



# Penerapan K-Means Clustering Untuk Mengelompokkan Tingkat Kemiskinan Di Provinsi Kalimantan Barat

**Samsul Dwi Cahyo<sup>1</sup>, Irmawati Tri Wahyuni<sup>2</sup>, Revalyna Octavia Maharani<sup>3</sup>, Maulana Nurfaizi<sup>4</sup>, Rujianto Eko Saputro<sup>5</sup>, Tarwoto<sup>6</sup>**

<sup>1,2,3,4,5,6</sup> Sistem Informasi, FIK, Universitas AMIKOM Purwokerto, Indonesia

Email: [samsuldwicahyo@gmail.com](mailto:samsuldwicahyo@gmail.com)<sup>1</sup>, [irmawahyunikl@gmail.com](mailto:irmawahyunikl@gmail.com)<sup>2</sup>, [revalynarevalyna@gmail.com](mailto:revalynarevalyna@gmail.com)<sup>3</sup>, [maulanafaizi81@gmail.com](mailto:maulanafaizi81@gmail.com)<sup>4</sup>, [rujianto@amikompurwokerto.ac.id](mailto:rujianto@amikompurwokerto.ac.id)<sup>5</sup>, [tarwoto@amikompurwokerto.ac.id](mailto:tarwoto@amikompurwokerto.ac.id)<sup>6</sup>

## Abstract

*This study aims to use the K-means clustering algorithm to categorize poverty levels in the West Kalimantan province. The data used for clustering represents poverty levels across four districts: Melawi, Kapuas Hulu, Sekadau, and Kayong Utara. The K-means clustering method is employed to group these districts based on similarities in their poverty levels. The clustering results reveal four distinct categories of poverty levels: Cluster 0 represents areas with very high poverty rates; Cluster 1 shows Melawi with a high poverty rate; Cluster 2 includes Sambas, Kapuas Hulu, and Sintang, with relatively low poverty rates; and Cluster 3 includes Landak, Sanggau, and Ketapang, with high poverty rates. The analysis reveals interesting patterns in the distribution of poverty across West Kalimantan, which can assist local governments in designing more effective policies for poverty reduction. This study makes a significant contribution to understanding poverty dynamics in West Kalimantan and provides a basis for more efficient decision-making in poverty alleviation efforts.*

**Keywords:** Clustering, Data Mining, K-Means, West Kalimantan

## Abstrak

*Penelitian ini bertujuan menggunakan algoritma pengelompokan K-means untuk mengelompokkan tingkat kemiskinan di provinsi Kalimantan Barat. Data yang digunakan mengelompokkan tingkat kemiskinan dari empat kabupaten: Melawi, Kapuas Hulu, Sekadau, dan Kayong Utara. Metode pengelompokan K-means digunakan untuk mengelompokkan kabupaten-kabupaten ini berdasarkan kesamaan dalam tingkat kemiskinan. Hasil pengelompokan mengungkap empat kategori yang berlainan dalam tingkat kemiskinan: Cluster 0 menunjukkan tingkat kemiskinan yang amat tinggi; Cluster 1 menampilkan Melawi dengan tingkat kemiskinan yang tinggi; Cluster 2 melibatkan Sambas, Kapuas Hulu, dan Sintang dengan tingkat kemiskinan yang cukup rendah; Cluster 3 melibatkan Landak, Sanggau, dan Ketapang dengan tingkat kemiskinan yang tinggi. Analisis menunjukkan corak menarik dalam penyebaran kemiskinan di Kalimantan Barat, yang bisa membantu pemerintah daerah merancang kebijakan yang lebih efisien dalam pengurangan kemiskinan. Penelitian ini memberikan sumbangan penting dalam memahami dinamika kekurangan di Kalimantan Barat serta memberikan dasar bagi pengambilan keputusan yang lebih efisien dalam upaya pengurangan kemiskinan.*

**Kata kunci:** Clustering, Data Mining, K-Means, Kalimantan Barat

## 1. PENDAHULUAN

Kemiskinan adalah masalah besar yang mengkhawatirkan dunia. Saat ini, banyak negara non-industri mengalami pertumbuhan ekonomi yang cepat tetapi tidak memberikan manfaat bagi masyarakatnya, sementara negara-negara yang kurang beruntung justru menghadapi masalah distribusi pendapatan dan kemajuan yang tidak konsisten[1]. Meskipun Indonesia bukan negara industri,

kebutuhan tetap penting. Kemiskinan adalah suatu kondisi di mana seseorang tidak dapat menangani masalah dasar seperti makanan, pakaian, pendidikan, dan berbagai kebutuhan lainnya[2]. Ketidakmampuan seseorang atau sekelompok orang untuk memenuhi hak-hak dasar mereka untuk menjalani kehidupan yang baik dan bermartabat juga dikenal sebagai kemiskinan. Akses terhadap informasi kemiskinan yang tepat dan khusus merupakan bagian penting dari teknologi yang bertujuan untuk mengurangi kebutuhan bantuan[3]. Menurut data BPS tahun 2019, ada beberapa provinsi di mana angka kemiskinan meningkat. Di antara lima provinsi paling miskin di Indonesia, kemiskinan di Kalimantan Barat meningkat sebesar 4.444, yang terletak di bagian barat Kalimantan Barat. Sebanyak 4.444 orang yang tersisa berada di wilayah tengah dan 4.444 di wilayah timur Indonesia. Berdasarkan data dari 4.444 BPS, Kalimantan Barat masing-masing menempati peringkat ke-21 dan ke-4.444 dengan proporsi penduduk miskin sebesar 7,49%. Lima provinsi lain di Kalimantan menempati peringkat pertama, dan Kalimantan Barat menempati peringkat satu, yang berarti ada 4.444 orang miskin yang tinggal di Kalimantan Barat. Misalnya, kemiskinan di Kalimantan meningkat pada tingkat berikut: Kalimantan Selatan 4,55%, Kalimantan Tengah 5,98%, Kalimantan Timur 5,94%, Kalimantan Utara 6,63%, dan Kalimantan Barat 7,49%[4]. Dalam penelitian sebelumnya yang dilakukan oleh Yunita et al. (2020), yang disebut sebagai "Penerapan algoritma K-Means pada klaster data kemiskinan Provinsi Banten menggunakan RapidMiner," ditemukan bahwa terdapat tiga klaster. Kota Tangerang, Kota Cilegon, dan Kota Serang termasuk dalam Klaster 0, sedangkan Pandeglang, Lebak, dan Serang termasuk dalam Klaster 1, dan Tangerang dan Tangerang Selatan termasuk dalam Klaster 2[5]. Berdasarkan masalah di atas, penulis sangat tertarik untuk melakukan penelitian yang akan mengumpulkan data kemiskinan di setiap kabupaten di Provinsi Kalimantan Barat. Untuk menentukan tingkat kemiskinan di suatu wilayah, perlu dibuat cara yang lebih mudah untuk mengkategorikan daerah yang masih sangat miskin. Studi ini menggunakan metode *clustering* untuk data mining, yang menggambarkan pola nilai dengan mengelompokkan nilai yang memiliki kesamaan. Penelitian ini bertujuan untuk menggambarkan setiap kabupaten di Provinsi Kalimantan Barat dengan tingkat kebutuhan yang rendah, sedang, dan tinggi, sehingga pemerintah dapat membuat kebijakan dan strategi pembangunan yang sesuai dengan keadaan masing-masing. Berdasarkan masalah di atas, penulis ingin melakukan penelitian yang akan mengumpulkan informasi tentang kemiskinan di setiap kabupaten di Provinsi Kalimantan Barat. Perlu dibuat metode yang lebih mudah untuk mengkategorikan wilayah yang masih sangat miskin untuk menentukan tingkat kemiskinan di suatu wilayah. Metode data mining yang digunakan dalam penelitian ini adalah *clustering*, yang menggambarkan pola nilai dengan mengelompokkan nilai yang memiliki kesamaan. Tujuan dari penelitian ini adalah untuk menggambarkan setiap kabupaten di Provinsi Kalimantan Barat dengan tingkat kebutuhan rendah, sedang, dan tinggi, sehingga pemerintah dapat membuat kebijakan dan strategi pembangunan yang sesuai dengan situasinya.

## 2. METODOLOGI PENELITIAN

Data yang digunakan berasal dari dokumen yang menunjukkan jumlah penduduk miskin di masing-masing Kabupaten atau Kota Provinsi Kalimantan Barat dari tahun 2019 hingga 2021, yang dihasilkan oleh publikasi Statistik Indonesia yang diolah oleh BPS. Setelah diproses, data dikelompokkan menggunakan algoritma K-Means dan dibagi menjadi empat *cluster*. Gambar 1 menunjukkan langkah-langkah yang diambil dalam penelitian ini.

### 2.1. Pengambilan Data

Penelitian mengumpulkan data untuk mendapatkan informasi penting untuk mencapai tujuan penelitian[6]. Data yang kami gunakan berasal dari dokumen yang menunjukkan jumlah penduduk miskin di Provinsi Riau di masing-masing wilayah regional atau perkotaan, Data dari tahun 2019–2021 digunakan dari publikasi statistik Indonesia dan diolah oleh Badan Pusat Statistik. Setelah itu, data akan diolah sebelum *clustering* dilakukan. Untuk membuat informasi lebih akurat dan dapat diakses oleh orang lain, pengolahan data digunakan[7]. Pada langkah ini, semua data penggunaan dikategorikan berdasarkan tiga kriteria penilaian dengan menghitung rata-ratanya:

- 1) Jumlah Penduduk Miskin
- 2) Persentase Penduduk Miskin
- 3) Garis Kemiskinan



Gambar 1. Alur Proses Penelitian

### 2.2. Clustering

*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. Ada dua metode *clustering*, yaitu hirarkis *clustering* dan pengelompokan partisi. Data dikelompokkan berdasarkan grafik hierarkis dalam metode pengelompokan hierarkis, dimana dua grup terdekat digabung atau semua data dikelompokkan menjadi *cluster*. Dalam pengelompokan partisi pengelompokan data tidak memiliki hierarki apapun, setiap *cluster* memiliki *centroid*, tujuannya adalah untuk meminimalkan jarak semua data ke *centroid*[8].

### 2.3. K-Means

K-Means adalah sebuah metode pengelompokan data menjadi dua atau lebih dari kelompok. Algoritma *K-Means* merupakan metode analisis kelompok yang membagi subjek penelitian ke dalam kelompok-kelompok, dimana setiap subjek yang akan diamati berada dalam satu kelompok data dengan rata-rata yang berdekatan satu sama lain. Seperti yang kita inginkan, K digunakan untuk konstanta *clustering* total dan Means berarti mean dari dataset, dalam hal ini sebagai *cluster*, jadi *K-means Clustering* adalah metode analisis data yang menggunakan sistem partisi untuk pengelompokan data[9]. *K-Means* merupakan metode pengelompokan data non-hierarkis bagi data menjadi satu atau satu bentuk lebih banyak cluster atau grup untuk data dengan karakteristik yang sama dikelompokkan dalam *cluster* data identik dan karakteristik kelompok yang berbeda adalah kelompok lain[10].

## 3. HASIL DAN PEMBAHASAN

### 3.1. Pengolahan Data

Adapun data yang digunakan adalah data kabupaten yang ada di Provinsi Kalimantan Barat dengan atribut berupa Jumlah Penduduk Miskin, Persentase Penduduk Miskin, Garis Kemiskinan tahun 2023, yang diperoleh dari Badan Pusat Statistik Provinsi Kalimantan barat yang telah dilakukan proses *cleaning*.

**Tabel 1. Pengolahan Data**

Kabupaten	Jumlah Penduduk Miskin	(%) Penduduk Miskin	P1	P2	Garis Kemiskinan
Sambas	38,17	7,08	1,10	0,26	498.209
Bengkayang	16,87	6,28	0,98	0,26	416.148
Landak	38,92	9,97	1,32	0,28	451.681
Mempawah	14,15	5,21	0,52	0,08	444.790
Sanggau	23,34	4,79	0,55	0,11	416.112
Ketapang	49,95	9,25	1,84	0,50	539.108
Sintang	35,49	8,18	1,16	0,31	644.010
Kapuas Hulu	22,59	8,16	1,51	0,46	543.301
Sekadau	12,09	5,90	0,91	0,19	392.195
Melawi	24,18	11,12	2,47	0,83	640.608
Kayong Utara	10,77	9,13	1,71	0,47	365.262
Kubu Raya	25,33	4,23	0,56	0,15	492.642
Pontianak	29,92	4,45	0,49	0,08	677.521
Singkawang	11,06	4,70	0,74	0,18	588.561
Total	353,35	6,71	1,03	0,24	563.288

### 3.2. Clustering

Tahapan selanjutnya yaitu melakukan *clustering* tingkat kemiskinan di Provinsi Kalimantan Barat berdasarkan Jumlah Penduduk Miskin, Persentase Penduduk Miskin, Kedalaman, Keparahan, dan Garis Kemiskinan di Provinsi Kalimantan Barat dengan *tools Rapidminer*. Adapun algoritma yang digunakan yaitu *K-Means* dengan nilai K yang ditetapkan yaitu 4 untuk menghasilkan 4 *cluster*. Penggunaan algoritma *K-Means* membentuk 4 *cluster* dan menghasilkan jumlah Kabupaten pada setiap *cluster* yaitu *cluster 0* sebanyak 1 Kabupaten, *cluster 1* sebanyak 2 Kabupaten, *cluster 2* sebanyak 6 Kabupaten dan *cluster 3* sebanyak 4 Kabupaten. Adapun nilai *centroid* yang dihasilkan dapat dilihat pada Tabel 2.

Tabel 2. Clustering

Kabupaten	Jumlah Penduduk Miskin	Persentase Penduduk Miskin	P1	P2	Garis Kemiskinan	Cluster
Sambas	38.71	07.08	01.01	00.26	498.209	Cluster3
Bengkayang	0,727083	06.28	0,068056	00.26	416.148	Cluster2
Landak	38.92	0,442361	01.32	00.28	451.681	Cluster3
Mempawah	14.15	05.21	00.52	00.08	444.79	Cluster2
Sanggau	23.34	0,221528	00.55	00.11	416.112	Cluster2
Ketapang	49.95	09.25	0,1	00.05	539.108	Cluster3
Sintang	35.49.00	08.18	01.16	00.31	644.000	Cluster1
Kapuas-Hulu	22.59	08.16	01.51	00.46	543.301	Cluster1
Sekadau	12.09	05.09	0,063194	00.19	392.195	Cluster2
Melawi	24.18.00	11.12	02.47	0,057639	640.608	Cluster0
Kayong-Utara	0,470139	09.13	0,090972	00.47	365.262	Cluster3
Kubu-Raya	25.33.00	04.23	00.56	00.15	492.642	Cluster2
Pontianak	29.92	04.45	00.49	00.08	677.512	Cluster2
Singkawang	11.06	04.07	0,051389	00.18	588.561	Cluster2
Sambas	38.71	07.08	01.01	00.26	498.209	Cluster3

### 4. SIMPULAN

Berdasarkan hasil *clustering* data kemiskinan di Provinsi Kalimantan Barat, terdapat empat *cluster* dengan karakteristik berbeda. *Cluster 0*, yang mencakup Kabupaten Melawi, memiliki tingkat kemiskinan sangat tinggi dengan jumlah penduduk miskin mencapai 24,18% dan garis kemiskinan sebesar 640.608. *Cluster 1*, yang terdiri dari Kabupaten Sintang dan Kapuas Hulu, menunjukkan tingkat kemiskinan tinggi dengan persentase penduduk miskin masing-masing 8,18% dan 8,16%, serta garis kemiskinan sebesar 644.01 dan 543.301. *Cluster 2*, yang mencakup Bengkayang, Mempawah, Sanggau, Sekadau, Kubu Raya, Pontianak, dan Singkawang, memiliki tingkat kemiskinan relatif lebih rendah dengan persentase penduduk miskin berkisar antara 4,07% hingga 6,28%, dan garis kemiskinan antara 392.195 hingga 677.512. *Cluster 3*, yang terdiri dari Kabupaten Sambas, Landak, Ketapang, dan Kayong Utara, menunjukkan tingkat kemiskinan sedang dengan persentase penduduk miskin antara 7,08% hingga 9,25%, dan garis kemiskinan antara 365.262 hingga 539.108. Pemahaman tentang distribusi kemiskinan ini dapat membantu merancang kebijakan yang lebih efektif dan tepat sasaran untuk mengurangi kemiskinan di setiap *cluster*.

#### DAFTAR PUSTAKA

- [1] T. Agus Triono and R. C. Sangaji, "Faktor Mempengaruhi Tingkat Kemiskinan di Indonesia: Studi Literatur Laporan Data Kemiskinan BPS Tahun 2022," *J. Soc. Bridg.*, vol. 1, no. 1, pp. 59–67, 2023, doi: 10.59012/jsb.v1i1.5.
- [2] N. Islami and A. Anis, "Pengaruh Upah Minimum Provinsi, Pendidikan Dan Kesehatan Terhadap Kemiskinan Di Indonesia," *J. Kaji. Ekon. dan Pembang.*, vol. 1, no. 3, p. 939, 2019, doi: 10.24036/jkep.v1i3.7721.
- [3] G. Anuraga, "Hierarchical Clustering Multiscale Bootstrap untuk Pengelompokkan Kemiskinan di Jawa Timur," *Statistika*, vol. 1, pp. 27–33, 2015.
- [4] S. Beni, "Pemberdayaan Anggota Credit Union Sebagai Upaya Pengentasan Kemiskinan Di Kalimantan Barat," *Sebatik*, vol. 25, no. 1, pp. 195–200, 2021, doi: 10.46984/sebatik.v25i1.1339.
- [5] Y. R. Sari, A. Sudewa, D. A. Lestari, and T. I. Jaya, "Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 5, no. 2, p. 192, 2020, doi: 10.24114/cess.v5i2.18519.
- [6] A. Dubey and M. Arif Kamal, "Evaluating Energy Efficiency in Traditional Architecture with Reference to Criterion 8 of GRIHA," *Am. J. Civ. Eng. Archit.*, vol. 10, no. 1, pp. 16–22, 2022, doi: 10.12691/ajcea-10-1-3.
- [7] A. Rusmana, "Pengantar Pengolahan Data dan Informasi," *Skom4437/Modul*, p. 2, 2015.
- [8] H. Nabila, D. Retno, and S. Saputro, "Clustering Data Campuran Numerik dan Kategorik Menggunakan Algoritme Ensemble Quick ROBust Clustering using linKS ( QROCK )," *Prism. Pros. Semin. Nas. Mat.*, vol. 5, no. 1, pp. 716–720, 2022.
- [9] S. Regina, E. Sutinah, and N. Agustina, "Clustering Kualitas Kinerja Karyawan Pada Perusahaan Bahan Kimia Menggunakan Algoritma K-Means," *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 573, 2021, doi: 10.30865/mib.v5i2.2909.
- [10] E. Muningsih, I. Maryani, and V. R. Handayani, "Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa," *J. Sains dan Manaj.*, vol. 9, no. 1, p. 96, 2021.