



Penerapan model *based-clustering* kriteria *Integrated Completed Likelihood* (ICL) untuk mengelompokkan kabupaten/kota di Provinsi Jambi berdasarkan indikator pendidikan

MAWADDAH KHAIRANI, DIAN CAHYAWATI*, OKI DWIPURWANI, DAN HERLINA HANUM

Jurusan Matematika, Fakultas Matematika & Ilmu Pengetahuan Alam, Universitas Sriwijaya

Kata kunci:

indikator pendidikan,
Integrated Completed Likelihood,
model *based-clustering*

ABSTRAK: Kondisi objektif mutu pendidikan di Provinsi Jambi masih memiliki banyak kelemahan yang harus diperbaiki dan memiliki tingkat partisipasi pendidikan yang berbeda-beda pada tiap wilayah. Tujuan penelitian ini adalah mengelompokkan wilayah di Provinsi Jambi untuk mendapatkan kemiripan karakteristik antar kelompok berdasarkan indikator pendidikan pada setiap level pendidikan yaitu tingkat pendidikan dasar, menengah, dan tinggi. Metode pengelompokan yang diterapkan adalah metode *Based-Clustering t* multivariat kriteria *Integrated Completed Likelihood* (ICL). Data indikator pendidikan dari 11 wilayah di Provinsi Jambi yang dianalisis adalah Angka Partisipasi Kasar (APK), Angka Partisipasi Sekolah (APS), Angka Partisipasi Murni (APM), Angka Melek Huruf (AMH), dan Rata-rata Lama Belajar (RLB). Data diperoleh dari laman resmi publikasi BPS Provinsi Jambi. Hasil penelitian menunjukkan bahwa pengelompokan 11 wilayah untuk tingkat pendidikan dasar hanya dikelompokkan menjadi satu kelompok. Setiap wilayah di Provinsi Jambi memiliki kualitas pendidikan yang tidak jauh berbeda. Pengelompokan wilayah untuk tingkat pendidikan menengah dan pendidikan tinggi menghasilkan dua kelompok. Kedua kelompok yang dihasilkan memiliki kualitas pendidikan yang berbeda, Kedua kelompok tersebut memberikan perbedaan yang signifikan berdasarkan hasil uji perbedaan rata-rata dengan *Manova*. Kelompok wilayah yang memiliki nilai rata-rata dari indikator pendidikan lebih tinggi memiliki kualitas pendidikan yang lebih baik dibandingkan dengan kelompok wilayah yang memiliki nilai rata-rata dari indikator pendidikan yang lebih rendah.

Keywords:

education indicators,
Integrated Completed Likelihood,
based-clustering model

ABSTRACT: The objective condition of the quality of education in Jambi Province still has many weaknesses that must be improved and has different levels of educational participation in each region. The purpose of this research is to cluster regions in Jambi Province to obtain similar characteristics between groups based on educational indicators at each educational level, namely primary, middle and higher education levels. The clustering method applied is the *Based-Clustering t* multivariate *Integrated Completed Likelihood* (ICL) criteria. Data on education indicators from 11 regions in Jambi Province that were analyzed were Gross Enrollment Ratio (GER), School Participation Rate (SPR), Net Enrollment Rate (NER), Literacy Rate (LR), and Average Years of Schooling (AYS). Data was obtained from the official website of the Jambi Province BPS publication. The results show that the grouping of 11 regions for the primary education level is only grouped into one group. Each region in Jambi Province has a similar quality of education. Grouping areas for middle and higher education levels resulted in two groups. The two groups produced have different educational quality, both groups provide significant differences based on the results of the *Manova* mean difference test. The group of areas that has a higher average value of education indicators has a better quality of education than the group of areas that has a lower average value of education indicators.

1 PENDAHULUAN

Analisis *cluster* adalah metode untuk mengelompokkan satuan objek pengamatan ke dalam beberapa kelompok berdasarkan variabel yang diamati [1]. Metode *cluster* dengan basis ukuran jarak

terdiri dari metode *cluster* berhierarki dan metode *cluster* tak berhierarki [2]. Metode *cluster* yang memperhatikan aspek statistik yaitu metode *cluster* berbasis model (model *based-clustering*). Model *based-clustering* mengasumsikan data berasal dari distribusi campuran (*mixture distribution*) [3]. Asumsi

* Corresponding Author: dianc_mipa@unsri.ac.id

ini mengarah kepada sebuah model probabilitas matematika data yaitu *finite mixture model* dan sebuah kelompok yang berbeda ini diwakili oleh setiap komponen pada *mixture model*. *Mixture model* memiliki dua proses utama yaitu estimasi parameter yang dilakukan dengan *Maximum Likelihood* (ML) dan pemilihan model terbaik yang dipilih berdasarkan ukuran kecocokan model seperti *Bayesian Information Criterion* (BIC), *Integrated Completed Likelihood* (ICL), dan *Minimum Message Length* (MML). Model *based-clustering* pertama kali digunakan oleh Banfield dan Raftery untuk mengelompokkan objek dalam populasi [3].

Model *based-clustering* sudah digunakan pada beberapa penelitian, salah satunya pada penelitian yang dilakukan oleh [4]. Hasil dari penelitian yang dilakukan oleh [4], menunjukkan bahwa model *based-clustering* memberikan struktur yang konsisten pada kelompok yang terbentuk dan menghasilkan tingkat akurasi pengelompokan yang lebih tinggi dibandingkan metode *K-means*. Persentase kesalahan pengelompokan pada metode *K-means* lebih besar daripada model *based-clustering* yaitu 10,67% dan 3,33%. Penelitian dengan model *based-clustering* juga dilakukan oleh [1] yang melakukan pengelompokan provinsi di Indonesia berdasarkan faktor penyebaran *Covid-19* dengan kriteria ICL. [5] melakukan penelitian untuk mengelompokkan saham berdasarkan rasio keuangan dengan model *based-clustering*. Penelitian tentang pengelompokan kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator pendidikan dilakukan oleh [6] menggunakan metode *clustering hierarki*.

Berdasarkan Data Pokok Pendidikan, saat ini mutu pendidikan di Provinsi Jambi pada berbagai tingkatan masih menunjukkan banyak kekurangan, antara lain terdapat 3.786 sekolah yang belum memenuhi standar nasional dan 1.952 sekolah belum terakreditasi oleh BAN-SM/M, banyak guru yang belum memenuhi standar (sarjana) dan belum sesuai dengan bidang studi yang diajar, program wajib belajar 9 tahun belum selesai sepenuhnya. Kekurangan tersebut perlu segera diperbaiki dan ditingkatkan melalui pendekatan yang sistematis dan terencana. Tinggi rendahnya mutu pendidikan diukur dari indikator pendidikan. Indikator pendidikan menurut UNESCO yang telah dihitung oleh BPS dan Kementerian Pendidikan Nasional antara lain, Angka Melek Huruf (AMH), Angka Partisipasi Murni (APM), Angka Partisipasi Kasar (APK), Angka Putus Sekolah, Angka Partisipasi Sekolah (APS), Angka Mengulang, Rasio Murid Guru, dan Rata-rata Lama Belajar (RLB).

Penelitian sebelumnya mengenai pengelompokan wilayah berdasarkan indikator pendidikan menggunakan analisis *cluster* hierarki metode *single linkage*, *average linkage*, dan *complete linkage*. Penelitian yang akan dilakukan yaitu mengelompokkan wilayah di Provinsi Jambi berdasarkan indikator pendidikan APK, APS, APM, AMH, dan RLB menurut tingkat dasar, menengah, dan tinggi pada tahun 2022. Wilayah Provinsi Jambi terdiri dari 11 kabupaten/kota. Kondisi indikator pendidikan yang beragam antar wilayah menyebabkan data cenderung mengandung *outlier* dan tidak berdistribusi normal.

2 METODE

Pengelompokan sebelas wilayah di Provinsi Jambi berdasarkan indikator pendidikan diterapkan dengan metode *based-clustering* distribusi *t* multivariat kriteria *Integrated Completed Likelihood* (ICL). Level pendidikan dibedakan pada tingkat pendidikan dasar, menengah, dan pendidikan tinggi. Metode distribusi campuran *t* multivariat digunakan untuk menyelesaikan data yang memuat *outlier*, menghasilkan dugaan [7].

[7] yang dianalisis adalah indikator-indikator pendidikan pada tahun 2022 untuk setiap wilayah di Provinsi Jambi. Sebelas wilayah tersebut adalah Kerinci, Merangin, Sarolangun, Batanghari, Muaro Jambi, Tanjung Jabung Timur, Tanjung Jabung Barat, Tebo, Bungo, Kota Jambi, dan Kota Sungai Penuh. Sumber data diperoleh dari hasil publikasi BPS (Badan Pusat Statistik). Alat bantu yang digunakan adalah *software R studio Build 463*.

Analisis Data

Variabel yang dianalisis dalam mengelompokkan wilayah di Provinsi Jambi untuk masing-masing tingkat pendidikan dituliskan pada Tabel 1.

1. Pengelompokan wilayah di Provinsi Jambi berdasarkan tingkat pendidikan dasar menggunakan variabel $X_1, X_2, X_5, X_8, X_9, X_{12}$, dan X_{13} .
2. Pengelompokan berdasarkan tingkat pendidikan menengah menggunakan variabel X_3, X_6, X_{10}, X_{12} , dan X_{13} .
3. Pengelompokan wilayah di Provinsi Jambi berdasarkan tingkat pendidikan tinggi menggunakan variabel X_4, X_7, X_{11}, X_{12} , dan X_{13} .

Tabel 1. Variabel Penelitian Tingkat Pendidikan Dasar, Menengah, dan Tinggi

| Variabel dalam persen (%) | Indikator Pendidikan | X | Penggunaan Variabel pada Tingkat Pendidikan | | |
|---------------------------------|------------------------|-----------------|---|----------|--------|
| | | | Dasar | Menengah | Tinggi |
| Angka Partisipasi Kasar (APK) | APK SD/ sederajat | X ₁ | ✓ | | |
| | APK SMP/ sederajat | X ₂ | ✓ | | |
| | APK SM/ sederajat | X ₃ | | ✓ | |
| | APK PT | X ₄ | | | ✓ |
| Angka Partisipasi Sekolah (APS) | APS Usia 13 – 15 Tahun | X ₅ | ✓ | | |
| | APS Usia 16 – 18 Tahun | X ₆ | | ✓ | |
| | APS Usia 19 – 23 Tahun | X ₇ | | | ✓ |
| Angka Partisipasi Murni (APM) | APM SD/ sederajat | X ₈ | ✓ | | |
| | APM SMP/ sederajat | X ₉ | ✓ | | |
| | APM SM/ sederajat | X ₁₀ | | ✓ | |
| | APM PT | X ₁₁ | | | ✓ |
| Angka Melek Huruf (AMH) | AMH Usia > 15 Tahun | X ₁₂ | ✓ | ✓ | ✓ |
| Rata-rata Lama Belajar (RLB) | RLB | X ₁₃ | ✓ | ✓ | ✓ |

Langkah-langkah menerapkan metode *based-clustering* kriteria *Integrated Completed Likelihood* (ICL) adalah sebagai berikut.

1. Pendeteksian keberadaan *outlier* multivariat menggunakan ukuran jarak Mahalanobis dan jarak *Robust*. Hasil pendeteksian disebut data *outlier* jika nilai akar kuadrat jarak Mahalanobis lebih besar dari $X^2_{p;1-\frac{\alpha}{2}}$ dan nilai jarak *Robust* lebih besar dari $\sqrt{X^2_{p;1-\frac{\alpha}{2}}}$.

Jarak Mahalanobis:

$$MD_i = [(\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})]^{\frac{1}{2}},$$

$$i = 1, 2, \dots, n, n = 11$$

Jarak *Robust*:

$$RD_i = [(\mathbf{x}_i - \mathbf{T})^T \mathbf{C}^{-1} (\mathbf{x}_i - \mathbf{T})]^{\frac{1}{2}},$$

$$i = 1, 2, \dots, n, n = 11$$

2. Pengujian normalitas multivariat dilakukan dengan uji mardia multivariat *skewness* dan *kurtosis*. Data disebut berdistribusi normal jika $b_{1,p} = 0$ dan $b_{2,p} = p(p + 2)$ atau jika nilai $p_{value} > \alpha = 0,05$.

Ukuran multivariat *skewness* ($b_{1,p}$):

$$b_{1,p} = \frac{1}{n^2} \left[\sum_{i,j=1}^n \{(\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_j - \bar{\mathbf{x}})\}^3 \right]$$

Ukuran multivariat *kurtosis* ($b_{2,p}$):

$$b_{2,p} = \frac{1}{n} \left[\sum_{i,j=1}^n \{(\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_j - \bar{\mathbf{x}})\}^2 \right]$$

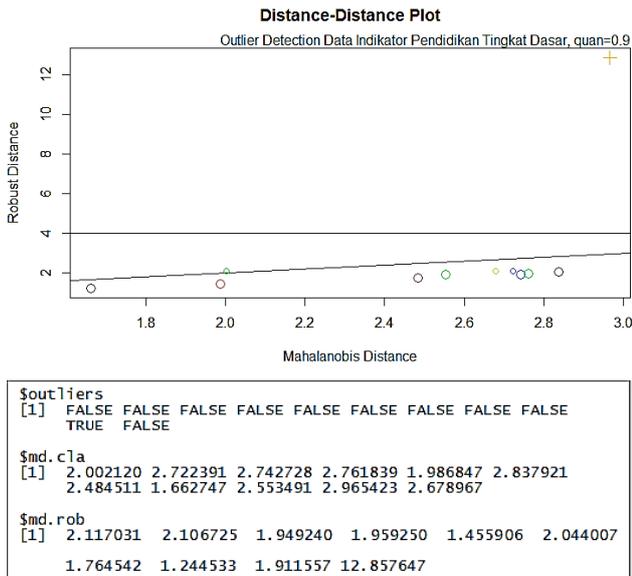
3. Pengelompokan wilayah di Provinsi Jambi menggunakan model *based-clustering* dengan kriteria *Integrated Completed Likelihood* (ICL) menggunakan *package teigen software R*.
4. Pemeriksaan kelompok optimal yang dipilih dari nilai ICL terbesar.
5. Pengujian perbedaan rata-rata antar kelompok menggunakan uji Manova dengan kriteria *Wilk's Lambda* $\Lambda = \frac{|W|}{|W+B|}$. Jika $F_{hitung} > F_{tabel}$ atau $p_{value} < \alpha(0,05)$ artinya terdapat perbedaan antar kelompok yang terbentuk.

3 HASIL DAN PEMBAHASAN

Pengelompokan dilakukan dengan menggunakan model *based-clustering*. Kriteria yang digunakan dalam pemilihan model terbaik adalah *Integrated Completed Likelihood* (ICL) yang ditentukan dari nilai ICL terbesar.

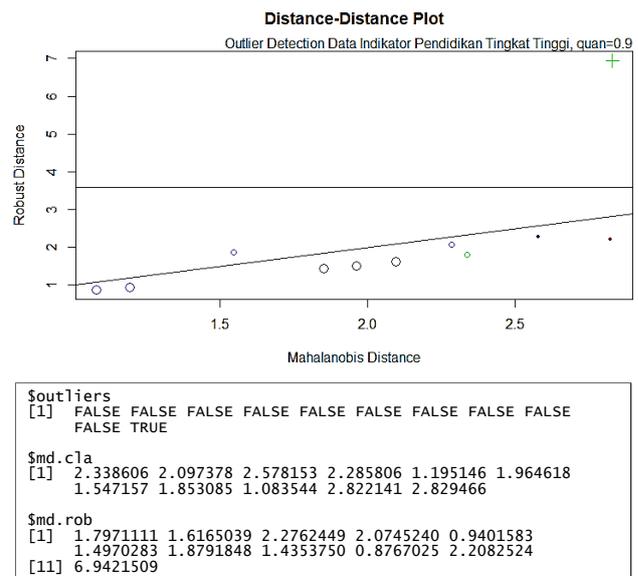
3.1 Deteksi Outlier Multivariat

Langkah awal dalam mendeteksi *outlier* multivariat yaitu menghitung jarak Mahalanobis dan jarak *Robust* untuk seluruh objek pengamatan. Pendeteksian *outlier* multivariat bertujuan untuk menguji asumsi bahwa data indikator pendidikan di Provinsi Jambi cenderung mengandung *outlier*. Setelah itu, melakukan perbandingan dengan kuantil dari distribusi $X_{p;0.975}$ sebagai nilai *cut off*. Nilai *cut off* data indikator pendidikan tiap tingkatan dapat berbeda-beda, tergantung jumlah variabel (p) yang dimiliki masing-masing tingkat pendidikan. Pendeteksian *outlier* multivariat menggunakan fungsi *dd-plot* pada *software R package mvoutlier*. Jika data tidak mengandung *outlier*, maka semua titik akan berada di sekitar garis lurus yang merupakan perpotongan dari jarak Mahalanobis dan jarak *Robust*. Berikut hasil plot jarak Mahalanobis terhadap jarak *Robust* pada data indikator pendidikan untuk tingkat pendidikan dasar, menengah, dan tinggi.



Gambar 1. *dd-plot* Tingkat Pendidikan Dasar

Gambar 1 menyajikan plot jarak Mahalanobis terhadap jarak *Robust* data indikator pendidikan untuk tingkat pendidikan dasar menggunakan fungsi *distance-distance plot (dd-plot)*. Titik-titik pada plot menggambarkan wilayah di Provinsi Jambi, titik yang terletak sangat menyimpang dari garis menunjukkan bahwa wilayah tersebut adalah outlier. Pada tingkat pendidikan dasar wilayah yang menjadi *outlier* adalah Kota Jambi.



Gambar 3. *dd-plot* Tingkat Pendidikan Tinggi

Gambar 3 menyajikan plot jarak Mahalanobis terhadap jarak *Robust* data indikator pendidikan untuk tingkat pendidikan tinggi menggunakan fungsi *dd-plot*. Berdasarkan hasil plot jarak Mahalanobis terhadap jarak *Robust*, *outlier* pada data indikator pendidikan untuk tingkat pendidikan tinggi adalah Kota Sungai Penuh.

3.2 Uji Normalitas Multivariat

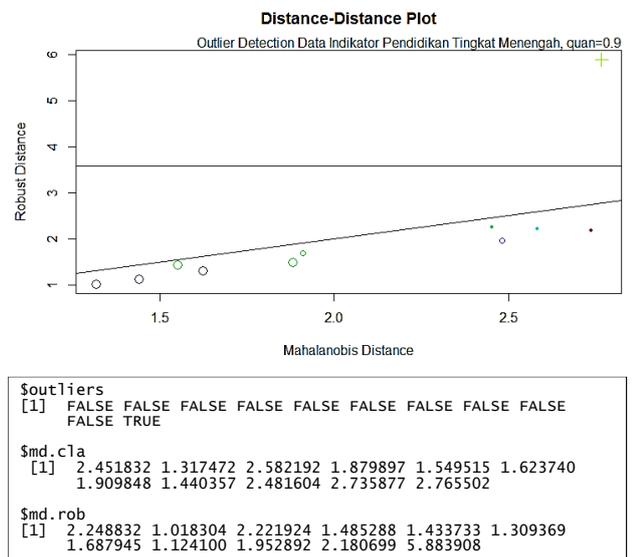
Pengujian normalitas multivariat menggunakan bantuan *software R package psych* dengan melakukan uji mardia multivariat *skewness* dan *kurtosis* dengan hipotesis uji:

H_0 : Data berdistribusi normal multivariat.

H_1 : Data tidak berdistribusi normal multivariat.

Keputusan hipotesis uji tolak H_0 terjadi jika $pvalue < \alpha = 0,05$. Selain ditentukan oleh nilai $pvalue$, keputusan untuk tolak H_0 juga dapat berdasarkan nilai kemencengan atau *skewness* ($b_{1,p}$) > 0 dan nilai kelandaian atau *kurtosis* ($b_{1,p}$) > $p(p + 2)$, p menunjukkan banyaknya variabel. Berdasarkan dari *output mardia*, dapat diketahui *skewness* data indikator pendidikan untuk tingkat pendidikan dasar memiliki nilai *probability* ($pvalue$) = 1 (lebih besar dari $\alpha = 0,05$). *Small sample skewness* memiliki nilai *probability* = 0,92 (lebih besar dari $\alpha = 0,05$) sehingga didapatkan keputusan hipotesis gagal tolak H_0 artinya data indikator pendidikan untuk tingkat pendidikan dasar berdistribusi normal multivariat.

Hasil uji *mardia test* untuk tingkat pendidikan menengah menghasilkan nilai *probability skewness* = 0,94 dan nilai *probability small sample skewness* =



Gambar 2. *dd-plot* Tingkat Pendidikan Menengah

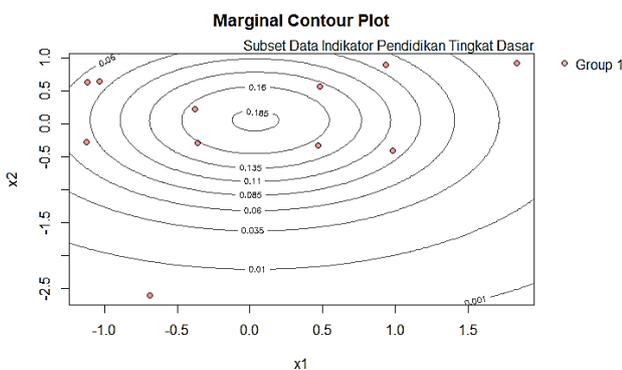
Gambar 2 menyajikan plot jarak Mahalanobis terhadap jarak *Robust* data indikator pendidikan untuk tingkat pendidikan menengah menggunakan fungsi *dd-plot*. Berdasarkan letak titik yang ditunjukkan pada plot, terdapat 1 wilayah yang menjadi *outlier* pada indikator pendidikan tingkat menengah yaitu Kota Sungai Penuh.

0,63, kedua nilai *probability* tersebut lebih besar dari $\alpha = 0,05$, sehingga keputusan hasil uji yaitu gagal tolak H_0 yang artinya data indikator pendidikan untuk tingkat pendidikan menengah berdistribusi normal multivariat.

Pada uji *mardia test* untuk tingkat pendidikan tinggi diperoleh hasil nilai *probability skewness* = 0,92 dan nilai *probability small sample skewness* = 0,55, kedua nilai *probability* tersebut lebih besar dari $\alpha = 0,05$. Keputusan hasil uji yaitu gagal tolak H_0 yang artinya data indikator pendidikan untuk tingkat pendidikan tinggi berdistribusi normal multivariat. Setelah dilakukan uji *mardia test*, didapatkan kesimpulan bahwa data indikator pendidikan untuk tingkat pendidikan dasar, menengah, dan tinggi berdistribusi normal dan pengelompokan dengan menggunakan model *based-clustering* distribusi *t* multivariat tetap dapat diaplikasikan karena terdapat *outlier* didalam data indikator pendidikan untuk tingkat pendidikan dasar, menengah, dan tinggi.

3.3 Pengelompokan Wilayah di Provinsi Jambi untuk Tingkat Pendidikan Dasar

Pengelompokan dilakukan dengan model *based-clustering* kriteria ICL menggunakan *software R package teigen*. Pengelompokan wilayah di Provinsi Jambi berdasarkan indikator pendidikan untuk tingkat pendidikan dasar memiliki *marginal contour plot* sebagai berikut.



Gambar 4. *Marginal Contour Plot* Tingkat Pendidikan Dasar

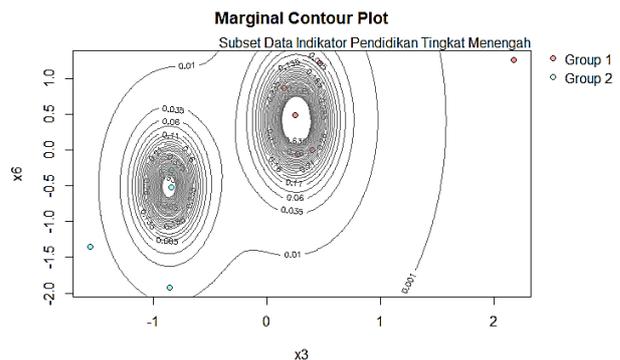
Berdasarkan Gambar 4, model terbaik dari pengelompokan adalah UIIU dengan jumlah kelompok optimal yaitu satu dan nilai ICL terbesar -232,2622. Pemilihan terhadap variabel untuk visualisasi kontur dilakukan dengan menggabungkan variabel X_1 dan X_2 . Variabel X_1 adalah APK SD/ sederajat dan variabel X_2 merupakan APK SMP/ sederajat. Pengelompokan ini hanya menghasilkan satu kelompok optimal, sehingga kelompok ini terdiri dari 11 wilayah yang ada di Provinsi Jambi

yaitu Kerinci, Merangin, Sarolangun, Batanghari, Muaro Jambi, Tanjung Jabung Timur, Tanjung Jabung Barat, Tebo, Bungo, Kota Jambi, Kota Sungai Penuh.

Pada pengelompokan wilayah untuk tingkat pendidikan dasar hanya menghasilkan satu kelompok wilayah, sehingga anggota kelompok terdiri dari 11 wilayah kabupaten/kota yang ada di Provinsi Jambi dan memiliki karakteristik pendidikan yang tidak jauh berbeda antar wilayah. Nilai rata-rata APK SD/ sederajat, APS usia 13-15 tahun, APM SD/ sederajat, APM SMP/ sederajat, dan AMH usia > 15 tahun yang dimiliki kelompok wilayah lebih tinggi daripada nilai rata-rata yang dimiliki Indonesia. Namun, pada nilai APK SMP/ sederajat dan RLB, masih di bawah nilai rata-rata Indonesia. Nilai APK, APS, serta APM yang tinggi menunjukkan bahwa Provinsi Jambi memiliki akses pendidikan pada tingkat pendidikan dasar yang baik, ketersediaan pendidikan yang lebih baik, kemampuan membaca dan menulis lebih tinggi serta lebih sedikit siswa yang meninggalkan pendidikan pada tingkat tertentu dan lebih banyak menyelesaikan tingkat pendidikan mereka. Nilai RLB yang rendah menunjukkan terdapat sejumlah siswa yang tidak menyelesaikan pendidikan mereka.

3.4 Pengelompokan Wilayah di Provinsi Jambi untuk Tingkat Pendidikan Menengah

Pengelompokan wilayah di Provinsi Jambi berdasarkan indikator pendidikan untuk tingkat pendidikan menengah memiliki *marginal contour plot* sebagai berikut.



Gambar 5. *Marginal Contour Plot* Tingkat Pendidikan Menengah

Berdasarkan Gambar 5, model terbaik yang dihasilkan adalah CICC dengan jumlah kelompok optimal yaitu dua dan nilai ICL terbesar -140,6367. Variabel yang dipilih untuk visualisasi kontur dilakukan dengan cara menggabungkan dua dari lima variabel yang digunakan. Variabel X_3 adalah APK SM/ sederajat dan X_6 merupakan APS usia 16-18 tahun. Dua

variabel tersebut memegang peran penting dalam pengelompokan pada tingkat menengah, dimana gabungan yang terbentuk cenderung menciptakan jumlah kontur yang sama. Titik yang berwarna merah mewakili anggota kelompok 1, dan titik yang berwarna hijau mewakili anggota kelompok 2. Pengelompokan ini menghasilkan dua kelompok optimal, untuk anggota tiap kelompok secara lebih jelas ditampilkan pada Tabel 3.

Tabel 2. Anggota Kelompok Wilayah di Provinsi Jambi untuk Tingkat Pendidikan Menengah

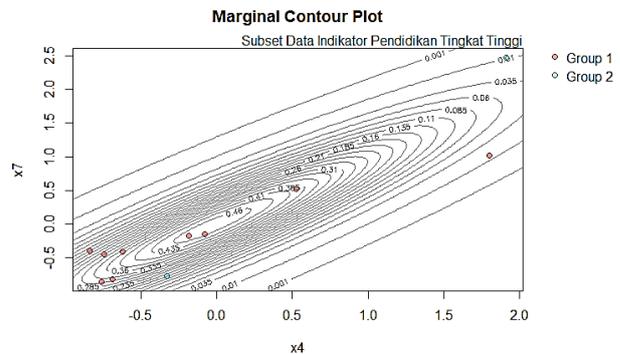
| Kelompok 1 | Kelompok 2 |
|-------------------------|-------------------------|
| 1. Merangin | 1. Kerinci |
| 2. Tanjung Jabung Timur | 2. Sarolangun |
| 3. Tebo | 3. Batanghari |
| 4. Bungo | 4. Muaro Jambi |
| | 5. Tanjung Jabung Barat |
| | 6. Kota Jambi |
| | 7. Kota Sungai Penuh |

Tabel 2. merinci anggota dari kelompok wilayah pada pengelompokan berdasarkan indikator pendidikan tingkat menengah. Pada pengelompokan wilayah untuk tingkat pendidikan menengah menghasilkan dua kelompok wilayah. Nilai rata-rata tiap indikator yang dimiliki kelompok 2 lebih tinggi daripada nilai rata-rata kelompok 1 dan Indonesia. Kelompok 2 merupakan wilayah dengan karakteristik aksesibilitas pendidikan yang baik, dimana sebagian besar populasi usia 16 – 18 tahun yang memenuhi syarat dapat mengakses pendidikan pada tingkat SM/ sederajat. Banyaknya peserta didik yang terlibat, ada potensi peningkatan kapasitas manusia suatu wilayah, pendidikan yang luas dapat membantu meningkatkan keterampilan dan pengetahuan masyarakat. Wilayah yang memiliki nilai RLB rendah menunjukkan bahwa aksesibilitas pendidikan di wilayah tersebut relatif tidak baik dan juga mencerminkan efisiensi sistem pendidikan di wilayah tersebut.

3.5 Pengelompokan Wilayah di Provinsi Jambi untuk Tingkat Pendidikan Tinggi

Pengelompokan wilayah di Provinsi Jambi berdasarkan indikator pendidikan untuk tingkat pendidikan tinggi memiliki *marginal contour plot* seperti Gambar 6. Berdasarkan Gambar 6 yang menggambarkan visualisasi anggota kelompok yang terbentuk, variabel yang dikombinasikan yaitu variabel X_4 dan X_7 . Variabel X_4 merupakan APK Perguruan Tinggi dan variabel X_7 merupakan APS Usia 19 – 23 Tahun. Model terbaik yang dihasilkan adalah CCC dengan jumlah kelompok optimal yaitu dua dan nilai ICL terbesar -102,5511. Titik yang berwarna merah mewakili anggota kelompok 1, dan titik yang berwarna

hijau mewakili anggota kelompok 2. Pengelompokan ini menghasilkan dua kelompok optimal, untuk anggota tiap kelompok secara lebih jelas ditampilkan pada Tabel 3.



Gambar 6. *Marginal Contour Plot* Tingkat Pendidikan Tinggi

Tabel 3. Anggota Kelompok Wilayah di Provinsi Jambi untuk Tingkat Pendidikan Tinggi

| Kelompok 1 | Kelompok 2 |
|-------------------------|----------------------|
| 1. Kerinci | 1. Kota Jambi |
| 2. Merangin | 2. Kota Sungai Penuh |
| 3. Sarolangun | |
| 4. Batanghari | |
| 5. Muaro Jambi | |
| 6. Tanjung Jabung Timur | |
| 7. Tanjung Jabung Barat | |
| 8. Tebo | |
| 9. Bungo | |

Tabel 3 merinci anggota dari kelompok wilayah pada pengelompokan berdasarkan indikator pendidikan untuk tingkat pendidikan tinggi. Pengelompokan wilayah untuk tingkat pendidikan tinggi menghasilkan dua kelompok wilayah. Kelompok 1 terdiri dari 9 kabupaten dan kelompok 2 terdiri dari 2 kota. Nilai rata-rata yang dimiliki kelompok 2 lebih tinggi daripada kelompok 1 dan Indonesia, dapat dilihat bahwa terdapat kesenjangan antara wilayah kabupaten dan kota. APK PT yang rendah dapat mengindikasikan bahwa jumlah perguruan tinggi pada wilayah tersebut sedikit atau bahkan tidak ada dan hanya sebagian kecil dari populasi usia yang memenuhi syarat di wilayah tersebut yang mendaftar atau terlibat di pendidikan tingkat tinggi, yang disebabkan oleh keterbatasan akses. Tingkat kesadaran pendidikan tingkat tinggi yang tidak optimal di masyarakat. APK PT yang rendah juga dapat mencerminkan kualitas pendidikan tingkat menengah yang tidak memadai dalam memberikan persiapan yang diperlukan untuk melanjutkan ke perguruan tinggi.

3.6 Uji Perbedaan Rata-rata antar Kelompok

Pengujian perbedaan rata-rata menggunakan uji Manova dengan bantuan *software* SPSS. Berikut adalah hipotesis untuk uji perbedaan rata-rata antar kelompok.

$H_0 : \tau_1 = \tau_2 = 0$ (tidak terdapat perbedaan antar kelompok)

$H_1 : \tau_i \neq \tau_j$ untuk $i \neq j$ (terdapat perbedaan antar kelompok)

Keputusan tolak H_0 jika $F_{hitung} > F_{tabel}$ atau $p_{value} < \alpha(0,05)$ artinya terdapat perbedaan antar kelompok yang terbentuk. Berikut adalah hasil pengujian perbedaan rata-rata antar kelompok yang terbentuk.

Table 4. Hasil Uji Perbedaan Rata-rata antar Kelompok (Uji Manova dengan Statistik Uji *Wilk's Lambda*)

| Education Indicators | Wilk's Lambda Value | F | pvalue |
|------------------------|---------------------|--------|--------|
| Middle Education Level | 0,081 | 11,325 | 0,009 |
| Higher Education Level | 0,061 | 15,494 | 0,005 |

Statistik *Wilk's Lambda* memiliki nilai p_{value} lebih kecil dari nilai $\alpha = 0,05$ untuk semua data indikator pendidikan, artinya penolakan H_0 menyatakan adanya perbedaan vektor *mean* antar kelompok di masing-masing data. Oleh karena itu, analisis kelompok untuk wilayah di Provinsi Jambi berdasarkan indikator pendidikan dapat dilakukan.

4 KESIMPULAN

Hasil pengelompokan wilayah kabupaten/kota di Provinsi Jambi berdasarkan indikator pendidikan untuk tingkat pendidikan dasar menghasilkan satu kelompok wilayah dengan jumlah anggota kelompoknya 11 wilayah, yang artinya kualitas pendidikan pada tingkat pendidikan dasar tiap wilayah di Provinsi Jambi tidak jauh berbeda. Nilai APK, APS, serta APM SD/ sederajat yang tinggi menunjukkan bahwa Provinsi Jambi memiliki akses pendidikan pada tingkat pendidikan dasar yang baik, ketersediaan pendidikan yang lebih baik. Pada hasil pengelompokan untuk tingkat pendidikan menengah menghasilkan dua kelompok wilayah, kelompok 1 terdiri dari 4 wilayah dan kelompok 2 terdiri dari 7 wilayah. Kelompok 2 dengan nilai APK, APS, serta APM SMP/ sederajat dan SM/ sederajat yang tinggi memiliki aksesibilitas pendidikan yang baik, dimana sebagian besar populasi usia 16-18 tahun yang memenuhi syarat dapat mengakses pendidikan pada tingkat SM/ sederajat.

Hasil pengelompokan wilayah di Provinsi Jambi untuk tingkat pendidikan tinggi menghasilkan 2 ke-

lompok, kelompok 1 terdiri dari 9 wilayah dan kelompok 2 terdiri dari 2 wilayah. Kelompok 2 yang beranggotakan wilayah kota memiliki akses pendidikan yang baik pada perguruan tinggi, dengan jumlah perguruan tinggi pada wilayah tersebut banyak dan sebagian besar dari populasi usia yang memenuhi syarat di wilayah tersebut dapat mendaftar atau terlibat di pendidikan tingkat tinggi. Hasil pengelompokan wilayah di Provinsi Jambi berdasarkan indikator pendidikan dapat menjadi pertimbangan bagi pemerintah Provinsi Jambi untuk melakukan penanganan pada pembangunan dan pemerataan kualitas pendidikan sesuai dengan tingkat kondisi pendidikan pada wilayah di Provinsi Jambi yang berbeda-beda.

REFERENSI

- [1] Hamidah, N., Santoso, R., & Rusgijono, A. (2022). Klasterisasi Provinsi di Indonesia Berdasarkan Faktor Penyebaran Covid-19 Menggunakan Model-Based Clustering t-Multivariat. *Jurnal Gaussian*, 10(1), 56–66.
- [2] Sumertajaya, I.M., Erfiani., & Putri, W.D.Y. (2007). Analisis Gerombol Menggunakan Metode Two Step Cluster. *Forum Statistika dan Komputasi*, 12(1), 18–23.
- [3] Agustini, M. (2017). Model Based-Clustering dengan Distribusi t Multivariat Menggunakan Kriteria Integrated Completed Likelihood dan Minimum Message Length. *Tesis*. Institut Teknologi Sepuluh Nopember.
- [4] Pardede, T. (2007). Perbandingan Metode Model-Based dengan Metode K-Mean dalam Analisis Cluster. *Jurnal Matematika, Sains, dan Teknologi*, 8(2), 98–108.
- [5] Hasnida, I.S.D., & Kusumawati, R. (2023). Penerapan Model-Based Clustering pada Pengelompokan Saham Berdasarkan Rasio Keuangan. *Jurnal Aplikasi Statistika & Komputasi Statistik*, 15(1).
- [6] Nafkiyah, D., Rifatin, L., Rozikin, M.R., et al. (2022). Analisis Cluster dalam Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Pendidikan. *Jurnal Ilmiah Matematika dan Pendidikan Matematika*, 12(1).
- [7] McLachlan, G., & Peel, D. (2000). *Finite Mixture Models*. New York: John Wiley & Sons.
- [8] Supranto, J. (2004). *Analisis Multivariat: Arti dan Interpretasi*. Cetakan 1. Jakarta: PT. Rineka Cipta.
- [9] Jain, A.K. (2010). Data Clustering: 50 years Beyond K-Means. *Pattern Recognition Letters*, 31(8), 651–666.
- [10] Johnson, R.A., & Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis*. 6th edition. New Jersey: Pearson Education, Inc.
- [11] Andrews, J.L., McNicholas, P.D., & Subedi, S. (2011). Model-Based Classification via Mixtures of Multivariate t-Distributions. *Computational Statistics and Data Analysis*, 55(1), 520–529.
- [12] Casella, G., & Berger, R.L. (2002). *Statistical Inference*. 2nd edition. Pacific Grove: Thomson Learning Inc.

- [13] Cain, M., Zhang, Z., & Yuan, K. (2017). Univariate and Multivariate Skewness and Kurtosis for Measuring Nonnormality: Prevalence, Influence, and Estimation. *Behavior Research Methods*, 49(5), 1716–1735.
- [14] Kassambara, A. (2017). *Practical Guide To Cluster Analysis in R: Unsupervised Machine Learning (Multivariate Analysis I)*. 1st edition. STHDA.
- [15] Aldini, U., & Pramesti, W. (2020). Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Mutu Pendidikan Sekolah Menengah Pertama Tahun 2016-2018 Menggunakan Model *Based-Clustering*. *J Statistika*, 13(2), 23–38.
- [16] Fraley, C., & Raftery, A.E. (2002). Model-Based Clustering, Discriminant Analysis, and Density Estimation. *Journal of the American Statistical Association*, 97(458), 611–631.
- [17] Biernacki, C., Celeux, G., & Govaert, G. (2000). *Assessing a Mixture Model for Clustering with the Integrated Classification Likelihood*. France: Institut National De Recherche En Informatique Et En Automatique.
- [18] Baudry, J.P., Cardoso, M., Celeux, G., et al. (2013). Enhancing the Selection of a Model-Based Clustering with External Categorical Variables. *Economics at Your Fingertips*, 9(2), 177–196.
- [19] Badan Pusat Statistik Provinsi Jambi. (2022). *Statistik Daerah Provinsi Jambi 2022*. BPS Provinsi Jambi. Jambi.
- [20] Badan Pusat Statistik Provinsi Jambi. (2022). *Statistik Pendidikan Provinsi Jambi 2022*. BPS Provinsi Jambi. Jambi.
- [21] Badan Pusat Statistik Republik Indonesia. (2022). *Statistik Pendidikan Indonesia 2022*. BPS RI. Jakarta.
- [22] Dinas Statistik Kesejahteraan Rakyat. (2010). *Profil Indikator Pendidikan di Indonesia (Laporan Sosial Indonesia 2009)*. Badan Pusat Statistik. Jakarta.
- [23] Hammouda, K., & Karray, F. (2000). A Comparative Study of Data Clustering Techniques. *Course Project SYDE 625*.

LAMPIRAN

```
> #Syntax Model Based-Clustering Kriteria ICL dengan Software R
> #Tingkat Pendidikan Dasar
> #Memanggil Data Indikator Pendidikan
> library(readxl)
> Data=read_excel("D:/Data Penelitian.xlsx")
> Data=data.frame(Data[,2:14])

> #Membentuk Kombinasi Data Indikator Pendidikan untuk Tingkat Pendidikan Dasar
> Datadasar=as.matrix(Data[,c(1,2,5,8,9,12,13)])

> #Mendeteksi Outlier Data Indikator Pendidikan untuk Tingkat Pendidikan Dasar
> library(mvoutlier)
> dd.plot(Datadasar, quan=0.9, alpha=0.05)
> mtext(text="Outlier Detection Data Indikator Pendidikan untuk
Tingkat Pendidikan Dasar, quan=0.9", side=3, adj=1)

> #Uji Normalitas Multivariat Data Indikator Pendidikan untuk Tingkat Pendidikan Dasar
> library(psych)
> mardia(Datadasar)

> #Klasifikasi dan Estimasi Parameter Cluster Optimal dengan Kriteria ICL
> library(teigen)
> teigen(Datadasar)$iclresult

> #Menampilkan Anggota Cluster Model UIIU dengan Jumlah Cluster G=1
> teigen(Datadasar)
> tDatadasar_UIIU <- teigen(Datadasar, models="UIIU", parallel.cores=2, Gs=1)
> tDatadasar_UIIU$iclresults$classification

> #Marginal Contour Plot Data Indikator Pendidikan untuk Tingkat Pendidikan Dasar
> tDatadasar <- teigen(Datadasar, models="UIIU", parallel.cores=2, Gs=1, init="hard")
> plot(tDatadasar, what="contour")
> mtext(text="Data Indikator Pendidikan untuk Tingkat
Pendidikan Dasar", side=3, adj=1)
```