

Penentuan Lokasi Baru Kantor Otoritas Jasa Keuangan (OJK) di Provinsi Sumatera Utara dengan Metode Kombinasi Analisis Faktor dan Analisis Kluster Berbasis Fuzzy

Santi Puteri Rahayu^{1*}, Mutiara Avista Candra Dewi Lasahido¹, dan Hendra Budi Kusuma¹

¹ Departemen Statistika, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

*Corresponding author: santi_pr@statistika.its.ac.id

Received: 23 March 2022

Accepted: 31 March 2022

Published: 15 April 2022

ABSTRAK – Mengingat pentingnya keberadaan kantor Otoritas Jasa Keuangan (OJK) untuk mendukung perekonomian di wilayah Sumatera Utara, maka perlu adanya penambahan kantor di lokasi baru. Penelitian ini, bertujuan untuk memperoleh hasil penentuan lokasi baru kantor OJK di Provinsi Sumatera Utara dengan menerapkan dua metode Analisis Kluster, yaitu metode Fuzzy C-Means (FCM) dan Fuzzy Gustafson-Kessel (FGK), yang secara umum merupakan suatu teknik pengelompokan observasi yang mempertimbangkan sifat keanggotaan fuzzy dalam suatu kelompok sebagai dasar pembobotan. Kedua metode Analisis Kluster mengelompokkan seluruh Kabupaten/Kota di Sumatera Utara berdasarkan variabel terpilih hasil Analisis Faktor diantara lima variabel, yaitu Produk Domestik Regional Bruto (PDRB), Pengeluaran Pemerintah, Jumlah Penduduk, Indeks Pembangunan Manusia (IPM), dan Piutang Perusahaan Pembiayaan. Berdasarkan hasil Analisis Kluster terbaik FCM dengan empat kelompok optimum dan dua variabel terpilih (PDRB dan IPM), maka Kabupaten/Kota di Provinsi Sumatera Utara yang disarankan menjadi lokasi kantor baru adalah Kabupaten Deli Serdang dan Simalungun, serta Kota Pematang Siantar. Selain itu, sebagai Kluster Kabupaten/Kota di Provinsi Sumatera Utara yang memiliki rata-rata PDRB dan IPM tertinggi, Kota Medan dikonfirmasi layak telah memiliki Kantor OJK.

Keywords – Faktor, Analisis Kluster Fuzzy C-Means, Analisis Kluster Fuzzy Gustafson-Kessel, Lokasi Kantor OJK.

I. PENDAHULUAN

Sebagai ilmu yang digunakan untuk mengumpulkan, menganalisis, menginterpretasikan, serta memberikan kesimpulan terhadap data, Statistika memiliki berbagai macam metode analisis yang diantaranya adalah Analisis Faktor dan Analisis Kluster. Pada penelitian ini, metode Analisis Kluster yang digunakan adalah Fuzzy C-Means (FCM) dan Fuzzy Gustafson-Kessel (FGK). FCM dan FGK merupakan suatu teknik pengelompokan dengan mempertimbangkan sifat keanggotaan fuzzy dalam suatu kelompok sebagai dasar pembobotan [1]. Analisis Faktor dan Analisis Kluster dapat digunakan untuk menentukan lokasi baru kantor Otoritas Jasa Keuangan (OJK) di Provinsi Sumatera Utara. Analisis faktor dapat mengetahui variabel terpilih yang dapat mewakili variabel yang lain dan analisis kluster dapat mengetahui kluster optimum sesuai dengan bentuk struktur kluster yang digunakan sebagai rekomendasi di provinsi Sumatera Utara.

Pulau Sumatera merupakan wilayah yang memiliki peranan tertinggi kedua setelah Pulau Jawa dalam pembentukan Produk Domestik Bruto (PDB) Nasional hingga mencapai 22,03% pada Tahun 2016 [2]. Sedangkan di dalam Pulau Sumatera, Provinsi Sumatera Utara merupakan provinsi dengan nilai Produk Domestik Regional Bruto (PDRB) berdasarkan harga konstan 2010 pada tahun 2016 tertinggi yaitu sebesar 463.775,85 miliar rupiah [3].

OJK merupakan lembaga yang bertugas sebagai pengawas industri jasa keuangan yang terpercaya, melindungi kepentingan konsumen dan masyarakat, dan mampu mewujudkan industri jasa keuangan menjadi pilar perekonomian nasional yang berdaya saing global serta dapat memajukan kesejahteraan [4].

Mengingat pentingnya keberadaan kantor OJK untuk mendukung perekonomian di wilayah Sumatera Utara, maka perlu adanya penambahan kantor di lokasi baru. Lokasi baru kantor OJK akan direkomendasikan melalui Analisis Kluster FCM dan FGK yang akan mengelompokkan seluruh Kabupaten/Kota di Sumatera Utara berdasarkan beberapa variabel terpilih hasil Analisis Faktor diantara lima variabel yaitu PDRB, Pengeluaran Pemerintah, Jumlah Penduduk, Indeks Pembangunan Manusia (IPM), dan Piutang Perusahaan Pembiayaan. Penelitian dilakukan pada 33 objek penelitian yang merupakan Kabupaten/Kota di Provinsi Sumatera Utara.

II. TINJAUAN PUSTAKA

A. Pemeriksaan Asumsi Kecukupan Korelasi

Pemeriksaan asumsi kecukupan korelasi antar variabel dapat diidentifikasi dengan nilai *Kaiser-Meyer-Olkin* (KMO) berikut.

$$KMO = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} q_{ij}^2} \quad (1)$$

dimana:

$r_{ij}^2 =$ kuadrat elemen dari \mathbf{R}

$q_{ij}^2 =$ kuadrat elemen dari \mathbf{Q} ; $\mathbf{Q} = \mathbf{DR}^{-1}\mathbf{D}$ dan $\mathbf{D} = \left[\text{diag } \mathbf{R}^{-1} \right]^{1/2}$

Nilai KMO yang lebih tinggi adalah sesuatu yang diharapkan pada suatu analisis. Secara umum nilai dari pemeriksaan diharapkan lebih besar dari 0,80, namun jika nilai KMO sebesar 0,5 maka hal tersebut masih diperbolehkan [5].

B. Pemeriksaan Asumsi Multivariat Normal

Distribusi multivariat normal dapat diuji menggunakan *Mardia's MVN Test* dengan (2) berikut.

$$\hat{\gamma}_{1,p} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n m_{ij}^3; \hat{\gamma}_{2,p} = \frac{1}{n} \sum_{i=1}^n m_{ii}^2 \tag{2}$$

dimana $m_{ij} = (x_i - \bar{x})' S^{-1} (x_j - \bar{x})$ adalah jarak *Mahalanobis* kuadrat, dan p adalah jumlah variabel [6].

C. Pemeriksaan Asumsi Dependensi Variabel

Uji *Bartlett* digunakan untuk mengetahui korelasi antar variabel. Untuk menguji kebebasan antar variabel ini dapat dilakukan uji *Bartlett sphericity* [7].

Hipotesis:

$H_0: \rho = \mathbf{I}$ (tidak ada korelasi antar variabel)

$H_1: \rho \neq \mathbf{I}$ (ada korelasi antar variabel)

Statistik Uji:

$$\chi_{hitung}^2 = - \left\{ n - 1 - \frac{2p + 5}{6} \right\} \ln |\mathbf{R}| \tag{3}$$

dimana:

n = banyak observasi

p = banyak variabel

$|\mathbf{R}|$ = nilai determinan matriks korelasi

Daerah kritis: tolak H_0 jika $\chi_{hitung}^2 > \chi_{\alpha; (\frac{1}{2}p(p-1))}^2$.

Matriks korelasi dapat dihitung dengan menggunakan rumus berikut [8].

$$r_{ik} = \frac{s_{ik}}{\sqrt{s_{ii}} \sqrt{s_{kk}}} = \frac{\sum_{j=1}^n (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k)}{\sqrt{\sum_{j=1}^n (x_{ji} - \bar{x}_i)^2} \sqrt{\sum_{j=1}^n (x_{jk} - \bar{x}_k)^2}} \tag{4}$$

Untuk $i = 1, 2, \dots, p$ dan $k = 1, 2, \dots, p$ dimana $r_{ik} = r_{ki}$ untuk semua i dan k .

D. Analisis Faktor

Analisis faktor dilakukan untuk melihat kedekatan antar variabel yang diduga memiliki korelasi sehingga dapat dikelompokkan dalam satu grup yang disebut sebuah faktor. Adapun model dari analisis faktor adalah sebagai berikut.

$$\begin{aligned} X_1 - \mu_1 &= \ell_{11}F_1 + \ell_{12}F_2 + \dots + \ell_{1m}F_m + \varepsilon_1 \\ &\dots \\ X_p - \mu_p &= \ell_{p1}F_1 + \ell_{p2}F_2 + \dots + \ell_{pm}F_m + \varepsilon_p \end{aligned} \tag{5}$$

atau dapat ditulis dengan notasi matriks berikut.

$$\mathbf{X}_{pxl} = \boldsymbol{\mu}_{(pxl)} + \mathbf{L}_{(pxm)} \mathbf{F}_{(mxl)} + \boldsymbol{\varepsilon}_{pxl} \tag{6}$$

dimana:

X_i = komponen utama ke- i

μ_i = rata-rata variabel ke- i

ε_i = faktor spesifik ke- i

F_j = *common factor* ke- j

ℓ_{ij} = loading dari variabel ke- i pada faktor ke- j

- L = matriks dari *loading factor*
- m = banyak faktor terbentuk
- p = banyak variabel

Bagian dari varian variabel ke $-i$ dari m *common factor* disebut komunalitas ke $-i$ yang merupakan jumlah kuadrat dari loading variabel ke $-i$ pada m *common faktor* [8].

$$h_i^2 = \ell_{i1}^2 + \ell_{i2}^2 + \dots + \ell_{im}^2; \ell_i = e_i \sqrt{\lambda_i} \tag{7}$$

dimana

- e_i = *eigen vektor ke-i*
- λ_i = *eigen value ke-i*

E. Analisis Kluster Fuzzy K-Means

Salah satu teknik pengelompokan adalah Analisis Kluster *K-Means*. Pada teknik ini, jumlah kelompok (c) dan anggota kelompok awal telah ditentukan terlebih dahulu. Berikut adalah prosedur pengelompokan dengan *K-Means Cluster* [8].

- a. Membagi data pengamatan ke dalam c kelompok awal
- b. Menghitung *centroid* setiap kelompok dengan (8)

$$\bar{x}_i = \frac{\sum_{k=1}^{n_i} x_{ki}}{n_i} \tag{8}$$

dimana:

- \bar{x}_i : *centroid* kelompok ke- i
- x_{ki} : data ke- k dari kelompok ke- i
- n_i : banyak data dari kelompok ke- i

- c. Menghitung jarak setiap data terhadap *centroid* menggunakan jarak *euclidian* pada (9) kemudian meletakkan data pengamatan ke dalam kelompok dengan jarak terdekat

$$d_{ki}^2 = \sum_{i=1}^c (x_{ki} - \bar{x}_i)^2 \tag{9}$$

Ulangi ke langkah (c) hingga tidak terdapat perubahan anggota kelompok.

F. Analisis Kluster Fuzzy C-Means (FCM)

Pengelompokan FCM memberikan hasil pengelompokan yang baik meskipun data pengamatan mengandung *outlier*. FCM mempertimbangkan sifat keanggotaan *fuzzy* dalam suatu kelompok sebagai dasar pembobotan [1].

Representasi Linier

Pemetaan input ke derajat keanggotaan pada representasi linear digambarkan sebagai suatu garis lurus. Himpunan *fuzzy* yang linear terdiri dari dua keadaan yaitu pertama, kenaikan himpunan dimulai pada nilai dominan yang memiliki derajat keanggotaan 0 bergerak ke nilai dominan yang memiliki derajat keanggotaan lebih tinggi. Fungsi keanggotaan representasi linear naik ditunjukkan pada (10) berikut.

$$\mu_{ik}(x_k) = \begin{cases} 0 & ; x \leq a \\ (x_k - a)/(b - a); & a \leq x \leq b \\ 1 & ; x \geq b \end{cases} \tag{10}$$

Keadaan kedua adalah garis lurus dimulai dari nilai dominan dengan derajat keanggotaan tertinggi pada sisi kiri, kemudian bergerak menurun ke nilai dominan yang memiliki derajat keanggotaan lebih rendah. Fungsi keanggotaan representasi linear turun ditunjukkan pada (11).

$$\mu_{ik}(x_k) = \begin{cases} (b - x_k)/(b - a); & a \leq x \leq b \\ 0 & ; x \geq b \end{cases} \tag{11}$$

Representasi Kurva Segitiga

Kurva segitiga merupakan gabungan antara 2 garis linear. Fungsi keanggotaan representasi kurva segitiga ditunjukkan pada (12).

$$\mu_{ik}(x_k) = \begin{cases} 0 & ; x \leq a \text{ atau } x \geq c \\ (x_k - a)/(b - a); & a \leq x \leq b \\ (x_k - b)/(c - b); & b \leq x \leq c \end{cases} \tag{12}$$

Representasi Kurva Trapesium

Kurva trapesium pada dasarnya seperti bentuk segitiga namun memiliki beberapa titik dengan nilai keanggotaan 1. Fungsi keanggotaan representasi kurva trapesium ditunjukkan pada (13).

$$\mu_{ik}(x_k) = \begin{cases} 0 & ; x \leq a \text{ atau } x \geq d \\ (x_k - a)/(b - a) & ; a \leq x \leq b \\ 1 & ; b \leq x \leq c \\ (d - x_k)/(d - c) & ; x \geq d \end{cases} \tag{13}$$

Fuzzy C-Means Cluster (FCM)

Konsep dasar FCM adalah menentukan pusat kelompok yang akan menandai lokasi rata-rata untuk masing-masing kelompok dimana pada kondisi awal pusat kelompok masih belum akurat. Hal tersebut diatasi dengan memperbaiki pusat kelompok dan nilai keanggotaan secara berulang, sehingga pusat kelompok bergerak menuju lokasi yang tepat. Perulangan ini didasarkan pada minimalisasi fungsi objektif. Fungsi objektif J_w FCM ditunjukkan pada (14) [9].

$$J_w(\mathbf{U}, \mathbf{v}) = \sum_{k=1}^n \sum_{i=1}^c (\mu_{ik})^w (d_{ik})^2 \tag{14}$$

$$d_{ik} = d(\mathbf{x}_k - \mathbf{v}_i) = \left[\sum_{j=1}^m (x_{kj} - v_{ij})^2 \right]^{\frac{1}{2}} \tag{15}$$

$$v_{ij} = \frac{\sum_{k=1}^n \mu_{ik}^w \cdot x_{kj}}{\sum_{k=1}^n \mu_{ik}^w} \tag{16}$$

dengan $w \in [1, \infty]$

\mathbf{U} : matriks partisi

μ_{ik} : nilai anggota dari kelompok ke- i pada data ke- k

d_{ik} : ukuran jarak dari data ke- k ke pusat kelompok ke- i

\mathbf{v}_i : vektor pusat kelompok ke- i ; $\mathbf{v}_i = \{v_{i1}, v_{i2}, \dots, v_{im}\}$

v_{ij} : nilai pusat kelompok ke- i pada variabel ke- j

Nilai J_w akan optimum jika memiliki nilai terkecil sehingga selanjutnya sebagai berikut.

$$J_w^* = (\mathbf{U}, \mathbf{v}) = \min J(\mathbf{U}, \mathbf{v}) \tag{17}$$

Algoritma pengelompokan dengan menggunakan metode FCM adalah sebagai berikut:

1. Mendefinisikan matriks berukuran $n \times m$, dimana n adalah banyaknya data dan m adalah banyaknya variabel.
2. Menentukan jumlah kelompok yang akan dibentuk (c) dan pembobot eksponen (w), nilai dari w yang paling optimal dan sering dipakai adalah $w=2$.
3. Membentuk matriks partisi awal ($\mathbf{U}^{(0)}$). Setiap langkah pada algoritma ini akan diberi label r , dimana $r=0,1,2,\dots$

$$\mathbf{U}^{(0)} = \begin{bmatrix} \mu_{11}(x_1) & \mu_{12}(x_2) & \dots & \mu_{1n}(x_n) \\ \mu_{21}(x_1) & \mu_{22}(x_2) & \dots & \mu_{2n}(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{c1}(x_1) & \mu_{c2}(x_2) & \dots & \mu_{cn}(x_n) \end{bmatrix} \tag{18}$$

Matriks ini biasanya disusun secara random namun dapat juga disusun menggunakan pendekatan fungsi keanggotaan (10) hingga (14) hingga diperoleh nilai $\sum_{i=1}^c \mu_{ik} = 1$.

4. Menghitung pusat kelompok dari masing-masing kelompok $\{\mathbf{v}_i^{(r)}\}$ menggunakan (16) untuk setiap langkah.
5. Menghitung fungsi obyektif ke- r , $P^{(r)}$, dengan (19).

$$P^{(r)} = \sum_{k=1}^n \sum_{i=1}^c (\mu_{ik})^w (d_{ik})^2 \tag{19}$$

6. Memperbaiki matriks partisi (fungsi keanggotaan) untuk setiap langkah ke- r , $\mathbf{U}^{(r)}$ menggunakan (20) berikut.

$$\mu_{ik}^{(r+1)} = \left[\sum_{j=1}^c \left(\frac{d_{jk}^{(r)}}{d_{ik}^{(r)}} \right)^{\frac{2}{(w-1)}} \right]^{-1} \tag{20}$$

7. Menentukan kriteria berhenti, yaitu jika perubahan fungsi obyektif sekarang dengan sebelumnya bernilai $\leq \epsilon$ atau $|P^{(r+1)} - P^{(r)}| \leq \epsilon$. Namun jika perubahan fungsi obyektif sekarang dengan sebelumnya bernilai $> \epsilon$ maka gunakan $r=r+1$ dan kembali ke langkah 4. Nilai *threshold* (ϵ) yang digunakan adalah 10^{-6} .

G. Analisis Klaster Fuzzy Gustafson-Kessel (FGK)

Menurut Balasko, Abonyi, dan Feil [10], pengelompokan *Fuzzy Gustafson-Kessel* merupakan pengembangan dari *Fuzzy C-Means*. Metode pengelompokan *Gustafson-Kessel*, nilai pembentuk matriks disebut dengan *adaptive distance norm* diperbarui pada setiap iterasi. Hal ini yang menyebabkan pengelompokan *Gustafson-Kessel* lebih dapat menyesuaikan bentuk geometris fungsi keanggotaan yang tepat untuk sebuah himpunan data.

Algoritma dari pengelompokan *Fuzzy Gustafson Kessel* adalah sebagai berikut [10]:

1. Input data yang akan dikelompokkan
2. Menentukan banyak kelompok yang akan dibentuk ($1 < c < N$), *weighting exponent* ($m > 1$) maksimum iterasi (Maslter), error terkecil ($\epsilon > 0$), fungsi objektif awal= 0, dan iterasi awal (t=1)
3. Membuat matriks partisi awal dengan empat fungsi keanggotaan.

$$\sum_{i=1}^c \mu_{ik} = 1, \\ i = 1, 2, \dots, c; 1 \leq k \leq N$$

Sehingga dibentuk matrik U seperti dibawah ini.

$$U = \begin{pmatrix} \mu_{11}(x_1) & \mu_{12}(x_2) & \dots & \mu_{1l}(x_l) \\ \mu_{21}(x_1) & \mu_{22}(x_2) & \dots & \mu_{2l}(x_l) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{c1}(x_1) & \mu_{c2}(x_2) & \dots & \mu_{cl}(x_l) \end{pmatrix}$$

4. Menghitung pusat kelompok ke -k

$$v_i = \frac{\sum_{k=1}^N \mu_{ik}^m x_k}{\sum_{k=1}^N \mu_{ik}^m}, i = 1, 2, \dots, c$$

5. Menghitung matriks kovarians kelompok.

$$F_i = \frac{\sum_{k=1}^N \mu_{ik}^m (x_k - v_i)(x_k - v_i)^T}{\sum_{k=1}^N \mu_{ik}^m}$$

6. Menghitung jarak $D_{ikA_i}^2$ dengan rumus:

$$D_{ikA_i}^2 = (x_k - v_i)^T [\rho_i \det(F_i)]^{1/n} F_i^{-1} (x_k - v_i)$$

7. Menghitung fungsi objektif pada iterasi ke-t.

$$J_{GK}(X, U, V, A) = \sum_{i=1}^c \sum_{k=1}^N (\mu_{ik})^m D_{ikA_i}^2$$

8. Menghitung nilai fungsi keanggotaan yang baru U_{t+1}

$$u_{ik} = \left[\sum_{j=1}^c \left(\frac{D_{ikA_i}}{D_{jkA_j}} \right)^{\frac{2}{m-1}} \right]^{-1}$$

9. Membandingkan nilai keanggotaan dalam matriks U, jika $\|U^{t+1} - U^t\| < \epsilon$ atau (t>Maksler) maka sudah konvergen.

Jika $\|U^{t+1} - U^t\| \geq \epsilon$ kembali ke langkah 4.

H. Pseudo F-Statistic

Metode yang digunakan untuk menentukan banyaknya kelompok optimum adalah *pseudo F-statistic*. Nilai *pseudo F-statistics* tertingggi menunjukkan bahwa jumlah kelompok telah optimal [11].

$$Pseudo F = \frac{\left(\frac{R^2}{c-1} \right)}{\left(\frac{1-R^2}{n-c} \right)} \tag{21}$$

$$R^2 = \frac{SSB}{SST} = \frac{(SST - SSW)}{SST} \tag{22}$$

$$SST = \sum_{k=1}^n \sum_{i=1}^c \sum_{j=1}^m (x_{kij} - \bar{x}_j)^2 ; SSW = \sum_{k=1}^n \sum_{i=1}^c \sum_{j=1}^m (x_{kij} - \bar{x}_{ij})^2$$

Keterangan:

- R^2 = rasio dari *Sum Square Between* (SSB) dan *Sum Square Total* (SST)
- SST = total jumlah dari kuadrat jarak data terhadap rata-rata seluruh data
- SSW = total jumlah dari kuadrat jarak data terhadap rata-rata kelompok
- x_{kij} = data ke-k pada kelompok ke-i dan variabel ke-j

- \bar{x}_j = rata-rata seluruh data pada variabel j
- \bar{x}_{ij} = rata-rata data pada kelompok ke- i dan variabel ke- j

I. Internal Cluster Dispersion Rate

Perbandingan metode pengelompokan dapat diukur dengan menghitung rata-rata persebaran dalam kelompok atau *internal cluster dispersion rate (icdrate)* terhadap partisi secara keseluruhan [12].

$$icdrate = 1 - \frac{(SST - SSW)}{SST} = 1 - \frac{SSB}{SST} = 1 - R^2 \quad (30)$$

Nilai *icdrate* yang semakin kecil menunjukkan perbedaan keanggotaan tiap kelompok kecil.

J. Otoritas Jasa Keuangan (OJK)

OJK merupakan lembaga independen dan bebas dari campur tangan pihak lain, yang mempunyai fungsi, tugas dan wewenang pengaturan, pengawasan, pemeriksaan dan penyidikan di Sektor Jasa Keuangan (SJK) sebagaimana dimaksud dalam Undang-Undang Republik Indonesia Nomor 21 Tahun 2011 tentang Otoritas Jasa Keuangan [4].

OJK telah memiliki beberapa kantor yang tersebar di Indonesia. OJK memiliki 1 kantor pusat dan 36 kantor perwakilan di 9 regional. Dari seluruh kantor tersebut, hanya ada 1 kantor perwakilan di Sumatera Utara yaitu Kantor Regional 5 Sumatera Bagian Utara yang terletak di Medan serta perwakilan yang belum memiliki kantor yakni Padang Sidempuan [13].

K. Provinsi Sumatera Utara

Sumatera Utara merupakan salah satu provinsi yang ada di Pulau Sumatera, Indonesia. Perekonomian Sumatera Utara pada tahun 2015 menunjukkan peningkatan meskipun pertumbuhan melambat. Sektor ekonomi yang memiliki peranan PDRB tertinggi adalah Pertanian, Kehutanan, dan Perikanan (22,02%); Industri Pengolahan (20,21%); dan Perdagangan Besar dan Eceran, Reparasi Mobil, dan Sepeda Motor (17,41%) [14]. Provinsi Sumatera Utara terdiri dari 33 Kabupaten/Kota.

III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan merupakan data *website* Badan Pusat Statistik (BPS) Pusat dan BPS Provinsi Sumatera Utara untuk seluruh Kabupaten/Kota di Provinsi Sumatera Utara serta data dari Kantor Regional IV OJK Jawa Timur pada tahun 2015.

B. Units

Variabel dalam penelitian ini adalah sebagai berikut.

Tabel 1 Variabel Penelitian

Variabel	Simbol	Satuan
PDRB	X_1	Miliyar Rupiah
Pengeluaran Pemerintah	X_2	Miliyar Rupiah
Jumlah Penduduk	X_3	Ribu Jiwa
IPM	X_4	Rasio (%)
Piutang Perusahaan Pembiayaan	X_5	Miliyar Rupiah

C. Langkah Analisis

Langkah analisis yang digunakan pada penelitian ini adalah sebagai berikut.

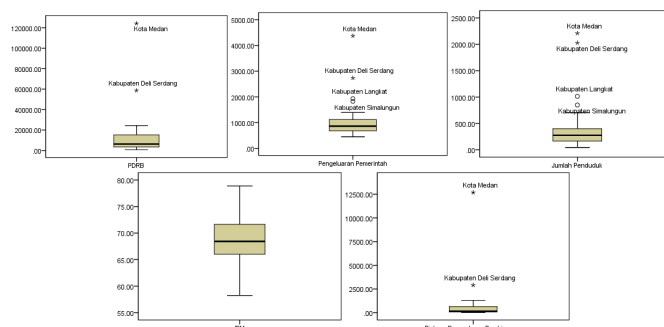
1. Mendeskripsikan karakteristik data Provinsi Sumatera Utara.
2. Melakukan pengelompokan variabel dan Kabupaten/Kota di Provinsi Sumatera Utara dengan variabel terpilih serta keseluruhan variabel.
 - a. Melakukan Pemeriksaan Asumsi Kecukupan Korelasi menggunakan (1), Pemeriksaan Asumsi Multivariat Normal menggunakan (2), dan Pemeriksaan Asumsi Dependensi Variabel menggunakan (3).
 - b. Melakukan Analisis Faktor dan menentukan variabel terpilih dari masing-masing faktor yang terbentuk menggunakan (6) dan (7). Variabel dipilih dari masing-masing faktor berdasarkan kontribusi terbesar variabel tersebut terhadap faktor.
 - c. Melakukan Analisis Klaster K-Means dan FCM terhadap variabel terpilih dan juga seluruh variabel menggunakan *software R* berdasarkan (8) hingga (20). Selanjutnya, membandingkan hasil klaster berdasarkan variabel terpilih dengan keseluruhan variabel.

IV. HASIL DAN PEMBAHASAN

A. Karakteristik Perekonomian Provinsi Sumatera Utara

Data yang digunakan merupakan data *website* Badan Pusat Statistik (BPS) Pusat dan BPS Provinsi Sumatera Utara untuk seluruh Kabupaten/Kota di Provinsi Sumatera Utara serta data dari Kantor Regional IV OJK Jawa Timur pada tahun 2015.

Analisis Statistika Deskriptif dilakukan untuk menjelaskan karakteristik data perekonomian Provinsi Sumatera Utara dengan menyajikan ringkasan data dan grafik-grafik.



Gambar 1 Karakteristik PDRB, Pengeluaran Pemerintah, Jumlah Penduduk, IPM, dan Piutang Perusahaan Pembiayaan Kabupaten/Kota di Provinsi Sumatera Utara Tahun 2015.

Gambar 1 memberikan informasi bahwa PDRB, Pengeluaran Pemerintah, Jumlah Penduduk, dan Piutang Perusahaan Pembiayaan Kota Medan dan Kabupaten Deli Serdang merupakan data ekstrim. Kota Medan memiliki maksimum pada seluruh variabel. Rata-rata PDRB Kabupaten/Kota di Provinsi Sumatera Utara adalah 13.418,14 miliar rupiah. Pengeluaran Pemerintah Kabupaten Langket dan Kabupaten Simalungun merupakan data *outlier* dengan rata-rata keseluruhan 1.077,57 miliar rupiah. Jumlah Penduduk Kabupaten Langket dan Kabupaten Simalungun merupakan data *outlier* dengan rata-rata keseluruhan 422,36 ribu jiwa. Data IPM tidak memiliki nilai ekstrim maupun *outlier*. Rata-rata keseluruhan IPM adalah 68,43%. Piutang Perusahaan Pembiayaan memiliki rata-rata keseluruhan 772 miliar rupiah.

B. Pengelompokan Variabel dan Kabupaten/Kota di Provinsi Sumatera Utara

Sebelum Analisis Faktor, dilakukan beberapa pemeriksaan asumsi. Pemeriksaan asumsi tersebut yaitu kecukupan korelasi, multivariat normal, dan dependensi variabel.

Pemeriksaan Asumsi Kecukupan Korelasi

Tabel 2 adalah nilai korelasi antar variabel yang diketahui bahwa X1, X2, X3, dan X5 yaitu PDRB, Pengeluaran Pemerintah, Jumlah Penduduk, dan Piutang Perusahaan Pembiayaan saling berkorelasi satu sama lain. Sedangkan X4 yaitu IPM tidak berkorelasi dengan yang lain.

Tabel 2. Korelasi antar Variabel

Variabel	X1	X2	X3	X4	X5
X2	0,955	1			
X3	0,926	0,962	1		
X4	0,477	0,491	0,432	1	
X5	0,946	0,876	0,792	0,484	1

Berdasarkan nilai korelasi tersebut, selanjutnya dihitung nilai KMO. Nilai KMO ditunjukkan pada Tabel 3 berikut.

Tabel 3 Nilai KMO

Variabel	Nilai KMO
Keseluruhan	0,713
PDRB	0,710
Pengeluaran Pemerintah	0,809
Jumlah Penduduk	0,643
IPM	0,954
Piutang Perusahaan Pembiayaan	0,645

Tabel 3 menunjukkan bahwa nilai KMO secara keseluruhan maupun masing-masing variabel lebih besar dari 0,5. Sehingga, korelasi antar variabel dikatakan cukup.

Pemeriksaan Asumsi Multivariat Normal

Pemeriksaan dilakukan untuk mengetahui apakah data telah berdistribusi normal secara multivariat menggunakan *Mardia's Multivariate Normality Test*.

Tabel 4 Mardia's Multivariate Normality Test

<i>Chi-Square</i>	304,0419
<i>P-Value</i>	0,000
Hasil	Data Tidak Multivariat Normal

Tabel 4 memberikan informasi bahwa data yang digunakan tidak normal secara multivariat. Hal tersebut disebabkan oleh adanya nilai ekstrim dan *outlier* pada data. Nilai ekstrim tersebut adalah data dari Kota Medan dan Kabupaten Deli Serdang.

Pemeriksaan Asumsi Dependensi Variabel

Pemeriksaan ini dilakukan untuk mengetahui apakah variabel saling dependen dan berkorelasi dengan menggunakan uji *Bartlett Sphericity*.

Tabel 5 Uji Bartlett Sphericity

<i>Chi-Square</i>	247,202
<i>Chi-Square Tabel</i>	18,307
Hasil	Variabel Saling Berkorelasi dan Dependen

Tabel 5 menunjukkan bahwa variabel yang digunakan dalam penelitian saling berkorelasi dan dependen. Sehingga, dapat dilanjutkan ke analisis selanjutnya yaitu Analisis Faktor.

Analisis Faktor

Analisis Faktor digunakan untuk membentuk faktor-faktor yang dapat mewakili keseluruhan variabel yang memiliki kemiripan pada masing-masing faktor.

Tabel 6. Nilai Eigen

Komponen	Nilai Eigen Awal		
	Nilai Eigen	% Variasi	% Kumulatif
1	4,024	80,480	80,480
2	0,710	14,201	94,680
3	0,224	4,477	99,158
4	0,030	0,605	99,763
5	0,012	0,237	100,000

Tabel 6 memberikan informasi bahwa nilai eigen yang berada di sekitar nilai 1 ada sebanyak 2 komponen. Kedua komponen memberikan keragaman masing-masing 80,48% dan 14,2% yang berbeda jauh dengan komponen selanjutnya. Apabila dilihat dari prosentase kumulatif, kedua komponen tersebut memiliki nilai sebesar 94,68%. Nilai tersebut cukup mendekati nilai keseluruhan yaitu 100%.

Tabel 7 Matriks Komponen

Variabel	Komponen	
	1	2
X1	0,982	-0,133
X2	0,976	-0,118
X3	0,940	-0,174
X4	0,597	0,802
X5	0,932	-0,074

Tabel 7 menunjukkan bahwa terbentuk 2 faktor. Faktor 1 terdiri dari variabel PDRB, Pengeluaran Pemerintah, Jumlah Penduduk, dan Piutang Perusahaan Pembiayaan. Variabel terpilih dari faktor 1 adalah PDRB (X_1) karena memiliki kontribusi tertinggi dibandingkan dengan variabel yang lain. Sedangkan faktor 2 terdiri dari IPM (X_4).

Analisis Klaster

Penelitian dilanjutkan dengan Analisis Klaster *K-Means* dan FCM menggunakan lima dan dua variabel hasil Analisis Faktor. FCM menunjukkan bahwa fungsi keanggotaan memberikan hasil yang sama setiap klaster, untuk itu digunakan fungsi keanggotaan linier naik sebagai perbandingan. Berikut ini adalah perbandingan hasil pengelompokan.

Tabel 8 Perbandingan Hasil Pengelompokan

Klaster	K-Means					
	Dua Variabel			Lima Variabel		
	2	3	4	2	3	4
1	32	31	22	32	31	22
2	1	1	1	1	1	1
3		1	1		1	1
4			9			9

Klaster	FCM					
	Dua Variabel			Lima Variabel		
	2	3	4	2	3	4
1	1	1	1	1	1	1
2	32	23	9	32	23	9
3		9	22		9	22
4			1			1

Tabel 8 menunjukkan bahwa hasil kelompok menggunakan lima dan dua variabel sama. Namun pengelompokan metode *K-Means* dan *FCM* berbeda pada tiga klaster.

Selanjutnya menentukan klaster optimum berdasarkan nilai *pseudo F-statistics* terbesar dan *icdrate* terkecil. Berikut adalah nilai *pseudo F-statistics* dan *icdrate* dari masing-masing klaster terbentuk.

Tabel 9 Perbandingan Pseudo F-Statistics dan Icdrate antara K-Means dan FCM

Jumlah Klaster	K-Means			
	Dua Variabel		Lima Variabel	
	<i>Pseudo F-Statistics</i>	<i>Icdrate</i>	<i>Pseudo F-Statistics</i>	<i>Icdrate</i>
2	100,016	0,237	100,720	0,235
3	154,048	0,089	154,816	0,088
4	662,007*	0,014**	650,884	0,015

Jumlah Klaster	FCM			
	Dua Variabel		Lima Variabel	
	<i>Pseudo F-Statistics</i>	<i>Icdrate</i>	<i>Pseudo F-Statistics</i>	<i>Icdrate</i>
2	100,016	0,237	100,720	0,235
3	136,174	0,099	136,394	0,099
4	662,007*	0,014**	650,884	0,015

*Nilai *Pseudo F-Statistics* Terbesar; **Nilai *Icdrate* Terkecil

Tabel 9 memberikan informasi bahwa metode *K-Means* dan *FCM* dengan dua variabel dan membentuk empat klaster terbukti lebih optimum berdasarkan nilai *Pseudo F-Statistics* yang paling besar dan *Icdrate* yang paling kecil. Hal tersebut menunjukkan bahwa lima variabel tersebut terwakili oleh dua variabel yang terbentuk dari hasil Analisis Faktor yaitu PDRB dan IPM. Selain itu ditunjukkan bahwa, pengelompokan dengan *FCM* lebih baik daripada *K-Means*.

Selanjutnya dilakukan perbandingan kebaikan hasil pengelompokan, antara *FCM* dan *FGK* dengan dua variabel dan kelompok optimum masing-masing metode adalah empat, seperti ditunjukkan pada Tabel 10.

Tabel 10 Perbandingan Pseudo F-Statistics dan Icdrate antara FCM dan FGK

Jumlah Klaster	FCM		FGK	
	<i>Pseudo F-Statistics</i>	<i>Icdrate</i>	<i>Pseudo F-Statistics</i>	<i>Icdrate</i>
	4	662,007*	0,014**	24,1606

Jumlah Anggota Setiap Klaster	(1,9,22,1)	(12,10,9,2)
-------------------------------	------------	-------------

Tabel 11 menunjukkan anggota kelompok akhir yang terbentuk, berdasarkan metode pengelompokan terbaik yaitu *FCM* (Tabel 10).

Tabel 11 Anggota Kluster Terbentuk

Kluster	Anggota Kluster
1	Deli Serdang
2	Langkat, Simalungun, Asahan, Batu Bara, Labuhanbatu, Serdang Bedagai, Labuhanbatu Selatan, Labuhanbatu Utara, Karo.
3	Pematang Siantar, Tapanuli Selatan, Mandailing Natal, Padanglawas Utara, Binjai, Padanglawas, Tapanuli Tengah, Dairi, Tapanuli Utara, Tanjung Balai, Toba Samosir, Nias Selatan, Padangsidempuan, Humbang Hasundutan, Tebingtinggi, Sibolga, Gunungsitoli, Samosir, Nias, Nias Utara, Nias Barat, Pakpak Bharat.
4	Medan

Karakteristik masing-masing kluster yang terbentuk dilihat dari PDRB dan IPM dari Kabupaten/Kota. Kluster yang memiliki rata-rata PDRB dan IPM tertinggi secara berurutan adalah kluster 4 yaitu 124.277,48 miliar rupiah dan 78,87%, kluster 1 yaitu 58.722,46 miliar rupiah dan 72,79%, kluster 2 yaitu 18.245,74 miliar rupiah dan 69,39%, dan kluster 3 yaitu 4.344,87 miliar rupiah dan 67,36%. Anggota dengan PDRB dan IPM tertinggi masing-masing kluster adalah Kota Medan (kluster 4), Kabupaten Deli Serdang (kluster 1), Kabupaten Simalungun yaitu 22.305,43 miliar rupiah dan 71,24% (kluster 2), serta Kota Pematang Siantar (kluster 3) yaitu 7.992,32 miliar rupiah dan 76,34%.

Merekomendasikan Kabupaten/Kota yang akan dibangun kantor OJK di Provinsi Sumatera Utara harus memperhatikan keberadaan Kantor OJK yang sudah ada dan jarak antara lokasi tersebut. Karakteristik kluster menunjukkan kluster 4 yaitu Kota Medan adalah kluster paling unggul. Hal tersebut memberikan konfirmasi bahwa Kota Medan merupakan Kabupaten/Kota yang menjadi prioritas OJK sehingga pembangunan kantor OJK yang telah dilakukan pertama kali di Provinsi Sumatera Utara adalah langkah yang tepat. Setelah kluster 4, rekomendasi lokasi kantor baru OJK adalah di Kabupaten Deli Serdang sebagai anggota kluster 1 terunggul yang memiliki jarak dengan Kota Medan kurang lebih 25,4 km. Sebagai alternatif, Kabupaten Simalungun direkomendasikan sebagai anggota kluster 2 terunggul yang memiliki jarak dengan Kota Medan kurang lebih 155 km. Alternatif selanjutnya adalah Kota Pematang Siantar sebagai anggota kluster 3 terunggul yang memiliki jarak dengan Kota Medan kurang lebih 128 km.

V. KESIMPULAN

Berdasarkan hasil analisis dapat disimpulkan bahwa PDRB, Pengeluaran Pemerintah, Jumlah Penduduk, dan Piutang Perusahaan Pembiayaan memiliki data ekstrim yaitu Kota Medan dan Kabupaten Deli Serdang. Variabel membentuk dua faktor yaitu faktor pertama terdiri dari variabel PDRB, Pengeluaran Pemerintah, Jumlah Penduduk, dan Piutang Perusahaan Pembiayaan. Variabel terpilih dari faktor pertama adalah PDRB dan faktor kedua yaitu IPM.

Dibandingkan dengan FGK, diperoleh hasil Analisis kluster terbaik dengan FCM pada pengelompokan optimum empat kluster dengan dua variabel. Kluster Kabupaten/Kota di Provinsi Sumatera Utara yang memiliki rata-rata PDRB dan IPM tertinggi adalah Kota Medan sehingga dikonfirmasi layak telah memiliki Kantor OJK. Maka dari itu, rekomendasi akhir lokasi Kantor OJK baru selanjutnya adalah Kabupaten Deli Serdang dengan alternatif Kabupaten Simalungun dan Kota Pematang Siantar.

REFERENSI

- [1] Suyanto, *Data Mining untuk Klasifikasi dan Klusterisasi Data*. Bandung: Informatika.(2017).
- [2] Badan Pusat Statistik, 1. *Laporan Perekonomian Indonesia 2017*. Jakarta: CV Nario Sari.(2017).
- [3] Badan Pusat Statistik. (30 Oktober 2017). [Seri 2010] *Produk Domestik Regional Bruto Atas Dasar Harga Konstan 2010 Menurut Provinsi, 2010-2016 (Miliar Rupiah)*. [Online]. Available: <https://www.bps.go.id/linkTableDinamis/view/id/956>.
- [4] Otoritas Jasa Keuangan. (2017, Oktober 30). *Booklet Perbankan Indonesia 2017*. [Online]. Available at: <http://www.ojk.go.id/id/kanal/perbankan/data-dan-statistik/booklet-perbankan-indonesia/Documents/pages/booklet-perbankan-indonesia-2017/BPI%202017%20Bahasa%20Indonesia.pdf>.
- [5] S. Sharma, *Applied Multivariate Techniques*. New York: John Willey & Sons, Inc.(1996).
- [6] S. Korkmaz, D. Goksuluk, dan G. Zararsiz, *MVN: An R Package for Assessing Multivariate Normality*. Ankara: s.n.(2016).
- [7] D. F. Morrison, *Multivariate Statistical Method Fourth Edition*. New York: McGraw-Hill, Inc.(2005).
- [8] R. A. Johnson dan D. W. Winchern, *Applied Multivariate Statistical Analysis Sixth Edition*. New Jersey: Pearson Prentice Hall.(2007).
- [9] T. J. Ross, *Fuzzy Logic with Engineering Applications Third Edition*. USA: John Wiley & Sons.(2010)
- [10] Balasko, B., Abonyi, J. & Feil, B. (2007). *Fuzzy Clustering and Data Analysis Toolbox For Use with Matlab*. University of Veszprem: Veszprem.
- [11] R. A. Orpin dan V. E. Kostylev, *Towards a Statistically Valid Method of Textural Sea Floor Characterization of Benthic Habitats, Marine Geology*. (2006) pp. 225:209-222.
- [12] S. A. Mingoti dan J. O. Lima, *Comparing SOM Neural Network with Fuzzy C-Means, K-Means, and Traditional Hierarchical Clustering Algorithms. European Journal of Operational Research*, (2006) pp. 1742-1759.
- [13] Otoritas Jasa Keuangan. (30 Oktober 2017). *Kontak Kami*. [Online] Available at: <http://www.ojk.go.id/id/contact.aspx>
- [14] BPS Sumatera Utara, *PDRB Menurut Lapangan Usaha Provinsi Sumatera Utara 2012 – 2016*. Medan: CV Rilis Grafika. (2017).