

**ANALISIS *CLUSTER* COVID-19 DI SUMATERA BARAT
DENGAN METODE *NON-HIRARKI (K-MEANS)***

SKRIPSI

*Diajukan sebagai salah satu persyaratan guna memperoleh gelar
sarjana sains*



Oleh:

**AFRIDHO AFNANDA
NIM 17030035**

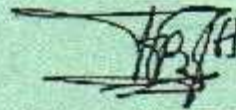
**PRODI MATEMATIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG**

2021

PERSETUJUAN SKRIPSI

Judul : Analisis *Cluster* COVID-19 di Sumatera Barat dengan Metode *Non-Hirarki (K-Means)*
Nama : Afridho Afnanda
NIM : 17030035
Program Studi : Matematika
Jurusan : Matematika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Padang, 31 Agustus 2021
Disetujui oleh,
Pembimbing



Dr. Arnellis, M.Si
NIP. 19610502 198703 2 002

HALAMAN PENGESAHAN LULUS UJIAN SKRIPSI

Nama : Afridho Afananda
NIM / TM : 17030035/2017
Program Studi : Matematika
Jurusan : Matematika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan Judul Skripsi

**ANALISIS *CLUSTER* COVID-19 DI SUMATERA BARAT
DENGAN METODE *NON-HIRARKI (K-MEANS)***

Dinyatakan lulus setelah dipertahankan di depan Tim Penguji Skripsi
Program Studi Matematika Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Padang

Padang, 31 Agustus 2021

Tim Penguji

	Nama
Ketua	: Dr. Arnellis, M.Si
Anggota	: Dra. Media Rosha, M.Si
Anggota	: Rara Sandhy Winanda, S.Pd., M.Sc

Tanda Tangan



SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan di bawah ini:

Nama : Afridho Afnanda
NIM : 17030035
Program Studi : Matematika
Jurusan : Matematika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan, bahwa skripsi saya dengan judul "**Analisis Cluster COVID-19 di Sumatera Barat dengan Metode Non-Hirarki (K-Means)**" 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 dan rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Padang, 31 Agustus 2021

Diketahui oleh,
Ketua Jurusan Matematika,



Dra. Media Rosha, M.Si
NIP. 19620815 198703 2 004

Saya yang menyatakan,



Afridho Afnanda
NIM. 17030035

ABSTRAK

Wabah virus corona baru SARSCoV-2 (penyakit coronavirus 2019; sebelumnya 2019nCoV), yang berpusat di Provinsi Hubei, Republik Rakyat Cina, telah menyebar ke banyak negara lain termasuk ke Indonesia. Januari 2020, Komite Darurat WHO mengumumkan keadaan darurat kesehatan global berdasarkan tingkat pemberitahuan kasus yang meningkat di Cina dan lokasi internasional. Di Indonesia sendiri virus ini pertama kali diumumkan pada tanggal 2 Maret 2020, Hingga 31 Maret 2020, terdapat 1,528 kasus COVID-19 yang dikonfirmasi di Indonesia dan 136 kematian terkait penyakit tersebut. Case fatality rate (CFR) negara ini juga jauh lebih tinggi daripada di Republik Rakyat Cina (8,9% vs 4%). Dalam upaya untuk menghentikan penyebaran COVID-19 pemerintah Indonesia membuat beberapa kebijakan yang diharapkan dapat menekan penyebaran virus ini. Pada akhir Maret 2020, presiden Indonesia akhirnya memutuskan untuk melaksanakan sosial berskala besar di kota dan provinsi, bukan karantina regional. Analisis klaster merupakan suatu kelas teknik, dipergunakan untuk mengklasifikasikan objek atau kasus kedalam kelompok yang relatif homogen (sama) yang disebut klaster (*cluster*).

Jenis data yang digunakan adalah data sekunder COVID-19 di kabupaten/kota Provinsi Sumatera Barat yang diperoleh dari Dinas Kesehatan Provinsi Sumatera Barat dan *website* gugus tugas COVID-19, dimana data yang diperoleh adalah data kumulatif dari bulan Maret 2020 sampai 14 Juli 2021. Pengoptimalan jumlah K cluster memakai metode elbow dengan menghasilkan K=3, hasil analisis cluster K-means dengan jumlah K=3 didapatkan kesimpulan bahwa jarak anggota dalam cluster yang rendah dan jarak antar cluster yang besar.

Kata Kunci: COVID-19, Non-Hierarchical Cluster Analysis, K-means, Elbow Method

ABSTRACT

The outbreak of the new coronavirus SARSCoV-2 (coronavirus disease 2019; previously 2019nCoV), which was centered in Hubei Province, People's Republic of China, has spread to many other countries including Indonesia. January 2020, the WHO Emergency Committee declared a global health emergency based on increasing case notification rates in China and international locations. In Indonesia alone this virus was first announced on March 2, 2020, Until March 31, 2020, there were 1,528 confirmed cases of COVID-19 in Indonesia and 136 deaths related to the disease. The country's case fatality rate (CFR) is also much higher than in the People's Republic of China (8.9% vs. 4%). In an effort to stop the spread of COVID-19, the Indonesian government has made several policies that are expected to suppress the spread of this virus. At the end of March 2020, the Indonesian president finally decided to implement large-scale social activities in cities and provinces, not regional quarantine. Cluster analysis is a technical class, used to classify objects or cases into relatively homogeneous (same) groups called clusters.

The type of data used is secondary data on COVID-19 in districts/cities of West Sumatra Province obtained from the West Sumatra Provincial Health Office and the COVID-19 task force website, where the data obtained is cumulative data from March 2020 to July 14, 2021. Optimization The number of K clusters using the elbow method produces $K = 3$, the results of the analysis of the K-means cluster with the number of $K = 3$, it is concluded that the distance of members in the cluster is low and the distance between clusters is large.

Keywords: COVID-19, Non-Hierarchical Cluster Analysis, K-means, Elbow Method

KATA PENGANTAR

Alhamdulillah puji syukur peneliti ucapkan kepada Allah SWT yang telah melimpahkan rahmat, taufik dan hidayah-Nya kepada peneliti sehingga peneliti dapat menyelesaikan penulisan tugas akhir dengan judul “**Analisis Cluster COVID-19 di Sumatera Barat dengan Metode *Non-Hirarki (K-Means)***”. Shalawat dan salam selalu tercurah terhadap junjungan nabi besar Muhammad SAW.

Penulisan skripsi ini bertujuan untuk memenuhi salah satu syarat dalam memperoleh gelar Sarjana Sains di Program Studi Matematika Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Negeri Padang (UNP). Dalam penulisan dan penyusunan skripsi ini peneliti banyak mendapatkan bantuan dari berbagai pihak, oleh sebab itu peneliti ingin mengucapkan terima kasih atas bantuan berupa dukungan, bimbingan, petunjuk, nasihat, arahan dan kerja sama dari berbagai pihak, yaitu kepada:

1. Ibu Dr. Arnellis, M.Si selaku Dosen Pembimbing dan Panesehat Akademik.
2. Ibu Dra. Media Rosha, M.Si dan Rara Sandhy Winanda, S.Pd, M.Si selaku Dosen Penguji
3. Ibu Dra. Media Rosha, M.Si selaku Ketua Prodi Matematika dan Ketua Jurusan Matematika
4. Bapak dan Ibu staf pengajar serta karyawan Jurusan Matematika FMIPA UNP

5. Orang tua yang selalu memberikana dukungan, motivasi dan doa restunya.
6. Sahabat, rekan-rekan Matematika angkatan 2017.
7. Semua pihak yang telah membantu selama studi dan penyelesaian skripsi ini yang tidak disebutkan satu persatu.

Semoga dukungan, bimbingan, doa, dan kerjasamanya dibalas oleh Allah SWT sebagai amal kebaikan dan bernilai ibadah, Aamiin Allahumma Aamiin.

Dalam penulisan skripsi ini masih jauh dari kata sempurna. Untuk itu peneliti memohon maaf kepada pembaca. Kritik, Saran dan masukan yang membangun dari pembaca sangat diharapkan untuk kesempurnaan tugas akhir ini. Semoga dapat bermanfaat dan dipergunakan sebagaimana mestinya serta dapat menjadi pedoman bagi peneliti selanjutnya. Atas perhatiannya penulis ucapkan terima kasih.

Padang, Juli 2021

Peneliti

DAFTAR ISI

ABSTRAK.....	i
KATA PENGANTAR.....	iii
DAFTAR ISI.....	v
DAFTAR GAMBAR.....	vii
DAFTAR TABEL	viii
DAFTAR LAMPIRAN	ix
A. Latar Belakang Masalah	1
B. Rumusan masalah	3
C. Pertanyaan penelitian.....	4
D. Tujuan Penelitian	4
E. Manfaat Penelitian.....	4
BAB II TEORI PENDUKUNG	5
A. COVID-19.....	5
B. Statistika deskriptif	6
1. Mengumpulkan data.....	7
2. Mengolah data	7
3. Menyajikan data	7
4. Analisis sederhana	7
C. Statistika Inferensia.....	9
D. Analisis Multivariat	9
1. metode dependensi	9
2. Metode independensi.....	10
E. Peubah baku dan Matriks Korelasi	10
F. Analisis Komponen Utama	12
G. Analisis kluster (<i>cluster</i>)	13
1. Merumuskan masalah.....	15
2. Memilih ukuran jarak atau similaritas	15
3. Memilih suatu prosedur pengklasteran.....	19
4. Menentukan banyaknya kluster.....	22
5. Menginterpretasikan dan memprofil kluster.....	23

6. Mengakses validitas kluster	24
BAB III METODE PENELITIAN.....	25
A. Jenis Penelitian.....	25
B. Jenis Data dan Sumber Data	25
C. Objek dan Variabel Pengamatan	25
D. Teknik Analisis Data	26
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	28
A. Deskripsi Data	28
B. Hasil Analisis data	30
1. Karakteristik Data.....	30
2. Pengelompokan Cluster	31
a Standardisasi data.....	31
b Korelasi Antar Variabel	31
c Analisis Komponen Utama (AKU).....	32
d Optimasi K Cluster dengan metode Elbow	34
e Analisis Cluster K-Means.....	36
C. Pembahasan.....	38
BAB V PENUTUP	41
A. Kesimpulan.....	41
B. Saran	43
Daftar Pustaka.....	44

DAFTAR GAMBAR

Gambar 1. Prosedur analisis karakter	14
<i>Gambar 2. Prosedur pengklasteran.....</i>	<i>19</i>
<i>Gambar 3. Plot metode Elbow</i>	<i>35</i>

DAFTAR TABEL

<i>Tabel 1. Kasus penyebaran COVID-19 di Provinsi Sumatera Barat sampai dengan tanggal 14 Juli 2021</i>	<i>28</i>
<i>Tabel 2. Variabel Penelitian</i>	<i>30</i>
<i>Tabel 3. Nilai korelasi dan P-Value antar variabel</i>	<i>31</i>
<i>Tabel 4. Nilai keragaman dan proporsi masing-masing variabel</i>	<i>33</i>
<i>Tabel 5. Nilai SSE setiap K</i>	<i>34</i>
<i>Tabel 6. Jumlah Pengamatan cluster</i>	<i>36</i>
<i>Tabel 7. Pengelompokkan Kabupaten/Kota di Sumatera Barat berdasarkan penyebaran COVID-19</i>	<i>37</i>

DAFTAR LAMPIRAN

Lampiran 1. Data Awal Penelitian	45
Lampiran 2. Data hasil Standardisasi	46
Lampiran 3. Pemeriksaan Korelasi	47
Lampiran 4. Output Analisis Komponen Utama	48
Lampiran 5. Score Analisis Komponen Utama	49
Lampiran 6. Plot hasil analisis utama.....	50
Lampiran 7. Pemeriksaan Korelasi Ke-2.....	53
Lampiran 8. Elbow Plot Dan Nilai Sum Square Of Error	54
Lampiran 9. Output Analisis Klaster K-Means.....	55
Lampiran 10. Rata-Rata Setiap Variabel	56

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

COVID-19 telah menyebar dan menjadi pandemi di hampir seluruh dunia termasuk di Indonesia. Di Indonesia sendiri virus ini pertama kali diumumkan pada tanggal 2 Maret 2020, sampai pada tanggal 14 sampai 27 Maret 2020 total kasus yang teridentifikasi sebanyak 77.261 kasus. Dalam upaya untuk menghentikan penyebaran COVID-19 pemerintah Indonesia membuat beberapa kebijakan yang diharapkan dapat menekan penyebaran virus ini. Ada beberapa kebijakan yang disahkan oleh pemerintah diantaranya adalah pembatasan sosial bersekala besar atau disingkat PSBB.

Kebijakan PSBB (Pembatasan sosial bersekala besar) merupakan salah satu upaya pemerintah untuk menghadang laju penyebaran COVID-19 di Indonesia. Pelaksanaan PSBB telah diterapkan di Indonesia sejak tanggal 2 Maret 2020. Kebijakan pembatasan sosial berskala besar ini diterapkan hampir keseluruhan daerah di Indonesia termasuk di provinsi Sumatera Barat.

Kasus covid-19 di Sumatera Barat pada tanggal 20 April 2020 mencapai 74 kasus teridentifikasi dan menempatkan Sumatera Barat menjadi peringkat 15 besar jumlah kasus penyebaran COVID-19 di Indonesia. Kasus ini menyebar di 19 kabupaten/kota yang ada di Sumatera Barat, dan menunjukkan kurva peningkatan kasus positif setiap harinya. Semenjak April untuk pertama kalinya salah satu daerah di Sumatera Barat yaitu kabupaten 50 kota berubah status menjadi zona merah, sekitar 9 daerah yang berada di zona *orange* dan 9 daerah lainnya berada di zona kuning. Melihat hal tersebut perlu adanya kebijakan yang tepat untuk menghambat laju penyebaran COVID-19.

Penerapan PSBB dimasing-masing daerah yang ada di Indonesia tentu berbeda-beda tergantung kebijakan pada daerah masing-masing dan sesuai syarat yang harus dipenuhi oleh suatu daerah seperti: jumlah kasus dan/atau jumlah kematian akibat penyakit meningkat dan menyebar secara signifikan dan cepat kebeberapa wilayah, dan teradapat kaitan epidemiologis dengan kebijakan serupa diwilayah atau negara lain.

Dengan adanya perbedaan kebijakan disetiap daerah dalam menerapkan kebijakan PSBB, maka akan ada perbedaan penanganan yang dilakukan dalam upaya penerapan kebijakan. Hal ini akan mempengaruhi cepat dan lambatnya penyebaran terjadi dari daerah tingkat penyebaran tinggi ke daerah yang tingkat penyebaran rendah. Kebijakan yang berlaku dilakukan sesuai dengan keadaan epidemiologis daerah masing-masing maka kebijakan didaerah dengan penyebaran yang relatif tinggi akan berbeda dengan kebijakan dan penerapannya pada daerah yang penyebaran yang relatif rendah, sehingga perlu dikolompokkan daerah mana saja tingkat penyebaran COVID-19 yang tinggi dan yang relatif rendah, untuk itu perlu dilakukan pengelompokan.

Penelitian ini merupakan penelitian pengembangan yang relevan dengan penelitian terdahulu yaitu penelitian dari Rizkiana Prima R, Yashintia Arien E, dan Sutikno yang berjudul Analisis Cluster virus Corona (COVID-19) di Indonesia pada 2 Maret 2020-12 April 2020 dengan Metode *K-Means Clustering* metode ini juga digunakan dalam pengelompokan COVID-19 di Sumatera Barat dimana data yang digunakan, cakupan dan pemilihan jarak *clustering* yang berbeda.

Untuk mengenali keadaan daerah yang terdampak COVID-19 di setiap kabupaten/kota yang ada di Sumatera Barat, maka dilakukan pengklasifikasian berdasarkan kondisi tingkat penyebaran COVID-19 di setiap kabupaten/kota yang ada di Sumatera Barat, pengelompokan pada penelitian ini berdasarkan persentase penyebaran COVID-19 dapat dilakukan dengan menggunakan analisis *clustering*.

Pengelompokan ini dilakukan untuk mengenali bagaimana tingkat penyebaran COVID-19 di setiap kabupaten/kota di provinsi Sumatera Barat, hal ini dapat dijadikan rekomendasi dan acuan bagi pembentuk kebijakan di setiap kabupaten/kota untuk menyesuaikan kebijakan atau melihat tingkat penyebaran guna membuat keefektifan kebijakan dalam menangani penyebaran COVID-19. Program PSBB di setiap kabupaten/kota memiliki efektifitas yang berbeda-beda, hal ini dikarenakan karakteristik dari penerapan kebijakan PSBB ini berbeda di setiap kabupaten/kota yang ada di Sumatera Barat. Dengan melihat penyebaran COVID-19 di Sumatera Barat akan diperoleh data yang cenderung lebih banyak sehingga peneliti menggunakan metode *K-Means* yang bisa mengolah data yang lebih banyak dan juga mempercepat mendapatkan hasil analisisnya, juga peneliti dapat menentukan jumlah cluster yang diinginkan. Oleh karena itu penulis tertarik dan ingin melakukan penelitian terkait masalah yang telah dijabarkan dengan judul: “**Analisis Cluster COVID-19 di Sumatera Barat dengan Metode Non-Hirarki (K-Means)**”.

B. Rumusan masalah

Berdasarkan latar belakang, peneliti mengemukakan rumusan masalah yaitu “bagaimana pengelompokan COVID-19 di Sumatera Barat dengan Metode *Non-Hirarki (K-Means)*”.

C. Pertanyaan penelitian

1. Apa kesamaan karakteristik data pasien COVID-19 di setiap kabupaten/ kota di Sumatera Barat.
2. Bagaimana pengelompokan kabupaten/ kota pada penyebaran COVID-19 di Sumatera Barat
3. Apa ciri khas dari tiap-tiap *cluster* yang telah dikelompokan.

D. Tujuan Penelitian

Berdasarkan rumusan masalah yang peneliti ajukan, maka penelitian ini ditujukan untuk :

1. Mengetahui kesamaan karakteristik data pasien COVID-19 di setiap kabupaten/ kota di Sumatera Barat.
2. Mengetahui pengelompokan kabupaten/ kota untuk penyebaran COVID-19 di Sumatera Barat.
3. Mengetahui ciri khas dari tiap-tiap *cluster* yang telah dikelompokan.

E. Manfaat Penelitian

1. Menambah pengetahuan, wawasan, dan pemahaman peneliti tentang penggunaan Metode non-hirarki (*K-Means*) dalam analisis *cluster*.
2. Sebagai masukan dan bahan pertimbangan bagi pemerintah kabupaten/kota di Sumatera Barat dalam menentukan kebijakan menangani penyebaran COVID-19.
3. Acuan dan bahan referensi terhadap ilmu pengetahuan dalam mengembangkan dan memperluas cakupan hasil penelitian.