

# Analisis Kemiskinan Menggunakan Metode Algoritma Clustering K-Means

Dwiki Rasya Rahadian<sup>1\*</sup>, Nurmalitasari<sup>2</sup>

<sup>1,2</sup>Sistem Informasi

Universitas Duta Bangsa Surakarta

<sup>1\*</sup>202021001@mhs.udb.ac.id,<sup>2</sup>nurmalitasari@udb.ac.id

*Abstrak*— Kemiskinan memiliki dampak yang luas dan serius terhadap kehidupan individu dan masyarakat. Ketika seseorang hidup dalam kemiskinan, mereka mungkin menghadapi kesulitan untuk memenuhi kebutuhan dasar, seperti makanan yang cukup, tempat tinggal yang layak, dan pendidikan yang layak. Keterbatasan ini dapat berdampak negatif pada kesehatan fisik dan mental, pendidikan, kesempatan kerja, dan kualitas hidup secara keseluruhan. Tujuan penelitian ini yaitu mengetahui pengelompokan kabupaten/kota yang mempunyai kesamaan karakteristik berdasarkan indikator kemiskinan tahun 2019. Penelitian ini menggunakan data yang didapatkan dari BPS (Badan Pusat Statistik). Metode yang digunakan ialah metode k-means clustering yang merupakan metode partisi klastering guna pengelompokan objek ke dalam k cluster. Berdasarkan hasil penelitian, diperoleh pengelompokan karakteristik masing-masing cluster yang dibentuk berdasarkan nilai indikator kemiskinan di beberapa kabupaten/kota pada tahun 2019 sebanyak 2 cluster. Terbentuk dari 20 kabupaten/kota pada klaster 1 dan 29 kabupaten/kota pada klaster 2. Klaster 1 memiliki karakteristik Tantangan Kerja Rendah, dengan Tingkat Pengeluaran Perkapita Rendah dan Tingkat Tidak Bekerja Rendah sedangkan pada Klaster 2 memiliki karakteristik Tantangan Kerja Tinggi, dengan Tingkat Pengeluaran Perkapita Tinggi dan Tingkat Tidak Bekerja Tinggi.  
*Kata kunci*— Klaster, K-means, Kemiskinan

*Abstrak*— Poverty has a broad and serious impact on the lives of individuals and society. When people live in poverty, they may face difficulties meeting basic needs, such as adequate food, adequate housing, and proper education. These limitations can negatively impact physical and mental health, education, employment opportunities, and overall quality of life. The purpose of this study is to find out the grouping of districts/cities that have similar characteristics based on the 2019 poverty indicators. This research uses data obtained from the BPS (Central Bureau of Statistics). The method used is the k-means clustering method which is a clustering partition method for grouping objects into k clusters. Based on the research results, the characteristics of each cluster were grouped based on the poverty indicator values in several districts/cities in 2019 as many as 2 clusters. Formed from 20 districts/cities in cluster 1 and 29 districts/cities in cluster 2. Cluster 1 has the characteristics of Low Work Challenges, with Low Per Capita Expenditure Rates and Low Unemployment Rates while Cluster 2 has the characteristics of High Job Challenges, with Per Capita Expenditure Levels High and High Non Working Rate.

*Keywords*— Clusters, K-means, Poverty.

## I. PENDAHULUAN

Kemiskinan adalah salah satu tantangan yang dihadapi oleh banyak negara di dunia. Meskipun manusia telah mencapai kemajuan yang luar biasa dalam berbagai aspek kehidupan, masih terdapat jutaan orang yang hidup dalam kondisi kemiskinan yang ekstrim. Kemiskinan bukan hanya tentang kurangnya pendapatan, tetapi juga mencakup akses terbatas terhadap pendidikan, layanan kesehatan, perumahan layak, air bersih, dan kebutuhan dasar lainnya. Fenomena ini memiliki dampak yang merusak pada individu, keluarga, dan masyarakat secara keseluruhan.

Kemiskinan dapat ditemukan di seluruh dunia, baik di negara berkembang maupun negara maju. Di negara-negara berkembang, masalah kemiskinan sering kali menjadi lebih mencolok karena kurangnya infrastruktur, akses terhadap sumber daya, serta ketimpangan sosial dan ekonomi yang besar. Di

sisi lain, di negara maju, kemiskinan sering kali terkait dengan masalah seperti pengangguran struktural, pekerjaan yang tidak stabil, serta kesenjangan pendapatan yang semakin melebar.

Kemiskinan memiliki dampak yang luas dan kompleks terhadap individu dan masyarakat. Orang yang hidup dalam kemiskinan sering kali menghadapi kesulitan untuk memenuhi kebutuhan dasar seperti makanan, perumahan yang layak, dan pendidikan. Kondisi ini dapat mempengaruhi kesehatan fisik dan mental, pertumbuhan anak-anak, dan peluang pendidikan. Selain itu, kemiskinan juga dapat menciptakan lingkaran setan di mana sulit bagi individu untuk keluar dari kemiskinan karena keterbatasan akses terhadap peluang dan sumber daya.

Berdasarkan permasalahan kemiskinan di Indonesia, telah dilakukan penelitian dari beberapa peneliti terdahulu diantaranya yaitu penelitian dari

(Wahyudi et al., 2019) yang bertujuan untuk melihat tinggi rendahnya tingkat kemiskinan di Indonesia dan diperoleh tingkat kemiskinan tertinggi di Indonesia beranggotakan 23 provinsi dan tingkat kemiskinan terendah beranggotakan 11 provinsi. Dari uraian di atas, membuat peneliti tertarik untuk mengelompokkan wilayah di beberapa daerah yang memiliki karakteristik sama berdasarkan indikator-indikator kemiskinan menggunakan metode K-Means Clustering, sehingga dapat dengan mudah menemukan solusi atau mengambil kebijakan untuk mengurangi tingkat kemiskinan.

## II. METODOLOGI PENELITIAN

Penelitian ini menggunakan Data dan Informasi Kemiskinan Kabupaten/Kota Tahun 2019 yang didapatkan dari Badan Pusat Statistik (BPS). Indikator kemiskinan yang digunakan ada enam indikator yaitu Jumlah Penduduk Miskin, Pendidikan, Tidak Bekerja, Air Layak, Pengeluaran Perkapita, dan Garis Kemiskinan. Metode yang digunakan ialah metode K-Means Clustering Analysis.

### a. K-means

Penambahan data adalah teknologi baru yang sangat berguna yang dapat membantu Perusahaan menemukan informasi yang sangat penting di gudang data mereka. Beberapa Aplikasi penambahan data fokus pada prediksi, mereka memprediksi apa yang akan terjadi di masa depan situasi baru dari data yang menggambarkan masa lalu.

Algoritma K-Means menghasilkan titik centroid Konsisten dengan tujuan dari algoritma K-Means. Ketika iterasi k-means berhenti, masing-masing objek dalam dataset menjadi anggota kluster. Nilai kluster ditentukan dengan mencari semua objek untuk menemukan kluster yang paling dekat dengan objek. Algoritma K-Means mengelompokkan elemen data dari sebuah dataset ke kluster berdasarkan jarak terpendek.

### b. Klustering

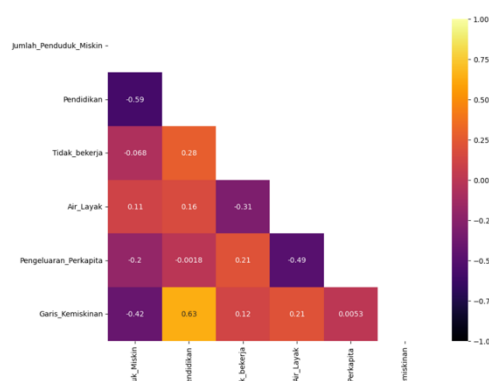
Klustering, dalam konteks analisis data, adalah proses mengelompokkan objek atau data menjadi kelompok-kelompok yang memiliki kesamaan tertentu. Tujuannya adalah untuk menemukan pola

atau struktur yang ada dalam data, tanpa memiliki label atau kategori sebelumnya.

Klustering biasanya digunakan dalam berbagai bidang, seperti ilmu komputer, statistik, bioinformatika, dan ilmu sosial. Beberapa contoh penggunaan clustering termasuk segmentasi pasar, analisis genetik, pengelompokan pengguna dalam aplikasi web, dan analisis tingkah laku konsumen.

## III. HASIL DAN PEMBAHASAN

Hasil yang diperoleh melalui Analisis untuk data kemiskinan yang didapat dari BPS (Badan Pusat Statistik) pada tahun 2019



Gambar 1. Plot korelasi antar variabel

Nilai Korelasi yang akan dibahas secara tuntas adalah nilai korelasi antara variabel Tidak bekerja dan Pengeluaran perkapita terhadap variabel-variabel lain.

Selanjutnya digunakan metode K-MEANS, untuk mengkluster tantangan kerja di Indonesia Per-Kabupaten/Kota Berdasarkan Pengeluaran Perkapita & Tingkat Tidak Bekerja.

K	Silhouette Scores
0	2 0.523291
6	8 0.433331
5	7 0.428369
8	10 0.416121
4	6 0.409883
7	9 0.384463
1	3 0.368901
3	5 0.331324
2	4 0.306834

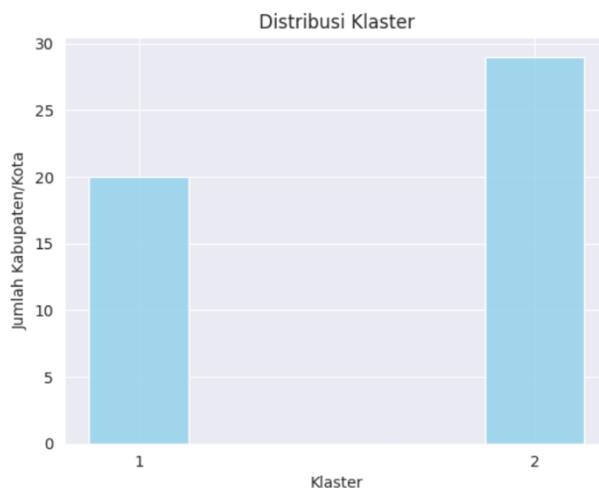
Gambar 2. Silhouette Scores untuk menentukan K Optimal

Berdasarkan kandidat  $K = 2, 3, \dots, 10$ , diperoleh Silhouette Scores tertinggi sebesar 0.523291 untuk  $K = 2$ . Oleh karena itu, akan dibentuk 2 kluster tantangan kerja.

Kluster	Pengeluaran_Perkapita		Tidak_bekerja	
	mean	median	mean	median
1	63.336000	64.03	33.122000	33.77
2	64.588621	64.90	47.347586	46.56

Gambar 3. Karakterisasi Tiap Kluster

Terbentuk 2 Kluster dengan:  
 Kluster 1: Tantangan Kerja Rendah, dengan Tingkat Pengeluaran Perkapita Rendah dan Tingkat Tidak Bekerja Rendah.  
 Kluster 2: Tantangan Kerja Tinggi, dengan Tingkat Pengeluaran Perkapita Tinggi dan Tingkat Tidak Bekerja Tinggi.



Gambar 4. Distribusi Kluster

Berdasarkan plot di atas, dapat dilihat bahwa masih sedikit Kabupaten/Kota di Indonesia yang masuk kedalam kluster tantangan kerja rendah, serta lebih di dominasi pada kluster tantangan kerja yang tinggi.

#### IV. KESIMPULAN

Dengan menggunakan analisis k-means klustering diperoleh kluster optimal sebanyak 2 kluster. Kluster 1 menunjukkan tantangan kerja yang rendah sedangkan kluster 2 menunjukkan tantangan kerja yang tinggi.

Hasil yang di dapat dari penelitian ini diharapkan mampu membantu pihak lainnya dalam mengambil keputusan dan menemukan informasi yang dibutuhkan untuk memecahkan masalah dalam hal kemiskinan berdasarkan kabupaten/kota yang berada di Indonesia serta dapat mengetahui mana saja yang harus mendapatkan perhatian lebih terhadap kluster yang ditentukan.

#### UCAPAN TERIMA KASIH

Alhamdulillah, segala puji dan rasa syukur penulis panjatkan kehadirat Allah SWT atas segala Rahmat dan karunia-nya sehingga karya ilmiah ini dapat terselesaikan dengan baik, lancar, dan tepat waktu

#### REFERENSI

- [1] A. D. A. F. E. W. Febiyanti Alfiah, "Analisis Clustering K-Medoids Berdasarkan Indikator Kemiskinan di Jawa Timur Tahun 2020," *Jurnal Ilmiah Sains*, pp. 1-7, 2022.

- [2] A. F. F. P. S. Achmad Bahauddin, "ANALISIS CLUSTERING PROVINSI DI INDONESIA BERDASARKAN TINGKAT KEMISKINAN MENGGUNAKAN ALGORITMA K-MEANS," *MISI(Jurnal Manajemen informatika & Sistem Informasi)*, 2021.
- [3] F. S. Kamila Aprilia, "ANALISIS GARIS KEMISKINAN MAKANAN MENGGUNAKAN METODE ALGORITMA K-MEANS CLUSTERING," *SISMATIK (Seminar Nasional Sistem Informasi dan Manajemen Informatika)*, 2021.
- [4] "ANALISIS KASUS KEMISKINAN DI PROVINSI KALIMANTAN TENGAH DENGAN PENDEKATAN PRINCIPAL COMPONENT ANALYSIS," *JURNAL AL-QARDH*, 2022.
- [5] D. V. Ferezagia, "Analisis Tingkat Kemiskinan di Indonesia," *Jurnal Sosial Humaniora Terapan*, 2018.
- [6] I. Juani, "DAMPAK PENERAPAN PRINCIPAL COMPONENT ANALYSIS (PCA) DALAM CLUSTERING CURAH HUJAN DI PULAU JAWA, BALI DAN LOMBOK," 2014.
- [7] G. M. Tinungki, "PENENTUAN INDIKATOR KEMISKINAN PENDUDUK INDONESIA TAHUN 2017 DENGAN PEMODELAN SPARSE PRINCIPAL COMPONENT ANALYSIS," *SEMNAS*, 2018.
- [8] F. R. E. M. Derita Lamtiar Pasaribu, "NALISIS FAKTOR KEMISKINAN KABUPATEN/KOTA DI KALIMANTAN, SULAWESI, BALI DAN NUSA TENGGARA," *Jurnal Dinamika Pertanian Edisi XXXV*, pp. 107-116, 2019.
- [9] D. K. Noor Zuhdiyaty, "ANALISIS FAKTOR - FAKTOR YANG MEMPENGARUHI KEMISKINAN DI INDONESIA SELAMA LIMA TAHUN TERAKHIR (Studi Kasus Pada 33 Provinsi)".
- [10] "Penerapan Sparse Principal Component Analysis dalam Menghasilkan Matriks Loading yang Spars," *JMSK*, pp. 44-54, 2019.