**PENGELOMPOKKAN POTENSI PADI DI INDONESIA**

**MENGGUNAKAN K-MEANS *CLUSTER***

**Eko Suprapto**

**Badan Pusat Statistik Provinsi Jambi**

[esuprapto@bps.go.id](mailto:esuprapto@bps.go.id)

***ABSTRACT***

*The Improvements to rice data in Indonesia have been carried out by BPS in collaboration with BPPT using the Area Sample Framework (KSA) method. The KSA survey for rice commodities began to be implemented nationally in 2018 and is included in a national project to support the achievement of one of the national priorities, namely food security, in order to improve food statistics data. Once the agricultural sector is so crucial, especially rice, regional-based planning is absolutely necessary. The purpose of this study is to classify rice potential by province in Indonesia based on the variables of harvested area, productivity, and rice production in 2021 using* K-means *Cluster analysis. The results of the study resulted in three provincial clusters based on rice potential. The first cluster is a provincial cluster with the highest rice potential or the highest contributor to rice production in Indonesia. The second cluster is a group of provinces that have production values ​​above the average, which is above 1.6 tons of GKG. While the third cluster is a group of provinces with below average rice potential.*

**Keywords:** *KSA, rice potential,*k-means *cluster*

**ABSTRAK**

Perbaikan data padi di Indonesia telah dilakukan BPS bekerjasama dengan BPPT mengunakan metode Kerangka Sampel Area (KSA). Survei KSA untuk komoditas padi mulai diimplementasikan secara nasional pada tahun 2018 dan termasuk dalam proyek nasional untuk mendukung pencapaian salah satu prioritas nasional, yaitu ketahanan pangan, dalam rangka perbaikan data statistik pangan. Melihat begitu krusialnya sektor pertanian, khususnya padi, maka perencanaan berbasis wilayah mutlak diperlukan. Tujuan penelitian ini adalah mengelompokkan potensi padi berdasarkan provinsi di Indonesia berdasarkan variabel luas panen, produktivitas, dan produksi padi tahun 2021 menggunakan analisis K-means *Cluster*. Hasil penelitian menghasilkan tiga *cluster* provinsi berdasarkan potensi padi. *Cluster* pertama merupakan *cluster* provinsi dengan potensi padi tertinggi atau penyumbang produksi padi tertinggi di Indonesia. *Cluster* kedua kelompok provinsi yang mempunyai nilai produksi di atas rata-rata, yaitu diatas 1,6 ton GKG. Sedangkan *cluster* ketiga merupakan kelompok provinsi dengan potensi padi di bawah rata-rata.

**Kata Kunci:** KSA, potensi padi, k-means *cluster*

1. **PENDAHULUAN**

Menurut Undang-Undang Nomor 7 Tahun 1996 tentang pangan, menyatakan bahwa pangan merupakan hak asasi bagi setiap individu di Indonesia. Pangan merupakan kebutuhan manusia paling mendasar, sehingga dalam pemenuhannya harus mendapat perhatian serius. Dalam hal ini sektor pertanian khususnya tanaman pangan.

Secara umum, makanan pokok penduduk Indonesia adalah beras. Produksi beras di dalam negeri diharapkan dapat memenuhi kebutuhan semua masyarakat Indonesia sehingga tidak lagi melakukan impor dari negara lain. Faktanya, Indonesia masih mengimpor beras dari negara lain untuk memenuhi kebutuhan beras nasional.

Produksi padi nasional tahun 2021 sebesar 54,42 juta ton GKG, dengan luas panen padi sebesar 10,41 juta hektar. Jika dikonversikan menjadi beras, produksi beras pada 2021 mencapai sekitar 31,36 juta ton (BPS, 2022).

Upaya untuk meningkatkan hasil produksi padi secara teknis terus dilakukan oleh Pemerintah khususnya Kementrian Pertanian dengan perluasan areal tanam, peningkatan produkstivitas dengan penggunaan benih varietas unggul dengan potensi produksi yang tinggi, pengembangan padi hibrida, anjuran penggunaan pupuk berimbang, pengamanan pertanaman melalui pengawalan pengendalian OPT dan menekan kehilangan hasil melalui penanganan pasca panen yang baik.

Selain dari sisi teknis upaya peningkatan hasil produksi juga harus diikuti dengan penyediaan data penghitungan produksi beras secara akurat, sehingga antara data dan fakta kondisi di lapangan sama.

Sejak tahun 2018, BPS bekerja sama dengan Badan Pengkajian dan Penerapan Teknologi (BPPT) dan Lembaga Penerbangan dan Antariksa Nasional (LAPAN) yang sekarang bergabung mejadi Badan Riset dan Inovasi Nasional (BRIN), Kementrian Agraria dan Tata Ruang/Badan Pertanahan Nasional (Kementerian ATR/BPN), serta Badan Informasi Geospasial (BIG) berupaya memperbaiki metodologi penghitungan luas panen padi melalui penerapan objective measurement dengan memanfaatkan kemajuan teknologi informasi dan komunikasi serta ketersediaan citra satelit resolusi tinggi.

Kolaborasi tersebut diwujudkan dalam satu kegiatan yang bertajuk “Pendataan Statistik Pertanian Tanaman Pangan Terintegrasi dengan Metode Kerangka Sampel Area (KSA)” atau lebih dikenal dengan Survei KSA.

Pelaksanaan Survei KSA untuk komoditas padi mulai diimplementasikan secara nasional pada tahun 2018 dan termasuk dalam proyek nasional untuk mendukung pencapaian salah satu prioritas nasional, yaitu ketahanan pangan, dalam rangka perbaikan data statistik pangan (BPS, 2022).

Melihat begitu krusialnya sektor pertanian, khusunya padi, sebagai pemenuhan pangan penduduk Indonesia, maka perencanaan berbasis wilayah mutlak diperlukan. Hal ini diperlukan guna melihat potensi padi di Indonesia berdasarkan wilayah (provinsi).

Pengelompokkan potensi padi berdasarkan provinsi di Indonesia bisa dilakukan dengan analisis *cluster*. Analisis *cluster* adalah metode statistik multivariat dimana objek penelitian diklasifikasikan berdasarkan karakteristik objek itu sendiri (Bangun, 2016).

Penelitian sebelumnya mengenai analisis *cluster* dalam pengelompokkan wilayah berdasarkan faktor produksi padi telah dilakukan antara lain oleh Bangun (2016), Sudono, dkk (2013), dan Sugiyani (2016).

Berdasarkan latar belakang tersebut, tujuan penelitian ini adalah untuk pengelompokkan provinsi di Indonesia berdasarkan variabel potensi padi.

**II. TINJAUAN PUSTAKA**

2.1 Produksi Padi

Padi di Indonesia terdiri dari padi sawah dan padi ladang. Produksi padi diperoleh dari hasil perkalian antara luas panen (bersih) dengan produktivitas. Luas panen tanaman padi di lahan sawah dikoreksi dengan besaran konversi galengan.

Sementara itu, untuk luas panen tanaman padi di lahan bukan sawah, luas galengan dianggap tidak ada (tidak dikoreksi dengan besaran konversi galengan).

Produksi beras didapatkan dari hasil perkalian produksi padi/gabah dengan angka konversi gabah ke beras. Produksi padi dan beras dihitung pada level kabupaten/kota.

Penyempurnaan dalam berbagai tahapan penghitungan produksi telah dilakukan secara komprehensif tidak hanya luas lahan baku sawah saja, tetapi juga perbaikan penghitungan konversi gabah kering menjadi beras. Secara garis besar tahapan penghitungan produksi beras:

1. Penetapan luas lahan baku sawah nasional yang digunakan untuk mengestimasi luas panen yang ditetapkan berdasarkan Keputusan Menteri ATR/Kepala BPN No. 686/SK-PG.03.03/XII/2019 tanggal 17 Desember 2019 yaitu sebesar 7.463.948 ha.
2. Pengamatan fase tumbuh padi untuk menghitung luas panen dengan KSA yang dikembangkan bersama BPPT dan telah mendapat pengakuan dari Lembaga Ilmu (LIPI), dimana saat ini LIPI juga bergabung dalam Badan Riset dan Inovasi Nasional (BRIN).
3. Pengukuran produktivitas padi per hektar melalui Survei Ubinan yang telah dilakukan penyempurnaan dengan mengganti metode ubinan berbasis rumah tangga menjadi berbasis sampel KSA.
4. Penetapan angka konversi dari gabah kering panen (GKP) ke gabah kering giling (GKG) dan angka konversi dari GKG ke beras yang berasal dari Survei Konversi Gabah ke Beras tahun 2018. Survei ini menyediakan angka konversi yang lebih akurat karena dilakukan di dua periode musim yang berbeda pada level provinsi sehingga didapatkan angka konversi untuk masing-masing provinsi yang memperhitungkan pengaruh musim.

2.2 Penelitian terdahulu

Maria Yunitasari (2007) melakukan penelitian tentang analisis hubungan tentang pertumbuhan ekonomi dengan pembangunan manusia provinsi jawa timur. Hasil penelitian menunjukkan bahwa pertumbuhan ekonomi memiliki pengaruh signifikan terhadap pembangunan manusia dengan nilai koefisien positif.

Novia Hera Pratami (2020) melakukan penelitian dengan judul Hubungan Indeks Pembangunan Manusia Terhadap Pertumbuhan Ekonomi dikabupaten Karanganyar. Hasil penelitian menunjukkan bahwa variabel indeks pembangunan manusia memiliki hubungan yang positif dan signifikan dengan laju pertumbuhan ekonomi dengan derajat hubungan sedang.

Moh Muqorobin (2017) melakukan penelitian dengan judul Pengaruh Indeks Pembangunan Manusia (IPM) Terhadap Pertumbuhan Ekonomi Jawa Timur dengan metode regresi linier sederhana. Hasil penelitian menunjukkan bahwa menunjukkan bahwa variabel indeks pembangunan manusia berpengaruh negatif dan signifikan terhadap pertumbuhan ekonomi provinsi Jawa Timur.

**III. METODOLOGI**

3.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data BPS mengenai luas panen, produksi, dan produktivitas padi di Indonesia tahun 2021.

3.2 Metode Pengolahan dan Analisis

Penelitian ini dilakukan di seluruh provinsi yang ada di Indonesia dengan menggunakan data sekunder. Data diperoleh dari tabel statis Badan Pusat Statistik.

Variabel-variabel yang menjadi penelitian antara lain:

X1 : Luas Panen padi tahun 2021

X2 : Produktivitas padi tahun 2021

X3 : Produksi padi tahun 2021

Metode yang digunakan dalam penelitian ini adalah analisis data dengan banyak peubah (multivariat) yaitu Analisis Komponen Utama (AKU) yang dirangkai dengan analisis faktor, dan menggunakan analisis *cluster* non hierarki.

Langkah-langkah analisis *cluster* non-hirearki dilakukan sebagai berikut:

1. Analisis Komponen Utama (AKU)

Analisis ini digunakan untuk mengetahui apakah kajian ini layak untuk dilakukan analisis lebih lanjut dengan menggunakan Analisis Faktor. Kelayakan tersebut dilakukan berdasarkan perhitungan metode Kaiser Meyer Olkin (KMO) dan uji Bartlett.

1. Analisis faktor

Merupakan suatu analisis data untuk mengetahui faktor-faktor yang dominan dalam menjelaskan suatu masalah.

1. Analisis *Cluster*

Merupakan salah satu metode *data mining* yang bersifat tanpa arahan (*unsupervised*). Ada dua jenis data *clustering* yang sering dipergunakan dalam proses pengelompokan data yaitu *hierarchical* (hierarki) data *clustering* dan *non-hierarchical* (non hierarki) data *clustering*.

Salah satu metode data *cluster*ing non hierarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*/kelompok.

Metode ini mempartisi data ke dalam *cluster*/kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang mempunyai karakteristik berbeda dikelompokkan ke dalam kelompok yang lain.

Adapun tujuan dari data *cluster*ing ini adalah untuk meminimalisasikan fungsi obyektif yang diset dalam proses *cluster*ing, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu *cluster*/kelompok dan memaksimalkan variasi antar *cluster*/kelompok.

Metode *cluster* K-means dimulai dengan menentukan sejumlah nilai *cluster*/kelompok awal sesuai dengan jumlah yang diingankan. Setelah ditentukan, objek kemudian digabungkan ke dalam *cluster*-*cluster* tersebut. Secara umum metode K-means *Cluster Analysis* menggunakan algoritma K-means.

**IV. HASIL PEMBAHASAN**

* 1. Analisis Deskriptif

Produksi padi di Indonesia pada tahun 2021 sebesar 54.415.294,22 ton Gabah Kering Giling. Daerah yang merupakan sentra produksi padi di Indonesia adalah Jawa Timur, Jawa Tengah, dan Jawa Barat yang memberikan kontribusi sebesar 52,4 persen terhadap produksi padi di Indonesia.

Produksi padi dan luas panen padi di Indonesia tahun 2021 berdasarkan provinsi dapat dilihat pada Gambar 1 dan Gambar 2.

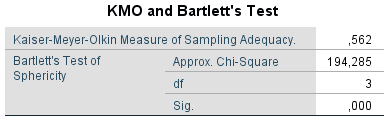
Gambar 1. Produksi Padi di Indonesia 2021

Gambar 2. Luas Panen Padi di Indonesia 2021

* 1. Analisis Komponen Utama dan Analisis Faktor

Berdasarkan hasil tabel 1, diketahui bahwa uji KMO and Barlett’s sebesar 0,562 dan nilai sig 0,000. Adapun nilai Measure of Sampling Adequacy (MSA) untuk semua variabel > 0,50. Hal ini dapat kita ambil kesimpulan bahwa data luas panen, produksi dan produktivitas dapat dilanjutkan untuk menggunakan teknik analisis faktor.

Tabel 1. KMO dan Uji Barlett’s

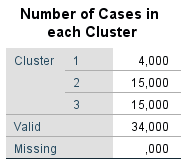


Sumber. Data Olahan

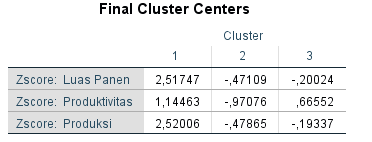
* 1. Analisis K-means *Cluster*

Hasil pengujian analisis *Cluster* K-means:

Tabel 2. Jumlah *Cluster*

****

Tabel 3. Hasil Pembentukan kluster

****

Sumber. Data Olahan

Dari tabel 2 dan tabel 3 dapat dijelaskan bahwa:

|  |  |
| --- | --- |
| CL | Nama Provinsi |
| 1 | Jawa Barat, Jawa Tengah, Jawa Timur, dan Sulawesi Selatan. |
| 2 | Riau, Kep. Bangka Belitung, Kep. Riau, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tenggara, Maluku, Maluku Utara, Papua Barat, dan Papua. |
| 3 | Aceh, Sumatera Utara, Sumatera Barat, Jambi, Sumatera Selatan, Bengkulu, Lampung, DKI Jakarta, DI Yogyakarta, Banten, Bali, Nusa Tenggara Barat, Sulawesi Tengah, Gorontalo, dan Sulawesi Barat. |

Ada sebanyak 4 provinsi yang masuk dalam *Cluster* 1, dalam *Cluster* ini dapat dikatakan bahwa provinsi dengan potensi padi tertinggi atau penyumbang produksi tertinggi di Indonesia.

Sedangkan untuk *Cluster* 2 ada sebanyak 15 provinsi yang masuk dalam *Cluster* ini, bisa dikatakan provinsi yang mempunyai produksi di atas rata-rata, yaitu di atas 1,6 ton Gabah Kering Giling.

Untuk *Cluster* 3 juga ada sebanyak 15 provinsi, bisa dikatakan juga bahwa provinsi yang mempunyai produksi di bawah rata-rata.

**V. KESIMPULAN**

5.1 Kesimpulan

Dari hasil analisis, dapat ditarik kesimpulan. Terdapat 3 *cluster* potensi padi di Indonesia, yaitu

1. *Cluster* 1 yang merupakan sentra padi di Indonesia, ada sebanyak 4 provinsi, antara lain: Jawa Barat, Jawa Tengah, Jawa Timur, dan Sulawesi Selatan.

2. *Cluster* 2 yang mempunyai nilai produksi di atas rata-rata, yaitu diatas 1,6 ton gabah kering giling, ada sebanyak 15 provinsi, antara lain: Riau, Kep. Bangka Belitung, Kep. Riau, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tenggara, Maluku, Maluku Utara, Papua Barat, dan Papua.

3. *Cluster* 3 yang mempunyai nilai di bawah rata-rata. Ada sebanyak 15 provinsi, antara lain: Aceh, Sumatera Utara, Sumatera Barat, Jambi, Sumatera Selatan, Bengkulu, Lampung, DKI Jakarta, DI Yogyakarta, Banten, Bali, Nusa Tenggara Barat, Sulawesi Tengah, Gorontalo, dan Sulawesi Barat.

Setiap *cluster* memiliki karakteristik masing-masing yang berbeda dengan *cluster* lainnya dan dapat dilihat pengelompokkan wilayah berdasarkan potensi padi. Hasil ini dapat dikembangkan sebagai bahan penelitian lanjutan sebagai program peningkatan dan pengembangan produksi tanaman pangan khususnya padi, dengan tujuan mencapai swasembada pangan dan ketahanan pangan.

5.2 Saran

Pemerintah dapat menentukan prioritas daerah mana yang dapat dijadikan potensi padi di Indonesia. Karena tersedianya data pertanian yang tepat waktu dan akurat merupakan pondasi untuk mewujudkan kebijakan pertanian yang tepat sasaran.

**DAFTAR PUSTAKA**

Sudono, dkk. 2013. Pengelompokan Produksi Padi Nasional dengan Pendekatan Data Mining Konsep K-means. Peneliti di Balai Irigasi, Pusat Litbang Sumber Daya Air, Kementerian Pekerjaan Umum.

Bangun, R. H. B. 2016. Analisis Klaster Non-Hierarki Dalam Pengelompokan Kabupaten/Kota Di Sumatera Utara Berdasarkan Faktor Produksi Padi. Jurnal Agribisnis Sumatera Utara. BPS. Sumatera Utara.

Fathia, dkk. 2016. Analisis Klaster Kecamatan Di Kabupaten Semarang Berdasarkan Potensi Desa Menggunakan Metode Ward Dan Single Linkage. Jurnal Gaussian. UNDIP. Semarang.

Alam, dkk. 2017. Analisis *Cluster* Pada Desa/Kelurahan Di Kabupaten Wonosobo Berdasarkan Data Potensi Desa Tahun 2015. The 5TH Urecol Proceeding. AIS Muhammadiyah. Semarang.

Safitri, dkk. 2012. Analisis *Cluster* Pada Kabupaten/Kota Di Jawa Tengah Berdasarkan Produksi Palawija. Media Statistika. UNDIP. Semarang.

Silvi, Rini. 2018. Analisis *Cluster* Dengan Data Outlier Menggunakan Centroid Linkage dan K-means *Cluster*ing Untuk Pengelompokan Indikator HIV/AIDS di Indonesia. Jurnal Matematika “MANTIK”. Universitas Padjajaran. Semarang

Sugiyani, Yani. 2016. Pengelompokan Wilayah Berdasarkan Potensi Hasil Pertanian Menggunakan K-means Di Kota Cilegon. Jurnal ProTekInfo. Universitas Serang Raya. Banten.