PEMETAAN WILAYAH POTENSI DEGRADASI HUTAN DAN DEFORESTASI DI KABUPATEN SANGGAU MENGGUNAKAN *K-MEANS CLUSTER* DENGAN EVALUASI *CALINSKI HARABASZ INDEX*

UMMI RAHIMAH
NIM H1091211040

SKRIPSI

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

PEMETAAN WILAYAH POTENSI DEGRADASI HUTAN DAN DEFORESTASI DI KABUPATEN SANGGAU MENGGUNAKAN *K-MEANS CLUSTER* DENGAN EVALUASI *CALINSKI HARABASZ INDEX*

UMMI RAHIMAH
NIM H1091211040

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	:Pemetaan Wilayah Potensi Degradasi Hutan dan Deforestasi di Kabupaten Sanggau menggunakan *K-Means Cluster* dengan Evaluasi *Calinski Harabasz Index*
Nama Mahasiswa	: Ummi Rahimah
NIM	: H1091211040
Jurusan/Program Studi	: Matematika/Statistika
Tanggal Lulus	: 3 Maret 2025
SK Pembimbing	: No. 2616/UN22.8/TD.06/2024/Tanggal 2 September 2024
SK Penguji	: No. 595/UN22.8/TD.06/2025/Tanggal 25 Februari 2025

Dosen Pembimbing

Pembimbing I	Pembimbing II
Shantika Martha, M.Si.	Nurfitri Imro'ah, M.Si.
NIP 198403082008122003	NIP 198907182019032021

Dosen Penguji

Ketua Penguji	Anggota Penguji
Dr. Evy Sulistianingsih, M.Sc.	Hendra Perdana, M.Sc.
NIP 198502172008122006	NIP 198810102019031020

Pimpinan Sidang (merangkap anggota penguji)	Sekretaris Sidang (merangkap anggota penguji)
Shantika Martha, M.Si.	Nurfitri Imro'ah, M.Si.
NIP 198403082008122003	NIP 198907182019032021

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	: Ummi Rahimah
NIM	: H1091211040
Program Studi/Jurusan	: 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 di kemudian hari dokumen ilmiah Tugas Akhir ini mengandung unsur 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, 19 Februari 2025



Ummi Rahimah

NIM H1091211040

Pemetaan Wilayah Potensi Degradasi Hutan dan Deforestasi di Kabupaten Sanggau Menggunakan *K-Means Cluster* dengan Evaluasi *Calinski Harabasz Index*

Abstrak

Indonesia memiliki kekayaan keanekaragaman hayati yang besar, dengan sebagian besar wilayah berupa kawasan hutan, termasuk hutan di Kalimantan yang dikenal sebagai paru-paru dunia. Namun, potensi besar ini menghadapi ancaman degradasi hutan dan deforestasi yang tinggi, terutama di Kalimantan Barat, dengan provinsi laju deforestasi tertinggi di Pulau Kalimantan. Kabupaten Sanggau, selain dikenal karena kekayaan hutannya, juga merupakan salah satu wilayah penyumbang Produk Domestik Regional Bruto (PDRB) sektor pertanian tertinggi di Kalimantan Barat, menjadikannya wilayah strategis dalam upaya pelestarian lingkungan dan pengembangan ekonomi berbasis sumber daya alam. Penelitian ini bertujuan untuk mengelompokkan wilayah-wilayah di Kabupaten Sanggau yang berpotensi mengalami degradasi hutan dan deforestasi menggunakan algoritma *K-Means Cluster* dan mengevaluasi *cluster* yang optimal menggunakan metode *Calinski Harabasz Index.* Langkah-langkah penelitian meliputi pengumpulan data, pengelompokan kecamatan menggunakan *K-Means Cluster,* serta evaluasi *cluster* menggunakan *Calinski Harabasz Index*. Hasil analisis menunjukkan bahwa pada tahun 2023 terdapat 5 *cluster* wilayah berpotensi degradasi hutan dan deforestasi, yaitu *cluster* pertama terdiri dari 1 kecamatan, *cluster* kedua terdiri dari 3 kecamatan, *cluster* ketiga terdiri dari 2 kecamatan, *cluster* keempat terdiri dari 5 kecamatan, dan *cluster* kelima terdiri dari 4 kecamatan. Kelima *cluster* ini merupakan *cluster* optimal dengan nilai *Calinski Harabasz Index* terbesar sebesar 3,87. Pengelompokkan ini juga menghasilkan peta tingkat urgensi potensi degradasi hutan dan deforestasi, dengan karakteristik wilayah berpotensi sangat tinggi, tinggi, sedang, rendah, dan sangat rendah. Peta ini dapat menjadi acuan bagi pemerintah Kabupaten Sanggau dalam upaya mitigasi bencana dan pengembangan wilayah sesuai karakteristik setiap kecamatan.

Kata Kunci: Ekologi, Pengelompokkan, Evaluasi *Cluster,* Pemantauan Lingkungan

Mapping Areas of Potential Forest Degradation and Deforestation in Sanggau Regency Using K-Means Cluster with Calinski Harabasz Evaluation Index

Abstract

Indonesia has a wealth of biodiversity, with most areas consisting of forest regions, including the forests in Kalimantan, which are known as the lungs of the world. However, this great potential faces significant threats of forest degradation and deforestation, especially in West Kalimantan, which has the highest deforestation rate on the Kalimantan Island. Sanggau Regency, in addition to being known for its rich forests, is also one of the regions contributing the highest Gross Regional Domestic Product (GRDP) in the agricultur sector in West Kalimantan, making it a strategic area in efforts tu preserve the environment and develop an economy based on natural resources. This research aims to cluster areas in Sanggau Regency that have the potential to experience forest degradation and deforestation using the K-Means Cluster algorithm and evaluate the optimal cluster using Calinski Harabasz Index method. The research steps include data collection, sub-district clustering using the K-Means, and cluster evaluation using the Calinski Harabasz Index. The analysis results show that in 2023, there are 5 clusters of areas at risk of forest degradation and deforestation, namely the first cluster consist of 1 sub-districts, the second cluster consists of 3 sub-districts, the third cluster consists of 2 sub-districts, the fourth cluster consists of 5 sub-districts, and the fifth cluster consists of 4 sub-districts. These five clusters are the optimal clusters with the highest Calinski Harabasz Index value of 3,87. This clustering also produced a map of urgency levels for potential forest degradation and deforestation, with characteristics of areas having very high, high, moderate, low, and very low potential. This map can be a reference for the Sanggau Regency government in disaster mitigation efforts and regional development according to the characteristics of each sub-district.

Keywords: *Ecology, Clustering, Cluster Evaluation, Environmental Monitoring*

PRAKATA

Puji syukur kehadirat Allah SWT, Tuhan yang Maha Esa atas rahmat dan karunia-Nya penulis dapat menyelesaikan penulisan Tugas Akhir yang berjudul "Pemetaan Wilayah Potensi Degradasi Hutan dan Deforestasi di Kabupaten Sanggau menggunakan *K-Means Cluster* dengan Evaluasi *Calinski Harabasz Index*". Adapun penelitian ini dibuat untuk memenuhi syarat penulis untuk memperoleh gelar Sarjana Statistika pada Program Studi Statistika di Fakultas Matematika dan Ilmu Pengetahuan Alam di Universitas Tanjungpura. Penulisan ini dapat dilakukan dan dapat diselesaikan tidak terlepas dari bimbingan, dukungan, saran, dan dukungan dari berbagai pihak. Oleh karena itu, penulis berterima kasih kepada:

1. Kedua orang tua yaitu Apak dan Umak yang telah memberikan semangat, motivasi, dan doa yang tak berhenti mengalir mendoakan kelancaran penulis dalam menempuh pendidikan. Terima kasih penulis ucapkan juga atas dukungan moral dan materi yang dilimpahkan dari kedua orang tua sehingga penulis dapat menempuh kuliah S1 hingga selesai.
2. Kepada Kakak, Ngangah, Teteh, dan Almahyra yang telah mendoakan dan mendukung penulis dalam menjalankan proses kuliah dengan baik pada saat awal penulis masih menjadi mahasiswa baru hingga lulus kuliah.
3. Bapak Ir. Dadan Kusnandar, Ph.D., dan Ibu Dr. Evy Sulistianingsih, M.Sc., selaku Dosen Pembimbing Akademik yang memberikan motivasi, saran, dan masukan selama masa perkuliahan.
4. Ibu Shantika Martha, M.Si., selaku Dosen Pembimbing Pertama yang telah meluangkan waktu untuk membimbing, memberi masukan, serta motivasi yang bermanfaat bagi penulis dalam menjalankan proses penulisan tugas akhir.
5. Ibu Nurfitri Imro'ah, M.Si., selaku Dosen Pembimbing Kedua yang telah meluangkan waktu untuk membimbing, memberi masukan, serta motivasi yang bermanfaat bagi penulis dalam menjalankan proses penulisan tugas akhir.

6. Ibu Dr. Evy Sulistianingsih, M.Sc., selaku Dosen Penguji Pertama yang telah memberikan saran dan masukan untuk melakukan perbaikan penulisan tugas akhir sehingga menjadi lebih baik.
7. 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.
8. Bapak Riza Adriat, M.Si., yang meluangkan waktu untuk membimbing, memberikan masukan, serta motivasi selama proses penelitian.
9. Adhelia Ananta, Fadhiilah Nurjannah, Khairunnisa Fiqri, Riska Nur Fitri, M. Yahya Ayyash, David Jordy Dhandio, Hesty Pratiwi, Noerul Hanin, Sy. Farini Nurhaliza, Endah Saraswi, dan Meftahudin yang selalu membersamai penulis serta menjadi *support system* terutama saat penulis menyusun tugas akhir.
10. Teman-teman "Expos" yang berjuang bersama di Program Studi Statistika dan selalu memberikan dorongan serta motivasi hingga bertahan sampai saat ini.

Penulis berharap semoga skripsi ini dapat bermanfaat dan dapat menjadi kemajuan ilmu pengetahuan. Skripsi ini jauh dari kata sempurna, sehingga penulis berharap dapat memperoleh kritik dan saran yang bersifat membangun dari semua pihak.

Pontianak, 19 Februari 2025

Ummi Rahimah

DAFTAR ISI

Abstrak ..... **iv**
Abstract ..... **v**
PRAKATA ..... **vi**
DAFTAR ISI ..... **viii**
DAFTAR GAMBAR ..... **x**
DAFTAR TABEL ..... **xi**
DAFTAR LAMPIRAN ..... **xii**
DAFTAR SIMBOL ..... **xiii**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang 1
1.2 Rumusan Masalah 3
1.3 Tujuan Penelitian 4
1.4 Batasan Masalah 4
1.5 Tinjauan Pustaka 4
1.6 Metodologi Penelitian 6
BAB II LANDASAN TEORI **12**
2.1 Analisis Multivariat 12
2.2 Analisis *Cluster* 12
2.2.1 Klasifikasi Analisis *Cluster* 13
2.2.2 Asumsi Analisis *Cluster* 13
2.3 Standardisasi Data 14
2.4 Jarak Antar Objek 15
BAB III EVALUASI *CLUSTER* **16**
3.1 *K-Means Cluster* 16
3.2 *Calinski Harabasz Index* 17
BAB IV HASIL DAN PEMBAHASAN **19**
4.1 Statistik Deskriptif 19
4.2 Asumsi Analisis *Cluster* 20
4.3 Standardisasi Data 22

4.4 Analisis *K-Means Cluster* 22
4.4.1 Penentuan *Centroid* Awal 22
4.4.2 Perhitungan Jarak terdekat 22
4.4.3 Pengelompokkan Objek Berdasarkan Jarak Terdekat..... 23
4.4.4 Penentuan *Centroid* Kembali 24
4.5 Evaluasi *Cluster* Menggunakan *Calinski Harabasz Index*..... 25
4.5.1 Perhitungan *Centroid* 25
4.5.2 Perhitungan *Grand Mean* 25
4.5.3 Perhitungan *Sum of Square Between-Cluster* 26
4.5.4 Perhitungan *Sum of Square Within-Cluster* 26
4.5.5 Perhitungan *Calinski Harabasz Index*..... 27
4.6 Penentuan *Cluster* Optimal 27
4.7 Karakteristik Setiap *Cluster* 28
BAB V PENUTUP..... 33
5.1 Kesimpulan..... 33
5.2 Saran..... 34
DAFTAR PUSTAKA..... 35
LAMPIRAN..... 39

DAFTAR GAMBAR

Gambar 1. 1 *Flowchar*t penelitian 11
Gambar 4. 1 Peta potensi degradasi hutan dan deforestasi..... 31

DAFTAR TABEL

Tabel 4. 1 Variabel penelitian 19
Tabel 4. 2 Statistik deskriptif 20
Tabel 4. 3 Nilai *Variance Inflation Factor*..... 21
Tabel 4. 4 Nilai *Variance Inflation Factor* setelah mengeluarkan $X5$ 21
Tabel 4. 5 Centroid awal 22
Tabel 4. 6 Centroid baru..... 25
Tabel 4. 7 Nilai *Calinski Harabasz Index* 27
Tabel 4. 8 Pengelompokkan kecamatan di Kabupaten Sanggau untuk 5 *cluster*. 28
Tabel 4. 9 Nilai rata-rata setiap variabel pada masing-masing *cluster*..... 28

DAFTAR LAMPIRAN

Lampiran I Data Penelitian 39
Lampiran II Standardisasi Data 40
Lampiran III Pengelompokkan Jarak Terdekat Iterasi 1 41
Lampiran IV Pengelompokkan Jarak Terdekat Iterasi 2..... 42

DAFTAR SIMBOL

VIF_n	:	Nilai *Variance Inflation Factor* variabel ke-n
${R^2}_n$	:	Koefisien determinasi variabel ke-n
$Z_{a;n}$	:	Standardisasi data untuk objek ke-a pada variabel ke-n, dengan $a = 1, 2, \dots, A$ dan $n = 1, 2, \dots, N$
$x_{a;n}$	:	Nilai dari objek ke-a pada variabel ke-n
$\bar{x}_n$	:	Rata-rata variabel ke-n
σ_n	:	Standar deviasi variabel ke-n
$d_{a;b}$	:	Jarak *euclidean* antara objek ke-a dan objek ke-b
N	:	Banyak variabel
$u_{k;n}$	:	*Centroid cluster* ke-k pada variabel ke-n
A_k	:	Banyak objek dalam *cluster* ke-k, dengan $k = 1, 2, \dots, K$
$x_{a;k;n}$	:	Nilai dari objek ke-a di *cluster* ke-k pada variabel ke-n
$\bar{x}$	:	Rata-rata keseluruhan objek
A	:	Banyak objek
K	:	Banyak *cluster*
SSB	:	Nilai *Sum of Square Between-cluster*
SSW	:	Nilai *Sum of Square Within-cluster*
$Z_{a;k;n}$	:	Nilai standardisasi dari objek ke-a di *cluster* ke-k pada variabel ke-n
CHI	:	Nilai *Calinski Harabasz Index*

BAB I
PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara yang kaya keanekaragaman hayati dengan 63,1% dari luas daratan berupa kawasan hutan (Kementerian Lingkungan Hidup dan Kehutanan, 2023). Luas hutan yang masif menyebabkan Indonesia memiliki peran penting sebagai penyangga keseimbangan ekosistem global. Hutan Kalimantan merupakan paru-paru dunia yang berfungsi sebagai sumber utama oksigen kehidupan di bumi (Ramdhoni, Fitriani, dan Afif, 2019). Kawasan hutan memiliki sumber daya yang sangat besar, permukaan hutan menyediakan berbagai jenis kayu komersial dan non-komersial, serta menjadi habitat bagi berbagai satwa yang dapat dimanfaatkan untuk keperluan lainnya seperti pariwisata dan penelitian.

Kalimantan Barat memiliki potensi sumber daya alam hayati yang sangat besar, termasuk potensi kayu sebesar 922,0 m^3 dengan nilai moneter mencapai 437,5 triliun pada tahun 2022 (Kementerian Lingkungan Hidup dan Kehutanan, 2023). Namun ironisnya, kekayaan ini terancam oleh kerusakan yang semakin meluas yang mana pada periode 2021-2022, Kalimantan Barat merupakan provinsi dengan laju deforestasi tercepat di pulau Kalimantan yang terjadi pada Kawasan hutan konservasi, hutan lindung, dan hutan produksi terbatas (Kementerian Lingkungan Hidup dan Kehutanan, 2023). Kabupaten Sanggau, sebagai salah satu wilayah agraris dengan kontribusi signifikan terhadap Produk Domestik Regional Bruto (PDRB) sektor pertanian di pemerintahan Provinsi Kalimantan Barat, turut terdampak (Badan Pusat Statistik Provinsi Kalimantan Barat, 2024). Berdasarkan proyeksi tahun 2023, penduduk Kabupaten Sanggau mencapai 503.937 jiwa, dengan mayoritas penduduk sebesar 64,64% bekerja di sektor pertanian (Badan Pusat Statistik Kabupaten Sanggau, 2024). Ketergantungan ekonomi yang tinggi pada sektor pertanian, di mana sebagian besar penduduk bekerja dengan metode pembukaan lahan yang masih mengandalkan pembakaran, telah menyebabkan terjadinya degradasi hutan dan deforestasi. Meskipun beberapa upaya pelestarian, telah dilakukan setelah pembukaan lahan, efektivitasnya belum optimal dalam

mencegah kerusakan lebih lanjut pada hutan dan lingkungan. Degradasi hutan adalah penurunan fungsi dan produktivitas hutan dan deforestasi merupakan hilangnya tutupan hutan secara permanen atau sementara (Fauzi dan Hazz, 2020).

Pertumbuhan populasi yang pesat dan konversi hutan untuk memenuhi kebutuhan pembangunan telah mempercepat laju degradasi hutan dan deforestasi (Hadiyan, Yuliah, dan Pambudi, 2017). Penebangan hutan secara besar-besaran dapat menghilangkan tutupan vegetasi yang berfungsi sebagai penahan erosi. Akibatnya, tanah menjadi mudah tererosi oleh air hujan, sehingga kesuburan tanah menurun dan lahan menjadi tandus (Fajeriana dan Risal, 2023). Degradasi hutan akibat erosi dan pencucian nutrisi membuat lahan tidak lagi produktif untuk pertanian atau kehutanan. Terjadinya penurunan luas dan kerusakan hutan telah menimbulkan dampak negatif antara lain terjadinya banjir lumpur, tanah longsor, erosi, siklus air terganggu, perubahan iklim, serta dampak lainnya (Jainuddin, 2023). Kondisi ini menciptakan lingkaran setan, di mana deforestasi menyebabkan degradasi dan degradasi memperparah deforestasi. Oleh sebab itu, diperlukan penanganan yang tepat agar tidak menyebabkan kerugian dan tidak ada masyarakat yang menderita. Langkah awal dalam mengatasi degradasi hutan dan deforestasi diperlukan pengelompokkan wilayah berdasarkan karakteristik penyebab degradasi hutan dan deforestasi di setiap wilayah.

Analisis untuk mengelompokkan suatu objek berdasarkan karakteristik yang sama dikenal dengan analisis *cluster.* Terdapat berbagai jenis analisis *cluster*, salah satunya adalah analisis *K-Means Cluster. K-Means Cluster* adalah algoritma yang digunakan untuk mengelompokkan data secara terpisah, di mana setiap kelompok memuat data yang berbeda-beda (Musharyadi, 2017). Algoritma ini menggunakan metode *clustering non-hierarchical* yang membagi data menjadi satu atau lebih *cluster* (Sukamto dan Angraini, 2018). Disebabkan oleh pembentukan *cluster* secara mandiri dalam analisis maka sulit memperoleh *cluster* yang optimal (Syahputri, Sutarman, dan Siregar, 2024). Terdapat beragam jenis metode yang dapat digunakan untuk mengevaluasi *cluster* seperti *Davies Bouldin Index, Silhouette, Krzanowski and Lai,* dan *Calinski Harabasz Index*. Pembentukan *cluster*

yang optimal pada penelitian ini dievaluasi menggunakan metode *Calinski Harabasz Index.*

Calinski Harabasz Index (CHI) merupakan metode yang menghitung rasio antar *cluster* dengan variasi dalam *cluster* (Lima dan Cruz, 2020). Suatu *cluster* dikatakan optimal ketika nilai CHI lebih tinggi dibandingkan *cluster* lainnya (Wang dan Xu, 2019). Menurut Milligan dan Cooper (1985) CHI memiliki beberapa keunggulan dibandingkan dengan metode lain, yaitu konsistensi yang tinggi dalam mengindentifikasi banyak *cluster* optimal pada berbagai dataset dan konfigurasi *cluster,* sehingga dapat memberikan hasil yang stabil bahkan pada tingkat *cluster* yang rendah. Selain itu, CHI tidak bergantung pada parameter kontrol data khusus, yang menjadikannya lebih fleksibel untuk diterapkan pada berbagai jenis data. Tingkat kesalahan CHI juga lebih rendah dibandingkan metode lainnya, karena mampu meminimalkan kesalahan baik untuk *cluster* yang "terlalu sedikit" maupun "terlalu banyak". Penerapan metode CHI pada analisis *K-Means Cluster* dapat membentuk kelompok optimal dari wilayah berpotensi terjadinya degradasi hutan dan deforestasi di Kabupaten Sanggau yang dapat dijadikan acuan pemerintahan kabupaten dalam pembangunan wilayah dan mitigasi bencana disesuaikan dengan karakteristik setiap kecamatan pada setiap kelompoknya.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang diuraikan, maka rumusan masalah yang dibahas sebagai berikut.

1. Bagaimana mengelompokkan wilayah potensi degradasi hutan dan deforestasi di Kabupaten Sanggau berdasarkan faktor yang mempengaruhinya dengan menggunakan analisis *K-Means Cluster*?
2. Bagaimana menentukan banyak *cluster* optimal berdasarkan metode evaluasi *Calinski Harabasz Index*?
3. Bagaimana memetakan tingkat urgensi wilayah potensi degradasi hutan dan deforestasi di Kabupaten Sanggau?

1.3 Tujuan Penelitian

Mengacu pada rumusan masalah, maka tujuan penelitian sebagai berikut.

1. Mengelompokan wilayah potensi degradasi hutan dan deforestasi di Kabupaten Sanggau berdasarkan faktor yang mempengaruhinya dengan menggunakan analisis *K-Means Cluster.*
2. Menentukan banyak *cluster* optimal berdasarkan metode evaluasi *Calinski Harabasz Index.*
3. Memetakan tingkat urgensi wilayah potensi degradasi hutan dan deforestasi di Kabupaten Sanggau.

1.4 Batasan Masalah

Batasan masalah dibuat untuk memastikan fokus penelitian ini tetap terjaga dan menghindari pembahasan yang terlalu luas. Adapun batasan masalah pada penelitian ini sebagai berikut.

1. Penelitian ini menggunakan data tahun 2023 dari 15 kecamatan di Kabupaten Sanggau, Kalimantan Barat. Data yang digunakan meliputi data curah hujan yang diperoleh dari *Climate Hazards Group Infrared Precipitation* (CHIRPS), data kelerengan yang diperoleh dari *Digital Elevation Model* (DEM) Nasional, data kepadatan penduduk/km^2, indikator tutupan lahan berupa sawah, tegal/kebun, ladang/huma, perkebunan, hutan rakyat, hutan negara, dan bukan pertanian yang diperoleh dari publikasi Badan Pusat Statistik Kabupaten Sanggau.
2. Banyaknya *cluster* (K) adalah 2, 3, 4, dan 5.
3. Perhitungan jarak yang digunakan yaitu jarak *euclidean.*

1.5 Tinjauan Pustaka

Silamon, Idris, dan Anggara (2015) melakukan penelitian yang bertujuan untuk mengetahui keselarasan lahan tanaman industri kehutanan dan kelas degradasi lahan di Desa Marga Karya Resort Semamung. Data yang digunakan pada penelitian berupa curah hujan, suhu udara, kelerengan, persentase C organik, tekstur tanah, kedalaman efektif tanah, dan persentase batuan. Dari penelitian tersebut diketahui bahwa tanaman *Akasia Mangium* paling banyak ditanam pada

lahan dengan tingkat kesesuaian yang sangat tinggi, mencapai 88,06% dari total luas area. Sedangkan, untuk degradasi lahan diperoleh 5 kelas, yaitu terdegradasi sangat ringan seluas 0,55 ha, terdegradasi ringan seluas 9,08 ha, terdegradasi sedang seluas 24,2 ha, terdegradasi berat seluas 31,8 ha, dan terdegradasi sangat berat seluas 31,8 ha.

Djaenudin, Oktaviani, Hartoyo, dan Dwiprabowo (2018) menganalisis faktor-faktor yang menyebabkan terjadinya penurunan laju deforestasi di Indonesia. Penelitian ini menggunakan data laju deforestasi, kepadatan penduduk jiwa/km^2, pangsa tutupan hutan, dan pendapatan perkapita. Hasil penelitian menggunakan *Ordered Logistic Regression* (OLR) menunjukkan bahwa provinsi-provinsi di Indonesia terbagi menjadi 3 kelompok, yaitu rendah, sedang, dan tinggi. Kepadatan penduduk dan pangsa tutupan hutan menjadi faktor yang dapat membantu mengurangi laju deforestasi. Sedangkan, pendapatan perkapita terindikasi menghambat penurunan laju deforestasi.

Pada penelitian lainnya, Solehuwey, Mardiatmoko, dan Boreel (2024) mempelajari faktor penyebab terjadinya degradasi dan deforestasi di Kecamatan Seram Utara Timur Kobi untuk mengetahui laju degradasi dan deforestasi yang terjadi. Pada penelitian ini melibatkan 7 jenis tutupan lahan yaitu hutan lahan kering sekunder, hutan lahan kering primer, lahan terbuka, semak belukar, sawah, perkebunan, dan pemukiman pada tahun 2013 hingga 2020. Sehingga, diperoleh laju degradasi hutan di Kecamatan Seram Utara Timur Kobi dalam rentang tersebut, yaitu hutan lahan kering sekunder 2,24% dan hutan lahan primer 0,14%. Sedangkan, laju deforestasi pada tutupan lahan hutan yang mengalami perubahan menjadi *non*-hutan pada tahun 2020 sebesar -0,25%.

Penelitian yang dilakukan oleh Santoso, Ratnawati, dan Riyanti (2022) yang menggunakan analisis *K-Means Cluster* untuk mengelompokkan tingkat deforestasi di Indonesia. Penelitian ini melibatkan 34 provinsi di Indonesia dengan penggabungan Kalimantan Timur dan Kalimantan Utara sehingga objek penelitian terdapat 33 data deforestasi di Indonesia. Data yang digunakan merupakan data periode 2013/2014 hingga 2018/2019. Hasil penelitian menunjukkan bahwa terbentuk 5 *cluster* tingkat deforestasi yaitu terdapat dua provinsi pada *cluster* 1,

satu provinsi pada *cluster* 2, 27 provinsi pada *cluster* 3, 1 provinsi pada *cluster* 4, dan 2 provinsi pada *cluster* 5.

Febryanti, Adiningsi, Saputra (2023) menerapkan *Fuzzy K-Means* dengan metode evaluasi *cluster* menggunakan *Silhouette Coeficient, Davies Bouldin Index,* dan *Calinski Harabasz Index.* Data yang digunakan meliputi hutan lindung, kawasan suaka alam, hutan produksi terbatas, hutan produksi tetap, dan hutan produksi yang dapat dikonversi di Sulawesi Tenggara. Dari penelitian tersebut diperoleh kesimpulan bahwa terdapat 2 *cluster* yaitu terdapat 13 provinsi pada *cluster* 1 dengan tingkat deforestasi tinggi dan terdapat 10 provinsi pada *cluster* 2 dengan tingkat deforestasi rendah. Hasil evaluasi menunjukkan bahwa *cluster* yang terbentuk memiliki pemisahan yang baik antar *cluster* dan dalam *cluster* ditunjukkan dengan nilai *Silhouette Score* sebesar 0,469, *Davies-Bouldin Index* sebesar 1,38, dan *Calinski Harabasz* sebesar 25,714.

1.6 Metodologi Penelitian

Penelitian ini menggunakan data sekunder yang melibatkan 10 variabel penyebab terjadinya degradasi hutan dan deforestasi. Variabel pertama (X_1) yang digunakan ialah data kelerengan wilayah setiap kecamatan di Kabupaten Sanggau tahun 2023. Data kelerengan diperoleh dari *Digital Elevation Model* (DEM) yang bersumber dari DEM Nasional. Data yang tersedia berupa mosaik dengan format raster, sehingga perlu untuk disatukan antara mosaik satu dengan mosaik yang lainnya. Kumpulan data mosaik raster perlu digabungkan untuk membentuk satu dataset yang utuh. Data kelerengan yang merupakan turunan dari data DEM dapat diperoleh dengan menerapkan algoritma perhitungan *slope* dengan satuan persen yang diproses menggunakan *software* QGis. Mengingat lokasi penelitian adalah Kabupaten Sanggau, maka data raster dipotong untuk wilayah tersebut menggunakan perintah *clip* dan di konversi ke dalam format poligon. Setelah data *slope* disesuaikan dengan batas wilayah penelitian, data numerik siap digunakan.

Variabel kedua (X_2) yang digunakan ialah data jumlah curah hujan yang diperoleh dari *Climate Hazards Group Infrared Precipitation* (CHIRPS). Data jumlah curah hujan yang digunakan pada awalnya berbentuk raster dan berjumlah 12 data yang merupakan data curah hujan bulanan di Indonesia pada tahun 2023.

Curah hujan bulanan tersebut diakumulasikan menjadi satu sehingga menjadi data curah hujan di Indonesia pada tahun 2023. Dari data tersebut, kemudian dicari jumlah curah hujan dengan menggunakan alat *cell statistics tool* kemudian memilih operasi *sum* yang diproses menggunakan *software* QGis. Setelah itu, dilakukan pemotongan data raster jumlah curah hujan tahunan dengan menggunakan batas wilayah Kabupaten Sanggau dan dikonversi ke dalam format poligon. Dengan demikian, diperoleh data numerik jumlah curah hujan tahunan yang hanya mencakup Kabupaten Sanggau.

Variabel lainnya yang digunakan pada penelitian ini diperoleh dari publikasi Badan Pusat Statistik Kabupaten Sanggau yaitu Kabupaten Sanggau Dalam Angka Tahun 2024. Adapun data yang digunakan adalah data untuk 15 kecamatan di Kabupaten Sanggau berupa data kepadatan penduduk/km^2 (X_3) dan data indikator tutupan lahan seperti sawah (X_4), tegal/kebun (X_5), ladang/huma (X_6), perkebunan (X_7), hutan rakyat (X_8), hutan negara (X_9), dan bukan pertanian (X_{10}).

Tahap berikutnya adalah memasukkan seluruh data yang diperlukan ke dalam basis analisis. Sebelum melangkah lebih jauh, dilakukan pemeriksaan awal terhadap data melalui analisis statistik deskriptif, mencakup nilai minimum, maksimum, dan rata-rata, untuk memastikan bahwa data yang diinput sesuai dengan yang diperlukan.

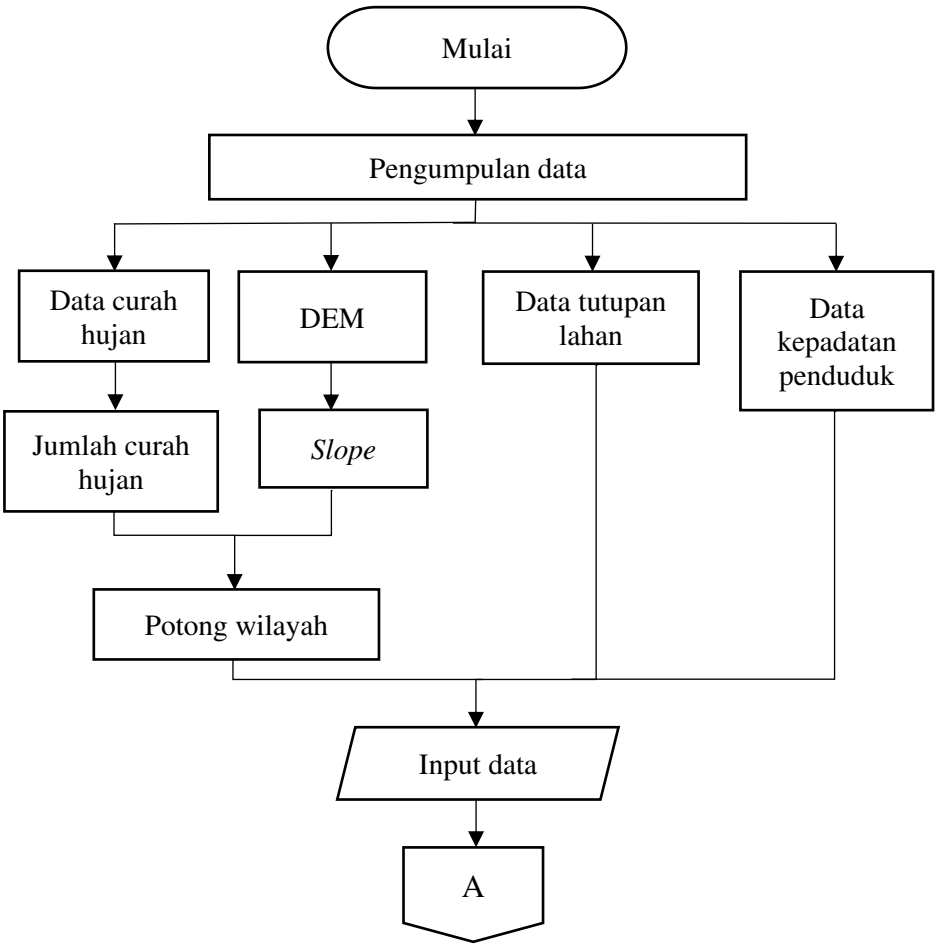
Pada *K-Means Cluster* penting untuk mengetahui apakah variabel yang digunakan saling berkorelasi. Oleh karena itu, dilakukan pengujian multikolinearitas untuk mengindentifikasi adanya korelasi antar variabel, menggunakan metode *Variance Inflation Factor* (VIF). Jika nilai VIF suatu variabel lebih dari 10, variabel tersebut harus dikeluarkan, dan pengecekan VIF diulang hingga semua variabel memiliki nilai VIF di bawah 10, menandakan bahwa data siap untuk analisis lebih lanjut.

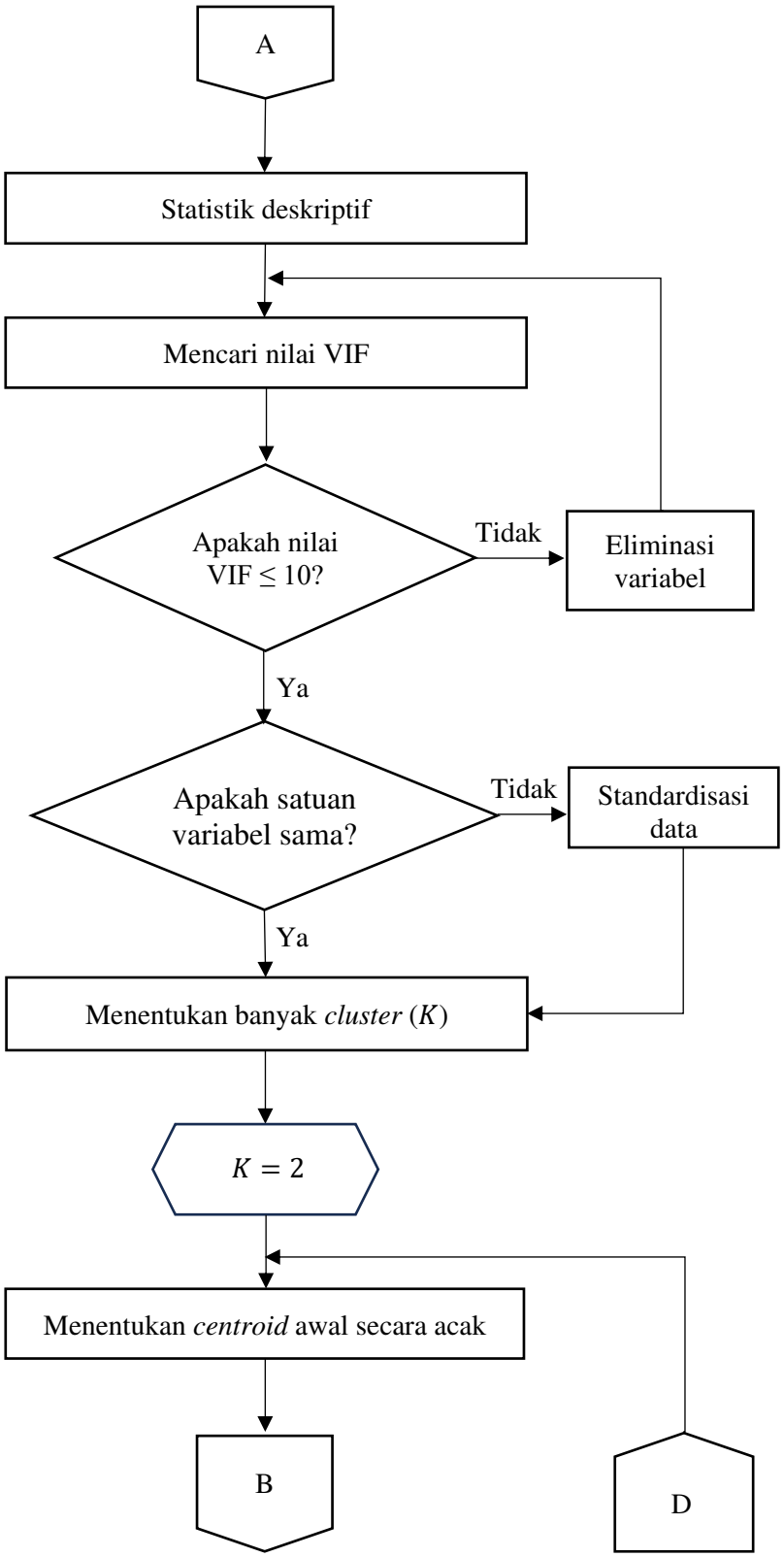
Variabel yang digunakan pada penelitian ini memiliki satuan data yang berbeda, maka dilakukan standardisasi data agar memiliki satuan yang sebanding, menggunakan metode *Z-score*. Data yang telah di standardisasi kemudian siap untuk analisis *K-Means Cluster.* Langkah pertama dalam metode *K-Means Cluster* adalah menentukan banyak *cluster* (K) yang akan digunakan. Setelah K ditetapkan,

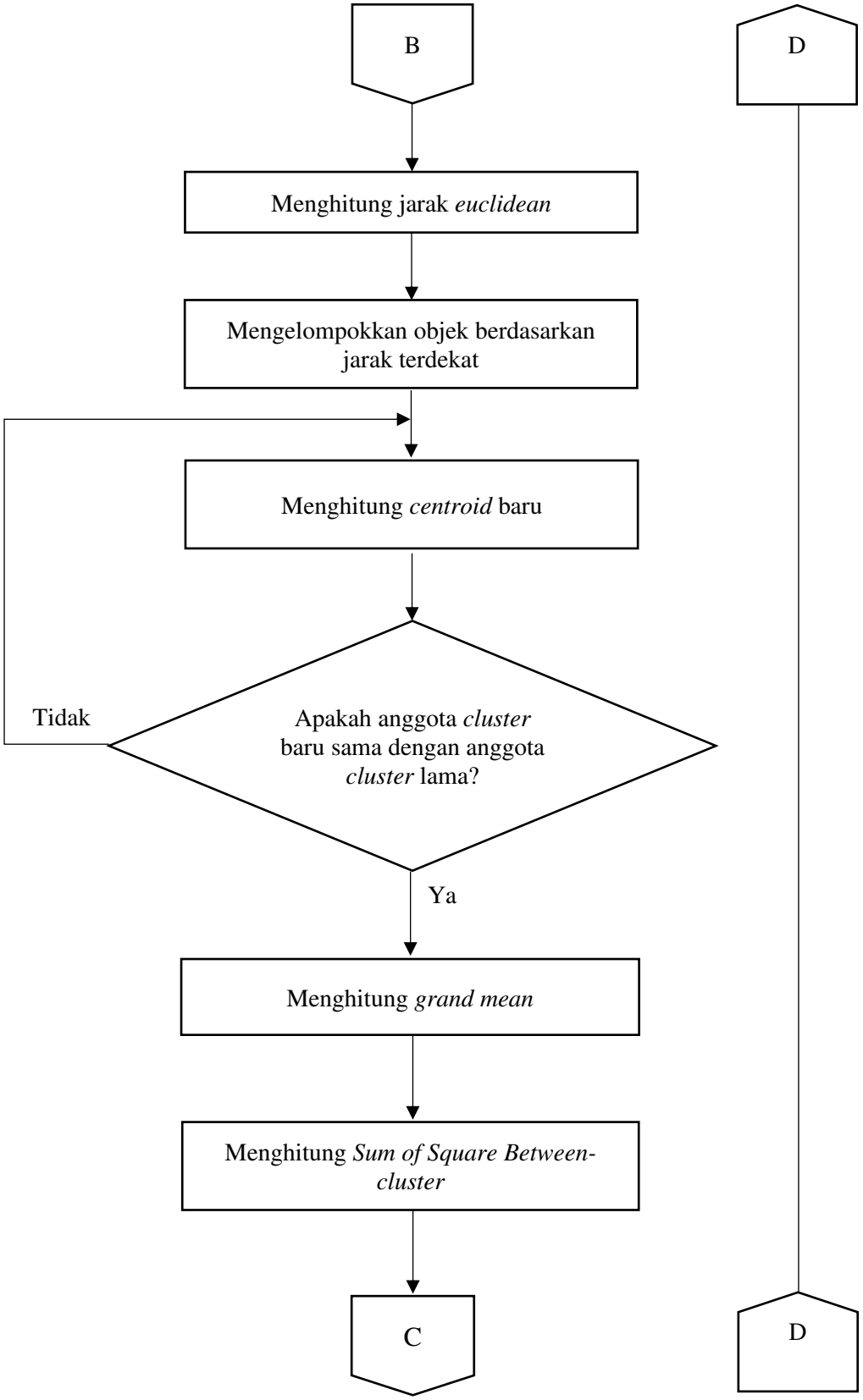
centroid awal sebanyak K ditentukan secara acak dan hitung jarak *euclidean* antara setiap objek dengan *centroid*-nya. Kemudian, objek-objek dikelompokkan ke dalam *cluster* berdasarkan *centroid* terdekat dan hitung kembali *centroid* baru. Langkah ini diulangi hingga *centroid* baru sama dengan *centroid* lama.

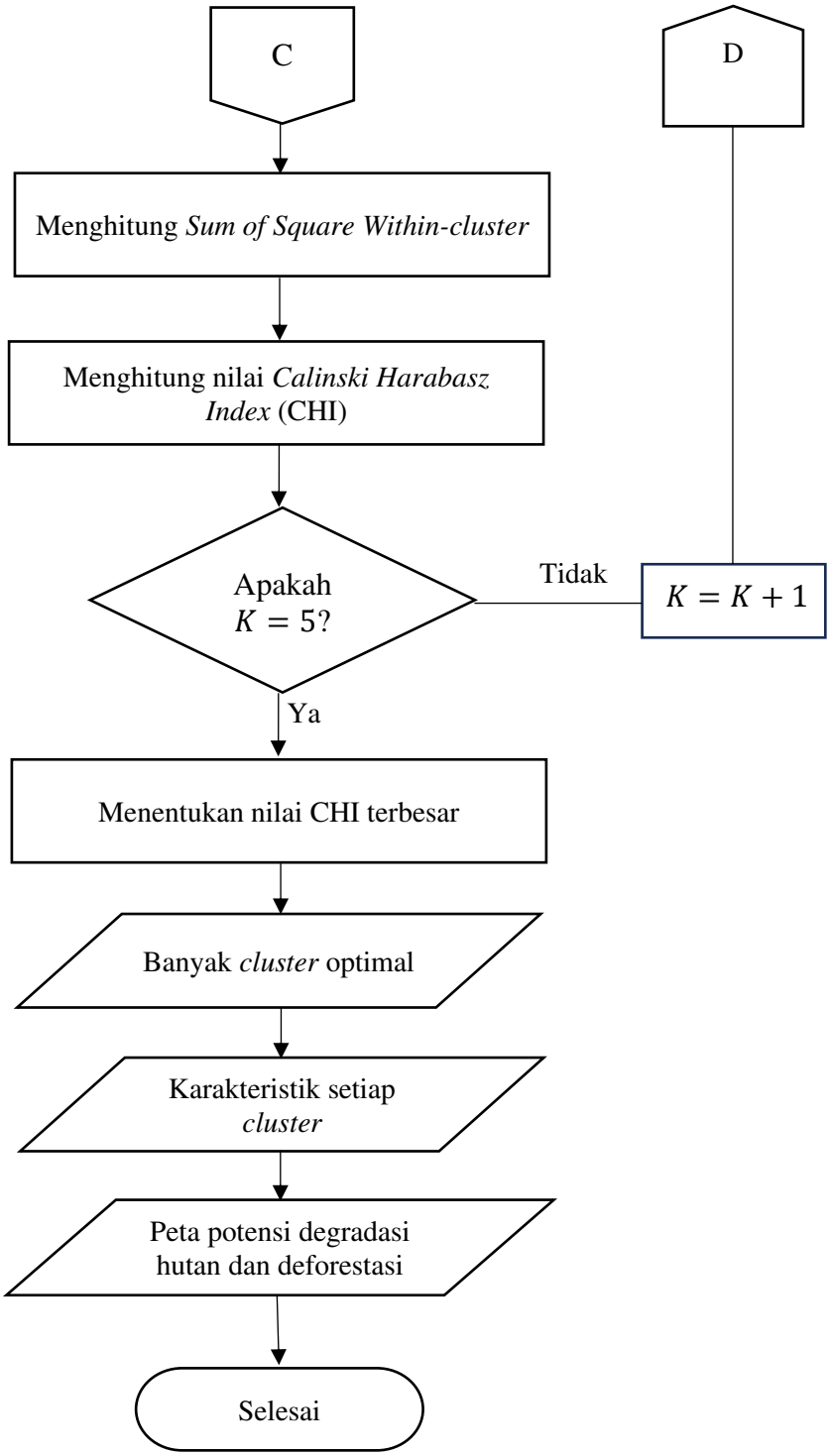
Konvergensi tercapai ketika *centroid* baru dan *centroid* lama telah sama. Proses ini dilakukan kembali hingga $K = 5$. Langkah berikutnya, dilakukan evaluasi *cluster* yang terbentuk menggunakan *Calinski Harabasz Index* untuk menilai kualitas pengelompokkan. Langkah evaluasi meliputi perhitungan rata-rata keseluruhan data, nilai *Sum of Square Between-cluster,* nilai *Sum of Square Within-cluster,* dan nilai *Calinski Harabasz Index*. Berdasarkan nilai indeks tersebut, jumlah *cluster* terbaik dapat ditentukan, dan dilanjutkan dengan analisis karakteristik setiap *cluster.*

Tahap terakhir adalah pemetaan wilayah yang berpotensi mengalami degradasi hutan dan deforestasi berdasarkan karakteristik *cluster* yang terbentuk. Hasil akhir penelitian ini diharapkan dapar memberikan pemetaan yang akurat mengenai potensi degradasi hutan dan deforestasi di Kabupaten Sanggau.




A
Statistik deskriptif
Mencari nilai VIF
Apakah nilai VIF ≤ 10?
Tidak
Eliminasi variabel
Ya
Apakah satuan variabel sama?
Tidak
Standardisasi data
Ya
Menentukan banyak *cluster* (K)
$K = 2$
Menentukan *centroid* awal secara acak
B
D


B
D
Menghitung jarak *euclidean*
Mengelompokkan objek berdasarkan jarak terdekat
Menghitung *centroid* baru
Apakah anggota *cluster* baru sama dengan anggota *cluster* lama?
Tidak
Ya
Menghitung *grand mean*
Menghitung *Sum of Square Between-cluster*
C
D



Gambar 1. 1 *Flowchart* penelitian