

BAB IV

HASIL DAN PEMBAHASAN

4.1 Penyebaran CoViD-19 Sumatera Utara

Data *clustering* menggunakan metode *K-Means*. *Clustering* ini dilakukan dengan menggunakan data CoViD-19 Sumatera Utara pada 24 Juni 2022. Data sekunder dari Dinas Kesehatan Provinsi Sumatera Utara yang beralamat di Jl. Prof. H.M Yamin SH. No. 41AA, Perintis, Kec. Medan Timur, Kota Medan Sumatera Utara 20232. Adapun data CoViD-19 Sumatera Utara pada 24 Juni 2022 adalah sebagai berikut:

Tabel 4.1 Data Penyebaran CoViD-19 Sumatera Utara

No	Kabupaten/Kota	Penduduk (Orang)	Kepadatan Penduduk (Orang/km ²)	Positif (Orang)	Sembuh (Orang)	Meninggal (Orang)
1	Asahan	769960	207,97	1718	1688	30
2	Batu Bara	410678	445,32	931	923	8
3	Dairi	30874	160,16	2963	2856	107
4	Deli Serdang	1931441	861,60	8806	8592	214
5	Humbang Hasundutan	197751	84,68	3615	3234	33
6	Karo	404998	190,41	1504	1466	58
7	Labuhanbatu	493899	229,08	2321	2237	84
8	Labuhanbatu Selatan	314094	87,35	1438	1367	71
9	Labuhanbatu Utara	381994	106,97	1773	1694	79
10	Langkat	1030202	164,52	1573	1511	49
11	Mandailing Natal	472886	77,09	1164	1119	45
12	Nias	146672	79,60	12	8	0

13	Nias Barat	89994	189,97	0	0	0
14	Nias Selatan	360531	197,53	0	0	0
15	Nias Utara	147274	122,44	37	33	0
16	Padang Lawas Utara	260720	66,54	513	484	29
17	Padang Lawas	261011	67,05	121	118	3
18	Pakpak Barat	52351	42,97	557	524	33
19	Samosir	136441	65,94	1849	1815	34
20	Serdang Bedagai	657490	346,01	3025	2769	256
21	Simalungun	990246	226,65	0	0	0
22	Tapanuli Selatan	300911	49,90	1935	1890	45
23	Tapanuli Tengah	365177	166,90	1179	1161	18
24	Tapanuli Utara	312758	82,49	5001	4921	80
25	Toba	206199	88,54	53	48	1
26	Binjai	291842	4930,60	25733	25590	143
27	Gunung Sitoli	136017	484,43	115	112	3
28	Medan	2435252	9189,63	72961	71929	1019
29	Padang Sidimpuan	225105	1963,24	519	516	3
30	Pematang Siantar	268254	4819,51	2550	2517	33
31	Sibolga	89584	2168,58	0	0	0
32	Tanjung Balai	176027	1632,45	614	585	29
33	Tebing Tinggi	172838	5575,42	1573	1498	75
	Total	14521471	35171,5	146153	143205	2582

Berdasarkan data diatas dilakukan *clustering* terhadap kasus CoViD-19.

Data yang ingin diproses dalam penelitian ini adalah data kasus CoViD-19 yang

meliputi X_1 adalah data positif CoViD-19, X_2 adalah data orang yang sembuh dari CoViD-19, X_3 adalah data orang meninggal karena CoViD-19 di Sumatera Utara. Data yang digunakan pada proses *clustering* di tabel 4.1 mempunyai satuan yang sama. Berdasarkan hal tersebut data dapat langung digunakan dan tidak perlu dilakukan standarisasi data.

4.2 Analisis Deskriptif

Analisis deskriptif pada penelitian ini digunakan untuk mengetahui gambaran secara umum mengenai data penyebaran CoViD-19. Analisis statistik deskriptif dilakukan guna mengetahui deskripsi suatu variabel yang dilihat dari nilai maksimum, nilai minimum, nilai rata-rata (mean), dan nilai standar deviasi. Dekriptif data CoViD-19 adalah sebagai berikut:

Tabel 4.2 Dekriptif Data

	N	Minimum	Maximum	Mean	Std. Deviation
Jumlah Penduduk	33	30874	2435252	440044.58	512984.386
Kepadatan Penduduk	33	42.97	9189.63	1065.8042	2080.63188
Positif CoViD-19	33	0	72961	4428.88	13123.267
Sembuh CoViD-19	33	0	71929	4339.55	12952.528
Meninggal CoViD-19	33	0	1019	78.24	179.162
Valid N (listwise)	33				

4.3 Uji Asumsi Analisis Klaster

Terdapat dua asumsi dalam analisis klaster yaitu sampel yang mewakili dan uji multikolieritas. Analisis klaster dikatakan baik apabila data yang digunakan memenuhi kedua asumsi tersebut.

Uji asumsi analisis klaster adalah sebagai berikut:

1. Sampel yang mewakili (*Sampel Representatif*)

Sampel yang mewakili atau representatif adalah sampel yang diambil dapat dikatakan mewakili populasi yang ada. Pengujian sampel yang mewakili (sampel representatif) dapat dilakukan dengan uji Kaiser-Mayer-Olkin (KMO). Hasil uji KMO adalah sebagai berikut:

Tabel 4.3 Hasil Uji Kaiser-Mayer-Olkin

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.696
Bartlett's Test of Sphericity	Approx. Chi-Square	441.433
	Df	10
	Sig.	.000

Sumber: Olah Data SPSS 24

Uji KMO ini mengukur kecukupan sampling secara menyeluruh dan mengukur kecukupan sampling untuk setiap indikator. Jika nilai KMO berkisar 0,5 sampai 1 maka sampel dapat dikatakan mewakili populasi atau sampel representatif. Karena nilai KMO sebesar 0,696050 dan nilai KMO berkisar 0,5 sampai 1 maka sampel dapat dikatakan mewakili populasi atau sampel representatif.

2. Uji Multikolinearitas

Multikolinearitas adalah adanya hubungan linear yang sempurna atau pasti di antara beberapa atau semua variabel (Gujarati, 1978). Terjadinya multikolinearitas apabila nilai $(VIF_j) \geq 10$. Jika terindikasi terjadi multikolinearitas maka harus dilakukan tindakan perbaikan multikolinearitas.

Hasil uji multikolinearitas adalah sebagai berikut:

Tabel 4.4 Uji Multikolinearitas

Variabel	Tolerance	VIF
Jumlah Penduduk	.176	5.692
Kepadatan Penduduk	.559	1.787
Positif CoViD-19	.166	6.027
Sembuh CoViD-19	.188	5.325
Meninggal CoViD-19	.181	5.514

Sumber: Olah Data SPSS 24

Nilai VIF semua variabel berada dibawah 10 maka dapat disimpulkan tidak terjadi multikolinieritas pada data. Karena kedua asumsi telah terpenuhi maka data dapat dilakukan *clustering K-Means*.

4.3 Jumlah Klaster

Penentuan klaster digunakan untuk membagi tingkatan pada data. Pada penelitian ini klaster dibagi menjadi tiga klaster. Adapun klaster yang digunakan pada penelitian ini adalah sebagai berikut:

- a. Klaster 1 atau C1 yaitu daerah penyebaran kasus CoViD-19 Tinggi
- b. Klaster 2 atau C2 yaitu daerah penyebaran kasus CoViD-19 Sedang
- c. Klaster 3 atau C3 yaitu daerah penyebaran kasus CoViD-19 Rendah

4.4 Menentukan Titik Pusat atau *Centroid* Secara Acak

Pada penentuan titik pusat atau *centroid* secara acak. Pada penelitian digunakan hasil penjumlahan data (Jumlah penduduk+Kepadatan penduduk+Positif CoViD-19+ Sembuh CoViD-19+Meninggal CoViD-19) untuk menentukan klaster. Klaster 1 diambil dari penjumlahan data yang paling tinggi, klaster 2 diambil dari penjumlahan data menengah, dan kluster 3 diambil dari penjumlahan data paling rendah.

Pembagian klaster adalah sebagai berikut:

- a. Klaster 1, *Centroid* pada data ke-28 (Medan)
- b. Klaster 2 *Centroid* pada data ke-21 (Simalungun)

c. Klaster 3 *Centroid* pada data ke-3 (Dairi)

Berdarkan nilai *centroid* awal penyebaran kasus CoViD-19 di Sumatera Utara yang telah ditentukan diperoleh tabel *centroid* awal sebagai berikut:

Tabel 4.5 Centroid Awal Penyebaran CoViD-19

Data	Klaster		
	1	2	3
Penduduk	2435252	990246	30874
Kepadatan Penduduk	9189.63	226.65	160.16
Positif	72961	0	2963
Sembuh	71929	0	2856
Meninggal	1019	0	107

4.5 K-Means Clustering

Pada bagian sebelumnya telah ditentukan nilai *centroid* awal penyebaran CoViD-19. Berdasarkan hal tersebut dilakukan *clustering* pada data penyebaran CoViD-19 di Sumatera Utara. Pengelompokan penyebaran CoViD-19 di Sumatera Utara menggunakan *K-Means Clustering* adalah sebagai berikut:

Iterasi 1

1. Menghitung jarak (kemiripan) antar objek menggunakan perhitungan jarak

Euclidean

a. Klaster 1

$$D(X_1, C_1) = \sqrt{(X_{11} - C_1)^2 + (X_{12} - C_1)^2 + (X_{13} - C_1)^2 + (X_{14} - C_1)^2 + (X_{15} - C_1)^2}$$

$$D(X_1, C_1)$$

$$= \sqrt{(769960 - 2435252)^2 + (207,97 - 9189.63)^2 + (178 - 72961)^2 + (1688 - 71929)^2 + (30 - 1019)^2}$$

$$D(X_1, C_1) = \sqrt{2783288456731.35}$$

$$D(X_1, C_1) = 1668319.05$$

Dengan menggunakan cara yang sama pada $D(X_2, C_1)$ sampai $D(X_{33}, C_1)$ hasil perhitungan jarak dapat dilihat pada tabel 4.6.

b. Klaster 2

$$D(X_1, C_2) = \sqrt{(X_{11} - C_2)^2 + (X_{12} - C_2)^2 + (X_{13} - C_2)^2 + (X_{14} - C_2)^2 + (X_{15} - C_2)^2}$$

$$D(X_1, C_2)$$

$$= \sqrt{(769960 - 990246)^2 + (207,97 - 226,65)^2 + (178 - 0)^2 + (1688 - 0)^2 + (30 - 0)^2}$$

$$D(X_1, C_2) = \sqrt{48531723912.94}$$

$$D(X_1, C_2) = 220299,17$$

Dengan menggunakan cara yang sama pada $D(X_2, C_2)$ sampai $D(X_{33}, C_2)$ hasil perhitungan jarak dapat dilihat pada tabel 4.6.

c. Klaster 3

$$D(X_1, C_3) = \sqrt{(X_{11} - C_3)^2 + (X_{12} - C_3)^2 + (X_{13} - C_3)^2 + (X_{14} - C_3)^2 + (X_{15} - C_3)^2}$$

$$D(X_1, C_3)$$

$$= \sqrt{(769960 - 30874)^2 + (207,97 - 160.16)^2 + (178 - 2963)^2 + (1688 - 2856)^2 + (30 - 107)^2}$$

$$D(X_1, C_3) = \sqrt{546251037859}$$

$$D(X_1, C_3) = 739087,98$$

Dengan menggunakan cara yang sama pada $D(X_2, C_3)$ sampai $D(X_{33}, C_3)$ hasil perhitungan jarak dapat dilihat pada tabel 4.6. Berdasarkan perhitungan diatas diperoleh jarak pada iterasi pertama sebagai berikut:

Tabel 4.6 Perhitungan Jarak Pada Iterasi Pertama

Kabupaten/Kota	Jarak K1	Jarak K2	Jarak K3
Asahan	1668319	220299.2	739088
Batu Bara	2027118	579569.5	379814.5
Dairi	2406405	959380.8	0
Deli Serdang	511881.8	941275.6	1900585
Humbang Hasundutan	2239648	792509.9	166878.7
Karo	2032753	585251.8	374129.4
Labuhanbatu	1943908	496357.5	463025.9
Labuhanbatu Selatan	2123556	676154.9	283228
Labuhanbatu Utara	2055712	608257	351123.9
Langkat	1408653	40015.57	999329.9
Mandailing Natal	1964977	517362.5	442019.1
Nias	2290890	843574	115870.7
Nias Barat	2347512	900252	59263.17

Nias Selatan	2077269	629715	329682.7
Nias Utara	2290287	842972	116471
Padang Lawas Utara	2176931	729526.4	229871.3
Padang Lawas	2176665	729235	230170.9
Pakpak Barat	2385088	937895.3	21737.24
Samosir	2300997	853808.9	105578.1
Serdang Bedagai	1780503	332781.4	626616.1
Simalungun	1448662	0	959380.8
Tapanuli Selatan	2136690	689340.3	270040.7
Tapanuli Tengah	2072548	625071.2	334312.1
Tapanuli Utara	2124658	677524.3	281898.9
Toba	2231422	784047	175371.7
Binjai	2144435	699362.1	262987.4
Gunung Sitoli	2301526	854229.1	105217.9
Medan	0	1448662	2406405
Padang Sidimpuan	2212499	765143.3	194268.9
Pematang Siantar	2169257	722015.5	237426.3
Sibolga	2347915	900664.1	58888.41
Tanjung Balai	2261522	814220.7	145197.3
Tebing Tinggi	2264639	817428.4	142080.5

2. Mengalokasikan jarak kemasing-masing klaster

Klaster yang memiliki jarak terpendek (terkecil) merupakan klaster data tersebut.

Pengelompokan klaster pada masing-masing data adalah sebagai berikut:

Tabel 4.7 Perhitungan Jarak dan Klaster Pada Iterasi Pertama

Kabupaten/Kota	Jarak K1	Jarak K2	Jarak K3	Klaster

Asahan	1668319	220299.2	739088	2
Batu Bara	2027118	579569.5	379814.5	3
Dairi	2406405	959380.8	0	3
Deli Serdang	511881.8	941275.6	1900585	1
Humbang Hasundutan	2239648	792509.9	166878.7	3
Karo	2032753	585251.8	374129.4	3
Labuhanbatu	1943908	496357.5	463025.9	3
Labuhanbatu Selatan	2123556	676154.9	283228	3
Labuhanbatu Utara	2055712	608257	351123.9	3
Langkat	1408653	40015.57	999329.9	2
Mandailing Natal	1964977	517362.5	442019.1	3
Nias	2290890	843574	115870.7	3
Nias Barat	2347512	900252	59263.17	3
Nias Selatan	2077269	629715	329682.7	3
Nias Utara	2290287	842972	116471	3
Padang Lawas Utara	2176931	729526.4	229871.3	3
Padang Lawas	2176665	729235	230170.9	3
Pakpak Barat	2385088	937895.3	21737.24	3
Samosir	2300997	853808.9	105578.1	3
Serdang Bedagai	1780503	332781.4	626616.1	2
Simalungun	1448662	0	959380.8	2
Tapanuli Selatan	2136690	689340.3	270040.7	3
Tapanuli Tengah	2072548	625071.2	334312.1	3
Tapanuli Utara	2124658	677524.3	281898.9	3
Toba	2231422	784047	175371.7	3
Binjai	2144435	699362.1	262987.4	3
Gunung Sitoli	2301526	854229.1	105217.9	3
Medan	0	1448662	2406405	1
Padang Sidempuan	2212499	765143.3	194268.9	3
Pematang Siantar	2169257	722015.5	237426.3	3

Sibolga	2347915	900664.1	58888.41	3
Tanjung Balai	2261522	814220.7	145197.3	3
Tebing Tinggi	2264639	817428.4	142080.5	3

3. Menentukan Posisi Klaster Baru

Posisi klaster baru diperoleh dengan cara menghitung nilai rata-rata dari data-data yang ada pada klaster. Hasil perhitungan rata-rata pada pusat klaster dapat dilihat pada lampiran 2. Dari hasil pemasatan klaster didapatkan pusat klaster baru sebagai berikut:

Tabel 4.8 Centroid Baru Iterasi Pertama

Data	Klaster		
	1	2	3
Penduduk	2183346.5	861974.5	248402.963
Kepadatan Penduduk	5025.62	236.2875	895.3762963
Positif	40883.5	1579	2150.740741
Sembuh	40260.50	1492	2100.592593
Meninggal	616.5	83.75	37.55555556

4. Karena *centroid* awal klaster pada Tabel 4.5 dan *centroid* klaster baru pada iterasi pertama (Tabel 4.8) mengalami perubahan, maka proses klasterisasi berulang dan kembali ke langkah 4.

SUMATERA UTARA MEDAN

Iterasi 2

1. Menghitung jarak (kemiripan) antar objek menggunakan perhitungan jarak *Euclidean*
 - a. Klaster 1

$$D(X_1, C_1) = \sqrt{(X_{11} - C_1)^2 + (X_{12} - C_1)^2 + (X_{13} - C_1)^2 + (X_{14} - C_1)^2 + (X_{15} - C_1)^2}$$

$$\begin{aligned}
 D(X_1, C_1) &= \sqrt{(769960 - 2183346,5)^2 + (207,97 - 5025,62)^2 + (178 - 40883,2)^2 + (1688 - 40260,5)^2 +} \\
 &\quad (30 - 616,5)^2 \\
 D(X_1, C_1) &= \sqrt{2000706726214,35} \\
 D(X_1, C_1) &= 1414463,41
 \end{aligned}$$

Dengan menggunakan cara yang sama pada $D(X_2, C_1)$ sampai $D(X_{33}, C_1)$ hasil perhitungan jarak dapat dilihat pada tabel 4.9.

b. Klaster 2

$$\begin{aligned}
 D(X_1, C_2) &= \sqrt{(X_{11} - C_2)^2 + (X_{12} - C_2)^2 + (X_{13} - C_2)^2 + (X_{14} - C_2)^2 + (X_{15} - C_2)^2} \\
 D(X_1, C_2) &= \sqrt{(769960 - 861974,5)^2 + (207,97 - 236,2875)^2 + (178 - 1579)^2 + (1688 - 1492)^2 +} \\
 &\quad (30 - 83,75) \\
 D(X_1, C_2) &= \sqrt{8466729638,19} \\
 D(X_1, C_2) &= 92014,83
 \end{aligned}$$

Dengan menggunakan cara yang sama pada $D(X_2, C_2)$ sampai $D(X_{33}, C_2)$ hasil perhitungan jarak dapat dilihat pada tabel 4.9.

c. Klaster 3

$$\begin{aligned}
 D(X_1, C_3) &= \sqrt{(X_{11} - C_3)^2 + (X_{12} - C_3)^2 + (X_{13} - C_3)^2 + (X_{14} - C_3)^2 + (X_{15} - C_3)^2} \\
 D(X_1, C_3) &= \sqrt{(769960 - 248402,963)^2 + (207,97 - 895,37)^2 + (178 - 2150,74)^2 + (1688 - 2100,59)^2 +} \\
 &\quad (30 - 37,55)^2 \\
 D(X_1, C_3) &= \sqrt{272022572964,55} \\
 D(X_1, C_3) &= 521557,83
 \end{aligned}$$

Dengan menggunakan cara yang sama pada $D(X_2, C_3)$ sampai $D(X_{33}, C_3)$ hasil perhitungan jarak dapat dilihat pada tabel 4.9.

Berdasarkan perhitungan diatas diperoleh jarak pada iterasi pertama sebagai berikut:

Tabel 4.9 Perhitungan Jarak Pada Iterasi Kedua

Kabupaten/Kota	Jarak K1	Jarak K2	Jarak K3
Asahan	1414463	92014.83	521557.8
Batu Bara	1773561	451297.4	162284.5
Dairi	2153137	831102.8	217533
Deli Serdang	255940.9	1069515	1683064
Humbang Hasundutan	1986297	664228.9	50692.28
Karo	1779214	456976.5	156599.2
Labuhanbatu	1690322	368077	245497
Labuhanbatu Selatan	1870080	547880.6	65703.98
Labuhanbatu Utara	1802197	479980.6	133594.5
Langkat	1154475	168227.5	781799.8
Mandailing Natal	1711377	389088.9	224488.8
Nias	2037488	715305.8	101778.2
Nias Barat	2094144	771983.6	158439.1
Nias Selatan	1823725	501448.2	112170.5
Nias Utara	2036885	714703.7	101175.1
Padang Lawas Utara	1923468	601256.3	12557.55
Padang Lawas	1923193	600966.9	12949.92
Pakpak Barat	2131753	809624.7	196066.6
Samosir	2047645	725533.6	111965.8
Serdang Bedagai	1526794	204493.7	409088.9
Simalungun	1194489	128289.9	741849.4
Tapanuli Selatan	1883236	561063.8	52515.71
Tapanuli Tengah	1819030	496797.8	116784.1
Tapanuli Utara	1871273	549237.9	64484.98
Toba	1977984	655778.9	42313.6
Binjai	1891622	571171.8	54873.67
Gunung Sitoli	2048134	725960.3	112422.7
Medan	255940.9	1576496	2189125
Padang Sidimpuan	1959063	636873.5	23433.1

Pematang Siantar	1915848	593739.9	20243.4
Sibolga	2094551	772396	158852.5
Tanjung Balai	2008118	685950.2	72411.89
Tebing Tinggi	2011266	689157.2	75714.36

2. Mengalokasikan jarak ke masing-masing klaster

Klaster yang memiliki jarak terpendek (terkecil) merupakan klaster data tersebut. Pengelompokan klaster pada masing-masing data adalah sebagai berikut

Tabel 4.10 Perhitungan Jarak dan Klaster Pada Iterasi Kedua

Kabupaten/Kota	Jarak K1	Jarak K2	Jarak K3	Klaster
Asahan	1414463	92014.83	521557.8	2
Batu Bara	1773561	451297.4	162284.5	3
Dairi	2153137	831102.8	217533	3
Deli Serdang	255940.9	1069515	1683064	1
Humbang Hasundutan	1986297	664228.9	50692.28	3
Karo	1779214	456976.5	156599.2	3
Labuhanbatu	1690322	368077	245497	3
Labuhanbatu Selatan	1870080	547880.6	65703.98	3
Labuhanbatu Utara	1802197	479980.6	133594.5	3
Langkat	1154475	168227.5	781799.8	2
Mandailing Natal	1711377	389088.9	224488.8	3
Nias	2037488	715305.8	101778.2	3
Nias Barat	2094144	771983.6	158439.1	3
Nias Selatan	1823725	501448.2	112170.5	3
Nias Utara	2036885	714703.7	101175.1	3
Padang Lawas Utara	1923468	601256.3	12557.55	3
Padang Lawas	1923193	600966.9	12949.92	3
Pakpak Barat	2131753	809624.7	196066.6	3
Samosir	2047645	725533.6	111965.8	3

Serdang Bedagai	1526794	204493.7	409088.9	2
Simalungun	1194489	128289.9	741849.4	2
Tapanuli Selatan	1883236	561063.8	52515.71	3
Tapanuli Tengah	1819030	496797.8	116784.1	3
Tapanuli Utara	1871273	549237.9	64484.98	3
Toba	1977984	655778.9	42313.6	3
Binjai	1891622	571171.8	54873.67	3
Gunung Sitoli	2048134	725960.3	112422.7	3
Medan	255940.9	1576496	2189125	1
Padang Sidimpuan	1959063	636873.5	23433.1	3
Pematang Siantar	1915848	593739.9	20243.4	3
Sibolga	2094551	772396	158852.5	3
Tanjung Balai	2008118	685950.2	72411.89	3
Tebing Tinggi	2011266	689157.2	75714.36	3

3. Menentukan Posisi Klaster Baru

Posisi Klaster baru diperoleh dengan cara menghitung nilai rata-rata dari data-data yang ada pada klaster. Hasil perhitungan rata-rata pada pusat klaster dapat dilihat pada lampiran 5.

UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

Dari pemusatan didapatkan pusat klaster baru sebagai berikut:

Tabel 4.11 Centroid Baru Iterasi Kedua

Data	Klaster		
	1	2	3
Penduduk	2183346.5	861974.5	248402.963
Kepadatan Penduduk	5,025.62	236.2875	895.3762963
Positif	40883.5	1579	2150.740741
Sembuh	40,260.50	1492	2100.592593

Meninggal	616.5	83.75	37.55555556
-----------	-------	-------	-------------

4. Karena *centroid* awal klaster pada Tabel 4.5 dan *centroid* klaster baru pada iterasi kedua (Tabel 4.11) mempunyai nilai yang sama, maka *K-Means Clustering* telah selesai.

4.6 Tingkat Signifikansi Antar Klaster

Untuk melihat pengujian tingkat signifikansi antar klaster dan mengetahui perbedaan di setiap klaster, maka perlu dilakukan uji ANOVA. Ketentuan penggunaan angka F dalam analisis klasterialah bahwa semakin besar angka F hitung (jika dilakukan uji Hipotesis, maka F hitung akan lebih besar dari F tabel) dan tingkat signifikansi (sig) $<0,05$; maka semakin besar perbedaan antara ketiga klaster yang terbentuk (Bastian *et al.*, 2018).



Berikut adalah hasil dari uji ANOVA:

Tabel 4.12 ANOVA

	Cluster		Error		F	Sig.
	Mean Square	Df	Mean Square	df		
Jumlah Penduduk	3890959175695.8	2	21299233425.82	30	182.681	0.000
Kepadatan Penduduk	17448414.033	2	3454403.359	30	5.051	0.013
Positif CoViD-19	1415246850.915	2	89351694.723	30	15.839	0.000

Sembuh CoViD-19	1374206289.582	2	87338749.834	30	15.734	0.000
Meninggal CoViD-19	312130.072	2	13430.264	30	23.241	0.000

Sumber: Olah Data SPSS 24

Berdasarkan tabel yang didapat, pada kolom klaster adalah besaran *between cluster means* dan kolom *Error* adalah besaran *within cluster means*. Sementara pada kolom F didapat dari rumus sebagai berikut:

$$F = \frac{\text{between cluster mean}}{\text{within cluster mean}}$$

Hipotesis:

H_0 : Ketiga klaster tidak mempunyai perbedaan yang signifikan.

H_1 : Ketiga klaster mempunyai perbedaan yang signifikan.

Jika angka signifikan $> 0,05$; H_0 diterima H_1 ditolak.

Jika angka signifikan $< 0,05$; H_0 ditolak H_1 diterima.

Pada tabel 4.12 dapat dilihat variabel yang membedakan berturut-turut:

1. Jumlah Penduduk dengan nilai F = 182,681 dan sig. = 0,000.
2. Kepadatan Penduduk dengan nilai F = 5,051 dan sig. = 0,015.
3. Positif dengan nilai F = 15,839 dan sig. = 0,001.
4. Sembuh dengan nilai F = 15,734 dan sig. = 0,000.
5. Meninggal dengan nilai F = 23,241 dan sig. = 0,000.

Berdasarkan hasil yang didapatkan bahwa semua variabel memiliki nilai signifikansi (sig.) $< 0,05$, maka ketiga klaster mempunyai perbedaan yang signifikan. bahwa variabel jumlah penduduk adalah variabel yang paling membedakan anggota dari ketiga klaster karena mempunyai nilai F terbesar diantara variabel yang lain yaitu 182,681 dan nilai signifikan 0,000.

4.7 Pembahasan

Pada bagian sebelumnya telah dilakukan pengelompokan *K-Means Clustering* pada kasus penyebaran CoViD-19. Tahap pertama yang dilakukan

adalah penentuan klaster. Pada penelitian ini klaster dibagi menjadi tiga klaster yaitu daerah penyebaran kasus CoViD-19 tinggi, daerah penyebaran kasus CoViD-19 sedang dan daerah penyebaran kasus CoViD-19 rendah. Setelah klaster telah ditentukan selanjutnya ditentukan titik pusat secara acak, Penentuan titik pusat atau *centroid* secara acak yaitu klaster 1, *centroid* pada data ke-28 (Medan), klaster 2 *centroid* pada data ke-21 (Simalungun) dan klaster3 *Centroid* pada data ke-3 (Dari). Setelah *centroid* awal ditentukan kemudian dilakukan penghitungan jarak dengan jarak *Euclidean* pada masing-masing data terhadap ketiga klaster. Jarak terkecil pada antara data dengan tiap klaster menandakan bahwa data tersebut termasuk kedalam klaster dengan jarak terkecil tersebut. Kemudian data dilakukan pengelompokan klaster.

Centroid baru diperoleh dari rata-rata data tiap klaster. Jika *centroid* baru sama dengan *centroid* sebelumnya maka *clustering* selesai dilakukan. Jika *centroid* baru tidak sama dengan *centroid* sebelumnya maka ulangi langkah penghitungan nilai jarak dan pengelompokan klaster untuk memperoleh nilai *centroid* baru. Kemudian lakukan berulang kali hingga *centroid* baru sama dengan *centroid* sebelumnya.

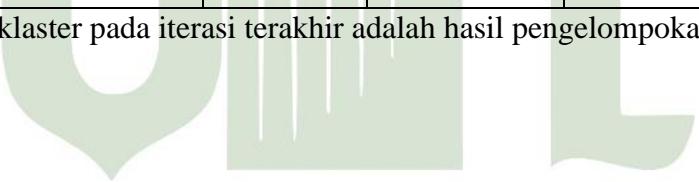
Centroid yang diperoleh pada penelitian ini adalah sebagai berikut:

Tabel 4.13 Centroid Pada Tiap Iterasi

Centroid Awal			
Klaster	K1	K2	K3
Penduduk	2435252	990246	30874
Kepadatan Penduduk	9,189.63	226.65	160.16
Positif	72961	0	2963

Sembuh	71929	0	2856
Meninggal	1019	0	107
<i>Centroid Baru Iterasi 1</i>			
Klaster	K1	K2	K3
Penduduk	2183346.5	861974.5	248402.963
Kepadatan Penduduk	5,025.62	236.2875	895.3762963
Positif	40883.5	1579	2150.740741
Sembuh	40,260.50	1492	2100.592593
Meninggal	616.5	83.75	37.55555556
<i>Centroid Baru Iterasi 2</i>			
Klaster	K1	K2	K3
Penduduk	2183346.5	861974.5	248402.963
Kepadatan Penduduk	5,025.62	236.2875	895.3762963
Positif	40883.5	1579	2150.740741
Sembuh	40,260.50	1492	2100.592593
Meninggal	616.5	83.75	37.55555556

Pengelompokan klaster pada iterasi terakhir adalah hasil pengelompokan optimal.



 UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

Hasil pengelompokan daerah penyebaran kasus CoViD-19 adalah sebagai berikut:

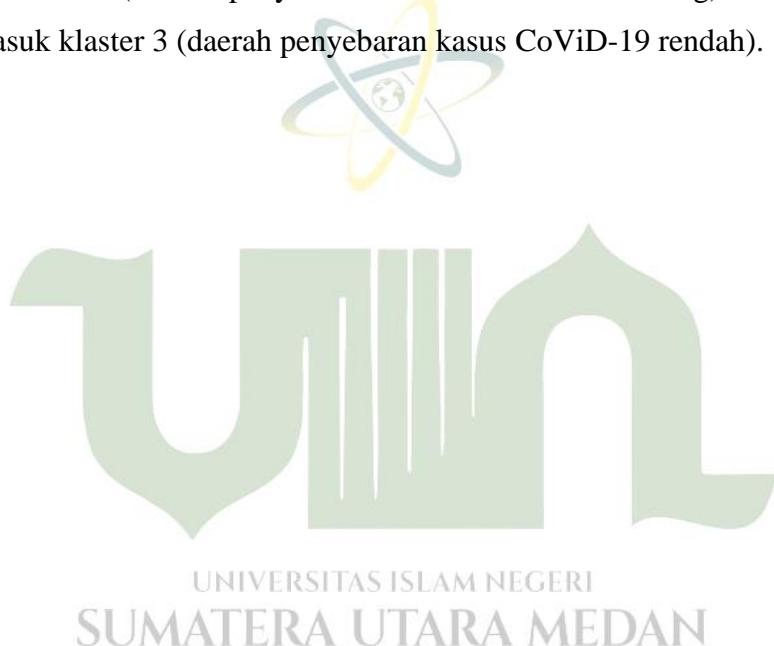
Tabel 4.14 Pengelompokan Penyebaran Kasus CoViD-19 Menggunakan *K-Means*

Kabupaten/Kota	Penduduk	Kepadatan Penduduk	Positif	Sembuh	Meninggal	Klaster
Asahan	769960	207.97	1718	1688	30	Sedang
Batu Bara	410678	445.32	931	923	8	Rendah

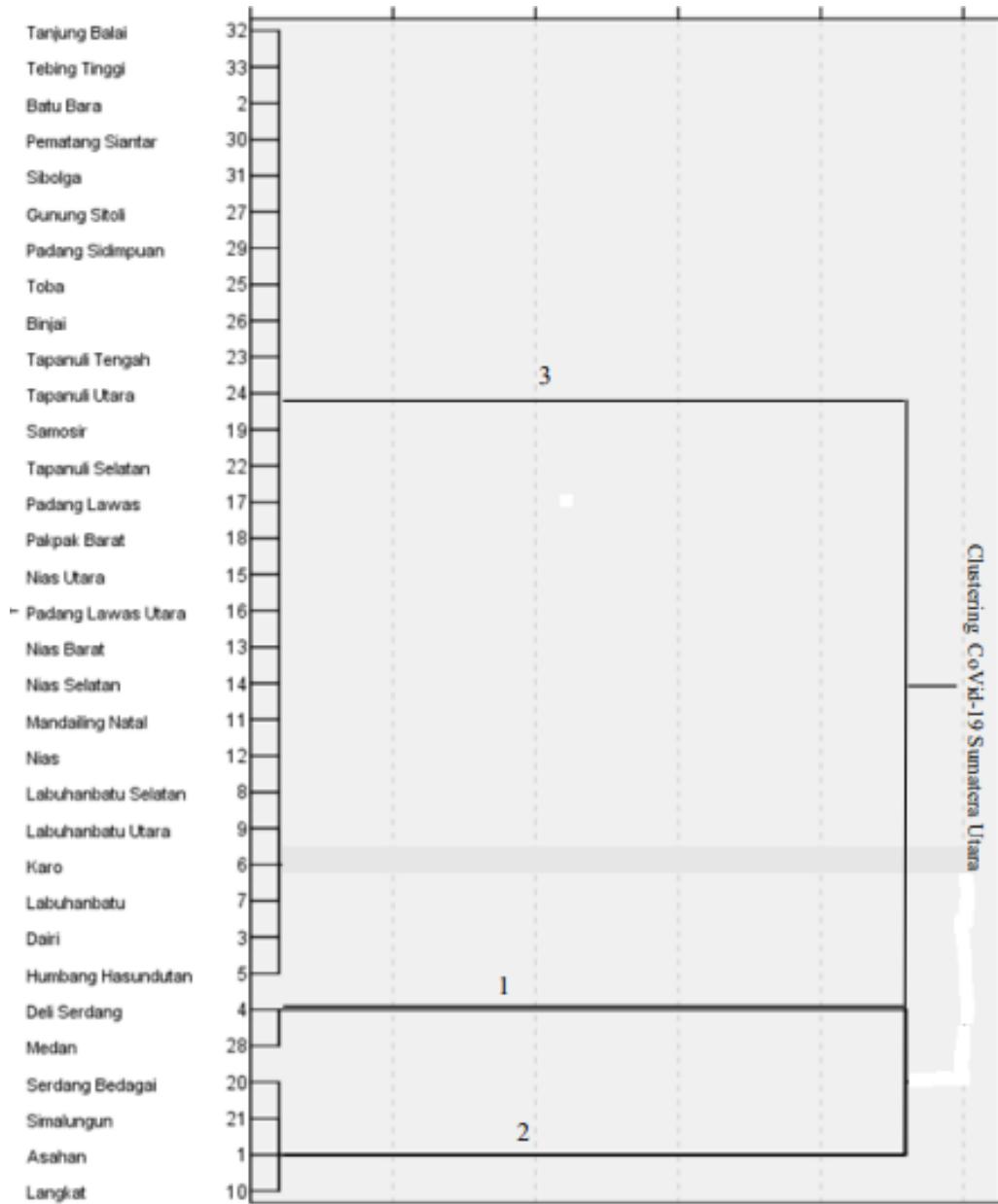
Dairi	30874	160.16	2963	2856	107	Rendah
Deli Serdang	1931441	861.6	8806	8592	214	Tinggi
Humbang Hasundutan	197751	84.68	3615	3234	33	Rendah
Karo	404998	190.41	1504	1466	58	Rendah
Labuhanbatu	493899	229.08	2321	2237	84	Rendah
Labuhanbatu Selatan	314094	87.35	1438	1367	71	Rendah
Labuhanbatu Utara	381994	106.97	1773	1694	79	Rendah
Langkat	1030202	164.52	1573	1511	49	Sedang
Mandailing Natal	472886	77.09	1164	1119	45	Rendah
Nias	146672	79.6	12	8	0	Rendah
Nias Barat	89994	189.97	0	0	0	Rendah
Nias Selatan	360531	197.53	0	0	0	Rendah
Nias Utara	147274	122.44	37	33	0	Rendah
Padang Lawas Utara	260720	66.54	513	484	29	Rendah
Padang Lawas	261011	67.05	121	118	3	Rendah
Pakpak Barat	52351	42.97	557	524	33	Rendah
Samosir	136441	65.94	1849	1815	34	Rendah
Serdang Bedagai	657490	346.01	3025	2769	256	Sedang
Simalungun	990246	226.65	0	0	0	Sedang
Tapanuli Selatan	300911	49.9	1935	1890	45	Rendah
Tapanuli Tengah	365177	166.9	1179	1161	18	Rendah
Tapanuli Utara	312758	82.49	5001	4921	80	Rendah
Toba	206199	88.54	53	48	1	Rendah
Binjai	291842	4930.6	25733	25590	143	Rendah
Gunung Sitoli	136017	484.43	115	112	3	Rendah
Medan	2435252	9189.63	72961	71929	1019	Tinggi
Padang Sidimpuan	225105	1963.24	519	516	3	Rendah

Pematang Siantar	268254	4819.51	2550	2517	33	Rendah
Sibolga	89584	2168.58	0	0	0	Rendah
Tanjung Balai	176027	1632.45	614	585	29	Rendah
Tebing Tinggi	172838	5575.42	1573	1498	75	Rendah

Berdasarkan tabel diatas setelah dilakukan pengelompokan dengan menggunakan *K-MeansClustering* dengan total jumlah penduduk 14490597 dan kepadatan penduduk 35011.38 orang/km² diperoleh Sebanyak 2 daerah yang termasuk klaster 1 (daerah penyebaran kasus CoViD-19 Medan), 4 daerah yang termasuk klaster 2 (daerah penyebaran kasus CoViD-19 sedang) dan 27 daerah yang termasuk klaster 3 (daerah penyebaran kasus CoViD-19 rendah).



Dalam bentuk dendrogram dapat dilihat sebagai berikut:



Gambar 4.1 Dendogram *Clustering CoViD-19 Sumatera Utara*

Hasil klaster yang didapat berdasarkan uji ANOVA menunjukkan bahwa semua variabel yang digunakan mempunyai perbedaan yang signifikan. Pada ketiga klaster yang terbentuk dan variabel yang memiliki perbedaan paling besar adalah variabel jumlah penduduk. Variabel ini yang paling membedakan anggota dari ketiga klaster karena mempunyai nilai F terbesar diantara variabel yang lain yaitu 182,681 dan nilai signifikan 0,000. Karena variabel jumlah penduduk

memiliki perbedaan paling besar terhadap klaster yang terbentuk dalam tingkat kasus pernyebaran CoViD-19, maka Provinsi yang memiliki jumlah penduduk terbesar di Sumatera Utara yaitu Kota Medan termasuk dalam klaster 1 yang memiliki tingkat kasus penyebaran CoViD-19 yang tinggi.

