PENGELOMPOKAN KABUPATEN/KOTA DI KALIMANTAN BARAT BERDASARKAN TINGKAT KEMISKINAN DENGAN *K-MEDOIDS* DENGAN JARAK *MANHATTAN*

RIZHA AYNUL FIQRIANI
NIM H1091211031

SKRIPSI

PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK
2025

PENGELOMPOKAN KABUPATEN/KOTA DI KALIMANTAN BARAT BERDASARKAN TINGKAT KEMISKINAN DENGAN *K-MEDOIDS* DENGAN JARAK *MANHATTAN*

RIZHA AYNUL FIQRIANI
NIM H1091211031

Skripsi
Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Statistika pada Program Studi Statistika

PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK
2025

LEMBAR PENGESAHAN

Judul Tugas Akhir	:	Pengelompokan Kabupaten/Kota di Kalimantan Barat Berdasarkan Tingkat Kemiskinan dengan *K-Medoids* dengan Jarak *Manhattan*
Nama Mahasiswa	:	Rizha Aynul Fiqriani
NIM	:	H1091211031
Jurusan/Program Studi	:	Matematika/Statistika
Tanggal Lulus	:	11 Juli 2025
SK Pembimbing	:	No. 2609/UN22.8/TD.06/2024 /Tanggal 2 September 2024
SK Penguji	:	No. 2061/UN22.8/TD.06/2025 /Tanggal 4 Juli 2025

Dosen Pembimbing

Pembimbing I	Pembimbing II
Nurfitri Imro'ah, M.Si.	Dr. Evy Sulistianingsih, M.Sc.
NIP 198907182019032021	NIP 198502172008122006

Dosen Penguji

Ketua Penguji	Anggota Penguji
Shantika Martha, M.Si	Hendra Perdana, M.Sc
NIP 198403082008122003	NIP 198810102019031020

Pimpinan Sidang (merangkap anggota penguji)	Sekretaris Sidang (merangkap anggota penguji)
Nurfitri Imro'ah, M.Si.	Dr. Evy Sulistianingsih, M.Sc.
NIP 198907182019032021	NIP 198502172008122006

Mengesahkan
Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Tanjungpura

Prof. Dr. Gusrizal, M.Si.
NIP 197108022000031001

PERNYATAAN INTEGRITAS AKADEMIK

Saya yang bertanda tangan di bawah ini:

Nama	: Rizha Aynul Fiqriani
NIM	: H1091211031
Jurusan/Program Studi	: Matematika/Statistika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam

dengan ini menyatakan bahwa dokumen ilmiah Tugas Akhir yang disajikan ini tidak mengandung unsur pelanggaran integritas akademik sesuai Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia Nomor 39 Tahun 2021. Apabila dikemudian hari dokumen ilmiah Tugas Akhir ini mengandung pelanggaran integritas akademik sesuai ketentuan perundangan tersebut, maka saya bersedia menerima sanksi akademik dan/atau sanksi hukum yang berlaku.

Demikian pernyataan ini untuk digunakan sebagaimana mestinya.

Pontianak, 11 Juli 2025



Rizha Aynul Fiqriani

NIM H1091211031

Pengelompokan Kabupaten/Kota di Kalimantan Barat Berdasarkan Tingkat Kemiskinan Menggunakan *K-Medoids* dengan Jarak *Manhattan*

Abstrak

Kemiskinan merupakan salah satu permasalahan utama yang dihadapi oleh banyak negara di dunia, termasuk Indonesia. Ada beberapa faktor yang menyebabkan terjadinya kemiskinan yaitu keterbatasan kesempatan kerja dan persaingan yang tinggi. Proses pembangunan ekonomi di Kalimantan Barat masih belum mencapai peningkatan kesejahteraan masyarakat secara merata, yang menunjukkan kemiskinan di wilayah tersebut. Penelitian ini bertujuan untuk mengelompokkan daerah di Kalimantan Barat berdasarkan tingkat kemiskinan menggunakan metode *cluster*. Data yang digunakan merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Kalimantan Barat tahun 2023 dengan beberapa indikator kemiskinan. Pengelompokan dilakukan menggunakan metode *K-Medoids,* yaitu metode *cluster* non hirarki yang merupakan varian dari *K-Means*. Proses pengelompokan dimulai dengan pemilihan *medoid* secara acak, kemudian perhitungan jarak dilakukan menggunakan jarak *manhattan* dan hasil pengelompokan di evaluasi dengan *Davies Bouldin Index* (DBI) sehingga diperoleh jumlah *cluster* optimal yaitu empat *cluster*. *Cluster* 1 terdiri dari Kota Pontianak yang memiliki tingkat kemiskinan sangat rendah. *Cluster* 2 terdiri dari Bengkayang, Sanggau dan Sekadau dengan tingkat kemiskinan rendah. *Cluster* 3 terdiri dari Sambas, Landak, Mempawah, Sintang, Kapuas Hulu, Kayong Utara, Kubu Raya, dan Kota Singkawang yang memiliki tingkat kemiskinan tinggi. Dan *cluster* 4 terdiri dari Ketapang dan Melawi yang memiliki tingkat kemiskinan sangat tinggi. Berdasarkan hasil ini, diperlukan kebijakan yang disesuaikan dengan karakteristik masing-masing *cluster*, seperti pemberian bantuan sosial, peningkatan akses pendidikan, serta pembangunan infrastruktur dan penciptaan lapangan kerja.

Kata Kunci: Kemiskinan, Pengelompokan , *K-Medoids*, *Davies Bouldin Index*

Grouping Districts/Cities in West Kalimantan Based on Proverty Level Using K-Medoids with Manhattan Distance

Abstract

Poverty is one of the main problems faced by many countries in the world, including Indonesia. There are several factors that cause poverty, namely limited job opportunities and high competition. The economic development process in West Kalimantan has not yet achieved an increase in community welfare evenly, which indicates poverty in the region. This study aims to group areas in West Kalimantan based on poverty levels using the cluster method. The data used are secondary data obtained from the West Kalimantan Central Statistics Agency (BPS) in 2023 with several poverty indicators. Clustering is carried out using the K-Medoids method, which is a non-hierarchical cluster method which is a variant of K-Means. The clustering process begins with random selection of medoids, then the distance calculation is carried out using the Manhattan distance and the clustering results are evaluated with the Davies Bouldin Index (DBI) so that the optimal number of clusters is four clusters. Cluster 1 consists of Pontianak City which has a very low poverty rate. Cluster 2 consists of Bengkayang, Sanggau and Sekadau with low poverty rates. Cluster 3 consists of Sambas, Landak, Mempawah, Sintang, Kapuas Hulu, Kayong Utara, Kubu Raya, and Singkawang City which have high poverty rates. And cluster 4 consists of Ketapang and Melawi which have very high poverty rates. Based on these results, policies are needed that are adjusted to the characteristics of each cluster, such as providing social assistance, increasing access to education, as well as infrastructure development and job creation.

Keywords*: Poverty, Grouping, K-Medoids, Davies Bouldin Index*

PRAKATA

Puji syukur atas kehadiran Allah SWT, berkat limpahan rahmat, hidayah dan karunia-Nya, penulis bisa menyelesaikan skripsi dengan judul "Pengelompokan Kabupaten/Kota di Kalimantan Barat Berdasarkan Tingkat Kemiskinaan Menggunakan *K-Medoids* dengan Jarak *Manhattan*". Skripsi ini adalah salah satu syarat dari beberapa persyaratan untuk memperoleh gelar sarjana Statistika pada Program Studi Statistika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Tangjungpura.

Penyusunan skripsi ini tidak lepas dari bimbingan, bantuan, arahan dan dukungan dari berbagai pihak terkait. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terima kasih kepada:

1. Kedua orang tua penulis yang tercinta, yang selalu memberikan semangat dan doa yang selalu mengalir mendoakan kelancaran penulis dalam menyelesaikan pendidikan.
2. Abang dan adik penulis yang selalu memberikan dukungan dan menjadi *support system* penulis selama penulisan skripsi.
3. Ibu Nurfitri Imro'ah, M.Si., selaku Dosen Pembimbing Pertama yang telah meluangkan waktu untuk membimbing, memberi masukan, serta motivasi yang bermanfaat bagi penulis dalam menjalankan proses penulisan skripsi.
4. Ibu Evy Sulistianingsih, M.Si., selaku Dosen Pembimbing Kedua yang telah memberikan saran dan masukan untuk melakukan perbaikan penulisan skripsi sehingga menjadi lebih baik.
5. Ibu Shantika Martha, M.S., selaku Dosen Penguji Pertama yang telah memberikan saran dan masukan untuk melakukan perbaikan penulisan tugas akhir sehingga menjadi lebih baik.
6. Bapak Hendra Perdana, M.Sc., selaku Dosen Penguji Kedua yang telah memberikan saran dan masukan untuk melakukan perbaikan penulisan tugas akhir sehingga menjadi lebih baik.
7. Ibu Nur'ainul Miftahul Huda, M.Si., selaku Dosen Pembimbing Akademik yang telah memberikan saran dan motivasi kepada penulis selama proses perkuliahan.

8. Adelea Delvi Deanovela, Rahmanisa Cahyani, Dini Dwi Savitri, dan Mely Amara Putri yang telah menemani penulis selama masa perkuliahan dan menjadi *support system* saat penulis menyusun tugas akhir.
9. Teman-teman "Expos" yang berjuang bersama di Program Studi Statistika dan selalu memberikan dorongan serta motivasi hingga bertahan sampai saat ini.
10. Terakhir, penulis ingin mengucapkan terima kasih kepada diri sendiri, Rizha Aynul Fiqriani, terima kasih karena telah bertahan sejauh ini. Terima kasih karena tetap memilih untuk melanjutkan, walau seringkali tidak tahu pasti kemana arah ini akan membawa. Terima kasih karena telah menjadi teman paling setia bagi diri sendiri. Terima kasih karena sudah mempercayai proses, meski hasil belum sesuai harapan. Terima kasih karena tetap jujur pada rasa takut, namun tidak membiarkan rasa takut itu membatasi langkah, karena keberanian bukanlah ketiadaan rasa takut, melainkan keinginan untuk tetap bergerak meski takut masih melekat erat, dan paling penting, terima kasih karena sudah berani memilih, memilih untuk mencoba, memilih untuk belajar, dan memilih untuk menyelesaikan apa yang telah kamu mulai.

Penulis berharap semoga skripsi ini dapat bermanfaat dan dapat menjadi kemajuan ilmu pengetahuan.

Pontianak, 11 Juli 2025

Rizha Aynul Fiqriani
H1091211031

DAFTAR ISI

Abstrak ..... **iv**

Abstract ..... **v**

PRAKARTA ..... **vi**

DAFTAR ISI ..... **viii**

DAFTAR GAMBAR ..... **x**

DAFTAR TABEL ..... **xi**

DAFTAR SIMBOL ..... **xii**

DAFTAR LAMPIRAN ..... **xiii**

BAB I PENDAHULUAN **1**

1.1 Latar Belakang 1

1.2 Rumusan Masalah 3

1.3 Tujuan Penelitian 3

1.4 Batasan Masalah 4

1.5 Tinjauan Pustaka 4

1.6 Metodologi Penelitian 5

BAB II LANDASAN TEORI **10**

2.1 Kemiskinan 10

2.2 Analisis Multivariat 10

2.3 Analisis *Cluster* 11

2.4 Pendeteksian *Outlier* 12

2.5 Standarisasi Data 12

2.6 Uji Multikolinearitas 13

2.7 Jarak *Manhattan* 13

BAB III *K-MEDOIDS CLUSTERING* DAN *DAVIES BOULDIN INDEX* (DBI) 15

3.1 *K-Medoids* 15

3.2 Validasi Hasil *Cluster* 16

BAB IV HASIL DAN PEMBAHASAN 18

4.1 Statistik Deskriptif 18

4.2 Pendeteksian *Outlier* 19

4.3 Standarisasi Data 20

4.4 Uji Multikolinearitas 21

4.5 Perhitungan Jarak *Manhattan* 22

4.6 Pembentukan *Cluster* dengan Metode *K-Medoids* 23

4.7 Evaluasi *Cluster* dengan Metode *Davies Bouldin Index* (DBI) 24

4.8 Interpretasi *Cluster* 27

BAB V KESIMPULAN 30

DAFTAR PUSTAKA 31

LAMPIRAN 33

DAFTAR GAMBAR

Gambar 1. 1 *Flowchart* Penelitian 9
Gambar 4. 1 *Boxplot* Indikator Tingkat Kemiskinan di Kalimantan Barat ... 20
Gambar 4. 2 Peta Tingkat Kemiskinan di Kalimantan Barat 29

DAFTAR TABEL

Tabel 4. 1 Variabel Penelitian..... 18
Tabel 4. 2 Statistik Deskriptif..... 18
Tabel 4. 3 Nilai VIF Data Standarisasi..... 21
Tabel 4. 4 Nilai VIF Data Standarisasi Setelah Penghapusan Variabel 22
Tabel 4. 5 Nilai *Davies Bouldin Index* (DBI)..... 26
Tabel 4. 6 Nilai Rata-rata Setiap Variabel pada Masing-masing *Cluster*..... 27
Tabel 4. 7 Pengelompokkan Kabupaten/Kota di Kalimantan Barat..... 28

DAFTAR SIMBOL

Z	: nilai standarisasi
x_a	: data ke-*a*, dengan *a* = 1, 2, 3,…, *N*
$\bar{x}$	: rata-rata data
s	: standar deviasi data
$d_{a,b}$	: jarak antara objek ke-*a* dengan objek ke-*b*, dengan *b* = 1, 2, 3, …, *N*
x_{al}	: nilai objek ke-*a* pada variabel ke-*l*, dengan *l* = 1, 2, 3, …, n
x_{bl}	: nilai objek ke-*b* pada variabel ke-*l*
m_j	: banyaknya objek pada *cluster* ke-*j*
$d(\ c_i, c_j)$	: jarak *centroid cluster* ke-*i* ke *centroid cluster* ke-*j*, dengan *i*,*j*=1,2,.., *K*
K	: jumlah *cluster*
N	: banyaknya objek
$R_{i,j}$	: nilai rasio
S	: total jarak
n	: banyak variabel yang diamati

DAFTAR LAMPIRAN

Lampiran 1 Data Penelitian 33
Lampiran 2 Standarisasi Data 34
Lampiran 3 Jarak *Manhattan* 35

BAB I
PENDAHULUAN

1.1 Latar Belakang

Kemiskinan merupakan salah satu permasalahan utama yang dihadapi oleh banyak negara di dunia, termasuk Indonesia. Kondisi ini terjadi ketika individu atau kelompok masyarakat tidak mampu memenuhi kebutuhan dasar mereka, seperti makanan yang cukup, pakaian yang layak, tempat tinggal yang memadai, akses terhadap pendidikan, serta layanan kesehatan yang berkualitas (Widodo, Ermayani, Laila, dan Madani, 2021). Untuk mengatasi permasalahan ini, pemerintah menjalankan berbagai program dan kebijakan, seperti bantuan sosial, subsidi kebutuhan pokok yang bertujuan meningkatkan kesejahteraan masyarakat miskin.

Berdasarkan Badan Pusat Statistik (BPS), jumlah penduduk miskin di Indonesia pada tahun 2022 adalah 26,36 juta orang pada September 2022. Jumlah ini meningkat 0,20 juta orang dari Maret 2022. Indonesia telah mencapai kemajuan besar dalam mengurangi kemiskinan dalam sepuluh tahun terakhir. Berdasarkan tujuan ini, berbagai kegiatan pembangunan harus dilakukan untuk mengembangkan wilayah yang masih tertinggal. Setiap daerah harus menetapkan tujuan pembangunan nasional dengan mempertimbangkan pembangunan jangka panjang dan jangka pendek, serta melaksanakan pembangunan sesuai dengan kebutuhan daerah masing-masing. Oleh karena itu, kemiskinan tidak hanya menjadi masalah nasional saja, tetapi juga meluas ke seluruh wilayah Indonesia.

Ada beberapa faktor yang menyebabkan kemiskinan yaitu taraf hidup masyarakat yang buruk, upah minimum yang tidak memadai dan peningkatan jumlah pengangguran tanpa kesempatan kerja tambahan setiap tahunnya (Made, Kusumawati, dan Yasin, 2024). Salah satu penyebab utama kemiskinan yang berasal dari dalam penduduk miskin adalah kualitas sumber daya manusia yang rendah untuk mengelola sumber daya alam. Kualitas sumber daya manusia yang rendah menyebabkan keterampilan masyarakat untuk mengelola sumber daya alam berkurang, sehingga pengetahuan dan teknik untuk mengolah sumber daya alam tidak digunakan secara optimal. Selain itu, beberapa penyebab kemiskinan yang

berasal dari luar penduduk miskin yaitu keterbatasan sumber daya alam, keterbatasan kesempatan kerja dan persaingan yang tinggi.

Proses pembangunan ekonomi di Kalimantan Barat masih belum mencapai peningkatan kesejahteraan masyarakat secara merata, yang menunjukkan kemiskinan di wilayah tersebut (Sari & Novianti, 2024). Maka diperlukan penelitian untuk mengetahui faktor yang memengaruhi kemiskinan dalam upaya mengurangi kemiskinan. Salah satu langkah penting untuk mencapai kesejahteraan masyarakat yang merata adalah memahami masalah yang dihadapi oleh masing-masing daerah dan menemukan solusi yang tepat. Pengelompokan kabupaten/kota di Kalimantan Barat berdasarkan tingkat kemiskinan masyarakat dapat dilakukan untuk mengetahui kabupaten/kota mana yang berada di kelompok atau *cluster* dengan tingkat kemiskinan tinggi atau rendah. Hal ini dilakukan agar pemerintah dapat membuat berbagai kebijakan yang lebih efisien dan efektif untuk mengatasi berbagai masalah kemiskinan di Kalimantan Barat.

Pengelompokan kabupaten/kota di Kalimantan Barat berdasarkan tingkat kemiskinan harus dilakukan dengan metode yang efektif dan tepat. Analisis *cluster* adalah teknik yang bertujuan untuk mengelompokkan objek ke dalam kelompok yang lebih mirip satu sama lain. Metode ini dapat digunakan untuk melakukan pengelompokan ini. Secara umum, analisis *cluster* dibagi menjadi dua metode yaitu metode *non-hierarki* dan metode *hierarki*. Metode yang termasuk dalam metode *non-hierarki* adalah metode *K-Means* (Paramadina, Sudarmin, dan Aidid, 2019). *K-Means* adalah metode *cluster non-hierarki* yang digunakan untuk mengelompokkan data ke dalam suatu *cluster* sehingga data yang memiliki karakteristik yang sama dapat dikelompokan ke dalam satu *cluster* yang sama (Zaki, Irwan, dan Sembe, 2022). Sedangkan metode yang termasuk dalam metode *hierarki* yaitu metode *ward,* metode *centroid linkage,* metode *complete linkage,* dan metode *single linkage*. Namun, metode *K-Means* memiliki kelemahan karena sensitif terhadap *outlier*. Kelemahan ini dapat diatasi dengan menggunakan metode *K-Medoids*. *K-Medoids* atau *Partitioning Around Method* (PAM), merupakan salah satu metode *cluster* non-hierarki yang merupakan varian dari *K-Means* yang *robust* terhadap *outlier*. Dalam meto de clustering, pemilihan ukuran kedekatan antar objek sangat

memengaruhi hasil pengelompokan. Salah satu ukuran kedekatan yang dapat digunakan adalah jarak *manhattan,* yang menghitung jumlah nilai *absolut* dari selisih antar atribut. Jarak ini dianggap lebih sesuai untuk data yang memiliki skala yang seragam serta lebih *robust* terhadap pencilan dibandingkan jarak *euclidean* (Siagian, Rikki, Barita, Simangunsong, 2024) . Oleh karena itu, dalam penelitian ini digunakan jarak *manhattan* sebagai dasar perhitungan kedekatan antar kabupaten/kota dalam proses pengelompokan menggunakan algoritma *K-Medoids*. Untuk mengukur keakuratan hasil pengelompokan yang terbentuk, maka perlu dilakukan evaluasi hasil pengelompokan. Pada penelitian ini, evaluasi dilakukan dengan metode *Davies Bouldin Index* (DBI).

Berdasarkan latar belakang tersebut, metode *K-Medoids* dapat digunakan untuk mengelompokkan kabupaten/kota di Kalimantan Barat berdasarkan tingkat kemiskinan. Selanjutnya, hasil pengelompokan dievaluasi menggunakan metode *Davies Bouldin Index* (DBI) untuk memperoleh pengelompokan yang optimal. Hasil pengelompokan ini diharapkan dapat menjadi dasar bagi pemerintah merumuskan dan menerapkan kebijakan yang efektif untuk mengatasi permasalahan kemiskinan yang terjadi di Kalimantan Barat.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka permasalahan yang akan dikaji dalam penelitian ini adalah:

1. Bagaimana pengelompokan kabupaten/kota di Kalimantan Barat berdasarkan tingkat kemiskinan menggunakan metode *K-Medoids?*
2. Bagaimana menentukan *cluster* optimal yang dibentuk oleh algoritma *K-Medoids* dengan jarak *manhattan* dan validasi *cluster* menggunakan metode *Davies Bouldin Index (DBI)* pada pengelompokan kabupaten/kota di Kalimantan Barat berdasarkan tingkat kemiskinan?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang dikemukakan, maka tujuan yang akan dicapai dari penelitian ini adalah:

1. Mengelompokkan kabupaten/kota di Kalimantan Barat menggunakan *K-Medoids* berdasarkan tingkat kemiskinan

2. Menentukan *cluster* optimal yang dibentuk oleh algoritma *K-Medoids* dengan jarak *manhattan* dan validasi *Davies Bouldin Index (DBI)* pada pengelompokan kabupaten/kota di Kalimantan Barat berdasarkan tingkat kemiskinan.

1.4 Batasan Masalah

Batasan masalah dalam penelitian ini yaitu:

1. Penelitian ini menggunakan data indikator yang memengaruhi kemiskinan yaitu persentase penduduk miskin, tingkat pengangguran terbuka , indeks kedalaman kemiskinaan, indeks keparahan kemiskinan, rata-rata lama sekolah, dan harapan lama sekolah.
2. Jumlah *cluster* (*K*) yang digunakan yaitu 2,3, dan 4.

1.5 Tinjauan Pustaka

Abdurrahman, Agus, dan Putra (2021) melakukan penelitian yang bertujuan untuk mengoptimalkan jumlah produksi komoditi perkebunan di Kalimantan Timur. Pada penelitian ini digunakan metode *clustering Partitioning Around Medoids* (PAM) dengan 3 *distance measure* yaitu *Euclidean Distance, Manhattan Distance,* dan *Chebyshev Distance.* Untuk mengukur kualitas hasil *cluster* digunakan metode *Silhouette Coefficient* (SC). Berdasarkan perhitungan hasil *clustering Partitioning Around Medoids* (PAM) menggunakan 3 *distance measure* diperoleh hasil *cluster* optimal yaitu 5 *cluster* dan *Manhattan* Distance menghasilkan nilai yang paling optimal.

Haumahu dan Matdoan (2022) melakukan pengelompokan tingkat kemiskinan pada kabupaten/kota di Kepulauan Maluku dan Papua dengan menggunakan algoritma *K-Medoids clustering*. Berdasarkan hasil penelitian dapat disimpulkan bahwa terdapat 4 *cluster* dalam pengelompokan tingkat kemiskinan di Kepulauan Maluku dan Papua. *Cluster* 0 terdiri dari 15 kabupaten/kota, *cluster* 1 terdiri dari 18 kabupaten/kota, *cluster* 2 terdiri dari 8 kabupaten/kota, dan *cluster* 3 terdiri dari 22 kabupaten/kota.

Penelitian yang dilakukan oleh Asmiatun, Wakhidah, dan Putri (2020) menggunakan metode *K-Medoids* untuk melakukan pengelompokan kondisi jalan di Kota Semarang. Hasil dari aplikasi *roadroid* didapatkan data sejumlah 638 data dengan atribut nama jalan, presentase kondisi jalan baik, sedang, rusak ringan, dan

rusak berat. Data dikelompokkan menjadi 4 *cluster* yang dapat menunjukkan tingkat frekuensi jalan kondisi baik, sedang, rusak ringan dan berat. Berdasarkan hasil pengujian kualitas *cluster* dengan *silhouette efficient* dapat disimpulkan bahwa jumlah *cluster* 2 merupakan *cluster* yang kualitasnya paling baik karena paling mendekati angka 1.

Pada penelitian lainnya, Ayuni, Kusnandar, dan Martha (2024) melakukan penelitian menggunakan algoritma *K-Medoids* dan *Clustering Large Applications* (CLARA) dengan optimasi *Silhouette Coefficient*. Penelitian ini bertujuan untuk menganalisis hasil pengelompokan kabupaten/kota berlandaskan indikator pembentuk Indeks Pembangunan Manusia (IPM). Berdasarkan hasil analisis, diperoleh pengelompokkan menggunakan algoritma *K-Medoids* menghasilkan 358 kabupaten/kota pada *cluster* 1 dan 156 kabupaten/kota pada *cluster* 2. Sementara itu, pengelompokan dengan CLARA menghasilkan 426 kabupaten/kota yang tergabung dalam *cluster* 1 dan 88 kabupaten/kota pada *cluster* 2. Algoritma CLARA merupakan algoritma yang optimal dengan nilai *silhouette coefficient* paling tinggi yakni sebesar 0,438.

1.6 Metodologi Penelitian

Metode yang digunakan dalam penelitian ini adalah metode *K-Medoids* dan evaluasi hasil *cluster* dengan menggunakan metode *Davies Bouldin Index* (DBI) untuk mengevaluasi hasil *cluster*. Studi kasus pada penelitian ini adalah dengan menganalisis 6 indikator kemiskinan di Kalimantan Barat pada tahun 2023. Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS). Dalam penelitian ini, terdapat 6 variabel yang digunakan yaitu persentase penduduk miskin, tingkat pengangguran terbuka, indeks kedalaman kemiskinan, indeks keparahan kemiskinan, rata-rata lama sekolah, dan harapan lama sekolah.

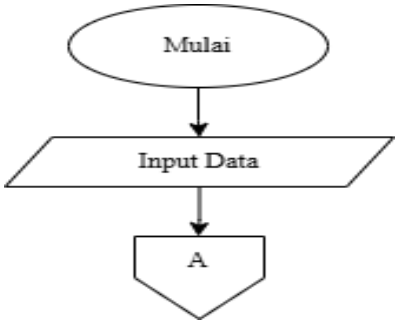
Tahapan analisis data dalam pengelompokan kabupaten di Kalimantan Barat menggunakan algoritma *K-Medoids* dengan jarak *manhattan* dan evaluasi hasil *cluster* dengan *Davies Bouldin Index* (DBI) dilakukan secara sistematis. Langkah pertama adalah melakukan analisis deskriptif terhadap data tiap indikator kemiskinan di Kalimantan Barat tahun 2023 untuk memahami karakteristik data.

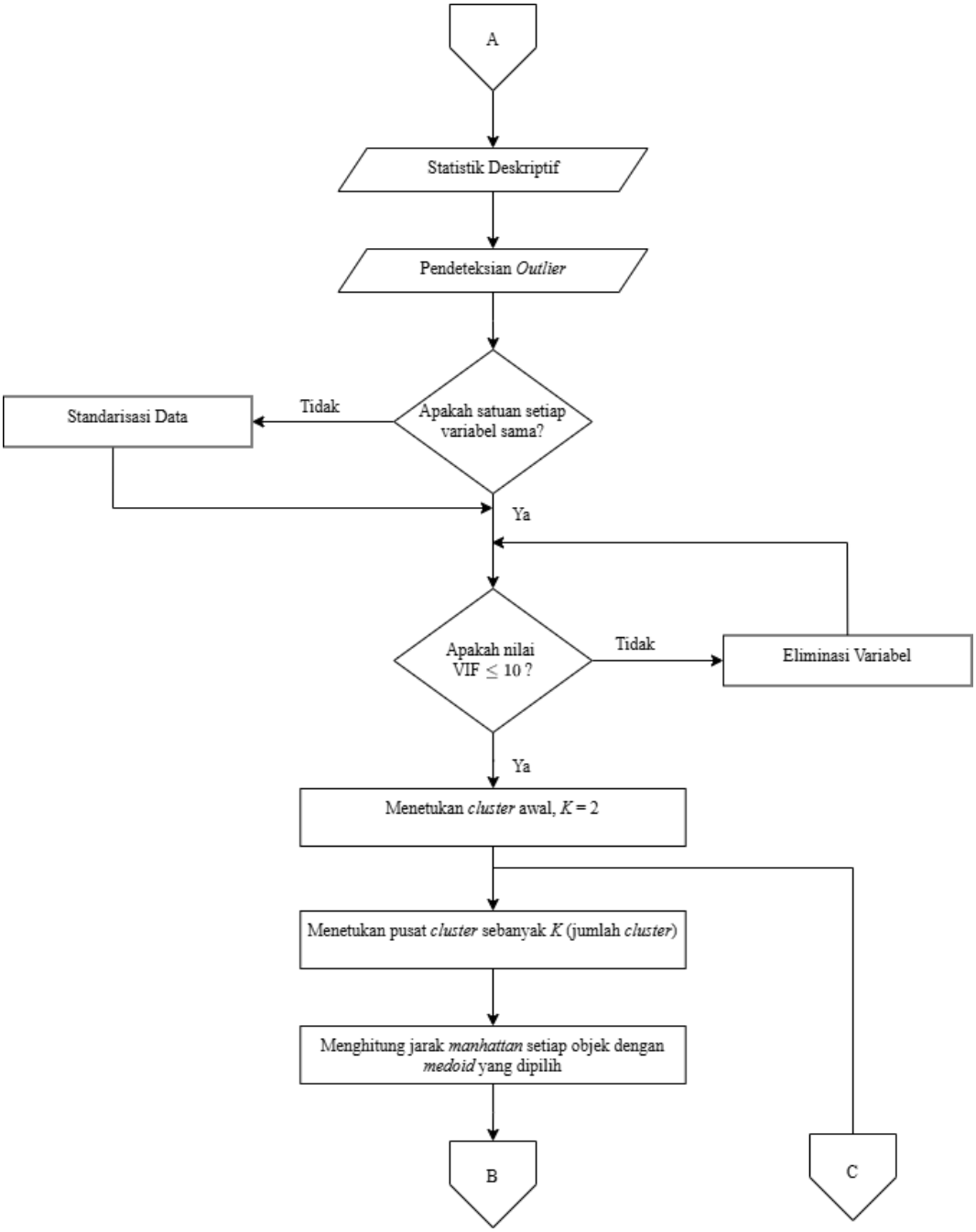
Kemudian dilakukan pendeteksian *outlier* menggunakan *boxplot*. Selanjutnya, dilakukan standarisasi data menggunakan *z-score* agar perbedaan satuan pada tiap indikator tidak memengaruhi hasil analisis.

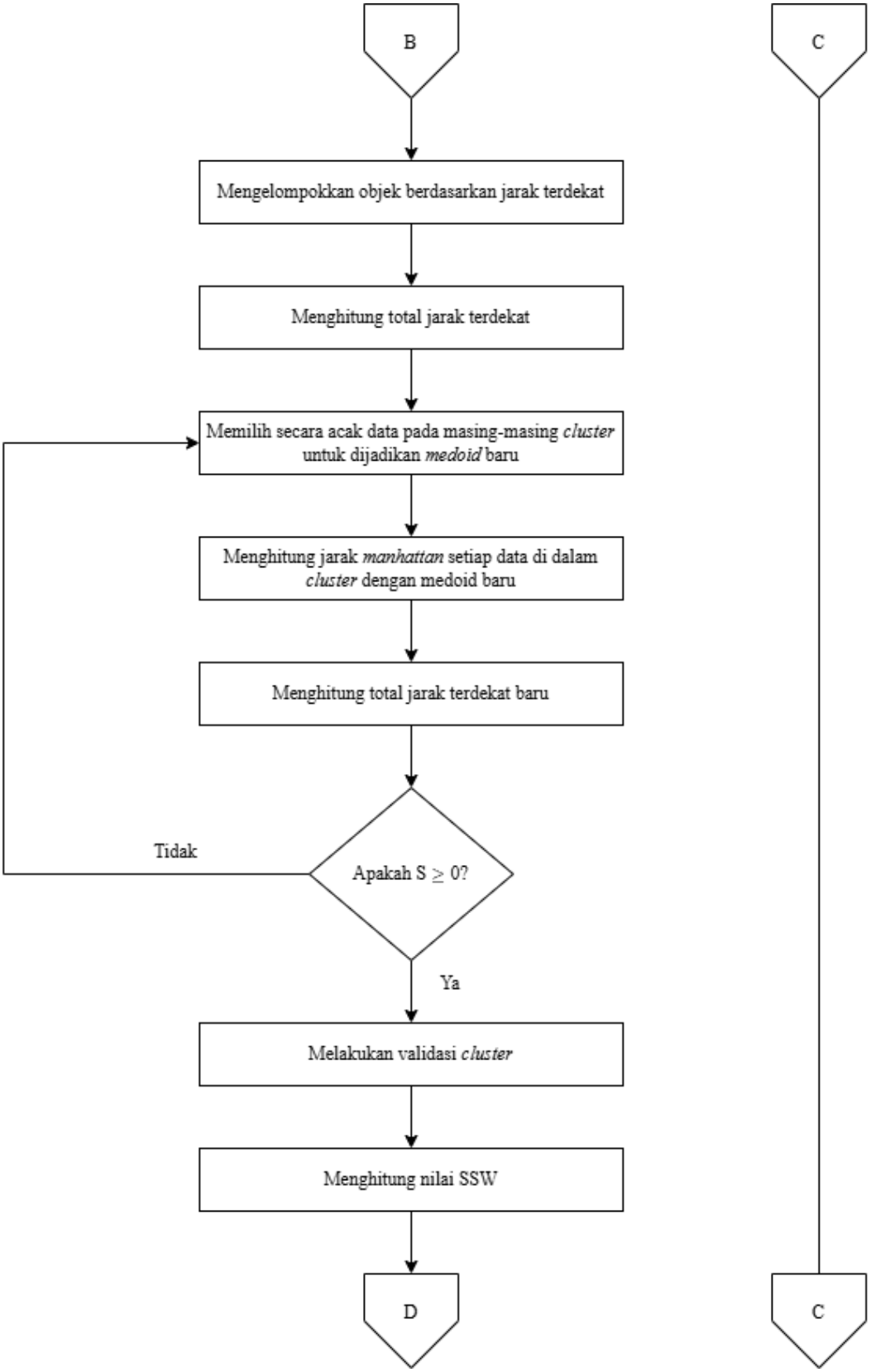
Setelah data distandarisasi, dilakukan uji multikolinearitas, jika nilai VIF > 10 artinya tidak terjadi multikolinearitas. Setelah itu, proses pengelompokan dilakukan menggunakan algoritma *K-Medoids* untuk jumlah *cluster K* = 2, 3, dan 4. Langkah-langkah dalam algoritma ini meliputi inisialisasi pusat *cluster* sebanyak *K*, perhitungan jarak setiap objek dengan medoid menggunakan *manhattan distance*, serta pengelompokan objek berdasarkan jarak terdekat. Setelah itu, dilakukan perhitungan total jarak terdekat dan pemilihan kandidat *medoid* baru secara acak dari masing-masing *cluster*. Kemudian, jarak setiap data dalam *cluster* dihitung terhadap kandidat *medoid* baru, diikuti dengan perhitungan total jarak terdekat dengan *medoid* baru. Selanjutnya, dilakukan perhitungan total simpangan (*S*) sebagai selisih antara total *distance* baru dan total *distance* lama. Jika nilai $S > 0$, maka iterasi dihentikan, namun jika tidak, langkah pemilihan medoid baru diulang hingga tidak terjadi perubahan medoid.

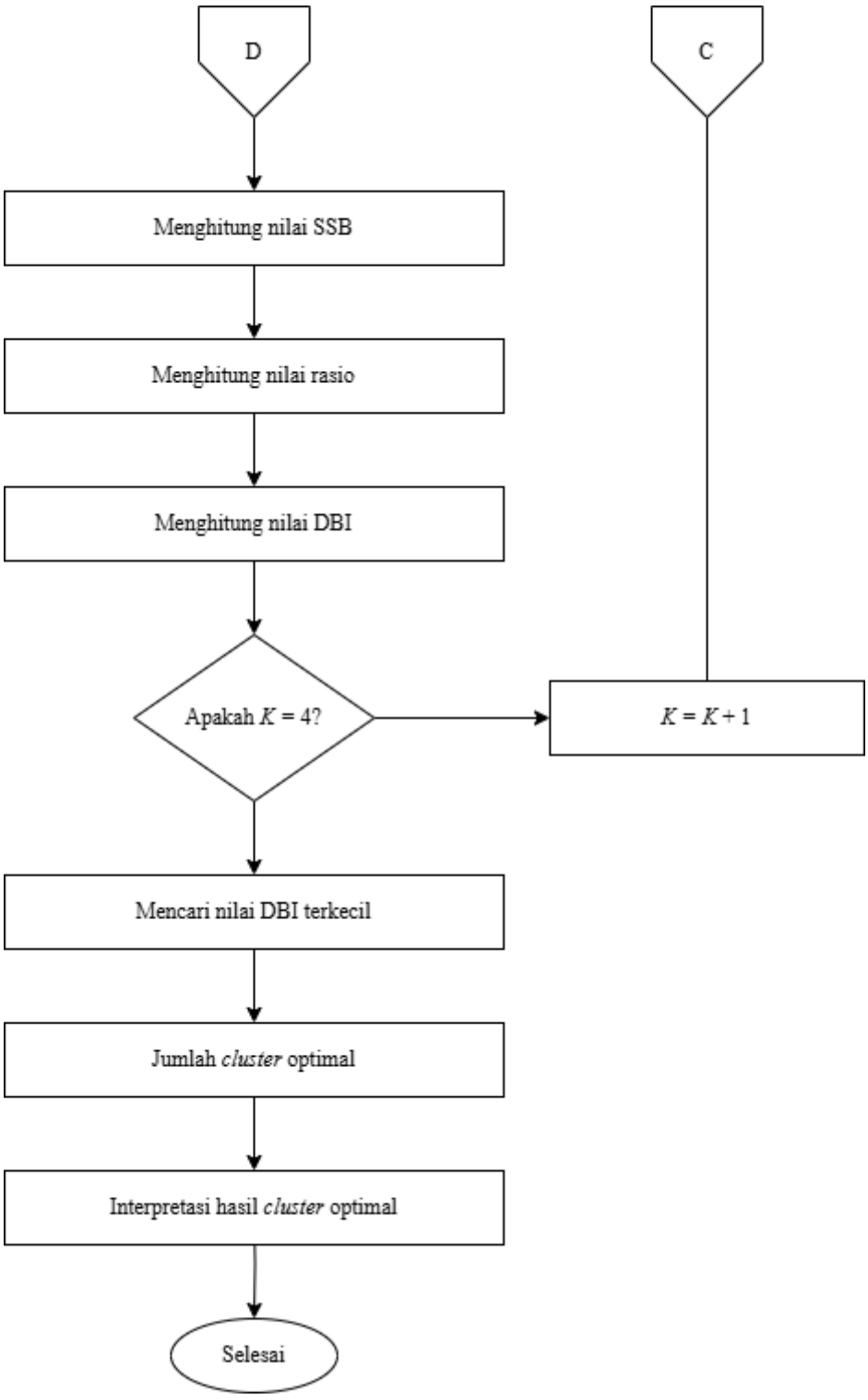
Evaluasi hasil *clustering* dilakukan dengan *Davies Bouldin Index* (DBI) untuk menentukan jumlah *cluster* yang optimal. Tahapan dalam evaluasi ini mencakup perhitungan nilai *Sum of Square Within* (SSW) dan *Sum of Square Between* (SSB), kemudian dihitung nilai rasio untuk memperoleh nilai DBI. Cluster optimal ditentukan berdasarkan nilai DBI terkecil. Setelah cluster optimal diperoleh, langkah selanjutnya adalah melakukan interpretasi terhadap hasil pengelompokan dan menarik kesimpulan berdasarkan hasil analisis

Flowchart penelitian dapat dilihat pada Gambar 1.1




A
Statistik Deskriptif
Pendeteksian *Outlier*
Apakah satuan setiap variabel sama?
Tidak
Standarisasi Data
Ya
Apakah nilai VIF ≤ 10 ?
Tidak
Eliminasi Variabel
Ya
Menetukan *cluster* awal, *K* = 2
Menetukan pusat *cluster* sebanyak *K* (jumlah *cluster*)
Menghitung jarak *manhattan* setiap objek dengan *medoid* yang dipilih
B
C


B
C
Mengelompokkan objek berdasarkan jarak terdekat
Menghitung total jarak terdekat
Memilih secara acak data pada masing-masing *cluster* untuk dijadikan *medoid* baru
Menghitung jarak *manhattan* setiap data di dalam *cluster* dengan medoid baru
Menghitung total jarak terdekat baru
Apakah S ≥ 0?
Tidak
Ya
Melakukan validasi *cluster*
Menghitung nilai SSW
D
C



Gambar 1. 1 *Flowchart* Penelitian