

46-2011-Analisis klaster

by Irmeilyana Irmeilyana

Submission date: 08-May-2023 11:24AM (UTC+0700)

Submission ID: 2087160022

File name: 46.Analisis_Cluster_terhadap_Tingkat_Pencemaran_-2011.pdf (190.41K)

Word count: 3749

Character count: 21050

Analisis Cluster terhadap Tingkat Pencemaran Udara pada Sektor Industri di Sumatera Selatan

ROBINSON SITEPU, IRMEILYANA, DAN BERRY GULTOM

Jurusan Matematika, Universitas Sriwijaya, Sumatera Selatan, Indonesia

15
INTISARI: Analisis cluster merupakan suatu teknik yang dipergunakan untuk mengklasifikasikan objek ke dalam kelompok yang relatif homogen. Analisis cluster terbagi atas dua metode, hirarki dan non-hirarki. Penelitian ini dibahas pengelompokan 10 jenis industri yang ada di Sumatera Selatan berdasarkan jenis polutan yang dihasilkan dan mengetahui ciri-ciri dari setiap kelompok industri. Berdasarkan hasil analisis cluster metode hirarki, ada 3 kelompok industri (cluster). Cluster pertama yaitu industri industri karet, industri sawit, industri pengalengan ikan, industri listrik, industri pertambangan, dan industri semen. Cluster kedua terdiri dari industri migas, industri minyak goreng, dan industri makanan. Cluster ketiga yaitu industri pupuk. Pada metode non-hirarki, cluster pertama yaitu industri yang memiliki rata-rata polutan yang lebih besar dari pada cluster kedua, dapat digolongkan menjadi kelompok industri dengan tingkat pencemaran tinggi. Anggotanya adalah industri migas, industri minyak goreng, industri makanan, dan industri pupuk. Cluster kedua yaitu industri yang memiliki rata-rata polutan yang lebih kecil dari pada cluster pertama, dapat digolongkan menjadi kelompok industri dengan tingkat pencemaran rendah. Anggotanya adalah industri karet, industri sawit, industri pengalengan ikan, industri listrik, industri pertambangan, dan industri semen.

KATA KUNCI: analisis kluster, pencemaran, industri

ABSTRACT: Cluster analysis is a technique used to classify objects into relatively homogeneous groups. Cluster analysis is divided into two methods, hierarchical and non-hierarchical. This research covered 10 types of industrial clustering in the southern Sumatran based on the type of pollutants generated and knowing the characteristics of each industry group. Based on the results of hierarchical cluster analysis method, there are three industry groups (clusters). The first cluster of industrial rubber industry, oil industry, fish canning industry, power industry, mining industry and cement industry. The second cluster consists of oil and gas industry, oil industry, and food industries. The third cluster is the fertilizer industry. In non-hierarchical method, the first clusters of industrial pollutants have on average larger than in the two clusters, can be classified into groups of industries with high pollution levels. The members are oil and gas industry, oil industry, food industry, and fertilizer industries. The second cluster of industrial pollutants have on average smaller than the first cluster, can be classified into groups of industries with low pollution levels. Its members are the rubber industry, oil industry, fish canning industry, power industry, mining industry and cement industry

KEYWORDS: cluster analysis, pollution, industry

Juli 2011

1 PENDAHULUAN

Industri menempati posisi yang sentral dalam ekonomi masyarakat modern dan merupakan penggerak paling dasar dalam peningkatan kemakmuran, terutama pada negara-negara maju. Sedangkan pada negara berkembang, industri sangat penting untuk memperluas landasan pembangunan dan memenuhi kebutuhan masyarakat yang semakin meningkat.

Pembangunan pada sektor industri membawa dampak bagi kehidupan manusianya, baik dampak positif maupun dampak negatif. Dampak positif dari kegiatan industri adalah dapat mengurangi persentase pengangguran, menambah devisa negara, serta

menarik minat investor untuk menanamkan modal. Selain dampak positif ini, kegiatan industri juga memiliki dampak negatif, **19** adapun salah satunya adalah pencemaran udara. Pencemaran udara dirasakan semakin hari semakin meningkat, terutama di daerah yang kepadatan lalu lintasnya cukup tinggi, serta di lokasi industri yang kurang memperhatikan dampak yang buruk pada lingkungannya. Udara yang tercemar dapat merusak lingkungan sekitarnya dan berpotensi mengganggu kesehatan masyarakat sekitarnya.

Dalam analisis statistik, kriteria polutan tersebut dinyatakan sebagai variabel. Semakin banyak kriteria polutan akan semakin rumit analisis statistika yang

akan dilakukan. Analisis¹⁴ multivariat merupakan salah satu metode statistika yang cocok untuk meringkas data dengan peubah banyak. Salah satu analisis multivariat yang dapat digunakan untuk memahami dan mempermudah interpretasi data adalah analisis cluster¹⁵.

Analisis cluster bertujuan untuk mengelompokkan objek-objek berdasarkan karakteristik²⁰ diantara objek-objek itu, sehingga dapat diketahui ciri khas dari tiap kelompok. Banyak objek yang dapat dikelompokkan dengan analisis cluster, diantaranya adalah produk (barang dan jasa), benda, manusia (responden konsumen)²¹.

2 TINJAUAN PUSTAKA

2.1 Analisis Multivariat

Analisis multivariat merupakan analisis yang meneliti masalah yang bersifat multidimensional dan menggunakan tiga atau lebih variabel. Analisis multivariat dibagi menjadi dua kategori utama, yakni:

1. Dependence Method.

Analisis ketergantungan digunakan apabila tujuan dari analisis adalah menjelaskan atau memprediksi variabel terikat berdasarkan dua atau lebih variabel bebas. Metode ini terdiri atas empat jenis yaitu:

- Analisis Regresi Berganda (*Multiple Regression Analysis*)
- Analisis Diskriminan Berganda (*Multiple Discriminant Analysis*)
- Analisis Multivariat Varians (*Multivariate analysis of Variance*)
- Analisis Korelasi Kanonikal (*Canonical Correlations Analysis*)

2. Metode Saling Ketergantungan (*Interdependence Method*)

Metode yang digunakan untuk menjelaskan seperangkat variabel atau pengelompokan berdasarkan variabel-variabel tertentu. Metode ini dikelompokkan menjadi tiga yaitu;

- Analisis Faktor (*Factor Analysis*)
- Analisis Cluster (*Cluster Analysis*)
- Skala Multidimensional (*Multidimensional Scalling*)

2.2 Analisis Cluster

Analisis cluster adalah analisis untuk mengelompokkan elemen yang mirip sebagai objek penelitian untuk menjadi kelompok (*cluster*) yang berbeda dan *mutually exclusive*²².

Analisis cluster termasuk dalam analisis statistik multivariat metode interdependen, dan oleh karena itu tujuan analisis cluster tidak untuk menghubungkan ataupun membedakan dengan sampel/variabel lain. Analisis cluster berguna untuk meringkas data dengan jalan mengelompokkan objek-objek berdasarkan kesamaan karakteristik tertentu diantara objek-objek yang akan diteliti. Analisis cluster terbagi atas 2 metode, yaitu:

1. Metode hirarki
2. Metode non-hirarki

Pada metode hirarki ini dimulai dengan mengelompokkan dua atau lebih objek yang mempunyai kesamaan yang paling dekat. Kemudian proses diteruskan ke objek lain yang mempunyai kedekatan kedua. Demikian seterusnya sampai cluster akan membentuk semacam "pohon" hierarki (tingkatan) yang jelas antar objek dari yang paling mirip sampai yang paling tidak mirip. Deondogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut.

Pada metode non-hirarki, digunakan jarak *Euclidian*, untuk menetapkan nilai kedekatan antara objek. Bakal cluster pertama adalah observasi pertama dalam set data. Bakal kedua adalah observasi lengkap berikutnya yang dipisahkan dari bakal pertama oleh jarak minimum khusus.

Analisis cluster adalah suatu alat untuk mengelompokkan sejumlah n objek berdasarkan p variabel yang secara relatif mempunyai kesamaan karakteristik diantara objek-objek tersebut, sehingga keragaman dalam suatu kelompok tersebut lebih kecil dibandingkan dengan keragaman antar kelompok. Objeknya dapat berupa barang, jasa, hewan, manusia (responden, konsumen, atau yang lain). Objek tersebut akan diklasifikasikan dalam satu atau lebih cluster (kelompok) sehingga objek-objek yang berada dalam satu cluster akan mempunyai kemiripan atau kesamaan karakter.

Jika terdapat n objek dan p variabel, maka observasi x_{ij} dengan $i = 1, 2, \dots, n$ dan $j = 1, 2, \dots, p$ dan dapat dirumuskan sebagai berikut:

Adapun ciri-ciri cluster (gerombol) adalah:

1. *Homogenitas* (kesamaan) yang tinggi antar anggota dalam satu cluster
2. *Heterogenitas* (perbedaan) yang tinggi antar cluster yang satu dengan cluster yang lain

2.3 Tujuan Analisis Cluster

Setelah mengelompokkan n buah objek pengamatan kedalam m kelompok berdasarkan p variabel dapat diketahui bahwa tujuan utama dari penggolongan objek adalah untuk memperoleh kelompok objek yang memiliki nilai relatif sama. Sehingga nantinya dalam

TABEL 1: Tabel data observasi x_{ij}

	Var	Var	...	Var	...	Var
	1	2		j		p
Objek 1	X_{11}	X_{12}	...	X_{1j}	...	X_{1p}
Objek 2	X_{21}	X_{22}	...	X_{2j}	...	X_{2p}
...
Objek i	X_{i1}	X_{i2}	...	X_{ij}	...	X_{ip}
...
Objek n	X_{n1}	X_{n2}	...	X_{nj}	...	X_{np}

²interpretasi, objek-objek yang berada pada satu cluster memiliki peluang yang cukup tinggi akan muncul bersamaan pada satu individu

2.4 Konsep Dasar dalam Analisis Cluster

Proses analisis cluster tersebut meliputi:

1. Menentukan ukuran ketakmiripan antara kedua objek

Proses pertama yaitu mengukur seberapa jauh ada kesamaan antar objek. Dengan memiliki sebuah ukuran kuantitatif untuk mengatakan dua objek tertentu lebih mirip dibandingkan dengan objek lain, akan mempermudah proses peng-cluster-an. Salah satu yang bisa menjadi ukuran ketidakmiripan adalah fungsi jarak antara objek a dan b, yang biasa dinotasikan dengan $d(a,b)$. Adapun sifat-sifat ukuran ketidakmiripan adalah:

- $d(a,b) \geq 0$
- $d(a,b) = 0$
- $d(a,b) = d(b,a)$
- (a,b) meningkat seiring semakin tidak miripnya kedua objek a dan b
- $d(a,c) \leq d(a,b) + d(b,c)$

Jarak paling umum digunakan adalah jarak *Euclidian*. Ukuran jarak atau ketidaksamaan antar objek ke- i dengan objek ke- h , dapat disimbolkan dengan d_{ih} . Adapun nilai d_{ih} diperoleh melalui perhitungan jarak kuadrat sebagai berikut:

$$d_{ih} = \sum (x_{ij} - x_{hj})^2$$

di mana ¹³ d_{ih} = jarak kuadrat Euclidian antar objek ke- i dengan objek ke- h p = jumlah variabel cluster x_{ij} = nilai dari objek ke- i pada variabel ke- j x_{hj} = nilai dari objek ke- h pada variabel ke- j

2. Membuat cluster

Proses pembuatan cluster dapat dilakukan dengan dua tahap, yaitu:

- (a) Metode Hirarki

Pada metode ini⁷ dimulai dengan mengelompokkan data yang mempunyai kesamaan yang paling dekat. Kemudian diteruskan ke objek lain yang mempunyai kedekatan kedua.⁴ Demikian seterusnya sehingga kelompok akan membentuk semacam "pohon", dimana ada hierarki yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Metode yang digunakan adalah *single linkage* (pautan tunggal). Metode ini akan mengelompokkan dua objek yang mempunyai jarak terdekat lebih dahulu. Jadi pada setiap tahapan, banyaknya cluster akan berkurang satu. Hasil berupa *single linkage clustering* dapat disajikan dalam bentuk suatu dendogram atau diagram pohon. Cabang-cabang pohon menunjukkan cluster/gerombol. Cabang-cabang tersebut bertemu bersama-sama pada simpul yang posisinya sepanjang suatu sumbu jarak (kemiripan). Ini menunjukkan tingkat dimana penggabungan terjadi.

- (b) Metode Non-Hirarki

Metode ini dimulai dengan proses penentuan jumlah gerombol terlebih dahulu, d_3 yang digunakan adalah metode *k-means*. Metode *k-means* digunakan sebagai alternatif metode cluster untuk data dengan ukuran yang besar karena memiliki kecepatan yang lebih tinggi dibandingkan metode hierarki.

Mac Queen menyarankan bahwa penggunaan *k-means* untuk menjelaskan algoritma dalam penentuan suatu objek ke dalam cluster tertentu berdasarkan rata-rata terdekat. Proses pengelompokan dengan *k-means* adalah:

- Menentukan besarnya k , yaitu banyaknya cluster dan menentukan *centroid* (rata-rata) di tiap cluster.
- Menghitung jarak tiap objek dengan setiap *centroid*⁶
- Menghitung kembali rata-rata (*centroid*) untuk cluster yang baru terbentuk.
- Mengulangi langkah 2 sampai tidak ada lagi pemindahan objek antar cluster.

3. Setelah cluster terbentuk,⁵ itu dengan kedua metode yang disebutkan, langkah selanjutnya adalah melakukan interpretasi terhadap cluster yang terbentuk.

4. Melakukan validasi cluster.

Untuk menguji validasi cluster digunakan uji parsial F dengan taraf signifikansi α .

Hipotesis:

H_0 : Variabel i bukan variabel pembeda dalam penggerombolan. H_1 : Variabel i merupakan variabel pembeda dalam penggerombolan.

Statistik uji

$$F = \frac{\text{means square cluster}}{\text{means square error}}$$

10
Kriteria uji:

Tolak H_0 jika $F > F_{\alpha, k-1, n-k}$

3 METODOLOGI

10
Langkah-langkah yang digunakan dalam menyusun penelitian ini adalah:

1. Pengambilan data pencemaran udara pada sektor industri dari BLH (Badan Lingkungan Hidup) Provinsi Sumatera Selatan.
2. Menentukan Variabel dan Objek
Adapun variabel-variabel yang digunakan dalam skripsi ini ada 5 variabel, dimana merupakan jenis polutan yang meliputi debu (TSP=*total suspended particulate*), *Carbon Monoksida* (CO), *Amonia* (NH₃), *Nitrogen Dioksida* (NO₂), dan *Sulfur Dioksida* (SO₂). Sedangkan objek-objek yang digunakan yaitu 10 jenis industri di Sumatera Selatan, yaitu industri karet, industri sawit, industri pengalengan ikan, industri migas, industri listrik, industri minyak goreng, industri makanan, industri pupuk, industri pertambangan, dan industri semen.
3. Menerapkan analisis cluster pada data penelitian yaitu dengan:

- Membentuk matriks ketidakmiripan antar objek dengan menggunakan *jarak Euclidian*.
- Membuat cluster dengan dua metode, yaitu metode hirarki dan non-hirarki.
 - a Metode hirarki dengan menggunakan metode pautan tunggal atau *single linkage*.
 - Menentukan jarak ketakmiripan antara sebuah objek dalam gerombol yang satu dengan sebuah objek dalam gerombol yang lain. Adapun ukuran ketidakmiripan yang digunakan adalah $h\{B_r, B_s\} = \min \{d(x_i, x_h); x_i \text{ anggota } B_r, x_h \text{ anggota } B_s\}$
 - b Metode non hirarki dengan menggunakan metode k-means.

i. Menentukan besarnya k , yaitu banyaknya cluster dan menentukan *centroid* (rata-rata) di tiap cluster.

ii. Menghitung jarak tiap objek dengan setiap *centroid*.

iii. Menghitung kembali rata-rata (*centroid*) untuk cluster yang baru terbentuk.

iv. Mengulangi langkah (ii) sampai tidak ada lagi pemindahan objek antar cluster.

– Memberi nama spesifik pada tiap cluster, untuk menggambarkan karakter tiap cluster.

– Melakukan validasi cluster.

4. Interpretasi hasil cluster yang diperoleh dari dua metode pada langkah kedua.

4 HASIL DAN PEMBAHASAN

1
Dari data awal, tabel data beban pencemaran udara pada sektor industri di Sumatera Selatan pada tahun 2009.

4.1 Menentukan Ukuran Ketakmiripan antara Dua Objek

Jarak tiap objek (jenis industri) yang dihitung dengan jarak *Euclidian* yang **22** ditampilkan dalam Tabel *Proximity Matrix* dimana semakin kecil jarak antara dua objek, maka semakin mirip kedua objek tersebut.

4.2 Proses Clustering secara Hirarki

Proses penggerombolan yaitu dengan menggunakan matriks jarak yang ada pada Tabel *Proximity Matrix*. Proses clustering dengan cara hirarki yaitu:

1. Bandingkan jarak Euclidian yang ada pada table *Proximity matriks*, dimana menentukan jarak yang terdekat antara dua objek dari sekian banyak kombinasi jarak dari kesepuluh objek yang ada.
2. Kemudian dilakukan perbaikan matriks jarak menggunakan metode pautan tunggal dengan persamaan: $h\{B_r, B_s\} = \min \{d(x_i, x_h); x_i \text{ anggota } B_r, x_h \text{ anggota } B_s\}$, sehingga diperoleh table matriks jarak yang baru.
3. Kembali menentukan jarak terdekat antara dua objek pada tabel matriks jarak yang baru dan mengulangi langkah 2, sampai tersisa satu gerombol

Pada awal proses ada sepuluh gerombol. Tahap pertama yang dilakukan adalah menentukan jarak yang

TABEL 2: Tabel Beban ¹ Pencemaran Udara Pada Sektor Industri di Sumatera Selatan

No	Jenis Industri	NO ₂	SO ₂	CO	TSP	NH ₃
1	Karet	630.545	1140.640	8620.366	948.708	2.102
2	Sawit	696.770	1841.352	1177.563	610.747	129.210
3	Pengalengan ikan	398.580	360.824	4380.000	206.999	0.876
4	Migas	1003.108	1672.109	16175.603	824.930	0.00
5	Listrik	370.110	405.150	5689.620	428.714	0.00
6	Minyak goreng	974.200	1641.624	17041.529	530.944	1.972
7	Makanan	956.855	1511.012	16950.600	1123.470	2.278
8	Pupuk	1097.453	2157.413	22250.400	1071.790	1.752
9	Pertambangan	605.841	449.563	4174.140	476.325	0.262
10	Semen	777.45	1727.910	9592.200	1151.414	0.00

terdekat antara dua objek dari sekian banyak kombinasi jarak dari kesepuluh objek yang ada. Jarak antara industri pengalengan ikan dengan industri tambang merupakan jarak yang terdekat yaitu sebesar

407,120 sehingga kedua industri tersebut menjadi satu kelompok. Sekarang tersisa sembilan gerombol.

Kemudian dilakukan perbaikan matriks jarak menggunakan metode pautan tunggal dengan persamaan:

$$h\{B_r, B_s\} = \min \{d(x_i, x_h); x_i \text{ anggota } B_r, x_h \text{ anggota } B_s\}$$

sehingga terjadi perubahan jarak yang melibatkan gerombol baru (gerombol yang anggotanya industri pengalengan ikan dan tambang) sehingga diperoleh:

$$\begin{aligned} h\{(3, 9), (1)\} &= \min \{d(3, 1), d(9, 1)\} \\ &= \min 4.830, 954, 4.524, 409 \\ &= 4.524, 409 \end{aligned}$$

$$\begin{aligned} h\{(3, 9), (2)\} &= \min \{d(3, 2), d(9, 2)\} \\ &= \min 3.565, 945, 3.310, 516 \\ &= 3.310, 516 \end{aligned}$$

$$\begin{aligned} h\{(3, 9), (4)\} &= \min \{d(3, 4), d(9, 4)\} \\ &= \min 11.899, 706, 12.075, 143 \\ &= 11.899, 706 \end{aligned}$$

$$\begin{aligned} h\{(3, 9), (5)\} &= \min \{d(3, 5), d(9, 5)\} \\ &= \min 1.329, 3, 1.535, 086 \\ &= 1.329, 300 \end{aligned}$$

$$\begin{aligned} h\{(3, 9), (6)\} &= \min \{d(3, 6), d(9, 6)\} \\ &= \min 12.743, 275, 12.927, 853 \\ &= 12.743, 275 \end{aligned}$$

$$\begin{aligned} h\{(3, 9), (7)\} &= \min \{d(3, 7), d(9, 7)\} \\ &= \min 12.668, 643, 12.841, 597 \\ &= 12.668, 643 \end{aligned}$$

$$\begin{aligned} h\{(3, 9), (8)\} &= \min \{d(3, 8), d(9, 8)\} \\ &= \min 17.994, 866, 18.173, 172 \\ &= 17.994, 866 \end{aligned}$$

$$\begin{aligned} h\{(3, 9), (10)\} &= \min \{d(3, 10), d(9, 10)\} \\ &= \min 5.483, 741, 5.610, 235 \\ &= 5.483, 741 \end{aligned}$$

Kemudian lakukan lagi penentuan jarak terdekat pada objek berdasarkan tabel matriks jarak yang baru.

4.3 Proses Penggerombolan secara non-hirarki

1. Tahap pertama, menentukan besarnya k .
2. Tahap kedua, menghitung jarak setiap objek dengan tiap centroid (rata-rata) dengan rumus jarak kuadrat seperti ya disebutkan diatas, lalu dibandingkan untuk mengetahui kelompok objek.
3. Tahap ketiga, menghitung kembali centroid yang baru yang merupakan rata-rata kelima variabel dari tiap cluster.
4. Tahap keempat, mengulangi tahap kedua dan ketiga. Jika tidak ada lagi perubahan hasil pengcluster-an, maka proses terhenti

TABEL 3: Tabel Jarak Euclidian, yaitu dari Tabel proximity Matriks

1	2	3	4	5	6	7	8	9	10	
1	0	7484.723	4830.954	7584.075	3077.083	8453.378	8346.676	13676.433	4524.409	1162.765
2		0	3565.945	15004.208	4751.618	15868.360	15787.487	21084.442	3310.516	8434.128
3			0	11899.706	1329.300	12743.275	12668.643	17994.866	407.120	5483.741
4				0	10588.612	915.437	847.256	6099.879	12075.143	6595.591
5					0	11435.475	11351.633	16681.500	1535.08	4203.337
6						0	613.772	5263.657	12927.853	7478.211
7							0	5341.175	12841.597	7363.835
8								0	18173.172	12669.777
9									0	5610.235
10										0

Dari kesepuluh jenis Industri kita bandingkan centroid yang terbesar dan terkecilnya. Adapun centroid terbesar (c_1) adalah rata-rata polutan pada industri pupuk dengan nilai $c_1 = (1.097,453; 2.157,413; 22.250,400; 1.071,790; 1,752)$ dan centroid terkecil (c_2) adalah nilai rata-rata polutan pada industri sawit dengan nilai $c_2 = (696,770; 1.841,352; 1.177,563; 610,747; 129,210)$.

Lalu penghitungan jarak setiap objek dari centroid (pusat cluster) pertama dengan menggunakan kuadrat jarak Euclidian. Gunakan rumus jarak kuadrat, sehingga diperoleh Tabel jarak setiap objek dari sentroid Pertama dan kedua sebagai berikut:

TABEL 4: Jarak Setiap Objek dari Sentroid Pertama

Industri	c_1
Karet	$181,04 \times 10^6$
Sawit	$444,55 \times 10^6$
Pengalengan ikan	$323,79 \times 10^6$
Migas	$37,44 \times 10^6$
Listrik	$278,27 \times 10^6$
Minyak goreng	$27,70 \times 10^6$
Makanan	$28,53 \times 10^6$
Pupuk	0
Pertambangan	$329,91 \times 10^6$
Semen	$160,52 \times 10^6$

Lalu kita bandingkan kedua jarak tiap objek (industri) ke pusat cluster (centroid).

4.4 Validasi cluster

Untuk melihat apakah variabel-variabel yang telah membentuk cluster tersebut merupakan variabel pembeda dalam peng-cluster-an dapat dilihat pada tabel Anova.

Hipotesis:

TABEL 5: Tabel Jarak Setiap Objek dari Sentroid Kedua

Industri	c_2
Karet	$56,02 \times 10^6$
Sawit	0
Pengalengan ikan	$15,8 \times 10^6$
Migas	$225,13 \times 10^6$
Listrik	$22,6 \times 10^6$
Minyak goreng	$251,8 \times 10^6$
Makanan	$249,24 \times 10^6$
Pupuk	$444,54 \times 10^6$
Pertambangan	$10,96 \times 10^6$
Semen	$71,13 \times 10^6$

H_0 : Variabel i bukan pembeda dalam peng-cluster-an

H_1 : Variabel i merupakan pembeda dalam peng-cluster-an

Kriteria uji

Tolak H_0 jika $F > F_{\alpha, k-1, n-k}$

Dipilih $\alpha = 0.05$ dan nilai $n = 10$, dan $k = 2$, sehingga dalam tabel distribusi F diperoleh nilai $F_{\alpha, k-1, n-k}$ adalah 5,32. Nilai F hitung pada tabel anova, variabel $NO_2 = 24,326$, variabel $SO_2 = 4,304$, variabel $CO = 41,94$, variabel $TSP = 1,44$, variabel $NH_3 = 0,590$. Karena nilai F hitung dari kelima variabel ada dua yang bernilai $>$ dari nilai $F_{\alpha, k-1, n-k}$ yaitu variabel NO_2 dan variabel CO , dengan demikian kedua variabel tersebut merupakan variabel pembeda dalam peng-cluster-an.

Dari tabel anova, juga dapat dilihat nilai signifikansi, dimana ada dua nilai, yaitu variabel $NO_2 = 0,001 < \alpha = 0,05$ dan $CO = 0,00 < \alpha = 0,05$, sehingga pada kedua variabel itu tolak H_0 . Hal ini berarti hanya ada dua variabel, yaitu NO_2 dan CO , yang merupakan variabel pembeda dalam peng-cluster-an

TABEL 6: Tabel Jarak Setiap Objek (Industri) ke Pusat Cluster

Industri	c_1	c_2
Karet	$181,04 \times 10^6$	$56,02 \times 10^6$
Sawit	$444,55 \times 10^6$	0
Pengalengan ikan	$323,79 \times 10^6$	$15,8 \times 10^6$
Migas	$37,44 \times 10^6$	$225,13 \times 10^6$
Listrik	$278,27 \times 10^6$	$22,6 \times 10^6$
Minyak goreng	$27,70 \times 10^6$	$251,8 \times 10^6$
Makanan	$28,53 \times 10^6$	$249,24 \times 10^6$
Pupuk	0	$444,54 \times 10^6$
Pertambangan	$329,91 \times 10^6$	$10,96 \times 10^6$
Semen	$160,52 \times 10^6$	$71,13 \times 10^6$

5 KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil analisis dan pembahasan dapat diambil keputusan sebagai berikut:

1. Hasil analisis cluster dari 10 jenis industri berdasarkan jenis polutan TSP, CO, NH₃, NO₂, dan SO₂,
 - a. Untuk metode hirarki dapat dikelompokkan menjadi tiga kelompok, yaitu cluster (kelompok industri) pencemar rendah, cluster pencemar sedang, dan cluster pencemar tinggi
 - b. Untuk metode non hirarki dapat dikelompokkan menjadi dua kelompok, yaitu cluster (kelompok industri) pencemar tinggi dan cluster (kelompok industri) pencemar rendah.
2. Ciri dari setiap kelompok yang terbentuk yaitu:
 - a. Untuk metode hirarki:
 1. cluster pencemar rendah: memiliki rerata nilai variabel (polutan) yang paling kecil dan anggotanya adalah industri karet, sawit, pengalengan ikan, listrik, pertambangan, dan semen.
 2. cluster pencemar sedang: memiliki rerata nilai variabel (polutan) yang lebih besar dari cluster pertama, namun lebih kecil dari cluster ketiga. Anggotanya adalah industri migas, minyak goreng, dan makanan.
 3. cluster pencemar tinggi: memiliki rerata nilai variabel terbesar dan anggotanya adalah industri pupuk
 - b. Untuk metode non hirarki:

1. cluster pencemar tinggi: kelima polutannya memiliki rerata lebih besar dari cluster kedua. Anggotanya adalah industri migas, minyak goreng, makanan, dan pupuk.
2. cluster pencemar rendah: kelima polutannya memiliki rerata lebih kecil dari cluster pertama. Anggotanya adalah industri karet, sawit, pengalengan ikan, listrik, pertambangan, dan semen

5.2 Saran

Berdasarkan dari kesimpulan yang telah dikemukakan, maka peneliti mencoba untuk memberikan saran-saran sebagai berikut:

1. Untuk meningkatkan pengaruh polutan terhadap pencemaran udara pada sektor industri, maka harus diarahkan lagi ke polutan yang memberikan pengaruh besar ke sektor industri itu.
2. Untuk peneliti selanjutnya, agar menambah variabel lain yang mempengaruhi tingkat pencemaran udara pada sektor industri di Sumatera Selatan.
3. Peranan industri dapat ditingkatkan untuk relokasi industri pencemar udara ke kawasan atau zona industri.

DAFTAR PUSTAKA

- [1] Tim Penelitian dan Pengembangan Wahana Komputer, 2005, *Pengembangan Analisis Multivariat dengan SPSS 12*, Edisi Pertama, Salemba Infotek, Jakarta, Hal.120
- [2] Supranto, J., 2000, *Teknik Sampling untuk Survei dan Eksperimen*, Edisi Baru, PT. Rineka Cipta, Jakarta
- [3] _____, 2010, Analisis Peubah Ganda, http://www.adia08.files.wordpress.com/2008/06/analisis_peubah_ganda.pdf, Diakses pada 20 February 2010
- [4] _____, 2010, Pencemaran Udara, http://www.sabah.org.my/bm/kenali_sabah/as_pencemaran_udara.asp, Diakses pada 20 February 2010
- [5] Kristanto, P., 2004, *Ekologi Industri*, Andi Offset, Yogyakarta
- [6] Perkins, H.C., 1980, *Air Pollution*, Mcgraw-Hill Higher Education, New York

46-2011-Analysis klaster

ORIGINALITY REPORT

16%

SIMILARITY INDEX

13%

INTERNET SOURCES

6%

PUBLICATIONS

4%

STUDENT PAPERS

PRIMARY SOURCES

1	ebookmarket.org Internet Source	1%
2	scholar.unand.ac.id Internet Source	1%
3	repository.uksw.edu Internet Source	1%
4	etd.repository.ugm.ac.id Internet Source	1%
5	repository.stieyapan.ac.id Internet Source	1%
6	Melisa Matahari, John Kekenusa, Yohanes Langi. "Pengelompokan Sekolah Dasar Di Siau berdasarkan Indikator Mutu Sekolah dengan Menggunakan Analisis Gerombol", d'CARTESIAN, 2015 Publication	1%
7	sis.binus.ac.id Internet Source	1%
8	Submitted to Sultan Agung Islamic University Student Paper	

1 %

9

repository.unmuhpnk.ac.id

Internet Source

1 %

10

repository.unpas.ac.id

Internet Source

1 %

11

Zaifullah Zaifullah, Tony Yulianto. "Analisis Cluster Untuk Pengelompokan Prestasi Mahasiswa Angkatan 2013 Fakultas MIPA Universitas Islam Madura", Zeta - Math Journal, 2022

Publication

1 %

12

cafestat.wordpress.com

Internet Source

1 %

13

Submitted to Padjadjaran University

Student Paper

1 %

14

Dewi Rachmatin. "APLIKASI METODE-METODE AGGLOMERATIVE DALAM ANALISIS KLASER PADA DATA TINGKAT POLUSI UDARA", Infinity Journal, 2014

Publication

1 %

15

Submitted to Universitas Bina Darma

Student Paper

1 %

16

eigen.unram.ac.id

Internet Source

1 %

17	eprints.unm.ac.id Internet Source	1 %
18	exsight.id Internet Source	1 %
19	jurnal.stmik-dci.ac.id Internet Source	1 %
20	repository.ipb.ac.id Internet Source	1 %
21	Submitted to Forum Komunikasi Perpustakaan Perguruan Tinggi Kristen Indonesia (FKPPTKI) Student Paper	1 %
22	ejournal.unitomo.ac.id Internet Source	1 %

Exclude quotes Off

Exclude matches < 1%

Exclude bibliography Off