PENGELOMPOKAN PROVINSI DI INDONESIA BERDASARKAN KASUS KRIMINALITAS MENGGUNAKAN METODE *WARD* DENGAN OPTIMALISASI *GAMMA INDEX*

DEA OKTITANNIA
NIM H1091211062

SKRIPSI

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

PENGELOMPOKAN PROVINSI DI INDONESIA BERDASARKAN KASUS KRIMINALITAS MENGGUNAKAN METODE *WARD* DENGAN OPTIMALISASI *GAMMA INDEX*

DEA OKTITANNIA
NIM H1091211062

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 Provinsi Di Indonesia Berdasarkan Kasus Kriminalitas Menggunakan Metode *Ward* Dengan Optimalisasi *Gamma Index*
Nama Mahasiswa	: Dea Oktitannia
NIM	: H1091211062
Jurusan/Program Studi	: Matematika/ Statistika
Tanggal Lulus	: 31 Januari 2025
SK Pembimbing	: No. 2636/UN22.8/TD.06/2024 /Tanggal 2 September 2024
SK Penguji	: No. 102 /UN22.8/TD.06/2025 /Tanggal 23 Januari 2025

Dosen Pembimbing

Pembimbing I	Pembimbing II
Dr. Evy Sulistianingsih, M.Sc.	Shantika Martha, M.Si.
NIP 198502172008122006	NIP 198403082008122003

Dosen Penguji

Ketua Penguji	Anggota Penguji
Neva Satyahadewi, M.Sc.,CRA.,CRP.,CRMP	Hendra Perdana, M.Sc.
NIP 198212042005012001	NIP 198810102019031020
Pimpinan Sidang (merangkap anggota penguji)	Sekretaris Sidang (merangkap anggota penguji)
Dr. Evy Sulitianingsih, M.Sc.	Shantika Martha, M.Si.
NIP 198502172008122006	NIP 198403082008122003

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 : Dea Oktitannia

NIM : H1091211062

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 pernyatan ini untuk digunakan sebagaimana mestinya.

Pontianak, 31 Januari 2025

Dea Oktitannia
H1091211062

Pengelompokan Provinsi Di Indonesia Berdasarkan Kasus Kriminalitas Menggunakan Metode *Ward* Dengan Optimalisasi *Gamma Index*

Abstrak

Kriminalitas menjadi permasalahan yang serius di Indonesia yang masih terjadi hingga saat ini dan dapat mempengaruhi kehidupan sosial masyarakat. Oleh karena itu, upaya menurunkan angka kriminalitas menjadi hal terpenting dilakukan, salah satunya dengan meningkatkan kewaspadaan pemerintah maupun masyarakat terhadap potensi kejahatan di lingkungan sekitar. Tujuan dari penelitian ini untuk pengelompokan provinsi di Indonesia berdasarkan kasus kriminalitas dan menentukan jumlah cluster optimal yang terbentuk. Salah satu metode yang digunakan dalam proses pengelompokan ini adalah analisis cluster, yang memungkinkan objek dikelompokkan berdasarkan informasi yang terkandung dalam data tersebut, termasuk hubungan antar objek. Dengan demikian, karakteristik setiap cluster dapat mengidentifikasi, sehingga tingkat keadaan darurat serta kebutuhan di masing-masing cluster dapat dianalisis dan di tindaklanjuti secara lebih efektif. Metode Ward merupakan metode pengelompokan yang bertujuan meminimalkan varians di dalam cluster. Selanjutnya, untuk menentukan jumlah cluster optimal digunakan Gamma Index. Hasil penelitian menunjukkan bahwa terbentuknya dua cluster optimal untuk pengelompokan provinsi di Indonesia berdasarkan kasus kriminalitas dengan nilai Gamma Index sebesar 0,9824 yang merupakan nilai lebih tertinggi dibandingkan cluster lainnya. Cluster 1 merupakan daerah karakteristik kriminalitasnya rendah terdiri dari 29 provinsi. Cluster 2 merupakan daerah karakteristik kriminalitasnya tinggi yang terdiri dari 5 provinsi. Untuk mengatasinya, diperlukan peningkatan penegakan hukum, analisis faktor pemicu kriminalitas, serta program rehabilitasi dan edukasi bagi pelaku kejahatan. Kolaborasi antara pemerintah, kepolisian dan masyarakat juga perlu diperkuat untuk menciptakan lingkungan yang lebih aman dan kondusif.

Kata kunci: *cluster* optimal*, concordant, discordant*

Clustering Provinces in Indonesia Based on Criminality Cases Using Ward Method with Optimization of Gamma Index

Abstract

Crime is a serious problem in Indonesia that still occurs today and can affect the social life of the community. Therefore, efforts to reduce the crime rate are the most important thing to do, one of which is by increasing the vigilance of the government and the community towards potential crimes in the surrounding environment. The purpose of this research is to group provinces in Indonesia based on crime cases and determine the optimal number of clusters formed. One of the methods used in this clustering process is cluster analysis, which allows objects to be grouped based on the information contained in the data, including the relationship between objects. Thus, the characteristics of each cluster can be identified, so that the level of emergency and needs in each cluster can be analyzed and followed up more effectively. Ward's method is a clustering method that aims to minimize the variance within the cluster. Furthermore, to determine the optimal number of clusters, the Gamma Index is used. The results showed that the formation of two optimal clusters for grouping provinces in Indonesia based on crime cases with a Gamma Index value of 0.9824 which is the highest value compared to other clusters. Cluster 1 is an area of low crime characteristics consisting of 29 provinces. Cluster 2 is an area of high crime characteristics consisting of 5 provinces. To overcome this, it is necessary to increase law enforcement, analyze the factors that trigger criminality, as well as rehabilitation and education programs for criminals. Collaboration between the government, police and communities also needs to be strengthened to create a safer and more conducive environment.

Keywords: *optimal cluster, concordant, discordant*

PRAKATA

Alhamdulilah, puji dan Syukur kita panjatkan kepada Allah SWT, karena berkat Rahmat-Nya sehingga saya dapat menyelesaikan penulisan tugas akhir yang berjudul "**Pengelompokan Provinsi Di Indonesia Berdasarkan Kasus Kriminalitas Menggunakan Metode *Ward* Dengan Optimalisasi *Gamma Index***" ini disusun sebagai salah satu syarat memperoleh gelar Sarjana Statistika bagi Mahasiswa Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Tanjungpura Pontianak.

Sepanjang proses penyusunan tugas akhir ini tidak lepas dari bantuan, bimbingan, dan dukungan banyak pihak. Oleh karena itu, pada kesempatan ini, saya ingin menyampaikan rasa terima kasih yang tulus kepada:

1. Kedua orang tua penulis, yang senantiasa memberikan doa, semangat dan dukungan tak terhingga selama proses penyusunan tugas akhir ini. Tanpa dukungan mereka, saya tidak akan bisa menyelesaikan tugas akhir ini.
2. Adik penulis yaitu Michele dan Naufar yang memberikan doa, semangat serta kebersamaan yang menjadi sumber energi positif selama penyusunan tugas akhir ini.
3. Ibu Dr. Evy Sulistianingsih, M.Sc. selaku dosen pembimbing pertama yang telah memberikan bimbingan dan arahan dalam menyelesaikan tugas akhir ini.
4. Ibu Shantika Martha, M.Si. selaku dosen pembimbing kedua yang telah memberikan bimbingan dan arahan dalam menyelesaikan tugas akhir ini.
5. Ibu Neva Satyahadewi, M.Sc.,CRA.,CRP.,CRMP. selaku dosen penguji pertama yang telah memberikan saran dan masukan kepada penulis.
6. Bapak Hendra Perdana, M.Sc. selaku dosen penguji kedua yang telah memberikan saran dan masukan kepada penulis.
7. Bapak Fransiskus Fran, S.Si.,M.Si. selaku dosen pembimbing akademik yang telah memberikan arahan dan membimbing penulis selama proses penyelesaian skipsi ini.
8. Kawan-kawan yang seperjuangan penulis, Setio, Nirwani, dan Syakira yang telah memberikan semangat dan dukungan dalam menyelesaikan tugas akhir.

9. Rekan-rekan seperjuangan Mahasiswa Statistika Universitas Tanjungpura Angkatan 2021.
10. Seluruh pihak yang telah memberikan doa, dukungan, dan motivasi dalam proses penyelesaian tugas akhir ini, namun tidak dapat disebutkan satu persatu.

Demikian, semoga tugas akhir ini bisa bermanfaat bagi semua pihak yang membutuhkan.

Pontianak, 31 Januari 2025

Dea Oktitannia

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 3
1.4 Batasan Masalah 3
1.5 Tinjauan Pustaka 4
1.6 Metodologi Penelitian 5
BAB II LANDASAN TEORI 9
2.1 Kriminalitas 9
2.2 Analisis Multivariat 10
2.3 Analisis *Cluster* 11
2.3.1 Analisis Hirarki 12
2.3.2 Analisis Non Hirarki 13
2.4 Asumsi Analisis *Cluster* 14
2.5 Tahapan Analisis *Cluster* 15
BAB III PENENTUAN JUMLAH *CLUSTER* OPTIMAL 19
3.1 *Ward Clustering* 19
3.4 Gamma Index 19
BAB IV HASIL DAN PEMBAHASAN 25
4.1 Deskipsi Data 25
4.2 Asumsi Analisis *Cluster* 27
4.3 Standarisasi Data 28

4.4 Perhitungan Jarak *Euclidean* Antar Objek 29
4.5 Proses Pengelompokan Menggunakan Metode *Ward*.....30
4.6 Menentukan Banyaknya *Cluster* Menggunakan *Gamma Index*.....31
4.7 Interpretasi Karakteristik *Cluster* 37
BAB V PENUTUP..... 41
5.1 Kesimpulan.....41
5.2 Saran.....41
DAFTAR PUSTAKA 42
LAMPIRAN..... 47

DAFTAR GAMBAR

Gambar 1.1 *Flowchart* Penentuan Jumlah *Cluster* Optimal 6
Gambar 2.1 Klasifikasi Prosedur *Clustering* 17
Gambar 4.1 Dendogram Pengelompokan Provinsi Indonesia 32
Gambar 4.2 Peta Hasil Pengelompokan Provinsi di Indonesia 39

DAFTAR TABEL

Tabel 3.1 Data 21
Tabel 3.1 Data (Lanjutan) 22
Tabel 3.2 Pengkategorian *Cluter* 1 Dengan *Cluster* 1 Dan *Cluster* 2 23
Tabel 3.3 Pengkategorian *Cluster* 2 Dengan *Cluster* 1 dan *Cluster* 2 24
Tabel 4.1 Variabel Penelitian 25
Tabel 4.2 Statistik Deskriptif 26
Tabel 4.3 Nilai VIF 27
Tabel 4.4 Nilai VIF mengeliminasi variabel X_6 28
Tabel 4.4 Nilai VIF mengeliminasi variabel X_9 28
Tabel 4.5 Pengelompokan Provinsi Indonesia untuk Dua *Cluster* 32
Tabel 4.6 Kategori Pada *Cluster* 1 dengan *Cluster* 1 dan *Cluster* 2 34
Tabel 4.6 Kategori Pada *Cluster* 1 dengan *Cluster* 1 dan *Cluster* 2 (Lanjutan) ... 35
Tabel 4.7 Kategori Pada *Cluster* 2 dengan *Cluster* 1 dan *Cluster* 2 36
Tabel 4.8 *Gamma Index* 37
Tabel 4.9 Karakteristik Setiap *Cluster* 37

DAFTAR LAMPIRAN

Lampiran I. Data Jumlah Kriminalitas Menurut Jenis Kejahatan Tahun 2022 ... 47
Lampiran II. Standarisasi Data ... 48
Lampiran III.Matrik Jarak *Euclidean* ... 49

DAFTAR SIMBOL

r_{ij}^2 : Koefisian korelasi sederhana variabel ke-*i* dan *j* dengan *i* = 1,2, ... ,*n* dan *j* = 1,2, ... ,*n*

a_{ij}^2 : Koefisiean korelasi parsial variabel ke-*i* dan *j*.

R_j^2 : Nilai koefisien determinasi antara X_j dengan variabel lainnya.

x_{ir} : Nilai objek ke-*i* pada variabel ke-*r* dengan *i* = 1,2, ... , *n* dan *r* = 1,2, ... , *p*

μ_r : Rata-rata variabel ke-*r*

σ_r : Simpangan baku dari variabel ke-*r*

d_{ij} : Jarak antara objek ke-*i* dan ke-*j* dengan *i*=1,2, ... , *n* dan *j*=1,2, ... , *n*

x_{jr} : Nilai objek ke-*j* pada variabel ke-*r* dengan *k* = 1,2, ... , *p*

p : Jumlah variabel *cluster.*

$d_{(AB)C}$: Jarak antara *cluster AB* dan *C*

d_{AC} : Jarak antara *cluster A* dan *C*

d_{BC} : Jarak antara *cluster B* dan *C*

$n_{(AB)}$: Banyak objek pada *cluster AB*

$n_{(C)}$: Banyak objek pada *cluster C*

GI : *Gamma Index*

s(+) : Jumlah pasangan objek yang *concordant*

s(-) : Jumlah pasangan objek yang *discordant*

$(u, v) \in I_W$: Pasangan dari objek dari satu *cluster* yang sama

$(r, s) \in I_B$: Pasangan dari objek antar *cluster* yang berbeda

BAB I
PENDAHULUAN

1.1 Latar Belakang

Kriminalitas atau kejahatan merupakan suatu tindakan atau perbuatan yang dapat menimbulkan masalah serta mengakibatkan keresahaan di tengah-tengah masyarakat. Berbagai macam tindakan kriminalitas dapat dilakukan oleh siapa pun, baik wanita maupun pria, bahkan terjadi pada berbagai kelompok usia, mulai dari anak-anak, dewasa, hingga usia lanjut (Septiani, Erwansyah dan Suryanata, 2020). Berbagai faktor mempengaruhi seseorang melakukan tindakan kriminalitas, seperti kondisi sosial, ekonomi, lingkungan, hingga faktor psikologis. Tindak kriminalitas dapat terjadi secara sadar, di mana tindakan tersebut dipikirkan, direncanakan, dan diarahkan untuk mencapai tujuan tertentu dengan kesadaran penuh. Namun, tindak kriminalitas juga dapat dilakukan dalam keadaan setengah sadar, dipengaruhi oleh dorongan kuat dari dalam diri, seperti impuls yang tak terkendali, tekanan paksaan yang besar, atau obsesi yang mendominasi (Ningsih dan Kuncoro, 2017).

Permasalahan kriminalitas di Indonesia akan menjadi isu yang dapat memepengaruhi kehidupan masyarakat (Rahmadayanti dan Rahayu, 2023). Berbagai upaya yang telah dilakukan oleh pemerintah maupun masyarakat untuk mengatasi permasalahan kriminalitas. Sejumlah program dan kebijakan terus diterapkan dan dikembangkan guna menemukan strategi yang paling tepat dan efektif dalam menekan angka kejahatan serta menciptakan lingkungan yang lebih aman dan kondusif (Rohman, 2016). Menurut Badan Pusat Statistik (BPS), jumlah kejadian kejahatan di Indonesia pada tahun 2022 mencapai 372.962 kejadian (BPS, 2023). Setiap provinsi menghadapi permasalahan kriminalitas yang berbeda, sehingga di perlukan pengelompokan berdasarkan kasus kriminalitas terjadi. Pengelompokan ini penting untuk meningkatkan kewaspadaan masyarakat serta membantu pemerintah dalam pengawasan dan penanganan daerah dengan tingkat kriminalitas yang tinggi. Metode statistik yang dapat digunakan untuk pengelompokan yaitu analisis *cluster.*

Analisis *cluster* adalah metode yang digunakan untuk mengelompokkan objek berdasarkan informasi yang terdapat dalam data tersebut, termasuk hubungan antar objek. Metode ini bertujuan untuk menghasilkan kelompok yang anggotanya memiliki kemiripan tinggi di dalam satu *cluster* dan berbeda secara signifikan dengan anggota *cluster* lainnya, sehingga terjadi tingkat homogenitas yang tinggi

dalam *cluster* dan perbedaan yang jelas antar *cluster* (Ningrat, Maruddani dan Wuryandari, 2016). Analisis *cluster* terbagi menjadi dua metode yang sering digunakan dalam pengelompokan data yaitu metode hirarki dan metode non hirarki. Kedua jenis analisis *cluster* memiliki perbedaan dalam membentuk kelompok. Pada analisis *cluster* hirarki dengan memulai menggabungkan dua atau lebih objek yang memiliki kesamaan dekat hingga *cluster* akan membentuk pohon yang terdapat tingkatan jelas antar objek seprti dari objek yang mempunyai kesamaan dekat yang paling dekat. Sementara itu, metode analisis non hirarki dengan menentukan terlebih dahulu jumlah *cluster* yang akan diinginkan, kemudian membagi objek ke dalam *cluster* tersebut (Lawuna, Zega dan Mendrofa, 2024). Salah satu metode hirarki yang dapat digunakan dalam mengelompokkan kasus kriminalitas adalah metode *Ward*.

Metode *Ward* adalah metode pengelompokan yang menggunakan pendekatan analisis varians untuk menghitung jarak antar *cluster* dengan meminimumkan jumlah kuadrat. Tujuan dari metode ini adalah meminimalkan varians di dalam *cluster*, artinya objek-objek dalam satu *cluster* memiliki varians yang relatif kecil, sehingga mereka semakin mirip *cluster* satu sama lain. Hal ini akan meningkatkan homogenitas dalam setiap *cluster* (Ramadhani, Purnamasari dan Amijaya, 2018). Terdapat beberapa metode dalam menentukan *cluster* optimal seperti *Calinski – Harabasz Index (CH Index), Gamma Index,* dan *Silhouette Index* (Doi, Rusgiyono dan Wuryandari, 2023). Pada penelitian ini, digunakan *Gamma Index* sebagai menentuan *cluster* optimal.

Gamma Index adalah metode yang digunakan sebagai acuan untuk menentukan jumlah *cluster* optimal. Selain itu, *Gamma Index* juga membantu menilai sejauh mana metode *clustering* berhasil mengelompokkan data berdasarkan kemiripan atau jarak antar data (Baker dan Hubert, 1975). Metode ini memiliki perhitungan yang lebih kompleks karena memperhitungkan seluruh pasangan dalam pengelompokan. Setiap pasangan data dibandingkan untuk menentukan termasuk kedalam kategori *concordant* atau *discordant*, yang membuat perhitungan meningkat seiring bertambahnya jumlah data. antara dua urutan dengan cara menghitung setiap pasangan. Oleh karena itu, dalam penelitian ini dilakukan

pengelompokan provinsi berdasarkan kasus kriminalitas di Indonesia menggunakan metode *Ward* dengan optimalisasi *Gamma Index*.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana mengelompokkan provinsi di Indonesia berdasarkan kasus kriminalitas pada tahun 2022 menggunakan metode *Ward*?
2. Bagaimana menentukan jumlah *cluster* optimal dalam mengelompokkan provinsi di Indonesia berdasarkan kasus kriminalitas pada tahun 2022 dengan *Gamma Index*?

1.3 Tujuan

Dari latar belakang dan rumusan masalah yang telah disampaikan, maka tujuan penelitian ini adalah:

1. Mengelompokkan provinsi di Indonesia berdasarkan kasus kriminalitas pada tahun 2022 menggunakan metode *Ward*.
2. Menentukan jumlah *cluster* optimal dalam mengelompokkan provinsi di Indonesia berdasarkan kasus kriminalitas pada tahun 2022 dengan *Gamma Index.*

1.4 Batasan Masalah

Batasan dalam penelitian ini antara lain:

1. Penelitian ini hanya mencakup data jumlah kriminalitas di setiap provinsi Indonesia menurut jenis kejahatan pada tahun 2022.
2. Variabel yang digunakan pada penelitian ini yaitu kejahatan terhadap nyawa; kejahatan terhadap fisik/badan; kejahatan kesesuliaan; kejahatan terhadap kemerdekaan orang; kejahatan terhadap hak milik/barang dengan penggunaan kekerasan; kejahatan terhadap hak milik/barang; kejahatan terkait narkotika; kejahatan terkait penipuan; penggelapan, dan korupsi; dan kejahatan ketertiban umum.
3. *Cluster* optimal yang akan terbentuk berjumlah 2, 3, 4, dan 5 *cluster.*

1.5 Tinjauan Pustaka

Penelitian Doi, Rusgiyono, dan Wuryandari pada tahun 2023 yang bertujuan untuk mengelompokkan daerah 3T di Indonesia berdasarkan Indeks Pembangunan Manusia (IPM) menggunakan *K-Medoid*. Besar kecilnya kemiripan objek dihitung dengan menggunakan jarak *euclidean* dan jarak manhattan. Untuk pemilihan jumlah *cluster* yang optimal digunakan validasi *cluster* internal, seperti *Calinski – Harabasz Index* (CH *Index*), *Gamma Index*, dan *Silhouette Index*. Hasil penelitian ini menunjukkan bahwa *cluster* optimal berjumlah sebesar 4 *cluster* dengan menggunakan pengukuran jarak *euclidean*, memiliki nilai CH *Index* sebesar 37,15764, *Gamma Index* sebesar 0,7821181, dan *Silhouette Index* sebesar 0,3354435.

Penelitian yang dilakukan Lestari, Utami, dan Haris pada tahun 2018 bertujuan untuk membandingkan pengelompokan provinsi di Indonesia berdasarkan kriminalitas dengan metode *Ward* dan metode *K-Medoids*. Hasil penelitian menujukkan bahwa nilai rasio simpangan baku dalam *cluster* dan antar *cluster* pada metode *Ward* lebih kecil jika dibandingkan metode *K-Medoids* sebesar 0,2764. Oleh karena itu, metode *Ward* dianggap sebagai metode terbaik untuk tujuan pengelompokan.

Saidah, Santoso dan Widiharih pada tahun 2022 melakukan penelitian untuk mengelompokkan provinsi di Indonesia berdasarkan indikator kesehatan lingkungan dengan metode *K-Medoids* menggunakan lima validasi indeks internal, yaitu validasi *Calinski-Harabasz Index*, *Baker-Hubert Index (Gamma Index), Silhouette Index, C-Index,* dan *Davies-Bouldin Index* untuk jumlah *cluster* optimal. Hasil penelitian ini terdapat dua *cluster* dengan pengukuran jarak manhattan, karena memiliki nilai *Calinski-Harabasz Index*, *Baker-Hubert Index (Gamma Index)*, dan *Silhouette Index* yang terbesar, serta nilai *C-index* dan *Davies-Bouldin Index* yang terkecil jika dibandingkan dengan jumlah *k* yang lain. *Cluster* satu terdiri dari 10 provinsi, *cluster* dua terdiri dari 24 provinsi.

Hapsari dan Widodo (2017) melakukan penelitian untuk pengelompokan daerah rawan kriminalitas dengan menggunakan analisis *K-Means clustering* dan pemetaan. Variabel yang digunakan merupakan jenis-jenis kriminalitas, yaitu

kejahatan terhadap nyawa, kejahatan terhadap fisik atau badan, kejahatan terhadap kesusilaan, kejahatan terhadap kemerdekaan orang, kejahatan terhadap kemerdekaan orang, kejahatan terhadap hak miliki atau barang, kejahatan terkait narkotika, kejahatan terkait penipuan, penggelapan, dan korupsi, dan kejahatan terhadap ketertiban umum. Penelitian memperoleh hasil bahwa terdapat empat kelompok, dimana kelompok I masuk ke dalam kategori daerah yang cukup aman, kelompok II masuk ke dalam kategori daerah rawan, kelompok III masuk ke dalam kategori daerah cukup rawan, dan kelompok IV masuk ke dalam kategori aman.

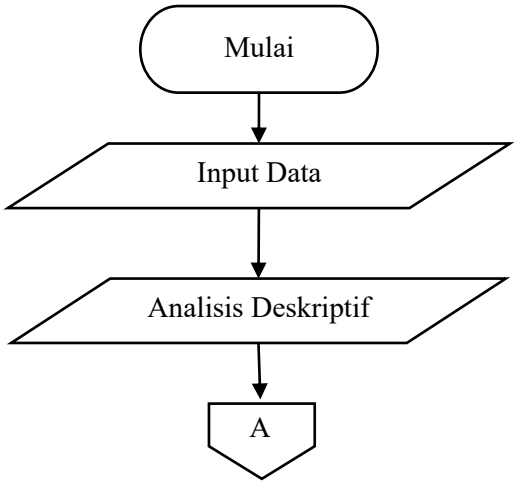
1.6 Metodologi Penelitian

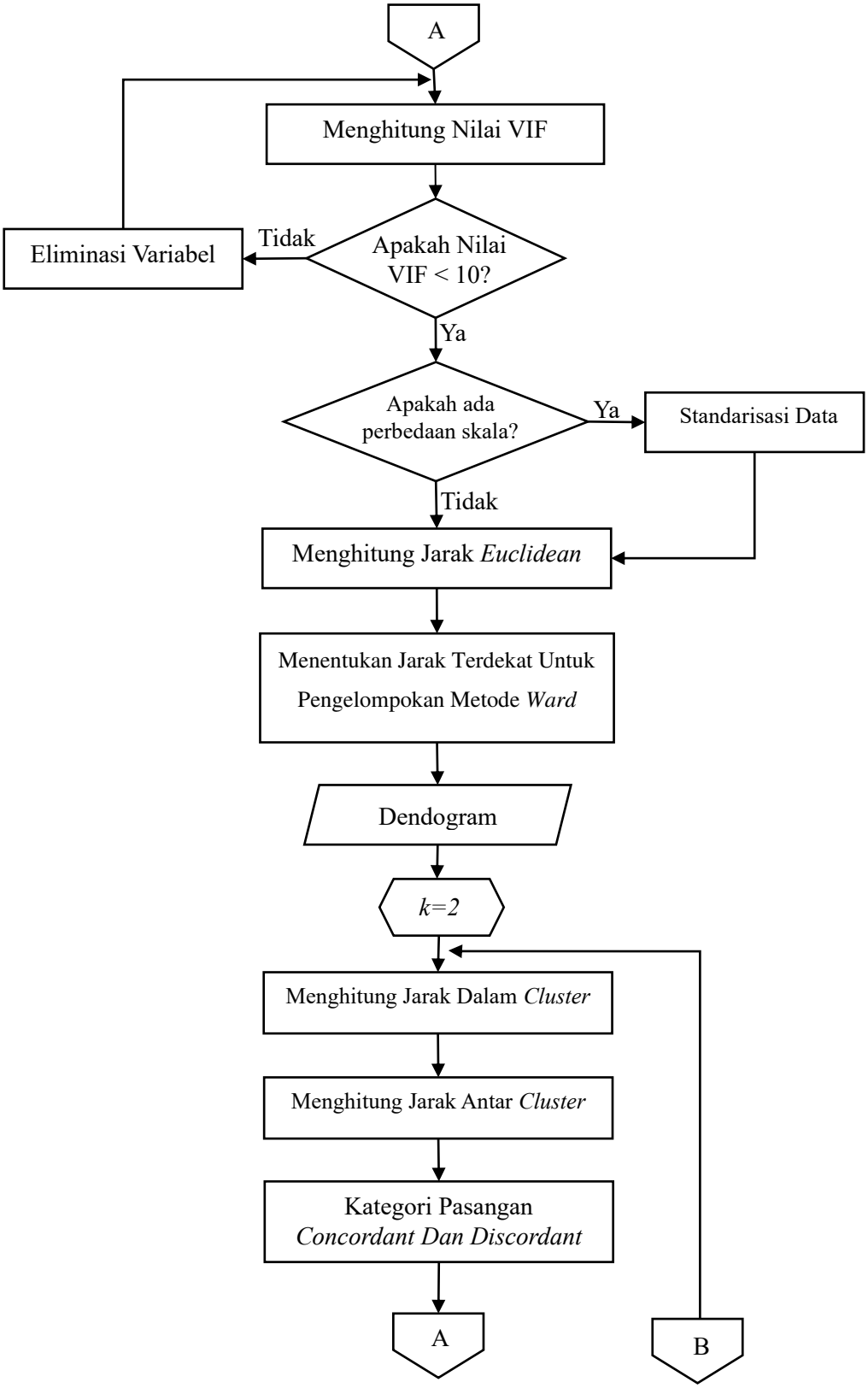
Penelitian ini menggunakan data sekunder yang diperoleh dari website Badan Pusat Statistik (BPS) Indonesia dalam buku Statistik Kriminal 2023. Metode yang digunakan yaitu metode *Ward* dengan optimalisasi *Gamma Index***.** Penelitian ini dilakukan dengan memberikan gambaran yang jelas mengenai langkah-langkah yang diperlukan selama proses penelitian ini.

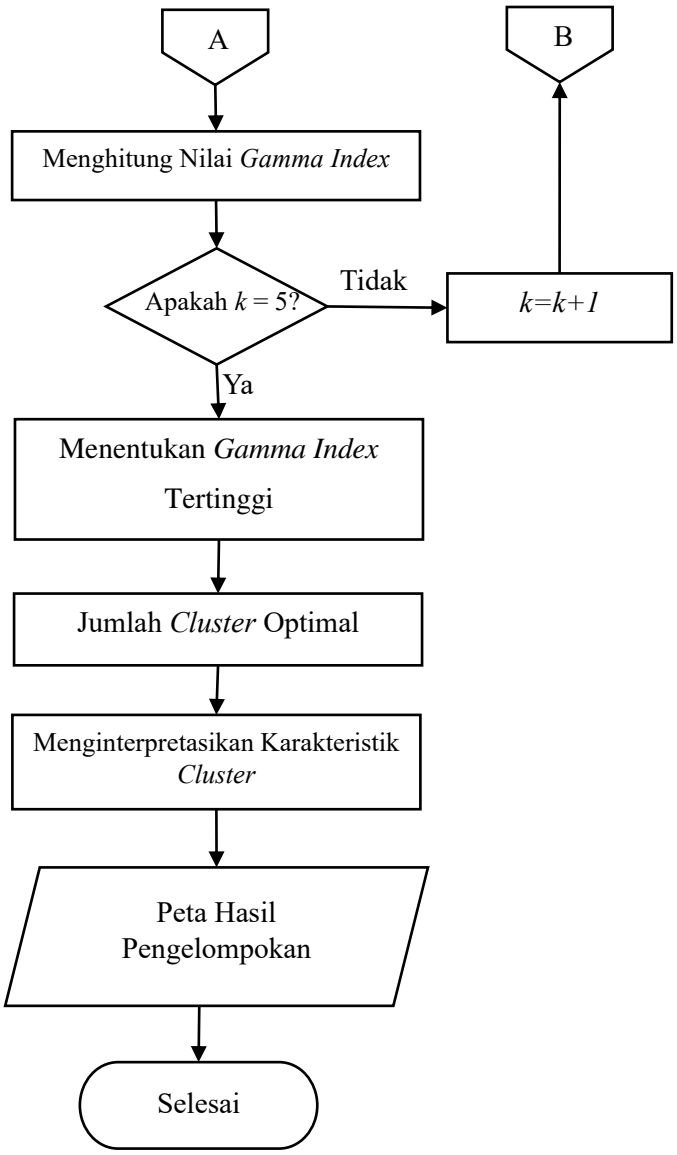
Tahapan pertama dalam metode *Ward* dengan optimalisasi *Gamma Index* adalah analisis deskriptif terhadap data yang digunakan. Selanjutnya, dilakukan asumsi analisis *cluster.* Dalam asumsi analisis *cluster* terbagi mencari dua, yaitu uji sampel representatif dan uji multikolinearitas. Uji sampel representatif tidak dilakukan dalam penelitian ini karena seluruh data mencakup 34 provinsi yang dijadikan sebagai objek analisis. Kemudian, dilakukan uji multikolinearitas untuk melihat apakah korelasi antar variabel. Apabila nilai VIF pada semua variabel kurang dari 10, maka tidak terjadi multikolinearitas dan analisis dapat berlanjut ke tahap selanjutnya. Sebaliknya, jika ditemukan variabel dengan nilai VIF lebih dari 10, maka dilakukan eliminasi variabel tersebut sebelum melanjutkan proses analisis ke tahap berikutnya. Standarisasi data perlu dilakukan karena data memiliki perbedaan skala yang besar. Kemudian dihitung jarak *euclidean* untuk mengukur kemiripan antar objek. Setelah itu, menentukan jarak terdekat dari hasil jarak *euclidean* untuk pengelompokan menggunakan metode *Ward* dan hasilnya disajikan dalam bentuk dendogram.

Selanjutnya, untuk *cluster* berjumlah dua, dilakukan perhitungan jarak dalam *cluster* yang sama dan jarak antar *cluster* yang berbeda. Setelah memperoleh nilai jarak dalam *cluster* yang sama dan jarak antar *cluster* yang berbeda, maka dilakukan kategori sebagai *concordant* dan *discordant*. Setelah jumlah pasangan *concordant* dan *discordant* diperoleh, dilakukan perhitungan *Gamma Index*. Jika perhitungan sebelumnya dilakukan untuk *cluster* yang berjumlahnya kurang dari lima, maka tahapan perhitungan jarak dalam *cluster* dan jarak antar *cluster* sampai nilai *Gamma Index* dilakukan kembali untuk *cluster* yang berjumlah *k+1*. Namun, jika perhitungan sebelumnya dilakukan untuk *cluster* yang jumlahnya lima. Langkah selanjutnya adalah mencari nilai *Gamma Index* tertinggi dari semua nilai *Gamma Index* yang telah diperoleh pada jumlah *cluster* sebelumnya. Setelah jumlah *cluster* optimal ditentukan berdasarkan *Gamma Index* tertinggi, langkah terakhir adalah menginterpretasikan karakteristik setiap *cluster* yang terbentuk serta membuat visualisasi hasil pengelompokan dalam bentuk peta.

Langkah-langkah analisis dalam penelitian ini dapat dilihat pada diagram alir yang ditampilkan pada Gambar 1.1




A
Menghitung Nilai VIF
Apakah Nilai VIF < 10?
Tidak
Eliminasi Variabel
Ya
Apakah ada perbedaan skala?
Ya
Standarisasi Data
Tidak
Menghitung Jarak *Euclidean*
Menentukan Jarak Terdekat Untuk Pengelompokan Metode *Ward*
Dendogram
k=2
Menghitung Jarak Dalam *Cluster*
Menghitung Jarak Antar *Cluster*
Kategori Pasangan *Concordant Dan Discordant*
A
B



Gambar 1.1 *Flowchart* Penentuan Jumlah *Cluster* Optimal