

ANALISIS PENGELOMPOKAN DAERAH MENGGUNAKAN METODE *NON-HIERARCHICAL PARTITIONING K-MEDOIDS* DARI HASIL KOMODITAS PERTANIAN TANAMAN PANGAN

(Studi Kasus Kabupaten/Kota Se-Jawa Tengah Tahun 2009 – 2013)

Etik Setyowati¹, Agus Rusgiyono², Moch. Abdul Mukid³

¹Mahasiswa Jurusan Statistika FSM Universitas Diponegoro

^{2,3}Staff Pengajar Jurusan Statistika FSM Universitas Diponegoro

ABSTRACT

Non-Hierarchical K-Medoids Partitioning is a clustering method for classifying objects based on the characteristics possessed by the object, wherein the object k randomly selected to be medoids is the center of the cluster. After medoids selected then other objects that have similarities with medoids made in one cluster. Medoids is the object which is considered to represent a cluster. Similarity between objects is calculated using euclidean distance. One application grouping method Non-Hierarchical K-Medoids Partitioning is to classify District in Central Java is based on the production of rice and pulses. Grouping Regency / City in Central Java using Non-Hierarchical Partitioning K-Medoids obtained information that rice production by Regency / City in Central Java can be grouped into seven clusters, but because of a case in 2010 and in 2011 the number of clusters that formed are two clusters, while the production of food crops by Regency / City in Central Java can be grouped into two clusters.

Keywords: k-medoids, Non-Hierarchical, Euclidean distance, Similarities.

1. PENDAHULUAN

Negara Indonesia merupakan negara agraris dengan sebagian besar penduduknya bermata pencaharian sebagai petani. Salah satu Provinsi yang cocok untuk daerah pertanian adalah Provinsi Jawa Tengah. Daerah Jawa Tengah bagian utara dan beberapa daerah di Jawa Tengah bagian tengah cocok untuk pertanian tanaman pangan padi dan palawija. Adanya alih fungsi lahan dan berkurangnya jumlah petani tiap tahun dapat menyebabkan ketahanan pangan nasional terancam. Upaya pemerintah untuk mewujudkan ketahanan pangan nasional agar tetap stabil adalah dengan memberlakukan kebijakan kebijakan tertentu dibidang pertanian.

Berdasarkan fakta-fakta diatas maka penulis mencoba membuat pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan hasil komoditas pertanian tanaman pangan tahun 2009-2013. Pengelompokan ini dilakukan karena beragamnya potensi hasil produksi pertanian tanaman pangan di dimasing-masing Kabupaten/Kota di Jawa Tengah, sehingga perlu dilakukan pengelompokan untuk mengetahui Kabupaten/Kota mana yang memiliki hasil produksi pertanian tertinggi sehingga dapat membantu untuk mengoptimalkan program-program pemerintah dibidang pertanian tanaman pangan. Hasil produksi pertanian di Kabupaten/Kota se-Jawa Tengah berbeda-beda, ada daerah yang memiliki produksi rendah ada juga daerah yang memiliki produksi padi sangat tinggi, untuk itu analisis yang dilakukan untuk mengelompokkan daerah tersebut adalah dengan

menggunakan metode *non-hierarchical partitioning k-medoids*. Metode ini merupakan salah satu metode yang bertujuan untuk mengelompokkan objek berdasarkan karakteristik yang dimiliki objek tersebut. Metode *non-hierarchical partitioning k-medoids* lebih baik dibandingkan metode *non-hierarchical partitioning k-means* karena tidak sensitif terhadap *outlier*.

2. TINJAUAN PUSTAKA

2.1. Analisis Cluster

Menurut Hair, *et al.* (2006), analisis cluster adalah teknik pengelompokan untuk mengelompokkan objek berdasarkan karakteristik yang dimiliki oleh objek tersebut. Analisis cluster mengelompokkan objek yang paling mirip ke dalam satu kelompok yang sama sehingga, karakteristik objek dalam satu cluster bersifat homogen sedangkan karakteristik objek antar cluster yang lain bersifat heterogen.

2.2. Objek *Outlier*/pencilan

Menurut Hair, *et al.* (2006), adanya *outlier* dapat menyebabkan pengamatan menyimpang dan tidak mewakili keadaan populasi. *Outlier* juga menyebabkan struktur yang tidak jelas sehingga cluster yang terbentuk menjadi tidak representatif. Cara untuk mendeteksi adanya *outlier univariat* dapat dilakukan dengan mengubah nilai data ke dalam nilai standar/z-score.

$$Z_{ij} = \frac{x_{ij} - m_j}{s_j}$$

dengan:

- x_{ij} = nilai objek ke- i pada variabel ke- j
- i = 1, 2, 3, . . . , n
- Z_{ij} = data x_{ij} yang sudah terstandarkan
- s_j = simpangan baku dari variabel ke- j
- m_j = rata-rata dari variabel ke- j

Data dikatakan *outlier* jika nilai $Z_{ij} > 4$ untuk ukuran sampel besar (diatas 80 observasi) atau nilai $Z_{ij} > 2,5$ untuk ukuran data kecil. Cara untuk mendeteksi adanya *outlier multivariat* dapat dilakukan dengan menggunakan jarak *mahalanobis* yang dirumuskan sebagai berikut:

$$D^2 = (\mathbf{X} - \boldsymbol{\mu})^T \mathbf{A}^{-1} (\mathbf{X} - \boldsymbol{\mu})$$

dengan :

- D^2 = jarak *mahalanobis*
- \mathbf{X} = matrik data
- $\boldsymbol{\mu}$ = vektor rata-rata dari matrik \mathbf{X}
- \mathbf{A}^{-1} = invers matrik kovarian dari \mathbf{X}

Untuk sampel kecil (kurang dari 80) data dikatakan *outlier* jika $\frac{D^2}{\text{banyak variabel}} > 2,5$

sedangkan untuk sampel besar (lebih dari 80) data dikatakan *outlier* jika $\frac{D^2}{\text{banyak variabel}} > 4$.

Menurut Santoso (2002), *outlier* bisa terjadi karena beberapa sebab antara lain :

1. Kesalahan dalam memasukkan data.
2. Kesalahan dalam pengambilan sampel.
3. Memang ada data ekstrim yang tidak bisa dihindarkan keberadaannya.

Data *outlier* dapat dihilangkan jika terjadi karena kesalahan pengambilan data atau input data. Data outlier dapat tetap dipertahankan jika disebabkan karena memang ada data *outlier* yang tidak bisa dihindari keberadaannya

2.3. Standarisasi/Pembakuan Data

Menurut Hair, *et al.* (2006), pembakuan data adalah proses mengkonversi nilai masing-masing data awal menjadi nilai standar dengan rata-rata 0 dan standar deviasi 1 untuk menghilangkan bias yang disebabkan karena perbedaan skala dari beberapa variabel yang digunakan dalam analisis. Untuk menghitung nilai standar tersebut digunakan rumus sebagai berikut:

$$Z_{ij} = \frac{x_{ij} - m_j}{s_j}$$

2.4. Asumsi dalam Analisis Cluster

Menurut Hair, *et al.* (2006), asumsi yang harus dipenuhi dalam analisis cluster adalah kecukupan sampel untuk mewakili populasi dan pengaruh multikolinieritas.

1. Kecukupan Sampel Mewakili Populasi

Sampel yang digunakan dalam analisis cluster harus dapat mewakili populasi yang ingin dijelaskan, karena hasil analisis cluster akan baik jika sampel yang digunakan representatif.

2. Pengaruh Multukolinieritas

Menurut Hair, *at al.* (1998), multikolinieritas antar variabel adalah suatu pelanggaran dalam analisis cluster karena dapat mempengaruhi hasil cluster.

Untuk mengetahui ada tidaknya multikolinieritas dapat digunakan nilai *variance inflation factor* (VIF).

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

dengan:

R_j^2 = koefisien determinasi variabel ke-j

j = 1, 2, 3, . . . , p

VIF_j = *variance inflation factor* variabel j

2.5. Metode *Non-Hierarchical Partitioning K-Medoids*

Menurut Han dan Kamber (2006), dalam prosedur *partitioning* objek data yang akan di cluster dipisah atau disekat menjadi k-cluster dimana jumlah $k \leq$ banyaknya data yang akan di cluster. Analisis cluster yang menggunakan prosedur partisi salah satunya adalah *k-medoids*. *Medoids* adalah objek yang dianggap mewakili cluster sekaligus sebagai pusat cluster. Analisis ini mencoba meminimumkan ketidaksamaan setiap objek dalam satu cluster dengan meminimumkan nilai *absolute error*. Nilai dari *absolute error* dirumuskan sebagai berikut:

$$E = \sum_{c=1}^k \sum_{i=1}^{n_c} |p_{ic} - O_c|$$

dengan:

E = *absolute error*

n_c = banyak objek dalam cluster ke-c,

p_{ic} = objek *non-medoids* i dalam cluster ke-c

O_c = *medoids* di cluster ke-c

i = 1, 2, 3, . . . , n_c

c = 1, 2, 3, . . . , k

Adapun algoritma analisis cluster *non-hierarchical K-medoids* adalah sebagai berikut:

- Memilih k objek untuk menjadi O_c , dengan O_c adalah objek yang menjadi *medoid* di cluster ke- c dan $c = 1, 2, 3, \dots, k$
 - Menghitung kemiripan antara objek *medoid* dengan objek *non-medoid* menggunakan jarak *euclidean*
 - Menempatkan objek *non-medoids* ke dalam kelompok yang paling dekat dengan *medoids*
 - Secara acak memilih O_{random} , dengan O_{random} adalah sebuah objek *non-medoids* untuk menggantikan O_c awal
 - Menghitung kemiripan antara objek *non- O_{random}* dengan objek O_{random} menggunakan jarak *euclidean*
 - Menempatkan objek *non- O_{random}* ke dalam kelompok yang paling mirip dengan O_{random}
 - Menghitung nilai *absolut error* sebelum dan sesudah pertukaran O_c dengan O_{random} , jika $E_{random} < E_c$ maka tukar O_j dengan O_{random} tetapi jika $E_{random} > E_c$ maka O_c tetap.
 - Mengulangi langkah d sampai g hingga semua objek *non-medoids* terpilih menjadi O_{random} dan tidak terjadi perubahan pada O_c
- Mengulangi langkah d sampai g hingga semua objek *non-medoids* terpilih menjadi O_{random} dan tidak terjadi perubahan pada O_j

2.6. Interpretasi dan Pembuatan Profil Cluster

Menurut Supranto (2010), menginterpretasi dan pembuatan profil cluster meliputi pengkajian rata-rata nilai objek yang terdapat dalam satu cluster pada setiap variabel. Menurut Hair, *et al.* (2006), interpretasi dilakukan dengan menghitung rata-rata hasil pengclusteran sehingga akan memberikan gambaran yang logis tentang hasil cluster, perhitungan rata-rata tersebut dilakukan pada data asli, jika data telah distandarisasi atau jika analisis cluster dilakukan dengan menggunakan hasil analisis analisis komponen utama, peneliti harus mengembalikan ke nilai asli, baru kemudian menghitung rata-rata menggunakan data tersebut.

2.7. Validasi Hasil Cluster

Menurut Dalton, *et al.* (2009), salah satu pengukuran validasi internal adalah *Silhouette Index*. *Silhouette Index* mengevaluasi objek secara visual baik di dalam cluster maupun yang berada diluar, berdasarkan rentang nilai *silhouettenya*. Nilai *silhoutte* objek mendefinisikan kedekatannya dengan kelompok sendiri dengan kelompok lain.

Rumus rentang *silhoutte* adalah :

$$sc_i = \frac{b_i - a_i}{\max[b_i, a_i]}$$

$$b_i = \min r_{(i,j)}$$

$$SC = \frac{\sum_{i=1}^n sc_i}{n}$$

dengan:

- a_i = rata-rata kemiripan antara objek ke- i dengan objek lainnya dalam satu cluster
- b_i = minimum rata-rata kemiripan antara objek ke- i dengan objek lain di masing-masing cluster
- $r_{(i,j)}$ = rata-rata kemiripan antara objek ke- i dengan objek lain di cluster ke- j , dengan i bukan objek dari cluster ke- j dan $j = 1, 2, 3, \dots, k$
- k = banyak cluster
- sc_i = *coefisien silhoutte* objek- i , dengan $i = 1, 2, \dots, n$
- SC = rata-rata rentang nilai *coefisien silhoutte* objek secara keseluruhan

Menurut Lewis (2010), jika nilai *coefisien silhoutte* (sc_i) < 0 maka objek berada dalam kelompok yang salah sedangkan jika nilai *coefisien silhoutte* (sc_i) > 0 maka objek berada dalam kelompok yang benar dan jika nilai *coefisien silhoutte* (sc_i) $= 0$ maka objek berada diantara dua cluster sehingga objek tersebut belum dapat diputuskan masuk di dalam cluster yang mana. Ketika nilai rata-rata rentang *coefisien silhoutte* objek secara keseluruhan (SC) $> 0,51$ menunjukkan bahwa struktur hasil pengelompokan sudah dianggap baik, sedangkan rata-rata rentang *coefisien silhoutte* objek secara keseluruhan (SC) $> 0,71$ menunjukkan bahwa struktur hasil pengelompokan sangat baik.

3. METODE PENELITIAN

3.1. Data Penelitian

Data yang digunakan dalam penelitian ini diperoleh dari Badan Pusat Statistik Provinsi Jawa Tengah berupa data sekunder hasil pertanian komoditas pertanian tanaman pangan tahun 2009-2013. Untuk penelitian ini data yang digunakan adalah data sampel sebanyak 251 kecamatan yang diambil dari 573 kecamatan di Jawa Tengah.

3.2. Variabel Penelitian

Variabel yang digunakan dalam pengelompokan Kabupaten/Kota berdasarkan produksi padi menggunakan variabel produksi padi, sedangkan untuk pengelompokan Kabupaten/Kota berdasarkan produksi palawija menggunakan variabel jagung, kedelai, ubi jalar, ubi kayu, kacang tanah, dan kacang hijau.

3.3. Tahapan Analisis

Pengelompokan Kabupaten/Kota di Jawa Tengah dalam penelitian ini dilakukan berdasarkan produksi padi dan berdasarkan produksi palawija dari tahun 2009-2013. Tahap-tahap pengolahan data adalah sebagai berikut:

1. Pemeriksaan kelengkapan data
Pemeriksaan kelengkapan data dengan cara pengecekan *missing value*.
2. Standarisasi/pembakuan data.
Pembakuan data ini merupakan proses pengubahan dari data awal menjadi data standar dengan mean 0 dan varian 1.
3. Deteksi *outlier*/pencilan
Outlier adalah suatu objek hasil pengamatan yang secara nyata berbeda dengan objek lain.
4. Asumsi dalam analisis cluster
Asumsi dalam analisis cluster yang harus dipenuhi adalah sampel harus mewakili populasi yang ada dan tidak terjadi masalah multikolinieritas.
5. Melakukan analisis cluster
Metode cluster yang digunakan dalam penelitian ini adalah metode *non-hierarchical partitioning k-medoids* dengan langkah sebagai berikut:
 - a. Memilih k objek untuk menjadi O_c , dengan O_c adalah objek yang menjadi *medoid* di cluster ke- c dan $c = 1, 2, \dots, k$
 - b. Menghitung kemiripan antara objek *medoid* dengan objek *non-medoid* menggunakan jarak *euclidean*
 - c. Menempatkan objek *non-medoids* ke dalam kelompok yang paling dekat dengan *medoids*

- d. Secara acak memilih O_{random} , dengan O_{random} adalah sebuah objek *non-medoids* untuk menggantikan O_c awal)
- e. Menghitung kemiripan antara objek *non- O_{random}* dengan objek O_{random} menggunakan jarak *euclidean*
- f. Menempatkan objek *non- O_{random}* ke dalam kelompok yang paling mirip dengan O_{random}
- g. Menghitung nilai *absolut error* sebelum dan sesudah pertukaran O_c dengan O_{random} , jika $E_{random} < E_c$ maka tukar O_c dengan O_{random} tetapi jika $E_{random} > E_c$ maka O_c tetap.
- h. Mengulangi langkah d sampai g hingga semua objek *non-medoids* terpilih menjadi O_{random} dan tidak terjadi perubahan pada O_c

4. HASIL DAN PEMBAHASAN

4.1. Pengelompokan Kabupaten di Jawa Tengah Berdasarkan Produksi Padi

Dari data hasil pertanian tanaman padi tahun 2009-2013 yang di peroleh dari Badan Pusat Statistik Jawa Tengah dilakukan pengecekan *missing value*, dari hasil yang diperoleh di simpulkan bahwa tidak terdapat *missing value* pada data tersebut.

Pembakuan data dilakukan apabila variabel yang digunakan dalam pengelompokan mempunyai satuan yang berbeda. Pada penelitian ini tidak dilakukan standarisasi data karena variabel yang digunakan untuk pengelompokan daerah berdasarkan hasil produksi padi hanya terdiri dari satu variabel yaitu variabel produksi tanaman padi dengan satuan ton. Nilai z-score dari masing-masing produksi padi tahun 2009 tahun 2010, tahun 2011, tahun 2012 dan tahun 2013 adalah kurang dari 4 sehingga dapat disimpulkan bahwa tidak ada objek *outlier* pada data tanaman padi tahun 2009, tahun 2010, tahun 2011, tahun 2012 dan tahun 2013.

Dengan menggunakan teknik pengambilan sampel *proporsional stratified random sampling* untuk pengambilan sampel dengan populasi seluruh kecamatan yang ada di Jawa Tengah peneliti yakin bahwa sampel yang diambil dapat merepresentasikan/mewakili populasi. Dalam pengelompokan kabupaten berdasarkan produksi padi, variabel yang digunakan hanya satu yaitu produksi padi sehingga tanpa pengecekan menggunakan nilai VIF dapat dikatakan tidak terjadi masalah multikolinieritas dari hasil produksi padi tahun 2009, 2010, 2011, 2012 maupun 2013.

4.1.1. Hasil Pengelompokan Kabupaten di Jawa Tengah Berdasarkan Produksi Padi

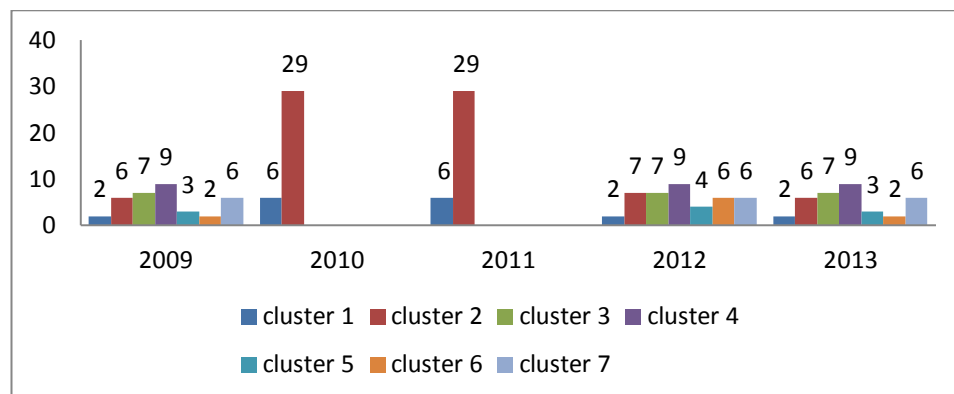
Hasil pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi padi menggunakan metode *non-hierarchical partitioning k-medoids* adalah sebagai berikut:

1. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi padi tahun 2009 menghasilkan 7 cluster, dengan rata-rata rentang nilai *coefisien silhouette* (SC) 0,67.
2. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi padi tahun 2010 menghasilkan 2 cluster, dengan rata-rata rentang nilai *coefisien silhouette* (SC) 0,71.
3. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi padi tahun 2011 menghasilkan 2 cluster, dengan rata-rata rentang nilai *coefisien silhouette* (SC) 0,68.
4. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi padi tahun 2012 menghasilkan 6 cluster, dengan rata-rata rentang nilai *coefisien silhouette* (SC) 0,69.
5. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi padi tahun 2013 menghasilkan 7 cluster, dengan rata-rata rentang nilai *coefisien silhouette* (SC) 0,73.

Rata-rata produksi padi masing-masing cluster dari tahun 2009-2013 adalah sebagai berikut:

Tabel 1. Rata-rata Produksi Padi Hasil Pengklusteran Tahun 2009-2013

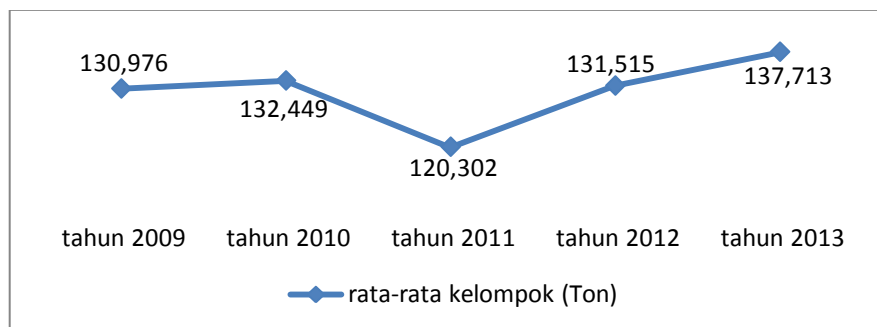
	2009	2010	2011	2012	2013
cluster 1	337.461	314.042	282.736	353.280	388.812
cluster 2	166.177	94.878	86.695	180.120	171.907
cluster 3	76.084			75.154	72.182
cluster 4	108.660			111.451	118.399
cluster 5	251.720			262.175	260.610
cluster 6	294.825			9.634	298.902
cluster 7	9.474				10.069



Gambar 1. Jumlah Kabupaten/Kota dimasing-masing Cluster Tahun 2009-2013

Dari Tabel 1 dan Gambar 1 dapat disimpulkan bahwa:

1. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi padi tahun 2009 menunjukkan sebagian besar Kabupaten memiliki rata-rata sedang sampai tinggi, hanya ada 6 Kabupaten yang memiliki rata-rata produksi terendah.
2. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi padi tahun 2010 menunjukkan sebagian besar Kabupaten di Jawa Tengah masuk dalam cluster ke-2 yang memiliki rata-rata produksi sebesar 94.878 ton pertahun, rata-rata ini jauh lebih kecil dibandingkan rata-rata produksi padi di cluster ke-1 yaitu sebesar 314.042 ton pertahun.
3. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi padi tahun 2011 menunjukkan hasil yang tidak jauh berbeda dengan tahun 2010, namun rata-rata produksi padi di cluster ke-1 maupun ke-2 mengalami penurunan.
4. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi padi tahun 2012 menghasilkan 6 cluster, ini menunjukkan bahwa hasil produksi padi mulai merata, hanya ada 6 Kabupaten yang memiliki rata-rata produksi terendah.
5. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi padi tahun 2013 menunjukkan hasil yang tidak jauh berbeda dengan tahun 2012. Sebagian besar kabupaten di Jawa tengah tahun 2013 memiliki rata-rata produksi sedang sampai tinggi.



Gambar 2. Rata-Rata Kelompok Hasil Pengclusteran Kabupaten Berdasarkan Produksi Padi Tahun 2009-2013

Dari Gambar 2 dapat disimpulkan bahwa perubahan jumlah cluster tahun 2010 dari 6 cluster di tahun 2009 menjadi 2 cluster terjadi karena kenaikan produksi padi hanya terjadi di beberapa daerah seperti Kabupaten Demak, Sragen, Brebes dan Pati.

Ditahun 2011 terjadi penurunan yang cukup signifikan dengan jumlah cluster sama dengan tahun 2010, hal ini menunjukkan bahwa di tahun 2011 terjadi penurunan hampir di seluruh Kabupaten/Kota di Jawa Tengah. Penurunan jumlah produksi padi tersebut terjadi karena berkurangnya luas panen.

4.2. Pengelompokan Kabupaten di Jawa Tengah Berdasarkan Produksi Palawija

Dari data hasil pertanian tanaman palawija tahun 2009-2013 yang diperoleh dari Badan Pusat Statistik Jawa Tengah dilakukan pengecekan *missing value*, dari hasil yang diperoleh di simpulkan bahwa tidak terdapat *missing value* pada data tersebut.

Pembakuan data dilakukan apabila variabel yang digunakan dalam pengelompokan mempunyai satuan yang berbeda. Pada penelitian ini tidak dilakukan standarisasi data karena variabel yang digunakan untuk pengelompokan daerah berdasarkan hasil produksi palawija memiliki satuan yang sama yaitu ton.

Untuk mendeteksi objek *outlier* dalam pengelompokan Kabupaten berdasarkan produksi palawija adalah dengan menggunakan jarak mahalnobis, jika jarak $\frac{D^2}{\text{banyak variabel}} > 4$ maka dapat dikatakan terdapat *outlier*. Berdasarkan nilai $\frac{D^2}{\text{banyak variabel}}$ dapat disimpulkan bahwa terdapat *outlier* pada data tanaman palawija tahun 2009, tahun 2010, tahun 2011, tahun 2012 dan tahun 2013 karena ada objek yang nilai jarak mahalnobis/jumlah variabelnya > 4 untuk masing-masing produksi tanaman palawija setiap tahunnya.

Dengan menggunakan teknik pengambilan sampel *proporsional stratified random sampling* untuk pengambilan sampel dengan populasi seluruh kecamatan yang ada di Jawa Tengah peneliti yakin bahwa sampel yang diambil dapat merepresentasikan/mewakili populasi.

Tabel 2. Nilai VIF Dari Hasil Produksi Pertanian Tanaman Palawija Tahun 2009-2013

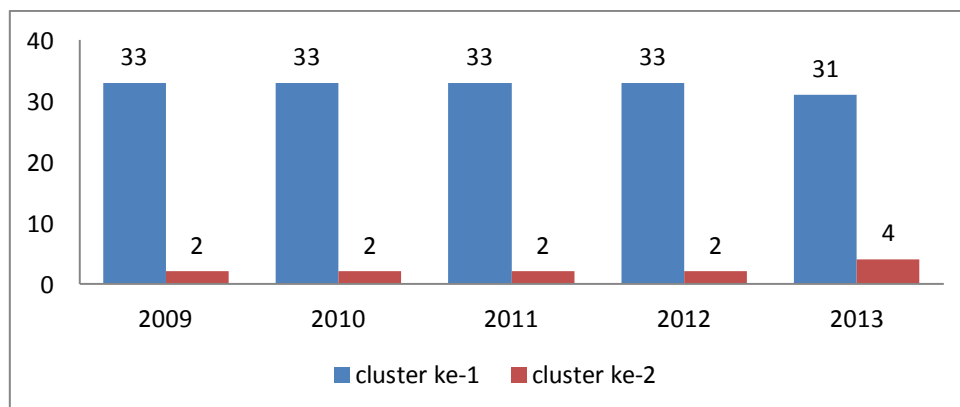
	Nilai VIF				
	2009	2010	2011	2012	2013
Jagung	3,095	6,550	1,760	6,250	3,287
Ubi kayu	2,469	1,973	1,868	2,203	1,771
Ubi jalar	1,165	1,165	1,053	1,141	1,065
Kedelai	3,993	7,694	2,706	6,877	3,947
Kacang tanah	3,0052	1,766	2,586	2,139	1,763
Kacang hijau	1,554	1,554	1,247	1,420	1,456

Dilihat dari tabel 2 nilai VIF dari hasil pertanian tanaman palawija dimasing-masing Kabupaten tahun 2009-2013 tidak terjadi masalah multikolinieritas karena nilai VIF < 10.

4.2.1. Hasil Pengelompokan Kabupaten di Jawa Tengah Berdasarkan Produksi Palawija

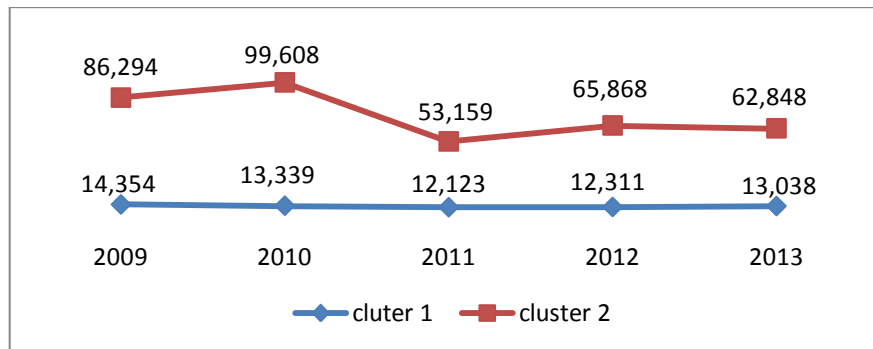
Hasil pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi palawija menggunakan metode *non-hierarchical partitioning k-medois* adalah sebagai berikut:

1. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi palawija tahun 2009 menghasilkan 2 cluster, dengan rata-rata rentang nilai *coefisien silhoutte* (SC) 0,78.
2. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi palawija tahun 2010 menghasilkan 2 cluster, dengan rata-rata rentang nilai *coefisien silhoutte* (SC) 0,83.
3. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi palawija tahun 2011 menghasilkan 2 cluster, dengan rata-rata rentang nilai *coefisien silhoutte* (SC) 0,67.
4. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi palawija tahun 2012 menghasilkan 2 cluster, dengan rata-rata rentang nilai *coefisien silhoutte* (SC) 0,82.
5. Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan produksi palawija tahun 2013 menghasilkan 2 cluster, dengan rata-rata rentang nilai *coefisien silhoutte* (SC) 0,76.



Gambar 3. Jumlah Kabupaten/Kota Dimasing-Masing Pengklasteran Tahun 2009-2013

Dilihat dari Gambar 3 dapat disimpulkan bahwa dari tahun 2009-2013 jumlah anggota cluster di cluster ke-1 selalu lebih banyak dari pada anggota cluster ke-2.



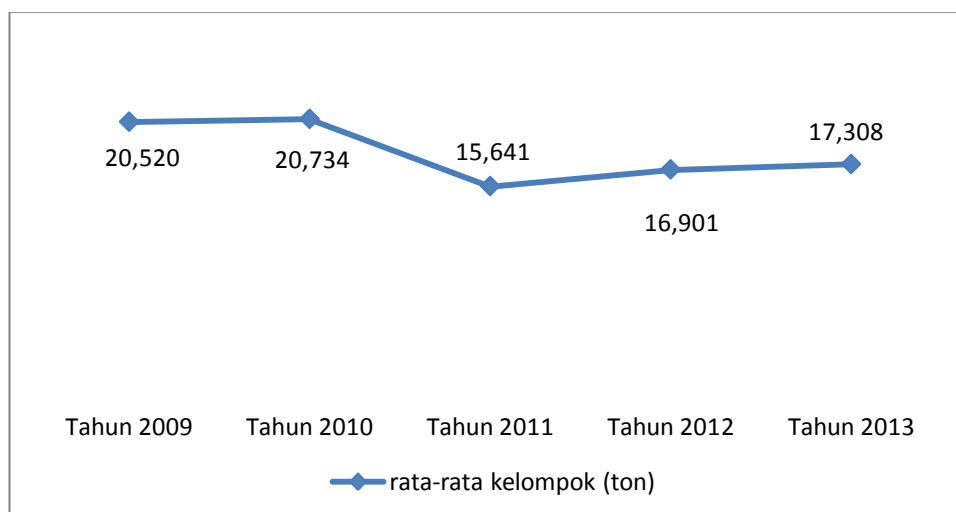
Gambar 4. Rata-rata Perkembangan Produksi Palawija Tahun 2009-2013

Dilihat dari Gambar 4 cluster ke-2 selalu memiliki rata-rata produksi palawija tertinggi dibandingkan cluster ke-1. Perkembangan rata-rata produksi palawija di cluster ke-1 tidak terlalu signifikan, sedangkan perkembangan rata-rata produksi palawija di cluster ke 2 cenderung mengalami fluktuasi naik turun dengan rata-rata produksi cukup besar.

Tabel 3. Rata-rata Produksi Palawija Hasil Pengklusteran Tahun 2009-2013

	Tahun 2009		Tahun 2010		Tahun 2011		Tahun 2012		Tahun 2013	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
Jagung	37.597	94.623	39.242	91.824	34.187	66.508	36.156	77.372	38.772	68.394
Ubi kayu	40.824	401.423	33.019	486.189	32.009	239.635	30.560	302.337	33.181	295.548
Ubi jalar	2.916	1.777	2.598	2.366	2.287	1.052	2.333	1.677	2.605	1.626
Kedelai	1.846	6.773	2.462	5.307	1.340	3.126	1.904	3.189	1.009	2.870
Kacang tanah	1.633	9.537	1.654	9.095	1.348	5.589	1.445	7.347	1.928	5.615
Kacang hijau	1.307	3.628	1.061	2.868	1.570	3.042	1.465	3.283	733	3.036

Dari Tabel 3 dapat disimpulkan bahwa Cluster ke-1 adalah kelompok Kabupaten yang memiliki hasil produksi ubi jalar tinggi sedangkan cluster ke-2 adalah kelompok Kabupaten yang memiliki hasil produksi jagung, ubi kayu, ubi jalar, kedelai, kacang tanah dan kacang hijau. Anggota cluster ke-2 antara lain Kabupaten Wonogiri, Pati, Jepara dan Kebumen, sedangkan Kabupaten yang lainnya masuk dalam cluster ke-1.



Gambar 5. Perkembangan Rata-rata Kelompok Produksi Palawija Tahun 2009-2013

Dilihat dari Gambar 5 dapat disimpulkan bahwa rata-rata produksi palawija hasil pengelompokan menunjukkan terjadi penurunan pada tahun 2011. Dilihat dari Gambar 4 dan Gambar 5 dapat disimpulkan bahwa perkembangan rata-rata produksi palawija di Jawa Tengah dipengaruhi oleh hasil produksi palawija di Kabupaten/Kota yang berada di cluster ke-2.

5. KESIMPULAN

Berdasarkan hasil dari pembahasan didapatkan kesimpulan bahwa:

1. Hasil pengelompokan Kabupaten/Kota di Jawa Tengah dari tahun 2009-2013 berdasarkan hasil produksi padi menunjukkan bahwa secara umum Kabupaten/Kota di Jawa Tengah dapat dikelompokkan menjadi 7 cluster, hal ini menunjukkan bahwa rata-rata Kabupaten/Kota di Jawa Tengah memiliki produksi padi baik. Pada tahun 2010 dan tahun 2011 jumlah cluster yang terbentuk hanya terdiri dari 2 cluster, hal ini terjadi karena ditahun 2010 rata-rata peningkatan jumlah produksi padi hanya terjadi di daerah tertentu seperti Demak, Sragen, Brebes dan Pati. Pada tahun 2011 rata-rata terjadi produksi padi di seluruh Kabupaten/Kota di Jawa Tengah, penurunan jumlah produksi padi ini disebabkan karena luas panen yang juga turun signifikan. Penurunan luas panen ini disebabkan pada tahun 2011 banyak petani yang mengalami gagal panen.
2. Hasil pengelompokan Kabupaten/Kota di Jawa Tengah dari tahun 2009-2013 berdasarkan hasil produksi palawija menghasilkan 2 cluster, dengan jumlah anggota cluster ke-1 selalu lebih banyak dari pada anggota cluster ke-2, rata-rata kelompok dari hasil produksi padi untuk masing-masing cluster menunjukkan rata-rata produksi cluster ke-2 selalu lebih tinggi dibandingkan cluster ke-1, ini menunjukkan bahwa hanya sebagian kecil Kabupaten/Kota yang memiliki produksi palawija tinggi khususnya untuk produksi jagung, ubi kayu, kedelai, kacang tanah dan kacang hijau. Kabupaten yang memiliki rata-rata produksi palawija tinggi tersebut adalah kabupaten Banyumas, Pati dan Purbalingga, sedangkan Kabupaten yang lainnya memiliki rata-rata produksi palawija yang rendah.

DAFTAR PUSTAKA

- Dalton, et.al.2009.*Clustering Algorithms: On Learning, Validation, Performance, and Applications to Genomics* <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2766793/>. Diakses 15 Februari 2015.
- Hair, et al.2006. *Multivariate data analysis*. edisi ke-6.pearson education.
- Han, Jiawai. Kamber, Micheline.2006.*Data Mining Concepts and Techniques*. second Edition. San Francisco:Elsever.
- Lewis, Paul D. 2010.*R For Medicine and Biology*. Jones and Bartlett Publisher
- Supranto, J.2010.*Analisis Multivariat Arti & Interpretasi*. Jakarta: Rineka Cipta.
- Susanto, Singgih.2014.*Satistik Multivariat Konsep Dan Aplikasi Dengan Spss*. Edisi Revisi. Jakarta:PT Elex media komputinda.