PENGELOMPOKAN INDEKS PEMBANGUNAN MANUSIA DI INDONESIA MENGGUNAKAN METODE *K-MEDOIDS* DENGAN EVALUASI *DAVIES BOULDIN INDEX*

AISNA ERSAWAHYUNI
NIM. H1091211008

SKRIPSI

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

PENGELOMPOKAN INDEKS PEMBANGUNAN MANUSIA DI INDONESIA MENGGUNAKAN METODE *K-MEDOIDS* DENGAN EVALUASI *DAVIES BOULDIN INDEX*

AISNA ERSAWAHYUNI
NIM. H1091211008

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 Indeks Pembangunan Manusia di Indonesia Menggunakan Metode *K-Medoids* dengan Evaluasi *Davies Bouldin Index*
Nama Mahasiswa	: Aisna Ersawahyuni
NIM	: H1091211008
Jurusan/Program Studi	: Matematika/Statistika
Tanggal Lulus	: 24 Maret 2025
SK Pembimbing	: No. 2587/UN22.8/TD.06/2024/Tanggal 2 September 2024
SK Penguji	: No. 822/UN22.8/TD.06/2025/Tanggal 11 Maret 2025

Dosen Pembimbing

Pembimbing I	Pembimbing II
Shantika Martha, M.Si.	Hendra Perdana, M.Sc.
NIP. 198403082008122003	NIP. 198810102019031020

Dosen Penguji

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

Pimpinan Sidang (merangkap anggota penguji)	Sekretaris Sidang (merangkap anggota penguji)
Shantika Martha, M.Si.	Hendra Perdana, M.Sc.
NIP. 198403082008122003	NIP. 198810102019031020

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	: Aisna Ersawahyuni
NIM	: H1091211008
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, 25 Maret 2025

Aisna Ersawahyuni
NIM. H1091211008

Pengelompokan Indeks Pembangunan Manusia di Indonesia Menggunakan Metode *K-Medoids* dengan Evaluasi *Davies Bouldin Index*

Abstrak

Indeks Pembangunan Manusia (IPM) adalah salah satu indikator penting dalam menilai kemajuan suatu negara. Pengelompokan IPM berguna memberikan wawasan mengenai pemahaman tentang pola dan karakteristik perkembangan kesejahteraan masyarakat. Tujuan penelitian ini adalah mengelompokkan provinsi-provinsi di Indonesia berdasarkan IPM masing-masing provinsi tahun 2023 menggunakan metode *K-Medoids* dan menentukan jumlah *cluster* yang optimal dengan evaluasi *Davies Bouldin Index* (DBI). Data yang digunakan berasal dari Publikasi BPS Pusat Tahun 2023, variabel yang digunakan yaitu umur harapan hidup, harapan lama sekolah, rata-rata lama sekolah, pengeluaran perkapita, PDRB, jumlah penduduk, dan persentase penduduk miskin. Proses analisis dimulai dengan statistik deskriptif dan pengujian multikolinearitas *Variance Inflation Factor* (VIF). Variabel dengan nilai *VIF > 10* dieliminasi secara bertahap. Selanjutnya, standarisasi data dan menentukan jumlah *cluster* awal, k yang digunakan $(k = 2, 3, 4, 5, dan\ 6)$, dilanjutkan pemilihan *medoids* acak dan perhitungan jarak *euclidean* untuk menentukan kedekatan antar objek. Proses dilakukan hingga total jarak $(S) \geq 0$. Setelah pengelompokan selesai dilakukan evaluasi jumlah *cluster* dengan menggunakan DBI, dengan cara meminimalkan nilai *Sum of Square Within* (SSW) dan memaksimalkan *Sum of Square Between* (SSB) dilanjutkan dengan menghitung rasio antara SSW dan SSB sehingga diperoleh nilai DBI. Hasil pengelompokan optimal yaitu dua *cluster* berdasarkan nilai DBI sebesar 0,94, namun dipilih *cluster* optimal empat *cluster* sebagai alternatif karena memiliki nilai DBI terendah kedua sebesar 1,07. Hasil pengelompokan empat *cluster* yaitu *cluster* 1 IPM kategori sangat tinggi berjumlah empat provinsi, *cluster* 2 IPM kategori tinggi berjumlah 14 provinsi, *cluster* 3 IPM kategori rendah berjumlah 10 provinsi, dan *cluster* 4 IPM kategori sangat rendah berjumlah enam provinsi.

Kata kunci: Jarak *Euclidean*, DBI, VIF, *Hierachical*.

Grouping of Human Development Index in Indonesia Using K-Medoids Method with Davies Bouldin Index Evaluation

Abstract

The Human Development Index (HDI) is one of the important indicators in assessing the progress of a country. The HDI grouping is useful for providing insight into understanding the patterns and characteristics of the development of community welfare. The purpose of this study is to group provinces in Indonesia based on the HDI of each province in 2023 using the K-Medoids method and determine the optimal number of clusters by evaluating the Davies Bouldin Index (DBI). The data used comes from the 2023 Central BPS Publication, the variables used are life expectancy, expected years of schooling, average years of schooling, per capita expenditure, GRDP, population, and percentage of poor people. The analysis process begins with descriptive statistics and Variance Inflation Factor (VIF) multicollinearity testing. Variables with VIF values > 10 are eliminated gradually. Furthermore, data standardization and determining the initial number of clusters, *k* used *(k = 2,3,4,5, and 6)*, followed by random medoids selection and calculation of Euclidean distance to determine the proximity between objects. The process is carried out until the total distance $(S) \geq 0$. After the clustering is complete, an evaluation of the number of clusters is carried out using DBI, by minimizing the Sum of Square Within (SSW) value and maximizing the Sum of Square Between (SSB) followed by calculating the ratio between SSW and SSB to obtain the DBI value. The optimal clustering result is two clusters based on the DBI value of 0.94, but the optimal cluster of four clusters was chosen as an alternative because it has the second lowest DBI value of 1.07. The results of the 4 cluster grouping are cluster 1 HDI category very high totaling four provinces, cluster 2 HDI category high totaling 14 provinces, cluster 3 HDI category low totaling 10 provinces, and cluster 4 HDI category very low totaling six provinces.

Keywords: *Euclidean Distance, DBI, VIF, Hierarchical.*

PRAKATA

Puji syukur kehadirat Allah SWT karena atas limpahan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan penulisan skripsi ini. Skripsi ini berjudul **"Pengelompokan Indeks Pembangunan Manusia di Indonesia Menggunakan Metode *K-Medoids* dengan Evaluasi *Davies Bouldin Index*"**. 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.

Dengan selesainya penulisan skripsi ini, penulis menyadari bahwa selama proses penulisan skripsi ini tidak lepas dari bimbingan, arahan, dan dukungan dari berbagai pihak. Maka dari itu penulis ingin mengucapkan terima kasih kepada:

1. Allah SWT yang telah melimpahkan rahmat, berkah, dan karunia-Nya memberikan kelancaran dalam penyusunan skripsi sehingga skripsi ini dapat diselesaikan sesuai dengan waktu yang diharapkan.
2. Kedua orang tua penulis yang tercinta, Bapak Yuniansyah dan Ibu Nani Oktaleni yang senantiasa memberikan doa, semangat, dan motivasi serta dukungan moril dan material kepada penulis selama proses perkuliahan dan dalam menyelesaikan skripsi ini.
3. Ibu Shantika Martha, M.Si. selaku dosen pembimbing pertama yang telah banyak memberikan bimbingan serta arahan selama penulisan skripsi ini.
4. Bapak Hendra Perdana, M.Sc. selaku dosen pembimbing kedua yang telah banyak memberikan bimbingan serta arahan selama penulisan skripsi ini.
5. Bapak Ir. Dadan Kusnandar, M.Sc., Ph.D. dan Ibu Dr. Evy Sulistianingsih, M.Sc. selaku Dosen Pembimbing Akademik yang telah meluangkan waktu, memberikan saran, masukan, dan motivasi kepada penulis selama proses perkuliahan.
6. Ibu Nurfitri Imro'ah, M.Si. selaku dosen penguji pertama yang telah memberikan masukan dan saran kepada penulis.

7. Ibu Dr. Evy Sulistianingsih, M.Sc. selaku dosen penguji kedua yang telah memberikan masukan dan saran kepada penulis.
8. Adik yang penulis sayangi, Erick Alsa Ansera yang selalu memberikan semangat dan dukungannya selama ini.
9. Teman seperjuangan penulis yang selalu menemani selama proses perkuliahan dan menyelesaikan skripsi ini yaitu Eri Sindia, Rizha Aynul Fiqriani, Nabila Izza Harnanta, dan Alsa Muarti yang telah membantu dan selalu memberikan dukungan, semangat, dan motivasi kepada penulis.
10. Teman-teman Statistika 2021 yang telah menemani dan memberikan motivasi serta semangat selama penulisan skripsi ini.
11. 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 atas segala bantuan yang diberikan dalam penulisan skripsi ini. Penulis berharap skripsi ini dapat bermanfaat bagi semua pihak. Penulis menyadari bahwa skripsi ini jauh dari kata sempurna, oleh karena itu, kritik dan saran yang bersifat membangun sangat diharapkan oleh penulis untuk menjadikan skripsi ini lebih baik.

Pontianak, 25 Maret 2025

Aisna Ersawahyuni

DAFTAR ISI

LEMBAR PENGESAHAN ii
PERNYATAAN INTEGRITAS AKADEMIK iii
Abstrak iv
Abstract v
PRAKATA vi
DAFTAR ISI viii
DAFTAR TABEL x
DAFTAR GAMBAR 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 Pembangunan Manusia 8
2.2 Analisis Multivariat 9
2.3 Analisis *Cluster* 9
2.3.1 Metode *Hierarchical* 10
2.3.2 Metode *non-hierarchical* 11
2.4 Uji Multikolinearitas 12
2.5 Standarisasi *Z-Score* 12
2.6 Jarak *Euclidean* 13
BAB III ALGORITMA *K-MEDOIDS* 14
3.1 Algoritma *K-Medoids* 14
3.2 *Davies Bouldin Index* 15

BAB IV HASIL DAN PEMBAHASAN 17
4.1 Deskripsi Data 17
4.2 Pendeteksian Multikolinearitas 18
4.3 Standarisasi Data 19
4.4 *Clustering* dengan *K-Medoids* 19
4.5 Penentuan Banyak *Cluster* dengan *Davies Bouldin Index* 22
4.6 Interpretasi Hasil *Clustering* 27
BAB V KESIMPULAN 31
DAFTAR PUSTAKA 32
LAMPIRAN 35

DAFTAR TABEL

Tabel 4.1 Variabel Penelitian..... 17
Tabel 4.2 Statistik Deskriptif..... 18
Tabel 4.3 Nilai *Variance Inflation Factor*..... 18
Tabel 4.4 Nilai *Davies Bouldin Index*..... 25
Tabel 4.5 Pengelompokan Provinsi di Indonesia pada Dua *Cluster*..... 26
Tabel 4.6 Karakteristik Pengelompokan *K-Medoids Clustering*..... 27

DAFTAR GAMBAR

Gambar 1.1 *Flowchart* Penelitian.....7
Gambar 4.1 Peta Hasil Pengelompokan dengan K-Medoids…….26

DAFTAR LAMPIRAN

Lampiran 1 Data Indikator IPM Provinsi di Indonesia Tahun 2023..... 35
Lampiran 2 Standarisasi Data..... 36

DAFTAR SIMBOL

n	: banyaknya objek yang diamati
$Z_{i,l}$	: nilai *z-score* objek ke-i pada variabel ke-l
p	: banyaknya variabel yang diamati
R_l^2	: koefisien determinasi antara X_l dengan variabel independen lainnya
VIF_l	: *Variance Inflation Factor* (VIF) untuk variabel independen ke-l, dengan $l = 1, 2, \ldots, p$
X_l	: nilai variabel ke-l, dengan $l = 1, 2, \ldots, p$
p	: banyaknya variabel yang diamati
μ_l	: nilai rata-rata dari variabel ke-l, dengan $l = 1, 2, \ldots, p$
σ_l	: simpangan baku dari variabel ke-l, dengan $l = 1, 2, \ldots, p$
d_{ij}	: jarak *euclidean* antara objek ke-i dan objek ke-j, dengan dengan $i, j = 1, 2, \ldots, n$
x_{il}	: nilai objek ke-i pada variabel ke-l, dengan $i = 1, 2, \ldots, n$ dan $l = 1, 2, \ldots, p$
x_{jl}	: nilai objek ke-j pada variabel ke-l, dengan $j = 1, 2, \ldots, n$ dan $l = 1, 2, \ldots, p$
S	: total jarak
k	: banyaknya *cluster* yang digunakan
m_j	: banyaknya objek pada *cluster* ke-j, dengan $j = 1, 2, \ldots, k$
$d(x_i, c_j)$	: jarak objek ke-i ke *centroid* yang berada dalam *cluster* ke-j
$d(c_i, c_j)$	: jarak *centroid cluster* ke-i ke *centroid cluster* ke-j
SSW_j	: nilai *Sum of Square Within* (SSW) dalam *cluster* ke-j
$SSB_{i,j}$	: nilai *Sum of Square Between* (SSB) antara *cluster* ke-i dengan *cluster* ke-j
$R_{i,j}$	: nilai rasio antara *cluster* ke-i dengan *cluster* ke-j
DBI	: nilai *Davies Bouldin Index*

BAB I
PENDAHULUAN

1.1 Latar Belakang

Seiring dengan berjalannya waktu, dunia semakin berkembang pesat dan penuh dengan teknologi, sehingga untuk mencari dan menerima informasi semakin mudah. Berkembangnya teknologi sekarang ini dan keterbukaan dalam mendapatkan informasi yang difasilitasi dengan internet mendorong manusia agar bisa memanfaatkan hal tersebut untuk berkembang. Perkembangan manusia membutuhkan waktu untuk berproses tentang bagaimana seseorang bisa memperoleh lebih banyak kemampuan dan menikmati kesempatan tersebut. Pembangunan nasional adalah salah satu upaya untuk meningkatkan ketentraman dan kesejahteraan masyarakat. Artinya, masyarakat dapat berperan sebagai pelaku dan pelaksana pembangunan (Siswati dan Hermawati, 2018).

Pembangunan nasional merupakan upaya yang dilakukan secara berkelanjutan untuk meningkatkan kualitas hidup masyarakat, yang mencakup berbagai aspek seperti kesehatan, pendidikan, serta kesejahteraan ekonomi. Indeks Pembangunan Manusia (IPM) adalah salah satu indikator penting dalam mengukur kemajuan suatu negara. IPM pertama kali dikenalkan pada tahun 1990 oleh *United Nations Development Programme* (UNDP). Indeks Pembangunan Manusia (IPM) terdiri dari tiga aspek utama yang berkaitan dengan kesejahteraan manusia yaitu aspek umur panjang dan hidup sehat, tingkat pendidikan, dan standar hidup layak yang diukur melalui pendapatan nasional bruto per kapita (Yektiningsih, 2018).

Meskipun IPM di Indonesia terus mengalami peningkatan sejak tahun 2020 hingga 2023, masih ada provinsi seperti Papua yang memiliki IPM terendah, yaitu sebesar 63,01, jauh lebih rendah dibandingkan dengan DKI Jakarta yang mencapai 83,5 (Badan Pusat Statistik, 2024). Papua dan Papua Barat juga mencatat angka kemiskinan yang tinggi, masing-masing sebesar 26,03 dan 20,49, angka ini cukup tinggi dibandingkan dengan provinsi-provinsi lainnya. Hal ini

menjadi perhatian utama bagi pemerintah dalam upaya memperbaiki pembangunan manusia di wilayah tersebut. Oleh karena itu, pengelompokan provinsi di Indonesia berdasarkan IPM sangat penting agar pemerintah dapat merumuskan kebijakan yang tepat dan sesuai. (Badan Pusat Statistik, 2024).

Pengelompokan Indeks Pembangunan Manusia di suatu negara dapat memberikan wawasan yang penting terhadap pemahaman dengan pola dan karakteristik perkembangan kesejahteraan masyarakat di berbagai wilayah. Upaya yang dapat dilakukan untuk pengelompokan IPM provinsi yang ada di Indonesia adalah dengan analisis *cluster*. Analisis *cluster* merupakan analisis multivariat yang memiliki tujuan untuk mengelompokkan objek berdasarkan dengan karakteristik yang dimiliki (Luthfi dan Wijayanto, 2021). Salah satu metode *cluster* yang digunakan dalam mengelompokkan Indeks Pembangunan Manusia (IPM) adalah metode *K-Medoids*.

Algoritma *K-Medoids* atau dikenal dengan Algoritma *Partitioning Around Medoid* (PAM) merupakan sebuah algoritma yang menggunakan metode partisi *clustering* untuk mengelompokkan sekumpulan *n* objek menjadi sejumlah *k cluster* (Prakasawati, Chrisnanto, dan Hadiana, 2019). Algoritma *K-Medoids* merupakan salah satu metode yang bisa membantu untuk mengatasi kelemahan dari *K-Means* yang sensitif terhadap pencilan atau *outlier* karena objek-objek tersebut letak dan karakteristiknya sangat berbeda dari yang lainnya (Wira, Budianto, dan Wiguna 2019). Untuk menentukan *cluster* optimal pada metode *K-Medoids* dapat menggunakan evaluasi, seperti *Silhouette*, *Dunn Index*, *Elbow*, dan *Davies Bouldin Index* (DBI). Pada penelitian ini digunakan evaluasi *Davies Bouldin Index* (DBI) untuk menentukan *cluster* yang paling optimal.

Davies Bouldin Index (DBI) digunakan untuk mengevaluasi algoritma klastering dengan skema evaluasi internal dan berfungsi untuk memvalidasi seberapa baik klastering yang telah dilakukan (Rinaldi, Surlanto, Sudrajat, dan Kurnia, 2019). Semakin kecil nilai *Davies Bouldin Index* menunjukkan hasil *cluster* yang optimal (Dewi dan Pramita, 2019). Oleh karena itu, pada penelitian ini dilakukan pengelompokan provinsi di Indonesia berdasarkan Indeks

Pembangunan Manusia (IPM) dengan menggunakan metode *K-Medoids cluster* dengan optimasi *Davies Bouldin Index*.

1.2 Rumusan Masalah

Berdasarkan latar belakang, rumusan masalah dalam penelitian ini, yaitu:

1. Bagaimana mengelompokkan provinsi di Indonesia berdasarkan Indeks Pembangunan Manusia (IPM) tahun 2023 dengan algoritma *K-Medoids*?
2. Bagaimana menentukan banyaknya *cluster* optimal pada pengelompokan provinsi di Indonesia berdasarkan Indeks Pembangunan Manusia (IPM) tahun 2023 dengan menggunakan evaluasi *Davies Bouldin Index*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dibuat, maka tujuan penelitian yang ingin dicapai dalam penelitian ini, yaitu:

1. Mengelompokkan provinsi di Indonesia berdasarkan Indeks Pembangunan Manusia (IPM) tahun 2023 dengan algoritma *K-Medoids.*
2. Menentukan banyaknya *cluster* optimal pada pengelompokan provinsi di Indonesia berdasarkan Indeks Pembangunan Manusia (IPM) dengan menggunakan evaluasi *Davies Bouldin Index.*

1.4 Batasan Masalah

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

1. Data yang digunakan merupakan data Indeks Pembangunan Manusia (IPM) 2023 yang diperoleh dari publikasi Badan Pusat Statistik Indonesia.

2. Terdapat tujuh variabel yang digunakan yaitu Umur Harapan Hidup (UHH), Harapan Lama Sekolah (HLS), Rata-rata Lama Sekolah (RLS), Pengeluaran per Kapita, Produk Domestik Regional Bruto (PDRB), Jumlah Penduduk, dan Persentase Penduduk Miskin.
3. Pembentukan *cluster* optimal maksimal berjumlah 6.
4. Perhitungan jarak menggunakan jarak *Euclidean*.

1.5 Tinjauan Pustaka

Dalam penelitian Septiani, Fauzan, dan Huda (2022), dilakukan pengelompokan harapan hidup pasca operasi pada pasien penderita kanker paru-paru menggunakan *K-Medoids* dan evaluasi *Davies Bouldin Index*. Pengelompokan menggunakan *K-Medoids* terbagi menjadi dua *cluster* yaitu meninggal dalam kurun waktu satu tahun dan mampu bertahan hidup. Hasil dari penelitian ini diperoleh nilai *Davies Bouldin Index* yaitu 0,93543.

Penelitian yang dilakukan oleh Sapaat, Lapian, dan Tumangkeng (2020) tentang analisis faktor-faktor yang mempengaruhi IPM di Provinsi Sulawesi Utara tahun 2005-2019. Hasil dari penelitian menunjukkan bahwa PDRB, jumlah penduduk, dan tingkat kemiskinan berpengaruh signifikan terhadap IPM. Namun secara parsial tingkat kemiskinan tidak berpengaruh signifikan terhadap IPM di Sulawesi Utara.

Luthfi dan Wijayanto (2021) melakukan analisis perbandingan antara tiga metode dalam *clustering* yaitu *Hierarchical*, *K-Means*, dan *K-Medoids*. Hasil dari ketiga metode tersebut dibandingkan untuk memperoleh metode yang terbaik dengan melihat ukuran validitas dari nilai *Dunn Index* (*DN*), *Davies Bouldin Index* (*DBI*), *Calinski-Harabasz Index* (*CH*). Hasil analisisnya menunjukkan bahwa model terbaik yang diperoleh yaitu menggunakan *K-Medoids* lebih baik dilihat dari perbandingan nilai rasio simpangan baku.

Implementasi *K-Medoids Clustering* untuk mengetahui pola pemilihan program studi mahasiswa baru tahun 2018 di Universitas Kanjuruhan Malang pada penelitian Wira, Budianto, dan Wiguna (2019). Hasil pengujian *clustering K-Medoids* dengan validasi *Silhouette Coefficient* diperoleh hasil nilai *Silhouette Coefficient* terbaik sebesar 0.690754 dengan jumlah *cluster* sebanyak tiga.

Pada penelitian Okta, Satyahadewi, dan Debataraja (2019) dilakukan penerapan metode *K-Medoids* pada pengelompokan daerah penghasil kelapa sawit dengan validasi *Silhouette*. Hasil dari analisis terbentuk tiga *cluster*, yaitu *cluster* pertama kategori rendah terdapat tujuh kabupaten/kota, *cluster* kedua kategori sedang terdapat empat kabupaten/kota, dan *cluster* ketiga kategori tinggi terdapat dua kabupaten/kota. Nilai koefisien *Silhouette* masing-masing objek yaitu objek ke-1 sebesar 0,55884, objek ke-2 sebesar 0,42447, sampai dengan objek ke-13 sebesar 0,69756.

1.6 Metodologi Penelitian

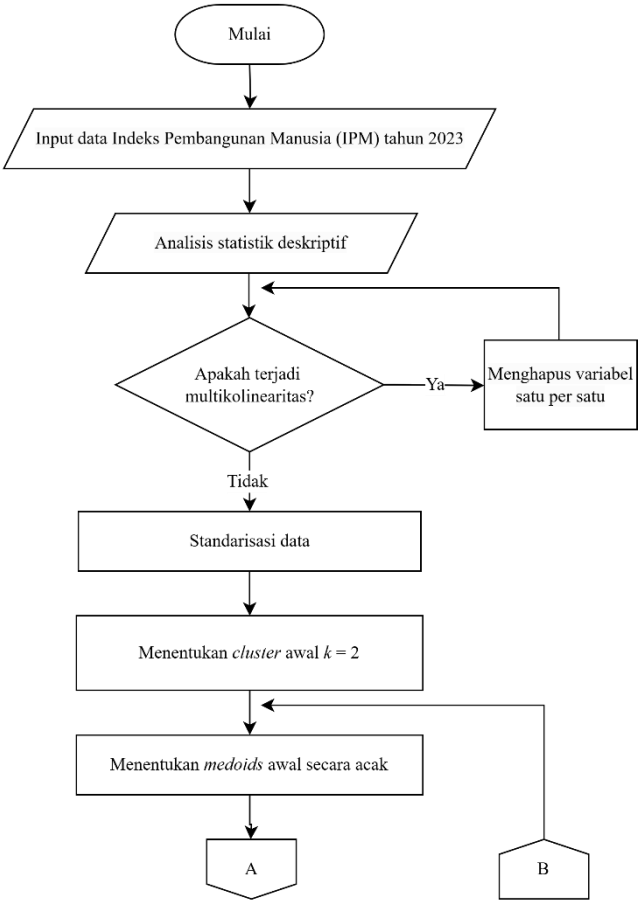
Penelitian ini dimulai dengan studi kasus yang menggunakan data sekunder yang didapatkan dari *website* Badan Pusat Statistik (BPS) Indonesia. Variabel penelitian ini terdiri dari empat indikator pembentuk IPM dan tiga indikator tambahan yang diteliti oleh peneliti lainnya memiliki pengaruh terhadap IPM. Metode yang digunakan adalah *K-Medoids cluster* dengan optimasi *Davies Bouldin Index* (DBI).

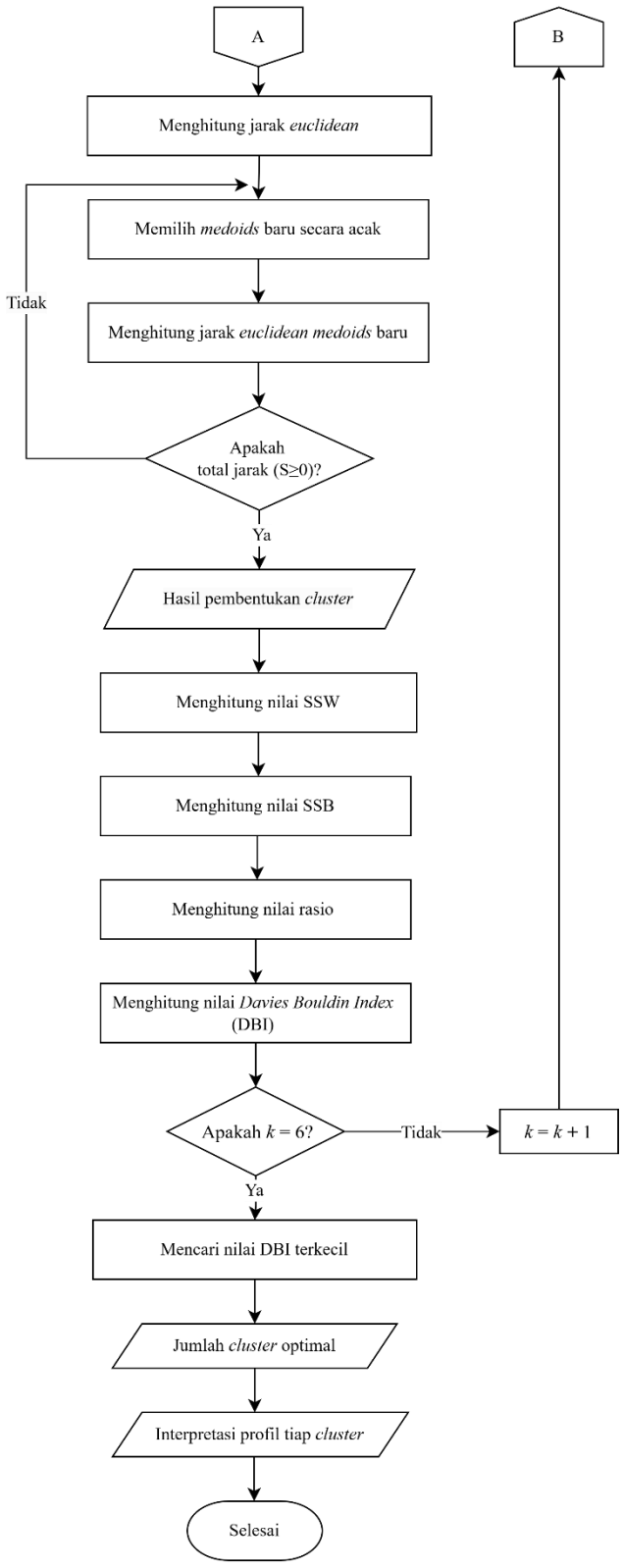
Proses analisis dilakukan dengan alat bantu *software* RStudio dan *Microsoft excel*. Langkah pertama yang dilakukan adalah analisis statistik deskriptif. Kemudian pengujian multikolinearitas untuk melihat korelasi antar variabel-variabel dengan menggunakan nilai *Variance Inflation Factor* (VIF). Jika nilai VIF pada semua variabel kurang dari sama dengan 10, maka tidak terjadi multikolinearitas sehingga proses penelitian dapat diteruskan ke langkah selanjutnya. Namun, jika nilai VIF pada semua variabel terdapat nilai VIF yang lebih dari 10, maka dilakukan eliminasi satu per satu variabel dengan nilai VIF lebih dari 0. Kemudian dilakukan pengujian multikolinearitas kembali hingga tidak terjadi multikolinearitas. Selanjutnya melakukan standarisasi, dikarenakan satuan variabel-variabel yang digunakan berbeda.

Selanjutnya menentukan jumlah *cluster* awal atau kelompok yang akan dibentuk, k yang digunakan adalah $(k = 2, 3, 4, 5, dan\ 6)$. Pada algoritma *K-Medoids* tahapan selanjutnya adalah memilih objek amatan secara acak untuk digunakan sebagai *medoids* awal. Kemudian menghitung jarak *euclidean* untuk

melihat ukuran kedekatan antar objek hitung. Selanjutnya memilih *medoids* baru secara acak dan menghitung jarak *euclidean medoids* yang baru. Setelah itu, menghitung total jarak (S), Jika $(S) < 0$, maka bentuk k objek baru sebagai *medoids* dan ulangi tahapannya hingga nilai total jarak $(S) \geq 0$.

Setelah diperoleh hasil *clustering*, kemudian menentukan jumlah *cluster* optimal dengan optimasi DBI. Langkah pertama adalah meminimumkan jarak antar setiap titik *cluster* dengan menghitung nilai *Sum of Square Within* (SSW). Selanjutnya menghitung nilai *Sum of Square Between* (SSB) untuk memaksimumkan perbedaan antar *cluster*. Setelah itu menghitung nilai rasio menggunakan nilai *SSW* dan *SSB* yang diperoleh sebelumnya untuk mendapatkan perbandingan antara satu *cluster* dengan lainnya. Setelah nilai rasio diperoleh, selanjutnya menghitung nilai DBI. Kemudian mencari nilai DBI terkecil dari masing-masing *cluster* yang digunakan. *Cluster* yang optimal dilihat dari nilai DBI terkecil. Setelah diperoleh jumlah *cluster* optimal, selanjutnya melakukan interpretasi dari setiap *cluster* yang diperoleh.





Gambar 1.1 *Flowchart* Penelitian