

## PENGELO MPOKAN KUALITAS UDARA AMBIEN MENURUT KABUPATEN/KOTA DI JAWA TENGAH MENGUNAKAN ANALISIS KLASTER

Rizki Taher Dwi Kurniawati<sup>1</sup>, Rita Rahmawati<sup>2</sup>, Yuciana Wilandari<sup>3</sup>

<sup>1</sup>Mahasiswa Jurusan Statistika FSM Universitas Diponegoro

<sup>2,3</sup>Staf Pengajar Jurusan Statistika FSM Universitas Diponegoro

### ABSTRACT

Ambient air is free of air inhaled daily by living creatures. Ambient air quality can be said to be decreased which is known from the results of measuring the quality of ambient air. The measurements carried out on residential areas, industrial areas, and traffic congested area and to SO<sub>2</sub>, NO<sub>2</sub>, CO, and HC. To help find solutions used cluster analysis of air pollution. Cluster analysis classifying objects based on object similarity. Similarities object seen by the small size of the Euclidean distance. The process of clustering with average linkage method performed on the data type of the region and type of pollutants. Clustering process produces two clusters for different kinds of land and 2 clusters for these types of pollutants. From the analysis on the type of region, cluster 1 is composed of 33 districts/cities with the results of measuring between 507  $\mu\text{g}/\text{N m}^3$  to 6760  $\mu\text{g}/\text{N m}^3$  can be said to have a good air quality conditions and in cluster 2 consists of two districts/cities with the results of measuring 11856.6  $\mu\text{g}/\text{N m}^3$  and 10594.8  $\mu\text{g}/\text{N m}^3$  is said to have poor air quality conditions. On the type of pollutant, Cluster 1 consists of 34 districts/cities with the measuring between 30  $\mu\text{g}/\text{N m}^3$  to 10810  $\mu\text{g}/\text{N m}^3$  which is said to have good air condition and the second cluster consists of one district/cities that have poor air conditions with a value of 20095 HC pollutants  $\mu\text{g}/\text{N m}^3$ .

**Keywords:** ambient air, euclidean, average linkage, cluster analysis.

### 1. PENDAHULUAN

Udara dibedakan menjadi udara emisi dan udara ambien. Udara emisi yaitu udara yang dikeluarkan oleh sumber emisi seperti knalpot kendaraan bermotor dan cerobong gas buang industri. Sedangkan udara ambien adalah udara bebas di permukaan bumi yang sehari-hari dihirup oleh makhluk hidup (PP No.41 Tahun 1999). Untuk mendapatkan udara ambien yang berkualitas baik perlu dilakukan pengendalian pencemaran udara. Pengendalian pencemaran udara dapat dilakukan salah satunya dengan memantau atau mengukur kualitas udara, baik udara ambien ataupun udara emisi. Pengukuran kualitas udara ambien dilakukan di kawasan perumahan, kawasan industri, dan kawasan padat lalu lintas dimana di kawasan-kawasan tersebut banyak terjadi kegiatan manusia. pengukuran kualitas udara ambien juga dilakukan terhadap zat-zat yang dapat menjadi polutan seperti SO<sub>2</sub>, NO<sub>2</sub>, CO, dan HC.

Dari hasil pengukuran kualitas udara ambien tersebut dapat diketahui kualitas udara ambien jenis kawasan dan jenis polutan. Untuk mengetahui jenis kawasan dan

jenis polutan yang dapat mempengaruhi kualitas udara ambien, salah satunya dapat dilakukan dengan mengelompokkan jenis kawasan dan jenis polutan tersebut. Dari hasil pengelompokan kemudian dapat digunakan sebagai bahan pertimbangan dalam mencari solusi untuk mengatasi kualitas udara ambien yang mulai menurun.

Salah satu analisis yang dapat digunakan untuk melakukan proses pengelompokan tersebut adalah analisis klaster (*cluster analysis*). Analisis klaster melakukan pengelompokan objek berdasarkan karakteristik yang dimiliki objek (Hair *et al.*, 2010). Setiap kelompok (klaster) terdiri dari beberapa objek yang memiliki karakteristik yang sama dengan prinsip homogenitas grup.

## **2. TINJAUAN PUSTAKA**

### **2.1 Udara Ambien**

Menurut PP No. 41 Tahun 1999 tentang Pengendalian Pencemaran Udara, udara ambien adalah udara bebas di permukaan bumi pada lapisan troposfir yang berada di dalam wilayah yurisdiksi Republik Indonesia yang dibutuhkan dan mempengaruhi kesehatan manusia, makhluk hidup, dan unsur lingkungan hidup lainnya. Unsur-unsur berbahaya yang masuk ke dalam atmosfer dapat berupa Karbonmonoksida (CO), Nitrogen dioksida (NO<sub>2</sub>), Sulfur dioksida (SO<sub>2</sub>), Hidrokarbon (HC), dan lain-lain.

Menurut PP No. 41 Tahun 1999 tentang Pengendalian Pencemaran Udara, baku mutu udara ambien adalah ukuran batas atau kadar zat, energi dan/atau komponen yang ada atau yang seharusnya ada dan/atau unsur pencemar yang ditenggang keberadaannya dalam udara ambien. Udara yang melebihi baku mutu dapat merusak lingkungan sekitarnya dan berpotensi mengganggu kesehatan masyarakat sekitarnya.

Baku mutu udara ambien untuk SO<sub>2</sub> sebesar 632  $\mu\text{g}/\text{Nm}^3$ , NO<sub>2</sub> sebesar 316  $\mu\text{g}/\text{Nm}^3$ , CO sebesar 15000  $\mu\text{g}/\text{Nm}^3$ , dan HC sebesar 160  $\mu\text{g}/\text{Nm}^3$ . Pencemaran udara ambien dirasakan semakin hari semakin meningkat, terutama di kawasan perumahan, kawasan industri, dan kawasan padat lalu lintas, dimana di kawasan-kawasan tersebut banyak terjadi kegiatan manusia. Pencemaran udara ambien dapat pula menimbulkan dampak terhadap lingkungan alam, antara lain: hujan asam, penipisan lapisan ozon dan pemanasan global.

### **2.2 Analisis Klaster (*Cluster Analysis*)**

Pengklasteran adalah pengelompokan objek atau kasus menjadi kelompok-kelompok yang lebih kecil dimana setiap kelompok berisi objek atau kasus yang mirip satu sama lain (Supranto, 2004). Analisis klaster merupakan kelompok teknik multivariat yang bertujuan untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimiliki (Hair *et al.*, 2010).

Setiap kelompok terdiri dari objek-objek yang relatif mirip atau sama yang disebut klaster (*cluster*). Klaster atau kelompok yang terbentuk memiliki homogenitas internal yang tinggi dan heterogenitas eksternal yang tinggi (Supranto, 2004). Objek-objek dalam satu *cluster* atau kelompok akan menjadi dekat bersama-sama ketika diplotkan, dan kelompok yang berbeda akan jauh terpisah.

Beberapa tujuan dari analisis klaster yang dipergunakan dalam berbagai bidang menurut Dillon (1984) yaitu:

1. Dalam bidang psikologi digunakan untuk mengelompokkan individu berdasarkan tipe kepribadian seseorang.
2. Dalam bidang kimia digunakan untuk mengelompokkan senyawa berdasarkan sifat.
3. Dalam bidang pemasaran digunakan untuk membuat segmen pasar, memahami perilaku pembeli, mengenali peluang produk baru, dan lain-lain.

## 2.3 Tahapan Analisis Klaster

### 2.3.1 Perumusan Masalah

Dalam perumusan masalah pada analisis klaster, harus diketahui terlebih dahulu tujuan utama analisis klaster yaitu mengelompokkan objek-objek ke dalam dua atau lebih kelompok berdasarkan kemiripan karakteristik objek. Dalam analisis klaster dilakukan dengan mengelompokkan pengamatan yang kemudian dapat digunakan untuk analisis selanjutnya. Dari klaster atau kelompok yang telah terbentuk dan struktur data yang mendasari diperlihatkan di dalam klaster, peneliti mendapat informasi tentang hubungan antar pengamatan (Hair *et al.*, 2010).

### 2.3.2 Asumsi Analisis Klaster

#### a. Sampel yang Mewakili (*Sample Representative*)

Sampel yang mewakili atau sampel representatif adalah sampel yang diambil dapat dikatakan mempresentasikan atau mewakili populasi yang ada. Pengujian sampel yang mewakili (sampel representatif) dapat dilakukan dengan uji Kaiser-Mayer-Olkin (KMO). Uji Kaiser-Mayer-Olkin (KMO) banyak digunakan untuk melihat syarat kecukupan suatu sampel. Uji KMO ini mengukur kecukupan *sampling* secara menyeluruh dan mengukur kecukupan *sampling* untuk setiap indikator. Jika nilai KMO berkisar 0,5 sampai 1 maka sampel dapat dikatakan mewakili populasi atau sampel representatif. Dirumuskan sebagai berikut (Widarjono, 2010):

$$KMO = \frac{\sum_{j=1}^p \sum_{k=1}^p r_{X_j X_k}^2}{\sum_{j=1}^p \sum_{k=1}^p r_{X_j X_k}^2 + \sum_{j=1}^p \sum_{k=1}^p \rho_{X_j X_k, X_l}^2}$$

dengan:

$$r_{X_j X_k} = \frac{\frac{1}{n} \sum_{i=1}^n (X_{ij} - \bar{X}_j)(X_{ik} - \bar{X}_k)}{\sqrt{\frac{\sum_{i=1}^n (X_{ij} - \bar{X}_j)^2}{n}} \sqrt{\frac{\sum_{i=1}^n (X_{ik} - \bar{X}_k)^2}{n}}} \quad \text{dan}$$

$$\rho_{X_j X_k, X_l} = \frac{r_{X_j X_k} - r_{X_j X_l} r_{X_k X_l}}{\sqrt{(1 - r_{X_j X_l}^2)(1 - r_{X_k X_l}^2)}}$$

dimana:

- $r_{X_j X_k}$  = korelasi antar variabel  $X_j$  dan  $X_k$
- $\bar{X}_j$  = rata-rata variabel  $X_j$
- $\bar{X}_k$  = rata-rata variabel  $X_k$
- $n$  = banyaknya observasi (objek)

$\rho_{X_j X_k, X_l}$  = korelasi parsial antara variabel  $X_j$  dan  $X_k$  dengan menjaga agar  $X_l$  konstan

b. Multikolinearitas

Multikolinearitas adalah adanya hubungan linear yang sempurna atau pasti di antara beberapa atau semua variabel (Gujarati, 1978). Untuk mengetahui adanya multikolinieritas salah satunya adalah dengan menghitung nilai *Variance Inflation Factor* ( $VIF_j$ ) dengan rumus :

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

Dengan  $R_j^2$  adalah nilai koefisien determinasi variabel dependen ke-j dengan variabel independen selain variabel ke-j. Menurut Widarjono (2010), terjadinya multikolinearitas apabila nilai ( $VIF_j$ )  $\geq 10$ . Jika terindikasi terjadi multikolinieritas maka harus dilakukan tindakan perbaikan multikolinearitas salah satunya dengan menggunakan analisis komponen utama (*principal components analysis*) (Hair *et al.*, 2010).

**2.3.3 Pemilihan Ukuran Jarak**

Menurut Hair *et al.* (2010) beberapa jarak yang umum digunakan dalam analisis kluster yaitu:

- a) Jarak *Euclidean* adalah akar dari jumlah kuadrat perbedaan deviasi di dalam nilai untuk setiap variabel (Supranto, 2004). Jarak *Euclidean* antara kelompok objek ke-i dan kelompok objek ke-g dari p variabel didefinisikan sebagai berikut:

$$d(X_i, X_g) = \left[ \sum_{j=1}^p |X_{ij} - X_{gj}|^2 \right]^{1/2}$$

- b) Jarak *Squared Euclidean* adalah jumlah kuadrat perbedaan deviasi di dalam nilai untuk setiap variabel (Hair *et al.*, 2010). Jarak *Squared Euclidean* antara kelompok objek ke-i dan kelompok objek ke-g dari p variabel didefinisikan sebagai berikut:

$$d(X_i, X_g) = \sum_{j=1}^p (X_{ij} - X_{gj})^2$$

- c) Jarak *Cityblock* atau Jarak *Manhattan* adalah jumlah perbedaan mutlak atau absolut di dalam nilai untuk setiap variabel (Supranto, 2004). Jarak *Cityblock* atau Jarak *Manhattan* antara kelompok objek ke-i dan kelompok objek ke-g dari p variabel didefinisikan sebagai berikut:

$$d(X_i, X_g) = \left[ \sum_{j=1}^p |X_{ij} - X_{gj}| \right]^{1/2}$$

d) Jarak Mahalanobis ( $D^2$ )

Jarak Mahalanobis merupakan bentuk standar dari jarak *Euclidean*. Respon pengukuran dalam bentuk standar deviasi yang menstandarkan data dengan penyesuaian yang dibuat untuk korelasi antar variabel (Hair *et al.*, 2010), didefinisikan sebagai berikut:

$$D^2 = (X - \mu)^T S^{-1} (X - \mu)$$

dimana:

$D^2$  = Jarak Mahalanobis

$S^{-1}$  = Invers dari Matriks Varian Kovarian

Dalam tugas akhir ini, pada tahap pemilihan ukuran jarak dipilih jarak *Euclidean* sebagai ukuran kemiripan dari suatu objek. Dikatakan memiliki kemiripan apabila memiliki jarak yang pendek. Semakin pendek jarak *Euclidean* antar objek maka dapat dikatakan kedua objek memiliki kemiripan.

### 2.3.4 Pemilihan Prosedur Pengklasteran

Menurut Hair *et al.* (2010) terdapat dua prosedur pengklasteran yaitu:

a. Metode Hirarki (*Hierarchical Method*)

Prosedur hirarki melibatkan serangkaian keputusan N-1 pengelompokan (dimana N sama dengan jumlah pengamatan) yang menggabungkan pengamatan hirarki atau struktur mirip pohon (*treelike structure*). Terdapat dua tipe dalam metode hirarki, yaitu aglomeratif (*agglomerative*) dan devisif (*divisive*).

Metode aglomeratif yang biasa digunakan dalam analisis kluster yaitu:

1. *Single linkage*

Mendefinisikan kesamaan atau kemiripan antar kluster dari jarak terdekat objek-objek dalam satu kluster ke objek-objek dalam kluster lainnya (Johnson dan Wichern, 2007). Dihitung dengan cara:

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

2. *Complete linkage*

Mendefinisikan kesamaan atau kemiripan antar kluster dari jarak maksimal atau terjauh dari objek-objek dalam satu kluster ke objek-objek dalam kluster lainnya (Johnson dan Wichern, 2007). Dihitung dengan cara:

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

3. *Average linkage*

Mendefinisikan kesamaan antar kluster dari rata-rata kesamaan dari semua objek dalam satu kluster dengan semua objek dalam kluster lainnya (Johnson dan Wichern, 2007). Dihitung dengan cara:

$$d_{(UV)W} = \text{average}\{d_{UW}, d_{VW}\} = \frac{\sum_{i=1}^n \sum_{g=1}^n d_{ig}}{N_{(UV)}N_W}$$

4. *Centroid method*

Mendefinisikan kesamaan antar kluster dari jarak di antara dua *centroid* kluster-kluster yang ada. Centroid adalah rata-rata jarak, yang didapat dengan melakukan rata-rata pada semua anggota suatu kluster tertentu (Johnson dan Wichern, 2007). Dihitung dengan cara:

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

### 5. *Ward's method*

Mendefinisikan kesamaan antar kluster dengan menghitung jumlah kuadrat antar kluster yang didapat dari penjumlahan seluruh variabel (Johnson dan Wichern, 2007). Dihitung dengan cara:

$$ESS = \sum_{i=1}^n (x_i - \bar{x})^T (x_i - \bar{x})$$

Dimana:

ESS = *Error Sum of Square*

$x_i$  = Vektor objek ke-i berukuran (px1)

$\bar{x}$  = Vektor rata-rata semua objek berukuran (px1)

n = banyaknya objek

### b. Metode Nonhirarki (*Nonhierarchical Method*)

Menurut Hair *et al.* (2010) prosedur nonhirarki tidak melakukan proses struktur mirip pohon (*treelike structure*). Prosedur nonhirarki menempatkan objek ke dalam *cluster* dengan jumlah atau nomor kluster telah ditentukan sebelumnya. Metode nonhirarki yang biasa digunakan adalah metode *sequential threshold*, *parallel threshold*, dan *optimizing partitioning*.

Dalam tugas akhir ini, pada tahap pemilihan prosedur pengklasteran dipilih metode hirarki *average linkage* untuk proses pengklasteran.

### 2.3.5 Menginterpretasikan Kluster

Tahap interpretasi kluster atau kelompok mencakup memeriksa setiap kelompok dari segi variasi kluster dalam hal memberikan nama atau menandai dengan suatu label secara tepat yang dapat menggambarkan sifat dari suatu kluster atau kelompok. Ukuran yang biasa dipakai untuk proses interpretasi yaitu *cluster centroid* (titik tengah dari suatu kluster) (Hair *et al.*, 2010).

### 2.3.6 Validasi dan Pemrofilan Kluster

Validasi meliputi upaya peneliti untuk meyakinkan bahwa kluster atau kelompok yang terbentuk dapat menjelaskan populasi dan dapat digeneralisasikan untuk objek lain serta stabil untuk waktu tertentu. Validasi dapat dilakukan dengan melakukan proses pengklasteran dengan jarak dan metode pengklasteran yang berbeda yang kemudian dibandingkan hasilnya. Pemrofilan kluster meliputi penggambaran karakteristik dari masing-masing kluster untuk menjelaskan perbedaan antar kluster (Hair *et al.*, 2010).

## 3. METODOLOGI PENELITIAN

### 3.1 Sumber Data

Data yang digunakan adalah data hasil pengukuran kualitas udara ambien di kabupaten atau kota di Jawa Tengah tahun 2013. Data yang digunakan merupakan data dari Badan Lingkungan Hidup (BLH) Provinsi Jawa Tengah bidang Pengendalian Pencemaran Lingkungan.

### 3.2 Variabel Data

Data pengukuran kualitas udara ambien dikelompokkan menurut 35 kabupaten/kota di Jawa Tengah berdasarkan dari 3 jenis kawasan yaitu kawasan perumahan, kawasan industri, dan kawasan padat lalu lintas serta diukur dari beberapa jenis udara yang dapat mencemari yaitu: SO<sub>2</sub> (Sulfurdioksida), NO<sub>2</sub> (Nitrogen dioksida), CO (Karbonmonoksida), dan HC (Hidro Karbon).

### 3.3 Metode Analisis Data

Langkah-langkah yang dilakukan dalam pengolahan data adalah sebagai berikut:

1. Melakukan penginputan data jenis kawasan dan data jenis polutan masing-masing kabupaten/kota di Provinsi Jawa Tengah.
2. Melakukan Analisis Kluster

Proses analisis kluster dilakukan terhadap dua jenis data yaitu terhadap jenis kawasan dan terhadap jenis polutan yang masing-masing dilakukan dengan tahapan sebagai berikut:

- a. Melakukan dua uji asumsi analisis kluster yaitu asumsi sampel representatif dan asumsi multikolinearitas.
- b. Melakukan proses pengklasteran dengan menggunakan metode *average linkage*:
  - i. Menghitung jarak *Eucliden* antar kabupaten/kota:

$$d(X_i, X_g) = \left[ \sum_{j=1}^p |X_{ij} - X_{gj}|^2 \right]^{1/2}$$

Dimana:

i, g = 1, 2, ..., 35 menunjukkan kabupaten/kota di Provinsi Jawa Tengah.

j  $\left\{ \begin{array}{l} = 1, 2, 3, \text{ untuk jenis kawasan} \\ \text{atau} \\ = 1, 2, 3, 4, \text{ untuk jenis polutan} \end{array} \right.$

- ii. Membentuk matriks jarak (*proximity matrix*)  $\mathbf{D} = \{d_{ig}\}$ .
- iii. Melakukan proses pengklasteran menggunakan metode *average linkage*:

$$d_{(UV)W} = \frac{\sum_{i=1}^{35} \sum_{g=1}^{35} d_{ig}}{N_{(UV)} N_W}$$

- c. Interpretasi *cluster* atau kluster atau kelompok yang terbentuk untuk memberi nama atau menandai dengan suatu label secara tepat yang dapat menggambarkan sifat dari kluster yang terbentuk.
- d. Validasi dan memprofilkan kluster untuk meyakinkan bahwa kluster atau kelompok yang terbentuk dapat menjelaskan populasi serta kluster yang terbentuk dapat digeneralisasikan untuk objek lain serta stabil untuk waktu tertentu.

## 4. HASIL dan PEMBAHASAN

### 4.1 Deskripsi Data

**Tabel 1** Gambaran Umum Kualitas Udara Ambien Menurut Jenis Kawasan

	Jenis Kawasan		
	Kawasan Perumahan	Kawasan Industri	Kawasan Padat Lalu Lintas
<b>Maksimum</b>	5899,2 $\mu\text{g}/\text{N m}^3$ (Kab. Magelang)	11856,6 $\mu\text{g}/\text{N m}^3$ (Kota Salatiga)	11322 $\mu\text{g}/\text{N m}^3$ (Kota Salatiga)
<b>Minimum</b>	506,81 $\mu\text{g}/\text{N m}^3$ (Kab. Klaten)	634,91 $\mu\text{g}/\text{N m}^3$ (Kab. Purbalingga)	2094,77 $\mu\text{g}/\text{N m}^3$ (Kab. Banyumas)
<b>Rata-rata</b>	2749,07 $\mu\text{g}/\text{N m}^3$	3467,28 $\mu\text{g}/\text{N m}^3$	3830,71 $\mu\text{g}/\text{N m}^3$

**Tabel 2** Gambaran Umum Kualitas Udara Ambien Menurut Jenis Polutan

Jenis Polutan	Maksimum	Minimum	Rata-rata
SO <sub>2</sub>	136,47 $\mu\text{g}/\text{N m}^3$ (Kota Semarang)	75 $\mu\text{g}/\text{N m}^3$ (terdiri dari 24 Kab/Kota)	84,98 $\mu\text{g}/\text{N m}^3$
NO <sub>2</sub>	257,58 $\mu\text{g}/\text{N m}^3$ (Kota Semarang)	30 $\mu\text{g}/\text{N m}^3$ (terdiri dari 5 Kab/Kota)	88,02 $\mu\text{g}/\text{N m}^3$
CO	10810 $\mu\text{g}/\text{N m}^3$ (Kab. Magelang)	1782 $\mu\text{g}/\text{N m}^3$ (Kab. Klaten)	5725,73 $\mu\text{g}/\text{N m}^3$
HC	20095 $\mu\text{g}/\text{N m}^3$ (Kota Salatiga)	1161 $\mu\text{g}/\text{N m}^3$ (Kab. Purbalingga)	3149,90 $\mu\text{g}/\text{N m}^3$

### 4.2 Asumsi Analisis Klaster

Pada Uji sampel yang mewakili menggunakan uji KMO didapat nilai KMO untuk jenis kawasan 0,527 dan untuk jenis polutan 0,505 yang berkisar antara 0,5 sampai 1 maka sampel representatif.

Pada Uji multikolinearitas menggunakan nilai VIF, didapat nilai VIF untuk masing-masing jenis kawasan yaitu 1,036; 1,832; 1,807 dan untuk masing-masing jenis polutan 1,136; 1,444; 1,274; 1,018 yang mana nilai VIF<10 maka tidak terjadi multikolinearitas.

### 4.3 Hasil dan Interpretasi Klaster

Hasil klaster berdasarkan jenis kawasan dan jenis polutan masing-masing menghasilkan 2 klaster dari proses pemotongan dendogram. Untuk jenis kawasan:

1. Klaster 1 terdiri dari 33 kabupaten/kota yaitu Kab. Brebes, Kab. Sukoharjo, Kab. Boyolali, Kab. Kudus, Kota Magelang, Kab. Grobogan, Kab. Demak, Kab. Banjarnegara, Kab. Pekalongan, Kab. Purbalingga, Kab. Wonogiri, Kab. Blora, Kab. Pemalang, Kab. Kebumen, Kab. Wonosobo, Kab. Temanggung, Kab. Batang, Kab. Klaten, Kab. Sragen, Kab. Cilacap, Kab. Kendal, Kab. Pati, Kab. Rembang, Kab. Banyumas, Kab. Purworejo, Kota Semarang, Kota Tegal, Kab. Jepara, Kota Pekalongan, Kab. Semarang, Kab. Tegal, Kota Surakarta, dan Kab. Magelang.
2. Klaster 2 terdiri dari 2 kabupaten/kota yaitu Kota Salatiga, Kab. Karanganyar.



Penamaan klaster dilihat berdasarkan nilai *centroid* klaster dapat diketahui bahwa:

1. Klaster 1 mempunyai nilai *centroid* yang paling tinggi (3524) pada kawasan padat lalu lintas maka klaster 1 diberi nama “kawasan padat lalu lintas”.
2. Klaster 2 mempunyai nilai *centroid* tertinggi (11225) pada kawasan industri sehingga klaster 2 diberi nama “kawasan industri”.

Untuk jenis polutan:

1. Klaster 1 terdiri dari 34 kabupaten/kota yaitu Kab. Kudus, Kab. Rembang, Kab. Purworejo, Kab. Blora, Kab. Pati, Kab. Kebumen, Kab. Banyumas, Kota Tegal, Kota Magelang, Kab. Grobogan, Kab. Boyolali, Kab. Sukoharjo, Kab. Kendal, Kab. Tegal, Kab. Brebes, Kab. Pemalang, Kab. Jepara, Kab. Banjarnegara, Kab. Wonosobo, Kab. Batang, Kab. Temanggung, Kab. Pekalongan, Kab. Cilacap, Kab. Purbalingga, Kab. Demak, Kab. Wonogiri, Kab. Klaten, Kab. Sragen, Kota Surakarta, Kab. Karanganyar, Kota Semarang, Kota Pekalongan, Kab. Magelang, dan Kab. Semarang.
2. Klaster 2 terdiri dari 1 kabupaten/kota yaitu Kota Salatiga.

Penamaan klaster dilihat berdasarkan nilai *centroid* klaster dapat diketahui bahwa:

1. Klaster 1 mempunyai nilai *centroid* tinggi (5731,13) pada jenis polutan CO maka klaster 1 diberi nama “Karbonmonoksida”.
2. Klaster 2 mempunyai nilai *centroid* yang paling tinggi (20095) pada jenis polutan HC sehingga klaster 2 dapat diberi nama “Hidro Karbon”.

#### 4.4 Validasi dan Pemrofilan Klaster

Validasi klaster dilakukan dengan cara melakukan proses pengklasteran kembali terhadap data jenis kawasan dan jenis polutan menggunakan ukuran jarak *squared Euclidean* dan metode pengklasteran *centroid linkage*. Pemrofilan klaster dilakukan secara bersama dengan proses validasi. Hasil dari proses validasi pengklasteran tersebut menghasilkan jumlah klaster yang terbentuk dan anggota dari setiap klaster sama dengan hasil dari proses pengklasteran menggunakan ukuran jarak *Euclidean* dan metode pengklasteran *average linkage* sehingga dapat dikatakan bahwa klaster yang telah terbentuk untuk jenis kawasan dan jenis polutan sudah valid.

## 5. KESIMPULAN

Berdasarkan hasil analisis dapat diambil kesimpulan sebagai berikut:

1. Pengklasteran jenis kawasan menurut kabupaten atau kota Provinsi Jawa Tengah menggunakan *average linkage* didapatkan hasil 2 klaster. Pada klaster 1 terdiri dari 33 kabupaten atau kota dan klaster 2 terdiri dari 2 kabupaten atau kota. Klaster 1 terdiri dari nilai kualitas udara di kawasan perumahan, kawasan industri, dan kawasan padat lalu lintas berkisar antara  $507 \mu\text{g}/\text{N m}^3$  hingga  $6760 \mu\text{g}/\text{N m}^3$  dan mempunyai ciri dengan nilai *centroid* tertinggi sebesar 3524,43 terdapat pada kawasan padat lalu lintas sehingga klaster 1 diberi nama dengan klaster kawasan padat lalu lintas. Hal ini menunjukkan bahwa kualitas udara ambien pada klaster 1 dapat dikatakan dalam kondisi yang baik. Sedangkan klaster 2 terdiri dari nilai kualitas udara ambien di kawasan industri yang tinggi

yaitu untuk Kota Salatiga sebesar  $11856,6 \mu\text{g}/\text{N m}^3$  dan Kabupaten Karanganyar sebesar  $10594,8 \mu\text{g}/\text{N m}^3$  dan mempunyai ciri dengan nilai dengan nilai *centroid* tertinggi sebesar 11225,70 terdapat pada kawasan industri sehingga klaster 2 diberi nama dengan klaster kawasan industri. Hal ini menunjukkan bahwa kualitas udara ambien pada klaster 2 dapat dikatakan dalam kondisi yang buruk.

2. Pengklasteran jenis polutan menurut kabupaten atau kota Provinsi Jawa Tengah menggunakan *average linkage* didapatkan hasil 2 klaster. Pada klaster 1 terdiri dari 34 kabupaten atau kota dan klaster 2 terdiri dari 1 kabupaten atau kota. Klaster 1 terdiri dari nilai kualitas udara pada jenis polutan SO<sub>2</sub>, NO<sub>2</sub>, CO, dan HC berkisar antara  $30 \mu\text{g}/\text{N m}^3$  hingga  $10810 \mu\text{g}/\text{N m}^3$  dan mempunyai ciri dengan nilai *centroid* tertinggi sebesar 5731,13 terdapat pada polutan CO sehingga klaster 1 diberi nama dengan klaster Karbonmonoksida. Hal ini menunjukkan bahwa kualitas udara yang terdapat pada klaster 1 dalam kondisi baik, karena masih dalam ambang batas baku mutu udara ambien. Sedangkan klaster 2 terdiri dari nilai kualitas udara pada jenis polutan HC (Hidro Karbon) sebesar  $20095 \mu\text{g}/\text{N m}^3$  dan mempunyai ciri dengan nilai *centroid* tertinggi sebesar 20095 terdapat pada polutan HC sehingga klaster 2 diberi nama dengan klaster Hidro Karbon. Hal ini menunjukkan bahwa kualitas udara yang terdapat pada klaster 2 dalam kondisi buruk karena telah melebihi nilai ambang batas baku mutu udara ambien untuk HC sebesar  $160 \mu\text{g}/\text{N m}^3$ .

## 6. DAFTAR PUSTAKA

- Dillon, W.R., dan M. Goldstein. 1984. *Multivariate Analysis Methods and Applications*. New York: Wiley.
- Gujarati, D. 1978. *Basic Econometrica*. McGraw-Hill. New York
- Hair, J.F.Jr., R.E. Anderson, R.L. Thatham dan W.C. Black. 2010. *Multivariate Data Analysis Seventh Edition*. New Jersey: Pearson Education, Inc.
- Johnson, R. A. dan Wichern, D. W. 2007. *Applied Multivariate Statistical Analysis. Sixth Edition*. New Jersey : Pearson International Edition.
- Republik Indonesia. 1999. Peraturan Pemerintah No. 41 Tahun 1999 tentang Pengendalian Pencemaran Udara. Lembaran Negara RI Tahun 1999, No. 86. Sekretariat Negara. Jakarta.
- Supranto, J. 2004. *Analisis Multivariat: Arti dan Interpretasi*. Jakarta: Rineka Cipta.
- Widarjono, A. 2010. *Analisis Statistika Multivariat Terapan*. Yogyakarta: UPP STIM YKPN.