

**PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI  
SUMATERA BARAT BERDASARKAN INDIKATOR  
KESEHATAN MASYARAKAT TAHUN 2020  
MENGUNAKAN ANALISIS *CLUSTER***

**TUGAS AKHIR**

*Untuk memenuhi persyaratan memperoleh gelar Ahli madya*



Oleh  
**SINTA DWI PUSPITA SARI**  
**NIM 18037070**

**PROGRAM STUDI DIPLOMA III STATISTIKA  
JURUSAN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS NEGERI PADANG  
2022**

**PERSETUJUAN UJIAN TUGAS AKHIR**

**PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI SUMATERA  
BARAT BERDASARKAN INDIKATOR KESEHATAN MASYARAKAT  
TAHUN 2020 MENGGUNAKAN ANALISIS *CLUSTER***

Nama : Sinta Dwi Puspita Sari  
NIM/Tahun Masuk : 18037070/2018  
Program Studi : DIII Statistika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Padang, 9 Februari 2022

Disetujui oleh :

Pembimbing Akademik



Admi Salma, S.Pd., M.Si  
NIP. 182005

**HALAMAN PENGESAHAN LULUS UJIAN TUGAS AKHIR**

Nama : Sinta Dwi Puspita Sari  
TM/NIM : 2018/18037070  
Program Studi : DIII Statistika  
Jurusan : Statistika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam

**PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI SUMATERA  
BARAT BERDASARKAN INDIKATOR KESEHATAN MASYARAKAT  
TAHUN 2020 MENGGUNAKAN ANALISIS *CLUSTER***

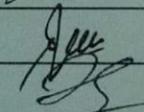
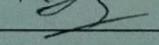
**Dinyatakan lulus setelah dipertahankan di depan Tim Penguji Tugas Akhir  
Program Studi DIII Statistika Jurusan Statistika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Negeri Padang**

Padang, 9 Februari 2022

Tim Penguji

Nama

1. Ketua : Admi Salma, S.Pd., M.Si
2. Anggota : Dr. Dony Permana, M.Si
3. Anggota : Dra. Nonong Amalita, M.Si

1.   
2.   
3. 

## SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan di bawah ini :

Nama : Sinta Dwi Puspita Sari  
NIM/TM : 18037070/2018  
Program Studi : DIII Statistika  
Jurusan : Statistika  
Fakultas : MIPA UNP

Dengan ini menyatakan, bahwa Tugas Akhir saya dengan judul "**Pengelompokan Kabupaten/Kota Di Provinsi Sumatera Barat Berdasarkan Indikator Kesehatan Masyarakat Tahun 2020 Menggunakan Analisis Cluster**" adalah benar merupakan hasil karya saya dan bukan merupakan plagiat dari karya orang lain atau pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam tradisi keilmuan. Apabila suatu saat terbukti saya melakukan plagiat maka saya bersedia diproses dan menerima sanksi akademis maupun hukum sesuai dengan hukum dan ketentuan yang berlaku, baik di institusi UNP maupun di masyarakat dan Negara.

Demikianlah pernyataan ini saya buat dengan penuh kesadaran rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Diketahui oleh,

Ketua Jurusan Statistika,

  
Dr. Dedy Permana, M.Si  
NIP. 197501272006041001

Saya yang menyatakan,  
  
Sinta Dwi Puspita Sari  
NIM. 18037070

## ABSTRAK

### **Sinta Dwi Puspita S: Pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat Berdasarkan Indikator Kesehatan Masyarakat Tahun 2020 Menggunakan Analisis *Cluster***

Beberapa wilayah di Indonesia masih ada yang memiliki ketidakmerataan indikator kesehatan masyarakat, salah satunya adalah Provinsi Sumatera Barat. Hal ini dapat dilihat dari masih adanya daerah di Provinsi Sumatera Barat yang memiliki indikator kesehatan masyarakat berada di bawah capaian Provinsi nya. Oleh karena itu, perlu dilakukan pengelompokan kabupaten/kota di Sumatera Barat berdasarkan indikator kesehatan masyarakat sehingga memudahkan pemerintah dalam mengambil kebijakan untuk pembangunan kesehatan di daerah tersebut.

Jenis penelitian ini merupakan penelitian terapan. Data yang digunakan berupa data sekunder yang diperoleh dari Badan Pusat Statistik Sumatera Barat. Variabel yang digunakan adalah angka harapan hidup, angka morbiditas, persentase rumah tangga yang memiliki akses sanitasi layak, persentase rumah tangga yang memiliki akses terhadap sumber air minum layak.

Analisis *cluster* hierarki dengan metode *average linkage* telah digunakan untuk melakukan pengelompokan dan diperoleh hasil bahwa terdapat 4 *cluster* berdasarkan indikator kesehatan masyarakat. *Cluster* 1 termasuk kategori tinggi, dengan kabupaten/kotanya adalah Kota Padang, Kota Payakumbuh, Kota Solok, Kota Bukittinggi, Kota Sawah Lunto, Kota Pariaman, Kabupaten Agam, Kota Padang Panjang, Kabupaten Tanah Datar, Kabupaten Padang Pariaman, Kabupaten Pesisir Selatan, Kabupaten Dharmasraya, *cluster* 2 termasuk kategori rendah dengan kabupaten/kotanya adalah Kabupaten Solok, Kabupaten Lima Puluh Kota, Kabupaten Sijunjung, Kabupaten Pasaman Barat, Kabupaten Pasaman, *cluster* 3 termasuk kategori sangat rendah dengan kabupaten/kotanya Kabupaten Kepulauan Mentawai, dan *cluster* 4 termasuk kategori sedang dengan kabupaten/kotanya adalah Kabupaten Solok Selatan.

**Kata Kunci:** Indikator Kesehatan Masyarakat, *Cluster Analysis*, *Average Linkage*

## KATA PENGANTAR



Puji dan syukur penulis ucapkan kehadiran Allah SWT karena berkat rahmat dan karunia-Nya peneliti dapat menyelesaikan tugas akhir yang berjudul **“Pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat Berdasarkan Indikator Kesehatan Masyarakat Tahun 2020 Menggunakan Analisis Cluster”**. Penulisan Tugas Akhir ini merupakan salah satu syarat dalam menyelesaikan pendidikan pada Program Studi Diploma III Statistika Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang. Dalam penulisan dan penyusunan Tugas Akhir ini peneliti banyak mendapatkan dukungan, bimbingan, arahan serta saran dari berbagai pihak. Untuk itu peneliti mengucapkan terima kasih kepada:

1. Ibu Admi Salma, S.Pd, M.Si., pembimbing dan penasehat akademik yang telah meluangkan waktunya dalam membimbing dan memberikan saran dan arahan yang bermanfaat bagi peneliti.
2. Bapak Dr. Dony Permana, M.Si., dosen penguji dan sekaligus Ketua Jurusan Statistika Fakultas Matematika Ilmu Pengetahuan Alam Universitas Negeri Padang.
3. Ibu Dra. Nonong Amalita, M.Si., dosen penguji sekaligus Sekretaris Jurusan Statistika Fakultas Matematika Ilmu Pengetahuan Alam Universitas Negeri Padang.
4. Bapak dan Ibu Dosen Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam.

5. Teristimewa untuk kedua orang tua yang selalu memberikan dukungan baik moril maupun materil serta motivasi dan doa.
6. Sahabat, teman, dan rekan-rekan yang telah banyak membantu dan memberikan dukungan kepada peneliti.

Semoga semua bimbingan, bantuan, dan kerjasamanya dapat dibalas oleh Allah SWT sebagai amal ibadah. Dalam penyusunan Tugas Akhir ini peneliti telah berusaha semaksimal mungkin untuk memberikan yang terbaik, namun peneliti menyadari bahwa penulisan Tugas akhir ini belum sempurna, karena keterbatasan pengetahuan dan kemampuan yang peneliti miliki. Untuk itu kritik dan saran sangat peneliti harapkan demi kesempurnaan penyusunan Tugas Akhir ini. Semoga Tugas Akhir ini dapat bermanfaat bagi peneliti dan pembaca.

Padang, 9 Februari 2022

Sinta Dwi Puspita Sari

## DAFTAR ISI

<b>ABSTRAK</b> .....	i
<b>KATA PENGANTAR</b> .....	ii
<b>DAFTAR ISI</b> .....	iv
<b>DAFTAR TABEL</b> .....	v
<b>DAFTAR GAMBAR</b> .....	vi
<b>DAFTAR LAMPIRAN</b> .....	vii
<b>BAB I PENDAHULUAN</b> .....	1
A. Latar Belakang .....	1
B. Batasan Masalah.....	6
C. Pertanyaan Penelitian .....	6
D. Tujuan Penelitian .....	6
E. Manfaat Penelitian .....	7
<b>BAB II KAJIAN TEORI</b> .....	8
A. Kesehatan .....	8
B. Indikator Kesehatan Masyarakat.....	8
C. Standardisasi Data.....	11
D. Analisis <i>Cluster</i> .....	11
<b>BAB III METODE PENELITIAN</b> .....	20
A. Jenis Penelitian.....	20
B. Jenis dan Sumber Data .....	20
C. Variabel Penelitian .....	20
D. Struktur Data Penelitian .....	21
E. Teknik Analisis Data.....	21
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	23
A. Hasil Penelitian .....	23
B. Pembahasan.....	33
<b>BAB V PENUTUP</b> .....	35
<b>DAFTAR PUSTAKA</b> .....	37
<b>LAMPIRAN</b> .....	39

## **DAFTAR TABEL**

<b>Tabel</b>	<b>Halaman</b>
1. Struktur Data .....	21
2. Statistik Deskriptif Indikator Kesehatan Masyarakat tahun 2020.....	23
3. Cluster dengan Masing-masing Anggota .....	31
4. Rata-rata Variabel Setiap Indikator Kesehatan Masyarakat .....	32

## DAFTAR GAMBAR

<b>Gambar</b>	<b>Halaman</b>
1. Contoh Dendogram .....	18
2. Angka Harapan Hidup per kabupaten/kota di Provinsi Sumatera Barat Tahun 2020.....	24
3. Grafik Angka Morbiditas per kabupaten/kota di Provinsi Sumatera Barat Tahun 2020.....	25
4. Grafik Persentase Rumah Tangga yang Memiliki Akses Sanitasi Layak per kabupaten/kota di Provinsi Sumatera Barat Tahun 2020 .....	26
5. Dendogram Average Linkage .....	30

## DAFTAR LAMPIRAN

<b>Lampiran</b>	<b>Halaman</b>
1. Data Angka Harapan Hidup di Indonesia tahun 2020 .....	39
2. Data Persentase Rumah Tangga yang Memiliki Akses Sanitasi Layak di Indonesia Tahun 2020 .....	40
3. Data Persentase Rumah Tangga yang Memiliki Akses terhadap Sumber Air Minum Layak di Indonesia tahun 2020 .....	41
4. Data Indikator Kesehatan Masyarakat di Provinsi Sumatera Barat Tahun 2020 .....	42
5. Standarisasi Data .....	43
6. Jarak Euclid Indikator Kesehatan Masyarakat .....	44

## **BAB I**

### **PENDAHULUAN**

#### **A. Latar Belakang**

Kesehatan merupakan salah satu pondasi suatu negara dan menjadi faktor penting dalam masyarakat. Kesehatan dan kesejahteraan masyarakat merupakan salah satu usaha pemerintah untuk membangun kesehatan bangsa agar terwujud negara yang sehat. Kesehatan juga merupakan salah satu faktor yang berperan penting dalam investasi pembangunan sumber daya manusia berkualitas, oleh karena itu diperlukan adanya pembangunan di bidang kesehatan untuk meningkatkan pelayanan dan kesehatan masyarakat. Pembangunan kesehatan merupakan upaya yang dilaksanakan oleh semua komponen bangsa agar terwujud derajat kesehatan masyarakat setinggi-tingginya.

Untuk mengetahui keberhasilan pembangunan kesehatan yang dilaksanakan secara berkesinambungan diperlukan suatu indikator kesehatan. Indikator kesehatan merupakan petunjuk yang memberikan gambaran tentang kesehatan dan keadaan pada suatu daerah. Indikator kesehatan terdiri atas angka-angka mortalitas, angka-angka morbiditas, indikator status gizi, indikator keadaan lingkungan, indikator perilaku hidup masyarakat, indikator akses dan mutu pelayanan kesehatan, indikator pelayanan kesehatan, indikator sumber daya kesehatan, indikator manajemen kesehatan serta indikator kontribusi sektor-sektor terkait (Dinas Kesehatan RI, 2010). Menurut Dinas Kesehatan Provinsi Sumatera Barat ada beberapa indikator yang menjadi prioritas utama dalam hal pembangunan kesehatan. Indikator yang dimaksud adalah angka harapan hidup, angka morbiditas,

rumah tangga yang memiliki akses sanitasi layak, rumah tangga yang memiliki akses terhadap layanan sumber air minum layak.

Angka harapan hidup merupakan indikator yang dapat mencerminkan derajat kesehatan suatu wilayah, baik dari sarana prasarana, akses, hingga kualitas kesehatan (Indeks Pembangunan Manusia, 2020). Menurut Dinas Kesehatan Provinsi Sumatera Barat Tahun 2020 Angka harapan hidup di Provinsi Sumatera Barat sebesar 69,47 tahun, masih dibawah angka nasional yang mencapai 71,47 tahun, angka harapan hidup Provinsi Sumatera Barat berada pada urutan ke-22 secara nasional. Angka harapan hidup di Provinsi Sumatera Barat termasuk kelompok rendah jika dibandingkan provinsi lain di Pulau Sumatera. Data lengkap mengenai angka harapan hidup berdasarkan Provinsi di Indonesia dapat dilihat di Lampiran 1.

Indikator lain ialah angka morbiditas, semakin tinggi morbiditas berarti semakin banyak penduduk yang mengalami keluhan kesehatan, hal ini menunjukkan rendahnya tingkat kesehatan masyarakat (Profil Kesehatan Sumbar, 2020). Berdasarkan data Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2020, angka morbiditas penduduk di Provinsi Sumatera Barat adalah 14,39 persen. Angka tersebut lebih besar dibandingkan tahun 2019 yaitu sebesar 13,54 persen. Hal ini menunjukkan bahwa tingkat kesehatan penduduk Provinsi Sumatera Barat sedikit menurun dalam satu tahun terakhir.

Indikator lain adalah akses sanitasi layak, fasilitas sanitasi yang memenuhi syarat kesehatan jika dilengkapi dengan kloset leher angsa dan tempat pembuangan tangki septik. Fasilitas Sanitasi merupakan hal yang sangat penting bagi masyarakat

karena dibutuhkan dalam kehidupan sehari-hari. Buruknya kondisi sanitasi akan berdampak negatif di banyak aspek kehidupan, mulai dari turunnya kualitas lingkungan hidup masyarakat, dan tercemarnya sumber air minum (Profil Kesehatan Provinsi Sumbar, 2020). Menurut Dinas Kesehatan Provinsi Sumatera Barat Tahun 2020, penduduk yang memiliki akses sanitasi layak di Provinsi Sumatera Barat sebesar 69,93 persen, dimana angka nasionalnya bernilai 79,53 persen, hal ini menunjukkan bahwa Provinsi Sumatera Barat masih berada jauh dibawah angka nasionalnya. Data lengkap mengenai persentase rumah tangga yang memiliki akses sanitasi layak berdasarkan Provinsi di Indonesia dapat dilihat di Lampiran 2.

Indikator lain adalah rumah tangga yang memiliki akses terhadap sumber air minum yang layak. Air minum yang dikonsumsi perlu ditetapkan persyaratan kualitas air minum sehingga tidak menimbulkan gangguan kesehatan. Menurut Dinas Kesehatan Provinsi Sumatera Barat Tahun 2020 persentase rumah tangga yang memiliki akses terhadap sumber air minum layak di Provinsi Sumatera Barat sebesar 83,37 persen, masih dibawah angka nasional yang mencapai 90,21 persen. Data lengkap mengenai persentase rumah tangga yang memiliki akses terhadap sumber air minum layak berdasarkan Provinsi di Indonesia dapat dilihat di Lampiran 3.

Berdasarkan data tersebut terlihat bahwa Provinsi Sumatera Barat masih memiliki kesenjangan di bidang kesehatan. Oleh karena itu diperlukan pengelompokkan kabupaten/kota berdasarkan indikator kesehatan. Pengelompokkan wilayah kabupaten/kota didasari dari kemiripan karakteristiknya.

Hal ini dilakukan agar pemerintah dapat melihat kabupaten/kota yang perlu diprioritaskan terlebih dahulu untuk meningkatkan pembangunan di bidang kesehatan. Untuk mengelompokkan kabupaten/kota berdasarkan indikator kesehatan masyarakat diperlukan suatu analisis. Salah satu analisis yang dapat digunakan untuk memecahkan permasalahan tersebut adalah analisis *cluster*.

Menurut (Johnson, 2007), analisis *cluster* ialah teknik statistik yang digunakan untuk mengelompokkan sejumlah objek dalam kelompok-kelompok, dimana dalam satu kelompok pengamatan memiliki sifat kemiripan, sedangkan antar kelompok memiliki sifat ketidakmiripan. Analisis *cluster* sering digunakan untuk menganalisis kesamaan karakteristik antar wilayah. Analisis ini diawali dengan pemahaman bahwa sejumlah data tertentu mempunyai kemiripan diantara anggotanya, karena itu dimungkinkan untuk mengelompokkan anggota-anggota yang mirip atau mempunyai karakteristik yang serupa tersebut dalam satu atau lebih kelompok (Santoso, 2015). Dalam analisis *cluster* terdapat dua metode pengelompokkan, yaitu metode hierarki dan metode non-hierarki (Mattjik, 2011).

Analisis *cluster* dengan metode hierarki digunakan untuk mengelompokkan objek secara bertahap/terstruktur berdasarkan kemiripan sifatnya dan banyak kelompok belum diketahui. Sedangkan metode non-hierarki digunakan untuk pengelompokkan objek dimana dan biasanya digunakan untuk jumlah objek yang lebih besar, salah satu metode penggerombolan non-hierarki yaitu *k-means* (Mattjik,2011). Dalam metode hierarki terdapat beberapa metode yang sering digunakan dalam pengklasteran yaitu *single linkage*, *complete linkage*, *average linkage* dan metode *ward*.

Penelitian yang dilakukan oleh (Prasetyo et al., 2016) tentang pengelompokan kabupaten/kota di Provinsi Aceh menggunakan Analisis *Cluster*. Variabel yang digunakan adalah penduduk yang mengalami keluhan kesehatan, penduduk yang mengalami keluhan kesehatan dan berobat sendiri, penduduk yang mengalami keluhan kesehatan dan berobat jalan, penduduk yang berobat sendiri dan menggunakan obat tradisional, bayi lahir ditolong tenaga kesehatan, balita yang sudah imunisasi lengkap, balita yang diberi ASI tanpa makanan tambahan selama 0-4 bulan, rumah tangga yang menggunakan sumber air bersih, rumah tangga yang memiliki akses sanitasi layak, rumah tangga yang memiliki akses air minum, angka kematian bayi. Hasil analisis menghasilkan 3 klaster, yang terdiri dari 16 kabupaten/kota untuk klaster 1 dengan kualitas kesehatan cukup baik, 4 kabupaten/kota untuk klaster 2 dengan kualitas kesehatan kurang baik, dan 15 kabupaten/kota untuk klaster 3 dengan kualitas kesehatan baik.

Pada penelitian ini menggunakan analisis *cluster* dengan metode *average linkage*. Hal ini karena metode *average linkage* dianggap lebih stabil dibandingkan dengan metode *single linkage*, *complete linkage* dan metode *ward* (Mattjik, 2011). Pada penelitian ini variabel yang akan digunakan dalam penelitian ini adalah angka harapan hidup, angka morbiditas, persentase rumah tangga yang memiliki akses sanitasi layak, persentase rumah tangga yang memiliki akses terhadap layanan sumber air minum layak.

Berdasarkan latar belakang di atas, maka penelitian ini diberi judul **“Pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat Berdasarkan Indikator Kesehatan Masyarakat Tahun 2020 Menggunakan Analisis *Cluster*”**.

## **B. Batasan Masalah**

Berdasarkan latar belakang masalah yang telah dikemukakan sebelumnya penelitian ini difokuskan pada pengelompokan wilayah di Sumatera Barat berdasarkan indikator kesehatan tahun 2020, dimana indikator kesehatan adalah angka harapan hidup, angka morbiditas, persentase rumah tangga yang memiliki akses terhadap sanitasi layak, dan persentase rumah tangga yang memiliki akses terhadap layanan sumber air minum layak.

## **C. Pertanyaan Penelitian**

Pertanyaan penelitian ini adalah.

1. Bagaimana hasil pengelompokan kabupaten/kota di Sumatera Barat berdasarkan indikator kesehatan masyarakat menggunakan analisis *cluster*?
2. Bagaimana kategori dari kelompok kabupaten/kota di Sumatera Barat berdasarkan indikator kesehatan masyarakat?

## **D. Tujuan Penelitian**

Berdasarkan pertanyaan penelitian di atas, maka tujuan yang ingin dicapai dalam penelitian ini adalah.

1. Untuk mengetahui pengelompokan kabupaten/kota di Sumatera Barat berdasarkan indikator kesehatan masyarakat menggunakan analisis *Cluster*.
2. Untuk mengetahui kategori kelompok kabupaten/kota di Sumatera Barat berdasarkan indikator kesehatan masyarakat.

### **E. Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut.

1. Bagi penulis, dapat menambah wawasan dan memperdalam pengetahuan metode analisis *cluster*.
2. Bagi pemerintah, sebagai sumber informasi dan bahan pertimbangan dalam mengambil kebijakan mengenai pemerataan indikator kesehatan masyarakat di Provinsi Sumatera Barat.
3. Peneliti selanjutnya, sebagai bahan referensi dalam melakukan penelitian selanjutnya

## **BAB II**

### **KAJIAN TEORI**

#### **A. Kesehatan**

Menurut (Robert, 2017), kesehatan adalah sebuah sumber daya yang dimiliki semua manusia dan bukan merupakan suatu tujuan hidup yang perlu dicapai. Kesehatan tidak terfokus kepada fisik yang bugar meliputi jiwa yang sehat dimana individu dapat bersikap toleran dan dapat menerima perbedaan. Menurut Undang-Undang No.36 Tahun 2009 tentang kesehatan adalah keadaan baik secara fisik, mental, spiritual maupun sosial yang memungkinkan setiap orang untuk hidup produktif secara sosial dan ekonomis. Sehingga dari dua pengertian diatas penulis dapat menyimpulkan bahwa kesehatan adalah bukan hanya ketiadaan penyakit atau kelemahan melainkan keadaan yang seimbang dan dinamis antara fisik, mental, kesejahteraan sosial, fungsi tubuh dan berbagai faktor yang mempengaruhi.

#### **B. Indikator Kesehatan Masyarakat**

Indikator kesehatan masyarakat adalah ukuran yang digunakan untuk melihat apakah derajat kesahatan masyarakat sudah optimal, yang dilihat dari unsur kualitas hidup (Dinas Kesehatan Sumbar, 2019). Indikator kesehatan dapat digunakan untuk mengetahui kondisi kesehatan masyarakat di suatu wilayah, serta terdapat variabel-variabel yang dapat membentuk indikator kesehatan masyarakat di suatu wilayah.

Berikut adalah indikator kesehatan masyarakat di Provinsi Sumatera Barat.

##### **1. Angka Harapan Hidup**

Angka harapan hidup merupakan rata-rata perkiraan lamanya waktu (dalam tahun) yang dapat dijalani oleh seseorang selama hidupnya. Penghitungan angka

harapan hidup dilakukan melalui pendekatan tidaklangsung (*indirect estimation*). Data yang digunakan adalah Anak Lahir Hidup (ALH) dan Anak Masih Hidup (AMH) yang bersumber dari hasil Sensus Penduduk tahun 2020. Metode estimasi yang digunakan adalah metode *Trussel* dengan model *West*, yang sesuai dengan sejarah kependudukan dan kondisi Indonesia dan negara-negara Asia Tenggara umumnya.

## 2. Angka Morbiditas

Angka morbiditas adalah angka kesakitan, morbiditas menggambarkan kejadian penyakit dalam suatu populasi pada kurun waktu tertentu. Tingkat kesakitan suatu wilayah juga mencerminkan situasi derajat kesehatan masyarakat yang ada didalamnya (Dinas Kesehatan Sumbar, 2019). Rumus angka morbiditas adalah sebagai berikut:

$$AM = \left( \frac{JPKK}{JP} \right) \times 100$$

dengan:

$AM$  = Angka Morbiditas

$JPKK$  = Jumlah Penduduk yang Mengalami Keluhan Kesehatan dan Terganggunya Aktivitas

$JP$  = Jumlah Penduduk

## 3. Persentase Rumah Tangga yang Memiliki Akses Sanitasi Layak

Definisi sanitasi dari WHO merujuk kepada penyediaan sarana pelayanan pembuangan limbah kotoran manusia. Istilah sanitasi juga mengacu kepada pemeliharaan kondisi higienis kesehatan lingkungan yang mempengaruhi derajat kesehatan masyarakat. Buruknya kondisi sanitasi akan berdampak negatif di

banyak aspek kehidupan, mulai dari turunnya kualitas lingkungan hidup masyarakat, dan menimbulkan berbagai macam penyakit. Rumah tangga dikatakan memiliki akses terhadap layanan sanitasi layak apabila rumah tangga memiliki fasilitas tempat Buang Air Besar (BAB) yang digunakan sendiri atau bersama rumah tangga tertentu, menggunakan jenis kloset leher angsa, dan tempat pembuangan akhir tinja di tangki septik atau bisa juga dengan lubang tanah jika wilayah tempat tinggalnya di perdesaan (BPS Sumbar, 2020).

#### 4. Persentase Rumah Tangga yang Memiliki Layanan Sumber Air Minum Layak

Rumah tangga harus memiliki akses air minum layak dan bersih dalam mendukung kesehatan lingkungan dan kesehatan masyarakat. Kebutuhan air minum tidak hanya dilihat dari kuantitasnya tetapi juga dari kualitas air minum. Menurut BPS (Badan Pusat Statistik) yang dipublikasikan melalui Indikator Perumahan dan Kesehatan Lingkungan 2020, beberapa faktor yang berpengaruh terhadap akses air minum layak diantaranya adalah:

- a. Jenis sumber air utama yang digunakan oleh rumah tangga untuk diminum berasal dari air leding, sumur bor atau sumur pompa, sumur terlindung, mata air terlindung dan air hujan
- b. Jenis sumber air utama yang digunakan oleh rumah tangga berasal dari air kemasan bermerk atau air isi ulang namun sumber air utama untuk mandi, mencuci dan lain-lain yang digunakan berasal dari air leding, sumur bor atau sumur pompa, sumur terlindungi, mata air terlindungi, dan air hujan.

### C. Standardisasi Data

Menurut (Kassambara, 2017) standardisasi data dilakukan pada data yang memiliki perbedaan nilai yang cukup besar antar variabel. Perbedaan satuan yang besar akan menyebabkan perhitungan jarak pada analisis cluster menjadi tidak valid (Nugroho, 2008). Oleh sebab itu perlu dilakukan proses standardisasi data melalui transformasi pada data aktual kedalam bentuk *z-skor* dengan rumus sebagai berikut:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (1)$$

dengan:

$x_{ij}$  = nilai peubah baku ke-j untuk pengamatan ke-i

$\bar{x}_j$  = nilai rata-rata peubah ke-j

$s_j$  = standar deviasi peubah ke-j

(Kassambara, 2017)

### D. Analisis Cluster

#### 1. Pengertian Analisis Cluster

Analisis *cluster* merupakan analisis yang digunakan untuk mencari pola dalam gugus data dengan mengelompokkan objek atau pengamatan kedalam *cluster*. Tujuan analisis *cluster* ini yaitu untuk menemukan kelompok yang optimal dari objek atau pengamatan, dimana objek dalam setiap *cluster* memiliki karakteristik yang berbeda (Rencher & Christensen, 2012). Jadi, analisis *cluster* adalah analisis multivariat yang digunakan untuk mengelompokkan objek yang memiliki kesamaan karakteristik kedalam satu *cluster* dan objek yang memiliki karakteristik yang berbeda dimasukkan kedalam *cluster* lainnya, sehingga antar *cluster* memiliki

karakteristik yang tidak sama. Menurut (Hair, 2013) ciri-ciri *cluster* yang baik dari analisis *cluster* adalah:

- a. Homogenitas atau kesamaan yang tinggi antar objek dalam satu *cluster* (*Within-Cluster*)
  - b. Heterogenitas atau perbedaan yang tinggi antar *cluster* yang satu dengan *cluster* yang lain (*Between-Cluster*)
2. Ukuran Jarak Analisis Cluster

Menurut (Rencher & Christensen, 2012) untuk menggunakan analisis *cluster* dapat diidentifikasi dari objek yang sama dan mengelompokkan kedalam satu *cluster*, banyak teknik yang dapat mengukur kesamaan atau kedekatan antara setiap pasang dari objek. Pengukuran kedekatan jarak yang tepat antara dua objek, kemudian objek yang memiliki jarak yang jauh termasuk kedalam pengukuran dari ketidaksamaan. Terdapat beberapa macam ukuran jarak menunjukkan kemiripan antar objek. Terdapat beberapa macam ukuran jarak yang dapat digunakan, yaitu:

- a. Jarak *Euclid* (*Euclidean Distance*)

Jarak *Euclid* memiliki kelebihan yaitu jarak dari dua objek apa saja tidak dipengaruhi oleh penambahan dari objek baru untuk di analisis, yang mungkin merupakan pencilan (Mattjik, 2011).

Pengukuran jarak yang biasa digunakan dalam analisis *cluster* adalah jarak *Euclidean* dengan rumus sebagai berikut.

$$d_{ik} = \sqrt{\sum_{j=1}^p (x_{ij} - y_{kj})^2} \quad (2)$$

dimana:

$d_{ik}$  = Kuadrat jarak euclid antar objek x dan objek pada y

$x_{ij}$  = Nilai dari objek  $i$  pada variabel ke- $j$

$y_{kj}$  = Nilai dari Objek  $k$  pada variabel ke- $j$

(Rencher & Christensen, 2012)

b. Jarak *City-Block* (Manhattan)

Jarak ini memudahkan jarak rata-rata dimensi-dimensi secara menyilang. Ukuran jarak ini mendapatkan hasil yang mirip terhadap jarak *Euclidean* sederhana. Pada ukuran ini, efek dari perbedaan besar yang tunggal (pencilan) dibuang, karena tidak dikuadratkan. Jarak *city-block* ditulis sebagai berikut:

$$\sum_{i=1}^n |x_i - y_i|$$

(Mattjik, 2011)

c. Jarak *Mahalanobis*

Jarak *mahalanobis* dapat digunakan jika terdapat korelasi antar variabel. Jarak *mahalanobis* didefinisikan:

$$d(x, y) = \sqrt{(x - y)' \Sigma^{-1} (x - y)}$$

dengan  $\Sigma$  adalah matriks ragam-peragam gabungan antara X dan Y dan  $\Sigma$  diduga dengan menggunakan S, sehingga jarak yang dipakai dikenal dengan jarak *mahalanobis*, dengan formula:

$$d(x, y) = \sqrt{(x - y)' S^{-1} (x - y)}$$

dimana:

$d(x, y)$  = jarak antar objek kedua, objek x dan objek y

**S** = matriks ragam peragam

Untuk mengukur tingkat kemiripan antar cluster maka dilakukan perhitungan sebagai berikut :

$$S_{ik} = 100\left(1 - \left(\frac{d_{ik}}{d_{max}}\right)\right)$$

dimana

$S_{ik}$  = Kemiripan antara cluster  $i$  dan cluster  $k$

$d_{ik}$  = jarak antara cluster  $i$  dan  $k$

$d_{max}$  = nilai maksimum dari matriks jarak  $D$

### 3. Analisis *Cluster* Hierarki

Menurut (Kassambara, 2017) analisis *cluster* terbagi menjadi 2 bagian yaitu hierarki dan non hierarki. Metode *cluster* hierarki adalah pendekatan alternatif yang digunakan untuk mengelompokkan objek berdasarkan kesamaannya. Pada metode hierarki tidak memerlukan terlebih dahulu banyaknya kelompok atau *cluster* yang akan dihasilkan. Metode hierarki terbagi menjadi dua tipe yaitu *agglomerative* (penggabungan) dan *divisive* (pemecahan). Tipe *agglomerative* yaitu teknik pengelompokkan, dimana setiap objek atau pengamatan *cluster* yang memiliki kemiripan atau kesamaan digabungkan sampai menjadi satu *cluster* yang besar. Sedangkan tipe *divisive* yaitu teknik pengelompokkan, dimana semua objek berada dalam satu *cluster* besar, kemudian *cluster* yang memiliki karakteristik yang berbeda membentuk *cluster* yang lain sampai semua objek atau pengamatan berada pada *cluster* nya masing-masing. Kedua cara dalam metode berhierarki ini tidak berbeda dalam pembentukan kelompok yang terjadi tetapi hanya berbeda dalam tahapan pembentukan kelompok saja (Mattjik, 2011).

Analisis *cluster* bagian hierarki dengan metode *agglomerative* dibagi kedalam beberapa metode sebagai berikut.

a. *Single Linkage* atau Pautan Tunggal

Menurut (Johnson & Wichern, 2014) metode pautan tunggal dilakukan dengan meminimumkan jarak antar *cluster* yang digabungkan, atau dengan kata lain metode yang mengelompokkan dua objek yang mempunyai jarak terdekat terlebih dahulu. Jarak antar *cluster* di bentuk dari individu-individu dalam dua *cluster* yang mempunyai jarak terkecil atau kemiripan terbesar. Metode ini dimulai dengan menentukan jarak terkecil dalam matriks *proximity*  $D = \{d_{ik}\}$  dan gabungan objek yang memiliki jarak terkecil tersebut, misal  $U$  dan  $V$ , sehingga diperoleh *cluster*  $UV$ . Untuk mencari *cluster*  $UV$  dan  $W$  lainnya dapat diperoleh dengan cara sebagai berikut:

$$d_{(UV)W} = \min\{d_{UW}, d_{VW}\}$$

dimana  $d_{UW}$  adalah jarak terdekat dari *cluster*  $U$  dan  $W$  sedangkan  $d_{VW}$  merupakan jarak terdekat dari *cluster*  $V$  dan  $W$ .

(Johnson & Wichern, 2014)

b. *Complete Linkage* atau Pautan Lengkap

Metode ini hampir sama dengan metode pautan tunggal hanya saja metode pautan lengkap menggunakan jarak terjauh. Metode pautan lengkap adalah metode yang mengklasterkan dua objek yang mempunyai jarak terjauh terlebih dahulu. *Cluster* dimulai dengan mencari jarak pada matriks *proximity*  $D = \{d_{ik}\}$  dan penggabungan antara  $U$  dan  $V$  untuk mendapatkan *cluster* pertama atau  $UV$ . Selanjutnya jarak antara  $UV$  dan setiap *cluster* dihitung dengan:

$$d_{(UV)W} = \max\{d_{UW}, d_{VW}\}$$

dimana  $d_{UW}$  adalah jarak terjauh dari *cluster* U dan W sedangkan  $d_{VW}$  merupakan jarak terjauh dari *cluster* V dan W.

(Johnson & Wichern, 2014)

c. *Average Linkage* atau Pautan Rata-rata

Metode *average linkage* memperlakukan jarak antar dua *cluster* sebagai jarak antar rata-rata antara semua pasangan dari objek dimana salah satu anggota dari pasangan dimiliki oleh setiap *cluster*. Metode ini dimulai dengan mencari matriks jarak  $D = \{d_{ik}\}$  untuk memperoleh objek-objek paling dekat (yang memiliki kemiripan), misalnya U dan V. Objek-objek ini digabungkan sehingga akan terbentuk *cluster* W, dengan menggunakan rumus sebagai berikut:

$$d_{(UV)W} = \frac{\sum_{i=1}^n \sum_{k=1}^n d_{ik}}{N_{(UV)}N_W} \quad (3)$$

dimana:

$d_{(UV)W}$  = merupakan jarak antara *cluster* UV dengan *cluster* V

$d_{ik}$  = merupakan jarak antara objek  $i$  pada *cluster* (UV) dengan objek  $k$  pada *cluster*

$N_{(UV)}$  = merupakan jumlah objek pada *cluster* (UV)

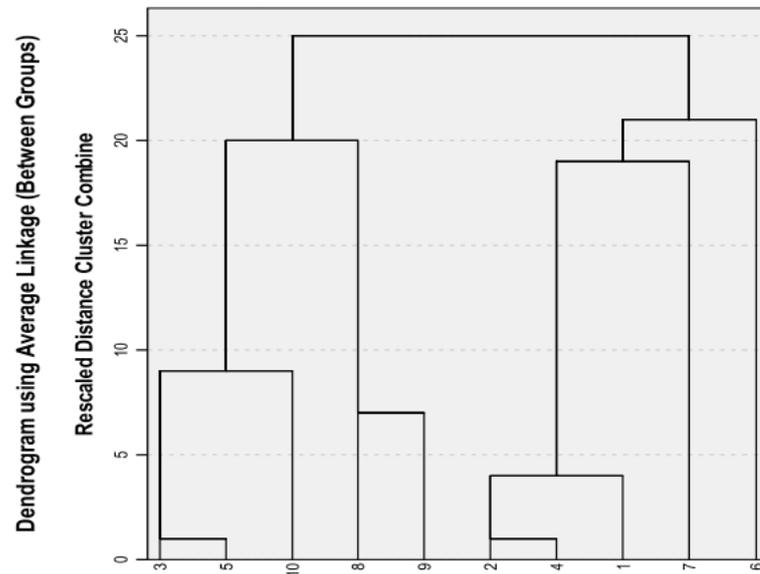
$N_W$  = merupakan jumlah objek pada *cluster* W

(Johnson & Wichern, 2014)

Menurut Yim & Ramdeen (2015), metode *average linkage* digunakan ketika metode *single linkage* dan *complete linkage* memiliki keterbatasan saat menentukan ukuran jarak dalam mengelompokkan karakteristik yang sama. Metode ini juga lebih stabil karena menghitung jarak berdasarkan rata-rata jarak dari dua metode *single linkage* dan *complete linkage* memiliki keterbatasan saat menentukan ukuran jarak dalam mengelompokkan karakteristik yang sama. Metode ini juga lebih stabil karena menghitung jarak berdasarkan rata-rata jarak dari dua objek.

#### 4. Dendogram

Dendogram merupakan salah satu bentuk dari diagram pohon yang memperhatikan hubungan antar objek. Tujuan dari pembuatan dendogram adalah untuk memberikan penelusuran pengelompokkan objek yang lebih mudah dan informatif. Menurut (Rencher & Christensen, 2012) pemisahan kelompok biasanya ditentukan berdasarkan selisih jarak penggabungan terbesar. Berikut adalah salah satu contoh dendogram dengan metode *average linkage*.



**Gambar 1. Contoh Dendrogram**

Menurut (Supranto, 2010) Sumbu tegak lurus (*vertical*) pada dendrogram menunjukkan objek pengamatan dan sumbu mendatar (*horizontal*) menunjukkan jarak (*distance*).

#### 5. Menginterpretasikan *cluster*

Tahap interpretasi *cluster* adalah untuk mencari karakteristik setiap kelompok dengan cara membandingkan rata-rata pada masing-masing kelompok. Selain itu juga dilakukan untuk menjelaskan karakteristik dari setiap kelompok tersebut (Rizki, 2013). Adapun cara menghitung rata-rata (*centroid*) yaitu:

$$C = \frac{\sum_{j=1}^n \sum_{k=1}^p x_{jk}}{N}$$

dimana:

$C$  = Nilai rata-rata (*centroid*) cluster

$x_{jk}$  = Nilai dari objek  $j$  pada variabel  $k$  dalam cluster yang dicari nilai *centroid*

$N$  = Jumlah objek dalam cluster yang dicari nilai *centroid*

## **BAB V PENUTUP**

### **A. KESIMPULAN**

Berdasarkan analisis yang telah dilakukan diperoleh kesimpulan sebagai berikut:

1. Indikator kesehatan menghasilkan 4 *cluster*. *Cluster* 1 yaitu Kota Padang, Kota Payakumbuh, Kota Solok, Kota Bukittinggi, Kota Sawah Lunto, Kota Pariaman, Kabupaten Agam, Kota Padang Panjang, Kabupaten Tanah Datar, Kabupaten Padang Pariaman, Kabupaten Pesisir Selatan, Kabupaten Dharmasraya, *cluster* 2 yaitu Kabupaten Solok, Kabupaten Lima Puluh Kota, Kabupaten Sijunjung, Kabupaten Pasaman Barat, Kabupaten Pasaman, *cluster* 3 yaitu Kabupaten Kepulauan Mentawai, dan *cluster* 4 yaitu Kabupaten Solok Selatan
2. Kategori yang diperoleh masing-masing cluster pada indikator kesehatan masyarakat:
  - a. *Cluster* 1 dikategorikan tinggi untuk variabel angka harapan hidup, persentase rumah tangga yang memiliki akses sanitasi layak, persentase rumah tangga yang memiliki akses terhadap layanan sumber air minum layak memiliki rata-rata yang tinggi, sedangkan untuk variabel angka morbiditas rata-ratanya lebih rendah.
  - b. *Cluster* 2 dikategorikan rendah untuk variabel angka harapan hidup, persentase rumah tangga yang memiliki akses terhadap layanan sumber air minum layak memiliki rata-rata yang tinggi sedangkan untuk variabel angka

morbiditas dan persentase rumah tangga yang memiliki akses sanitasi layak memiliki rata-rata yang rendah.

- c. *Cluster* 3 dikategorikan sangat rendah untuk variabel angka harapan hidup, persentase rumah tangga yang memiliki akses terhadap layanan sumber air minum layak memiliki rata-rata yang sangat rendah dibandingkan dengan *cluster* lainnya. Sedangkan untuk variabel angka morbiditas, dan persentase penduduk yang memiliki akses sanitasi layak memiliki rata-rata yang cukup tinggi.
- d. *Cluster* 4 dikategorikan sedang untuk variabel persentase penduduk yang memiliki akses sanitasi layak, dan persentase rumah tangga yang memiliki akses terhadap layanan sumber air minum layak memiliki rata-rata lebih rendah dari *cluster* 1 tetapi lebih tinggi dari *cluster* 2 dan 3.

## **B. SARAN**

Adapun saran dari penelitian ini sebagai berikut:

1. Setelah mengetahui hasil pengelompokan dan karakteristik masing-masing cluster, diharapkan pemerintah dapat memberikan perhatian dan memberikan upaya pemerataan terhadap indikator kesehatan masyarakat.
2. Bagi yang akan melakukan penelitian selanjutnya dapat menambah variabel lain yang mungkin berkaitan dengan indikator kesehatan masyarakat.

## DAFTAR PUSTAKA

- Badan Pusat Statistik. (2011). *Kajian Indikator Kesehatan (Laporan Sosial 2010)*. Jakarta: Badan Pusat Statistik.
- Badan Pusat Statistik. (2019). *Profil Kesehatan Provinsi Sumatera Barat 2019*. Padang: Badan Pusat Statistik.
- Badan Pusat Statistik. (2020). *Provinsi Sumatera Barat Dalam Angka 2020*.
- Dinas Kesehatan Provinsi Sumatera Barat. (2020). *Profil Kesehatan Provinsi Sumatera Barat 2020*. Padang: Dinas Kesehatan Provinsi Sumatera Barat.
- Hair, dkk. (2013). *Multivariate Data Analysis, 7<sup>th</sup> Edition*. New Jersey: Pearson Prentice Hall.
- Johnson, A.R dan Dean W.Wichern. (2007). *Applied Multivariat Statistica Analysis*. United States of America: Pearson Prentice Hall.
- Johnson, R, & Wichern, D. (2014). *Applied Multivariat Statistical Analysis*. 6th ed. United States of America : Pearson Education.
- Kassambara, A. (2017). *Practical Guide To Cluster Analysis in R*. 1st ed.sthd.
- Mattjik, A.A., Sumertajaya IM. (2011). *Sidik Peubah Ganda dengan Menggunakan SAS*. Wibawa GNA, Hadi AF, editor. Bogor (ID): IPB Press.
- Prasetyo, R. D., Suryowati, K., & Bekti, R. D. (2016). *Pengelompokan kabupaten/kota di Jawa Tengah berdasarkan Variabel Indikator Kesehatan Menggunakan Analisis Cluster*. *Jurnal Statistika Industri dan Komputasi*, 1(01), 70-79.
- Rencher, A. C., & Christensen, W.F. (2012). *Methods of Multivariate Analysis*. 3rd ed. New Jersey : John Wiley & Sons, Inc.
- Rizki, M.R.,dkk. (2013). *Analisis Cluster dalam Mengidentifikasi Tipe Kawasan Berdasarkan Karakteristik Timbunan Sampah Rumah Tangga di Perkotaan Kabupaten Jember*. *Jurnal Teknik POMITS*, Vol 2(1)
- Santoso, S. (2015). *Menguasai Statistik Multivariat*. Jakarta: PT. Elex Media Komputindo.
- Supranto, J. (2010). *Analisis Multivariate Arti dan Interpretasi*. Jakarta: Rineka Cipta.