



# Pendekatan *K-Means Clustering* untuk Mengidentifikasi Wilayah Berpotensi Lumbung Pangan di Jawa Tengah

Nur Hanifah Miftahul Jannah<sup>1</sup>, Naila Kamilia Hasna Safitri<sup>2</sup>, Khrisna Aji Pamungkas<sup>3</sup>,  
Robert Kurniawan<sup>4</sup>

<sup>1, 2, 3</sup>Program Studi Statistika, Politeknik Statistika STIS, Jakarta, Indonesia

<sup>4</sup>Program Studi Komputasi Statistik, Politeknik Statistika STIS

<sup>1</sup>[212112270@stis.ac.id](mailto:212112270@stis.ac.id)

<sup>3</sup>[212112136@stis.ac.id](mailto:212112136@stis.ac.id)

<sup>4</sup>[robertk@stis.ac.id](mailto:robertk@stis.ac.id)

Corresponding author email: [212112239@stis.ac.id](mailto:212112239@stis.ac.id)

**Abstract:** Population dynamics along with the times certainly have an impact on the environment, one of which is the issue of vegetation which arises from various factors. The aim of this research is to determine the regional grouping of potential food production areas in Central Java Province based on NDVI data, Population Density, Average Expenditure on Food, Rice Productivity and Temperature. The analytical method used in this research is *K-Means Clustering* by determining the best number of clusters using the *Silhouette* method. The research results showed that 2 clusters were formed for Central Java Province based on several variables used. Most areas in Central Java are included in the green cluster, namely clusters with the characteristics of low population density, low average expenditure on food, quite high rice productivity, quite high vegetation, and varying temperatures.

**Keywords:** NDVI, *K-Means*, *Clustering*, *Silhouette*, Climate, Social

**Abstrak:** Dinamika populasi seiring dengan perkembangan zaman tentu memiliki dampak terhadap lingkungan, salah satunya adalah terhadap isu vegetasi yang muncul dari berbagai macam faktor. Tujuan dari penelitian ini adalah untuk mengetahui pengelompokan wilayah berpotensi lumbung pangan di Provinsi Jawa Tengah berdasarkan data NDVI, Kepadatan Penduduk, Rata-rata Pengeluaran untuk Makanan, Produktivitas Padi, dan Suhu. Metode analisis yang digunakan pada penelitian ini adalah menggunakan *K-Means Clustering* dengan penentuan jumlah cluster terbaik menggunakan metode *Silhouette*. Hasil penelitian menunjukkan bahwa terbentuk 2 cluster untuk Provinsi Jawa Tengah berdasarkan beberapa variabel yang digunakan. Sebagian besar wilayah di Jawa Tengah termasuk ke dalam cluster hijau, yaitu cluster dengan karakteristik kepadatan penduduk cenderung rendah, rata-rata pengeluaran untuk makanan rendah, produktivitas padi cukup tinggi, vegetasi cukup tinggi, dan suhu yang bervariasi.

**Kata kunci:** NDVI, *K-Means*, *Clustering*, *Silhouette*, Iklim, Sosial

## I. PENDAHULUAN

Urbanisasi merupakan peristiwa yang bisa berdampak pada kualitas lingkungan. Urbanisasi yang terjadi saat ini bisa menurunkan kualitas ekologi, khususnya pada wilayah yang memiliki pendapatan tinggi [1]. Salah satu dampaknya adalah permasalahan mengenai vegetasi. Vegetasi sendiri merupakan kumpulan dari berbagai macam jenis tumbuhan yang tumbuh bersama-sama di dalam suatu ekosistem. Vegetasi banyak menjadi perhatian karena saat ini banyak hutan maupun tutupan lahan hijau lainnya yang terkena dampak dari masalah lingkungan maupun campur tangan manusia.

Salah satu alat ukur kualitas vegetasi adalah indeks vegetasi. Indeks vegetasi merupakan penerapan dari suatu algoritma dengan menggunakan citra satelit yang dapat memperlihatkan aspek vegetasi ataupun aspek lainnya [2]. NDVI atau *Normalized Difference Vegetation Index* merupakan indeks vegetasi yang bisa menggambarkan vegetasi dengan sederhana dan bisa meningkatkan analisis tentang vegetasi dengan data penginderaan jarak jauh, serta efektif untuk membedakan berbagai jenis lahan atau hutan [3]. Saat ini, NDVI menjadi alat ukur kualitas vegetasi yang paling sering digunakan dalam berbagai analisis di mana menjadi alat untuk menganalisa, memantau, dan mengelola vegetasi sebagai acuan guna memperbaiki berbagai macam masalah.



Berdasarkan data yang diperoleh dari [4], Indonesia merupakan negara keempat dengan jumlah penduduk terbanyak di dunia, yaitu lebih dari 270 juta penduduk. Namun, sebagian besar penduduk terpusat tinggal di Pulau Jawa. Hal ini disebabkan oleh banyaknya penduduk yang bermigrasi ke Pulau Jawa dengan tujuan untuk mendapatkan sumber penghasilan dan fasilitas penunjang kehidupan yang lebih baik. Menurut [4], jumlah penduduk yang tinggal di Pulau Jawa pada tahun 2022 sebesar 154.280 orang. Padahal, jika dibandingkan dengan keempat pulau besar lainnya yang ada di Indonesia, Pulau Jawa merupakan pulau dengan luas wilayah terkecil, yaitu sekitar 128.297 km<sup>2</sup>.

Provinsi Jawa Tengah menduduki urutan ke-4 provinsi dengan penduduk terpadat di Pulau Jawa tahun 2021. Meskipun begitu, Provinsi Jawa Tengah juga merupakan provinsi ketiga dengan luas tutupan lahan (hutan dan non hutan) terbesar di Pulau Jawa, sehingga hasil dari kekayaan alam yang dapat dimanfaatkan dari Provinsi Jawa Tengah juga tinggi. Provinsi Jawa Tengah juga menjadi sandaran produksi nasional dalam produksi berbasis tanaman pangan dan hortikultura [5]. Semakin tinggi produksi yang dihasilkan oleh suatu wilayah, ketahanan pangan di wilayah tersebut juga semakin baik. Jika dilihat dari angka produksi padi terbanyak di Indonesia, maka Provinsi Jawa Timur menduduki urutan pertama provinsi dengan jumlah produksi beras terbanyak di Indonesia [6]. Namun, bila dilihat dari Indeks Ketahanan Pangan (IKP), nilai IKP yang dimiliki Provinsi Jawa Tengah lebih tinggi dibandingkan nilai yang dimiliki Provinsi Jawa Timur, yaitu menduduki peringkat kedua teratas setelah Provinsi Bali pada tahun 2020 sampai dengan 2022 [7]. Hal ini mengindikasikan bahwa masih banyak sektor produksi dan wilayah yang berpotensi untuk menjadi lumbung pangan di Provinsi Jawa Tengah.



Gambar 1. Grafik NDVI menurut kabupaten/kota di Provinsi Jawa Tengah tahun 2023

Angka NDVI kabupaten/kota di Provinsi Jawa Tengah nilainya berkisar antara 0.1 - 0.24. Nilai NDVI ini masih sangat kecil atau tergolong di bawah vegetasi yang jarang [8]. Rata-rata wilayah yang memiliki nilai NDVI cenderung kecil merupakan wilayah perkotaan, sedangkan wilayah kabupaten nilainya cenderung lebih tinggi. Namun, Kota Tegal merupakan salah satu wilayah perkotaan yang memiliki nilai NDVI yang cukup tinggi jika dibandingkan dengan wilayah perkotaan lainnya. Hal ini tentu disebabkan oleh beberapa faktor lain yang memiliki pengaruh terhadap tinggi/rendahnya nilai NDVI di suatu daerah, tidak hanya berdasarkan kepadatan penduduk saja. Vegetasi di suatu wilayah juga bisa dipengaruhi oleh letak wilayah tersebut. Pada daerah dataran rendah, penurunan pertumbuhan tanaman bisa disebabkan oleh terjadinya kenaikan suhu pada wilayah tersebut [9]. Artinya, setiap wilayah tentu memiliki peluang untuk mengembangkan potensi alam yang dimilikinya.

## NDVI

Penelitian dengan analisis *clustering* indeks vegetasi yang dilakukan oleh [10, p. 8] menggunakan metode K-Means untuk melihat pengelompokan wilayah berdasarkan curah hujan, sehingga nantinya



bisa menjadi dasar bagi pemerintah untuk melakukan pengelolaan vegetasi per wilayah karena indeks vegetasi yang mengalami perubahan dari tahun ke tahun tidak terlepas dari faktor curah hujan.

### **Kepadatan Penduduk**

Indonesia merupakan negara dengan kekayaan alam yang sangat berlimpah, khususnya kekayaan alam yang dihasilkan dari sektor kehutanan. Kekayaan yang terkandung tentu memiliki manfaat tersendiri bagi daerah tersebut dan tentunya hal ini akan bervariasi pada setiap wilayah, bergantung pula dengan karakteristik lain yang dimiliki wilayah tersebut. Daerah yang memiliki banyak lahan subur tentu bisa memanfaatkan hasil kekayaan alamnya menjadi keuntungan bagi mereka. Namun, daerah yang mengalami kelangkaan lahan subur akibat dari tingginya tekanan populasi, tentu akan mengalami penurunan pada jumlah lahan yang bisa ditanami [11].

### **Rata-rata pengeluaran untuk makanan**

Dari sisi ekonomi, pemanfaatan sumber daya potensial yang dimiliki oleh masing-masing wilayah bisa membantu alam untuk menjaga keanekaragaman hayatinya dan juga memberi dampak positif terhadap perekonomian lokal [12]. Namun demikian, pertumbuhan ekonomi yang pesat juga bisa menimbulkan berbagai macam permasalahan lingkungan hidup, seperti degradasi lingkungan atau menurunnya kualitas lingkungan akibat kerusakan yang terjadi [13]. Hal ini menunjukkan bahwa permasalahan yang timbul berkaitan dengan lingkungan tidak akan terlepas dari permasalahan ekonomi yang disebabkan oleh aktivitas atau perilaku dari penduduk yang tinggal di wilayahnya.

### **Produktivitas Padi**

Hasil dari penelitian yang pernah dilakukan oleh [14] yang menyatakan bahwa NDVI memiliki korelasi kuat dengan fase pertumbuhan padi. Prediksi hasil padi petani kecil dapat dilakukan dengan menggunakan NDVI yang nantinya bisa membantu dalam pengoptimalan penggunaan pupuk dan pengelolaan airnya [15]. Hal ini mengindikasikan bahwa tingginya nilai NDVI menunjukkan keberadaan kualitas vegetasi yang baik pada daerah tersebut.

### **Suhu**

Penelitian lainnya yang dilakukan oleh [16] untuk membandingkan tiga metode *clustering* untuk melihat pengelompokan wilayah berdasarkan kehomogenan suhu dan curah hujan. Metode K-Means yang digunakan pada penelitian ini menunjukkan hasil yang terbaik untuk pengelompokan data suhu dan curah hujan di Amerika Selatan, sehingga bisa terlihat regionalisasi suhu dan curah hujan masing-masing wilayah. Hasil dari penelitian ini berpotensi untuk prakiraan iklim yang lebih baik.

Beberapa studi yang telah dilakukan sebelumnya hanya menganalisis *clustering* vegetasi dengan beberapa faktor tertentu saja. Namun, masih sangat jarang ditemui literatur yang membahas mengenai *clustering* dengan mengombinasikan beberapa faktor tersebut sehingga terlihat kondisi vegetasi di beberapa wilayah berdasarkan beberapa faktor yang memiliki keterkaitan. Selain itu, masih sangat jarang ditemukan literatur studi yang melakukan perbandingan *clustering* pada setiap tahun di beberapa tahun terakhir. Padahal, hal ini sangat penting untuk dilakukan guna mendapatkan gambaran bagaimana keadaan vegetasi beserta faktor pendukungnya di Provinsi Jawa Tengah. Harapannya *clustering* ini dapat bermanfaat dalam penentuan kebijakan berkaitan ketahanan pangan di Jawa Tengah dan sekitarnya. Ini juga bisa menjadi gambaran bagi pengelola sumber daya alam agar bisa mengambil langkah sesuai dengan kondisi di masing-masing wilayah agar kondisi vegetasi di setiap wilayah tetap terjaga dengan baik serta dengan mempertimbangkan faktor-faktor lainnya. Oleh karena itu, penelitian ini bertujuan untuk mengidentifikasi pola kelompok yang terbentuk dari kabupaten/kota di Provinsi Jawa Tengah berdasarkan beberapa variabel yang digunakan dan bagaimana perubahan yang terjadi setiap tahunnya serta melihat kabupaten/kota apa saja yang memiliki potensi sebagai lumbung pangan.

## II. METODE PENELITIAN

### 2.1 Wilayah

Wilayah yang dipilih dalam penelitian ini adalah seluruh kabupaten dan kota yang ada di Provinsi Jawa Tengah, terdiri atas 29 kabupaten dan 6 kota. Dengan luas 32.800,69 km<sup>2</sup>.

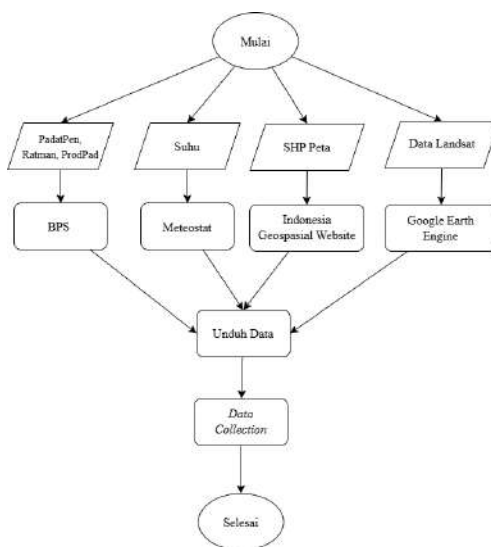
### 2.2 Data dan Sumber Data

Sumber data yang digunakan dalam penelitian ini berasal dari laman resmi BPS, Meteostat, dan dataset Landsat 8 dengan batas Shp wilayah kabupaten/kota dari Indonesia Geospasial [17-21]. Data diolah menggunakan R Studio dan *platform* komputasi berbasis *cloud* yaitu Google Earth Engine.

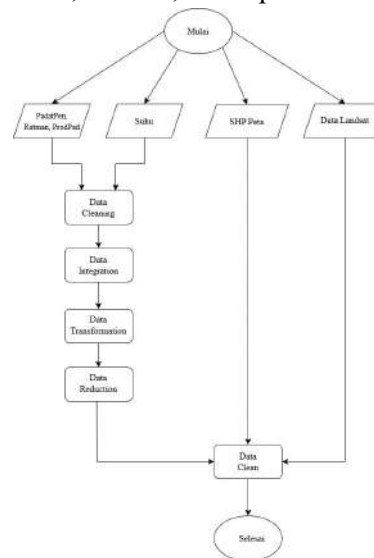
Tabel 1. Definisi Variabel Operasional

Variabel	Definisi	Satuan
Kepadatan Penduduk	Kepadatan penduduk adalah banyaknya penduduk per satuan luas.	per km <sup>2</sup>
Rata-Rata Pengeluaran untuk Makanan	Rata-rata Pengeluaran per Kapita per Bulan Makanan	Rupiah
Produktivitas Padi	Jumlah produksi padi yang dihasilkan per satuan luas lahan dalam suatu periode tertentu.	Kuwintal/ Hektar
Normalized Difference Vegetation Index (NDVI)	Indeks yang digunakan untuk mengukur kesehatan vegetasi atau tingkat kehijauan tanaman.	-
Suhu	Suhu rata-rata suatu wilayah	Celcius

Data yang diperoleh perlu dilakukan preprocessing karena jenis data statistik dan citra satelit memiliki karakteristik berbeda. Data tersebut melalui *data cleaning*, *data integration*, *data transformation*, dan *data reduction* agar analisisnya dapat akurat, relevan, dan representatif.



Gambar 2a.



Gambar 2b.

Gambar 2 merupakan Tahapan Pengerjaan: (2a) Alur Pengumpulan Data (2b) Preprocessing Data

Terdapat beberapa metode untuk memetakan kerapatan vegetasi menggunakan data citra satelit, salah satunya menggunakan analisis NDVI. NDVI dihitung dari data citra landsat-8 dengan melalui kode di *Google Earth Engine* yang telah dilakukan *preprocessing data*. Analisis NDVI dihitung menggunakan pantulan piksel pita merah (*visible*) dan pita *Near Infrared* (NIR). Nilai indeks NDVI digunakan untuk melihat kerapatan vegetasi karena menangkap jumlah reflektansi produksi klorofil pada tanaman [22]. Vegetasi dengan produksi klorofil yang baik akan menyerap gelombang *Red* dan



memancarkan gelombang NIR lebih tinggi, begitupun sebaliknya. Nilai tersebut mempunyai rentang antara -1,0 hingga 1.0. Semakin tinggi derajat aktivitas klorofil, nilai NDVI juga semakin besar.

$$NDVI = \frac{NIR-Red}{NIR + Red} \quad (1)$$

### 2.3 Metode

Metode analisis yang digunakan dalam penelitian ini adalah analisis *cluster*. Analisis *cluster* dilakukan untuk mengetahui bagaimana proses setelah persiapan data kemudian dihasilkan kelompok-kelompok pengamatan yang optimal. Dalam konteks ini, hasil yang diharapkan adalah terciptanya kelas-kelas dengan anggota yang memiliki karakteristik serupa (homogen) di dalam kelas, namun berbeda (heterogen) antar kelas yang satu dengan yang lain. Asumsi yang perlu untuk dilakukan adalah multikolinearitas dengan menggunakan matriks korelasi. Adanya multikolinearitas ketika nilai korelasi antar variabel bernilai lebih dari 0.95 [23].

Secara umum, terdapat dua metode analisis *cluster*, yaitu hierarki dan non-hierarki. Metode hierarki mengelompokkan pengamatan secara terstruktur berdasarkan kesamaan antar pengamatan. Sebaliknya, metode non-hierarki mengelompokkannya secara tidak terstruktur dengan jumlah kelas yang sudah ditentukan sejak awal analisis. Seluruh kabupaten/kota dikelompokkan berdasarkan karakteristik variabel-variabel yang telah ditentukan dan selanjutnya akan dipetakan sehingga terlihat apakah wilayah yang dekat cenderung mempunyai karakteristik yang sama atau tidak. Salah satu metode *clustering* non hierarki adalah metode K-Means. Tujuannya untuk mengelompokkan pengamatan sedemikian hingga jarak antar pengamatan dengan centroid seminimum mungkin. Pemilihan algoritma *clustering* ini didasarkan kapasitas kinerja yang tinggi pada kumpulan data besar dengan tetap sederhana dalam proses komputasinya tetapi sensitif terhadap outlier [24].

### 2.4 Evaluasi Hasil

#### 1. Index Connectivity

Nilai Indeks Connectivity berkisar dari 0 hingga tak terhingga. Semakin kecil nilai *Index Connectivity*, semakin baik kualitas *cluster* yang terbentuk.

$$Conn = \sum_{i=1}^N \sum_{j=1}^L X_{i,nni(j)} \quad (2)$$

dengan:

Conn : indeks Connectivity

$nni(j)$  : pengamatan tetangga terdekat dari data ke-j ke data ke-i

N : banyak pengamatan

L : banyak *cluster*

#### 2. Silhouette Coefficient

*Silhouette Coefficient* menggunakan selisih antara jarak rata-rata dalam *cluster* dengan jarak minimum antar *cluster* untuk memperoleh jumlah *cluster* yang optimal. Semakin besar nilai *silhouette* menunjukkan efek *clustering* yang lebih baik. perhitungannya sebagai berikut:

$$S_i = b_i - a_i / \max[a_i, b_i] \quad (4)$$

$$SC = \frac{1}{n} \sum_{i=1}^n S_i \quad (5)$$

dengan:

$S_i$  : indeks *silhouette* pada objek ke-i, dengan  $i = 1, 2, \dots, n$

$a_i$  : rata-rata kemiripan antara objek ke-i dengan objek lain di dalam klasternya

$b_i$  : nilai minimum dari rata-rata kemiripan antara objek ke-i dengan objek lain diluar *cluster*-nya



SC : rata-rata rentang nilai indeks silhouette

### 3. *Dunn Index* (DI)

*Dunn Index* adalah ukuran untuk mengevaluasi *cluster* yang mengukur tingkat *compactness* dan tingkat pemisahan antar *cluster*. DI mengukur perbandingan jarak antar *cluster* dibagi dengan jarak intra *cluster* atau *compactness* [25]. Semakin besar nilai DI, semakin besar partisi yang dianggap lebih baik dan dapat mewakili *cluster* yang terpisah dengan baik.

$$DI = \frac{\min_{1 \leq i \neq j \leq m} \delta(C_i, C_j)}{\max_{1 \leq j \leq m} \Delta_j} \quad (6)$$

dengan:

DI : indeks dunn

$\delta(C_i, C_j)$  : ukuran kedekatan antara cluster i dan cluster j

$\Delta_j$  : ukuran kedekatan antara anggota dalam cluster j

### 4. *Davies-Bouldin Index*

*Index Davies-Bouldin* adalah salah satu metode evaluasi internal yang mengukur kualitas *cluster* dalam suatu metode pengelompokan berdasarkan nilai kohesi dan separasi.

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (7)$$

$$DBI = \frac{1}{K} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (8)$$

dengan:

$SSW_i$  : sum Of Square Within-Cluster pada centroid i

$SSB_{i,j}$  : sum of Square Between Cluster data ke i dengan j pada cluster yang berbeda

$R_{i,j}$  : rasio dari nilai SSW dan SSB

K : jumlah *cluster*

DBI : index Davies-Bouldin

## III. HASIL DAN PEMBAHASAN

### 3.1 *Preprocessing*

Data yang digunakan tidak memiliki missing value pada setiap variabelnya dari tahun 2021-2023. Tidak terdapat pula duplicate pada data tersebut. Sebelumnya, terdapat perbedaan satuan waktu yang digunakan pada masing-masing variabel. Variabel suhu dari setiap kabupaten kota merupakan data harian, sedangkan variabel lainnya merupakan data tahunan. Oleh karena itu, untuk menyamakan satuan waktunya, dilakukan rata-rata pada data harian untuk variabel suhu menjadi rata-rata tahunan suhu pada masing-masing wilayah. Proses tersebut dilakukan sebelum akhirnya proses pengolahan data dilakukan lebih lanjut. Kemudian, data dengan masing-masing variabel dipisahkan berdasarkan tahun. Pemisahan variabel berdasarkan tahunnya dilakukan untuk melihat keterbandingan dari setiap variabel di setiap wilayah pada setiap tahunnya.

**Tabel 2.** Ringkasan Statistik Data Tahun 2021

Variabel	Minimal	Maksimal	Mean	Median
Kepadatan Penduduk	491	11361	2102	1124
Rata-Rata Pengeluaran untuk Makanan	363805	740318	529322	523244
Produktivitas Padi	44.42	66.28	54.82	54.55
NDVI	0.07819	0.24733	0.17624	0.17934
suhu	27.14	28.35	27.73	27.57



Berdasarkan Tabel 2, diperoleh nilai minimum Kepadatan Penduduk tahun 2021 sebesar 491 dimiliki oleh Blora, sedangkan nilai maksimumnya dimiliki Kota Surakarta. Nilai minimum Rata-Rata Pengeluaran untuk Makanan Jawa Tengah dimiliki oleh Temanggung dan nilai maksimumnya dimiliki Kota Salatiga. Lalu, nilai minimum Produktivitas Padi dimiliki oleh Pekalongan dan nilai maksimumnya dimiliki Sragen. NDVI terendah dimiliki oleh Kota Semarang dan tertingginya dimiliki Blora. Banyumas memiliki suhu terendah dan Blora memiliki suhu tertinggi.

**Tabel 3.** Ringkasan Statistik Data Tahun 2022

Variabel	Minimal	Maksimal	Mean	Median
Kepadatan Penduduk	495.0	11878.0	2148.5	1188.0
Rata-Rata Pengeluaran untuk Makanan	437931	858244	585627	564773
Produktivitas Padi	44.52	64.03	55.40	55.60
NDVI	0.09896	0.25055	0.19547	0.20045
suhu	26.91	28.27	27.66	27.77

Nilai minimum dan maksimum setiap variabel pada tahun 2022 dimiliki oleh kabupaten/kota yang sama dengan tahun 2021, kecuali nilai maksimum pada variabel Produktivitas Padi dan NDVI, dan nilai terendah pada variabel suhu. Sukoharjo memiliki nilai Produktivitas Padi tertinggi dan NDVI tertinggi dimiliki oleh Purworejo. Suhu terendah dimiliki oleh Pekalongan.

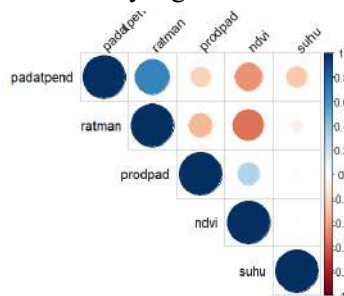
**Tabel 4.** Ringkasan Statistik Data Tahun 2023

Variabel	Minimal	Maksimal	Mean	Median
Kepadatan Penduduk	461	11277	2066	1172
Rata-Rata Pengeluaran untuk Makanan	480598	827635	612896	593017
Produktivitas Padi	44.24	65.43	55.34	54.95
NDVI	0.1004	0.2352	0.1893	0.1973
suhu	26.82	28.94	28.13	28.12

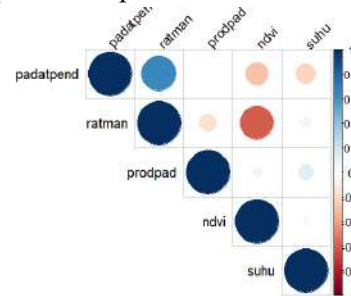
Nilai minimum dan maksimum sebagian besar variabel pada tahun 2023 dimiliki oleh kabupaten/kota yang sama dengan tahun sebelumnya. Perbedaannya ada pada nilai terendah Rata-Rata Pengeluaran untuk Makanan dimiliki oleh Purworejo dan tertingginya dimiliki oleh Kota Tegal. Produktivitas Padi tertinggi dimiliki oleh Sukoharjo. NDVI terendah dimiliki oleh Kota Pekalongan dan tertinggi dimiliki oleh Sragen.

### 3.2 Processing

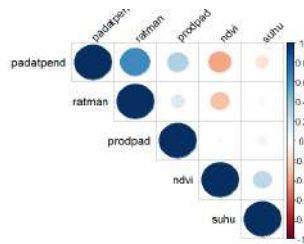
Selanjutnya, dilakukan uji multikolinearitas pada setiap data yang digunakan. Data yang digunakan telah memenuhi bahwa tidak terjadi pelanggaran asumsi multikolinearitas. Ini karena tidak ada nilai korelasi antar variabel yang bernilai lebih dari 0.95, serta dapat dilihat pada Gambar 3.



**Gambar 3a.**



**Gambar 3b.**



**Gambar 3c.**

**Gambar 3.** merupakan Matriks Korelasi : (3a) Data Tahun 2021 (3b) Data Tahun 2022 (3c) Data Tahun 2023

Dalam menentukan jumlah *cluster* optimal, metode yang digunakan adalah metode silhouette. Silhouette merupakan metode yang populer untuk digunakan dalam menentukan nilai  $k$  pada K-Means *Clustering* selain metode Elbow. Metode ini lebih disukai untuk memperkirakan potensi jumlah *cluster* optimal [26].

**Tabel 5.** Evaluasi *Cluster* Tahun 2021

<b>k</b>	<b>Silhouette</b>	<b>Connectivity</b>	<b>Dunn Index</b>	<b>DBI</b>
2	0.704	5.878	0.198	0.472
3	0.551	11.047	0.032	0.658
4	0.569	17.827	0.061	0.573
5	0.564	19.467	0.062	0.446

Berdasarkan Tabel 5, diperoleh bahwa jumlah *cluster* optimal untuk data tahun 2021 adalah 2. Hal ini ditunjukkan oleh  $k=2$  yang memiliki Silhouette coefficient terbesar, Index Connectivity terkecil, dan Dunn Index terbesar dibandingkan  $k$  lainnya, yaitu masing-masing sebesar 0.704, 5.878, dan 0.198. Namun, DBI terkecil dimiliki oleh  $k=5$  yang berarti bahwa menurut DBI jumlah *cluster* optimal adalah 5 dengan DBI yaitu sebesar 0.446. Karena  $k=2$  lebih banyak memenuhi syarat untuk evaluasi *cluster* optimal, jumlah *cluster* terbaik untuk data tahun 2021 sebesar 2.

**Tabel 6.** Evaluasi *Cluster* Tahun 2022

<b>k</b>	<b>Silhouette</b>	<b>Connectivity</b>	<b>Dunn Index</b>	<b>DBI</b>
2	0.588	7.725	0.037	0.742
3	0.367	26.238	0.007	1.094
4	0.244	40.824	0.009	1.176
5	0.241	43.070	0.010	1.434

Berdasarkan Tabel 6, diperoleh bahwa jumlah *cluster* optimal untuk data tahun 2022 adalah 2. Hal ini ditunjukkan oleh  $k=2$  yang memiliki Silhouette coefficient terbesar, Index Connectivity terkecil, Dunn Index terbesar, dan DBI terkecil dibandingkan  $k$  lainnya.

**Tabel 7.** Evaluasi *Cluster* Tahun 2023

<b>k</b>	<b>Silhouette</b>	<b>Connectivity</b>	<b>Dunn Index</b>	<b>DBI</b>
2	0.482	17.596	0.0009	0.890
3	0.299	33.216	0.0008	1.009
4	0.162	44.149	0.0009	1.029
5	0.015	63.828	0.0010	1.351

Berdasarkan Tabel 7, diperoleh bahwa jumlah *cluster* optimal untuk data tahun 2023 adalah 2. Hal ini ditunjukkan oleh  $k=2$  yang memiliki Silhouette coefficient terbesar, Index Connectivity terkecil, dan DBI terkecil dibandingkan  $k$  lainnya, yaitu masing-masing sebesar 0.482, 17.596, dan 0.890. Namun, Dunn Index terbesar dimiliki oleh  $k=5$  yang berarti bahwa menurut Dunn Index jumlah *cluster* optimal adalah 5 dengan Dunn Index sebesar 1.351. Karena  $k=2$  lebih banyak memenuhi syarat untuk evaluasi *cluster* optimal, jumlah *cluster* terbaik untuk data tahun 2023 sebesar 2.



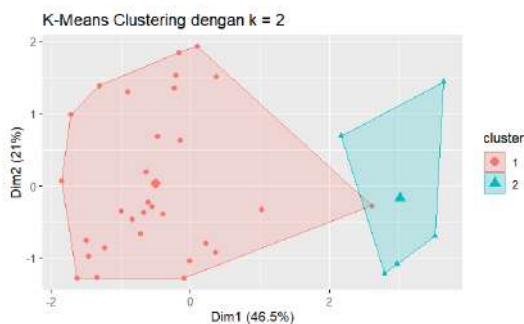
Tabel 8. Centroid

Variabel	Cluster Tahun 2021		Cluster Tahun 2022		Cluster Tahun 2023	
	1	2	1	2	1	2
Kepadatan Penduduk	1354.4	6586.0	1177.5	4953.2	1103.1	4166.2
Rata-Rata Pengeluaran untuk Makanan	501258.0	697702.8	536840.4	726566.1	569328.7	707952.3
Produktivitas Padi	55.38	51.45	55.54	55.00	55.36	55.28
NDVI	0.18	0.13	0.21	0.16	0.19	0.17
suhu	27.75	27.55	27.62	27.74	28.12	28.14
Label	Hijau	Kurang Hijau	Hijau	Kurang Hijau	Hijau	Kurang Hijau

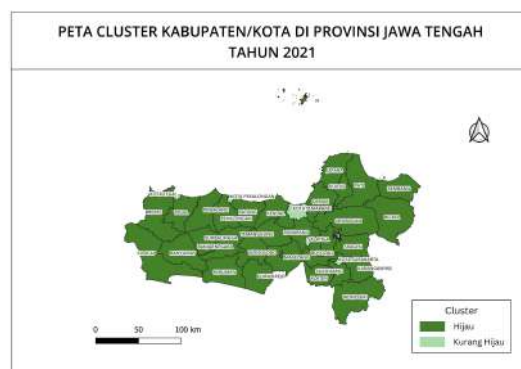
Setelah didapatkan jumlah *cluster* terbaik untuk setiap data yang digunakan, centroid untuk setiap *cluster* diperoleh pada Tabel 8. Kemudian, 68.3% dari total variasi dalam data tahun 2021 dijelaskan oleh perbedaan antar *cluster*. Untuk tahun 2022, 71.1% dari total variasi dalam data dijelaskan oleh perbedaan antar *cluster*. Total variasi dalam data tahun 2023 dijelaskan oleh perbedaan antar *cluster* sebesar 65.9%. Setiap tahun memiliki 2 *cluster*, yaitu hijau dan kurang hijau. *Cluster* hijau memiliki karakteristik di mana kepadatan penduduknya cenderung rendah, rata-rata pengeluaran perkapita per bulan untuk makanannya rendah, produktivitas padinya cukup tinggi, vegetasi cukup tinggi, dan suhu yang bervariasi. Sedangkan, *cluster* kurang hijau memiliki karakteristik sebaliknya, yaitu kepadatan penduduknya cenderung tinggi, rata-rata pengeluaran perkapita per bulan untuk makanannya tinggi, produktivitas padinya cukup rendah, vegetasinya cukup rendah, dan suhunya juga bervariasi.

Karakteristik yang terdapat pada masing-masing *cluster* tentu bersesuaian satu dengan yang lainnya. Wilayah yang kepadatan penduduknya cenderung rendah, tentu akan cenderung memiliki lahan yang lebih untuk bisa ditanami berbagai jenis tanaman untuk kebutuhan vegetasi di daerah tersebut. Sebagian besar kabupaten/kota yang terdapat di Provinsi Jawa Tengah termasuk ke dalam *cluster* hijau yang angka produktivitas padinya tergolong tinggi. Hal ini sesuai karena memang Provinsi Jawa Tengah merupakan salah satu provinsi penghasil padi terbesar di Indonesia.

Pada Gambar 4, tahun 2021 memiliki *cluster* hijau berisi 30 kabupaten/kota dan kurang hijau berisi 5 kabupaten/kota di Jawa Tengah. *Cluster* hijau terdiri dari Cilacap, Banyumas, Purbalingga, Banjarnegara, Kebumen, Purworejo, Wonosobo, Magelang, Boyolali, Klaten, Sukoharjo, Wonogiri, Karanganyar, Sragen, Grobogan, Blora, Rembang, Pati, Kudus, Jepara, Demak, Semarang, Temanggung, Kendal, Batang, Pekalongan, Pemalang, Tegal, Brebes, dan Kota Magelang. Lalu, kurang hijau berisi Kota Surakarta, Kota Salatiga, Kota Semarang, Kota Pekalongan, dan Kota Tegal.



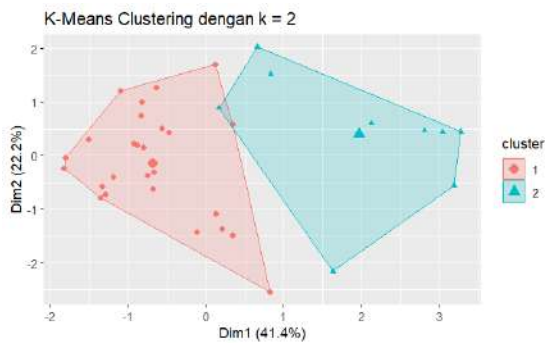
Gambar 4a.



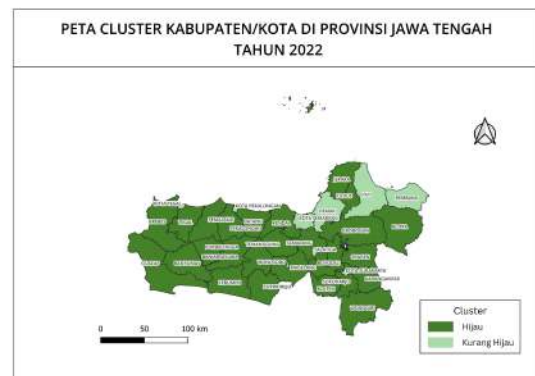
Gambar 4b.

Gambar 4. merupakan *Cluster* Tahun 2021: (4a) Visualisasi (4b) Peta

Pada tahun 2022 yang terlihat pada Gambar 5, *cluster* hijau berisi 26 kabupaten/kota dan *cluster* kurang hijau berisi 9 kabupaten/kota di Jawa Tengah. *Cluster* hijau terdiri dari Cilacap, Banyumas, Purbalingga, Banjarnegara, Kebumen, Purworejo, Wonosobo, Magelang, Boyolali, Klaten, Sukoharjo, Wonogiri, Karanganyar, Sragen, Grobogan, Blora, Kudus, Jepara, Semarang, Temanggung, Kendal, Batang, Pekalongan, Pemalang, Tegal, dan Brebes. Sedangkan, *cluster* kurang hijau terdiri dari Rembang, Pati, Demak, Kota Magelang, Kota Surakarta, Kota Salatiga, Kota Semarang, Kota Pekalongan, dan Kota Tegal. Terdapat perbedaan jumlah setiap *cluster* pada tahun 2022 dibandingkan tahun 2021. Ada 4 daerah yang mengalami perubahan *cluster*, di mana pada tahun 2021 daerah tersebut masuk dalam *cluster* hijau sedangkan pada tahun 2022 masuk dalam *cluster* kurang hijau. Keempat daerah tersebut adalah Rembang, Pati, Demak, dan Kota Magelang.



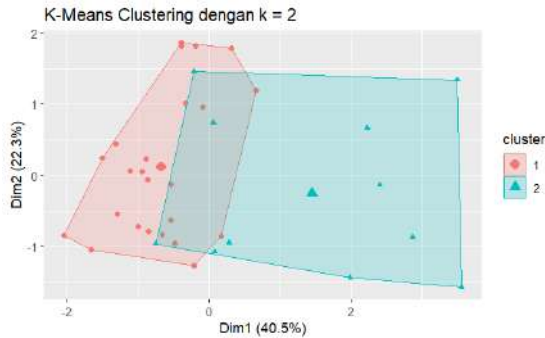
Gambar 5a.



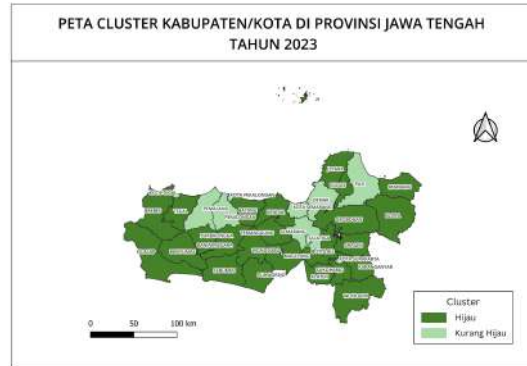
Gambar 5b.

Gambar 5. merupakan *Cluster* Tahun 2022: (5a) Visualisasi (5b) Peta

Pada tahun 2023 yang terlihat pada Gambar 6, *cluster* hijau berisi 24 kabupaten/kota dan *cluster* kurang hijau berisi 11 kabupaten/kota. *Cluster* hijau terdiri dari Cilacap, Banyumas, Purbalingga, Banjarnegara, Kebumen, Purworejo, Wonosobo, Magelang, Boyolali, Klaten, Sukoharjo, Wonogiri, Karanganyar, Sragen, Grobogan, Blora, Rembang, Kudus, Jepara, Temanggung, Kendal, Batang, Tegal, dan Brebes. Sedangkan, *cluster* kurang hijau terdiri dari Pati, Demak, Semarang, Pekalongan, Pemalang, Kota Magelang, Kota Surakarta, Kota Salatiga, Kota Semarang, Kota Pekalongan, dan Kota Tegal. Terdapat perbedaan jumlah setiap *cluster* pada tahun 2023 dibandingkan tahun 2022. Ada 4 daerah yang mengalami perubahan *cluster*. Tiga daerah mengalami perubahan *cluster* di mana pada tahun 2022 daerah tersebut masuk dalam *cluster* hijau sedangkan pada tahun 2023 masuk dalam *cluster* kurang hijau, yaitu Semarang, Pekalongan, dan Pemalang. Terdapat juga satu daerah yang mengalami perubahan *cluster* di mana pada tahun 2022 daerah tersebut masuk dalam *cluster* kurang hijau sedangkan pada tahun 2023 masuk dalam *cluster* hijau, yaitu Rembang.



Gambar 6a.

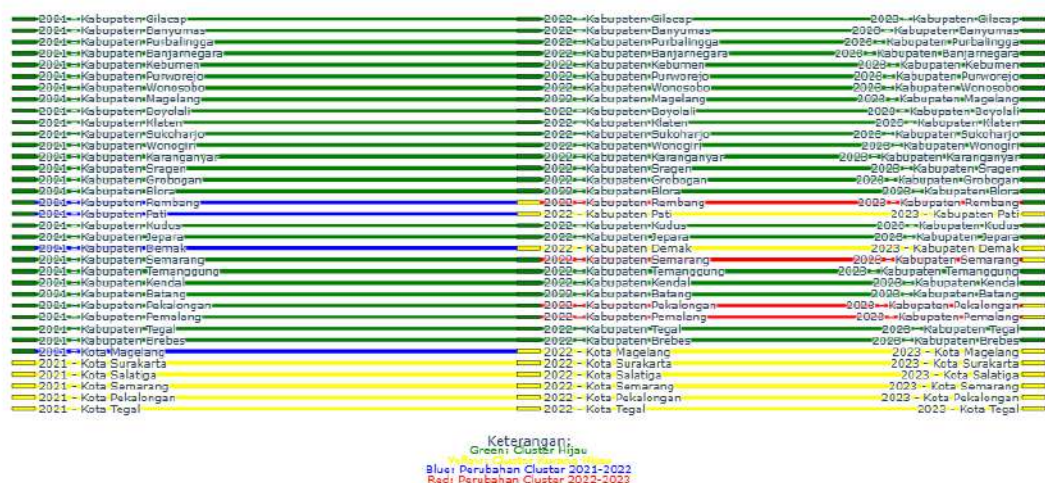


Gambar 6b.

Gambar 6 . merupakan Cluster Tahun 2023: (6a) Visualisasi (6b) Peta

Berdasarkan Gambar 7, diperoleh bahwa terdapat 7 kabupaten/kota di Jawa Tengah yang mengalami perubahan *cluster* dari tahun 2021-2023, yaitu Rembang, Pati, Demak, Semarang, Pekalongan, Pemalang, dan Kota Magelang. Rembang mengalami perubahan *cluster* dari tahun ke tahun di mana pada tahun 2021 masuk dalam *cluster* hijau, tahun 2022 masuk dalam *cluster* kurang hijau, dan tahun 2023 masuk dalam *cluster* hijau. Pati, Demak, dan Kota Magelang mengalami perubahan *cluster* yang sama dari tahun 2021-2023 di mana pada tahun 2021 kabupaten/kota tersebut masuk dalam *cluster* hijau dan pada tahun 2022-2023 masuk dalam *cluster* kurang hijau. Semarang, Pekalongan, dan Pemalang mengalami perubahan *cluster* yang sama dari tahun 2021-2023 di mana pada tahun 2021-2022 kabupaten/kota tersebut masuk dalam *cluster* hijau dan pada tahun 2023 masuk dalam *cluster* kurang hijau.

Sankey Diagram Perubahan Cluster Kabupaten/Kota di Provinsi Jawa Tengah Tahun 2021-2023



Gambar 7. Diagram Sankey Perubahan Cluster Kabupaten/Kota di Provinsi Jawa Tengah Tahun 2021-2023

Walaupun terdapat perubahan *cluster*, *cluster* hijau tetap mendominasi daerah di Jawa Tengah. Tentu ini merupakan sebuah potensi bagi provinsi Jawa Tengah untuk bisa berkontribusi menjadi lumbung pangan nasional. Sesuai dengan penelitian yang dilakukan oleh [27], bahwa penataan hutan dapat memengaruhi keragaman pola makan rumah tangga, cara masyarakat mengakses sumber daya dan mengelola produksinya. Meskipun suhu di setiap *cluster* bervariasi, para petani produktif di Kebumen setuju bahwa perubahan klimatologi, terutama suhu dan curah hujan, mempengaruhi kegiatan



pertanian [28]. Maka dari itu, Jawa Tengah menjadi provinsi yang memiliki angka kontribusi dalam pertanian terbesar terhadap ekonomi Indonesia [29].

Kemudian, *cluster* kurang hijau didominasi oleh daerah perkotaan. Perluasan perkotaan yang tidak terkendali akibat urbanisasi dapat memberikan dampak buruk terhadap lahan pertanian [30]. Hal ini dilihat dari semakin banyaknya tutupan lahan di daerah tersebut untuk pembangunan rumah dan bangunan pendukung lainnya. Namun, kurangnya lahan hijau yang terdapat di daerah perkotaan juga bisa terjadi karena perlu ada pertimbangan serius untuk menentukan jenis vegetasi yang ditambahkan ke daerah perkotaan disebabkan oleh perbedaan jenis vegetasi yang memiliki kemampuan berbeda dalam penyediaan fungsi ekologisnya [31]. Oleh karena itu, penting bagi para pemangku kebijakan untuk membuat kebijakan yang tepat agar kondisi vegetasi yang ada di Provinsi Jawa Tengah bisa terus membaik, sejalan dengan faktor sosial ekonomi lainnya yang juga terus berkembang, sehingga bisa memberi dampak positif, khususnya bagi masyarakat di sekitarnya.

#### IV. KESIMPULAN

Provinsi Jawa Tengah terbagi menjadi dua *cluster* utama, yaitu *cluster* hijau dan kurang hijau, yang terbentuk berdasarkan variabel NDVI, Kepadatan Penduduk, Rata-rata Pengeluaran untuk Makanan, Produktivitas Padi, dan Suhu. Sebagian besar kabupaten/kotanya masuk ke *cluster* hijau yang menunjukkan potensi provinsi ini menjadi lumbung pangan nasional. Dari tahun 2021 hingga 2023, karakteristik dari *cluster* hijau yaitu wilayah yang memiliki kepadatan penduduk dan rata-rata pengeluaran untuk makanannya rendah, serta vegetasi dan produktivitas padinya cukup tinggi. Suhu untuk setiap *cluster* cukup bervariasi. Di sisi lain, *cluster* kurang hijau didominasi oleh daerah perkotaan yang mengalami urbanisasi cepat serta berdampak pada pengurangan lahan pertanian. Untuk mendukung Jawa Tengah sebagai lumbung pangan nasional, dapat dilakukan beberapa langkah praktis. Pengelolaan area dan identifikasi dapat dilakukan di *cluster* hijau dengan potensi pertanian tinggi untuk meningkatkan hasil panen. Penggunaan teknologi pertanian modern dapat meningkatkan efisiensi dan produktivitas lahan. Diversifikasi tanaman untuk mengurangi risiko ketergantungan pada satu jenis tanaman dan memaksimalkan penggunaan lahan dapat dilakukan pada wilayah ini. Bahkan, pelatihan bagi petani mengenai praktik pertanian berkelanjutan dan manajemen sumber daya alam yang efektif guna membantu meningkatkan produktivitas pertanian di Jawa Tengah.

#### REFERENSI

1. X. Chen, C. Liu, and X. Yu, “Urbanization, Economic Development, and Ecological Environment: Evidence from Provincial Panel Data in China,” *Sustainability*, vol. 14, no. 3, p. 1124, Jan. 2022, doi: 10.3390/su14031124.
2. A. Hardianto, P. U. Dewi, T. Feriansyah, N. F. S. Sari, and N. S. Rifiana, “Pemanfaatan Citra Landsat 8 Dalam Mengidentifikasi Nilai Indeks Kerapatan Vegetasi (NDVI) Tahun 2013 dan 2019 (Area Studi: Kota Bandar Lampung),” *J. Geosains Remote Sens.*, vol. 2, no. 1, pp. 8–15, May 2021, doi: 10.23960/jgrs.2021.v2i1.38.
3. S. Huang, L. Tang, J. P. Hupy, Y. Wang, and G. Shao, “A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing,” *J. For. Res.*, vol. 32, no. 1, pp. 1–6, Feb. 2021, doi: 10.1007/s11676-020-01155-1.
4. Badan Pusat Statistik, “Hasil Sensus Penduduk 2020,” Jan. 2021.
5. A. Bahij and T. Rosy, “Mengungkit Ekonomi Jawa Tengah Melalui Agroindustri: Simulasi Injeksi Investasi Menggunakan Tabel Input Output,” *BESTARI BPS Kalimantan Timur*, vol. 2, no. 02, 2022.
6. D. R. Safitri and E. D. Sihaloho, “Lumbung padi Indonesia dan kemiskinan: studi kasus kabupaten kota di Jawa Timur,” *Ekonomis: Journal of Economics and Business*, vol. 4, no. 1, pp. 56–61, 2020.
7. Kementerian Pertanian, “Analisis Komoditas Pangan Strategis Tahun 2023,” Pusat Data dan Sistem Informasi Pertanian Kementerian Pertanian, 2022.



8. T. A. Akbar, Q. K. Hassan, S. Ishaq, M. Batool, H. J. Butt, and H. Jabbar, “Investigative Spatial Distribution and Modelling of Existing and Future Urban Land Changes and Its Impact on Urbanization and Economy,” *Remote Sensing*, vol. 11, no. 2, p. 105, Jan. 2019, doi: 10.3390/rs11020105.
9. M. Sharma, P. Bangotra, A. S. Gautam, and S. Gautam, “Sensitivity of normalized difference vegetation index (NDVI) to land surface temperature, soil moisture and precipitation over district Gautam Buddha Nagar, UP, India,” *Stoch Environ Res Risk Assess*, vol. 36, no. 6, pp. 1779–1789, Jun. 2022, doi: 10.1007/s00477-021-02066-1.
10. T. Mahesti, E. Umar, A. Ariadi, S. Y. J. Prasetyo, and C. Fibriani, “Identifikasi Perubahan Tutupan Vegetasi dan Curah Hujan Kabupaten Semarang Menggunakan Citra Saltelit Lansat 8,” vol. 3, 2020.
11. M. M. Maja and S. F. Ayano, “The Impact of Population Growth on Natural Resources and Farmers’ Capacity to Adapt to Climate Change in Low-Income Countries,” *Earth Syst Environ*, vol. 5, no. 2, pp. 271–283, Jun. 2021, doi: 10.1007/s41748-021-00209-6.
12. L. Serra-Majem *et al.*, “Updating the Mediterranean Diet Pyramid towards Sustainability: Focus on Environmental Concerns,” *IJERPH*, vol. 17, no. 23, p. 8758, Nov. 2020, doi: 10.3390/ijerph17238758.
13. L. Jiang, Y. Liu, S. Wu, and C. Yang, “Analyzing ecological environment change and associated driving factors in China based on NDVI time series data,” *Ecological Indicators*, vol. 129, p. 107933, Oct. 2021, doi: 10.1016/j.ecolind.2021.107933.
14. N. A. Nafarin and N. Novitasari, “Relationship Between Normalized Difference Vegetation Index (NDVI) and Rice Growth Phases in Danda Jaya Swamp Irrigation Area Regency Barito Kuala,” *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 1184, no. 1, p. 012019, May 2023, doi: 10.1088/1755-1315/1184/1/012019.
15. O. D. Kimaro, S. L. Gebre, P. Hieronimo, N. Kihupi, K.-H. Feger, and D. N. Kimaro, “Handheld NDVI sensor-based rice productivity assessment under combinations of fertilizer soil amendment and irrigation water management in lower Moshi irrigation scheme, North Tanzania,” *Environ Earth Sci*, vol. 82, no. 3, p. 78, Feb. 2023, doi: 10.1007/s12665-022-10730-0.
16. L. A. Pampuch, R. G. Negri, P. C. Loikith, and C. A. Bortolozzo, “A Review on Clustering Methods for Climatology Analysis and Its Application over South America,” *IJG*, vol. 14, no. 09, pp. 877–894, 2023, doi: 10.4236/ijg.2023.149047.
17. Badan Pusat Statistik, “Kepadatan Penduduk menurut Kabupaten/Kota.” [Online]. Available: <https://www.bps.go.id>. (accessed on 1 Mei 2024)
18. Badan Pusat Statistik, “Luas Panen, Produksi, dan Produktivitas Padi menurut Kabupaten/Kota.” [Online]. Available: <https://www.bps.go.id>. (accessed on 1 Mei 2024)
19. Badan Pusat Statistik, “Rata-Rata Pengeluaran per Kapita per Bulan Makanan dan Bukan Makanan (Rupiah) menurut Kabupaten/Kota.” [Online]. Available: <https://www.bps.go.id>. (accessed on 1 Mei 2024)
20. Meteostat, “Curah Hujan.” [Online]. Available: <https://meteostat.net>. (accessed on 1 Mei 2024)
21. Google Earth Engine, “NDVI.” [Online]. Available: <https://earthengine.google.com>. (accessed on 1 Mei 2024)
22. Z. Davis, L. Nesbitt, M. Guhn, and M. Van Den Bosch, “Assessing changes in urban vegetation using Normalised Difference Vegetation Index (NDVI) for epidemiological studies,” *Urban Forestry & Urban Greening*, vol. 88, p. 128080, Oct. 2023, doi: 10.1016/j.ufug.2023.128080.
23. F. P. Hidayat, R. P. Putra, M. D. Alfitriah, and E. Widodo, “Implementasi Clustering K-Medoids dalam Pengelompokan Kabupaten di Provinsi Aceh Berdasarkan Faktor yang Mempengaruhi Kemiskinan,” *Indonesian Journal of Applied Statistics*, vol. 5, no. 2, pp. 121–130, 2023.
24. S. Zelasky, C. L. Martin, C. Weaver, L. K. Baxter, and K. M. Rappazzo, “Identifying groups of children’s social mobility opportunity for public health applications using k-means clustering,” *Heliyon*, vol. 9, no. 9, p. e20250, Sep. 2023, doi: 10.1016/j.heliyon.2023.e20250.
25. C.-E. Ben Ncir, A. Hamza, and W. Bouaguel, “Parallel and scalable Dunn Index for the validation of big data clusters,” *Parallel Computing*, vol. 102, p. 102751, May 2021, doi: 10.1016/j.parco.2021.102751.
26. C. Shi, B. Wei, S. Wei, W. Wang, H. Liu, and J. Liu, “A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm,” *J Wireless Com Network*, vol. 2021, no. 1, p. 31, Dec. 2021, doi: 10.1186/s13638-021-01910-w.
27. L. V. Rasmussen *et al.*, “Forest pattern, not just amount, influences dietary quality in five African countries,” *Global Food Security*, vol. 25, p. 100331, Jun. 2020, doi: 10.1016/j.gfs.2019.100331.
28. A. B. Sekaranom, E. Nurjani, and F. Nucifera, “Agricultural Climate Change Adaptation in Kebumen, Central Java, Indonesia,” *Sustainability*, vol. 13, no. 13, p. 7069, Jun. 2021, doi: 10.3390/su13137069.
29. M. S. Mu’min and M. Yaqin, “Pengembangan Strategi Ketahanan Pangan Melalui Peran Teknologi, Informasi, dan Komunikasi: Studi Kasus Sektor Pertanian di Jawa Tengah”.



Seminar Nasional Sains Data 2024 (SENADA 2024)  
UPN "Veteran" Jawa Timur

E-ISSN 2808-5841  
P-ISSN 2808-7283

30. L. Gandharum, D. M. Hartono, A. Karsidi, and M. Ahmad, "Monitoring urban expansion and loss of agriculture on the north coast of west java province, Indonesia, using Google Earth engine and intensity analysis," *The Scientific World Journal*, vol. 2022, no. 1, p. 3123788, 2022.
31. D. R. Richards and R. N. Belcher, "Global Changes in Urban Vegetation Cover," *Remote Sensing*, vol. 12, no. 1, p. 23, Dec. 2019, doi: 10.3390/rs12010023.