PENGELOMPOKAN KABUPATEN/KOTA DI KALIMANTAN BARAT BERDASARKAN FAKTOR PENYEBAB *STUNTING* MENGGUNAKAN METODE *K-HARMONIC MEANS*

(Studi Kasus: Faktor Penyebab *Stunting* di Kalimantan Barat Tahun 2021)

RAHMANIA ANDARINI HATTI IMANNI

NIM. H1091191048

SKRIPSI

PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK
2023

PENGELOMPOKAN KABUPATEN/KOTA DI KALIMANTAN BARAT BERDASARKAN FAKTOR PENYEBAB *STUNTING* MENGGUNAKAN METODE *K-HARMONIC MEANS*

(Studi Kasus: Faktor Penyebab *Stunting* di Kalimantan Barat Tahun 2021)

RAHMANIA ANDARINI HATTI IMANNI

NIM. H1091191048

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar

Sarjana Statistika pada Program Studi Statistika

PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK
2023

PENGELOMPOKAN KABUPATEN/KOTA DI KALIMANTAN BARAT BERDASARKAN FAKTOR PENYEBAB *STUNTING* MENGGUNAKAN METODE *K-HARMONIC MEANS* (Studi Kasus: Faktor Penyebab *Stunting* di Kalimantan Barat Tahun 2021)

Tanggung Jawab Yuridis Material Pada

Rahmania Andarini Hatti Imanni
H1091191048

Disetujui Oleh

Pembimbing 1	Pembimbing 2
Dr.Evy Sulistianingsih, M.Sc.	Hendra Perdana, M.Sc.
NIP. 198502172008122006	NIP. 198810102019031020

Disahkan
Oleh

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Tanjungpura Pontianak

Dr. Gusrizal, M.Si.
NIP. 197108022000031001

KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET DAN TEKNOLOGI
UNIVERSITAS TANJUNGPURA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PONTIANAK

TIM PENGUJI SKRIPSI

PENGELOMPOKAN KABUPATEN/KOTA DI KALIMANTAN BARAT BERDASARKAN FAKTOR PENYEBAB *STUNTING* MENGGUNAKAN METODE *K-HARMONIC MEANS* (Studi Kasus: Faktor Penyebab *Stunting* di Kalimantan Barat Tahun 2021)

RAHMANIA ANDARINI HATTI IMANNI
NIM. H1091191048

NAMA/NIP	TIM PENGUJI	GOLONGAN/ JABATAN	TANDA TANGAN
Dr. Evy Sulistianingsih, M.Sc. NIP. 198502172008122006	Pemimpin sidang merangkap anggota penguji	III/b Lektor	
Hendra Perdana, M.Sc. NIP. 198810102019031020	Sekretaris sidang merangkap anggota penguji	III/b Tenaga Pengajar	
Nurfitri Imro'ah, M.Si. NIP. 198907182019032021	Ketua penguji	III/b Asisten Ahli	
Neva Satyahadewi, M.Sc., CRA., CRP., CRMP. NIP. 198212042005012001	Anggota penguji	III/d Lektor	

Berdasarkan Surat Keputusan
Dekan Fakultas Matematika dan Ilmu Pengatahuan Alam
Universitas Tanjungpura Pontianak

Nomor : 1900/UN22.8/TD.06/2023
Tanggal : 30 Mei 2023
Tanggal Lulus : 8 Juni 2023

HALAMAN PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi. Sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Pontianak, Juni 2023



Rahmania Andarini Hatti Imanni

PENGELOMPOKAN KABUPATEN/KOTA DI KALIMANTAN BARAT BERDASARKAN FAKTOR PENYEBAB *STUNTING* MENGGUNAKAN METODE *K-HARMONIC MEANS*

(Studi Kasus: Faktor Penyebab *Stunting* di Kalimantan Barat Tahun 2021)

INTISARI

Stunting merupakan masalah gizi kronis yang disebabkan oleh asupan gizi yang tidak mencukupi dalam jangka waktu yang lama. Menurut hasil Survei Status Gizi Indonesia (SSGI) 2021, persentase *stunting* di Kalimantan Barat adalah 29,8% dimana angka ini lebih tinggi dari rata-rata persentase *stunting* nasional yang bernilai 24,4%. Berdasarkan tingginya kasus *stunting* di Kalimantan Barat maka diperlukan pengelompokan Kabupaten/Kota di Kalimantan Barat berdasarkan faktor penyebab *stunting* menggunakan analisis *cluster*. Analisis *cluster* adalah teknik pengelompokan yang digunakan untuk mengelompokkan suatu objek berdasarkan karakteristik yang dimiliki oleh objek tersebut. Tujuan dari penelitian ini menganalisis pengelompokan Kabupaten/Kota di Kalimantan Barat berdasarkan faktor penyebab *stunting* menggunakan *metode K-Harmonic Means* dan menganalisis banyaknya *cluster* optimal dengan menggunakan *silhouette coefficient*. Pada penelitian ini digunakan *K-Harmonic Means* untuk pengelompokan dan *silhouette coefficient* untuk mengevaluasi hasil *clustering*. Variabel yang dianalisis pada penelitian ini adalah persentase rumah tangga yang tidak memiliki akses air minum bersih (X_1), persentase kurangnya pemberian Air Susu Ibu (ASI) eksklusif (X_2), persentase Bayi Berat Lahir Rendah lahir dengan selamat (X_3), persentase rumah tangga yang tidak memiliki fasilitas sanitasi yang layak (X_4). Berdasarkan hasil analisis, jumlah *cluster* optimal yaitu 4 *cluster* dengan nilai *silhouette coefficient* terbesar sebesar 0,744 yang berarti pengelompokan yang dihasilkan adalah struktur yang kuat. *Cluster* 1 merupakan *cluster* yang memiliki faktor penyebab stunting yang sangat tinggi. faktor yang bepengaruh pada *cluster* 1 adalah rumah tangga yang tidak memiliki air minum bersih, pemberian ASI eksklusif dan bayi lahir rendah lahir dengan selamat. Anggota dari *cluster* 1 adalah Bengkayang, Ketapang dan Melawi.

Kata Kunci: *Stunting*, *Cluster*, Optimal

GROUPING OF REGENCIES/CITIES IN WEST KALIMANTAN BASED ON FACTORS THAT CAUSE STUNTING USING K-HARMONIC MEANS METHODS

(Case Study: Factors Causing *Stunting* in West Kalimantan in 2021)

ABSTRACT

Stunting is a chronic nutritional problem caused by insufficient nutritional intake over a long period of time. According to the results of the 2021 Indonesian Nutritional Status Survey (SSGI), the percentage of stunting in West Kalimantan is 29.8%, which is higher than the national average stunting percentage of 24.4%. Based on the high number of stunting cases in West Kalimantan, it is necessary to group districts/cities in West Kalimantan based on the factors causing stunting using cluster analysis. Cluster analysis is a grouping technique used to group an object based on the characteristics possessed by that object. The purpose of this study is to analyze the grouping of districts / cities in West Kalimantan based on the factors causing stunting using the K-Harmonic Means method and analyze the number of optimal clusters using silhouette coefficient. In this study, K-Harmonic Means were used for grouping and silhouette coefficient to evaluate clustering results. The variables analyzed in this study are the percentage of households that do not have access to clean drinking water (X_1), the percentage of lack of exclusive breastfeeding (X_2), the percentage of Low Birth Weight Babies (BBLR) ≤ 2,500-grams born safely (X_3), the percentage of households that do not have proper sanitation facilities (X_4). Based on the results of the analysis, the optimal number of clusters is 4 clusters with the largest silhouette coefficient value of 0.744 which means that the resulting grouping is a strong structure. Cluster 1 is a cluster that has very high stunting factors covering the Bengkayang, Ketapang and Melawi areas. Cluster 2 consists of 4 districts, Cluster 3 consists of 4 districts, and Cluster 4 consists of 3 districts/cities.

Keywords: Stunting, Cluster, Optimal

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Alhamdulillahirobbil'alamin, Puji dan Syukur kita panjatkan kepada Allah SWT atas segala pertolongan, rahmat dan kasih sayang-Nya sehingga penulisa dapat menyelesaikan skripsi yang berjudul **"Pengelompokan Kabupaten/Kota Di Kalimantan Barat Berdasarkan Faktor Penyebab *Stunting* Menggunakan Metode *K-Harmonic Means* (Studi Kasus : Faktor Penyebab *Stunting* Di Kalimantan Barat Tahun 2021)"**.

Tugas akhir ini merupakan salah satu syarat untuk memperoleh gelar Sarjana Statistika pada Program Studi Statisitika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Tanjungpura. Penulis menyadari selama penyusunan tugas akhir ini tidak terlepas dari bantuan, bimbingan dan motivasi dari berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih yang sebesar-besarnya kepada :

1. Kedua orang tua penulis, Drs. Toto Sulistianto dan Analita, S.Pd yang selalu memberi kasih sayang, doa, nasehat serta atas kesabaran yang luar biasa dalam setiap langkah hidup penulis.
2. Ibu Nurfitri Imro'ah, M.Si selaku dosen Pembimbing Akademik yang telah memberikan bimbingan, arahan, dukungan dan motivasi untuk penulis selama perkuliahan.
3. Ibu Dr. Evy Sulistianingsih, M.Sc selaku Dosen Pembimbing I yang telah banyak memberikan bimbingan, saran waktu, tenaga dan arahan serta motivasi untuk menyelesaikan skripsi ini.
4. Bapak Hendra Perdana, M.Sc selaku Dosen Pembimbing II yang telah banyak memberikan bimbingan, saran waktu, tenaga dan arahan serta motivasi untuk menyelesaikan skripsi ini.
5. Ibu Nurfitri Imro'ah, M.Si selaku Dosen Penguji I yang telah memberikan arahan, saran dan masukan untuk memperbaiki skripsi ini agar menjadi lebih baik.

6. Neva Satyahadewi, M.Sc.,CRA.,CRP.,CRMP selaku Dosen Penguji II yang telah memberikan arahan, saran dan masukan untuk memperbaiki skripsi ini agar menjadi lebih baik lagi.
7. Seluruh teman-teman mahasiswa Statistika Angkatan 2019 yang tidak dapat disebutkan satu persatu, terima kasih atas dukungan, canda tawa, semangat dan kebersamaannya selama perkuliahan ini.
8. Teman-teman saya, Tasya, Zi, Nabila, Dinda dan Yesi, terima kasih atas kesenangan, canda tawa, saran, diskusi, serta kerja samanya.
9. Semua pihak yang tidak dapat disebutkan satu persatu yang dengan tulus dan ikhlas memberikan doa dan motivasi sehingga dapat terselesaikannya skripsi ini.

Penulis menyadari bahwa dalam penulisan tugas akhir ini masih banyak kekurangan dan jauh dari sempurna. Oleh karna itu, saran dan masukan yang mengarah pada perbaikan dan kesempurnaan dari tugas akhir ini sangat penulis harapkan. Semoga penulisan tugas akhir ini dapat menambah wawasan dan pengetahuan serta bermanfaat bagi pembaca dan pihak-pihak yang memerlukan informasi yang berhubungana dengan masalah yang sama. Terima Kasih,
Wassalamu'alaikum Warahmatuhllahi Wabarakatuh.

Pontianak, Maret 2023

Rahmania Andarini Hatti Imanni

DAFTAR ISI

INTISARI ..... **i**
ABSTRACT ..... **ii**
KATA PENGANTAR ..... **iii**
DAFTAR ISI ..... **v**
DAFTAR TABEL ..... **vii**
DAFTAR GAMBAR ..... **viii**
DAFTAR LAMPIRAN ..... **ix**
DAFTAR SIMBOL ..... **x**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang Masalah 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 4
BAB II LANDASAN TEORI **9**
2.1 Faktor -Faktor Penyebab *Stunting* 9
2.2 Analisis Multivariat 11
2.3 Standarisasi Data 11
2.4 Analisis *Cluster* 12
2.5 Jarak Antar Objek 13
2.6 *K-Means* 14
BAB III *K-HARMONIC MEANS* **16**
3.1 *K-Harmonic Means* 16
3.2 Evaluasi Data Hasil *Clustering* 18
BAB IV HASIL DAN PEMBAHASAN **21**
4.1 Statistik Deskriptif 21
4.2 Standarisasi Data 22
4.3 Pendeteksian Multikoliniearitas 23
4.4 Pengelompokan dengan Menggunakan Metode *K-Harmonic Means* 23

4.4.1 Analisis Pengelompokan dengan Metode K-Harmonic Means Menggunakan 2 *Cluster* 24
4.4.2 Analisis Pengelompokan dengan Metode *K-Harmonic Means* Menggunakan 3 *Cluster* dan 4 *Cluster* 29
4.5 Evaluasi Data Hasil *Clustering* 32
4.6 Perbandingan Data Hasil *Clustering* 36
BAB V KESIMPULAN 41
DAFTAR PUSTAKA 42
LAMPIRAN 46

DAFTAR TABEL

Tabel 3.1 Tabel Rentang Nilai *Silhouette Coefficient Global* 19
Tabel 4.1 Statistik Deskriptif Variabel Penelitian 21
Tabel 4.2 Koefisien Korelasi antar Variabel 23
Tabel 4.3 Nilai Fungsi Tujuan pada $K = 2$ menggunakan p = 2, 3 dan 4 25
Tabel 4.4 Nilai Keanggotaan Data Pengamatan Menggunakan $K = 2$ 26
Tabel 4.5 Nilai Bobot Data Pengamatan pada $K = 2$ 27
Tabel 4.6 Titik Pusat *Cluster* Baru yang Terbentuk Menggunakan p = 2,3 dan 428
Tabel 4.7 Hasil Penetapan Nilai Keanggotaan Data Pengamatan 2 *Cluster* 29
Tabel 4.8 Hasil Penetapan Nilai Keanggotaan Data Pengamatan 3 *Cluster* 30
Tabel 4.9 Hasil Penetapan Nilai Keanggotaan Data Pengamatan 4 *Cluster* 31
Tabel 4.10 Hasil Evaluasi Hasil *Clustering* Menggunakan Berdasarkan dari Nilai SCi, kuntuk Setiap Data Menggunakan $K = 2$ dan menggunakan $p = 2$ 34
Tabel 4.11 Hasil Evaluasi *Clustering* Berdasarkan dari Nilai $SC2$ dengan $K = 2$ dan menggunakan $p = 2$ 36
Tabel 4.12 Perbandingan Hasil *Clustering* Menggunakan Nilai SC *Global* 36
Tabel 4.13 Hasil pengelompokan untuk 4 *cluster* dengan $p = 2$ Kabupaten/Kota Menggunakan Metode K-Harmonic Means 37
Tabel 4.14 Nilai Rata-Rata Setiap Variabel Untuk Masing-Masing *Cluster* 38

DAFTAR GAMBAR

Gambar 1. 1 Flowchart Metode *K-Harmonic Means* 7

DAFTAR LAMPIRAN

Lampiran 1 Data Penelitian..... 46
Lampiran 2 Data Standarisasi..... 47
Lampiran 3 *Syntax* Metode *K-Harmonic Means*..... 48

DAFTAR SIMBOL

N	: Banyak data pengamatan
K	: Banyak *cluster*
k	: *Cluster* ke-k, dengan k = 1,2,3...,K
x_i	: Data pengamatan ke-i
X	: Variabel pengamatan
M	: Jumlah Variabel
X_j	: Variabel pengamatan ke-j
c_k	: Titik pusat *cluster* ke- k
p	: Parameter
$\|x_i - c_k\|$	: Jarak antara data x_i ke pusat *cluster* c_k
$KHM(X, C)$	: Nilai fungsi tujuan *K-Harmonic Means* untuk variabel pengamatan terhadap titik pusat *cluster*
$m(c_k \| x_i)$	: Nilai keanggotaan titik pusat *cluster* ke-k terhadap data pengamatan ke-i
$w(x_i)$	: Nilai bobot untuk data pengamatan ke-i
$a_i(k)$	: Rata-rata jarak ke-i dengan semua data pada *cluster* yang sama
$\|x_{i,k} - x_{r,k}\|$	: Jarak antara data ke-i dengan data ke-r pada *cluster* yang sama
$\|x_{i,k} - x_{r,l}\|$	: Jarak antara data ke-i dengan data ke-r pada *cluster* yang berbeda
$b(k)$	: Nilai terkecil dari jarak antara data ke-i dengan data ke-r pada *cluster* yang berbeda
$SC_{i,k}$	: Nilai *silhouette coefficient* pada data ke-i
$SC(k)$	: Nilai *silhouette coefficient* pada *cluster* ke-k
SC	: Nilai *silhouette coefficient global*
n_k	: Banyaknya anggota *cluster* dalam *cluster* ke-k
n_l	: Banyaknya anggota *cluster* pada *cluster* ke-l

$d_i(k)$: Jarak data ke-i dengan semua data pada *cluster* yang berbeda

$\hat{x}_{i,j}$: Standarisasi data ke-i pada variabel ke-j

$r_{x_j x_l}$: Nilai koefisien korelasi antara variabel ke-j dan ke-l

BAB I
PENDAHULUAN

1.1 Latar Belakang Masalah

Stunting merupakan masalah gizi kronis yang disebabkan oleh asupan gizi yang tidak mencukupi dalam jangka waktu yang lama. *Stunting* umumnya dikaitkan dengan asupan makanan yang tidak memenuhi kebutuhan gizi. Masalah *stunting* dimulai sejak dalam kandungan dan baru akan terlihat saat anak sudah menginjak usia dua tahun (Promkes, 2018).

Anak *stunting* memiliki risiko lebih tinggi terkena penyakit degeneratif seperti kanker, diabetes dan obesitas. Kebutuhan gizi makro dan mikro pada tubuh tidak terpenuhi secara optimal, sehingga perkembangan tubuh dan fungsi sel tubuh lainnya tidak terbentuk dengan sempurna. *Stunting* tidak hanya mempengaruhi kesehatan tetapi juga kecerdasan. *Intelligence Quotient* (IQ) anak *stunting* lebih rendah dari rata-rata IQ anak normal (NestleHealthSciene, 2022). Masalah *stunting* dipengaruhi oleh empat faktor utama yang terdiri dari rumah tangga yang tidak memiliki akses air minum bersih, kurangnya pemberian ASI eksklusif, bayi Berat Badan Lahir Rendah lahir dengan selamat dan rumah tangga yang tidak memiliki fasilitas sanitasi yang layak (Kemenkes, 2018).

Menurut Databoks (2022), pada tahun 2021 Kalimantan Barat merupakan salah satu provinsi dengan peringkat *stunting* yang tinggi di tingkat nasional. Karena tingginya peringkat *stunting*, Kalimantan Barat merupakan salah satu provinsi prioritas untuk mempercepat pengendalian *stunting*. Hasil Survei Status Gizi Indonesia (SSGI) 2021 menunjukkan peringkat *stunting* di Kalimantan Barat sebesar 29,8%, lebih tinggi dari rata-rata peringkat *stunting* nasional sebesar 24,4%. Namun, angka ini 1,7% lebih rendah dari Survei Status Gizi Balita Indonesia (SSGBI) 2019 yang berada di angka 31,5%.

Berdasarkan fakta tersebut, Tim Percepatan Penurunan *Stunting* (TP2S) Provinsi Kalimantan Barat berkomitmen untuk menurunkan angka *stunting* menjadi 25,49% pada akhir tahun 2022. Setelah mencapai angka 25,49%, pada

tahun 2023 ditargetkan angka *stunting* akan turun menjadi 21,28%. Selanjutnya pada tahun 2024 ditargetkan untuk turun sebesar 17,07% mendekati angka target nasional yaitu 14% (Muharrami,2022).

Penelitian ini dilakukan untuk mengetahui faktor penyebab *stunting* pada setiap Kabupaten/Kota di Kalimantan Barat. Dengan mengelompokkan Kabupaten/Kota mana yang memiliki kesamaan faktor penyebab *stunting* dapat memudahkan tugas pemerintah untuk memberikan bantuan, pertolongan serta kegiatan pencegahan terhadap kasus *stunting* di setiap Kabupaten/Kota di Kalimantan Barat. Pengelompokan ini dapat dilakukan dengan menggunakan analisis *cluster*.

Analisis *cluster* adalah teknik pengelompokan yang digunakan untuk mengelompokkan suatu objek berdasarkan karakteristik yang dimiliki oleh objek tersebut. Salah satu metode *cluster* yang sering digunakan adalah metode *non hierarki*. Dalam metode ini, objek dalam data dikelompokkan menjadi *k cluster*, dengan jumlah *cluster* ditentukan oleh peneliti (Salsabila,S.A, Widihari dan Sudarho, 2022).

Metode *non hierarki* yang umumnya digunakan adalah *K-Means* karena metode sangat mudah diimplementasikan. Namun, *K-Means* memiliki beberapa kelemahan, yaitu hasil *cluster* yang sensitif terhadap hasil penentuan awal (inisialisasi) di pusat *cluster*. Hasil yang berbeda dari proses *cluster* dengan menggunakan data yang sama sering kali disebabkan oleh inisialisasi titik pusat (*centroid*) yang berbeda (Wijaya, Mukid dan Ispriyanti, 2018).

Untuk mengatasi masalah yang terjadi, maka dikembangkan metode *K-Harmonic Means* yang merupakan pengembangan dari metode *K-Means*. Tujuan dari metode *K-Harmonic Means* adalah untuk meminimalkan rata-rata harmonik dari semua titik data ke semua pusat *cluster* yang ada. *K-Harmonic Means* tidak sensitif terhadap inisialisasi pusat *cluster* awal *(centroid*) dan secara signifikan meningkatkan kualitas pengelompokan dibanding dengan *K-Means* (Gunadi,I.G.A, Putri,D.P, Mistanada.I.N, Prayoga,G.S dan Rahayu,N.M.Y.D, 2019).

Hal lain yang perlu diperhatikan dalam analisis *cluster*, yaitu evaluasi hasil *clustering*. Evaluasi hasil *clustering* digunakan untuk mendapatkan pengelompokan yang paling sesuai dengan data. Jika tidak dilakukan evaluasi maka mempengaruhi

hasil analisis. Pada penelitian ini digunakan optimalisasi dengan pendekatan internal yaitu *silhouette coefficient*.

Berdasarkan latar belakang yang dikemukakan, metode *K-Harmonic Means* dapat digunakan untuk mengelompokkan Kabupaten/Kota mana saja yang memiliki faktor penyebab *stunting* yang sama. Kemudian dilakukan evaluasi hasil *clustering* untuk mendapatkan pengelompokan yang paling optimal dengan data, dengan menggunakan optimalisasi *silhouette coefficient*.

1.2 Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah:

1. Bagaimana analisis pengelompokan Kabupaten/Kota di Kalimantan Barat berdasarkan faktor penyebab *stunting* dengan menggunakan metode *K-Harmonic Means*?
2. Bagaimana analisis *cluster* optimal dengan menggunakan *silhouette coefficient* pada pengelompokan Kabupaten/Kota di Kalimantan Barat berdasarkan faktor penyebab *stunting*?

1.3 Tujuan Penelitian

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

1. Menganalisis pengelompokan Kabupaten/Kota di Kalimantan Barat berdasarkan faktor penyebab *stunting* di Kalimantan Barat dengan menggunakan metode *K-Harmonic Means*.
2. Menganalisis banyaknya *cluster* optimal dengan menggunakan *silhouette coefficient* pada pengelompokan Kabupaten/Kota di Kalimantan Barat berdasarkan faktor penyebab *stunting*.

1.4 Batasan Masalah

Batasan masalah dalam penelitian ini adalah:

1. Objek penelitian ini dibatasi dengan menggunakan 4 variabel berdasakan dari faktor penyebab *stunting* di Kalimantan Barat tahun 2021, yaitu persentase

rumah tangga yang tidak memiliki akses air minum bersih (X_1), persentase kurangnya pemberian Air Susu Ibu (ASI) eksklusif (X_2), persentase Bayi Berat Lahir Rendah lahir dengan selamat (X_3), persentase rumah tangga yang tidak memiliki fasilitas sanitasi yang layak(X_4).

2. Banyak *cluster* (K) yang digunakan dalam penelitian ini adalah 2, 3 dan 4.
3. Parameter (p) yang digunakan dalam penelitian ini adalah 2, 3 dan 4.
4. Evaluasi hasil *clustering* dengan menggunakan metode *silhouette coefficient*.

1.5 Tinjauan Pustaka

Widiartha (2011), melakukan penelitian tentang studi komparasi metode klasterisasi data *K-Means* dan *K-Harmonic Means*. Dalam penelitian ini, nilai simpangan baku dari metode *K-Harmonic Means* lebih kecil dibandingkan dengan menggunakan metode *K-Means* yang artinya metode *K-Harmonic Means* memiliki persebaran data yang kecil. Berdasarkan hasil penelitian didapatkan hasil bahwa metode *K-Harmonic Means* telah terbukti berhasil mengoptimalkan posisi titik pusat *cluster* dengan mengarahkan hasil *cluster* menuju solusi global optimal.

Handoyo, Rumani dan Nasution (2014), melakukan penelitian untuk mengukur kualitas *cluster* dengan menggunakan metode *silhouette coefficient*. *Silhouette coefficient* adalah salah satu metode yang dapat digunakan untuk mengevaluasi hasil *clustering* dengan menggabungkan metode separasi dan kohesi.

Dewi dan Pramita (2019), melakukan penelitian tentang membandingkan metode elbow dan *silhouette coefficient* menggunakan algoritma *cluster K-Medoids* dalam pengelompokan produksi kerajinan Bali untuk menentukan jumlah *cluster* yang tepat sehingga menghasilkan kualitas *cluster* yang optimal, didapatkan bahwa metode *silhouette coefficient* memiliki hasil yang lebih baik karena memiliki nilai DBI yang lebih rendah dibandingkan dengan metode *elbow*.

1.6 Metodologi Penelitian

Metodologi dalam penelitian ini berupa studi literatur dan studi kasus, Studi literatur dalam penelitian ini dilakukan dengan mencari teori-teori pendukung mengenai analisis *cluster*, metode *K-Harmonic Means cluster* dan evaluasi hasil

clustering dengan menggunakan metode *silhouette coefficient* melalui buku dan jurnal.

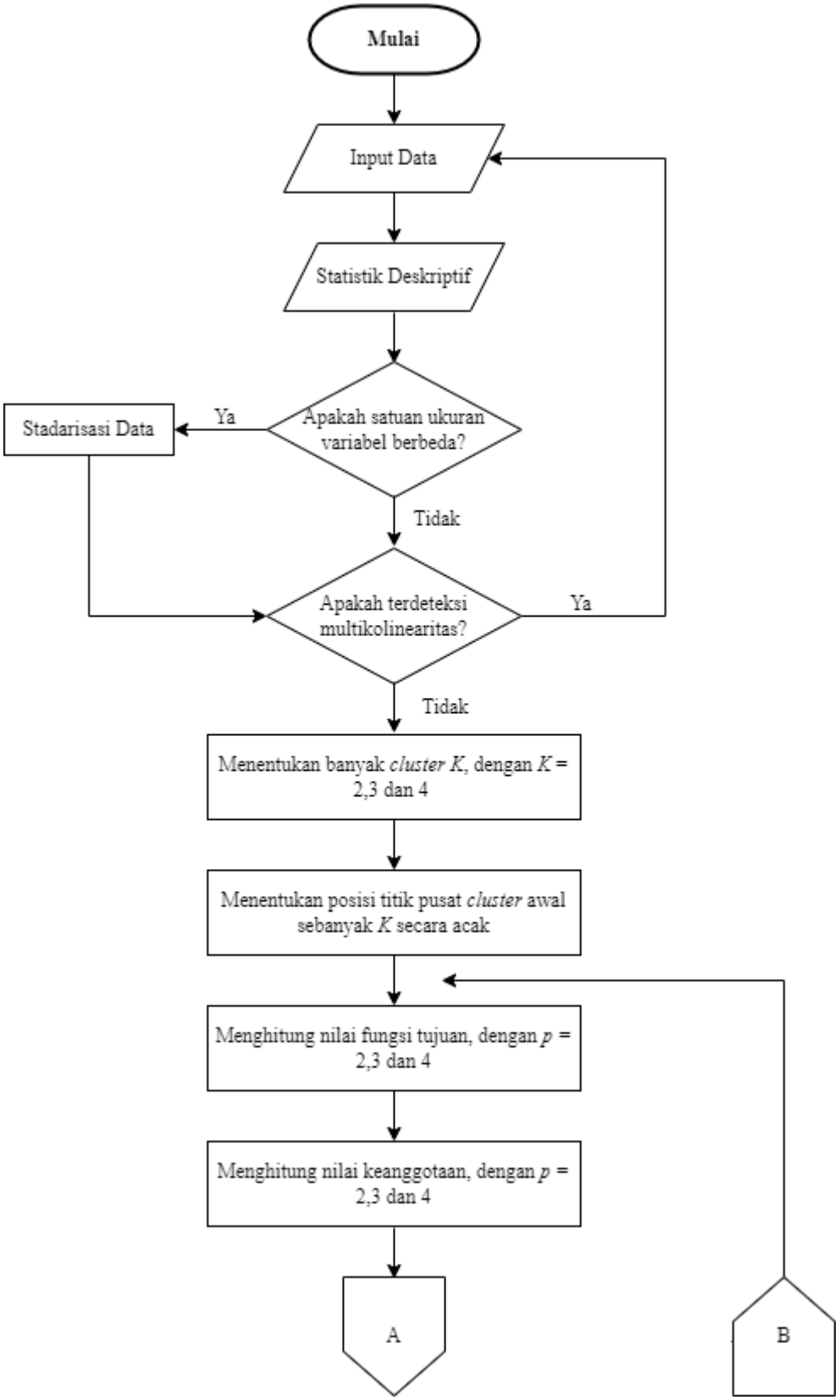
Studi kasus dalam penelitian ini dilakukan dengan menganalisis data empat faktor *stunting* di Kalimantan Barat tahun 2021. Data penelitian ini merupakan data sekunder yang dipublikasikan oleh Badan Pusat Statistik (BPS) Provinsi Kalimantan Barat dan data dari Dinas Kesehatan Provinsi Kalimantan Barat Tahun 2021. Dalam penelitian ini terdapat 4 variabel yang digunakan yaitu persentase rumah tangga yang tidak memiliki akses air minum bersih (X_1), persentase kurangnya pemberian Air Susu Ibu (ASI) eksklusif (X_2), persentase Bayi Berat Lahir Rendah lahir dengan selamat (X_3), persentase rumah tangga yang tidak memiliki fasilitas sanitasi yang layak(X_4).

Berikut ini adalah tahapan analisis data yang digunakan adalah sebagai berikut dengan menggunakan bantuan *software R*.

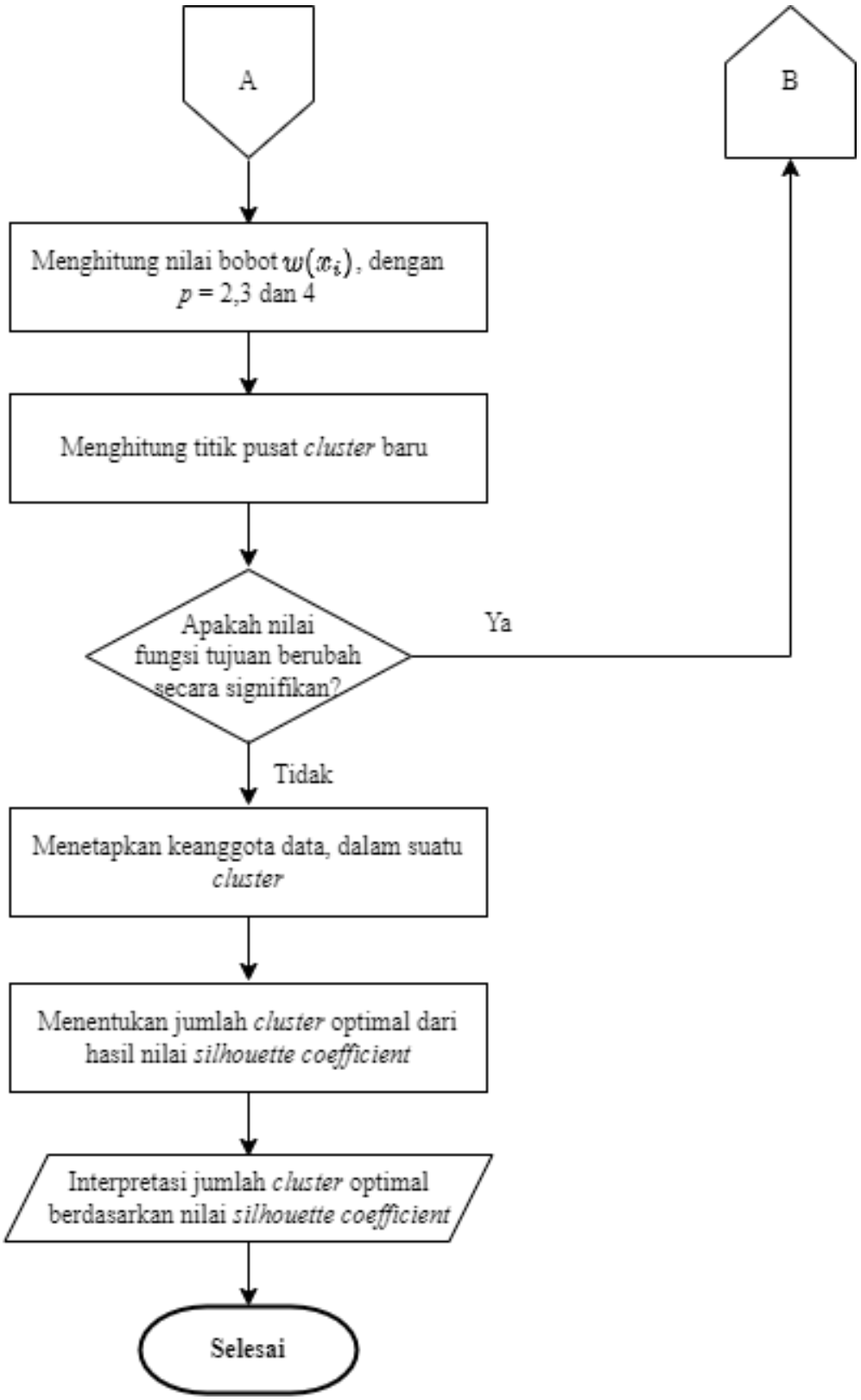
1. Melakukan analisis statistik deskriptif terhadap data faktor penyebab *stunting* di Kalimantan Barat tahun 2021 untuk mengetahui deskripsi data secara umum dan karakteristik data.
2. Melakukan standarisasi data
3. Deteksi multikolinearitas dengan menggunakan korelasi *pearson product momen*.
4. Melakukan pengelompokan data menggunakan metode *K-Harmonic Means* untuk banyak *cluster K* = 2, 3 dan 4. Tahapan metode *K-Harmonic Means* adalah sebagai berikut.
   a. Menentukan titik pusat *cluster* awal secara *random* dari data yang ada.
   b. Menghitung fungsi tujuan dengan menggunakan p = 2, 3 dan 4.
   c. Untuk setiap data x_i, menghitung nilai keanggotaan $m(c_k|x_i)$ pada setiap titik untuk pusat *cluster* c_k.
   d. Untuk setiap data x_i, menghitung nilai bobot $w(x_i)$.
   e. Untuk setiap titik pusat *cluster* c_k, dihitung kembali perhitungan untuk semua titik pusat *cluster* dari semua data berdasarkan dari nilai keanggotaan dan nilai bobot yang dimiliki oleh setiap data.

f. Mengulang langkah dari b hingga e sampai mendapatkan nilai fungsi tujuan yang tidak berubah signifikan.
g. Menetapkan keanggotaan data x_i dalam suatu *cluster* dengan titik pusat *cluster* c_k sesuai dengan nilai keanggotaan x_i terhadap titik pusat *cluster* c_k.

5. Menghitung nilai *silhouette coefficient* agar mengetahui kualitas dari hasil pengelompokan. Tahapan menghitung nilai *silhouette coefficient* adalah sebagai berikut.
   a. Menghitung rata-rata jarak dari suatu data ke-i dengan semua data yang terletak pada suatu *cluster* yang sama.
   b. Menghitung rata-rata jarak dari suatu data ke-i dengan semua data yang terletak pada *cluster* yang berbeda.
   c. Menghitung nilai *sillhouette coefficient* $(SC_{i,k})$ untuk setiap data ke-i.
   d. Menghitung nilai *sillhouette coefficient* $(SC(k))$ untuk *cluster* ke- k.
   e. Menghitung rata-rata dari $(SC(k))$ untuk mendapatkan nilai SC *global*.
   f. Mentukan nilai K yang optimal berdasarkan dari nilai *silhouette coeffiecient* terbesar.



Gambar 1. 1 *Flowchart* Metode *K-Harmonic Means*



Gambar 1. 1 *Flowchart* Metode *K-Harmonic Means* (Lanjutan)