

PENGELOMPOKAN PROVINSI DI INDONESIA BERDASARKAN RATA-RATA PRODUKSI TANAMAN PANGAN MENGGUNAKAN METODE KLASTER K-MEANS

Dian Juli Adisaputra¹ Sigit Nugroho² Pepi Novianti²

¹ Alumni Matematika FMIPA Universitas Bengkulu

² Staf Pengajar Matematika FMIPA Universitas Bengkulu

ABSTRACT

This research applying k-means cluster analysis into classification of provinces in Indonesia based on the average production of food crops. The results of this study indicate that members of the Province in cluster 4 (Riau, Jambi, Bengkulu, Bangka Belitung, Riau Islands, Yogyakarta, Bali, East Nusa Tenggara, West Kalimantan, Central Kalimantan, East Kalimantan, North Kalimantan, North Sulawesi, Central Sulawesi, Southeast Sulawesi, Gorontalo, West Sulawesi, Maluku, North Maluku, West Papua and Papua) is a region with low average production of food crops. While members of the Province in cluster 2 (East Java) is a region with higher average production of food crops.

Keywords: Cluster Analysis, Non Hierarchical Cluster, K-Means, food crops

PENDAHULUAN

Pertanian merupakan salah satu sektor utama dalam pembangunan ekonomi di Indonesia, mengingat Indonesia merupakan daerah agraris. Selain kontribusinya dalam Produk Domestik Bruto (PDB), peranan sektor pertanian dalam pembangunan ekonomi sangat luas, diantaranya sebagai sektor penyerap tenaga kerja terbesar, sebagai penghasil makanan penduduk, dan sebagai penentu stabilitas harga. Pemerintah juga mengharapkan produksi pertanian tanaman pangan mengalami peningkatan tiap tahunnya (Rivani, 2010).

Tanaman Palawija merupakan tanaman yang potensial untuk dikembangkan karena hasilnya dapat digunakan sebagai sumber karbohidrat, sumber protein nabati, dan bahan dasar berbagai industri. Tanaman palawija meliputi jagung, kedelai, kacang tanah, kacang hijau, ubi kayu, dan ubi jalar (Safitri, 2012).

Dalam Undang-Undang Republik Indonesia Nomor 7 Tahun 1996, dikenal dua istilah penting tentang pangan, yaitu sistem pangan dan ketahanan pangan. Sistem pangan diartikan sebagai segala sesuatu yang berhubungan dengan peraturan, pembinaan, dan pengawasan terhadap kegiatan atau produksi pangan dan peredaran pangan sampai dengan siap dikonsumsi oleh manusia. Sementara itu, ketahanan pangan diartikan sebagai kondisi terpenuhinya pangan bagi rumah tangga yang tercermin dari tersedianya pangan yang cukup, baik jumlah maupun mutunya, aman, merata, dan terjangkau. Pangan diartikan sebagai segala sesuatu yang bersumber dari sumber hayati dan air, baik yang diolah maupun yang tidak diolah. Pangan diperuntukkan bagi konsumsi manusia sebagai makanan atau minuman, termasuk bahan tambahan pangan, bahan baku pangan, dan bahan-bahan lain yang digunakan dalam proses penyiapan, pengolahan, dan pembuatan makanan atau minuman (Anonim, 2014).

Analisis kluster pertama kali digunakan oleh Tyron pada tahun 1939. Analisis kluster bertujuan untuk mengalokasikan sekelompok individu pada suatu kelompok-kelompok yang saling bebas sehingga individu-individu di dalam satu

kelompok yang sama mirip satu sama lain, sedangkan individu-individu di dalam kelompok yang berbeda tidak mirip (Rachmatin, 2014).

Terdapat beberapa metode pengelompokan dalam analisis kluster yaitu analisis kluster secara hierarki dan non hierarki, di dalam analisis kluster non hierarki penulis akan menerapkan metode kluster *K-Means*, metode ini dimulai dengan menentukan terlebih dahulu jumlah kluster yang diinginkan dan *centroid* di tiap kluster (Nugroho, 2008). *K-Means* merupakan salah satu metode kluster non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih kelompok, sehingga data yang memiliki karakteristik sama dikelompokkan ke dalam satu kelompok yang sama (Rismawan dan Kusumadewi, 2008).

Beberapa peneliti yang lebih dahulu membahas penelitian yang berkaitan dengan analisis Kluster *K-Means* diantaranya yaitu : Susanti (2012) melakukan penelitian dengan judul Pengklasteran Provinsi di Indonesia Berdasarkan Jumlah Penderita Penyakit Penyebab Kematian Tahun 2012 Dengan Menggunakan Metode *K-Means*. Safitri (2012) melakukan penelitian dengan judul Analisis *Cluster* Pada Kabupaten/Kota Di Jawa Tengah Berdasarkan Produksi Palawija. Rismawan dan Kusumadewi (2008) melakukan penelitian dengan judul Aplikasi *K-Means* Untuk Pengelompokan Mahasiswa Berdasarkan Nilai *Body Mass Index* (BMI) & Ukuran Kerangka.

Berdasarkan penjelasan sebelumnya penulis tertarik untuk melakukan penelitian dengan judul : Pengelompokan Provinsi Di Indonesia Berdasarkan Rata-Rata Produksi Tanaman Pangan Menggunakan Metode Kluster *K-Means*.

TINJAUAN PUSTAKA

2.1 Tanaman Pangan

Tanaman pangan adalah segala jenis tanaman yang dapat menghasilkan karbohidrat dan protein. Tanaman pangan terdiri dari padi, jagung, kedelai, kacang tanah, kacang hijau, ubi kayu, dan ubi jalar. Indonesia adalah negara agraris tempat tumbuh jenis tanaman pangan, dahulu indonesia pernah dikenal sebagai negara swasembada pangan (Hanum, 2008).

2.2 Analisis Kluster

Analisis kluster merupakan analisis yang digunakan untuk mengelompokkan pengamatan atau variabel menjadi beberapa kelompok pengamatan atau variabel yang jumlahnya lebih sedikit. Analisis kluster bertujuan untuk mengelompokkan n objek berdasarkan p variat yang memiliki kesamaan karakteristik diantara objek-objek tersebut. Objek tersebut akan diklasifikasikan ke dalam satu atau lebih kluster sehingga objek-objek yang berada dalam satu kluster akan mempunyai kemiripan atau kesamaan karakter (Yulianto, 2014).

2.3 Metode Kluster Non Hierarki

Metode ini dimulai dengan menentukan terlebih dahulu jumlah kluster yang diinginkan (dua kluster atau lebih). Setelah jumlah kluster ditentukan, baru proses kluster dilakukan tanpa mengikuti proses hierarki. Kelemahan pengklasteran dengan menggunakan metode ini adalah banyaknya kluster harus ditentukan sebelumnya dan pemilihan pusat kluster secara sembarang. Keunggulan pengklasteran dengan menggunakan metode ini lebih cepat daripada dengan pengklasteran dengan menggunakan metode hierarki dan lebih menguntungkan untuk jumlah objek/kasus yang besar (Rosiatun, 2010).

2.4 Proses Analisis Kluster

Langkah-langkah yang harus dilakukan untuk melakukan proses analisis kluster yaitu (Santoso, 2002) :

1. Mengukur kesamaan antar objek (*similarity*).
 1. Mengukur korelasi antar sepasang objek pada beberapa variabel.
 2. Mengukur jarak antara dua objek, pengukuran ada bermacam-macam dan yang paling populer adalah menggunakan metode *Euclidean Distance*.

2. Membuat kluster (dengan metode secara hierarki atau secara non hierarki).
3. Setelah kluster terbentuk, lakukan interpretasi terhadap kluster yang telah terbentuk, untuk menggambarkan isi kluster tersebut.
4. Melakukan validasi dan *profiling*.
Kluster yang terbentuk kemudian diuji apakah hasil pengklasteran tersebut valid, kemudian dilakukan proses *profiling* untuk menjelaskan karakteristik setiap kluster berdasarkan profil tertentu.

2.5 Kluster K-Means

Penggunaan pengklasteran *K-Means* untuk menjelaskan algoritma dalam

penentuan suatu obyek ke dalam kluster tertentu berdasarkan rataan terdekat. Asumsikan n adalah obyek dan p adalah variabel yang dinotasikan dengan $D(i, j)$ dimana $i = 1, 2, \dots, n$ dan $j = 1, 2, \dots, p$ dan dengan menggunakan jarak *euclid* antar obyek. Pengklasteran *K-Means* sangat cocok untuk data dengan ukuran yang besar karena memiliki kecepatan yang lebih tinggi dibandingkan metode hierarki. Namun, pemilihan banyaknya kluster dan *centroid* yang harus ditentukan lebih dahulu menjadi kelemahan metode ini (Nugroho, 2008).

K-Means merupakan metode kluster non hierarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih kelompok. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu kluster yang sama dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan kluster yang lain (Agusta, 2007).

2.6 Algoritma K-Means Kluster

Algoritma *K-Means* merupakan algoritma yang relatif sederhana untuk mengklasifikasikan atau mengelompokkan sejumlah besar obyek dengan atribut tertentu ke dalam kelompok-kelompok sebanyak K . Pada algoritma *K-Means*, jumlah kluster K sudah ditentukan lebih dahulu. Data clustering menggunakan metode *K-Means* ini secara umum dilakukan dengan algoritma dasar sebagai berikut (Agusta, 2007) :

1. Tentukan jumlah kluster k
2. Alokasikan data ke dalam kluster secara random
3. Hitung jarak setiap data yang ada terhadap setiap pusat kluster di masing-masing kluster menggunakan rumus *Euclidian Distance* (Nugroho, 2008) :

$$D(x_i, y_j) = \sqrt{(x_{i_1} - y_{j_1})^2 + (x_{i_2} - y_{j_2})^2 + \dots + (x_{i_p} - y_{j_p})^2} \quad (2.1)$$

$$D(x_i, y_j) = \sqrt{\sum_{l=1}^p (x_{i_l} - y_{j_l})^2} \quad (2.2)$$

Keterangan :

$D(x_i, y_j)$ = Jarak data ke- i ke pusat kluster ke- j
 x_{i_l} = Data ke- i pada variabel ke- l
 y_{j_l} = Titik pusat kluster ke- j pada variabel ke- l

4. Kelompokkan setiap data ke jarak terdekat pusatnya
5. Tentukan posisi pusat kluster baru C_{kj} dengan cara menghitung nilai rata-rata dari data-data yang ada pada kluster yang sama dengan rumus (Rivani, 2010) :

$$C_{kj} = \left(\frac{1}{n_k}\right) \sum_{l=1}^p d_{j_l} \quad (2.3)$$

Keterangan :

C_{kj} = Pusat kluster baru ke- j pada variabel ke- l
 n_k = Banyaknya anggota objek dalam kluster ke- k
 d_{j_l} = Data dalam kluster ke- k pada variabel ke- l

5. Jika pusat klaster tidak berubah lagi maka proses klaster selesai, atau kembali ke langkah 3 apabila masih ada data yang berpindah klaster.

2.7 Kelebihan dan Kekurangan *K-Means*

Kelebihan *K-Means* :

1. Algoritma *K-Means* ini relatif sederhana dan cepat.
2. Algoritma *K-Means* memberikan hasil yang relatif baik pada klaster.
3. Algoritma *K-Means* memiliki ketelitian yang cukup tinggi terhadap ukuran objek dan tidak berpengaruh pada urutan objek.

Kekurangan *K-Means* :

1. Menentukan jumlah klaster sebelum mengetahui jumlah klaster.
2. Hanya bekerja pada data numerik.
3. Sangat sensitif pada pembangkitan titik pusat awal secara random.

METODE PENELITIAN

3.1 Jenis Penelitian

Jenis penelitian ini adalah penelitian statistik terapan, penelitian ini dilakukan dengan menggunakan analisis Klaster *K-Means*.

3.2 Pengumpulan Data

Data yang digunakan pada penelitian ini adalah data sekunder, yang diperoleh dari Badan Pusat Statistik (BPS) yaitu data produksi tanaman pangan diseluruh Provinsi di Indonesia pada tahun 2015.

3.3 Variabel Penelitian

Objek yang digunakan dalam penelitian ini adalah Provinsi diseluruh Indonesia, dengan indikator penelitian adalah produksi tanaman pangan, dan variabel yang digunakan adalah sebagai berikut:

- X1 = Produksi Padi diseluruh Provinsi di Indonesia
- X2 = Produksi Jagung diseluruh Provinsi di Indonesia
- X3 = Produksi Kedelai diseluruh Provinsi di Indonesia
- X4 = Produksi Kacang Tanah diseluruh Provinsi di Indonesia
- X5 = Produksi Ubi Jalar diseluruh Provinsi di Indonesia

3.4 Metode Analisis

Metode analisis yang digunakan dalam penelitian adalah analisis Klaster *K-Means*, dengan mengambil data hasil produksi tanaman pangan (padi, jagung, kedelai, kacang tanah, ubi jalar) di Provinsi di Indonesia pada tahun 2015 yang diperoleh dari Badan Pusat Statistik. Pengolahan data dilakukan dengan bantuan *software Microsoft Office Excel 2007* dan *SPSS 16.0*.

3.5 Teknik Analisis Data

Dalam penelitian ini dianalisis melalui tahapan-tahapan sebagai berikut :

1. Menentukan jumlah klaster yang ingin dibentuk dengan metode *K-Means*
2. Memilih ukuran jarak antar objek yaitu dengan menggunakan jarak *Euclid*.
3. Melakukan proses pengklasteran, jika pada proses iterasi pengklasteran objek tidak ada yang berpindah maka proses iterasi pengklasteran selesai, jika masih ada objek yang berpindah maka lakukan proses iterasi pengklasteran.
4. Melakukan interpretasi pada klaster yang diperoleh.

BAB IV : HASIL DAN PEMBAHASAN

4.1 Deskripsi Data

Data yang digunakan dalam penelitian ini adalah data sekunder, yang diperoleh dari Badan Pusat Statistik (BPS) yaitu data produksi tanaman pangan diseluruh provinsi di Indonesia pada tahun 2015. Variabel data yang digunakan

dalam penelitian ini terdiri dari 5 variabel produksi tanaman pangan yaitu produksi padi, produksi jagung, produksi kedelai, produksi kacang tanah, dan produksi ubi jalar. Objek data yang digunakan dalam penelitian ini yaitu seluruh provinsi di Indonesia yang terdiri dari 33 provinsi.

4.2 Metode Pengklasteran *K-Means*

Metode pengklasteran *K-Means* memproses semua data secara sekaligus, dimana *K* adalah banyaknya klaster. Dengan menggunakan bantuan program aplikasi SPSS 16.0 semua objek telah diproses dan diperoleh bahwa tidak ada data yang hilang seperti tabel *number of cases in each cluster* di bawah ini (Lampiran 2) sehingga proses pengklasteran dapat dilanjutkan.

Tabel 1. Number of Cases in each Cluster

Cluster	1	5.000
	2	1.000
	3	4.000
	4	22.000
	5	2.000
Valid		34.000
Missing		.000

4.3 Pembentukan Klaster *K-Means*

Banyak klaster yang dibentuk peneliti pada proses pengklasteran dengan metode *K-Means* adalah sebanyak 5 klaster yaitu klaster 1, klaster 2, klaster 3, klaster 4, dan klaster 5. Dengan menggunakan bantuan program aplikasi SPSS 16.0, maka diperoleh :

Tabel 2. Initial Cluster Centers

	Cluster				
	1	2	3	4	5
Padi	2417392	13154967	5471806	959	11301422
Jagung	959973	6131163	1528414	473	3212391
Kedelai	125036	344998	67192	15	129794
KacangTanah	31142	191579	19024	153	109204
UbiJalar	19024	350516	71681	1795	151312

Tabel *Initial Cluster* di atas menunjukkan bahwa ada lima buah titik awal pusat klaster yang pertama kali terbentuk di setiap klaster, selanjutnya dengan metode klaster *K-Means* akan diuji dan dilakukan realokasi klaster yang ada (Lampiran 6, 7 dan 8). Proses tersebut dinamakan *Iteration*, yang memuat perubahan pada pusat klaster.

4.4 Analisis Pengelompokan Berdasarkan Produksi Tanaman Pangan

Dari lima variabel yang relevan, untuk membedakan isi klaster, dapat dianalisis provinsi yang termasuk klaster 1, klaster 2, klaster 3, klaster 4, dan klaster 5. Pengelompokan 34 provinsi adalah sebagai berikut:

Tabel 3. Hasil Pengelompokan Berdasarkan Produksi Tanaman Pangan

Klaster	Jumlah Anggota	Anggota
1	5	Aceh
		Sumatera Barat
		Banten
		Nusa Tenggara Barat
		Kalimantan Selatan
2	1	Jawa Timur
3	4	Sumatera Utara
		Sumatera Selatan
		Lampung
		Sulawesi Selatan
4	21	Riau
		Jambi
		Bengkulu
		Kepulauan Bangka Belitung
		Kepulauan Riau
		Daerah Istimewa Yogyakarta
		Bali
		Nusa Tenggara Timur
		Kalimantan Barat
		Kalimantan Tengah
		Kalimantan Timur
		Kalimantan Utara
		Sulawesi Utara
		Sulawesi Tengah
		Sulawesi Tenggara
		Gorontalo
		Sulawesi Barat
		Maluku
		Maluku Utara
Papua Barat		
Papua		
5	2	Jawa Barat
		Jawa Tengah

Berdasarkan tabel di atas terlihat bahwa pengelompokan 33 provinsi ke dalam lima klaster produksi tanaman pangan adalah sebagai berikut: Klaster 1 terdiri dari Provinsi Aceh, Sumatera Barat, Banten, Nusa Tenggara Barat, dan Kalimantan Selatan. Klaster 2 hanya terdiri oleh Provinsi Jawa Timur. Klaster 3 terdiri dari Provinsi Sumatera Utara, Sumatera Selatan, Lampung, dan Sulawesi Selatan. Klaster 4 terdiri dari Provinsi Riau, Jambi, Bengkulu, Kepulauan Bangka Belitung, Kepulauan Riau, Daerah Istimewa Yogyakarta, Bali, Nusa Tenggara

Timor, Kalimantan Barat, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, dan Papua. Klaster 5 terdiri dari Provinsi Jawa Barat dan Jawa Tengah.

4.5 Interpretasi Profil

Interpretasi profil klaster meliputi pengkajian mengenai *centroid* yaitu rata-rata nilai obyek yang terdapat dalam klaster pada tiap variabel (Nugroho, 2008). Interpretasi profil diperoleh dari menghitung nilai rata-rata obyek terhadap kelima klaster yang terbentuk dengan menganalisis variabel-variabel yang membedakan antara kelima klaster. Setelah diketahui variabel-variabel yang relevan untuk membedakan klaster 1, klaster 2, klaster 3, klaster 4, dan klaster 5, selanjutnya dianalisis rata-rata produksi masing-masing provinsi yang termasuk ke dalam masing-masing klaster yang dapat dilihat dari nilai akhir pusat klaster pada variabel yang membedakan.

Tabel 4. Final Cluster Centers

	Cluster				
	1	2	3	4	5
Padi	2325664	13154967	4351613	477130	11337328
Jagung	381604	6131163	1209851	116400	2086162
Kedelai	38225	344998	25094	4449	114366
KacangTanah	11952	191579	8631	5879	94962
UbiJalar	45389	350516	59838	38038	303744

Tabel *final cluster centers* diatas menunjukkan bahwa ada lima buah titik akhir pusat klaster yang terbentuk disetiap klaster. Karakteristik dalam setiap klaster ini berbeda, dapat dilihat dari nilai tabel *final cluster centers* di atas dari nilai satuan yang terkecil sampai nilai satuan yang terbesar disetiap variabel, maka berturut-turut adalah sebagai daerah produksi paling sedikit, sedikit, sedang, banyak, dan paling banyak untuk nilai satuan yang terbesar di setiap variabel.

Secara umum provinsi pada klaster 2 (Jawa Timur) memproduksi tanaman pangan dengan jumlah rata-rata produksi yang paling banyak dibandingkan dengan jumlah produksi tanaman pangan pada klaster lainnya

Provinsi pada klaster 5 (Jawa Barat dan Jawa Tengah) merupakan provinsi dengan jumlah rata-rata produksi tanaman pangan banyak untuk padi, jagung, kedelai, kacang tanah, dan ubi jalar. Provinsi pada klaster 3 (Sumatera Utara, Sumatera Selatan, Lampung, dan Sulawesi Selatan) merupakan provinsi dengan jumlah rata-rata produksi sedang untuk padi, jagung dan ubi jalar, sedangkan untuk jumlah rata-rata produksi kedelai dan kacang tanah sedikit. Provinsi pada klaster 1 (Aceh, Sumatera Barat, Banten, Nusa Tenggara Barat, dan Kalimantan Selatan) merupakan provinsi dengan jumlah rata-rata produksi sedikit untuk padi, jagung dan ubi jalar, sedangkan untuk jumlah rata-rata produksi kedelai dan kacang tanah sedang.

Provinsi pada klaster 4 (Riau, Jambi, Bengkulu, Kepulauan Bangka Belitung, Kepulauan Riau, Daerah Istimewa Yogyakarta, Bali, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, dan Papua) merupakan

provinsi dengan jumlah rata-rata produksi tanaman pangan sangat untuk padi, jagung, kedelai, kacang tanah, dan ubi jalar

BAB V : KESIMPULAN DAN SARAN

5.1 Kesimpulan

Pengelompokan Provinsi di Indonesia menggunakan analisis kluster *K-Means* berdasarkan rata-rata produksi tanaman pangan dikelompokkan menjadi 5 kluster, secara umum Provinsi pada kluster 2 (Jawa Timur) memproduksi tanaman pangan dengan jumlah rata-rata produksi yang paling banyak jika dibandingkan dengan jumlah produksi tanaman pangan pada kluster lainnya, sedangkan Provinsi pada kluster 4 (Riau, Jambi, Bengkulu, Kepulauan Bangka Belitung, Kepulauan Riau, Daerah Istimewa Yogyakarta, Bali, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, dan Papua) memproduksi tanaman pangan dengan jumlah rata-rata produksi yang paling sedikit.

5.2 Saran

Pada penelitian ini penulis menggunakan metode kluster *K-Means* untuk mengelompokkan provinsi di Indonesia berdasarkan rata-rata produksi tanaman pangan. Bagi peneliti selanjutnya yang tertarik dengan analisis kluster *K-Means* dapat menggunakan data dalam bidang lain, dan karenakan cakupan kluster sebenarnya sangat luas, maka bagi peneliti selanjutnya juga dapat menggunakan perbandingan kluster non hierarki dengan metode *K-Means* dengan kluster hierarki dengan metode lain.

DAFTAR PUSTAKA

- Agusta, Y. 2007. *K-Means – Penerapan, Permasalahan dan Metode Terkait. Jurnal Sistem dan Informatika*, Vol. 3 : 47-60.
- Anonim. 2014. Dinas Pertanian Tanaman Pangan dan Hortikultura Provinsi Kalimantan Selatan. <http://distantph.kalselprov.go.id/2014/02/03/mengenal-tanaman-pangan> diakses pada Sabtu, 06 Februari 2016.
- Hanum, C. 2008. *Teknik Budidaya Tanaman*, Jilid Dua. Penerbit Direktorat Pembinaan Sekolah Menengah Kejuruan. Jakarta.
- Martiana, E., N, R, Muhtada'i., dan E, Purnomo. 2011. Penggunaan Metode Pengklasteran Untuk Menentukan Bidang Tugas Akhir Mahasiswa Teknik Informatika. <http://repo.eepis-its.edu/130> diakses pada Sabtu, 06 Februari 2016.
- Nugroho, S. 2008. *Pengantar Statistika Matematika*, Edisi Pertama. UNIB Press.
- Nugroho, S. 2008. *Statistika Multivariat Terapan*, Edisi Pertama. UNIB Press.
- Rachmatin, D. 2014. Aplikasi Metode-Metode Agglomerative Dalam Analisis Kluster Pada Data Tingkat Polusi Udara. *Jurnal Ilmiah Program Studi Matematika STKIP Siliwangi Bandung*, Vol 3, No.2.
- Rismawan, T, dan Kusumadewi, S. 2008. Aplikasi K-Means Untuk Pengelompokan Mahasiswa Berdasarkan Nilai Body Mass Index (BMI) & Ukuran Kerangka. *Jurnal Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, ISSN: 1907-5022.
- Rivani, E. 2010. Aplikasi K-Means Cluster Untuk Pengelompokan Provinsi Berdasarkan Produksi Padi, Jagung, Kedelai, Dan Kacang Hijau Tahun 2009. *Jurnal Mat Stat*, Vol. 10 No. 2 : 122-134.
- Sa'adah, I, R., Supriyanta., dan Subejo. 2013. Keragaman Warna Gabah Dan Warna Beras Varietas Lokal Padi Beras Hitam (*Oryza Sativa L.*) Yang

- Dibudidayakan Oleh Petani Kabupaten Sleman, Bantul, Dan Magelang. *Jurnal Vegetalika*, Vol.2 No.3 : 13-20.
- Safitri, D. 2012. Analisis Cluster Pada Kabupaten/Kota Di Jawa Tengah Berdasarkan Produksi Palawija. *Jurnal Media Statistika*, Vol. 5, No. 1 : 11-16.
- Susanti, P. 2012. Pengklasteran Provinsi di Indonesia Berdasarkan Jumlah Penderita Penyakit Penyebab Kematian Tahun 2012 Dengan Menggunakan Metode K-Means. *Skripsi Universitas Bengkulu*.
- Yulianto, S. 2014. Analisis Klaster Untuk Pengelompokan Kabupaten/Kota Di Provinsi Jawa Tengah Berdasarkan Indikator Kesejahteraan Rakyat. *Jurnal Statistika*, Vol. 2, No. 1.