

**APLIKASI METODE *LINKAGE* DENGAN JARAK *EUCLID*, KUADRAT
EUCLID, *CITY-BLOCK* ATAU *MANHATTAN* DAN *CHEBYCHEV*
MENGUNAKAN PROGRAM R 2.14.**

(Studi Kasus: Hasil Produksi Pangan Di Indonesia Tahun 2013)

Masni Uli Br Purba¹, Sigit Nugroho², Jose Rizal²

1 Alumni Jurusan Matematika Fakultas MIPA Universitas Bengkulu

2 Staf Pengajar Jurusan Matematika Fakultas MIPA Universitas Bengkulu

Email : masniuli_b@yahoo.com

ABSTRACT

Penelitian ini bertujuan untuk mengelompokkan 34 Propinsi di Indonesia berdasarkan Data Hasil Produksi Pangan di Indonesia pada Tahun 2013 dan mengkaji secara statistik peran jarak dan metode pengklasteran pada pengklasteran Hirarki. Data Hasil Produksi Pangan di Indonesia pada Tahun 2013 diperoleh dari Badan Pusat Statistik (BPS). Dalam penelitian ini digunakan analisis kluster dengan metode hierarki (*Single linkage*, *Complete Linkage*, *Average Linkage*) dengan jarak *Euclid*, Kuadrat *Euclid*, *City-Block* atau *Manhattan* dan *Chebychev*. Hasil penelitian menunjukkan bahwa jika ditentukan tiga kluster yang terbentuk dengan metode pengklasteran yang berbeda, jarak *City-Block* atau *Manhattan* akan menghasilkan pengklasteran yang sama. Dalam kasus Data Hasil Produksi Pangan di Indonesia Tahun 2013, jarak *City-Block* atau *Manhattan* merupakan jarak yang paling baik.

Kata kunci: Analisis Kluster, Metode Hierarki, *Linkage*, Jarak (*Distance*)

1. PENDAHULUAN

Analisis Kluster adalah analisis statistika yang bertujuan untuk mengklasterkan data sedemikian sehingga data yang berada dalam kluster yang sama mempunyai sifat yang relatif homogen daripada data yang berada dalam kluster yang berbeda (Purnamasari, 2011). Terdapat beberapa metode pengklasteran dalam Analisis Kluster, yaitu metode Hierarki dan metode Non-Hierarki (pengklasteran *K-Means*). Pada metode Hierarki terdapat dua tipe dasar kluster yaitu *Agglomerative* (penggabungan) dan *Divisive* (pemecahan). Pengklasteran *Agglomerative* (penggabungan) terdiri dari tiga metode yaitu *Linkage*, *Ward's* dan *Centroid* (Nugroho, 2008). Pada analisis kluster ukuran kemiripan atau ketidakmiripan yang digunakan adalah jarak (*distance*). Menurut Nugroho

(2008), terdapat beberapa metode untuk mengukur jarak antara dua objek dengan ukuran kedekatan yang berbeda yaitu *Euclid*, Kuadrat *Euclid*, *City-Block* atau *Manhattan*, dan *Chebychev*.

Dalam penelitian ini akan dipelajari tentang pengklasteran hierarki khususnya metode *Linkage* (*Single Linkage*, *Complete Linkage* dan *Average Linkage*) menggunakan empat jarak yang berbeda yaitu *Euclid*, Kuadrat *Euclid*, *City-Block* atau *Manhattan* dan *Chebychev* dengan menggunakan Data Hasil Produksi Pangan Tahun 2013 di Propinsi yang ada di Indonesia.

2. LANDASAN TEORI

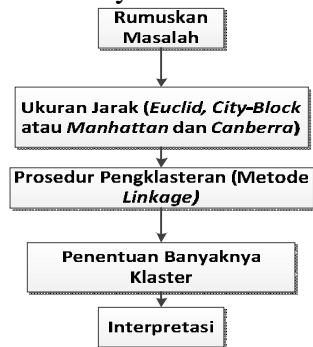
2.1 Analisis Kluster

Analisis kluster merupakan salah satu analisis multivariat yang termasuk dalam metode interdependensi yaitu variabel bebas x atau faktor penyebab

tidak dibedakan dengan variabel terikat y atau respon (Nugroho, 2008). Tujuan analisis kluster adalah mengelompokkan objek-objek berdasarkan kesamaan karakteristik diantara objek-objek tersebut. Objek dapat berupa produk, benda, serta orang. Objek tersebut akan diklasifikasikan ke dalam satu atau lebih kluster sehingga objek-objek yang berada dalam satu kluster akan mempunyai kemiripan satu dengan yang lain (Santoso, 2004). Dalam analisis kluster terdapat beberapa asumsi yang harus dipenuhi yaitu sampel yang diambil benar-benar dapat mewakili populasi yang ada dan tidak multikolinieritas.

2.2 Proses Analisis Kluster

Menurut Supranto (2010) untuk menentukan banyaknya kluster yang akan terbentuk, perlu dilakukan proses analisis kluster. Beberapa tahap yang harus dilakukan yakni:



Gambar 1. Proses Analisis Kluster

2.3 Pengukuran Konsep Kemiripan Antar Objek

Menurut Nugroho (2008) pengukuran jarak untuk Mengukur jarak (*distance*) antara dua objek terdiri dari jarak *Euclid*, *Kuadrat Euclid*, *City-Block* atau *Manhattan*, dan *Chebychev*. Pada penelitian ini jarak yang akan digunakan adalah *Euclid*, *Kuadrat Euclid*, *City-Block* atau *Manhattan* dan *Chebychev* dengan persamaan yang dimiliki masing-masing jarak sebagai berikut:

1. Jarak *Euclid*

$$d(x, y) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2}$$

2. Jarak kuadrat *Euclid*

$$d(x, y) = \sum_{i=1}^p (x_i - y_i)^2$$

3. Jarak *City-Block* atau *Manhattan*

$$d(x, y) = \sum_{i=1}^p |x_i - y_i|$$

4. Jarak *Chebychev*

$$d(x, y) = \max |x_i - y_i|$$

2.4 Metode Hierarki

Metode Hierarki adalah metode yang memulai pengelompokkan dengan dua atau lebih objek yang mempunyai kesamaan paling dekat. Menurut Nugroho (2008), metode hierarki lebih populer dibandingkan metode pengklasteran *K-means* (*Non-Hierarki*) namun metode ini membutuhkan waktu lama jika datanya besar. Dalam Metode hierarki terdapat dua pengklasteran yaitu *Agglomerative* (*pengabungan*) dan *Devisive* (*pemecahan*). Berdasarkan ukuran kemiripannya metode *Agglomerative* terdiri dari beberapa metode yaitu metode *Linkage*, metode *Ward*, metode *Centroid* dan metode *Median* (Nugroho, 2008). Metode yang akan dibahas dalam penelitian adalah metode *Linkage* (*Single linkage*, *Complete Linkage*, dan *Average Linkage*).

2.4.1 *Single Linkage*

Menurut Supranto (2010), metode *single linkage* adalah proses pengklasteran yang didasarkan pada jarak terdekat antar objek pada matriks *proximity* dan proses pengabungan objek-objek dilakukan berdasarkan objek-objek yang memiliki jarak terkecil (*minimum*) tersebut, misal U dan V , sehingga diperoleh kluster

(UV). Untuk mencari kluster antara (UV) dan kluster W lainnya dapat diperoleh dengan cara sebagai berikut:

$$d_{(UV)W} = \min\{d_{UW}, d_{VW}\}$$

d_{UW} adalah anggota jarak dari kluster U ke W dan d_{VW} adalah anggota jarak dari kluster V ke W.

2.4.2 Complete Linkage

Menurut Rachmatin (2014), pada metode *complete linkage* proses pengklasteran dimulai dengan menentukan jarak terdekat pada matriks *proximity* dan proses penggabungan objek-objek dilakukan berdasarkan objek-objek yang memiliki jarak terjauh (maksimum) tersebut, misal U dan V, sehingga diperoleh kluster (UV). Selanjutnya jarak antara (UV) dan kluster W lainnya dapat diperoleh dengan cara sebagai berikut:

$$d_{(UV)W} = \max\{d_{UW}, d_{VW}\}$$

d_{UW} adalah anggota jarak dari kluster U ke W dan d_{VW} adalah anggota jarak dari kluster V ke W.

2.4.3 Average Linkage

Menurut Rachmatin (2014), metode *average linkage* adalah metode pengklasteran yang didasarkan pada jarak rata-rata antar objek. Menurut Supranto (2010), metode *average linkage* merupakan gabungan dari *single linkage* dan *complete linkage*. Pada metode *average linkage* proses pengklasteran dimulai dengan menentukan jarak terdekat dalam matriks *proximity* dan proses penggabungan objek-objek dilakukan berdasarkan nilai rata-rata jarak dari objek-objek tersebut, misal U dan V, sehingga diperoleh kluster (UV). Selanjutnya jarak antara (UV) dan kluster W lainnya dapat diperoleh dengan cara sebagai berikut:

$$d_{(UV),W} = \frac{[\sum_{i=1}^n \sum_{k=1}^k d_{ik}]}{N_{(UV)}N_{(W)}}$$

dimana d_{ik} adalah jarak antara objek i dalam kelompok (UV) dan objek k

dalam kelompok W, dan $N_{(UV)}$ dan N_w berturut-turut adalah banyaknya objek-objek dalam kelompok (UV) dan W.

2.5 Interpretasi

Menurut Supranto (2010), interpretasi meliputi pengkajian mengenai centroids yaitu rata-rata nilai objek yang terdapat dalam kluster pada setiap variabel. Nilai centroid memungkinkan kita untuk menguraikan setiap kluster dengan cara memberikan suatu nama atau label.

3. Metode penelitian

Penelitian ini merupakan penelitian statistik terapan, yaitu suatu penelitian yang dilakukan dengan mengaplikasikan metode statistika ke dalam bidang-bidang tertentu. Pada penelitian ini diterapkan metode Analisis Kluster yaitu metode *Linkage* menggunakan tiga jarak berbeda yaitu jarak *Euclid*, Kuadrat *Euclid*, *City-Block* atau *Manhattan* dan *Chebyshev*. Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) yang telah dipublikasikan yaitu data hasil produksi pangan periode tahun 2013.

Variabel yang digunakan dalam penelitian ini adalah X_1 (Hasil produksi pangan jenis tanaman Padi dalam satuan ton), X_2 (Hasil produksi pangan jenis tanaman Jagung dalam satuan ton), X_3 (Hasil produksi pangan jenis tanaman Kedelai dalam satuan ton), X_4 (Hasil produksi pangan jenis tanaman Kacang Tanah dalam satuan ton), X_5 (Hasil produksi pangan jenis tanaman Kacang Hijau dalam satuan ton), X_6 (Hasil produksi pangan jenis tanaman Ubi Jalar dalam satuan ton), dan X_7 (Hasil produksi pangan jenis tanaman Ubi Kayu dalam satuan ton).

Tahapan dalam analisis dari penelitian ini adalah pertama melakukan uji asumsi yang harus dipenuhi dalam analisis kluster yaitu sampel yang diambil benar-benar dapat

mewakili populasi yang ada dan tidak multikolinieritas. Kedua mengeluarkan satu atau beberapa variabel yang mengalami masalah multikolinieritas untuk mengatasi masalah multikolinieritas pada data. Ketiga memilih ukuran jarak antar objek dengan menggunakan jarak *Euclid*, *Kuadrat Euclid*, *City-Block* atau *Manhattan* dan *Chebychev*. Keempat melakukan proses pengklasteran akan dilakukan pengelompokan data dengan mengolah data menggunakan metode *Linkage*, dan kelima melakukan interpretasi terhadap klaster yang telah terbentuk dengan cara memberi nama untuk menggambarkan karakteristik masing-masing klaster.

4. HASIL PENELITIAN

4.1 Asumsi Analisis Klaster

Menurut Santoso (2010), ada beberapa asumsi yang harus dipenuhi yaitu sampel yang diambil benar-benar dapat mewakili populasi yang ada dan tidak multikolinieritas. Penelitian ini menggunakan populasi sebagai total sampling. Dengan demikian asumsi pertama telah terpenuhi. Asumsi kedua yaitu tidak mengalami masalah Multikolinieritas. Pada penelitian ini untuk mengetahui ada tidaknya multikolinieritas, peneliti menggunakan uji korelasi dan uji *Variance Inflation Factor* (VIF) untuk mendeteksi masalah multikolinieritas. Berdasarkan uji korelasi dan uji *Variance Inflation Factor* (VIF) dari output program R menunjukkan bahwa variabel jagung dan kacang tanah jagung mengalami masalah multikolinieritas, untuk mengatasi masalah tersebut maka dapat diatasi dengan cara mengeluarkan satu atau beberapa variabel yang mengalami masalah multikolinieritas (Gujarati, 2004) maka dari itu variabel jagung dan kacang tanah harus dikeluarkan dari tahap analisis maka dari itu variabel padi, kedelai, kacang hijau, ubi kayu, dan ubi jalar dapat digunakan untuk tahap analisis selanjutnya.



Gambar 2. Flowchart (Diagram Alir)

4.2 Proses dan Hasil Pengklasteran

Proses pengklasteran hirarki dilakukan dengan menggunakan *software* R versi 2.14.1. Metode pengklasteran hirarki dilakukan dengan memperhatikan ukuran kemiripan antar objek yang terdapat dalam matriks *proximity*. Metode pautan (*Linkage Methods*) terdiri dari pautan tunggal (*Single Linkage*), pautan lengkap (*Complete Linkage*), dan pautan rata-rata (*Average Linkage*). Pada metode ini tidak ada sumber yang mengharuskan menggunakan jarak tertentu untuk mendapatkan matriks *proximity*nya.

Untuk itulah digunakan metode jarak *Euclid*, Kuadrat *Euclid*, *City-Block* atau *Manhattan* dan *Chebychev*. Dari analisis yang dilakukan diperoleh hasil pengklasteran dengan metode pengukuran jarak *City-Block* atau *Manhattan* memberikan hasil pengklasteran yang sama persis untuk metode *Linkage*. Sedangkan pengukuran jarak *Euclid*, Kuadrat *Euclid*, dan *Chebychev* menghasilkan pengklasteran yang berbeda, namun perbedaannya tidak terlalu mencolok.

Output metode pengklasteran metode *Linkage* disajikan secara visual berbentuk dendrogram yaitu suatu bagan yang menyajikan banyaknya kluster terbesar hingga terkecil. Dendrogram dibaca dari kiri ke kanan. Garis tegak lurus menunjukkan objek yang digabung membentuk satu kluster. Sedangkan posisi garis pada skala menunjukkan jarak kluster yang digabung. Tabel Agglomerasi di bawah ini menunjukkan cara penggabungan kluster setiap tahap pada metode pautan tunggal (*Single Linkage*) dengan menggunakan jarak *Euclid*.

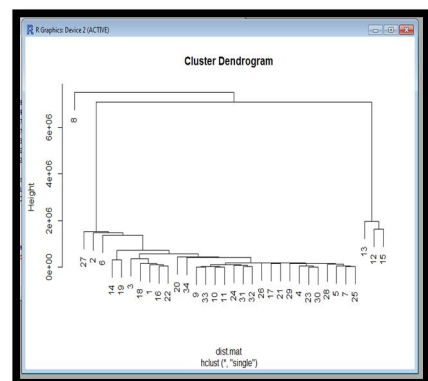
Tabel 2. Proses Agglomerative metode *Single Linkage* untuk jarak *Euclid*

Step	Clusters	Joined	Distance Level
1	23	30	1,19E+16
2	9	33	3,65E+15
3	10	11	2,51E+16
.	.	.	.
.	.	.	.
.	.	.	.
32	29	31	5,60E+19
33	32	8	5,60E+19

Pada Tahap Pertama proses pengabungan dilakukan terhadap Provinsi Kalimantan Timur (23) dan Provinsi Sulawesi Barat (30) dengan nilai jarak yang dihasilkan berdasarkan jarak *Euclid* adalah 109303,456 sehingga terbentuk kluster pada tahap pertama. Pada tahap kedua proses pengabungan dilakukan terhadap

Provinsi KEP. Bangka Belitung (9) dan Provinsi Papua Barat (33) dengan nilai jarak yang dihasilkan berdasarkan jarak *Euclid* adalah 60450,84 maka terbentuk kluster kedua dan begitu seterusnya sampai tahap ke 33.

Pada dendrogram metode pautan tunggal (*Single Linkage*) dengan jarak *Euclid*, kluster yang dapat dibentuk adalah 2 kluster, 3 kluster, 4 kluster dan seterusnya, dimana jika diambil 2 kluster maka anggota kluster tersebut adalah objek (Lampung) dan 33 objek lainnya. Jika diambil 3 kluster maka anggotanya adalah kluster dengan objek (Lampung), objek (Jawa Timur, Jawa Barat dan Jawa Tengah), dan 30 objek lainnya. Jika diambil 4 kluster maka anggotanya adalah kluster dengan objek (Lampung), objek (Jawa Tengah, Jawa Timur dan Jawa Barat), objek (Sulawesi Selatan, Sumatera Utara, Sumatera Selatan) dan 27 objek lainnya.



Gambar 3. Dendrogram Metode *Single Linkage* menggunakan Jarak *Euclid*

Jika yang diambil 3 kluster sebagai pemecahannya, maka:

- Untuk semua konsep pengukuran jarak pada metode *Complete Linkage* dan *Complete Linkage* didapat anggota pengklasteran sebagai berikut: (Jawa Timur, Jawa Barat dan Jawa Tengah), (Lampung) dan 30 objek lainnya.
- Metode *Average Linkage*, *Complete Linkage*, dan *Single Linkage*, berdasarkan konsep

pengukuran Jarak *City-Block* atau *Manhattan* anggota klasternya adalah (Jawa Timur, Jawa Barat dan Jawa Tengah), (Lampung) dan 30 objek lainnya.

- Metode *Single Linkage* dengan konsep pengukuran jarak *Euclid*, Kuadrat *Euclid* dan *Chebychev* anggota klasternya adalah objek (Lampung), objek (Jawa Timur, Jawa Barat dan Jawa Tengah), dan 30 objek lainnya.

4.3 Interpretasi

Interpretasi meliputi pengkajian mengenai *centroid* yaitu rata-rata nilai objek yang terdapat dalam kluster pada tiap variabel. Nilai *centroid* memungkinkan untuk menguraikan setiap kluster dengan cara memberi suatu label atau nama. Label suatu kluster juga dapat didasarkan pada manfaat yang akan dicari (Nugroho, 2008).

Untuk metode *Single Linkage* menggunakan Jarak *Euclid Kuadrat Euclid* dan *Chebychev*, kluster pertama dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Cukup Baik, kluster kedua dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Kurang Baik dan kluster ketiga dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Baik.

Untuk metode *Linkage* menggunakan Jarak *City-Block* atau *Manhattan*, kluster pertama dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Baik, kluster kedua dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Cukup Baik, kluster ketiga dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Kurang Baik.

Untuk Metode *Complete Linkage* dan *Average Linkage Euclid Kuadrat Euclid City-Block* atau *Manhattan* dan *Chebychev*, kluster pertama dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Baik, kluster kedua dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Cukup Baik, kluster ketiga dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Kurang Baik.

5. Kesimpulan dan Saran

Pengklastran pada data Hasil Produksi Pangan di Indonesia tahun 2013 menghasilkan matriks *proximity* dan proses *agglomerative* (penggabungan) yang berbeda untuk setiap konsep pengukuran dan metode pengklastran yang digunakan. Metode *Single Linkage* memiliki anggota pengklastran dan interpretasi yang berbeda untuk ketiga jarak yang digunakan. Sedangkan pada metode *Complete Linkage* dan *Average Linkage* untuk jarak *Euclid*, Kuadrat *Euclid*, *City-Block* atau *Manhattan* dan *Chebychev* memiliki anggota pengklastran dan interpretasi yang sama. Dalam kasus Data Hasil Produksi Pangan di Indonesia Tahun 2013, jarak *City-Block* atau *Manhattan* merupakan jarak yang paling baik karena memiliki anggota pengklastran, interpretasi dan dendogram yang sama pada setiap metode pengklastran *Linkage*, disarankan untuk melanjutkan penelitian ini pada metode pengklastran nonhirarki.

DAFTAR PUSTAKA

- Anonim. 2007. *Proyeksi Penduduk 2000-2025*. www.bps.go.id. Indonesia.
- Admaja, S. L. 2009. *Statistika Untuk Bisnis dan Ekonomi*. Andy Yogyakarta. Yogyakarta.
- Afriana, M. 2014. *Penentuan Awal Keanggotaan Klasteran Hirarki K-Means*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Apreity, H. 2015. *Pengklasteran Dengan Metode Average Linkage*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Carsono, N. 2009. Peran Pemuliaan dalam Meningkatkan Produksi Pertanian di Indonesia. *Faperta UNPAD*. Jatinagor.
- Dewi, S. Y. 2010. OLS, LASSO dan PLS Pada Data Mengandung Multikolinieritas. *Jurnal Ilmu Dasar* **Vol. 11** No. 1, Januari 2010 :83-91.
- Ferawati, L. 2006. *Pengklasteran Provinsi Diindonesia Menurut Status Daerah Rawan Pangan Berdasarkan Indikator Yang Mempengaruhi Tahun 2006*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Gujarati, D. 2010. *Ekonometri Dasar*. Selemba Empat. Jakarta.
- Hanani, N. 2012. Strategi Pencapaian Ketahanan Pangan Keluarga. *E-Journal Ekonomi Pertanian* **Vol. 1**, No.1.
- Izenman, A.J. 2008. *Modern Multivariate Statistical Techniques*. Springer Texts in Statistics: Philadelphia.
- Johnson, R.A. dan Wichern, D.W. 2002. *Applied Multivariate Statistical Analysis*. Prentice Hall International, New Jersey.
- Lestari, P. T. 2012. *Analisis Kluster Sekolah Menengah Atas (Sma) Berdasarkan Jumlah Tenaga Guru, Guru Sertifikasi, Pengawai Dan Siswa*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Lubis, A. I. 2008. *Kajian Metode Pengklasteran Hirarki Dengan Berbagai Jarak*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Mangkuatmodjo, S. 2004. *Statistika Lanjutan*. Rineka Cipta. Jakarta
- Marcus, G. L, Wattimanela, H. J. dan Lesnussa, Y. A. 2012. Analisis Regresi Komponen Utama Untuk Mengatasi Multikolinieritas Dalam Analisis Regresi Linier Berganda. *Jurnal Berekeng* **Vol. 6** No. 1 Hal. 31-40 (2012).
- Mattjik, A. A dan I. M. Sumertajaya. 2011. *Sidik Peubah Ganda Dengan menggunakan SAS*. IPB: Bandung.
- Mulyana, A. 2012. Pengutan Ketahanan Pangan Untuk Menekan Jumlah Penduduk Miskin dan Rentan Pangan di Tingkat Nasional Dan Regional. *E-Journal Ekonomi Pertanian* **Vol. 1**, No.1.
- Nugroho, S. 2008. *Statistika Multivariat Terapan*. UNIB Press. Bengkulu.
- Oktarina. 2014. *Metode Linkage Pada Persentase Kelahiran Balita Menurut Penolong Kelahiran Terakhir*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.

- Prabowo, R. 2010. *Kebijakan Pemerintah dalam Mewujudkan Ketahanan Pangan di Indonesia. Jurnal Pertanian* **Vol. 6**, No.2.
- Purnamasari, S. M. 2011. Analisis Kelompok. *Makalah II2092 Probabilitas dan Statistik*. ITB.
- 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, September 2014.
- Rencher, A.C., 2002. *Methods Of Multivariate Analysis Second Edition*. Wiley Series In Probability and Mathematical Statistics: Canada.
- Rosiatun, A. Widiariyah, T & Safitri, D. 2010. Analisis Kluster Untuk Segmentasi Pemirsya Program Berita Sore Stasiun TV Swasta. *Media Statistika*, **vol. 3**, No. 2, Desember 2010: 93-102.
- Santoso, S. 2010. *Statistika Multivariat Konsep & Aplikasi dengan SPSS*. Jakarta: PT. Elex Media Komputindo.
- Saputra, D. E. 2014. *Pengklasteran Hirarki Untuk Daerah Produksi Ikan Di Provinsi Bengkulu Pada Setiap Tempat Penangkapan Ikan Berdasarkan Jenis Alat Tangkapnya*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Sarwono, J. 2013. *Statistik Multivariat Aplikasi untuk Riset Skripsi*. Andy Yogyakarta. Yogyakarta.
- Sari, M. N. 2012. Analisis Kluster Hirarki Dalam Pengelompokan Panti Sosial Diprovinsi Bengkulu. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Simatupang, J. T. 2006. Pengembangan dan Aplikasi dalam Pembangunan Pertanian di Indonesia. *Jurnal Penelitian Bidang Ilmu Pertanian* **Vol. 4**, No.1.
- Suhartono. 2008. Analisis Data Statistik Dengan R. Lab Statistik Komputasi ITS. Surabaya.
- Supranto, J. 2010. *Analisis Multivariat Arti dan Interpretasi*. Cetakan Kedua. Rineka Cipta. Jakarta.
- Suryana, N. 2011. Penggunaan Metode Statistik K-Means Clustering Pada Analisis Peruntukan Lahan Usaha Tambang Berbasis Sistem Informasi Geografi. *Statistika*, **Vol. 11**. No. 1, Mei 2011.
- Susanti, P. 2015. Pengklasteran Provinsi Diindonesia Berdasarkan Jumlah Penderita Penyakit Penyebab Kematian Tahun 2012 Dengan Menggunakan Metode K-Means. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Yuliana, W. N, Sukarsa, G. K. I. dan Srinadi, M. A. G. I. 2013. Perbandingan Regresi Komponen Utama Dan Robpca Dalam Mengatasi Multikolinieritas Dan Pencilan Pada Regresi Linier Berganda. *E-Jurnal Matematika* **Vol. 2**, No. 4, Nopember 2013, 1-5.
- Yulianto, S & Hidayatullah, H. K. 2014. Analisis Kluster Untuk Pengelompokan Kabupaten/Kota Diprovinsi Jawa Tengah Berdasarkan Indikator Kesejahteraan Rakyat. *Statistika*, **Vol. 2**, No.1, Mei 2014