ANALISIS *K-MEANS CLUSTERING* DENGAN *BOOTSTRAP* PADA PENGELOMPOKAN DESA DI KABUPATEN MEMPAWAH BERDASARKAN INDEKS DESA MEMBANGUN

FADHELA TRIFAIZA
NIM H1091211032

SKRIPSI

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

ANALISIS *K-MEANS CLUSTERING* DENGAN *BOOTSTRAP* PADA PENGELOMPOKAN DESA DI KABUPATEN MEMPAWAH BERDASARKAN INDEKS DESA MEMBANGUN

FADHELA TRIFAIZA
NIM H1091211032

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

Tugas Akhir	: **Analisis *K-Means Clustering* dengan *Bootstrap* pada Pengelompokan Desa di Kabupaten Mempawah Berdasarkan Indeks Desa Membangun**
Nama Mahasiswa	: Fadhela Trifaiza
NIM	: H1091211032
Jurusan/Program Studi	: Matematika/Statistika
Tanggal Lulus	: 16 Juni 2025
SK Pembimbing	: No.2610/UN22.8/TD.06/2024/Tanggal 2 September 2024
SK Penguji	: No.1706/UN22.8/TD.06/2025/Tanggal 5 Juni 2025

Dosen Pembimbing

Pembimbing I	Pembimbing II
Hendra Perdana, M.Sc	Neva Satyahadewi, M.Sc., CRA., CRP., CRMP
NIP 198810102019031020	NIP 198212042005012001

Dosen Penguji

Ketua Penguji	Anggota Penguji
Dr. Evy Sulistianingsih, M.Sc	Nurfitri Imro'ah, M.Si
NIP 198502172008122006	NIP 198907182019032021

Pimpinan Sidang (merangkap anggota penguji)	Sekretaris Sidang (merangkap anggota penguji)
Hendra Perdana, M.Sc	Neva Satyahadewi, M.Sc., CRA., CRP., CRMP
NIP 198810102019031020	NIP 198212042005012001

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	: Fadhela Trifaiza
NIM	: H1091211032
Program Studi/ Jurusan	: Statistika/Matematika
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, 16 Juni 2025



Fadhela Trifaiza

NIM H1091211032

Analisis *K-Means Clustering* dengan *Bootstrap* pada Pengelompokan Desa di Kabupaten Mempawah Berdasarkan Indeks Desa Membangun

Abstrak

Pembangunan desa merupakan suatu upaya yang dilakukan dalam meningkatkan kesejahteraan desa yang diukur melalui status desa berdasarkan Indeks Desa Membangun atau IDM. Pembangunan desa melalui IDM dilakukan pemerintah sebagai upaya menciptakan kesejahteraan desa, mengurangi kesenjangan yang terjadi antar desa, dan dapat memberikan perhatian khusus bagi desa dengan status yang rendah. Sehingga diperlukan pengelompokan desa berdasarkan IDM menggunakan analisis *cluster.* Analisis *cluster* merupakan teknik statistik yang mengelompokkan objek berdasarkan kesamaan karakteristik tiap objek. Tujuan penelitian ini adalah menganalisis pengelompokan desa di Kabupaten Mempawah berdasarkan IDM menggunakan *K-Means Clustering* dengan *Bootstrap.* Data yang digunakan pada penelitian ini merupakan data indikator Indeks Ketahanan Ekonomi (IKE) berdasarkan IDM di Kabupaten Mempawah tahun 2022 yang terdiri dari 60 desa dan 12 variabel, yaitu Keragaman Produksi, Pertokoan, Pasar, Toko/Warung Kelontong, Kedai dan Penginapan, Pos dan Logistik, Bank dan BPR, Kredit, Lembaga Ekonomi, Moda Transportasi Umum, Keterbukaan Wilayah, dan Kualitas Jalan. Berdasarkan analisis yang dilakukan diperoleh nilai akurasi hasil perbandingan metode *K-Means Bootstrap* dengan status desa di IDM yaitu sebesar 68%. *Cluster* 1 dengan status desa tertinggal memiliki anggota sebanyak 4 desa. *Cluster* 2 dengan status paling tinggi yaitu mandiri terdiri dari 32 anggota *cluster*. *Cluster* 3 dengan status desa maju terdiri dari 8 anggota *cluster*. *Cluster* 4 dengan status desa berkembang memiliki anggota sebanyak 13 desa. *Cluster* 5 dengan status paling rendah yaitu sangat tertinggal terdiri dari 3 anggota *cluster,* yaitu desa Ansiap, Pentek, dan Suak Barangan.

Kata kunci: Desa, *Cluster*, Jarak *Gower*

K-Means Clustering Analysis with Bootstrap on Village Groupings in Mempawah District Based on The Village Development Index

Abstract

Village development is an efforts made to improve village welfare as measured through village status based on the Village Development Index or IDM. Village development through IDM is carried out by the government as an effort to create village prosperity and reduce disparities that occur between villages, and can provide spesial attention to villages with low status. Therefore it is necessary to group villages based on IDM using cluster analysis. Cluster analysis is a statistical technique that groups objects based on the similar characteristics of each object. The aim of this research is to analyze village groupings in Mempawah Regency based on IDM using K-Means Chistering with Bootstrap. The data used in this research is economic resilience index indicator data based on IDM in Mempawah regency in 2022 which consists of 60 villages and 12 variables, that is Production Diversity, Shops, Markets, Shops/Grocery Stalls. Shops and Lodging, Post and Logistics, Banks and BPR, Credit, Economic Institutions, Public Transportation Modes, Regional Openness, and Road Quality. Based on the analysis carried out, the accuracy value obtained from the comparison K-Means Bootstrap with the village status in IDM was 68%. Cluster 1 with underdeveloped village status has 4 members. Cluster 2 with the highest status, namely independent, consists of 32 cluster members. Cluster 3 with developed village status consists of 8 cluster members. Cluster 4 with developing village status has 13 members. Cluster 5 with the lowest status, namely very underdeveloped, consists of 3 cluster members, namely the villages of Ansiap, Pentek, and Suak Barangan.

Keywords: *Village, Cluster, Gower Distance*

PRAKATA

Puji syukur kehadirat Allah SWT atas segala limpahan rahmat serta karunia-Nya, sehingga penulis dapat menyelesaikan penulisan skripsi yang berjudul **"Analisis *K-Means Clustering* dengan *Bootstrap* pada Pengelompokan Desa di Kabupaten Mempawah Berdasarkan Indeks Desa Membangun**". Skripsi ini dibuat dengan tujuan untuk memenuhi salah satu syarat dalam memperoleh gelar Sarjana Statistika pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Tanjungpura Pontianak.

Selama proses penulisan skripsi ini, penulis menyadari bahwa tidak lepas dari bimbingan, arahan, serta dukungan dari berbagai pihak. Oleh karena itu penulis ingin mengucapkan terima kasih kepada:

1. Keluarga tercinta yang selalu memberikan doa, semangat, dan motivasi serta dukungan moril dan material kepada penulis selama proses perkuliahan dan dalam menyelesaikan skripsi ini.
2. Bapak Hendra Perdana, M.Sc. selaku dosen pembimbing pertama yang telah banyak memberikan bimbingan serta arahan selama penulisan skripsi ini.
3. Ibu Neva Satyahadewi, M.Sc., CRA., CRP., CRMP. selaku dosen pembimbing kedua yang telah banyak memberikan bimbingan serta arahan selama penulisan skripsi ini.
4. Ibu Nur'ainul Miftahul Huda, M.Si. selaku Dosen Pembimbing Akademik yang telah meluangkan waktu, memberikan saran, masukan, dan motivasi kepada penulis selama proses perkuliahan.
5. Ibu Dr. Evy Sulistianingsih, M.Sc. selaku dosen penguji pertama yang telah memberikan masukan, saran, serta arahan kepada penulis.
6. Ibu Nurfitri Imro'ah, M.Si. selaku dosen penguji kedua yang telah memberikan masukan, saran, serta arahan kepada penulis.
7. Rekan-rekan Statistika 2021 "Expos" yang telah menemani dan memberikan dukungan serta semangat selama penulisan skripsi ini.

8. Semua pihak yang tidak bisa penulis sampaikan satu persatu atas dukungan dan bantuannya baik dari segi material maupun non material.

Semoga Allah membalas semua kebaikan dan bantuan yang diberikan dalam penulisan skripsi ini. Penulis mengharapkan kritik dan saran yang bersifat membangun untuk menjadikan skripsi ini lebih baik. Akhir kata, semoga skripsi ini dapat memberikan wawasan dan informasi, serta memberikan manfaat bagi penulis dan semua pihak.

Pontianak, 16 Juni 2025

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN ii

PERNYATAAN INTEGRITAS AKADEMIK iii

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 3

1.4 Batasan Masalah 3

1.5 Tinjauan Pustaka 4

1.6 Metodologi Penelitian 5

BAB II LANDASAN TEORI 8

2.1 Indeks Desa Membangun 8

2.2 Dimensi Pembentukan IDM 9

2.3 Analisis Multivariat 10

2.4 Jarak *Gower (Gower Distance)* 11

2.5 *Confusion Matrix* 12

BAB III *K-MEANS BOOTSTRAP* 13

3.1 Analisis *Cluster* 13
3.2 *K-Means Clustering* 14
3.3 Metode *Bootstrap* 15
3.4 *Majority Voting* 16
BAB IV HASIL DAN PEMBAHASAN 18
4.1 Deskripsi Data 18
4.2 Pengelompokan dengan Metode *K-Means* 19
4.2.1 Penentuan *Centroid* Awal 19
4.2.2 Perhitungan *Centroid* Baru 21
4.3 Pengelompokan dengan Metode *K-Means Bootstrap* 22
4.4 Interpretasi Hasil *Clustering* 23
BAB V PENUTUP 29
5.1 Kesimpulan 29
5.2 Saran 29
DAFTAR PUSTAKA 30
LAMPIRAN 33

DAFTAR GAMBAR

Gambar 1.1 *Flowchart* Metode Penelitian 7
Gambar 4.1 Peta Pengelompokan Desa Menggunakan *K-Means Bootstrap* 26

DAFTAR TABEL

Tabel 2.1 Klasifikasi Status Desa Berdasarkan IDM 9
Tabel 2.2 *Confusion Matrix*..... 12
Tabel 4.1 Statistik Deskriptif Data 18
Tabel 4.2 *Centroid* Awal Secara Acak 20
Tabel 4.3 Hasil Pengelompokan Desa di Kabupaten Mempawah dengan 5 *Cluster* 23
Tabel 4.4 Karakteristik Masing-masing *Cluster* 25
Tabel 4.5 *Confusion Matrix* Metode *K-Means Bootstrap* dengan IDM..... 27

DAFTAR LAMPIRAN

Lampiran I. Data Indikator Indeks Ketahanan Ekonomi (IKE) Berdasarkan IDM di Kabupaten Mempawah Tahun 2022 33

Lampiran II. Prosedur Pemberian Skor Setiap Indikator Indeks Ketahanan Ekonomi (IKE) Berdasarkan IDM 35

DAFTAR SIMBOL

D	: Nilai dimensi IDM
I_a	: Nilai indikator ke-a
b	: Banyak indikator
K	: Banyak *cluster*
k	: *Cluster* ke-k dengan $k = 1, 2, 3, \dots, K$
I	: Banyak objek penelitian
i	: Objek ke-i dengan $i = 1, 2, 3, \ \dots, I$
x_i	: Data dari objek ke-i
$x_{i:j}$	: Data dari objek ke-i pada variabel ke-j
$C_{k:j}$	: *Centroid cluster* ke-k dari variabel ke-j
n_k	: Banyak objek pada *cluster* ke-k
$d_{(a,b)}$	: Jarak *Gower* antara *centroid* ke-a dan objek ke-b
r	: Jumlah variabel
$R_{qc}(a)$	: *Rank* data *centroid* ke-a pada variabel ke-q
$R_{qo}(b)$	: *Rank* data objek ke-b pada variabel ke-q
R_q	: *Rank* pada variabel ke-q
B	: Jumlah iterasi *Bootstrap*

BAB I
PENDAHULUAN

1.1 Latar Belakang

Pembangunan desa merupakan salah satu upaya yang dilakukan untuk mengurangi terjadinya tingkat kesenjangan antara wilayah kota dan desa akibat pembangunan ekonomi yang lebih berpusat di kota (Sari & Oktavianor, 2021). Pembangunan desa dilakukan pemerintah sebagai strategi menciptakan kesejahteraan desa. Untuk melihat tingkat kesejahteraan dan kemajuan pembangunan desa diperlukan suatu tolak ukur. Tingkat kesejahteraan suatu desa dapat diukur melalui status desa berdasarkan Indeks Desa Membangun atau IDM.

IDM disusun untuk membantu menangani desa tertinggal dan meningkatkan kemandirian desa berdasarkan beberapa indikator (Muhtarom et al., 2018). Terdapat 3 indeks dalam IDM, yaitu Indeks Ketahanan Sosial, Indeks Ketahanan Ekonomi, dan Indeks Ketahanan Lingkungan atau Ekologi. Indikator-indikator yang digunakan dalam IDM dikembangkan agar dapat mencapai tujuan dari pembangunan desa yang maju dan mandiri. Dalam hal ini diperlukan kerangka kerja pembangunan berkelanjutan yang memperhatikan aspek sosial, ekonomi, dan lingkungan (Kementerian Desa, Pembangunan Daerah Tertinggal, dan Transmigrasi, 2022).

Berdasarkan Peraturan Menteri Desa, Pembangunan Daerah Tertinggal dan Transmigrasi No. 2 Tahun 2016 Pasal 5, Ayat 3 mengenai Indeks Desa Membangun menjelaskan bahwa status kemajuan maupun kemandirian desa didasarkan pada IDM dan ditetapkan melalui Surat Keputusan Direktur Jenderal Pembangunan dan Pemberdayaan Masyarakat Desa Kementerian Desa, Pembangunan Daerah Tertinggal, dan Transmigrasi Republik Indonesia. Pada IDM status desa terbagi menjadi lima kategori: sangat tertinggal, tertinggal, berkembang, maju, dan mandiri. Pembagian status tersebut disusun untuk menunjukkan keberagaman karakteristik status kemajuan dan

kemandirian desa. Desa dengan status sangat tertinggal dan tertinggal memerlukan pendekatan dan perhatian kebijakan yang berbeda dari desa lain yang lebih maju dan mandiri (Xaverius, 2021).

Dalam upaya memberikan perhatian khusus untuk desa yang memiliki status yang rendah, perlu dilakukan pengelompokan desa berdasarkan status IDM. Pengelompokan ini dapat memudahkan dalam mengenali karakteristik desa berdasarkan kelompok status tersebut. Analisis *cluster* dapat digunakan untuk mengelompokkan desa berdasarkan kesamaan karakteristik desa tersebut (Imanni, Sulistianingsih, dan Perdana, 2023).

Analisis *cluster* adalah teknik statistik yang bertujuan untuk mengelompokkan objek berdasarkan kesamaan dalam karakteristik mereka. Tujuannya adalah untuk mengelompokkan objek-objek tersebut ke dalam satu atau lebih kelompok (*cluster*), sehingga objek yang berada dalam satu kelompok memiliki kemiripan satu sama lain (Ramadhani, Purnamasari, dan Amijaya 2018). Analisis *cluster* adalah teknik statistik yang memungkinkan pengelompokan objek berdasarkan tingkat kesamaan mereka, sehingga objek yang memiliki kesamaan tertinggi akan dikelompokkan bersama dalam *cluster* yang sama (Haumahu & Nanlohy, 2020). Analisis *cluster* merupakan langkah dalam menyusun atau mengkategorikan objek-objek berdasarkan informasi yang diambil dari data yang menjelaskan keterkaitan antara objek-objek tersebut. Tujuannya adalah untuk meningkatkan keseragaman di antara anggota-anggota dalam satu kelas dan mengurangi keseragaman antara kelas atau kelompok (Rahmawati, Sihwi, dan Suryani, 2016).

Terdapat dua teknik umum yang digunakan pada analisis *cluster*, ialah non-hirarki serta hirarki. Pada teknik non-hirarki metode yang paling umum digunakan yaitu metode *K-Means* (Syafiyah, Asrafi, Wicaksono, Puspitasari, dan Sirait, 2022). Teknik Non-hierarki dilakukan dengan menentukan berapa jumlah *cluster* yang diinginkan. Kemudian dilanjutkan dengan memilih pusat *cluster* sementara. Pusat *cluster* yang dipilih tersebut akan terus diperbaharui pada setiap iterasi hingga kriteria tercapai. Oleh karena itu objek dapat berpindah dari *cluster* yang satu ke *cluster* yang lain (Imanni, Sulistianingsih, dan Perdana, 2023).

Metode *Bootstrap* merupakan metode yang memanfaatkan teknik *resampling* data sampel dengan pengembalian untuk menyelesaikan statistik ukuran sampel. Prinsip metode *Bootstrap* adalah untuk memperkirakan parameter masing-masing sampel *bootstrap* sebanyak *B* yang merupakan sampel acak yang diambil dengan pengembalian dari populasi pengamatan. Tujuannya adalah agar sampel tersebut dapat mewakili data populasi yang lebih besar (Faisal & Rifai, 2023).

Berdasarkan dengan latar belakang yang telah di sajikan, tujuan utama penelitian ini adalah untuk pengelompokan desa di Kabupaten Mempawah berdasarkan indikator indeks desa membangun. Adapun manfaat yang diharapkan dengan dilakukan penelitian ini yaitu dapat memberikan informasi karakteristik masing-masing desa di Kabupaten Mempawah berdasarkan status desa pada indikator indeks desa membangun. Sehingga pemerintah dapat melakukan upaya untuk membangun desa yang tertinggal dan sangat tertinggal.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka terdapat rumusan masalah dari penelitian ini mencakup bagaimana analisis pengelompokan desa di Kabupaten Mempawah berdasarkan IDM menggunakan *K-Means Clustering* dengan *Bootstrap*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan diatas, maka penelitian ini bertujuan menganalisis pengelompokan desa di Kabupaten Mempawah berdasarkan IDM menggunakan *K-Means Clustering* dengan *Bootstrap.*

1.4 Batasan Masalah

Berdasarkan uraian dari latar belakang, maka batasan masalah pada penelitian ini yaitu:

1. Data yang digunakan merupakan data indikator Indeks Ketahanan Ekonomi (IKE) berdasarkan IDM Kabupaten Mempawah pada tahun 2022.
2. Data yang digunakan berjumlah 60 desa.
3. Banyak *cluster* yang dibentuk yaitu $K = 5$.
4. Perhitungan jarak menggunakan Jarak *Gower* (*Gower Distance*).

1.5 Tinjauan Pustaka

Ramly, Sudarmin, dan Poerwanto (2022) melakukan penelitian tentang analisis *Clustering Bootstrap* dengan metode *K-Means* yang bertujuan untuk pengelompokan status gizi anak *undersfive* yang *crosstabulated* dengan status gizi anak berdasarkan WHO-2005 di Puskesmas Ajung, Jember. Hasil penelitian menghasilkan total tingkat akurasi (TAR) dan Total Error Rate (TER) untuk menentukan status gizi metode *K-Means* adalah TAR=0,9 dan TER=0,1. Metode *K-Means Bootstrap* (B=25) dengan nilai TAR=0,925 dan TER=0,075, *K-Means Bootstrap* (B=50) dengan nilai TAR=0,9417 dan TER=0,0583, dan terakhir *K-Means Bootstrap* (B=75) dengan nilai TAR=0,9583 dan TER=0,0417 setelah *crosstabulated* dengan status gizi berdasarkan WHO-2005. *Clustering Bootstrap* dengan metode *k-means* dan *crosstabulated* dengan status gizi berdasarkan WHO-2005 menghasilkan akurasi yang sangat baik dalam menentukan status gizi anak dan metode terbaik yaitu *K-Means Bootstrap* (*B*=75). *Clustering Bootstrap* dengan metode *k-means* dan *crosstabulated* dapat digunakan sebagai salah satu cara untuk mengelompokkan status gizi anak.

Penelitian tentang analisis *cluster* menggunakan metode *K-Means* pada indikator Indeks Pembangunan Manusia (IPM) dari Talakua, Leleury, Talluta (2017). Penelitian ini bertujuan untuk pengelompokkan kabupaten/kota di Provinsi Maluku berdasarkan ukuran Indeks Pembangunan Manusia (IPM). Dalam analisis ini menghasilkan 3 *cluster* dalam pengelompokkan kabupaten/kota. *Cluster* pertama dengan angka IPM dan AHH maksimal sementara nilai indikator IPM lainnya berbanding terbalik terdiri dari 1 kabupaten/kota. Kemudian untuk *cluster* kedua dengan keseluruhan nilai indikator

IPM terjadi sangat rendah terdiri dari 6 kabupaten/kota. *Cluster* ketiga dengan nilai keseluruhan indikator IPM terjadi sangat banyak yang terdiri dari 4 kabupaten/kota.

Siswanto, Suprapedi, dan Purwanto (2018) melakukan penelitian tentang *Sample Bootstraping* dengan metode *K-Nearest Neighbor* (KNN) yang bertujuan untuk klasifikasi status desa di Kabupaten Jepara. Hasil penelitian menunjukkan bahwa penggunaan metode *Sampling Bootstrap* dapat meningkatkan nilai akurasi pada pengklasifikasian dengan metode *K-Nearest Neighbor*. Peningkatan nilai akurasi tersebut sebesar 5,41% dari tingkat akurasi pengujian dengan *K-Nearest Neighbor* sebesar 91,89% naik menjadi 97,30% saat menggunakan *Bootstrap*. Metode *K-Nearest Neighbor* dengan *Bootstrapping* dapat digunakan untuk pengklasifikasian lima status desa yaitu Desa Sangat Tertinggal, Desa Tertinggal, Desa Berkembang, Desa Maju dan Desa Mandiri dan memberikan nilai akurasi yang tinggi.

1.6 Metodologi Penelitian

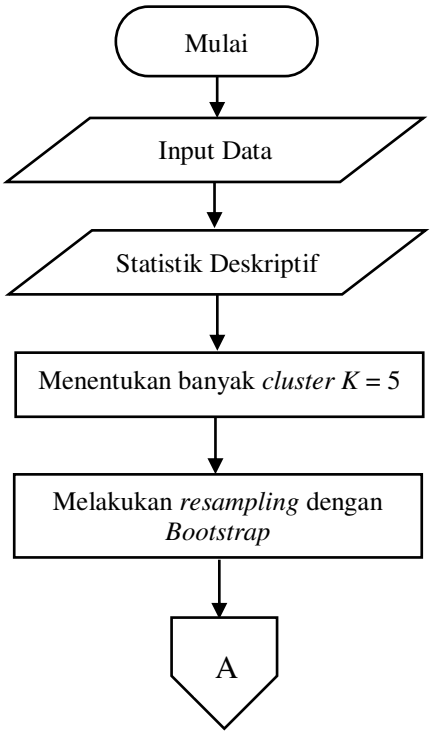
Metode yang digunakan dalam penelitian ini merupakan analisis *K-means Clustering* dengan *Bootstrap.* Studi kasus pada penelitian ini adalah menganalisis indikator Indeks Ketahanan Ekonomi berdasarkan IDM Kabupaten Mempawah tahun 2022. Data yang digunakan merupakan data sekunder yang diperoleh dari Satu Data Kalimantan Barat. Dalam penelitian ini terdapat 60 objek penelitian dengan 12 variabel yang digunakan, yaitu Keragaman Produksi (X_1), Pertokoan (X_2), Pasar (X_3), Toko/Warung Kelontong (X_4), Kedai & Penginapan (X_5), Pos & Logistik (X_6), Bank & BPR (X_7), Kredit (X_8), Lembaga Ekonomi (X_9), Moda Transportasi Umum (X_{10}), Keterbukaan Wilayah (X_{11}), dan Kualitas Jalan (X_{12}).

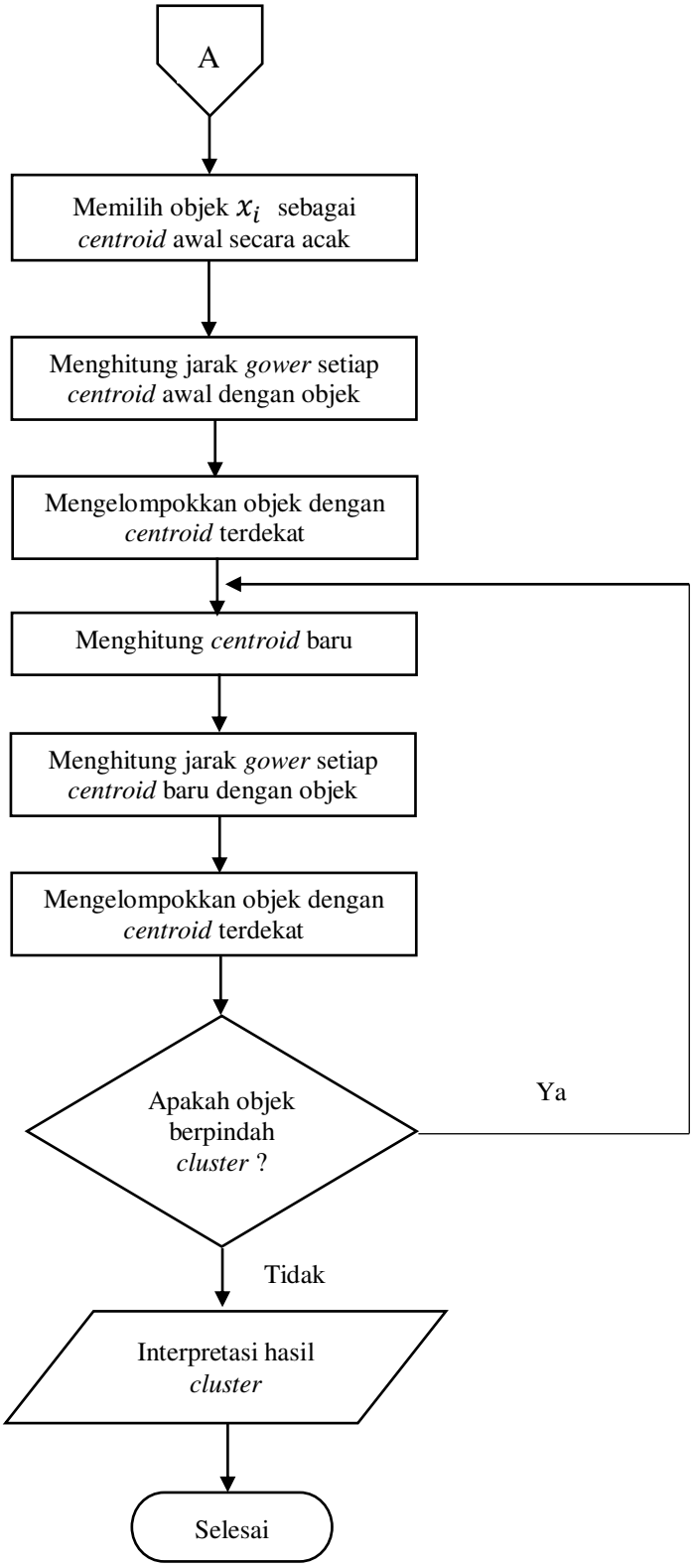
Tahapan analisis pengelompokan desa di Kabupaten Mempawah dengan algoritma *K-means Clustering* dengan *Bootstrap* dilakukan dengan langkah-langkah sebagai berikut (Jain & Moreau, 1987):

1. Melakukan analisis deskriptif tiap variabel.
2. Melakukan *resampling* data dengan *Bootstrap.*

3. Melakukan pengelompokan dengan metode *K-Means* untuk banyak *cluster* $K = 5$ dan iterasi *Bootstrap* $B = 100$. Berikut merupakan langkah-langkah pengelompokan data menggunakan metode *K-Means* (Fastaf & Yamasari, 2022).
   a. Menentukan *centroid* (titik pusat) awal secara acak sebanyak *K*.
   b. Menghitung jarak setiap objek dengan *centroid* awal menggunakan jarak *gower (gower distance).*
   c. Mengelompokkan setiap objek ke *centroid* terdekat.
   d. Mendefinisikan *centroid* baru dengan menghitung rata-rata objek pada setiap *cluster*.
   e. Menghitung jarak setiap objek dengan *centroid* baru menggunakan jarak *gower (gower distance).*
   f. Mengelompokkan setiap objek ke *centroid* terdekat.
   g. Mengulangi langkah d sampai f , hingga tidak ada perubahan dari anggota setiap *cluster*.
4. Menginterpretasi dan membuat kesimpulan.

Pada Gambar 1 menunjukkan diagram alir *(flowchart)* pada analisis *K-means Clustering* dengan *Bootstrap.*





Gambar 1.1 *Flowchart* Metode Penelitian