

## Klasterisasi Wilayah Kabupaten/Kota di Sulawesi Tenggara berdasarkan Produksi Bahan Pangan menggunakan Algoritma K-Means Clustering

Mursawal<sup>1</sup>, Rizal Adi Saputra<sup>2\*</sup>

<sup>1,2</sup> Jurusan Teknik Informatika, Fakultas Teknik, Universitas Halu Oleo

Email: mursyawal29@gmail.com<sup>1</sup>, rizaladisaputra@uho.ac.id<sup>2\*</sup>

### Abstrak

Bahan pangan merupakan bahan – bahan yang diproduksi dari hasil pertanian yang digunakan untuk membuat makanan. Bahan pangan tersebut terdiri dari sayuran, daging, kacang, ubi, dan lain sebagainya. Provinsi Sulawesi Tenggara menjadi salah satu provinsi yang memiliki jumlah produksi bahan pangan yang cukup tinggi di Indonesia. Penerapan algoritma K-Means clustering digunakan untuk mengelompokkan wilayah kabupaten/kota yang ada di Sulawesi Tenggara berdasarkan hasil produksi bahan pangan. Metode yang digunakan dalam penelitian ini adalah CRISP-DM dengan algoritma K-Means clustering. Terdapat 17 kabupaten/kota di Sulawesi Tenggara yang digunakan dalam penelitian ini. Jenis bahan pangan yang akan digunakan dalam penelitian ada 7 yaitu padi, jagung, ubi kayu, ubi jalar, kacang tanah, kacang kedelai, dan kacang hijau. Hasil dari penelitian ini adalah 1 kabupaten/kota yang memiliki tingkat produksi bahan pangan yang tinggi, 4 kabupaten/kota memiliki tingkat produksi bahan pangan yang sedang, dan 12 kabupaten/kota yang memiliki tingkat produksi bahan pangan yang rendah. Hasil pengujian menggunakan Davies bouldin index adalah cluster 2 yang mempunyai kualitas cluster yang terbaik karena hasil yang didapat dari cluster 2 yaitu 0.30 yang dimana semakin kecil hasil yang didapatkan maka semakin baik clusternya.

**Kata kunci:** Data Mining, K-Means, Pangan, Produksi, Wilayah

### Abstract

Food ingredients are ingredients produced from agricultural products that are used to make food. These foods consist of vegetables, meat, nuts, sweet potatoes, and so on. Southeast Sulawesi Province is one of the provinces that has a fairly high amount of food production in Indonesia. The application of the K-Means clustering algorithm is used to group districts/cities in Southeast Sulawesi based on food production results. The method used in this study is CRISP-DM with the K-Means clustering algorithm. There are 17 districts/cities in Southeast Sulawesi used in this study. There are 7 types of food ingredients that will be used in the study, namely rice, corn, cassava, sweet potatoes, peanuts, soybeans, and green beans. The results of this study are 1 district/city that has a high level of food production, 4 districts/cities have a moderate level of food production, and 12 districts/cities have a low level of food production. The test results using the Davies Bouldin index are cluster 2 which has the best cluster quality because the results obtained from cluster 2 are 0.30, where the smaller the results obtained, the better the cluster.

**Keywords:** Data Mining, K-Means, Food, Production, Region

### PENDAHULUAN

Sektor pertanian merupakan sektor yang mendapat perhatian besar dari pemerintah karena perannya dalam pembangunan ekonomi jangka panjang dan pemulihan ekonomi negara. Padi merupakan bahan pokok sehari – hari yang dibutuhkan setiap rumah tangga bahkan sebelum menjadi nasi. Sedangkan bahan lainnya seperti singkong, jagung, beras, kedelai, dan kacang tanah merupakan makanan pelengkap untuk setiap rumah tangga (Tendean & Purba, 2020). Masyarakat masih sangat bergantung pada

sektor pertanian, khususnya produksi pangan yang dikendalikan melalui penetapan harga pokok dan harga pangan maksimum. Tugas negara dan pemerintah daerah adalah menjaga stabilitas harga produk pertanian baik dalam kondisi kelebihan produksi maupun gagal panen, serta menjamin tidak adanya fluktuasi harga produk pertanian yang dapat berdampak buruk pada stabilitas harga barang dan jasa (Helbawanti dkk., 2021). Jumlah Penduduk yang semakin Bertambah memerlukan ketersediaan pangan yang dapat memenuhi kebutuhan

kelangsungan hidupnya. Pangan yang dibutuhkan penduduk meliputi hasil hortikultura karena merupakan salah satu dari sumber serat, vitamin, protein, dan bentuk nutrisi lain yang dibutuhkan tubuh manusia (Tando et al. 2019)

Bahan pangan merupakan salah satu dari kebutuhan pokok manusia dan tidak dapat tergantikan. Namun pertumbuhan penduduk yang dibarengi dengan berkurangnya luas lahan produktif dapat menyebabkan kerawanan pangan. Oleh karena itu, diperlukan upaya untuk menjaga ketahanan pangan guna memenuhi kebutuhan. Dalam beberapa tahun terakhir, isu ketahanan pangan telah menjadi isu yang sangat sensitif baik secara global maupun nasional (Ghazali & Nurhayati, 2018). K-Means adalah metode *clustering* yang termasuk bagian dari pendekatan partisi. Model *centroid* disertakan dalam algoritma K-Means. Model *centroid* adalah model yang menggunakan *centroid* untuk membuat *cluster*. *Centroid* digunakan untuk menghitung jarak ke *centroid* objek data. Suatu objek data dimasukkan ke dalam *cluster* jika mempunyai jarak terpendek ke pusat *cluster*. Algoritma K-Means dapat diartikan sebagai algoritma pembelajaran sederhana untuk menyelesaikan masalah pengelompokan yang bertujuan untuk meminimalkan kesalahan ganda (Hajar et al. 2020).

Temuan dari sistem penambahan dapat memberikan informasi penting. Metode pengelompokan data ini menggunakan metode data mining dengan menggunakan algoritma K-Means untuk mengelompokkan data (Nafsiah Muthmainnah et al. 2023). Dari beberapa penelitian sebelumnya yang menggunakan algoritma K-Means untuk analisis *clustering* yaitu (A'yuni et al. 2023), penelitian tersebut menggunakan metode algoritma K-Means yang dimana parameter yang digunakan berjumlah 4 yaitu program studi, semester, IPK, dan jenjang. Hasil pengujian kelayakan menggunakan RapidMiner dengan jumlah *cluster* 5 kemudian hasil pengujian yang dilakukan menggunakan metode DBI dengan nilai validitas sebesar 0.121 diperoleh *cluster* terbaik pada  $k=5$ . Penelitian selanjutnya dilakukan oleh (Fakhri dkk., 2021) dengan judul penelitian optimalisasi pelayanan perpustakaan terhadap minat baca menggunakan metode K-Means *clustering*. Hasil penelitian ini memiliki jumlah *cluster* 2 yang dimana penelitian ini dijadikan dasar rekomendasi terhadap pengoptimalisasian pelayanan perpustakaan baik dari tatak letak

buku, prioritas buku, dan lain sebagainya. Penelitian yang dilakukan oleh (Virgo, Defit, and Yuhandri 2020) dengan judul penelitian klasterisasi tingkat kehadiran dosen menggunakan algoritma K-Means *clustering*. Hasil penelitian tersebut memiliki jumlah *cluster* 3 yang dimana terdapat 72 matakuliah yang diampu dosen jarang melaksanakan pertemuan. Lalu terdapat 69 matakuliah yang tingkat pertemuannya sedang. Kemudian terakhir terdapat 1.370 matakuliah sering melakukan pertemuan.

Penelitian selanjutnya yang menggunakan algoritma K-Means dilakukan oleh (Hasanah, Defit, and Nurcahyo 2021) yang berjudul implementasi algoritma K-Means untuk klasterisasi peserta olimpiade sains nasional tingkat SMA. Berdasarkan penelitian tersebut hasil *cluster* yang didapatkan yaitu *cluster* 1 terdapat 7 siswa yang sangat berkompeten, kemudian *cluster* 2 terdapat 14 siswa yang berkompeten dan *cluster* 3 sebanyak 2 siswa yang kurang berkompeten. Penelitian sebelumnya tentang algoritma K-Means dilakukan juga oleh (Chandra et al. 2021) yang berjudul penerapan algoritma K-Means dalam mengelompokkan balita yang mengalami gizi buruk menurut provinsi. Dari hasil penelitian tersebut didapatkan 2 *cluster* yaitu tinggi dan rendah. *Cluster* dengan tingkat tinggi sebanyak 15 provinsi dan *cluster* dengan tingkat rendah sebanyak 19 provinsi. Hasil pengujian dilakukan melalui RapidMiner 5.3 dengan menggunakan algoritma K-Means. Penelitian sebelumnya tentang algoritma K-Means dilakukan juga oleh (Mulyadi 2022) dengan judul klasterisasi menggunakan metode algoritma K-Means dalam meningkatkan penjualan tupperware. Dari hasil pengujian data yang menggunakan perhitungan manual yaitu Z1 sama dengan *cluster* 0, Z2 data perhitungannya sama dengan *cluster* 1, dan Z3 data perhitungan manualnya sama dengan *cluster* 3 dengan pengujian dilakukan menggunakan RapidMiner. Hasil perhitungan akhir menunjukan Z1 adalah *cluster* Sangat Laris (SL) berjumlah 11 produk, Z2 adalah *cluster* Laris (L) yang berjumlah 12 produk dan Z3 adalah *cluster* Tidak Laris (TL) yang berjumlah 17 produk.

Dari permasalahan yang ada, begitu sangat diperlukannya proses pengelompokan yang nantinya dijadikan dasar dalam membuat kebijakan yang akan berpengaruh terhadap produksi bahan pangan setiap kabupaten/kota.

Pemerintah daerah juga nantinya akan mensosialisasikan ke wilayah tingkat produksi bahan pangan yang rendah untuk mendorong seluruh masyarakat dalam mengembangkan produksi bahan pangan mulai dari jagung, ubi kayu, ubi jalar, kacang tanah, kacang kedelai, kacang hijau, dan lain sebagainya. Oleh karena itu, penulis mencoba menerapkan metode algoritma K-Means *clustering* untuk mengelompokkan kabupaten/kota berdasarkan produksi bahan pangannya. Peneliti menggunakan K-Means karena relatif mudah dalam penerapannya dan mempunyai akurasi yang sangat tinggi terhadap ukuran benda sehingga relatif dapat diukur. Dengan metode ini dalam mengelompokkan data tingkat produksi bahan pangan setiap kabupaten/kota, akan lebih mudah untuk menganalisis daerah mana nantinya yang akan mempertimbangkan dukungan pengembangan produksi bahan pangan (Dwi Aulia dkk., 2023). Akurasi yang digunakan adalah metode *Davies Bouldin Index* (DBI) untuk mengetahui tingkat performa terbaik dari hasil Pengklasteran agar menjadi klaster yang terbaik.

## TINJAUAN PUSTAKA

Pada tahap awal penelitian ini, dilakukan tinjauan literatur untuk mencari penelitian terkait yang dapat mendukung penelitian ini. Beberapa penelitian sebelumnya menggunakan algoritma K-Means untuk menyelesaikan masalah yang berkaitan dengan produksi bahan pangan melalui pengelompokan wilayah berdasarkan produksi beras. Tujuan dari penelitian ini adalah melakukan klasterisasi wilayah berdasarkan produksi beras di Provinsi Sumatera Utara (Febriani, Safii, and Alfina 2022).

Penelitian lain dilakukan oleh (Novita 2022) yang melakukan pengelompokan menggunakan algoritma K-Means menggunakan metode elbow terhadap provinsi di Indonesia berdasarkan luas panen, produksi, dan produktivitas pangan komoditas. Penelitian ini dilakukan pengelompokan wilayah kota berdasarkan harga pangan pasar tradisional, sehingga luaran yang diperoleh dari hasil penelitian adalah jumlah klaster Provinsi di Indonesia dari tahun 2017 hingga 2019.

### 2.1 Bahan Pangan

Bahan pangan merupakan kebutuhan pokok setiap manusia, karena mengandung senyawa-senyawa yang sangat penting untuk pemulihan dan perbaikan jaringan tubuh yang rusak, reproduksi, dan produksi energi untuk berbagai aktivitas kehidupan. Komponen dari pangan terdiri dari protein, karbohidrat, lemak, vitamin, dan mineral (Prasetyo, Isdiana, and Sujadi 2019). Keanekaragaman hayati pangan sangat penting bagi struktur, fungsi, dan proses sistem produksi sebagai bagian penghidupan dan ketahanan pangan, serta berperan dalam penyediaan berbagai jasa ekosistem (Purwanto 2020).

### 2.2 Data Mining

Data Mining adalah alat yang memungkinkan pengguna mengakses data dalam jumlah besar dengan cepat. Definisi data mining yang lebih spesifik yaitu alat dan aplikasi yang menggunakan analisis data statistik. Data mining adalah proses mengekstraksi sejumlah data besar data atau informasi yang sebelumnya tidak diketahui dari database besar agar dapat dipahami, berguna, dan digunakan untuk membuat keputusan bisnis yang sangat penting (Zai 2022). Definisi sederhana dari data mining adalah ekstraksi informasi atau menarik dari data yang ada dalam database besar. Dalam jurnal ilmiah, data mining dikenal juga dengan *knowledge Discovery in Databases* (KDD). (Utami, 2020).

### 2.3 K-Means

K-Means adalah metode pengelompokan non-hierarki yang mempartisi data menjadi satu atau lebih grup. Ini mengelompokkan data dengan karakteristik yang sama ke dalam *cluster* yang sama, dan data dengan karakteristik berbeda ke dalam kelompok yang berbeda (Latipa Sari and Yati Beti 2023). K-Means *clustering* adalah algoritma yang dapat memecahkan kode kumpulan data dan mengungkapkan kecenderungan setiap tipe tertentu untuk digabungkan dengan tipe individual lainnya. Kecenderungan pengelompokan ini didasarkan pada kesamaan ciri-ciri individu dari data yang ada (Asyuti and Setyawan 2023). *Tools* yang akan digunakan adalah bahasa pemrograman *python*. K-Means berupaya mengelompokkan data yang ada ke dalam beberapa kelompok. Data pada satu kelompok mempunyai

karakteristik yang sama dan berbeda dengan data kelompok lain. Langkah-langkah algoritma pengelompokan K-Means (Haris Kurniawan, Sarjon Defit, and Sumijan 2020) :

Langkah 1 : Menentukan secara acak K buah data sebagai pusat *cluster*.

Langkah 2 : Menghitung jarak antara data dan pusat *cluster* menggunakan *Euclidian Distance*.

Langkah 3 : Menempatkan data pada jumlah terdekat dengan menghitung dari tengah *cluster*.

Langkah 4 : Menentukan pusat *cluster* baru jika semua data telah ditetapkan dalam *cluster* terdekat.

Langkah 5 : Menentukan dan menempatkan data dalam *cluster* kemudian diulangi sampai nilai *centroid* tidak berubah lagi.

## METODOLOGI PENELITIAN

Dalam Penelitian ini bisa dilakukan dengan menggunakan metode *Cross-Industry Standard Process for Data Mining* (CRISP-DM) Yang mempunyai 6 tahapan yaitu :

### 3.1 Business Understanding

Pemahaman bisnis pada penelitian ini yaitu berfokus pada rumusan masalah dari data 7 produksi bahan pangan pada setiap kabupaten/kota di Sulawesi Tenggara. Dari rumusan masalah ini , arah tujuan dari penelitian ini dapat ditentukan untuk membantu merumuskan kebijakan yang berkaitan dengan pengembangan produksi bahan pangan di Sulawesi Tenggara.

### 3.3 Data Preparation

Pada tahap ini data mentah dari laman <https://sultra.bps.go.id/> mengenai 7 produksi bahan pangan pada setiap wilayah kabupaten/kota di Sulawesi Tenggara. Kemudian dilakukan proses pembersihan data untuk membentuk dataset akhir yang optimal agar dapat menghasilkan pola pemodelan yang bagus. Proses ini terdiri dari integrasi data, reduksi data, pembersihan data, seleksi data, dan transformasi data.

### 3.4 Modelling

Pada tahap ini hal pertama yang dilakukan adalah menganalisis jumlah *cluster* yang optimal dengan menggunakan metode

*Elbow* dengan bahasa pemrograman *python* menggunakan *library Scikit-Learn* dan yang hal yang kedua melakukan proses pemodelan dengan metode *clustering* menggunakan algoritma K-Means.

### 3.5 Evaluation

Pada tahap ini, hasil pemodelan yang dilakukan menggunakan Davies-Bouldin Index (DBI) dievaluasi untuk mengetahui tingkat akurasi dan kinerja.

### 3.6 Deployment

Pada tahap yang terakhir ini akan melakukan penyusunan laporan dan memberikan hasil visualisasi dari hasil pengolahan data yang telah dilakukan.

## HASIL DAN PEMBAHASAN

### 4.1 Business Understanding

Penelitian ini memiliki fokus pada produksi bahan pangan disetiap wilayah kabupaten/kota di Sulawesi Tenggara. Permasalahan pengembangan dibidang pertanian di Sulawesi Tenggara adalah rendahnya tingkat produksi bahan pangan di Sulawesi Tenggara berdasarkan wilayah kabupaten/kota. Dari permasalahan ini tujuan yang akan dilakukan yaitu mengklasterisasi data produksi bahan pangan pada tiap wilayah kabupaten/kota di Sulawesi Tenggara dengan menggunakan algoritma K-Means.

### 4.2 Data Understanding

Pada tahap ini , yang dapat dilakukan ialah memahami dan menganalisis data yang didapatkan dan mengetahui apa yang dapat dilakukan pada data tersebut. Dataset yang digunakan dalam penelitian ini adalah data produksi bahan pangan tahun 2020 di Provinsi Sulawesi Tenggara. Sumber dataset berasal dari <https://sultra.bps.go.id/>. Dataset produksi bahan pangan terdiri dari 17 *records* yang dimana berisi , nama kabupaten/kota/kota dan disertai 7 atribut lainn yaitu padi, jagung, ubi kayu, ubi jalar, kacang tanah, kacang kedelai, dan kacang hijau. Berikut ini dataset yang dipakai dalam penelitian ini ditunjukkan pada Tabel 1. Dataset Awal.

Tabel 1. Dataset Awal

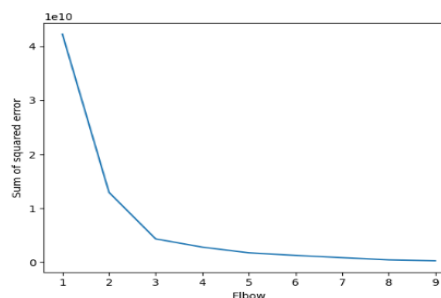
Kabupaten/kota/Kota	Padi	Jagung	Ubi_Kayu	Ubi_Jalar	Kacang_Tanah	Kacang_Kedelai	Kacang_Hijau
Buton	7598.08	2747	8543.55	1474.67	181.44	51.45	88.5
Muna	2898.34	23861	16555.28	3746.66	1087.3	7.14	50.6
Konawe	177444.57	11550	4415.18	1183.49	98.06	343.12	58.7
Kolaka	54657.11	9945	6691.22	1206.29	33.56	0	15.9
Konawe Selatan	82800.69	36493	26667.25	2236.66	153.67	210.01	24.7
Bombana	70459.83	2641	867.47	297.45	133.44	0	16
Wakatobi	0	163	13964.55	214.08	15.46	0	0
Kolaka Utara	6765.28	41808	469.53	401.92	29.17	37.72	1
Buton Utara	2238.57	811	10214.45	1486.04	46.4	87.81	45.5
Konawe Utara	7339.05	3539	797.38	515.72	210.92	61.66	74.7
Kolaka Timur	87867.03	15025	4049.48	1007.06	43.15	9.3	32.1
Konawe Kepulauan	835.73	195	5398.86	386.63	11.54	0	384.5
Muna Barat	5622.07	13681	12860.2	1938.68	1003.33	11.25	4.1
Buton Tengah	0	2867	17598.61	377.89	128.65	0	0
Buton Selatan	9.03	1500	24942.51	1867.45	9.22	0	2
Kota Kendari	2747.04	596	4184.11	456.67	6.49	0	2.9
Kota Bau-Bau	10424.51	395	1174.29	287.88	9.87	1.8	1

### 4.3 Data Preparation

Pada tahap ini mencakup semua aktivitas untuk membuat kumpulan data yang dimasukan ke dalam alat pemodelan dari data awal mentah awal atau untuk membuat database baru untuk pengaturan penambangan data. Ditahap ini akan dilakukan proses yang dimana terdiri dari pembersihan data, seleksi data, dan tranformasi data. Ditahap yang pertama nantinya akan dilakukan pembersihan data dilakukan menggunakan *python*, Kemudian ditahap yang kedua yaitu menyeleksi atribut yang akan digunakan, dan tahap ketiga mengubah yang semula nilai yang bertipe bilangan integer diubah menjadi bilangan real.

### 4.4 Modelling

Metode yang digunakan pada penelitian ini yaitu *clustering* dengan menggunakan algoritma K-Means. Sebelum melakukan pemodelan, harus dilakukan analisis untuk menentukan jumlah *cluster* yang optimal, yang dapat dilakukan dengan menggunakan metode *Elbow* diimplementasikan pada bahan pemrograman *python* dan hasilnya dapat dilihat pada Gambar 1.

Gambar 1. Hasil Metode *Elbow*

Pada Gambar 1. terlihat hasil kurva metode *Elbow* yang membentuk siku berada pada angka 2 dan 3, namun karena perbedaan garis kurva yang ada pada angka 3 maka jumlah *cluster* yang diambil adalah  $k=3$

Berdasarkan hasil metode *Elbow* jumlah *cluster* yang akan digunakan adalah 3 *cluster*. Sebelum melakukan *clustering* data dengan menggunakan K-Means, yang perlu dilakukan adalah menentukan titik pusat centroidnya. Tabel 2. menunjukan hasil *centroid* yang digunakan dengan pemrograman *python*.

Tabel 2. Titik Pusat *Centroid*

<i>Centroid</i>	Padi	Jagung	...	Kacang_Hijau
1	3.87E+03	7.68E+03	...	5.46E+01
2	7.39E+04	1.60E+04	...	2.22E+01
3	1.77E+05	1.16E+04	...	5.87E+01

Dari hasil *centroid* pada Tabel 2. menunjukkan ada 3 *cluster* akan dibentuk antara lain *cluster* 0, *cluster* 1, dan *cluster* 3. Hal selanjutnya yang harus dilakukan adalah proses *clustering* dengan menggunakan algoritma K-

Means pada *python*. Tabel 3. menunjukkan hasil hasil dari proses *clustering* yang menggunakan *python*.

Tabel 3. Hasil proses *clustering* dengan menggunakan *python*

Kab/Kota	Padi	Jagung	Ubi_Kayu	Ubi_Jalar	Kacang_Tanah	Kacang_Kedelai	Kacang_Hijau	Cluster
Buton	7598.08	2747	8543.55	1474.67	181.44	51.45	88.5	0
Muna	2898.34	23861	16555.28	3746.66	1087.3	7.14	50.6	0
Konawe	177444.6	11550	4415.18	1183.49	98.06	343.12	58.7	2
Kolaka	54657.11	9945	6691.22	1206.29	33.56	0	15.9	1
Konawe Selatan	82800.69	36493	26667.25	2236.66	153.67	210.01	24.7	1
Bombana	70459.83	2641	867.47	297.45	133.44	0	16	1
Wakatobi	0	163	13964.55	214.08	15.46	0	0	0
Kolaka Utara	6765.28	41808	469.53	401.92	29.17	37.72	1	0
Buton Utara	2238.57	811	10214.45	1486.04	46.4	87.81	45.5	0
Konawe Utara	7339.05	3539	797.38	515.72	210.92	61.66	74.7	0
Kolaka Timur	87867.03	15025	4049.48	1007.06	43.15	9.3	32.1	1
Konawe Kepulauan	835.73	195	5398.86	386.63	11.54	0	384.5	0
Muna Barat	5622.07	13681	12860.2	1938.68	1003.33	11.25	4.1	0
Buton Tengah	0	2867	17598.61	377.89	128.65	0	0	0
Buton Selatan	9.03	1500	24942.51	1867.45	9.22	0	2	0
Kota Kendari	2747.04	596	4184.11	456.67	6.49	0	2.9	0
Kota Bau-Bau	10424.51	395	1174.29	287.88	9.87	1.8	1	0

Dari hasil *clustering* pada Tabel 3. menunjukan bahwa produksi bahan pangan dengan tingkat rendah atau 0 adalah Buton, Muna, Wakatobi, Kolaka Utara, Buton Utara, Konawe Utara, Konawe Kepulauan, Muna Barat, Buton Tengah, Buton Selatan, Kota Kendari, Kota Bau-Bau. Selanjutnya untuk produksi bahan pangan dengan tingkat sedang atau 1 adalah Kolaka, Konawe Selatan, Bombana, dan Kolaka Timur. Kemudian terakhir produksi bahan pangan dengan tingkat tinggi adalah Konawe.

#### 4.5 Evaluation

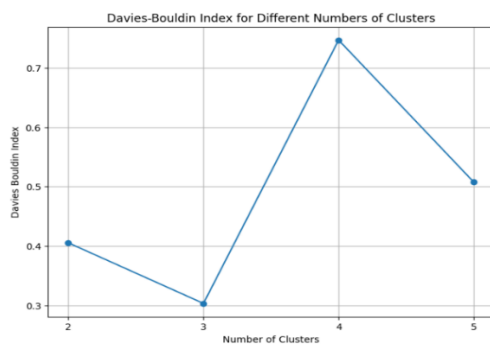
Pada tahap ini akan dilakukan evaluasi dengan menerapkan metode *Davies Bouldin Index* (DBI). Evaluasi ini dilakukan dengan menganalisis atau mengukur keakuratan model yang telah diterapkan. Evaluasi dilakukan untuk menganalisis hasil kinerja *cluster* produksi bahan pangan tahun 2020 dengan jumlah *cluster* 3 menggunakan metode *Davies Bouldin Index*. Adapun hasilnya yang memiliki nilai rendah maka kualitas dari *cluster* tersebut semakin baik. Tabel 4. adalah hasil pengujian dengan

menggunakan *Davies Bouldin Index* yang dijalankan pada bahasa pemrograman *python*.

Tabel 4. Hasil Pengujian menggunakan *Davies Bouldin Index* (DBI)

No	Jumlah Cluster	Davies Bouldin Index
1	2	0.405123
2	3	0.303042
3	4	0.766703
4	5	0.507467

Tabel 4. merupakan Hasil Pengujian menggunakan *Davies Bouldin Index* (DBI). didapatkan melalui pengujian *Davies Bouldin Index* dengan *cluster* 2 adalah dengan nilai terkecil sebesar 0.30 dimana jika nilai dari hasil pengujian melalui *Davies Bouldin Index* semakin rendah maka semakin baik *cluster* tersebut.

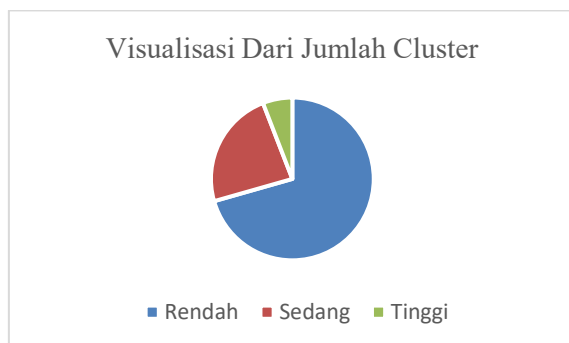


Gambar 4. Visualisasi terbaik dari hasil pengujian

Dari gambar visualisasi diatas dapat disimpulkan bahwa kurva yang mendekati nilai terkecil terlihat pada *cluster* 2 dengan jumlah *cluster* 3 yang dimana nilai *Davies Bouldin Index* (DBI) adalah 0.30 yang dimana semakin kecil nilai yang didapatkan maka semakin baik clusternya.

#### 4.6 Deployment

Setelah tahap evaluasi telah selesai untuk mengevaluasi hasil suatu model secara detail, maka akan dilakukan implementasi terhadap keseluruhan model yang dibangun. Dapat dilihat pada Gambar 5. Visualisasi jumlah *cluster*.



Gambar 5. Hasil visualisasi jumlah *cluster*

Pada Gambar 5. merupakan hasil dari dilakukannya proses dengan menggunakan bahasa pemrograman *python* dengan metode K-Means *Clustering* yang menjelaskan bahwa *cluster* 0 memiliki jumlah *cluster* sebanyak 71 % dengan tingkat produksi rendah, Lalu pada *cluster* 1 dengan jumlah *cluster* sebanyak 23% dengan tingkat produksi sedang, kemudian *cluster* 0 mendapatkan jumlah *cluster* sebanyak 6% dengan tingkat produksi tinggi. Hasil yang didapatkan ini bisa menjadi masukan pemerintah

daerah dalam membuat kebijakan yang berkaitan dengan pengembangan produksi bahan pangan di Sulawesi Tenggara.

#### SIMPULAN

Berdasarkan hasil penelitian yang dilakukan terhadap data produksi bahan pangan tiap wilayah kabupaten/kota di Sulawesi Tenggara menggunakan algoritma K-Means. Dalam penerapan metode algoritma K-Means untuk mengelompokkan kabupaten/kota menurut jumlah produksi bahan pangan dimulai dari atribut yang digunakan yang terdiri dari produksi Padi, produksi Jagung, produksi ubi kayu, produksi ubi jalar, produksi kacang tanah, produksi kacang kedelai, dan produksi kacang hijau. Dari atribut tersebut diperoleh 1 kabupaten/kota dengan produksi bahan pangan tingkat tinggi yaitu Konawe. 4 provinsi dengan produksi bahan pangan tingkat sedang yaitu Kolaka, Konawe Selatan, Bombana, dan Kolaka Timur dan 12 kabupaten/kota lainnya dengan tingkat produksi bahan yang rendah. Hasil pengujian menggunakan Davies bouldin index adalah *cluster* 2 yang mempunyai kualitas *cluster* yang terbaik karena hasil yang didapat dari *cluster* 2 yaitu 0.30 yang dimana semakin kecil hasil yang didapatkan maka semakin baik clusternya. *Clustering* yang dihasilkan nantinya dapat dijadikan masukan dalam membuat kebijakan yang berkaitan dengan produksi bahan di Sulawesi Tenggara yang dimana dapat dilihat dari kabupaten/kota mana yang menghasilkan jumlah produksi bahan pangan yang dominan terhadap padi, jagung, ubi kayu, ubi jalar, kacang tanah, kacang kedelai, dan kacang hijau.

Penelitian selanjutnya dapat mengintegrasikan data tambahan seperti luas lahan pertanian, tingkat penggunaan teknologi pertanian, jumlah tenaga kerja di sektor pertanian, dan faktor iklim untuk mendapatkan hasil klasterisasi yang lebih komprehensif. Selain itu, penerapan algoritma lain seperti DBSCAN atau Hierarchical Clustering dapat dipertimbangkan untuk membandingkan kualitas cluster dengan hasil dari algoritma K-Means. Pengembangan sistem berbasis dashboard analytics juga dapat dilakukan untuk memberikan visualisasi yang interaktif dan mempermudah pemangku kebijakan dalam mengambil keputusan strategis terkait produksi bahan pangan di wilayah tersebut.

## DAFTAR PUSTAKA

- Asyuti, Sulaiman, and Ahmad Aji Setyawan. 2023. 'DATA MINING DALAM PENGGUNAAN PRESENSI KARYAWAN DENGGA CLUSTER MEANS'. *Jurnal Ilmiah Sains Teknologi Dan Informasi* 1(1):01–10.
- A'yuni, Qurrata, Alwis Nazir, Lestari Handayani, and Iis Afrianty. 2023. 'Penerapan Algoritma K-Means Clustering Untuk Mengetahui Pola Penerima Beasiswa Bank Indonesia (BI)'. *Journal of Computer System and Informatics (JoSYC)* 4(3):530–39. doi: 10.47065/josyc.v4i3.3343.
- Chandra, Muhammad Dwi, Eka Irawan, Ilham Syahputra Saragih, Agus Perdana Windarto, and Dedi Suhendro. 2021. 'Penerapan Algoritma K-Means Dalam Mengelompokkan Balita Yang Mengalami Gizi Buruk Menurut Provinsi'. *BIOS: Jurnal Teknologi Informasi Dan Rekayasa Komputer* 2(1):30–38. doi: 10.37148/bios.v2i1.19.
- Dwi Aulia, Diana, Jurusan Sistem Informasi, Fakultas Ilmu Komputer, and Universitas Darwan Ali. 2023. 'Klasterisasi Pendidikan Masyarakat Untuk Mengetahui Daerah Dengan Pendidikan Terendah Menggunakan Algoritma K-Means'. *Jurnal Informatika Dan Rekayasa Perangkat Lunak* 5(1):38–44.
- Fakhri, Dwiki Aulia, Sarjon Defit, and Sumijan. 2021. 'Optimalisasi Pelayanan Perpustakaan Terhadap Minat Baca Menggunakan Metode K-Means Clustering'. *Jurnal Informasi Dan Teknologi* 3(3):160–66. doi: 10.37034/jidt.v3i3.137.
- Farida Utami, Sri. 2020. 'Penerapan Data Mining Algoritma Decision Tree Berbasis PSO'. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)* 677–81.
- Febriani, Intan, M. Safii, and Ommi Alfina. 2022. 'IMPLEMENTASI DATA MINING PENINGKATAN PRODUKSI BERAS MENGGUNAKAN METODE K-MEANS CLUSTERING'. *MAJALAH ILMIAH METHODA* 12(3):258–68. doi: 10.46880/methoda.Vol12No3.pp258-268.
- Ghazali, Mursal, and dan Nurhayati. 2018. 'PELUANG DAN TANTANGAN PENGEMBANGAN MAKROALGA NON BUDIDAYA SEBAGAI BAHAN PANGAN DI PULAU LOMBOK'. *AGROTEK* 5(2):135–40.
- Hajar, Siti, Asril Andi Novany, Agus Perdana Windarto, Anjar Wanto, and Eka Irawan. 2020. 'Penerapan K-Means Clustering Pada Ekspor Minyak Kelapa Sawit Menurut Negara Tujuan'. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)* 314–18.
- Haris Kurniawan, Sarjon Defit, and Sumijan. 2020. 'Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal'. *Journal of Applied Computer Science and Technology* 1(2):80–89. doi: 10.52158/jacost.v1i2.102.
- Hasanah, Miftahul, Sarjon Defit, and Gunadi Widi Nurcahyo. 2021. 'Implementasi Algoritma K-Means Untuk Klasterisasi Peserta Olimpiade Nasional Tingkat SMA'. *Jurnal Sistim Informasi Dan Teknologi* 1(3):30–35. doi: 10.35134/jsisfotek.v1i3.6.
- Helbawanti, Octaviana, Wahyu Adhi Saputro, and Amalia Nadipta Ulfa. 2021. 'PENGARUH HARGA BAHAN PANGAN TERHADAP INFLASI DI INDONESIA'. *Agrisaintifika Jurnal Ilmu-Ilmu Pertanian* 5(2):107–16.
- Latipa Sari, Herlina, and Ila Yati Beti. 2023. 'Penerapan Data Mining Dalam Pengelompokkan Buku Yang Dipinjam Menggunakan Algoritma K-Means'.



- (Media Online) 3(6):925–33. doi: 10.30865/klik.v3i6.826.
- Mulyadi, Iriene Putri. 2022. ‘Klasterisasi Menggunakan Metode Algoritma K-Means Dalam Meningkatkan Penjualan Tupperware’. *Jurnal Informatika Ekonomi Bisnis* 4(4):172–79. doi: 10.37034/infec.v4i4.164.
- Novita, Aditya. 2022. ‘KLAUSTERISASI PROVINSI DI INDONESIA BERDASARKAN PRODUKTIVITAS KOMODITAS PANGAN MENGGUNAKAN ALGORITMA K-MEANS’.
- Nafsiah Muthmainnah, Tasyifa, Siti Indriyana, Ultach Enri, Jurusan Sistem Informasi, Fakultas Ilmu Komputer, and Universitas Singaperbangsa Karawang. 2023. ‘Penerapan Algoritme K-Means Dalam Mengelompokkan Data Pengangguran Terbuka Di Provinsi Jawa Barat’. 5(2):122–29.
- Prasetyo, Tri Ferga, Abghi Firas Isdiana, and Harun Sujadi. 2019. ‘Implementasi Alat Pendeteksi Kadar Air Pada Bahan Pangan Berbasis Internet Of Things’. *SMARTICS Journal* 5(2):81–96. doi: 10.21067/smartics.v5i2.3700.
- Purwanto, Yohanes. 2020. ‘Penerapan Data Etnobiologi Sebagai Wahana Mendukung Pengelolaan Sumber Daya Hayati Bahan Pangan Secara Berkelanjutan’. *PROS SEM NAS MASY BIODIV INDON* 6(1):470–83. doi: 10.13057/psnmbi/m060101.
- Tando, Edi, Balai Pengkajian, Teknologi Pertanian, and Sulawesi Tenggara. 2019. ‘REVIEW : PEMANFAATAN TEKNOLOGI GREENHOUSE DAN HIDROPONIK SEBAGAI SOLUSI MENGHADAPI PERUBAHAN IKLIM DALAM BUDIDAYA TANAMAN HORTIKULTURA’. *Buana Sains* 19(1):91–102.
- Tendean, Tonny, and Windania Purba. 2020. ‘Analisis Cluster Provinsi Indonesia Berdasarkan Produksi Bahan Pangan Menggunakan Algoritma K-Means’. *Jurnal Sains Dan Teknologi* 1(2):5–11.
- Virgo, Ismail, Sarjon Defit, and Y. Yuhandri. 2020. ‘Klasterisasi Tingkat Kehadiran Dosen Menggunakan Algoritma K-Means Clustering’. *Jurnal Sistim Informasi Dan Teknologi* 23–28. doi: 10.37034/jsisfotek.v2i1.17.
- Zai, Charles. 2022. ‘IMPLEMENTASI DATA MINING SEBAGAI PENGOLAHAN DATA’. *Portaldata.Org* 2(3):1–12.