

The Application Of K – Harmonic Means Method In District/City Grouping (Case Study: Poverty in Kalimantan Island in 2020)

Penerapan Metode *K – Harmonic Means* Dalam Pengelompokan Kabupaten/Kota (Studi Kasus: Kemiskinan di Pulau Kalimantan Tahun 2020)

Dwi Indra Yunistya¹, Rito Goejantoro², Fidia Deny Tisna Amijaya³

^{1,2}Laboratorium Statistika Komputasi Fakultas MIPA, Universitas Mulawarman,

³Laboratorium Matematika Komputasi, Fakultas MIPA, Universitas Mulawarman

Email address: yunistyaindra77@gmail.com₁, rito.goejantoro@fmipa.unmul.ac.id₂,
fdiadta@fmipa.unmul.ac.id₃

Abstract

Poverty is one of the problems that faced by every country in the world, especially in developing countries, one of them is Indonesia. Poverty alleviation that is currently planned is no longer uniform, but it is necessary to pay attention to the condition of each dimension causing poverty in an area, so it is necessary to group districts/cities on the Kalimantan Island based on poverty. Cluster analysis is classifying the data (objects) only based on the information discovered in the data that describes the objects and the relations between them. The method used in this research is K-Harmonic Means method. K-Harmonic Means is a non-hierarchical clustering algorithm that uses the average harmonic distance from each data point to the cluster center. This study aims to classify the District/City in Kalimantan Island based on poverty indicators and obtain the silhouette coefficient value from the optimal cluster analysis. Based on the results of the analysis of the K-Harmonic Means method, the optimal number of clusters is 2 clusters with parameter (p) of 4. Cluster 1 consists of 11 Districts/Cities and Cluster 2 consists of 45 Districts/Cities. Silhouette coefficient value for data validation of District/City clustering results on Kalimantan Island using the K-Harmonic Means method, namely 2 clusters with parameter (p) of 4 is 0.323 which states that the resulting cluster structure in this grouping is a weak structure.

Keywords: Poverty, K-Harmonic Means, Silhouette Coefficient

Abstrak

Kemiskinan merupakan salah satu permasalahan yang dihadapi oleh setiap negara di dunia, terutama pada negara yang sedang berkembang, salah satunya adalah negara Indonesia. Pengentasan kemiskinan yang direncanakan saat ini tidak lagi bersifat seragam, namun perlu memperhatikan kondisi setiap dimensi penyebab kemiskinan suatu wilayah maka perlu dilakukan



JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Dwi Indra Yunistya, Rito Goejantoro, Fidia Deny Tisna Amijaya

pengelompokan Kabupaten/Kota di Pulau Kalimantan berdasarkan kemiskinan. Analisis kluster adalah mengelompokkan data (objek) yang didasarkan hanya pada informasi yang ditemukan dalam data yang menggambarkan objek tersebut dan hubungan di antaranya. *K-Harmonic Means* adalah algoritma pengelompokan non-hierarki yang menggunakan rata-rata harmonik jarak dari setiap data ke pusat kluster. Metode *K-Harmonic Means* memiliki nilai simpangan baku yang lebih kecil dibandingkan metode *K-Means* yang berarti hasil dari klusterisasi dari metode *K-Harmonic Means* memiliki persebaran data yang kecil. Penelitian ini bertujuan untuk mengelompokkan Kabupaten/Kota di Pulau Kalimantan berdasarkan indikator kemiskinan serta memperoleh nilai *silhouette coefficient* hasil analisis kluster optimal. Berdasarkan hasil analisis metode *K-Harmonic Means* didapatkan banyak kluster optimal yaitu 2 kluster dengan parameter (p) sebesar 4. Kluster 1 beranggotakan 11 Kabupaten/Kota dan kluster 2 beranggotakan 45 Kabupaten/Kota. Nilai *silhouette coefficient* untuk validasi data hasil klustering Kabupaten/Kota di Pulau Kalimantan menggunakan metode *K-Harmonic Means*, yaitu 2 kluster dengan parameter (p) sebesar 4 adalah 0,323 yang menyatakan bahwa struktur kluster yang dihasilkan pada pengelompokan ini adalah *weak structure*.

Kata kunci: Kemiskinan, *K-Harmonic Means*, *Silhouette Coefficient*

1. PENDAHULUAN

1.1 Latar Belakang

Kemiskinan merupakan salah satu permasalahan yang dihadapi oleh setiap negara di dunia, terutama pada negara yang sedang berkembang, salah satunya adalah negara Indonesia. Salah satu tujuan Negara Indonesia yaitu untuk menjamin dan memajukan kesejahteraan umum, di mana bebas dari kemiskinan merupakan salah satu indikator kesejahteraan rakyat [3]. Penanggulangan kemiskinan memerlukan upaya komprehensif untuk menghindari ketergantungan semata-mata pada indikator ekonomi dari pada mempertimbangkan faktor-faktor sosial, budaya, politik dan lainnya [10]. Faktor kemiskinan setiap wilayah berbeda, dan hal tersebut dipengaruhi oleh banyak indikator kemiskinan. Pengukuran dan penentuan indikator kemiskinan akan memudahkan pemerintah membedakan tingkat kemiskinan dan karakteristik daerah tersebut, sehingga pemerintah dapat membuat kebijakan yang lebih tepat untuk menanggulangi kemiskinan di wilayah tersebut [3].

Pada tanggal 26 Agustus 2019, Pulau Kalimantan secara resmi ditunjuk sebagai Ibu Kota Negara baru yaitu pada Provinsi Kalimantan Timur. Penelitian [2] mengatakan Jakarta sebagai Ibu Kota Negara merupakan pusat industri dan bisnis menjadi tujuan urbanisasi. Tren perpindahan penduduk dari pedesaan ke perkotaan akhir-akhir ini yang dikenal dengan urbanisasi menimbulkan banyak tantangan bagi masyarakat. Urbanisasi memberikan peluang besar bagi masyarakat untuk tumbuh dan berkembang, namun kesulitan yang di hadapi masyarakat di kawasan perkotaan ini juga banyak. Ini termasuk kemacetan dalam transportasi persaingan untuk layanan kesehatan, polusi, kelangkaan sumber daya tak terbarukan, pengelolaan limbah padat dan kurangnya kesempatan pendidikan [13]. Hal ini akan menjadi daya tarik bagi banyak orang untuk datang ke Pulau Kalimantan, sehingga dapat menimbulkan jumlah penduduk di pulau Kalimantan akan meningkat. Seperti pada penelitian [11] yang menyatakan Kedudukan Jakarta sebagai ibu kota Negara, menjadikan gengsi tersendiri yang menarik minat banyak orang datang ke Jakarta, membuat Jakarta semakin mengalami pertumbuhan dan perkembangan penduduk yang pesat. Pertumbuhan penduduk memegang peranan penting dalam mempengaruhi jumlah penduduk miskin [15]. Pengentasan kemiskinan yang disusun saat ini tidak lagi bersifat seragam namun perlu memperhatikan kondisi setiap dimensi penyebab kemiskinan suatu wilayah

[3]. Oleh karena itu, perlu dilakukan pengelompokan Kabupaten/Kota di Pulau Kalimantan berdasarkan kemiskinan.

Pengelompokan (*Clustering*) yaitu membagi kumpulan data ke dalam kelompok-kelompok dimana kesamaan dalam suatu kelompok lebih besar daripada antar kelompok [12]. Salah satu metode yang umum digunakan untuk melakukan pengelompokan adalah analisis klaster (*cluster analysis*). Analisis klaster (*cluster analysis*) adalah mengelompokkan data (objek) yang didasarkan hanya pada informasi yang ditemukan dalam data yang menggambarkan objek tersebut dan hubungan di antaranya [9]. *K-Means* adalah metode pengelompokan data non-hierarki yang mencoba untuk mempartisi data yang ada menjadi satu atau lebih bentuk cluster [12]. Algoritma *K-Means* telah banyak digunakan karena ide algoritmanya yang sederhana, implementasi yang mudah dan efisiensi yang tinggi saat memproses data skala besar. Ciri utama dari *clustering* algoritma *K-Means* adalah tingginya tingkat kemiripan data pada cluster yang sama dan rendahnya tingkat kemiripan data pada cluster yang berbeda [18]. Pada tahapan *K-Means* untuk mendapatkan klaster akhir terlihat kelemahan di mana keakuratan hasil klaster sangat tergantung dari penentuan titik awal pusat klaster sehingga permasalahan sensitifitas terhadap penentuan titik awal menjadi kelemahan metode ini [16].

Untuk mengatasi masalah mengenai penentuan pusat klaster, diusulkan sebuah metode baru yang bernama *K-Harmonic Means*. Metode ini bertujuan untuk meminimalisasi rata-rata harmonik dari seluruh titik data ke seluruh pusat klaster yang ada. Nilai fungsi tujuan dari algoritma *K-Harmonic Means* dihasilkan dengan mencari total rata-rata harmonik dari seluruh titik data terhadap jarak antara masing-masing titik data ke seluruh titik pusat klaster yang ada. Hal ini berbeda dengan *K-Means* di mana fungsi tujuan diperoleh dari total jarak seluruh data ke titik pusat klasternya. [17]. Pada *K-Means* diasumsikan tiap titik data mempunyai bobot yang sama sedangkan pada *K-Harmonic Means* tiap titik data diberi bobot yang dinamis berdasarkan suatu rata-rata harmoniknya [4]. Pada Penelitian [16], mengenai studi komparasi klasterisasi data *K-Means* dan *K-Harmonic Means* didapatkan bahwa metode *K-Harmonic Means* memiliki nilai simpangan baku yang lebih kecil dibandingkan metode *K-Means* yang berarti hasil dari klasterisasi dari metode *K-Harmonic Means* memiliki persebaran data yang kecil.

Untuk mengukur kekuatan dan kualitas hasil klaster maka perlu memvalidasi hasil klaster, salah satu metode untuk memvalidasi hasil klaster yaitu *silhouette coefficient*. *Silhouette coefficient* merupakan sebuah metode yang digunakan untuk memvalidasi klaster dengan menggabungkan metode separasi dan metode kohesi [5]. Metode *silhouette coefficient* dapat dikatakan bahwa metode ini paling banyak digunakan untuk validasi klaster [6]. Berdasarkan uraian di atas maka penulis tertarik melakukan penelitian mengenai pengelompokan Kabupaten/Kota di Pulau Kalimantan dengan menggunakan metode *K-Harmonic Means* berdasarkan kemiskinan.

1.2 *K-Harmonic Means*

K-Harmonic Means adalah algoritma pengelompokan berbasis pusat yang menggunakan rata-rata harmonik jarak dari setiap titik data ke pusat sebagai komponen untuk fungsi kinerjanya. Tujuan dari algoritma ini adalah meminimalisasi rata-rata harmonik dari semua titik pada data set ke seluruh pusat klaster [17].

K-Harmonic Means merupakan salah satu contoh *center-based cluster* dan merupakan sebuah metode di mana klaster-klaster dibentuk dengan penyempurnaan secara iteratif berdasarkan letak titik pusat dari masing-masing klaster. Rata-rata harmonik ini didefinisikan seperti Persamaan (1.1).

$$HA(X_j) = \frac{n}{\sum_{i=1}^n \frac{1}{x_{ij}}} \quad (1.1)$$

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Dwi Indra Yunistya, Rito Goejantoro, Fidia Deny Tisna Amijaya

Dalam fungsi harmonik, jika terdapat satu anggota dalam x_{1j}, \dots, x_{nj} bernilai kecil maka nilai rata-rata harmonik pun bernilai kecil, tetapi jika tidak ada anggota bernilai kecil maka nilainya pun besar [16].

Pada metode *K-Means*, penentuan suatu data ke dalam suatu kluster didasarkan pada jarak minimum ke suatu kluster, sedangkan pada *K-Harmonic Means* digunakan rata-rata harmoniknya. Fungsi tujuan pada *K-Means* adalah [17]:

$$KM(X, C) = \sum_{i=1}^n \text{MIN}\{\|x_i - c_l\|^2 | l = 1 \dots K\} \quad (1.2)$$

Dengan mengganti fungsi MIN () pada Persamaan (1.2) dengan fungsi HA diperoleh fungsi tujuan dari *K-Harmonic Means* sebagai berikut [17]:

$$KHM(X, C) = \sum_{i=1}^n \frac{K}{\sum_{l=1}^K \frac{1}{\|x_i - c_l\|^2}} \quad (1.3)$$

Zhang, Hsu, dan Dayal, (1999), memperkenalkan sebuah parameter p yang merupakan pemangkatan berasosiasi dengan pengukuran jarak antara data dan titik pusat kluster. Parameter p ini digunakan untuk mengganti *square* pada Persamaan (1.3), sehingga performansi dari *K-Harmonic Means* menjadi:

$$KHM(X, C) = \sum_{i=1}^n \frac{K}{\sum_{l=1}^K \frac{1}{\|x_i - c_l\|^p}} \quad (1.4)$$

Rata-rata harmonik sangat sensitif dengan keadaan di mana terdapat dua atau lebih titik pusat yang saling berdekatan. Metode ini secara natural menempatkan satu atau lebih titik pusat ke area titik data yang jauh dari titik-titik pusat yang ada sebelumnya. Hal ini akan membuat fungsi tujuan akan semakin kecil. Adapun langkah-langkah metode *K-Harmonic Means* adalah sebagai berikut [16]:

1. Menentukan posisi titik pusat kluster awal secara *random*
2. Menghitung nilai fungsi tujuan dengan Persamaan (1.4), di mana p adalah parameter. Nilai p biasanya ≥ 2 .
3. Menghitung nilai keanggotaan $m(c_l|x_i)$ untuk setiap titik data x_i pada setiap pusat kluster c_l berdasarkan Persamaan (1.5).

$$m(c_l|x_i) = \frac{\|x_i - c_l\|^{-p-2}}{\sum_{l=1}^K \|x_i - c_l\|^{-p-2}} \quad (1.5)$$

4. Menghitung nilai bobot $w(x_i)$ untuk setiap data x_i . Fungsi pembobotan *K-Harmonic Means* didefinisikan pada Persamaan (1.6).

$$w(x_i) = \frac{\sum_{l=1}^K \|x_i - c_l\|^{-p-2}}{(\sum_{l=1}^K \|x_i - c_l\|^{-p})^2} \quad (1.6)$$

5. Mengulang kembali perhitungan untuk posisi titik pusat kluster untuk setiap titik pusat c_l dari semua data berdasarkan nilai keanggotaan dan bobot yang dimiliki tiap data.

$$c_l = \frac{\sum_{i=1}^n m(c_l|x_i) \cdot w(x_i) \cdot x_i}{\sum_{i=1}^n m(c_l|x_i) \cdot w(x_i)} \quad (1.7)$$

6. Mengulang langkah 2 sampai 5 hingga mendapatkan nilai fungsi tujuan yang tidak terdapat perubahan yang signifikan.
7. Menetapkan keanggotaan data x_i pada suatu kluster dengan titik pusat kluster c_l sesuai dengan nilai keanggotaan x_i terhadap c_l .

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Dwi Indra Yunistya, Rito Goejantoro, Fidia Deny Tisna Amijaya

x_i merupakan anggota dari kluster dengan titik pusat kluster c_l apabila nilai keanggotaan $m(c_l|x_i)$ adalah yang terbesar dibandingkan dengan nilai keanggotaannya ke titik pusat kluster lain [16].

1.3 Evaluasi Hasil Klustering

Salah satu metode evaluasi yang dapat digunakan untuk melihat kualitas dan kekuatan kluster adalah metode *Silhouette Coefficient* (SC). Metode ini merupakan metode validasi kluster yang menggabungkan metode *cohesion* dan *separation*. Tahapan perhitungan *silhouette coefficient* adalah sebagai berikut [7]:

1. Menghitung rata-rata jarak dari suatu data ke- i dengan semua data yang berada pada satu kluster yang sama dengan menggunakan Persamaan (1.8)

$$a_i = \frac{1}{n_k - 1} \sum_{r=1}^{n_k-1} d(x_i, x_r), r \neq i \quad (1.8)$$

dengan $k= 1, 2, 3, \dots, K$.

2. Menghitung rata-rata jarak suatu data ke- i dengan semua data yang berada pada kluster yang berbeda dengan menggunakan Persamaan (1.10), kemudian ambil nilai terkecilnya berdasarkan Persamaan (1.9)

$$b_i = \min\{d_i(k)\}, r \neq i \quad (1.9)$$

dengan rumus jarak suatu data ke- i dengan semua data pada kluster yang berbeda adalah

$$d_i(k) = \frac{1}{n_k} \sum_{r=1}^{n_k} d(x_i, x_r) \quad (1.10)$$

3. Menghitung nilai *silhouette coefficient* untuk setiap data ke- i

$$SC_1(i) = \frac{b_i - a_i}{\max\{a_i, b_i\}}, i = 1, 2, \dots, n \quad (1.11)$$

Nilai SC dari sebuah kluster ($SC_2(k)$) diperoleh dengan menghitung rata-rata nilai $SC_1(i)$ semua data yang bergabung dalam kluster tersebut dengan menggunakan Persamaan (1.12).

$$SC_2(k) = \frac{1}{n_k} \sum_{x_i \in c_l}^{n_k} SC_1(i) \quad (1.12)$$

Setelah itu nilai SC *global* diperoleh dengan menghitung rata-rata nilai $SC_2(k)$ dari semua kluster dengan menggunakan Persamaan (1.13).

$$SC = \frac{\sum_{k=1}^K (n_k \times SC_2(k))}{\sum_{k=1}^K n_k} \quad (1.13)$$

Hasil perhitungan nilai SC dapat bervariasi antara -1 sampai dengan 1. Jika $SC = 1$ berarti objek x_i sudah berada dalam kluster yang tepat. Jika nilai $SC = 0$ maka objek x_i berada di antara dua kluster. Jika $SC = -1$ artinya struktur kluster yang dihasilkan objek x_i lebih tepat dimasukkan ke dalam kluster yang lain. Nilai *silhouette coefficient* yaitu [7]:

Tabel 2.1 Nilai *silhouette coefficient*

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Dwi Indra Yunistya, Rito Goejantoro, Fidia Deny Tisna Amijaya

No	Rentang Nilai SC	Keterangan
1	$0,7 < SC \leq 1$	<i>Strong Structure</i>
2	$0,5 < SC \leq 0,7$	<i>Medium Structure</i>
3	$0,25 < SC \leq 0,5$	<i>Weak Structure</i>
4	$SC \leq 0,25$	<i>No Structure</i>

1.4 Kemiskinan

Kemiskinan adalah suatu kondisi ketidakmampuan secara ekonomi untuk memenuhi standar hidup rata-rata masyarakat di suatu daerah. Kondisi ketidakmampuan ini ditandai dengan rendahnya pendapatan untuk memenuhi kebutuhan pokok [8]. Pada dasarnya kemiskinan terdapat empat pendekatan absolut, pendekatan relatif, kemiskinan kultural dan kemiskinan struktural. Pendekatan absolut dikaitkan dengan tingkat pendapatan yang dibutuhkan untuk memperoleh kebutuhan dasarnya yaitu makanan, pakaian dan perumahan agar dapat menjamin kelangsungan hidupnya. Sedangkan pendekatan relatif dilihat dari aspek ketimpangan sosial, semakin besar ketimpangan antara tingkat penghidupan golongan atas dan golongan bawah maka akan semakin besar pula jumlah penduduk miskin. Kemiskinan kultural, mengacu pada persoalan sikap seseorang atau masyarakat yang disebabkan oleh faktor budaya, seperti tidak mau berusaha memperbaiki tingkat kehidupan, malas, pemboros, tidak kreatif meskipun ada bantuan dari pihak luar. Kemiskinan struktural, situasi miskin yang disebabkan karena rendahnya akses terhadap sumber daya yang terjadi dalam suatu sistem sosial budaya dan sosial politik yang tidak mendukung pembebasan kemiskinan, tetapi seringkali menyebabkan suburnya kemiskinan [14].

Salah satu penyebab kemiskinan adalah karena kurangnya pendapatan dan aset (*lack of income and assets*) untuk memenuhi kebutuhan dasar seperti makanan, pakaian, perumahan dan tingkat kesehatan dan pendidikan yang dapat diterima (*acceptable*). Di samping itu kemiskinan juga berkaitan dengan keterbatasan lapangan pekerjaan dan biasanya mereka yang dikategorikan miskin (*the poor*) tidak memiliki pekerjaan (pengangguran), serta tingkat pendidikan dan kesehatan mereka pada umumnya tidak memadai. Mengatasi masalah kemiskinan tidak dapat dilakukan secara terpisah dari masalah-masalah pengangguran, pendidikan, kesehatan dan masalah-masalah lain yang secara eksplisit berkaitan erat dengan masalah kemiskinan. Dengan kata lain, pendekatannya harus dilakukan lintas sektor, lintas pelaku secara terpadu dan terkoordinasi dan terintegrasi [3].

2. HASIL DAN PEMBAHASAN

Analisis kluster dengan menggunakan metode *K-Harmonic Means* perlu menentukan parameter (p) yang digunakan dan banyak kluster yang akan terbentuk. Pada penelitian ini, parameter (p) yang digunakan yaitu 2, 3 dan 4. Banyak kluster yang akan dibentuk adalah 2 kluster, 3 kluster dan 4 kluster dan maksimal iterasi (*maxiter*) yang digunakan yaitu 100.

2.1 Analisis Pengelompokan dengan *K-Harmonic Means* 2 Kluster

1. Menentukan pusat kluster awal

Penentuan pusat kluster awal pada bagian ini digunakan banyak kluster (K) yaitu 2. Hasil penentuan pusat kluster awal yang akan digunakan dalam metode *K-Harmonic Means* dengan menggunakan *software R* didapatkan data ke-4 (Kabupaten Pontianak) dan data ke-25 (Kabupaten Gunung Mas).

2. Menghitung fungsi tujuan

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Dwi Indra Yunistya, Rito Goejantoro, Fidia Deny Tisna Amijaya

Menghitung nilai fungsi tujuan menggunakan Persamaan (1.4) dan parameter (p) yang digunakan yaitu 2, 3 dan 4. Berikut hasil nilai fungsi tujuan yang didapatkan pada saat iterasi berhenti, yang dapat dilihat pada Tabel 2.1 sebagai berikut:

Tabel 2.1 Nilai Fungsi Tujuan Pada 2 Klaster dengan Menggunakan $p = 2, 3$ dan 4

Parameter (p)	Nilai Fungsi Tujuan	Iterasi
$p = 2$	439,859	86
$p = 3$	1.317,001	19
$p = 4$	3.852,038	13

3. Menghitung Nilai Keanggotaan

Menghitung nilai keanggotaan di lakukan pada setiap data pengamatan terhadap pusat klaster. Nilai keanggotaan bernilai 0 hingga 1 dan dihitung menggunakan Persamaan (1.5) dan parameter (p) yang digunakan yaitu 2, 3 dan 4. Berikut hasil nilai keanggotaan yang didapatkan pada saat iterasi berhenti, yang dapat dilihat pada Tabel 2.2 sebagai berikut:

Tabel 2.2 Nilai Keanggotaan Pada 2 Klaster dengan Menggunakan $p = 2, 3$ dan 4

Data ke-	$p = 2$		$p = 3$		$p = 4$	
	K1	K2	K1	K2	K1	K2
1	0,233	0,767	0,021	0,979	0,007	0,993
2	0,234	0,766	0,014	0,986	0,003	0,997
⋮	⋮	⋮	⋮	⋮	⋮	⋮
56	0,688	0,312	0,977	0,023	0,989	0,011

4. Menghitung bobot

Menghitung bobot dari setiap data pengamatan dengan menggunakan Persamaan (1.6) dan parameter yang digunakan yaitu $p = 2, 3$ dan 4. Hasil perhitungan bobot ini mendefinisikan seberapa besar pengaruh data pengamatan sehingga ditetapkan keanggotaan data pada suatu klaster. Berikut hasil nilai bobot yang didapatkan pada saat iterasi berhenti, yang dapat dilihat pada Tabel 2.3 sebagai berikut:

Tabel 2.3 Nilai Bobot Data Pada 2 Klaster dengan Menggunakan $p = 2, 3$ dan 4

Data ke-	Nilai Bobot		
	$p = 2$	$p = 3$	$p = 4$
1	0,542	1,420	2,759
2	0,541	1,387	2,466
⋮	⋮	⋮	⋮
56	0,519	1,477	2,842

5. Menghitung pusat klaster baru

Perhitungan ulang pusat klaster baru ini dilakukan dengan menggunakan Persamaan (1.7) dan parameter (p) yang digunakan yaitu $p = 2, 3$ dan 4. Perhitungan ulang pusat klaster baru ini dilakukan secara berulang sesuai dengan tahapan analisis hingga iterasi berhenti.

6. Menetapkan keanggotaan data pengamatan pada suatu klaster

Penetapan keanggotaan data pada suatu klaster ini dilakukan dengan melihat nilai maksimum dari nilai keanggotaan pada saat iterasi berhenti dengan masing-masing parameter (p) yang digunakan. Berikut hasil penetapan keanggotaan data pengamatan pada suatu klaster, yang dapat dilihat pada Tabel 2.4 sebagai berikut:

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Dwi Indra Yunistya, Rito Goejantoro, Fidia Deny Tisna Amijaya

Tabel 2.4 hasil Penetapan Data Pada 2 Klaster dengan Menggunakan $p = 2, 3$ dan 4

Data ke-	Klaster		
	$p = 2$	$p = 3$	$p = 4$
1	2	2	2
2	2	2	2
⋮	⋮	⋮	⋮
56	1	1	1

Berdasarkan Tabel 2.4 hasil yang diperoleh dari penempatan data pengamatan pada suatu klaster dengan menggunakan $p = 2, 3$ dan 4 , maka dapat diketahui banyak data pengamatan yang berada pada masing-masing klaster, yang dapat dilihat pada Tabel 2.5 sebagai berikut:

Tabel 2.5 Hasil Pengelompokan dengan Metode *K-Harmonic Means* 2 klaster

Parameter (p)	Keanggotaan (Kab/Kota)	
	K1	K2
$p = 2$	23	33
$p = 3$	15	41
$p = 4$	11	45

2.2 Analisis Pengelompokan dengan *K-Harmonic Means* 3 Klaster

1. Menentukan pusat klaster awal

Penentuan pusat klaster awal pada bagian ini digunakan banyak klaster (K) yaitu 3. Hasil penentuan pusat klaster awal yang akan digunakan dalam metode *K-Harmonic Means* dengan menggunakan *software R* didapatkan data ke-50 (Kota Samarinda), data ke-48 (Kabupaten Mahakam Ulu) dan data ke-33 (Kabupaten Tapin).

2. Menghitung fungsi tujuan

Menghitung nilai fungsi tujuan menggunakan Persamaan (1.4) dan parameter (p) yang digunakan yaitu 2, 3 dan 4. Berikut hasil nilai fungsi tujuan yang didapatkan pada saat iterasi berhenti, yang dapat dilihat pada Tabel 2.6 sebagai berikut:

Tabel 2.6 Nilai Fungsi Tujuan Pada 3 Klaster dengan Menggunakan $p = 2, 3$ dan 4

Parameter (p)	Nilai Fungsi Tujuan	Iterasi
$p = 2$	427,614	100
$p = 3$	1.238,706	32
$p = 4$	3.510,008	32

3. Menghitung Nilai Keanggotaan

Menghitung nilai keanggotaan di lakukan pada setiap data pengamatan terhadap pusat klaster. Nilai keanggotaan bernilai 0 hingga 1 dan dihitung menggunakan Persamaan (1.5) dan parameter (p) yang digunakan yaitu 2, 3 dan 4. Berikut hasil nilai keanggotaan yang didapatkan pada saat iterasi berhenti, yang dapat dilihat pada Tabel 2.7 sebagai berikut:

Tabel 2.7 Nilai Keanggotaan Pada 3 Klaster dengan Menggunakan $p = 2, 3$ dan 4

Data ke-	$p = 2$			$p = 3$			$p = 4$		
	K1	K2	K3	K1	K2	K3	K1	K2	K3
1	0,025	0,763	0,212	0,001	0,979	0,019	0,0003	0,993	0,007

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Dwi Indra Yunistya, Rito Goejantoro, Fidia Deny Tisna Amijaya

2	0,021	0,684	0,295	0,003	0,867	0,129	0,001	0,914	0,085
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
56	0,923	0,035	0,042	0,959	0,014	0,027	0,953	0,012	0,035

4. Menghitung bobot

Menghitung bobot dari setiap data pengamatan dengan menggunakan Persamaan (1.6) dan parameter yang digunakan yaitu $p = 2, 3$ dan 4 . Hasil perhitungan bobot ini mendefinisikan seberapa besar pengaruh data pengamatan sehingga ditetapkan keanggotaan data pada suatu kluster. Berikut hasil nilai bobot yang didapatkan pada saat iterasi berhenti, yang dapat dilihat pada Tabel 2.8 sebagai berikut:

Tabel 2.8 Nilai Bobot Data Pada 2 Kluster dengan Menggunakan $p = 2, 3$ dan 4

Data ke-	Nilai Bobot		
	$p = 2$	$p = 3$	$p = 4$
1	0,450	0,945	1,201
2	0,436	0,935	1,699
⋮	⋮	⋮	⋮
56	0,547	1,220	2,818

5. Menghitung pusat kluster baru

Perhitungan ulang pusat kluster baru ini dilakukan dengan menggunakan Persamaan (1.7) dan parameter (p) yang digunakan yaitu $p = 2, 3$ dan 4 . Perhitungan ulang pusat kluster baru ini dilakukan secara berulang sesuai dengan tahapan analisis hingga iterasi berhenti.

6. Menetapkan keanggotaan data pengamatan pada suatu kluster

Penetapan keanggotaan data pada suatu kluster ini dilakukan dengan melihat nilai maksimum dari nilai keanggotaan pada saat iterasi berhenti dengan masing-masing parameter (p) yang digunakan. Berikut hasil penetapan keanggotaan data pengamatan pada suatu kluster, yang dapat dilihat pada Tabel 2.9 sebagai berikut:

Tabel 2.9 hasil Penetapan Data Pada 2 Kluster dengan Menggunakan $p = 2, 3$ dan 4

Data ke-	Kluster		
	$p = 2$	$p = 3$	$p = 4$
1	2	2	2
2	2	2	2
⋮	⋮	⋮	⋮
56	1	1	1

Berdasarkan Tabel 2.9 hasil yang diperoleh dari penempatan data pengamatan pada suatu kluster dengan menggunakan $p = 2, 3$ dan 4 , maka dapat diketahui banyak data pengamatan yang berada pada masing-masing kluster, yang dapat dilihat pada Tabel 2.10 sebagai berikut:

Tabel 2.10 Hasil Pengelompokan dengan Metode *K-Harmonic Means* 3 kluster

Parameter (p)	Keanggotaan (Kab/Kota)		
	K1	K2	K3
$p = 2$	12	23	21
$p = 3$	11	21	24
$p = 4$	9	21	26

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Dwi Indra Yunistya, Rito Goejantoro, Fidia Deny Tisna Amijaya

2.3 Analisis Pengelompokan dengan *K-Harmonic Means* 4 Klaster

1. Menentukan pusat klaster awal

Penentuan pusat klaster awal pada bagian ini digunakan banyak klaster (K) yaitu 4. Hasil penentuan pusat klaster awal yang akan digunakan dalam metode *K-Harmonic Means* dengan menggunakan *software R* didapatkan data ke-11 (Kabupaten Kayong Utara), data ke-38 (Kabupaten Tanah Bumbu), data ke-28 (Kota Palangka Raya) dan data ke-43 (Kabupaten Kutai Barat).

2. Menghitung fungsi tujuan

Menghitung nilai fungsi tujuan menggunakan Persamaan (1.4) dan parameter (p) yang digunakan yaitu 2, 3 dan 4. Berikut hasil nilai fungsi tujuan yang didapatkan pada saat iterasi berhenti, yang dapat dilihat pada Tabel 2.11 sebagai berikut:

Tabel 2.11 Nilai Fungsi Tujuan Pada 4 Klaster dengan Menggunakan $p = 2, 3$ dan 4

Parameter (p)	Nilai Fungsi Tujuan	Iterasi
$p = 2$	431,919	53
$p = 3$	1.211,336	32
$p = 4$	3.392,507	31

3. Menghitung Nilai Keanggotaan

Menghitung nilai keanggotaan di lakukan pada setiap data pengamatan terhadap pusat klaster. Nilai keanggotaan bernilai 0 hingga 1 dan dihitung menggunakan Persamaan (1.5) dan parameter (p) yang digunakan yaitu 2, 3 dan 4. Berikut hasil nilai keanggotaan yang didapatkan pada saat iterasi berhenti, yang dapat dilihat pada Tabel 2.12 dan Tabel 2.13 sebagai berikut:

Tabel 2.12 Nilai Keanggotaan Pada 4 Klaster dengan Menggunakan $p = 2$ dan 3

Data ke-	$p = 2$				$p = 3$			
	K1	K2	K3	K4	K1	K2	K3	K4
1	0,801	0,097	0,006	0,095	0,930	0,061	0,001	0,008
2	0,602	0,198	0,007	0,193	0,217	0,743	0,002	0,038
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
56	0,016	0,023	0,939	0,023	0,014	0,028	0,922	0,035

Tabel 2.13 Nilai Keanggotaan Pada 4 Klaster dengan Menggunakan $p = 4$

Data ke-	$p = 4$			
	K1	K2	K3	K4
1	0,951	0,045	0,0003	0,004
2	0,156	0,825	0,001	0,018
⋮	⋮	⋮	⋮	⋮
56	0,013	0,037	0,894	0,056

4. Menghitung bobot

Menghitung bobot dari setiap data pengamatan dengan menggunakan Persamaan (1.6) dan parameter yang digunakan yaitu $p = 2, 3$ dan 4. Hasil perhitungan bobot ini mendefinisikan

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Dwi Indra Yunistya, Rito Goejantoro, Fidia Deny Tisna Amijaya

seberapa besar pengaruh data pengamatan sehingga ditetapkan keanggotaan data pada suatu kluster. Berikut hasil nilai bobot yang didapatkan pada saat iterasi berhenti, yang dapat dilihat pada Tabel 2.14 sebagai berikut:

Tabel 2.14 Nilai Bobot Data Pada 2 Kluster dengan Menggunakan $p = 2, 3$ dan 4

Data ke-	Nilai Bobot		
	$p = 2$	$p = 3$	$p = 4$
1	0,394	0,752	1,154
2	0,329	0,692	1,316
3	0,318	1,121	2,290
⋮	⋮	⋮	⋮
56	0,513	1,051	2,618

5. Menghitung pusat kluster baru
 Perhitungan ulang pusat kluster baru ini dilakukan dengan menggunakan Persamaan (1.7) dan parameter (p) yang digunakan yaitu $p = 2, 3$ dan 4. Perhitungan ulang pusat kluster baru ini dilakukan secara berulang sesuai dengan tahapan analisis hingga iterasi berhenti.
6. Menetapkan keanggotaan data pengamatan pada suatu kluster
 Penetapan keanggotaan data pada suatu kluster ini dilakukan dengan melihat nilai maksimum dari nilai keanggotaan pada saat iterasi berhenti dengan masing-masing parameter (p) yang digunakan. Berikut hasil penetapan keanggotaan data pengamatan pada suatu kluster, yang dapat dilihat pada Tabel 2.15 sebagai berikut:

Tabel 2.15 hasil Penetapan Data Pada 2 Kluster dengan Menggunakan $p = 2, 3$ dan 4

Data ke-	Kluster		
	$p = 2$	$p = 3$	$p = 4$
1	1	1	1
2	1	2	2
⋮	⋮	⋮	⋮
56	3	3	3

Berdasarkan Tabel 2.15 hasil yang diperoleh dari penempatan data pengamatan pada suatu kluster dengan menggunakan $p = 2, 3$ dan 4, maka dapat diketahui banyak data pengamatan yang berada pada masing-masing kluster, yang dapat dilihat pada Tabel 2.16 sebagai berikut:

Tabel 2.16 Hasil Pengelompokan dengan Metode *K-Harmonic Means* 4 kluster

Parameter (p)	Keanggotaan (Kab/Kota)			
	K1	K2	K3	K4
$p = 2$	20	7	11	18
$p = 3$	16	13	10	17
$p = 4$	15	16	9	16

2.4 Evaluasi Hasil Klustering

Pada bagian ini digunakan untuk melihat kualitas dan kekuatan kluster. Metode yang digunakan, yaitu metode *Silhouette Coefficient* (SC). Berikut hasil kualitas hasil klustering berdasarkan hasil pengelompokan yang didapatkan pada masing-masing banyak kluster dan parameter (p) dengan menggunakan Persamaan (1.8), (1.9), (1.10), (1.11), (1.12) dan (1.13), yang dapat dilihat pada Tabel 2.17 sebagai berikut:

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Dwi Indra Yunistya, Rito Goejantoro, Fidia Deny Tisna Amijaya

Tabel 2.17 Hasil Validasi Klastering berdasarkan Nilai *Silhouette Coefficient Global*

Banyak Klaster	Parameter (p)	<i>Silhouette Coefficient</i>
2 Klaster	$p = 2$	0,179
	$p = 3$	0,262
	$p = 4$	0,323
3 Klaster	$p = 2$	0,158
	$p = 3$	0,162
	$p = 4$	0,165
4 Klaster	$p = 2$	0,117
	$p = 3$	0,186
	$p = 4$	0,181

Berdasarkan Tabel 2.17 dapat diketahui bahwa nilai *Silhouette Coefficient (SC)* untuk validasi data hasil pengelompokan data tingkat kemiskinan berdasarkan indikator-indikator kemiskinan di Kabupaten/Kota di Pulau Kalimantan dengan menggunakan metode *K-Harmonic Means* memiliki nilai *SC* yang berbeda-beda. Nilai *SC* yang terbesar adalah pengelompokan pada 2 klaster dengan menggunakan parameter (p) = 4 yaitu sebesar 0,323 yang berarti bahwa hasil pengelompokan tersebut bersifat *weak structure* (struktur yang lemah). Berikut hasil pengelompokan Kabupaten/Kota untuk 2 klaster dengan menggunakan $p = 4$, yang dapat dilihat pada Tabel 2.18 sebagai berikut:

Tabel 2.18 Hasil Validasi Klastering berdasarkan Nilai *Silhouette Coefficient Global*

Klaster	Kabupaten/Kota
1	Kota Pontianak, Kota Palangka Raya, Kota Banjarmasin, Kota Banjar Baru, Kab. Kutai Kartanegara, Kab. Kutai Timur, Kab. Berau, Kota Balikpapan, Kota Samarinda, Kota Bontang dan Kota Tarakan.
2	Kab. Sambas, Kab. Bengkayang, Kab. Landak, Kab. Pontianak, Kab. Sanggau, Kab. Ketapang, Kab. Sintang, Kab. Kapuas Hulu, Kab. Sekadau, Kab. Melawi, Kab. Kayong Utara, Kab. Kubu Raya, Kota Singkawang, Kota Waringin Barat, Kota Waringin Timur, Kab. Kapuas, Kab. Barito Selatan, Kab. Barito Utara, Kab. Sukamara, Kab. Lamandau, Kab. Seruyan, Kab. Katingan, Kab. Pulang Pisau, Kab. Gunung Mas, Kab. Barito Timur, Kab. Murung raya

Tabel 2.18 Hasil Validasi Klastering berdasarkan Nilai *Silhouette Coefficient Global* (lanjutan)

Klaster	Kabupaten/Kota
2	Kab. Tanah Laut, Kab. Kota Baru, Kab. Banjar, Kab. Barito Kuala, Kab. Tapin, Kab. Hulu Sungai Selatan, Kab. Hulu Sungai Tengah, Kab. Hulu Sungai Utara, Kab. Tabalong, Kab. Tanah Bumbu, Kab. Belangan, Kab. Paser, Kab. Kutai Barat, Kab. Penajam Paser Utara, Kab. Mahakam Ulu, Kab. Malinau, Kab. Bulungan, Kab. Tana Tidung dan Kab. Nunukan.

3. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan, maka kesimpulan yang dapat diambil adalah sebagai berikut.

1. Klaster optimal yang terbentuk pada pengelompokan Kabupaten/Kota di Pulau Kalimantan berdasarkan indikator kemiskinan dengan menggunakan metode *K-Harmonic Means* adalah sebanyak 2 klaster dan menggunakan parameter (p) sebesar 4. Hasil yang terbentuk yaitu

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Dwi Indra Yunistya, Rito Goejantoro, Fidia Deny Tisna Amijaya

klaster 1 beranggotakan 11 Kabupaten/Kota dan klaster 2 beranggotakan 45 Kabupaten/Kota.

2. Nilai *Silhouette Coefficient* untuk validasi data hasil klastering Kabupaten/Kota di Pulau Kalimantan berdasarkan indikator kemiskinan dengan menggunakan metode *K-Harmonic Means* yaitu pada 2 klaster dengan menggunakan parameter (p) sebesar 4 didapatkan nilai sebesar 0,323 yang menyatakan bahwa struktur klaster yang dihasilkan pada pengelompokan ini adalah *weak structure* (struktur klaster yang lemah).

REFERENCES

- [1] Agusta, I., 2007. Desa Tertinggal di Indonesia. *Jurnal Transdisiplin Sosiologi, Komunikasi dan Ekologi Manusia*. Vol 1, No 2, 233-252, ISSN:1978-4333.
- [2] Aziz, Nyimas Latifah, L., 2019. Relokasi Ibu Kota Negara: Lesson Learned Dari Negara Lain. *Jurnal Kajian Wilayah 10 (2019) 37-64*.
- [3] Bappenas, 2018. *Analisis Wilayah Dengan Kemiskinan Tinggi*. Jakarta: Kedepuitan Bidang Kependudukan dan Ketenagakerjaan Kementrian PPN/Bappenas.
- [4] Gungor, Z dan Unler, A., 2007. *K-Harmonic Means Data Clustering with Simulated Annealing Heuristic*. *Applied Mathematics and Computation*, Vol. 184, 199-209.
- [5] Jollyta, Deny., Siddik, Muhammad., Mawengkang, Herman., & Efendi, Syahril, 2021. Teknik Evaluasi *Cluster* Solusi Menggunakan Python dan RapidMiner. Yogyakarta: DEEPUBLISH.
- [6] Handoyo, Rendy., M Rumani, R., & Nasution, Surya Michrandi, 2014. Perbandingan Metode Clustering Menggunakan Metode Single Linkage dan K-Means Pada Pengelompokan Dokumen. *JSM STMIK Mikroskil ISSN.1412-0100 Vol 15, No 2 Hal. 73-82*.
- [7] Kaufman, L., & Rousseeuw, P. J., 1990. *Finding Groups in Data*. New York: John Wiley & Sons.
- [8] Marthalina, 2018. Peran Pemerintah Daerah Dalam Mengentaskan Kemiskinan di Kabupaten Tangerang Provinsi Banten. *TRANSFORMASI: JURNAL MANAJEMEN PEMERINTAHAN\VOL. 10 NO.1\ MARET 2018: 1-24*.
- [9] Prasetyo, E., 2012. *Data Mining: Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta: Andi.
- [10] Putri, R. F., Sumantyo, J. T. S., Sukamdi, & Harini, R., 2019. Human And Economic Resource Mapping Analysis to Evaluate the SDGs Accomplishment in South Kalimantan, Indonesia. *Jurnal Geografi Indonesia, 51, 364–384*.
- [11] Rahmatulloh, 2017. DINAMIKA KEPENDUDUKAN DI IBUKOTA JAKARTA (Deskripsi Perkembangan Kuantitas, Kualitas dan Kesejahteraan Penduduk di DKI Jakarta). *GENTA MULIA Volume VIII No. 2, Juli 2017 Page: 54-67*.

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Dwi Indra Yunistya, Rito Goejantoro, Fidia Deny Tisna Amijaya

- [12] Saptadi, N. T. S., Chyan, P., & Taga, V. P., 2021. Using K-Means Algorithm to Investigate Community Behavior in Treating Waste toward Smart City. *International Journal on Advanced Science, Engineering and Information Technology*, 11(4), 1455–1462.
- [13] Srivastava, P., & Mostafavi, A., 2018. Challenges and opportunities of crowdsourcing and participatory planning in developing infrastructure systems of smart cities. *Infrastructures*, 3(4).
- [14] Suryawati, Chriswardani, 2005. Memahami Kemiskinan Secara Multidimensional. *JMPK Vol. 08/No. 03/ September/ 2005, Hal. 121-129*.
- [15] Trisnu, Cokorda, G, S, P., & Suidiana, I Ketut, 2019. Pengaruh Pertumbuhan Penduduk, Pengangguran dan Pendidikan Terhadap Tingkat Kemiskinan Kabupaten/Kota Provinsi Bali. *E-Jurnal Unud*, 8 [11]: 2622 – 2655.
- [16] Widiartha, I. M., 2011. Studi Komparasi Metode Klasterisasi Data K-Means Dan K-Harmonic Means. *Jurnal Ilmu Komputer*, 30-34.
- [17] Zhang, B., Hsu, M., & Dayal, U., 1999. Clustering Algorithm K -Harmonic Means -A Data Clustering Algorithm. *Hewlett-Packard Labs Technical Report HPL-1999-124*.
- [18] Zhu, A., Hua, Z., Shi, Y., Tang, Y., & Miao, L., 2021. An improved k-means algorithm based on evidence distance. *Entropy*, 23(11).