
Pengelompokan Kabupaten/Kota Dalam Upaya Kie (Komunikasi, Informasi, dan Edukasi) Untuk Mengurangi Prevalensi Merokok

Athaya Raihan Zaki¹, Aprilia Noviyanti²

¹⁻² Statistika Bisnis, Institut Teknologi Sepuluh Nopember

Email: athayaraihanzaki24@gmail.com

Artikel info

Abstract. *The prevalence of smoking in Indonesia is increasing, especially among children and adolescents. Answering these problems, IEC (Communication, Information, and Education) is currently being pursued to reduce the prevalence of smoking in Indonesia that need special attention for the provincial governments of Lampung, Nusa Tenggara Barat, and Bengkulu. To find out which regional clusters need to strive for IEC, a grouping of districts/cities in the 3 provinces was carried out. Grouping can be done by non-hierarchical cluster analysis with the k-means method. Furthermore, discriminant analysis was carried out to identify smokers of what age were most influential in distinguishing between these cluster groups. This research is expected to be used as a policy consideration to find out which district/city clusters need to strive for IEC. The conclusions obtained from the cluster analysis were 14 members in the first cluster, 10 members in the second cluster and 11 members in the third cluster. The first cluster is a cluster with residents aged 15 years and over who have smoked in the past month higher than clusters 2 and 3 and the age that most affects the grouping is the age of 55 to 64 years with a high classification accuracy.*

Abstrak. Prevalensi merokok di Indonesia semakin meningkat terutama di kalangan usia anak dan remaja. Menjawab permasalahan tersebut, sedang diupayakan KIE (Komunikasi, Informasi, dan Edukasi) untuk mengurangi prevalensi merokok di Indonesia perlu menjadi perhatian khusus bagi pemerintah Provinsi Lampung, Nusa Tenggara Barat dan Bengkulu. Guna mengetahui klaster wilayah mana yang perlu mengupayakan KIE untuk mengurangi prevalensi merokok dilakukan pengelompokan kabupaten/kota di 3 provinsi tersebut. Pengelompokan dapat dilakukan dengan analisis klaster non hierarki dengan metode *k-means*. Selanjutnya, dilakukan analisis diskriminan untuk mengidentifikasi perokok umur berapa yang paling berpengaruh dalam membedakan antar kelompok klaster tersebut. Riset ini diharapkan digunakan sebagai pertimbangan kebijakan untuk mengetahui klaster kabupaten/kota mana yang perlu mengupayakan KIE untuk mengurangi prevalensi merokok sehingga angka kematian akibat merokok dapat ditekan. Kesimpulan yang diperoleh dari analisis

klaster didapatkan 14 anggota pada klaster pertama, 10 anggota pada klaster kedua dan 11 anggota pada klaster ketiga. Klaster pertama merupakan klaster dengan penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir lebih tinggi dibandingkan dengan klaster 2 dan 3 dan usia yang paling berpengaruh terhadap pengelompokan yaitu usia 55 hingga 64 tahun dengan ketepatan klasifikasi yang cukup tinggi.

Keywords:

*Analisis Diskriminan;
Analisis klaster; KIE;
Prevalensi Merokok.*

Corresponden author:

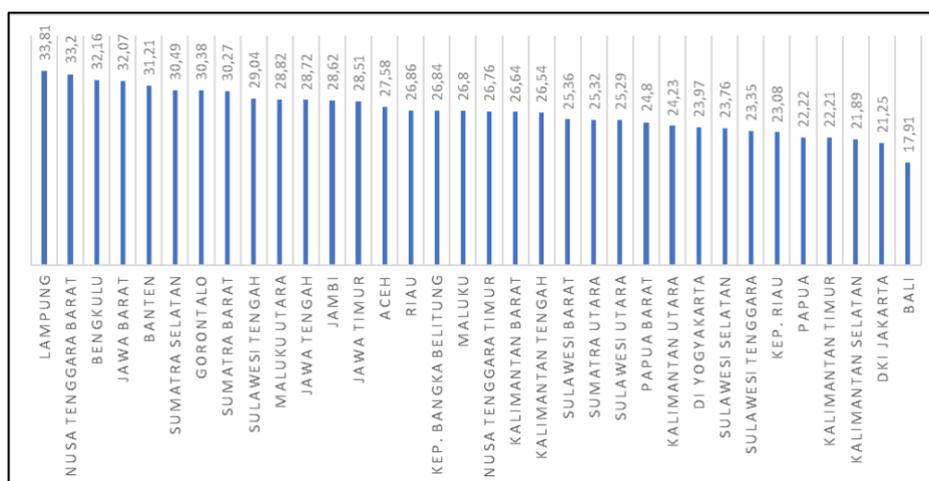
Email: athayaraihanzaki24@gmail.com



artikel dengan akses terbuka di bawah lisensi CC BY -4.0

PENDAHULUAN

Organisasi Kesehatan Dunia (WHO) menyatakan bahwa epidemi tembakau telah membunuh sekitar 8 juta orang setiap tahun. Lebih dari 7 juta kematian diakibatkan oleh perilaku merokok, sementara 1,2 juta kematian diakibatkan karena penyakit kardiovaskular dan gangguan pernapasan karena paparan asap rokok orang lain. Di Indonesia, konsumsi rokok juga menjadi masalah kesehatan masyarakat. Prevalensi merokok di Indonesia semakin meningkat terutama di kalangan usia anak dan remaja. Data Riskesdas (Riset Kesehatan Dasar) tahun 2013 hingga 2018 menunjukkan bahwa prevalensi perokok terjadi peningkatan terutama pada perempuan dan usia lebih muda. Selain itu, peningkatan prevalensi perokok usia 10-18 tahun juga meningkat dari 7,1% pada tahun 2013 menjadi 9,1% pada tahun 2018. Menjawab permasalahan tersebut, dewasa ini sedang diupayakan KIE (Komunikasi, Informasi, dan Edukasi) untuk mengurangi prevalensi merokok di Indonesia. KIE merupakan suatu upaya perubahan sosial yang diorganisasikan dengan baik oleh sekelompok orang (*change agent*) sebagai komunikator dalam jangka pendek maupun panjang dengan tujuan untuk mengubah, mengganti, atau memperkenalkan ide-ide, gagasan, kepercayaan, atau perilaku kepada sekelompok orang (*target adopter*) atau komunikan.



Grafik 1. Persentase Penduduk Usia 15 Tahun ke Atas yang Merokok

Pengupayaan KIE untuk mengurangi prevalensi merokok perlu menjadi perhatian khusus bagi pemerintah khususnya pemerintah Provinsi Lampung, Nusa Tenggara Barat dan Bengkulu. Gambar 1 menunjukkan bahwa Provinsi Lampung, Nusa Tenggara Barat dan Bengkulu merupakan 3 provinsi dengan persentase penduduk usia 15 tahun ke atas yang merokok yang paling tinggi. Pengelompokan kabupaten/kota di 3 provinsi tersebut perlu dilakukan guna mengetahui klaster wilayah mana yang perlu mengupayakan KIE untuk mengurangi prevalensi merokok. Pengelompokan dapat dilakukan

dengan metode statistika yaitu analisis klaster. Analisis klaster atau analisis kelompok merupakan teknik analisa data yang bertujuan untuk mengelompokkan individu atau objek ke dalam beberapa klaster yang memiliki sifat berbeda antar klaster, sehingga individu atau objek yang terletak di dalam satu klaster akan mempunyai sifat relatif homogen (Talakua *et al.*, 2017). Metode yang digunakan dalam analisis klaster pada riset ini yaitu analisis klaster non-hierarki menggunakan metode *k-means* karena dinilai sebagai metode yang sederhana dan umum digunakan. Berdasarkan hasil pemilihan klaster optimum, pada riset ini digunakan jumlah klaster sebanyak 3 klaster. Selanjutnya, dilakukan analisis diskriminan untuk mengidentifikasi perokok umur berapa yang paling penting dalam membedakan antar kelompok klaster tersebut.

Berdasarkan latar belakang yang telah diuraikan, pada riset ini diharapkan dapat mengelompokkan wilayah kabupaten/kota di Provinsi Lampung, Nusa Tenggara Barat dan Bengkulu berdasarkan persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir. Diharapkan hasil riset dapat digunakan sebagai pertimbangan kebijakan untuk mengetahui klaster kabupaten/kota mana yang perlu mengupayakan KIE untuk mengurangi prevalensi merokok sehingga angka kematian akibat merokok dapat ditekan.

Metode

1. Sumber Data

Data yang digunakan dalam riset ini adalah data sekunder yang diperoleh dari publikasi Provinsi Lampung, Nusa Tenggara Barat dan Bengkulu dalam angka 2023 yaitu persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022.

2. Variabel Riset

Variabel yang digunakan pada riset kali ini yaitu persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota yang ditunjukkan sebagai berikut.

Tabel 1. Variabel Riset

Variabel	Keterangan	Skala	Satuan
X ₁	Persentase penduduk usia 15-24 tahun yang merokok dalam sebulan terakhir menurut kabupaten/kota.	Rasio	Persen
X ₂	Persentase penduduk usia 25-34 tahun yang merokok dalam sebulan terakhir menurut kabupaten/kota	Rasio	Persen
X ₃	Persentase penduduk usia 35-44 tahun yang merokok dalam sebulan terakhir menurut kabupaten/kota	Rasio	Persen
X ₄	Persentase penduduk usia 45-54 tahun yang merokok dalam sebulan terakhir menurut kabupaten/kota	Rasio	Persen
X ₅	Persentase penduduk usia 55-64 tahun yang merokok dalam sebulan terakhir menurut kabupaten/kota	Rasio	Persen
X ₆	Persentase penduduk usia 65 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota	Rasio	Persen

3. K-Means Clustering

Analisis klaster atau analisis kelompok merupakan teknik analisa data yang bertujuan untuk mengelompokkan individu atau objek ke dalam beberapa klaster yang memiliki sifat berbeda antar klaster, sehingga individu atau objek yang terletak di dalam satu klaster akan mempunyai sifat relatif

homogen (Talakua, Leleury and Taluta, 2017). *K-means* adalah metode analisis data atau metode *data mining* yang melakukan proses pemodelan tanpa pengawasan, dan merupakan metode pengelompokan data ke dalam sistem partisi. Metode *K-means* mencoba untuk mengelompokkan data yang ada menjadi beberapa kelompok (Zikir, 2022). Proses analisis kluster menggunakan metode *k-means* adalah sebagai berikut (Ningrat, Maruddani and Wuryandari, 2016).

- a. Menentukan k sebagai jumlah kluster yang ingin dibentuk
- b. Mengalokasikan data ke dalam kluster secara acak
- c. Menentukan pusat kluster (*centroid*) dari data yang ada pada masing-masing kluster dengan persamaan sebagai berikut.

$$C_{kj} = \frac{X_{1j} + X_{2j} + \dots + X_{nj}}{n} \quad (1)$$

- d. Menentukan jarak setiap objek dengan setiap *centroid* dengan perhitungan jarak setiap objek dengan setiap *centroid* menggunakan jarak *euclidean*. Rumus jarak *euclidean* adalah sebagai berikut.

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

- e. Mengalokasikan masing-masing data ke *centroid*/rata-rata terdekat
- f. Mengulangi kembali langkah 3-5 sampai tidak ada lagi perpindahan objek.

4. Analisis Diskriminan

Analisis Diskriminan merupakan bentuk lain Regresi Berganda dengan variabel terikat berbentuk non-metrik atau kategori. Analisis Diskriminan adalah teknik *multivariate* yang termasuk *dependence method*, yakni adanya variabel dependen dan *independent*. Dengan demikian, ada variabel yang hasilnya tergantung dari data variabel independen. Ciri khusus adalah data variabel dependen harus dalam bentuk / berupa data kategorikal (*non metric*), sedangkan data independen justru berupa data non kategorikal (*metric*). Tujuan analisis ini adalah untuk mengetahui apakah ada perbedaan yang jelas antar grup pada variabel dependent Atau bisa dikatakan apakah ada perbedaan antara anggota grup 1 dengan anggota grup 2 (Isliko, 2016). Ada dua asumsi utama yang harus dipenuhi pada analisis diskriminan ini, yaitu: sejumlah variabel bebas harus berdistribusi normal, matriks varians-kovarian variabel bebas berukuran $p \times p$ pada kedua kelompok harus sama (Kinansi and Prihatin, 2020). Langkah-langkah dalam analisis diskriminan adalah sebagai berikut ('Adna, 2017).

A. Melakukan uji beda rata-rata kelompok.

Uji beda rata-rata kelompok dalam riset ini menggunakan *wilks' lambda* dan nilai signifikansi. Uji signifikansi statistik dari fungsi diskriminan dapat dilakukan dengan uji *wilks' lambda*, yaitu menguji perbedaan antar kelompok yang diamati dengan variabel independen yang digunakan secara bersama-sama (Kariman, 2016). Statistik *lambda* dari *wilk's* dinyatakan dengan L dengan rumus sebagai berikut.

$$L = \frac{1}{(1+k_1)(1+k_2)\dots(1+k_r)} \quad (3)$$

Jika *wilks' lambda* mendekati angka 0 maka cenderung ada perbedaan kelompok. Keputusan hipotesis dengan nilai signifikansi, yaitu jika signifikansi > 0.05 maka tidak ada perbedaan dalam kelompok dan jika signifikansi < 0.05 maka ada perbedaan dalam kelompok.

B. Kebaikan Fungsi Diskriminan

Uji signifikan diukur dengan menggunakan *eigenvalues*. Pada tabel *eigenvalues* terdapat nilai *canonical correlation*, untuk mengukur derajat hubungan antara hasil diskriminan atau besarnya variabilitas yang mampu diterangkan oleh variabel independen terhadap variabel dependen. Untuk mengetahui tingkat kebaikan fungsi diskriminan yang terbentuk dapat dilakukan dengan menggunakan kuadrat dari *canonical correlation* (CR) atau CR_2 (Hendayanti and Nurhidayati, 2021).

C. Pembentukan Fungsi Diskriminan

Pembuatan model persamaan dengan pembagian berdasarkan kode grup selanjutnya dicari selisihnya, maka model diskriminan telah terbentuk (Rofiq, Wuryandari and Rahmawati, 2016).

D. Perhitungan Centroids

Berdasarkan tabel centroids terdapat dua tipe grup yaitu positif dan negative. Nilai centroids pada tabel menunjukkan besaran Z yang memisahkan kedua grup tersebut (Rofiq, Wuryandari and Rahmawati, 2016).

E. Ketepatan Klasifikasi

Misclassification merupakan kesalahan dari pengklasifikasian suatu observasi baru ke dalam suatu kelompok. Penilaian kinerja dari setiap prosedur klasifikasi adalah dengan cara menghitung tingkat kesalahan atau probabilitas kesalahan klasifikasi. Salah satu metode yang dapat digunakan untuk menghitung probabilitas kesalahan klasifikasi adalah APER (*Apparent Error Rate*) yaitu perhitungan kinerja hasil klasifikasi yang dilakukan dengan menggunakan matriks konfusi (*confusion matrix*) (Rofiq, Wuryandari and Rahmawati, 2016).

Hasil Dan Pembahasan

Hasil analisis kluster, uji asumsi, dan analisis diskriminan persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 adalah sebagai berikut.

1. Analisis Kluster

Jumlah kluster yang akan dibentuk pada metode *k-means* adalah sebanyak 3 kluster. Metode *k-means* dalam pengelompokan persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota dimulai dari pembentukan *initial cluster center* sebagai berikut.

Tabel 2. Initial Cluster Center

Usia	Kluster			Usia	Kluster		
	1	2	3		1	2	3
15-24	24,97	13,96	19,21	45-54	46,58	17,07	33,20
25-34	39,44	29,51	31,24	55-64	41,46	11,31	22,50
35-44	41,82	24,41	35,13	65+	43,03	3,73	14,06

Tabel 2 menunjukkan bahwa kluster 1 memiliki persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir lebih tinggi dibandingkan dengan kluster 2 dan 3. Tahapan selanjutnya adalah melakukan iterasi hingga membentuk 3 kluster dengan karakteristik yang berdekatan sebagai berikut.

Tabel 3. Change in Cluster Centers

Iterasi ke-	Kluster		
	1	2	3
1	14,662	4,980	12,010
2	0,000	0,000	0,000

Tabel 3 menunjukkan bahwa untuk mendapatkan kluster yang dengan karakteristik yang berdekatan sebanyak 3 kluster, diperlukan proses iterasi sebanyak 2 kali. Konvergensi dicapai karena tidak ada atau perubahan kecil di pusat kluster. Perubahan koordinat absolut maksimum untuk pusat

mana pun adalah 0,000 dengan jarak minimum antara pusat awal adalah 25,250. Setelah tahapan iterasi, maka tahapan selanjutnya adalah menetapkan *Final Cluster Centers* sebagai berikut.

Tabel 4. Final Cluster Center

Usia	Klaster			Usia	Klaster		
	1	2	3		1	2	3
15-24	24,12	14,50	19,14	45-54	38,28	18,32	35,83
25-34	39,52	25,07	34,87	55-64	36,75	11,46	29,99
35-44	40,28	25,57	37,98	65+	32,04	5,07	21,80

Tabel 4 menunjukkan bahwa klaster 1 memiliki persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir lebih tinggi dibandingkan dengan klaster 2 dan 3. Anggota klaster pada masing-masing klaster yang terbentuk adalah sebagai berikut.

Tabel 5. Anggota Klaster yang Terbentuk

Klaster 1	Klaster 1	Klaster 2	Klaster 3	Klaster 3
Lampung Barat	Pringsewu	Bengkulu Selatan	Lampung Timur	Dompu
Tanggamus	Mesuji	Rejang Lebong	Lampung Utara	Bima
Lampung Selatan	Tulang Bawang	Bengkulu Utara	Pesisir Barat	Sumbawa Barat
Lampung Tengah	Lombok Barat	Kaur	Bandar Lampung	Kota Mataram
Way Kanan	Lombok Tengah	Seluma	Metro	Kota Bima
Tulang Bawang Pesawaran	Lombok Timur Lombok Utara	Mukomuko Lebong Kepahiang Bengkulu Tengah Kota Bengkulu	Sumbawa	

Tabel 5 menunjukkan bahwa terdapat 14 anggota pada klaster pertama, 10 anggota pada klaster kedua dan 11 anggota pada klaster ketiga.

2. Uji Asumsi

Pada analisis klaster dan diskriminan, asumsi distribusi normal multivariat dan homogenitas matriks varians-kovarian harus terpenuhi. Pengujian asumsi distribusi normal multivariat dan homogenitas matriks varians-kovarian dalam analisis klaster dan analisis diskriminan persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 adalah sebagai berikut.

A. Uji Asumsi Berdistribusi Normal Multivariat

Uji asumsi distribusi normal multivariat persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 adalah sebagai berikut.

Hipotesis:

H_0 : Persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 berdistribusi normal multivariat

H_1 : Persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 tidak berdistribusi normal multivariat

Taraf signifikan : 0,05

Daerah penolakan : Tolak H_0 jika $P\text{-Value} > 0,05$

Statistik uji : Nilai korelasi (ρ_{xy}) dari *mahalanobis distance* dan nilai kuantil cukup tinggi yaitu sebesar 0,982 yang mendekati nilai 1 dan diperkuat dengan $P\text{-Value}$ sebesar 0,000 yang kurang dari 0,05 maka diperoleh keputusan gagal tolak H_0 yang artinya persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 berdistribusi normal multivariat.

B. Uji Homogenitas Varians-Kovarian

Uji homogenitas varians-kovarian faktor yang memengaruhi hasil kluster persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 adalah sebagai berikut.

Hipotesis:

H_0 : $\Sigma_1 = \Sigma_2 = \Sigma_3$ (Matriks varians kovarian pada faktor yang memengaruhi hasil kluster homogen)

H_1 : Minimal satu $\Sigma_i \neq 0$ (Matriks varians kovarian pada faktor yang memengaruhi hasil kluster heterogen)

Taraf signifikan : 0,05

Daerah penolakan : Tolak H_0 jika $F_{hitung} \leq F_{0,05(42;2598)}$ atau $p\text{-value} < 0,05$

Statistik uji : Dari hasil uji homogenitas varians didapatkan nilai F_{hitung} sebesar 3,231 lebih besar dari nilai sebesar 1,388 sehingga diputuskan gagal tolak yang artinya matriks varians-kovarians pada faktor yang memengaruhi hasil kluster persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 homogen.

3. Analisis Diskriminan

Analisis diskriminan digunakan untuk mengetahui apakah ada perbedaan antara kelompok pada variabel respons dan variabel mana yang berpengaruh signifikan terhadap perbedaan tersebut. Setelah itu, dilanjutkan dengan membuat fungsi atau model diskriminan.

A. Uji Beda Rata-Rata Kluster Persentase Penduduk Usia 15 Tahun ke Atas yang Merokok Dalam Sebulan Terakhir Menurut Kabupaten/Kota Tahun 2022

Uji perbedaan faktor yang memengaruhi hasil kluster persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 dilakukan dengan hipotesis sebagai berikut.

Hipotesis :

a. Pengujian perbedaan rata-rata persentase penduduk usia 15-24 tahun yang merokok antara kluster 1, kluster 2 dan kluster 3.

H_0 : Tidak terdapat perbedaan rata-rata persentase penduduk usia 15-24 tahun yang merokok antara kluster 1, kluster 2 dan kluster 3

H_1 : Terdapat perbedaan rata-rata persentase penduduk usia 15-24 tahun yang merokok antara kluster 1, kluster 2 dan kluster 3

b. Pengujian perbedaan rata-rata persentase penduduk usia 25-34 tahun yang merokok antara kluster 1, kluster 2 dan kluster 3.

- H₀: Tidak terdapat perbedaan rata-rata persentase penduduk usia 25-34 tahun yang merokok antara klaster 1, klaster 2 dan klaster 3
- H₁: Terdapat perbedaan rata-rata persentase penduduk usia 25-34 tahun yang merokok antara klaster 1, klaster 2 dan klaster 3
- c. Pengujian perbedaan rata-rata persentase penduduk usia 35-44 tahun yang merokok antara klaster 1, klaster 2 dan klaster 3.
- H₀: Tidak terdapat perbedaan rata-rata persentase penduduk usia 35-44 tahun yang merokok antara klaster 1, klaster 2 dan klaster 3
- H₁: Terdapat perbedaan rata-rata persentase penduduk usia 35-44 tahun yang merokok antara klaster 1, klaster 2 dan klaster 3
- d. Pengujian perbedaan rata-rata persentase penduduk usia 45-54 tahun yang merokok antara klaster 1, klaster 2 dan klaster 3.
- H₀: Tidak terdapat perbedaan rata-rata persentase penduduk usia 45-54 tahun yang merokok antara klaster 1, klaster 2 dan klaster 3
- H₁: Terdapat perbedaan rata-rata persentase penduduk usia 45-54 tahun yang merokok antara klaster 1, klaster 2 dan klaster 3
- e. Pengujian perbedaan rata-rata persentase penduduk usia 55-64 tahun yang merokok antara klaster 1, klaster 2 dan klaster 3.
- H₀: Tidak terdapat perbedaan rata-rata persentase penduduk usia 55-64 tahun yang merokok antara klaster 1, klaster 2 dan klaster 3
- H₁: Terdapat perbedaan rata-rata persentase penduduk usia 55-64 tahun yang merokok antara klaster 1, klaster 2 dan klaster 3
- f. Pengujian perbedaan rata-rata persentase penduduk usia 65 tahun ke atas yang merokok antara klaster 1, klaster 2 dan klaster 3.
- H₀: Tidak terdapat perbedaan rata-rata persentase penduduk usia 65 tahun ke atas yang merokok antara klaster 1, klaster 2 dan klaster 3
- H₁: Terdapat perbedaan rata-rata persentase penduduk usia 65 tahun ke atas yang merokok antara klaster 1, klaster 2 dan klaster 3
- Taraf signifikan : 0,05
- Daerah penolakan : Tolak H₀ jika F_{hitung} > F_{0,05;(2;32)} atau P-value < 0,05
- Statistik uji :

Tabel 6. Statistik Uji Beda

Variabel	F _{hitung}	F _{0,05;(2;32)}	p-value	Variabel	F _{hitung}	F _{0,05;(2;32)}	p-value
15-24	37,366		0,000	45-54	148,946		0,000
25-34	68,474	3,294	0,000	55-64	152,323	3,294	0,000
35-44	83,263		0,000	65+	151,554		0,000

Tabel 6. menunjukkan bahwa pada setiap usia didapatkan F_{hitung} lebih dari F_{0,05;(2;32)} dan p-value kurang dari 0,05 sehingga diputuskan tolak H₀ yang artinya terdapat perbedaan rata-rata persentase penduduk pada tiap usia yang merokok antara klaster 1, klaster 2 dan klaster 3.

B. Kebaikan Model Klaster Persentase Penduduk Usia 15 Tahun ke Atas yang Merokok Dalam Sebulan Terakhir Menurut Kabupaten/Kota Tahun 2022

Tabel *eigen value* digunakan untuk berapa jumlah fungsi diskriminan yang terbentuk pada faktor yang memengaruhi hasil klaster persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 yang ditunjukkan pada Tabel 7 sebagai berikut.

Tabel 7. *Kebaikan Model Klaster Persentase Penduduk Usia 15 Tahun ke Atas yang Merokok Dalam Sebulan Terakhir*

Fungsi ke-	<i>Eigen Value</i>	%Varians	<i>Canonical Correlation</i>	R ²
1	19,037	95,0	0,975	95,0%
2	0,997	5,0	0,707	49,9%

Tabel 7 menunjukkan bahwa 95% keragaman mampu dijelaskan oleh model sebesar 95,00% pada fungsi ke-1 dan 5% keragaman mampu dijelaskan oleh model sebesar 49,90% pada fungsi ke-2, sisanya dijelaskan oleh variabel lain di luar model.

C. Tingkat Signifikansi Model Diskriminan

Wilk's Lambda digunakan untuk mengindikasikan tingkat signifikansi dari fungsi diskriminan. Nilai *Wilk's Lambda* yang terbentuk pada klaster persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 pada fungsi ke 1 hingga 2 didapatkan nilai *Wilk's Lambda* sebesar 0,025 dan diperkuat dengan *p-value* sebesar 0,000. Sehingga, perbedaan antar kelompok klaster persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 berbeda signifikan.

D. Tingkat Kepentingan Variabel Prediktor

Tingkat kepentingan usia dalam memengaruhi klaster persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 ditunjukkan pada Tabel 8 sebagai berikut.

Tabel 8. *Tingkat Kepentingan Variabel Prediktor*

Usia	Fungsi 1	Usia	Fungsi 1
15-24	0,106	45-54	0,350
25-34	0,264	55-64	0,408
35-44	0,092	65+	0,373

Berdasarkan tabel 8 dapat diketahui bahwa pada fungsi 1 usia 55-64 tahun merupakan variabel yang paling berpengaruh terhadap klaster persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022.

E. Tabel Koefisien Klasifikasi Fungsi Diskriminan

Tabel koefisien kanonik fungsi diskriminan digunakan untuk membentuk persamaan fungsi diskriminan pada faktor yang memengaruhi hasil klaster persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 sebagai berikut.

Tabel 9. Tabel Koefisien Klasifikasi

Koefisien	Fungsi 1	Koefisien	Fungsi 1
15-24	0,039	45-54	0,119
25-34	0,088	55-64	0,115
35-44	0,032	65+	0,100
		(Constant)	-13,931

Berdasarkan tabel 9 diketahui model yang terbentuk pada fungsi 1 adalah $Y = -13,931 + 0,039X_1 + 0,088X_2 + 0,032X_3 + 0,119X_4 + 0,115X_5 + 0,100X_6$.

F. Tabel Function at Group Centroids

Tabel *function at group centroids* yang terbentuk pada faktor yang memengaruhi hasil kluster persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 adalah sebagai berikut.

Tabel 10. Tabel Function at Group Centroids

Kluster	Fungsi 1
1	3,752
2	-6,336
3	0,985

Tabel 10 menunjukkan bahwa jika hasil prediksi pada fungsi 1 mendekati nilai 3,752 maka dapat dikelompokkan pada kluster 1. Jika hasil prediksi pada fungsi 1 mendekati nilai -6,336 maka dapat dikelompokkan pada kluster 2 dan jika hasil prediksi pada fungsi 1 mendekati nilai 0,985 maka dapat dikelompokkan pada kluster 3.

G. Ketepatan Klasifikasi Faktor yang Memengaruhi Hasil Kluster

Ketepatan klasifikasi dalam model diskriminan pada faktor yang memengaruhi hasil kluster persentase penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir menurut kabupaten/kota tahun 2022 adalah sebagai berikut.

Tabel 11. Ketepatan Klasifikasi Faktor yang Memengaruhi Hasil Kluster Persentase Penduduk Usia 15 Tahun ke Atas Yang Merokok Dalam Sebulan Terakhir

Klasifikasi	Prediksi		
	Kluster 1	Kluster 2	Kluster 3
Kluster 1	14	0	0
Kluster 2	0	10	0
Kluster 3	1	0	10

Tabel 11 menunjukkan bahwa kabupaten/kota yang diklasifikasikan pada kluster 1 yang diprediksikan benar sebanyak 14 kabupaten/kota. Kabupaten/kota yang diklasifikasikan pada kluster

2 yang diprediksi benar adalah 10 kabupaten/kota. Kabupaten/kota yang diklasifikasikan pada klaster 3 yang diprediksi benar adalah 10 kabupaten/kota. Sehingga, persentase kesalahan dalam mengklasifikasikan objek atau yang sering dikenal dengan APER (*Apparent Error Rate*) sebesar 2,9%.

Simpulan Dan Saran

Kesimpulan yang diperoleh dari analisis klaster didapatkan 14 anggota pada klaster pertama, 10 anggota pada klaster kedua dan 11 anggota pada klaster ketiga. Klaster pertama merupakan klaster dengan penduduk usia 15 tahun ke atas yang merokok dalam sebulan terakhir lebih tinggi dibandingkan dengan klaster 2 dan 3 dan usia yang paling berpengaruh terhadap pengelompokan yaitu usia 55 hingga 64 tahun dengan ketepatan klasifikasi yang cukup tinggi. Saran untuk pemerintah Provinsi Lampung, Nusa Tenggara Barat dan Bengkulu untuk dapat mengupayakan KIE untuk mengurangi prevalensi merokok pada klaster 1 guna menekan angka kematian akibat merokok.

Daftar Rujukan

- 'Adna, S.F. (2017) 'ANALISIS DISKRIMINAN UNTUK MEMPREDIKSI KELULUSAN NILAI AKHIR MAHASISWA', *Delta: Jurnal Ilmiah Pendidikan Matematika*, 5(1), p. 17. Available at: <https://doi.org/10.31941/delta.v5i1.390>.
- Dewi, R.A. (2015) *PENGELOMPOKAN KABUPATEN/KOTA DI JAWA TIMUR BERDASARKAN INDIKATOR KEMISKINAN MENGGUNAKAN METODE C-MEANS DAN FUZZY C-MEANS CLUSTERING*. Thesis SS141501. Surabaya: Institut Teknologi Sepuluh Nopember.
- Hendayanti, N.P.N. and Nurhidayati, M. (2021) 'KLASIFIKASI TINGKAT KEPARAHAN KEMISKINAN PROVINSI DI INDONESIA DENGAN ANALISIS DISKRIMINAN', *Math Educa Journal*, 5(1), pp. 14–21. Available at: <https://doi.org/10.15548/mej.v5i1.2510>.
- Isliko, T.W.A. (2016) 'KAJIAN ANALISIS DISKRIMINAN MENGUKUR LOYALITAS PELANGGAN TOKO BUKU SUCI KUPANG', *JOURNAL OF MANAGEMENT*, 3.
- Kariman, R. (2016) *Prediksi Kondisi Financial Distress Dengan Menggunakan Multiple Discriminant Analysis Pada Perusahaan Yang Terdaftar Di Bursa Efek Indonesia*.
- Kinansi, R.R. and Prihatin, M.T. (2020) 'Analisis Diskriminan Fisher untuk Hubungan Incidence Rate dengan Indeks Entomologi berdasarkan Klasifikasi Ekosistem', *Buletin Riset Sistem Kesehatan*, 22(4), pp. 282–291. Available at: <https://doi.org/10.22435/hsr.v22i4.2151>.
- Ningrat, D.R., Maruddani, D.A.I. and Wuryandari, T. (2016) 'ANALISIS CLUSTER DENGAN ALGORITMA K-MEANS DAN FUZZY C-MEANS CLUSTERING UNTUK PENGELOMPOKAN DATA OBLIGASI KORPORASI', 5(4).
- Rofiq, A., Wuryandari, T. and Rahmawati, R. (2016) 'PERBANDINGAN ANALISIS DISKRIMINAN FISHER DAN NAIVE BAYES UNTUK KLASIFIKASI RISIKO KREDIT (Studi Kasus Debitur di Koperasi Jateng Amanah Mandiri Cabang Sukorejo Kendal)', 5(1).
- Talakua, M.W., Leleury, Z.A. and Taluta, A.W. (2017) 'ANALISIS CLUSTER DENGAN MENGGUNAKAN METODE K-MEANS UNTUK PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI MALUKU BERDASARKAN INDIKATOR INDEKS PEMBANGUNAN MANUSIA TAHUN 2014', *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 11(2), pp. 119–128. Available at: <https://doi.org/10.30598/barekengvol11iss2pp119-128>.
- Zikir, A. (2022) 'Perbandingan Metode Clustering Dengan Menggunakan Metode Average Linkage Dan Metode K-Means Pada Industri Kecil Dan Menengah Di Kabupaten Wajo', (2).

