

**ANALISIS KLASSTER DENGAN METODE *K-MEANS* PADA
PENYEBARAN KASUS COVID-19 BERDASARKAN
KABUPATEN/KOTA DI SUMATERA UTARA**

SKRIPSI

SITI RAMADHANI NASUTION

NIM. 0703163047



**UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN**

**UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN
PROGRAM STUDI MATEMATIKA**

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUMATERA UTARA

MEDAN

2023

**ANALISIS KLASTER DENGAN METODE *K-MEANS* PADA
PENYEBARAN KASUS COVID-19 BERDASARKAN
KABUPATEN/KOTA DI SUMATERA UTARA**

SKRIPSI

*Diajukan untuk memenuhi Syarat Gelar Sarjana Matematika (S.Mat)
Dalam Sains dan Teknologi*

SITI RAMADHANI NASUTION

NIM. 0703163047




**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA
MEDAN
2023**

PENGESAHAN SKRIPSI

Nomor : B.160/ST/ST.V.2/PP.01.1/05/2023

Judul : Analisis Klaster Dengan Meode K-Means Pada
Penyebaran Kasus CoViD-19 Berdasarkan
Kabupaten/Kota Di Sumatera Utara
Nama : Siti Ramadhani Nasution
NIM : 070163047
Fakultas : Sains dan Teknologi
Telah dipertahankan dihadapan Dewan Penguji Skripsi Program Studi
Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Sumatera
Utara Medan dan dinyatakan **LULUS**.
Pada hari/Tanggal : Senin , 24 Agustus 2022
Tempat : Ruang Sidang Fakultas Sains dan Teknologi

Tim Ujian Munaqasyah
Ketua,



Dr.Riri Syafitri Lubis, S.Pd, M.Si
NIDN. 2013078402

Dewan Penguji,

Penguji I,


Dr. Rina Filia Sari, M.Si
NIDN.2001037703

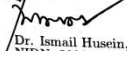
Penguji III,


Rima Aprilia, M.Si
NIDN. 2000130048801


Penguji II,


Rina Widyasari, M.Si
NIDN. 0118078801

Penguji IV,


Dr. Ismail Husein, M.Si
NIDN. 2022049101

Mengesahkan,
Dekan Fakultas Sains dan Teknologi
Universitas Islam Negeri Sumatera Utara Medan


Prof. Dr. Mhd. Syahnan, MA
NIP. 196609051991031002

v **PERSETUJUAN SKRIPSI**

Hal : Surat Persetujuan Skripsi

Lamp : -

Kepada Yth.,

Dekan Fakultas Sains dan Teknologi

Universitas Islam Negeri Sumatera Utara

Medan

Assalamu'alaykum Wr.Wb.

Setelah membaca, meneliti, memberikan petunjuk, dan mengoreksi serta mengadakan perbaikan, maka kami selaku pembimbing berpendapat bahwa skripsi saudara:

Nama : Siti Ramadhani Nasution

NIM : 0703163047

Program Studi : Matematika

Judul : Analisis Klaster Dengan Metode *K-Means* Pada Penyebaran Kasus CoViD-19 Berdasarkan Kabupaten/Kota Di Sumatera Utara

Dapat disetujui untuk segera dimunaqasyahkan. Atas perhatian kami ucapkan terimakasih

Disetujui oleh:

Pembimbing I



Dr. Rina Filia Sari, M.Si

NIDN. 2001037703

UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

Pembimbing II



Rina Widyasari, M.Si

NIDN. 0118078801

Mengetahui,

Ketua Program Studi Matematika

Fakultas Sains dan Teknologi

Universitas Islam Negeri Sumatera Utara Medan

Dr.

Dr. Riri Syafitri, S.Pd. M.Si

NIDN. 2013078402

SURAT PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan dibawah ini:

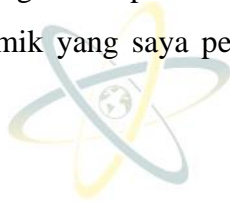
Nama : Siti Ramadhani Nasution

NIM : 0703163047

Prodi : Matematika

Judul : Analisis Kluster Dengan Metode *K-Means* Pada Penyebaran Kasus
CoViD-19 Berdasarkan Kabupaten/Kota Di SumateraUtara

Menyatakan bahwa skripsi ini adalah hasil karya saya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya. Apabila dikemudian hari ditemukan plagiat skripsi ini maka saya bersedia menerima sanksi pencabutan gelar akademik yang saya peroleh dan sanksi lainnya sesuai dengan peraturan yang berlaku.



Hormat Saya,



Siti Ramadhani Nasution

NIM. 0703163047

UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

KATA PENGANTAR

Assalamualaikum Wr. Wb

Dengan mengucapkan Syukur Alhamdulillah atas rahmat Allah SWT yang telah memberikan hidayah-Nya , sehingga penulis dapat menyelesaikan proposal skripsi yang berjudul **“Analisis Klaster Dengan Metode K-Means Pada Penyebaran Kasus CoViD-19 Berdasarkan Kabupaten/Kota Di Sumatera Utara”**. Selawat dan salam senantiasa tercurah kepada Nabi Muhammad *Shalallaahu Alaihi Wassalaam* yang diutus sebagai rahmat untuk sekalian alam serta keluarga dan sahabat beliau serta orang-orang yang berpegang teguh dengan petunjuk beliau hingga hari akhir.

Proposal skripsi ini disusun guna melengkapi persyaratan untuk dapat menyelesaikan pendidikan Strata Satu (S1) Fakultas Sains Dan Teknologi Jurusan Matematika Universitas Islam Negeri Sumatera Utara Medan. Penulis menyadari dalam penyusunan proposal skripsi ini tidak akan selesai tanpa bantuan dari berbagai pihak, untuk itu penulis ingin mengucapkan terimakasih yang sebesar-besarnya kepada:

1. Kepada yang tercinta Ayahanda Alm. Muhammad Zainuddin Nasution dan Ibunda Seriati yang telah membesarkan, mendidik, membimbing, melindungi, memberikan semangat yang tinggi, dan selalu memberikan dukungan kepada penulis, motivasi untuk terus berkarya dan adik-adik saya yang selalu menjadi penyemangat.
2. Bapak Prof. Dr. Syahrin Harahap, M.A., selaku Rektor UIN Sumatera Utara Medan
3. Bapak Dr. Mhd. Syahnun, M.A., selaku Dekan Fakultas Sains dan Teknologi UIN Sumatera Utara Medan.
4. Ibu Dr. Riri Syafitri Lubis, S.Pd., M.Si., selaku Ketua Program Studi Fakultas Sains dan Teknologi UIN Sumatera Utara Medan.
5. Ibu Rima Aprilia, M.Si., selaku Sekretaris Program Studi Matematika dan yang telah membantu mengarahkan dan membimbing penulis selama

6. menjalankan perkuliahan di Fakultas Sains dan Teknologi UIN Sumatera Utara.
7. Bapak Hendra Cipta, M.Si., selaku Penasehat Akademik yang telah membantu mengarahkan dan membimbing penulis selama menjalankan perkuliahan di Fakultas Sains dan Teknologi UIN Sumatera Utara.
8. IbuDr. Rina Filia Sari, M.Si., selaku Dosen Pembimbing I yang telah memberikan bimbingan dan arahan selama pelaksanaan penulisan skripsi ini. Sehingga penulis dapat menyelesaikan proposal ini dengan baik.
9. Ibu Rina Widyasari, M.Si., selaku Dosen Pembimbing II yang telah memberikan banyak arahan dan bimbingan kepada penulis dalam menyelesaikan skripsi ini.
10. Bapak/Ibu Dosen dan para staff pengajar di UIN Sumatera Utara Medan yang telah memberikan pendidikan dan pengetahuan kepada penulis.
11. Seluruh teman-teman Jurusan Matematika Stambuk 2016 yang telah membantu dan mendukung penulis dalam menyelesaikan skripsi ini

Kepada semua pihak yang membantu penyusunan proposal skripsi ini, penulis mengucapkan terima kasih dan hanya Allah SWT yang dapat memberikan balasan yang setimpal atas jasa dan bantuan yang telah diberikan kepada penulis. Penulis menyadari skripsi ini tidak luput dari kesalahan, dari itu penulis mengharapkan kritik dan saran demi kesempurnaannya sehingga skripsi ini dapat bermanfaat bagi semua pihak yang membaca dan memperluas cakrawala pemikiran dimasa yang akan datang. Akhir kata penulis mengucapkan terima kasih.

Wassalamualaikum Wr. Wb

Medan, Juni 2023

Penulis,

Siti Ramadhani Nasution

NIM.0703163047

DAFTAR ISI

PERSETUJUAN SKRIPSI	ii
PENGESAHAN SKRIPSI	iii
SURAT PERNYATAAN KEASLIAN SKRIPSI	iv
KATA PENGANTAR	v
DAFTAR ISI	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
ABSTRAK	xi
ABSTRACT	xii
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian	4
BAB II LANDASAN TEORI	
2.1 Jumlah Penduduk	5
2.2 Kepadatan Penduduk	5
2.3 Corona Virus <i>Desease</i> 2019 (CoViD-19)	5
2.3.1 Positif CoViD -19	6
2.3.2 Pasien yang sembuh.....	7
2.3.3 Pasien yang meninggal	7
2.4 Analisis Data	8
2.5 Analisis Multivariat	10
2.6 Analisis Kluster	10
2.6.1 Asumsi analisis kluster	11
2.7 Data Mining.....	13
2.7.1 <i>Clustering</i>	13
2.7.1.1 <i>K-Means</i>	14

2.7.2.2 Kelebihan dan Kelemahan <i>K-Means</i>	16
2.8 Penelitian Terdahulu	16
2.9 <i>Wahdaul Ulum</i>	16

BAB III METODOLOGI PENELITIAN

3.1 Tempat dan Waktu Penelitian	18
3.2 Jenis Penelitian dan Sumber Data	18
3.3 Variabel Penelitian	19
3.4 Analisis Data	20
3.5 Prosedur Penelitian.....	20
3.6 Diagram Alir.....	22

BAB IV PEMBAHASAN

4.1 Penyebaran CoViD-19 di Sumatera Utara	23
4.2 Analisis Deskriptif.....	25
4.3 Uji Asumsi Klaster	27
4.4 Jumlah Klaster	30
4.5 Menentukan Titik Pusat atau <i>Centroid</i> Secara Acak.....	30
4.6 <i>K-MeansClustering</i>	31
4.7 Tingkat Signifikansi Antar Klaster.....	40
4.8 Pembahasan.....	45

BAB V HASIL DAN KESIMPULAN

5.1 Kesimpulan	51
5.2 Saran.....	51

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR GAMBAR

- Gambar 3.1 Diagram Alir Analisis Klaster dengan Metode *K-Means*.....22
- Gambar 4.1 Dendogram *Clustering* CoViD-19 Sumatera Utara.....47



DAFTAR TABEL

Tabel 4.1 Data Penyebaran CoViD-19 di Sumatera Utara.....	23
Tabel 4.2 Deskriptif Data.....	25
Tabel 4.3 Hasil Uji KMO.....	26
Tabel 4.4 Hasil Uji Multikolinioritas.....	27
Tabel 4.5 <i>Centroid</i> Awal Penyebaran CoViD-19.....	31
Tabel 4.6 Perhitungan Jarak Pada Iterasi Pertama.....	32
Tabel 4.7 Perhitungan Jarak dan Klaster Pada Iterasi Pertama.....	34
Tabel 4.8 <i>Centroid</i> Baru Iterasi Pertama.....	32
Tabel 4.9 Perhitungan Jarak Pada Iterasi kedua.....	37
Tabel 4.10 Perhitungan Jarak dan Klaster Pada Iterasi kedua.....	38
Tabel 4.11 <i>Centroid</i> Baru Iterasi kedua.....	40
Tabel 4.12 ANOVA.....	41
Tabel 4.13 <i>Centroid</i> Pada Tiap Iterasi.....	46



ABSTRAK

Analisis kluster adalah teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek/*cases* berdasarkan karakteristik yang dimilikinya. Salah satu metode dalam analisis kluster adalah metode *K-Means*. Algoritma *K-Means* mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama antara satu dengan yang lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain. Kasus CoViD-19 di Indonesia mengalami peningkatan terkhususnya Sumatera Utara maka digunakan *K-Means clustering* untuk mengelompokkan daerah-daerah dengan tingkat penyebaran kasus CoViD-19 dengan tingkat penyebaran tinggi, sedang dan rendah. Adapun variabel yang diteliti dalam penelitian ini adalah X_1 = Jumlah Penduduk, X_2 = Kepadatan Penduduk, X_3 = positif CoViD-19, X_4 = pasien yang sembuh, X_5 = pasien yang meninggal berdasarkan Kabupaten/Kota yang ada di Sumatera Utara. Setelah dilakukan pengelompokan dengan menggunakan *K-Means Clustering* dengan total jumlah penduduk 14521471 dan kepadatan penduduk 35171,5 Orang/km² diperoleh Sebanyak 2 daerah yang termasuk kluster 1, 4 daerah yang termasuk kluster 2 dan 27 daerah yang termasuk kluster 3. Hasil kluster yang didapat berdasarkan uji ANOVA menunjukkan bahwa semua variabel yang digunakan mempunyai perbedaan yang signifikan pada ketiga kluster yang terbentuk dan variabel yang memiliki perbedaan paling besar pada ketiga yang terbentuk variabel jumlah penduduk adalah variabel yang paling membedakan anggota dari ketiga karena mempunyai nilai F terbesar diantara variabel yang lain yaitu 182,681 dan nilai signifikan 0,000.

Kata Kunci : CoViD-19, Analisis Kluster, *K-Means*

ABSTRACT

Cluster analysis is a multivariate technique that has the main objective of grouping objects/cases based on their characteristics. One of the methods in cluster analysis is the K-Means method. The K-Means algorithm groups the existing data into several groups, where the data in one group has the same characteristics as one another and has different characteristics from the data in the other group. The spread of CoViD-19 cases in Indonesia, especially North Sumatra, K Means clustering was used to group areas with a high, medium and low distribution rate of CoViD-19 cases. The variables studied in this study were $X_1 = \text{Total Population}$, $X_2 = \text{Population Density}$, $X_3 = \text{CoViD-19 positive}$, $X_4 = \text{patients who recovered}$, $X_5 = \text{patients who died based on districts/cities in North Sumatra}$. After grouping using K-Means Clustering with a total population of 14521471 and a population density of 35171.5 people/km², 2 areas are included in cluster 1 (low CoVid-19 case distribution areas), 4 areas including cluster 2 (CoVid case distribution areas). -19 moderate) and 27 areas included in cluster 3 (high CoVid-19 case distribution areas). The cluster results obtained based on the ANOVA test show that all the variables used have significant differences in the three clusters formed and the variable that has the greatest difference in the three clusters formed is the total population variable which is the variable that most distinguishes members of the three clusters because it has a value. The largest F among other variables is 182,681 and a significant value of 0.000.

Keywords:CoVid-19, Cluster Analysis, K Means