



Penerapan Metode *K-Medoids* Untuk Pengelompokan Tingkat Kemiskinan Pada Kabupaten/Kota Provinsi Kalimantan Timur

¹Gloria Luvita Diazmy

¹Jurusan Matematika dan Teknologi Informasi/Matematika, Institut Teknologi Kalimantan
Corresponding author email: 02211007@student.itk.ac.id

Abstract: Efforts to deal with poverty have become a priority in the government's work plan where this handling is expected to be a good way to overcome poverty. One way is to group the characteristics of each region based on poverty indicators. So that later we can find out which clusters have a big influence on poor communities. Number of poor people in East Kalimantan Province based on a Central Statistics Agency (BPS) survey. The percentage of poor people in East Kalimantan Province in September 2022 rose to 6.44 percent or an increase of 0.13 percentage points compared to March 2022. The problem of poverty is a problem that must be addressed immediately by the government. The term poverty appears when a group of people are unable to meet the level of economic well-being that is considered a minimum need. Therefore, this research was conducted to determine the grouping of Poverty Level Indicators in East Kalimantan Province using *K-Medoids* clustering analysis with the optimum number of *k* using the Silhouette Coefficient and as an illustration for the government for poverty reduction efforts. With the Silhouette coefficient, 2 Clusters are obtained, namely Cluster 1 High Poverty Level with 7 Districts/Cities, namely Paser Regency, West Kutai, Kutai Kartanegara, East Kutai, Berau, North Penajam Paser, and Mahakam Ulu and Cluster 2 Low Poverty Level with 3 Districts/Cities namely Balikpapan, Samarinda and Bontang.

Keywords: Sustainable Development Goals (SDGs), Poverty, Cluster, *K-Medoids*

Abstrak: Upaya penanganan kemiskinan telah menjadi prioritas dalam rencana kerja pemerintah dimana penanganan tersebut diharapkan menjadi cara yang baik untuk mengatasi kemiskinan. Salah satu caranya adalah dengan mengelompokkan karakteristik masing-masing daerah berdasarkan indikator kemiskinan. Sehingga nantinya dapat diketahui cluster yang sangat berpengaruh pada masyarakat kemiskinan. Jumlah penduduk miskin di Provinsi Kalimantan Timur berdasarkan survei Badan Pusat Statistik (BPS) Persentase penduduk miskin di Provinsi Kalimantan Timur pada September 2022 naik menjadi 6,44 persen atau meningkat 0,13 poin persen terhadap Maret 2022. Permasalahan kemiskinan menjadi masalah yang harus segera ditangani oleh pemerintah. Istilah kemiskinan muncul jika sekelompok orang tidak mampu mencukupi tingkat kesejahteraan ekonomi yang dianggap sebagai kebutuhan minimal. Oleh karena itu, penelitian ini dilakukan untuk mengetahui pengelompokan Indikator Tingkat Kemiskinan yang ada di Provinsi Kalimantan Timur dengan menggunakan analisis clustering *K-Medoids* dengan jumlah *k* optimum menggunakan Silhouette Coefficient dan sebagai gambaran bagi pemerintah untuk upaya penanggulangan kemiskinan. Dengan Silhouette coefficient diperoleh 2 Cluster yaitu Cluster 1 Tingkat Kemiskinan Tinggi dengan 7 Kab/Kota yaitu Kabupaten Paser, Kutai Barat, Kutai Kartanegara, Kutai Timur, Berau, Penajam Paser Utara, dan Mahakam Ulu dan Cluster 2 Tingkat Kemiskinan rendah dengan 3 Kab/Kota yaitu Balikpapan, Samarinda, dan Bontang.

Kata kunci: Sustainable Development Goals (SDGs), Kemiskinan, Cluster, *K-Medoids*

I. PENDAHULUAN

Salah satu permasalahan yang hampir dialami oleh semua negara adalah kemiskinan. Secara umum, kemiskinan ialah suatu keadaan seseorang atau kelompok orang yang tidak dapat mencukupi kebutuhan dasarnya. Salah satu penghambat laju pembangunan manusia adalah masalah kemiskinan. Kemiskinan sering kali menjadi penghalang bagi penduduk dalam mengakses kesehatan dan pendidikan. Pemberantasan kemiskinan akan berdampak pada peningkatan kualitas pembangunan manusia, terutama dari sisi ekonomi [1]. Upaya penanganan kemiskinan telah menjadi prioritas dalam rencana kerja pemerintah dimana penanganan tersebut diharapkan menjadi cara yang baik untuk mengatasi kemiskinan. Salah satu caranya adalah dengan mengelompokkan karakteristik masing-masing daerah berdasarkan indikator kemiskinan. Sehingga nantinya dapat diketahui cluster yang sangat berpengaruh pada masyarakat kemiskinan [5].



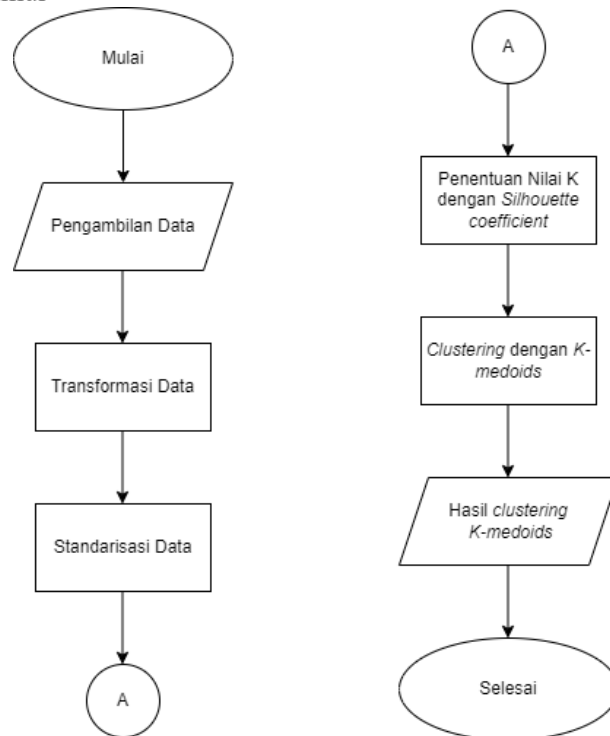
Jumlah penduduk miskin di Provinsi Kalimantan Timur berdasarkan survei Badan Pusat Statistik (BPS) Persentase penduduk miskin di Provinsi Kalimantan Timur pada September 2022 naik menjadi 6,44 persen atau meningkat 0,13 poin persen terhadap Maret 2022 [2]. Oleh karena itu diperlukan adanya pengelompokan Kabupaten/Kota di Kalimantan Timur berdasarkan Indikator Kemiskinan. Salah satu metode yang dapat digunakan untuk pengelompokan variabel atau objek adalah analisis *cluster*. *Clustering* merupakan teknik dalam data mining yang berguna untuk mengelompokkan sekumpulan objek ke dalam beberapa cluster dengan karakteristik yang sama, sehingga objek sebuah cluster mirip tetapi tidak mirip dengan objek dalam cluster yang berbeda. Cluster adalah kumpulan objek yang serupa tetapi berbeda dalam suatu grup dengan benda-benda milik kelompok lain [4]. Pada penelitian ini, metode analisis *clustering* yang dilakukan adalah metode *clustering K-Medoids*.

Pada penelitian ini digunakan untuk mengetahui pengelompokan tingkat kemiskinan pada Kabupaten/Kota di Kalimantan Timur dengan metode *K-Medoids* dan pengoptimalan jumlah *cluster* optimum menggunakan *Silhouette Coefficient* serta mengetahui pengelompokan tingkat kemiskinan pada Kabupaten/Kota di Kalimantan Timur dengan metode *K-Medoids* dapat menjadi gambaran untuk upaya penanggulangan kemiskinan di Kalimantan Timur. *K-Medoids Clustering* merupakan variasi metode *K-Means*. Perbedaan metode ini dibandingkan dengan *K-Means* yaitu pada penggunaan *medoids* bukan berdasarkan rata - rata setiap objek *cluster* [3].

II. METODE PENELITIAN

Adapun metode yang digunakan untuk mengelompokkan tingkat kemiskinan pada Kabupaten/Kota di Kalimantan Timur dengan menggunakan metode *K-Medoids*. Adapun prosedur penelitian yang dilakukan pada penelitian ini meliputi sebagai berikut :

1. Pengambilan Data, pada langkah ini yaitu melakukan pengambilan data yang berkaitan dengan indikator kemiskinan.
2. Statistika Deskriptif, pada langkah ini mengolah data yang disajikan secara deskriptif. Data yang disajikan berupa mean, median, standar deviasi.
3. Transformasi Data, pada langkah ini diambil data indikator kemiskinan dari tahun 2020 – 2022 dan dilakukan transformasi data untuk menggabungkan data tersebut.
4. Standarisasi Data, pada langkah ini dilakukan standarisasi data untuk menyamakan data yang satuannya berbeda.
5. Penentuan Nilai *K (cluster)* dengan *Silhouette Coefficient*, pada langkah ini dilakukan penentuan nilai *K* agar diketahui jumlah kluster yang terbentuk.
6. *Clustering* dengan *K-Medoids*, pada langkah ini dilakukan pengelompokan kabupaten/kota sesuai jumlah kluster yang terbentuk.
7. Hasil *clustering K-Medoids*, pada langkah ini dihasilkan pengelompokan kabupaten/kota sesuai jumlah kluster terbentuk.



Gambar 1. Flowchart Alur Pengolahan Data

III. HASIL DAN PEMBAHASAN

Variabel penelitian yang digunakan dalam penelitian ini adalah data Kabupaten/Kota di Kalimantan Timur yaitu data yang digunakan untuk menentukan tingkat kemiskinan di Kalimantan Timur. Adapun indikator yang digunakan pada tabel 1, sebagai berikut:

Tabel 1. Variabel Penelitian

Variabel	Keterangan
X_1	Presentase Jumlah Penduduk Miskin Kalimantan Timur berdasarkan Kabupaten/Kota
X_2	Pengeluaran Per Kapita Kalimantan Timur berdasarkan Kabupaten/Kota disesuaikan (Ribu Rupiah/Kapita/Tahun)
X_3	Presentase Harapan Lama Sekolah Kalimantan Timur berdasarkan Kabupaten/Kota
X_4	Presentase Rata –Rata Lama Sekolah Kalimantan Timur berdasarkan Kabupaten/Kota
X_5	Presentase Angka Harapan Hidup Kalimantan Timur berdasarkan Kabupaten/Kota
X_6	Presentase Pengangguran di Kalimantan Timur berdasarkan Kabupaten/Kota

Data yang digunakan pada penelitian ini adalah data yang diambil dari publikasi Statistik Indonesia dan diolah BPS. Data yang digunakan adalah dari tahun 2020 hingga 2022. Data yang diperoleh akan diolah dan di klaterisasi (*clustering*) menggunakan algoritma *K-Medoids*.



III.1. Transformasi Data

Pada langkah ini dilakukan transformasi data dari tahun 2020 – 2022. Langkah awal yang dilakukan pada transformasi data yaitu dengan menentukan rata-rata setiap variabel dari tahun 2020 hingga 2022. Sehingga didapatkan hasil dari rata-rata setiap variabel pada tabel 2 , sebagai berikut.

Tabel 2. Transformasi Data

Kabupaten/Kota	X_1	X_2	X_3	X_4	X_5	X_6
Paser	9,46	10733	13,21	8,71	72,80	5,03
Kutai barat	9,91	9991	13,00	8,65	73,02	4,91
Kutai Kartanegara	7,75	11148	13,61	9,23	72,54	5,17
Kutai Timur	9,55	10892	12,93	9,35	73,36	5,76
Berau	5,57	12516	13,33	9,53	72,23	5,31
Penajam Paser Utara	7,41	11591	12,61	8,38	71,60	3,76
Mahakam Ulu	11,63	7847	12,58	8,17	72,26	3,02
Balikpapan	2,64	14817	14,20	10,84	74,68	8,28
Samarinda	4,87	15341	15,03	10,56	74,46	7,73
Bontang	4,51	15273	13,13	10,80	74,48	9,06

III.2. Statistika Deskriptif

Karakteristik dan gambaran dari data Indikator Kemiskinan perlu diketahui sebelum melakukan analisis. Statistika deskriptif yang diperoleh meliputi rata-rata, standar deviasi, varians, nilai minimum, median, dan nilai maksimum yang ditunjukkan pada tabel 3, sebagai berikut.

Tabel 3. Statistika Deskriptif

Statistika Deskriptif	X_1	X_2	X_3	X_4	X_5	X_6
Min.	2,640	7847	12,58	8,170	71,60	3,020
Stadev	2,865	2474	0,7541	1,000	1,076	1,955
Kuartil 1	5,045	10772	12,95	8,666	72,33	4,941
Kuartil 2	7,580	11370	13,17	9,292	72,91	5,237
Mean	7,330	12015	13,36	9,423	73,14	5,804
Kuartil 3	9,527	14242	13,54	10,303	74,68	7,240
Max.	11,630	15341	15,03	10,837	74,68	9,063

III.3. Standarisasi Data

Pada *clustering* diperlukan standarisasi data untuk menyamakan data yang satuannya berbeda, data indikator kemiskinan di Provinsi Kalimantan Timur terdapat data dengan satuan yang berbeda, maka dilakukan standarisasi data menggunakan *Z-Score*. Langkah yang pertama untuk melakukan standarisasi data yaitu menghitung rata-rata dan standar deviasi pada setiap kolom variabel, didapatkan rata-rata dan standar deviasi pada tabel 4, sebagai berikut.

Tabel 4. Rata-Rata dan Standar Deviasi

X_n	Rata-Rata (μ)	Standar Deviasi (σ)
X_1	7,33	2,865352
X_2	12014,90	2473,9917
X_3	13,36	0,752793
X_4	9,42	0,99961
X_5	73,14	1,075258
X_6	5,80	1,95618

Selanjutnya dilakukan perhitungan *Z-score* dengan menggunakan rumus *Z-score* pada persamaan (1), yaitu:

$$Z = \frac{(X_n - \mu)}{\sigma} \quad (1)$$



Didapatkan hasil standarisasi data dari perhitungan menggunakan persamaan (1) pada tabel 5, sebagai berikut.

Tabel 5. Standarisasi Data

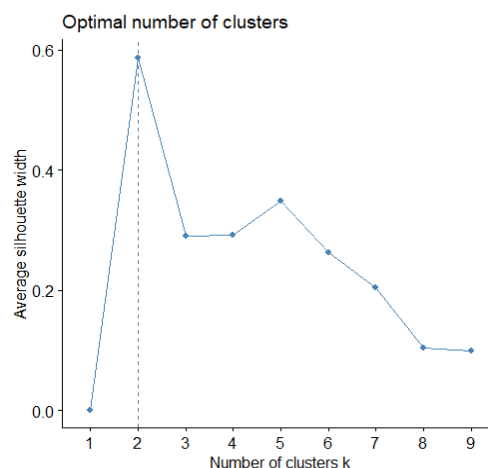
Kabupaten/Kota	ZX_1	ZX_2	ZX_3	ZX_4	ZX_5	ZX_6
Paser	0,743	-0,518	-0,206	-0,710	-0,322	-0,394
Kutai barat	0,900	-0,818	-0,476	-0,773	-0,115	-0,457
Kutai Kartanegara	0,147	-0,350	0,325	-0,193	-0,558	-0,326
Kutai Timur	0,775	-0,454	-0,574	-0,069	0,205	-0,022
Berau	-0,614	0,203	-0,038	0,107	-0,846	-0,254
Penajam Paser Utara	0,028	-0,171	-0,999	-1,040	-1,435	-1,043
Mahakam Ulu	1,501	-1,685	-1,039	-1,253	-0,822	-1,423
Balikpapan	-1,637	1,133	1,109	1,415	1,426	1,266
Samarinda	-0,859	1,344	2,211	1,138	1,221	0,986
Bontang	-0,984	1,317	-0,313	1,378	1,246	1,666

III.4. Penentuan Nilai-K Clustering dengan Silhouette Coefficient

Dalam menentukan jumlah *cluster* (k) dengan *Silhouette coefficient* digunakan pendekatan nilai rata – rata untuk memprediksi kualitas *cluster* yang terbentuk. Semakin tinggi nilai rata-rata yang diperoleh maka *cluster* yang terbentuk akan semakin baik. Didapatkan hasil menggunakan *silhouette coefficient* dilakukan dengan jumlah *cluster* $K=2$ dengan nilai *silhouette* = 0,605108. Perhitungan dan nilai rata-rata *silhouette coefficient* untuk semua *cluster* dapat dilihat pada Tabel 6 dan Gambar 2 dibawah ini.

Tabel 6. Perhitungan Silhouette Coefficient

Kabupaten/Kota	$a(i)$	$b(i)$	$s(i)$
Paser	1,3069	4,9937	1,3069
Kutai barat	1,4196	5,2626	1,4196
Kutai Kartanegara	1,6899	4,3449	1,6899
Kutai Timur	1,5937	4,5524	1,5937
Berau	2,4162	4,0151	2,4162
Penajam Paser Utara	2,2847	5,8256	2,2847
Mahakam Ulu	2,5791	6,9480	2,5791
Balikpapan	1,6969	5,0668	1,6969
Samarinda	2,7614	4,9755	2,7614
Bontang	1,9976	4,7180	1,9976
<i>Silhouette coefficient</i>			0,6051

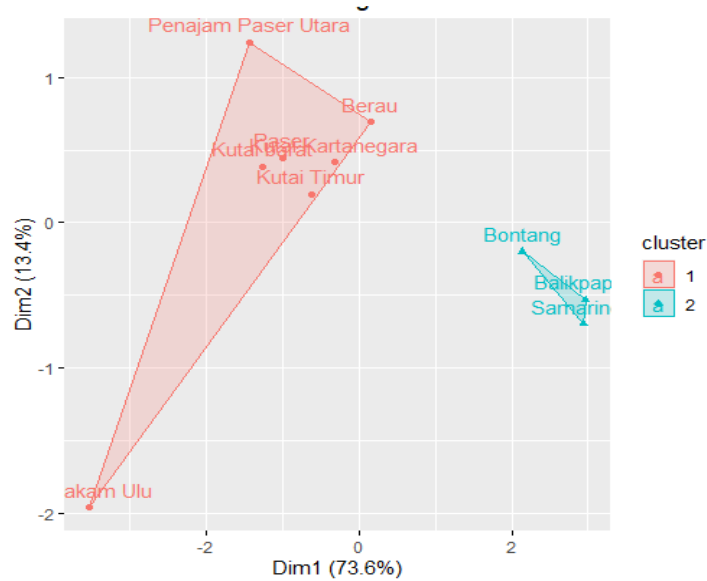


Gambar 2. Penentuan nilai – k Silhouette Coefficient

Berdasarkan Gambar 2, dapat diketahui bahwa nilai optimal dari *cluster* adalah 2. Sehingga, dengan menggunakan *Silhouette coefficient* ini diperoleh nilai *K* yang optimal sebesar 2 *cluster*.

III.5. Hasil Clustering K-Medoids

Adapun hasil *clustering* yang terbentuk dengan menggunakan nilai $K=2$, diperoleh bahwa *cluster* 1 (tingkat kemiskinan tinggi) terdiri 7 Kabupaten dan *cluster* 2 (tingkat kemiskinan rendah). Pengelompokan Kabupaten/Kota dapat dilihat pada gambar 3 dan tabel 7 sebagai berikut.



Gambar 3. Cluster Plot

Hasil *cluster* pada Gambar 3 yaitu warna merah merupakan *cluster* 1 dan bagian berwarna biru adalah *cluster* 2. Adapun hasil pengelompokan Kabupaten/Kota sebagai berikut.

Tabel 7. Pengelompokan Kabupaten/Kota

Kabupaten/Kota	Cluster
Paser	1
Kutai barat	
Kutai Kartanegara	
Kutai Timur	
Berau	
Penajam Paser	
Mahakam Ulu	
Balikpapan	2
Samarinda	
Bontang	

Berdasarkan gambar 3 dan tabel 7 diketahui bahwa *cluster* 1 (tingkat kemiskinan tinggi) yang terdiri dari 7 Kabupaten di Kalimantan Timur yaitu Paser, Kutai Barat, Kutai Kartanegara, Kutai Timur, Berau, Penajam Paser, Mahakam Ulu. Kemudian untuk *cluster* 2 (tingkat kemiskinan rendah) yang terdiri dari 3 Kota di Kalimantan Timur yaitu Balikpapan, Samarinda, dan Bontang.

Kemudian untuk melihat karakteristik dari setiap *cluster* dapat dilihat melalui nilai rata – rata setiap masing – masing indikator, sebagai berikut.



Tabel 8. Karakteristik *cluster*

<i>Cluster</i>	X_1	X_2	X_3	X_4	X_5	X_6
1	8,75	10674	13	8,86	72,5	4,71
2	4,01	15144	14,1	10,7	74,5	8,36

Pada tabel 8 dapat dilihat bahwa *cluster* 1 memiliki tingkat kemiskinan tinggi daripada *cluster* 2 yang memiliki tingkat kemiskinan yang rendah. Maka dari itu penanggulangan kemiskinan bisa lebih difokuskan terlebih dahulu ke daerah kabupaten dan tetap menjaga kestabilan pada daerah perkotaan.

IV. KESIMPULAN

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan pada pengolahan data, adapun kesimpulan yang diperoleh yaitu Pengelompokan tingkat kemiskinan yang terbentuk dengan melakukan penentuan jumlah *cluster* optimum pada *Silhouette Coefficient* diperoleh nilai $k = 2$, *Cluster* 1 merupakan kelompok dengan tingkat kemiskinan tinggi yang terdiri dari Paser, Kutai Barat, Kutai Kartanegara, Kutai Timur, Berau, Penajam Paser Utara, dan Mahakam Ulu, dan *cluster* 2 merupakan kelompok dengan tingkat kemiskinan rendah yang terdiri dari Balikpapan, Samarinda, dan Bontang. Pada hasil pengelompokan tingkat kemiskinan pada Kabupaten/Kota di Kalimantan Timur dengan metode *K-Medoids* dapat dilihat bahwa upaya penanggulangan kemiskinan dapat dimulai dengan meningkatkan Indeks Pembangunan Manusia pada *cluster* 1 terlebih dahulu, namun tetap mempertahankan Indeks Pembangunan Manusia pada *cluster* 2.

REFERENSI

1. Badan Pusat Statistik Provinsi Kalimantan Timur (2022). Indeks Pembangunan Manusia.
2. Badan Pusat Statistik Provinsi Kalimantan Timur (2023). Berita Resmi Statistik No. 36/07/64/Th.XXVI, 17 Juli 2023..
3. Farissa, R.A., Mayasari, R. & Umaidah, Y. 2021. Perbandingan Algoritma K-Means dan K-Medoids Untuk Pengelompokan Data Obat dengan Silhouette Coefficient. *Journal of Applied Informatics and Computing (JAIC)*, 5(2): 109-116.
4. Sepriyanti, N., Nahampun, R. S., Zikri, M. H., Ambarani, I., & Rahmadeyan, A. (2022, September). Penerapan K-Means Clustering Untuk Mengelompokkan Tingkat Kemiskinan di Provinsi Riau: Implementation of K-Means Clustering to Group Poverty Levels in Riau Province. In *SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat* (pp. 59-65).
5. Wira, B., Budianto, A.E. & Wiguna, A.S. 2019. Implementasi Metode K-Medoids Clustering Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru Tahun 2018 di Universitas Kanjuruhan Malang. *Jurnal Terapan Sains & Teknologi*, 1(3): 54-69.