# ANALISIS *CLUSTER* DENGAN METODE *WARD LINKAGE* PADA INDEKS PEMBANGUNAN MANUSIA DI PROVINSI KALIMANTAN BARAT

**FANNY RAHMASARI**

**H1011201067**

**SKRIPSI**

**PROGRAM STUDI MATEMATIKA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS TANJUNGPURA**

**PONTIANAK**

**2025**

# ANALISIS *CLUSTER* DENGAN METODE *WARD LINKAGE* PADA INDEKS PEMBANGUNAN MANUSIA DI PROVINSI KALIMANTAN BARAT

**FANNY RAHMASARI**

**H1011201067**

**SKRIPSI**

Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Matematika pada Program Studi matematika

**PROGRAM STUDI MATEMATIKA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS TANJUNGPURA**

**PONTIANAK**

**2025**

## LEMBAR PENGESAHAN

| | |
|---|---|
| Judul Tugas Akhir | : Analisis *Cluster* Dengan Metode *Ward Linkage* pada Indeks Pembangunan Manusia Di Provinsi Kalimantan Barat |
| Nama Mahasiswa | : Fanny Rahmasari |
| NIM | : H1011201067 |
| Jurusan/Program Studi | : Matematika/Matematika |
| Tanggal Lulus | : 21 Mei 2025 |
| SK Pembimbing | : No.706/UN22.8/TD.06/2024/Tanggal 1 Maret 2024 |
| SK Penguji | : No.1409/UN22.8/TD.06/2025/Tanggal 14 Mei 2025 |

Dosen Pembimbing

| Pembimbing I | Pembimbing II |
|---|---|
| Drs Helmi, M.Si. | Fransiskus Fran, M.Si. |
| NIP 196410171998021001 | NIP 198804152019031014 |

Dosen Penguji

| Ketua Penguji | Anggota Penguji |
|---|---|
| Meliana Pasaribu, M.Sc. | Dr. Nilamsari Kusumastuti, M.Sc. |
| NIP 198710192019032006 | NIP 198105102005012003 |

| Pimpinan Sidang (merangkap anggota penguji) | Sekretaris Sidang (merangkap anggota penguji) |
|---|---|
| Drs Helmi, M.Si. | Fransiskus Fran, M.Si. |
| NIP 196410171998021001 | NIP 198804152019031014 |

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 | : Fanny Rahmasari |
| NIM | : H1011201067 |
| Program Studi/ Jurusan | : Matematika/Matematika |
| Fakultas | : FMIPA |

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, 21 Mei 2025




Fanny Rahmasari

H1011201067

**Analisis Cluster Dengan Metode *Ward Linkage* pada Indeks Pembangunan Manusia Di Provinsi Kalimantan Barat**


## Abstrak

Indeks Pembangunan Manusia (IPM) adalah ukuran komposit yang digunakan untuk mengukur salah satu aspek penting yang berkaitan dengan kualitas dari hasil pembangunan manusia dan ekonomi, yakni derajat perkembangan manusia. IPM di Provinsi Kalimantan Barat tahun 2023 mencapai 70,47 poin, meningkat (1,09 %) dibanding tahun sebelumnya (69,71 poin). Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Provinsi Kalimantan Barat berdasarkan nilai Indeks Pembangunan Manusia (IPM) dengan menggunakan metode cluster hierarki, khususnya pendekatan *ward linkage* yang bertujuan meminimalkan variasi dalam cluster. Analisis cluster dilakukan melalui beberapa tahapan, yaitu pengumpulan data, analisis statistik deskriptif, standarisasi data, pengujian multikolinearitas, perhitungan jarak Euclidean, dan pengelompokan menggunakan metode *Ward linkage* yang diakhiri dengan interpretasi hasil. Proses pengolahan data dilakukan menggunakan software SPSS dengan menghasilkan dendogram sebagai dasar pembentukan cluster. Data yang digunakan merupakan data IPM tahun 2023 dari 14 kabupaten/kota. Hasil analisis menunjukkan bahwa seluruh wilayah dapat dikelompokkan ke dalam dua cluster. Cluster pertama terdiri dari kabupaten/kota dengan nilai IPM yang lebih tinggi dan karakteristik sosial ekonomi yang relatif baik, antara lain Kota Pontianak, Kota Singkawang, Kabupaten Mempawah, Sambas, Ketapang, Kubu Raya, dan Sanggau. Sementara itu, cluster kedua mencakup wilayah dengan memiliki IPM sedang atau relatif lebih rendah seperti Kabupaten Bengkayang, Landak, Melawi, Sintang, Kapuas Hulu, Sekadau, dan Kayong Utara.



**Kata Kunci**: *Ketimpangan pembangunan, Klasifikasi daerah, Dendogram*

***Cluster Analysis Using the Ward Linkage Method on the Human Development Index in West Kalimantan Province***


## *Abstract*

The Human Development Index (HDI) is a measuring tool used to measure one of the important aspects related to the quality of human economic development results, namely the degree of human development. The HDI in West Kalimantan Province in 2023 reached 70.47%, an increase (1.09%) compared to the previous year (69.71%). This study aims to group districts/cities in West Kalimantan Province based on the Human Development Index (HDI) value using the hierarchical cluster method, especially the *ward linkage* approach which aims to minimize variation in clusters. Cluster analysis is carried out through several stages, namely data collection, descriptive statistical analysis, data standardization, multicollinearity testing, Euclidean distance calculation, and grouping using the *Ward linkage* method which ends with interpretation of the results. The data processing process is carried out using SPSS software by producing a dendogram as the basis for forming clusters. The data used is the 2023 HDI data from 14 districts/cities. The results of the analysis show that all regions can be grouped into two main clusters. The first cluster consists of districts/cities with higher HDI values and relatively good socio-economic characteristics, including Pontianak City, Singkawang City, Mempawah Regency, Sambas, Ketapang, Kubu Raya, and Sanggau. Meanwhile, the second cluster includes areas with moderate or relatively lower HDI such as Bengkayang, Landak, Melawi, Sintang, Kapuas Hulu, Sekadau, and Kayong Utara Regencies.



***Keywords***: *Development inequality, Regional classification, Dendogram*

# PRAKATA

Puji dan syukur penulis panjatkan kehadirat Allah SWT, yang telah memberikan rahmat dan hidayah-Nya kepada penulis, sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis *Cluster* Dengan Metode *Ward Linkage* pada Indeks Pembangunan Manusia Di Provinsi Kalimantan Barat". Dengan segala kerendahan hati penulis mengucapkan terima kasih kepada:

1. Orang tua, kakak, adik dan keluarga besar yang tak henti-hentinya memberikan doa dan dukungan serta kasih sayang yang tulus sehingga penulis dapat menyelesaikan skripsi ini.
2. Drs. Helmi, M.Si. selaku Dosen Pembimbing I dan Fransiskus Fran, M.Si selaku Dosen Pembimbing II yang telah memberikan bimbingan, arahan, dan saran kepada penulis dalam menyelesaikan skripsi ini dengan baik.
3. Meliana Pasaribu, M.Sc selaku dosen penguji pertama dan Ibu Dr. Nilamsari Kusumastuti, M.Sc selaku dosen penguji kedua yang telah memberikan evaluasi dan saran perbaikan yang membangun dalam proses penyusunan skripsi ini.
4. Dr. Nilamsari Kusumastuti, M.Sc sebagai Dosen Pembimbing Akademik yang selalu memberikan arahan serta bimbingan kepada penulis selama masa perkuliahan.
5. Sahabat-sahabat seperjuangan: Amelia Vickylia Roos, Ayu Oktavia Tambunan, Elma Silvia dan Aisyah Syafina Putri yang selalu memberikan semangat serta dukungan kepada penulis.
6. Seluruh pihak yang telah membantu dan terlibat dalam menyelesaikan skripsi ini.

Penulis menyadari bahwa dalam penulisan skripsi ini masih jauh dari kesempurnaan, oleh karena itu kritik dan saran sangat penulis harapkan. Semoga skripsi ini dapat bermanfaat bagi penulis dan bagi para pembaca.

Penulis

## DAFTAR ISI

**LEMBAR PENGESAHAN** .....**Error! Bookmark not defined.**
**PERNYATAAN INTEGRITAS AKADEMIK**....i**Error! Bookmark not defined.**
**ABSTRAK** .....iiv
***ABSTRACT*** .....v
**PRAKATA**.....vi
**DAFTAR ISI**.....vii
**DAFTAR TABEL** .....ix
**DAFTAR GAMBAR**.....x
**DAFTAR SIMBOL**.....xi
**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 ..... 4
**BAB II LANDASAN TEORI** ..... 8
2.1 Analisis Cluster ..... 8
2.2 Indeks Pembangunan Manusia..... 9
2.3 Standarisasi Data ..... 11
2.4 Asumsi Analisis *Cluster*..... 11
2.5 Asumsi Multikolineritas..... 11
2.6 Ukuran Jarak Antar Data..... 12
2.7 Metode Pengelompokan..... 14
**BAB III HASIL DAN PEMBAHASAN** ..... 16
3.1 Deskripsi Data ..... 16
3.2 Hasil Statistik Deskriptif Variabel Penelitian..... 20

3.3 Standarisasi Data ..... 25
3.4 Asumsi Multikolineritas ..... 26
3.5 Perhitungan Jarak *Eucliden* ..... 27
3.6 Metode *Ward* Lingkage ..... 28
**BAB IV KESIMPULAN** ..... 32
**DAFTAR PUSTAKA** ..... 33
**LAMPIRAN** ..... 36

# DAFTAR TABEL

Tabel 3. 1 Deskripsi Data ..... 19
Tabel 3. 2 Sumber Data Indeks Pembangunan Manusia ..... 20
Tabel 3. 3 Statistik Deskriptif ..... 21
Tabel 3. 4 Standarisasi Data ..... 25
Tabel 3. 5 Uji Multikolinieritas ..... 27
Tabel 3. 6 Jarak *Euclidean* ..... 28
Tabel 3. 7 *Agglomeration Schedule* ..... 29

# DAFTAR GAMBAR

Gambar 1. 1 Flowchart Alur Metodologi Penelitian..... 7
Gambar 3. 2 Statistik Deskriptif dalam bentuk diagram batang ..... 21
Gambar 3. 3 Jumlah Anggota pada Setiap *Cluster* ..... 31

## DAFTAR SIMBOL

$d_{ij}$ : jarak antar objek ke-$i$ dan objek ke-$j$

$p$ : jumlah variabel cluster

$X_{ik}$ : data dari subjek ke-$i$ pada variabel ke-$k$

$X_{jk}$ : data dari subjek ke-$j$ pada variabel ke-$k$

$Z$ : nilai $z_{score}$

$x_i$ : nilai pengamatan ke-$i$

$\overline{x}$ : nilai rata-rata dari variabel penelitian

$S_x$ : simpangan baku

$k$ : banyaknya objek dalam cluster yang terbentuk

$SSE_{ij}$ : Jarak antara objek $i$ dan objek $j$

$x_i$ , $x_j$ : Nilai rata-rata objek $i$ dan $j$

$k$ : 1,2 … , $p$

$I_{(IJ)K}$ : Jarak antara cluster $IJ$ dan $K$

$I_{IK}$ : Jarak antara cluster $I$ dan $K$

$I_{JK}$ : Jarak antara cluster $J$ dan $K$

$I_{IJ}$ : Jarak antara cluster $I$ dan $J$

$n_I, n_J, n_K$ : Banyaknya objek pada cluster ke $I, J, K$

# BAB I
# PENDAHULUAN

## 1.1 Latar Belakang

Indeks Pembangunan Manusia (IPM) merupakan terobosan dalam mengevaluasi pembangunan manusia. Indeks pembangunan manusia ini merupakan ukuran untuk menilai kualitas pembangunan suatu wilayah, yang umumnya menunjukkan korelasi negatif terhadap tingkat kemiskinan dan pengangguran. Artinya, semakin tinggi nilai IPM suatu daerah, maka diharapkan semakin tinggi pula kualitas hidup masyarakatnya. Indeks Pembangunan Manusia (IPM) Menurut Saputra (2011) adalah angka yang mengukur kinerja pembangunan manusia berdasarkan sejumlah komponen fundamental kualitas hidup yang dapat mempengaruhi tingkat produktivitas individu. Sebagai ukuran kualitas hidup, IPM dibangun dengan menggunakan pendekatan dimensi hidup sehat, pengetahuan dan kehidupan yang bermartabat (Mahroji & Nurkhasanah, 2019). Salah satu faktor kunci dalam keberhasilan program pembangunan adalah ketepatan dalam mengidentifikasi kelompok sasaran dan wilayah sasaran. Oleh karena itu, pengelompokan 14 Kabupaten/Kota di Provinsi Kalimantan Barat pada indikator Indeks Pembangunan Manusia (IPM) menjadi penting untuk dipertimbangkan. Untuk mencapai tujuan ini, pendekatan yang tepat adalah analisis cluster (Wijayanto, 2022). Secara empiris dalam penelitian Maulana dan Bowow (2013) bahwa pendidikan sangat berpengaruh terhadap Indeks Pembangunan Manusia (IPM) di Indonesia pada periode 2007-2011, yang berarti semakin tinggi tingkat pendidikan maka semakin tinggi pula nilai IPM suatu wilayah.

Untuk memahami pola distribusi IPM antar wilayah, salah satu metode yang dapat digunakan adalah analisis cluster. Analisis cluster merupakan salah satu cabang statistika yang digunakan untuk mengelompokan data berdasarkan kemiripan karakteristik antar objek (Rahayu, 2013). Analisis cluster merupakan salah satu teknik dalam analisis multivariat metode interdependen yang tidak membedakan variabel bebas dan variabel terikat melainkan memfokuskan pada

hubungan kesamaan di antara seluruh variabel yang diamati dalam mengelompokkan suatu data. mengelompokkan objek-objek pengamatan menjadi beberapa kelompok yang mengacu pada karakteristik yang dipertimbangkan untuk diteliti merupakan tujuan utama dari analisis *cluster* (Supranto, 2004). Dalam analisis cluster, objek yang berada dalam kelompok yang sama merupakan objek yang paling dekat kesamaannya, serta mempunyai kemiripan satu dengan yang lain (Johnson & Wichern, 2007). Dengan demikian, penggunaan analisis cluster dalam studi IPM sangat relevan untuk mengelompokkan kabupaten/kota di Kalimantan Barat berdasarkan kemiripan nilai IPM yang dimiliki. Melalui clusterisasi, dapat diketahui pola-pola pembangunan manusia yang terjadi di wilayah tersebut, serta membantu dalam merancang kebijakan yang lebih tepat sasaran sesuai dengan karakteristik kelompok wilayah yang terbentuk. mengidentifikasi kelompok-kelompok wilayah yang memiliki pola pembangunan manusia yang serupa, yang pada akhirnya dapat membantu dalam perencanaan dan evaluasi kebijakan pembangunan yang lebih tepat sasaran

Pada penelitian ini digunakan metode *Ward* yaitu salah satu metode dalam analisis cluster hierarki yang digunakan untuk menentukan jarak antara dua cluster. Metode ini menghitung jarak antar objek berdasarkan jarak Euclidean terhadap nilai rata-rata (centroid) cluster. Setiap jarak dihitung dengan mempertimbangkan peningkatan total kuadrat galat (error sum of squares) dalam penggabungan cluster, sehingga penggunaan metode *Ward* dalam analisis cluster juga dinilai tepat karena mampu menghasilkan kelompok yang paling homogen secara internal. Pada setiap tahap pengelompokan, dua cluster yang memiliki kenaikan jumlah kuadrat sesatan dalam cluster yang terkecil digabungkan (Hartanto, dkk 2017).

Dalam metode *ward* didefinisikan sebagai metode pengclusteran dengan memaksimumkan kesamaan dalam cluster dengan menggunakan perhitungan yang lengkap. Pada metode *ward*, jarak antara dua cluster yang terbentuk adalah Sum of Square Error (SSE) (Rencher, 2002). Sesuai dengan tujuan utama analisis cluster maka metode *ward* dapat digunakan untuk mengelompokan masalah di

atas. Pengelompokkan data pada umumnya terfokus pada pengelompokan objek dengan
variabel data berskala numerik, sedangkan untuk data kategorik kurang difokuskan dan mulai berkembang sekitar tahun 1990-an. Pada umumnya analisis cluster hanya dikembangkan untuk memproses salah satu data yaitu data kategorik atau data numerik saja.

Berdasarkan uraian latar belakang tersebut peneliti tertarik mengangkat sebuah penelitian yang judul "Analisis *Cluster* dengan Metode *Ward Linkage* pada Indeks Pembangunan Manusia di Provinsi Kalimantan Barat".

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah dikemukakan sebelumnya, diperoleh rumusan masalahnya yaitu bagaimana menghasilkan pengelompokan kabupaten/kota di Kalimantan Barat pada tahun 2023 dengan metode *ward* pada analisis *cluster* dalam indeks pembangunan manusia.

## 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang sudah dikemukakan sebelumnya, maka tujuan penelitian yang ingin dicapai adalah sebagai berikut.

1. Mengelompokkan kabupaten/kota di Kalimantan Barat berdasarkan nilai IPM tahun 2023.
2. Menerapkan metode *Ward* dalam analisis *cluster* untuk mengetahui kesamaan karakteristik IPM antar wilayah.
3. Menyajikan hasil pengelompokan dalam bentuk visualisasi (dendrogram atau peta cluster) yang dapat digunakan sebagai dasar kebijakan pembangunan wilayah.

## 1.4 Batasan Masalah

Batasan masalah dalam penelitian ini adalah data yang digunakan merupakan data sekunder yang dibatasi pada data IPM yang tersedia dari Badan Pusat Statistik (BPS) Provinsi Kalimantan Barat pada tahun 2023, Analisis cluster dalam penelitian ini hanya akan menggunakan Metode *Ward Linkage* dengan perhitungan jarak Euclidean, tanpa membandingkan dengan metode clustering lainnya seperti k-means atau metode agglomerative lainnya dan hasil dari

penelitian ini akan dibatasi pada pembentukan cluster atau kelompok kabupaten/kota di Provinsi Kalimantan Barat berdasarkan kemiripan dalam IPM, dan tidak akan mengeksplorasi hubungan kausal atau prediktif antara variabel lain dengan IPM.

## 1.5 Tinjauan Pustaka

M. Y. Matdoan, dkk (2020) mengkaji klasifikasi kabupaten/kota di Provinsi Maluku berdasarkan Indeks Pembangunan Manusia (IPM) menggunakan analisis cluster hierarki. Penelitian menemukan metode *Ward*'s sebagai yang terbaik, membagi wilayah menjadi empat cluster IPM: tinggi (Kota Ambon), sedang, rendah, dan sangat rendah (Kepulauan Tanimbar, Seram Bagian Timur, dll.). Temuan ini diharapkan membantu kebijakan pembangunan yang lebih tepat sasaran di Maluku.

Yogi Isna Harnanto, dkk (2017) menggunakan metode *Ward* untuk mengelompokkan kabupaten/kota di Jawa Tengah berdasarkan penggunaan alat kontrasepsi seperti IUD, kondom, implan, suntikan, dan pil. Hasilnya, kabupaten/kota dikelompokkan menjadi tiga cluster berdasarkan tingkat penggunaan alat kontrasepsi, dengan cluster pertama memiliki penggunaan tertinggi dan cluster ketiga terendah.

Insiyah, dkk (2023) membahas pengelompokkan daerah rawan kriminalitas di Jawa Timur pada tahun 2021 menggunakan metode *ward clustering*. Berdasarkan data kriminalitas dari BPS, penelitian ini menghasilkan tiga cluster utama dari 38 kabupaten/kota. Penelitian ini membantu memetakan daerah dengan tingkat kriminalitas tinggi untuk mendukung upaya pencegahan.

## 1.6 Metodologi Penelitian

Hal pertama yang dilakukan sebelum melakukan penelitian adalah mencari topik yang akan diteliti. Adapun dalam penelitian ini menggunakan sistem studi literatur yang bersumber dari jurnal, buku ataupun artikel yang sesuai dengan topik penelitian.

Dalam penelitian kali ini membahas tentang masalah pengelompokan kabupaten/kota, yang dimana pada masalah ini metode yang digunakan yaitu

metode *agglomerative hierarchical* dengan sub metode *ward linkage*. Metode *ward* adalah salah satu metode yang dapat digunakan untuk menyelesaikan masalah analisis *cluster*. Pelaksanaan penelitian pada tahap kedua, pelaksanaan penelitian dilakukan dengan mengumpulkan sejumlah data yang dibutuhkan dalam penelitian, mengolah data penelitian, membuat rencana analisis data penelitian, selanjutnya menganalisis data penelitian dan menafsirkan hasil data penelitian.

Kemudian pada tahap ketiga, laporan penelitian menjadi tahap terakhir yang menentukan apakah suatu penelitian yang sudah dilakukan baik atau tidak. Pada tahap pembuatan laporan penelitian dilakukan penarikan kesimpulan dan menyusun laporan sebagai hasil penelitian dalam bentuk skripsi. Untuk mempermudah perhitungan, penelitian ini juga menggunakan bantuan software *SPSS*. Data yang digunakan dalam penelitian ini adalah data sekunder mengenai indeks pembangunan manusia pada tahun 2023 yang di ambil dari BPS Provinsi Kalimantan Barat. Pada bagian metodologi penelitian ini diuraikan langkah-langkah sebagai berikut :

1. Sumber Data

   Data yang digunakan dalam penelitian ini bersumber dari data BPS Provinsi Kalimantan Barat dengan menggunakan 1 variabel terikat yaitu Indeks Pembangunan Manusia dan 6 variabel bebas yaitu Persentase Penduduk Miskin, Tingkat Pengangguran Terbuka, Umur Harapan Hidup, Persentase Penduduk, Angka Partisipasi Sekolah dan Indeks Gizi Buruk.
2. Statistik Deskriptif

   Statistik deskriptif adalah teknik dalam statistik yang digunakan untuk menggambarkan atau meringkas karakteristik data dalam suatu penelitian. Tujuannya adalah untuk memberikan pandangan umum tentang karakteristik dasar dari data tersebut tanpa menarik kesimpulan lebih dalam. Analisis ini berfungsi sebagai langkah awal dalam pengolahan data, guna membantu peneliti memahami struktur data. Dalam penelitian ini, statistik deskriptif mencakup nilai

minimum, maksimum, rata-rata (mean), dan simpangan baku (standar deviasi) dari masing-masing variabel.

3. Uji Multikolinearitas

   Uji multikolinearitas adalah prosedur dalam analisis regresi untuk mendeteksi adanya hubungan linear yang tinggi antar variabel bebas (Gujarati & Porter, 2009). Multikolinearitas terjadi saat dua atau lebih variabel independen berkorelasi kuat, sehingga menyulitkan untuk mengukur pengaruh masing-masing variabel secara akurat terhadap variabel terikat.

   Model dikatakan bebas dari multikolinearitas jika seluruh nilai VIF < 10, dan lebih baik lagi jika VIF < 5.

4. Standarisasi Data

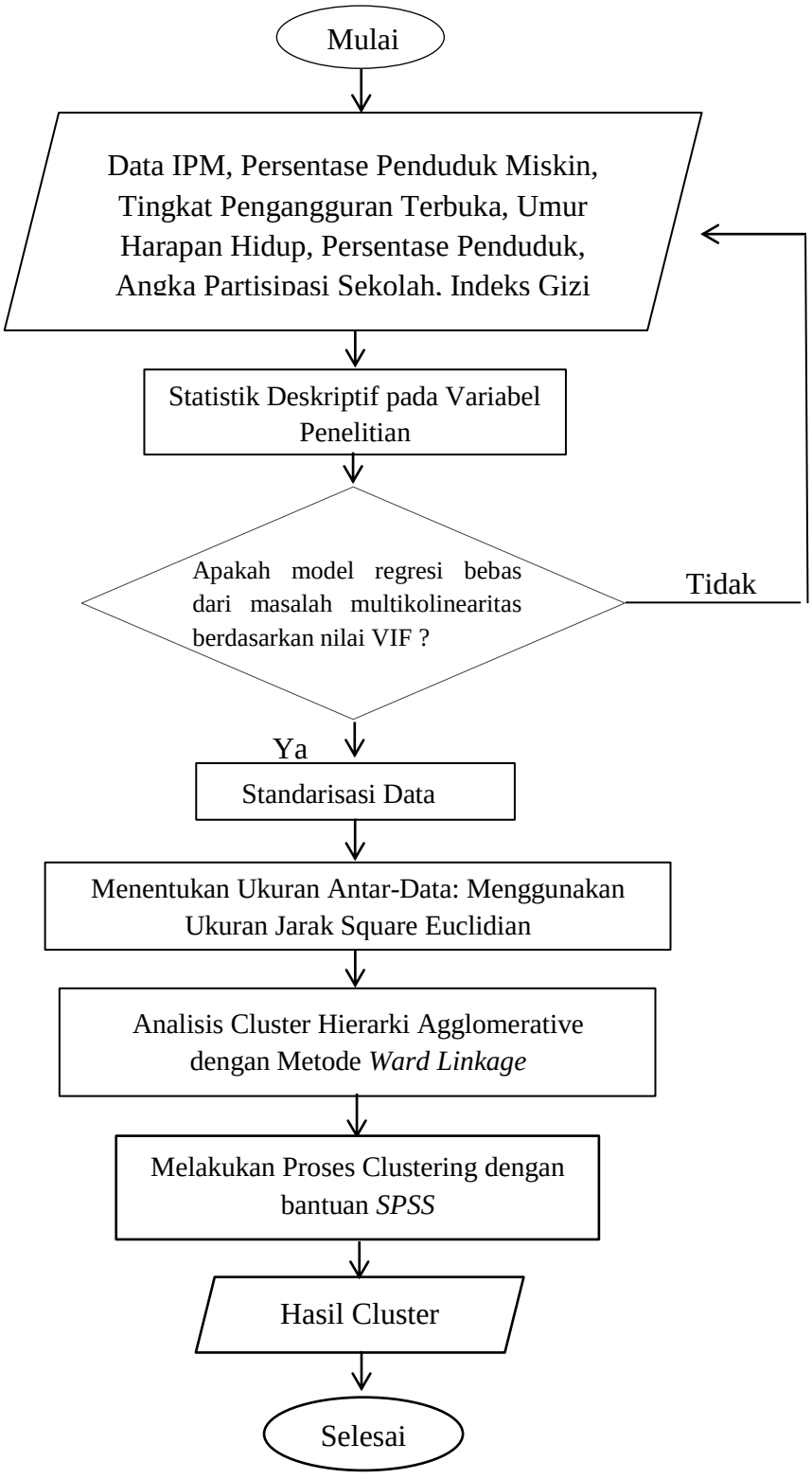
   Jika terdapat perbedaan satuan nilai dalam setiap variabel penelitian, maka data tersebut perlu di standarisasi. Perbedaan yang di maksud yaitu misalkan terdapat 2 variabel: Varibel pertama memiliki nilai dalam satuan persen sedangkan variabel kedua memiliki nilai satuan tahun. Kedua variabel jika tidak dilakukan standarisasi data, maka pada saat dilakukan pengelompokkan akan ada perbedaan yang besar dalam skala data.

5. Tahap *Hierarchical Clustering*

   Pada tahap ini data yang sudah siap dimodelkan menggunakan metode *Agglomerative Hierarchical Clustering*. Metode perhitungan dalam analisis cluster yaitu :

   1. Menghitung matriks jarak dengan menggunakan jarak Euclidean Menetapkan setiap objek adalah sebuah cluster
   2. Menggabungkan dua cluster berdasarkan ukuran jarak dan metode penggabungan yang digunakan
   3. Memperbarui matriks jarak dan menghitung jarak antar cluster baru dengan cluster awal
   4. Mengulangi langkah ke 3 dan 4 sampai terbentuk sejumlah cluster yang telah ditentukan.

Diagram alur penelitian ini dapat dilihat pada *flowchart* berikut:




**Gambar 1. 1 Flowchart Alur Metodologi Penelitian**