

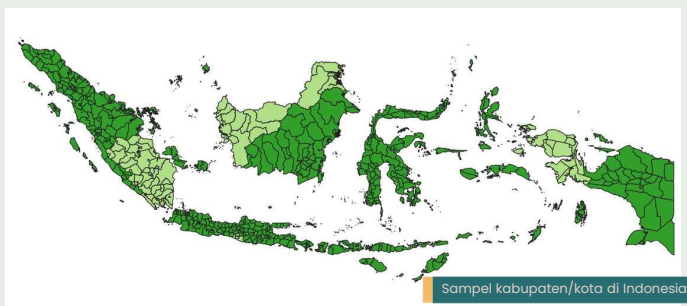


# PENGELOMPOKAN DAERAH POTENSIAL PENGHASIL PADI DI INDONESIA MENGGUNAKAN METODE K-MEANS CLUSTERING

## 01. PENDAHULUAN

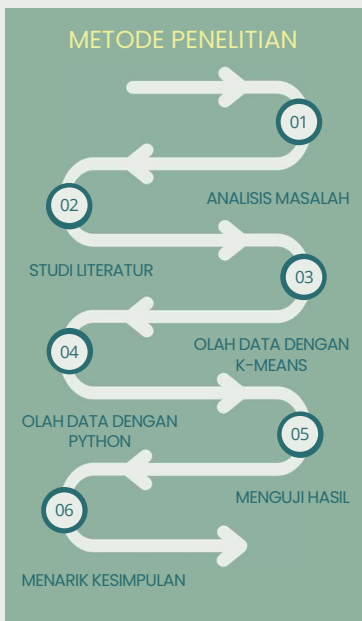
Padi merupakan salah satu sumber makanan pokok yang dikonsumsi masyarakat Indonesia. Namun, kebutuhan beras masyarakat yang semakin meningkat tidak diimbangi dengan pertumbuhan area pertanian. Luas panen kabupaten/kota di Indonesia pada tahun 2021 mengalami penurunan dibandingkan tahun sebelumnya. Selain itu, produksi padi juga mengalami penurunan. Oleh karena itu, Indonesia perlu mengelompokkan kabupaten/kota penghasil padi untuk mengetahui daerah yang memiliki potensi dalam memproduksi padi.

Metode yang dapat digunakan dalam mengelompokkan kabupaten/kota penghasil padi adalah *clustering*. *Clustering* dapat mengelompokkan data tersebut berdasarkan kesamaan jenis data ke dalam kluster. Salah satu metode *clustering* yang populer adalah *K-Means*. Pada penelitian ini, algoritma *K-Means* digunakan untuk mengelompokkan daerah penghasil padi di Indonesia sehingga diharapkan ditemukan pola-pola atau informasi tersembunyi dari data hasil produksi padi di Indonesia.



Sampel kabupaten/kota di Indonesia

## 02. DATA DAN METODE



Data yang digunakan dalam penelitian ini adalah data sekunder yang diambil dari Badan Pusat Statistik (BPS). Variabel yang digunakan adalah luas panen padi (ha), produksi padi (ton), dan produktivitas padi (ton/ha) menurut kabupaten/kota di Indonesia pada tahun 2021.

Luas panen padi diperoleh dari hasil Survei Kerangka Sampel Area (KSA). Produksi padi adalah hasil tanaman padi yang diambil berdasarkan luas yang dipanen. Produktivitas padi diperoleh dari hasil Survei Ubinan dengan melakukan pengukuran hasil panen pada plot berukuran 2,5 m x 2,5 m.

Pada tahun 2021, Indonesia terdiri atas 34 provinsi dan 514 kabupaten/kota. Setelah dilakukan penarikan sampel menggunakan cluster random sampling diperoleh sampel sebanyak 26 provinsi dan 394 kabupaten/kota.

## 04. KESIMPULAN

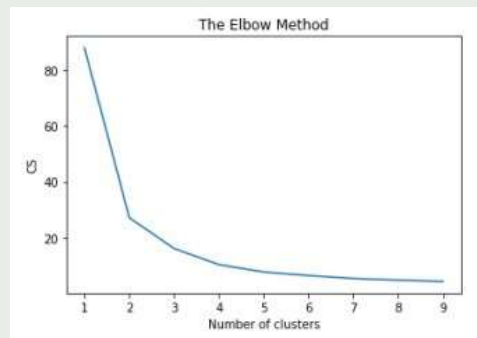
Berdasarkan hasil analisis data, perhitungan algoritma *K-Means* menggunakan pemrograman *python* berhenti pada iterasi ke-300. Jumlah kluster terbaik berdasarkan *the elbow method* dan *silhouette method* adalah 2 kelompok dengan nilai sebesar 0,63. Hasil kluster yang diperoleh adalah sebagai berikut:

- C0 menggambarkan kabupaten/kota yang memiliki produktivitas padi rendah. Kabupaten/kota yang termasuk dalam C0 berjumlah 295 daerah.
- C1 menggambarkan kabupaten/kota yang memiliki produktivitas padi tinggi. Kabupaten/kota yang termasuk dalam C1 berjumlah 99 daerah.

Selain itu, dapat juga disimpulkan bahwa penerapan metode *K-Means* untuk *clustering* daerah potensial penghasil padi di Indonesia berdasarkan luas panen, produksi, dan produktivitas padi berhasil diterapkan sehingga dapat dilakukan pengelolaan potensi pertanian dengan optimal sesuai sasaran.

## 03. HASIL DAN PEMBAHASAN

### JUMLAH KLASER TERBAIK



### CENTROID

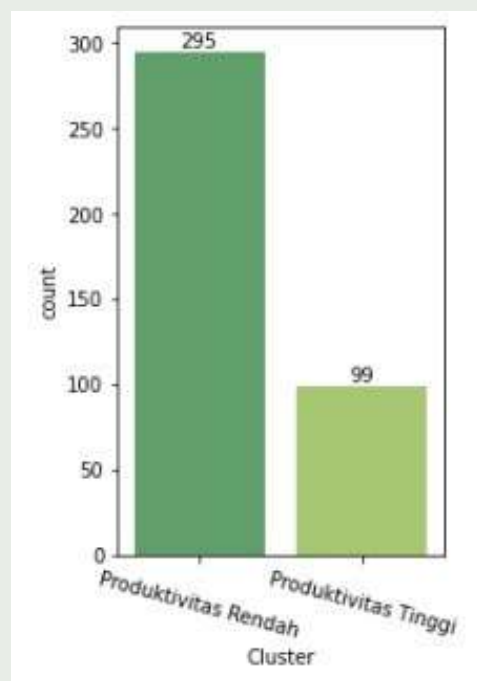
[0.1014949 , 0.09453351, 0.4801793 ]  
[0.7147018 , 0.7440232 , 0.6395027 ]

### JARAK EUCLIDEAN

$$D_{(x,y)} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

Dimana  $x$  adalah data produktivitas tanaman padi yaitu luas panen, produksi dan produktivitas, sedangkan  $y$  adalah centroid yang telah diperoleh sebelumnya.

### HASIL KLASER

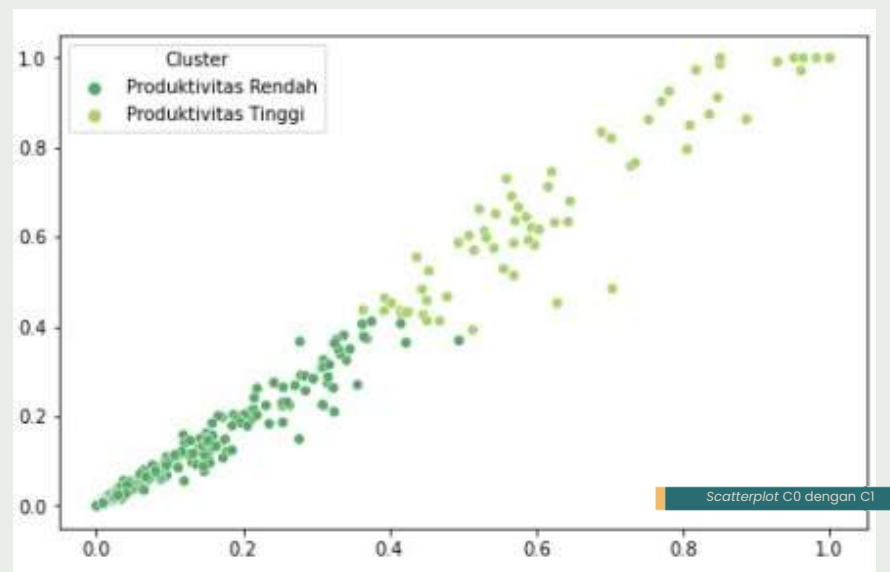


### HASIL KLASER

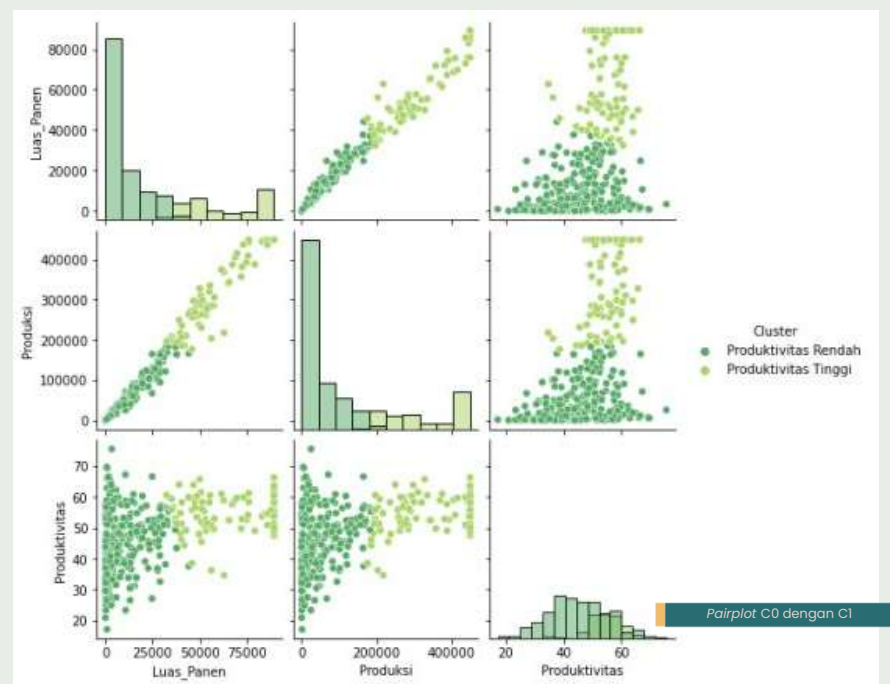
Kabupaten/kota yang memiliki produktivitas padi rendah berjumlah 295 daerah dan yang memiliki produktivitas padi tinggi berjumlah 99 daerah.

For n\_cluster=2, the silhouette score is 0.6274297755071471  
For n\_cluster=3, the silhouette score is 0.55147822859616  
For n\_cluster=4, the silhouette score is 0.4686752066859879  
For n\_cluster=5, the silhouette score is 0.48112419927312156  
For n\_cluster=6, the silhouette score is 0.46929067114096096  
For n\_cluster=7, the silhouette score is 0.42386004075171607  
For n\_cluster=8, the silhouette score is 0.42705056340092035

Berdasarkan hasil *the elbow method* dan *silhouette method*, diperoleh jumlah kluster terbaik adalah 2 dengan 300 kali iterasi. Maka, pengelompokan akan dibagi menjadi daerah dengan produktivitas padi rendah (C0) dan daerah produktivitas padi tinggi (C1).



Scatterplot C0 dengan C1

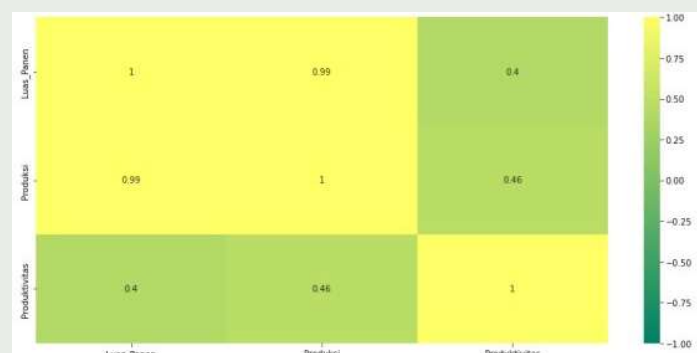


Pairplot C0 dengan C1

### ANALISIS DESKRIPTIF

Berdasarkan hasil pengelompokan menggunakan pemrograman *python* diperoleh ringkasan statistik sebagai berikut:

- Luas panen untuk daerah dengan produktivitas padi rendah rata-rata berkisar 9.062,82 ha. Luas panen terkecil adalah 0,11 ha dan terluas adalah 44.146,35 ha. Adapun untuk daerah produktivitas padi tinggi rata-rata berkisar 63.817,44 ha. Luas panen terkecil adalah 32.480,03 ha dan terluas adalah 89.292,36.
- Produksi padi untuk daerah dengan produktivitas rendah rata-rata berkisar 42.518,61 ton. Produksi padi paling sedikit adalah 0,23 ton dan paling banyak adalah 185218 ton. Adapun untuk daerah produktivitas padi tinggi rata-rata berkisar 334.639,87 ton. Produksi padi paling sedikit adalah 176.533,79 ton dan paling banyak adalah 449.770,67 ton.
- Produktivitas padi untuk daerah dengan produktivitas rendah rata-rata 45,19 ton/ha. Produktivitas padi terkecil adalah 17,16 ton/ha dan terbesar 75,53 ton/ha. Adapun untuk daerah produktivitas padi tinggi rata-rata 54,49 ton/ha. Produktivitas padi terkecil adalah 34,68 ton/ha dan terbesar adalah 66,8 ton/ha.



### KORELASI

- Luas panen memiliki korelasi yang sangat kuat dengan produksi padi. Korelasi yang terjadi adalah korelasi yang positif, artinya jika luas panen mengalami kenaikan, maka produksi padi juga akan mengalami kenaikan.
- Produktivitas padi tidak memiliki korelasi yang kuat dengan luas panen dan produksi padi.

### DAFTAR PUSTAKA

BPS. (2 April 2023). Luas Panen, Produksi, dan Produktivitas Padi Menurut Kabupaten/Kota di Indonesia Tahun 2021. Diakses dari <https://www.bps.go.id/>

Wijayanto, S., & Fathoni, M. Y. (2021). Pengelompokan Produktivitas Tanaman Padi di Jawa Tengah Menggunakan Metode Clustering K-Means. *Jurnal Penelitian Ilmu Dan Teknik Komputer*, 13, 212–219.

Wahyudi, M. (2020). *Data Mining: Penerapan Algoritma K-Means Clustering dan K-Medoids Clustering*. Penerbit Kita Menulis, Cengkareng.

### KELOMPOK 10 TEKNIK SAMPLING

Nurul Ummah (2108108010014)  
Husnul Khatimah (2108108010016)  
Riska Muliani Putri (2108108010067)

