K-Means," *Jurnal SIGMA*, vol. 9, no. 3, pp. 75-82, 2019.