

BAB IV

HASIL DAN PEMBAHASAN

4.1 Data dari Keempat Variabel

Data yang dikumpulkan, merupakan data sekunder yang diperoleh dari Kantor Badan Pusat Statistik Provinsi Nusa Tenggara Timur. Data diberikan oleh petugas pelayanan statistika terpadu. Berikut ini dipaparkan 4 variabel data yang digunakan (Riadi, 2023):

1. Data jumlah penduduk miskin (ribu jiwa).
2. Data rata-rata lamanya sekolah.
3. Data rata-rata pengeluaran perkapita sebulan makanan dan *non* makanan.
4. Data persentase penduduk yang sulit mengakses layanan kesehatan.

Data yang digunakan dari tahun 2017-2023. Keempat variabel data dapat dilihat pada Tabel 4.1, Tabel 4.2, Tabel 4.3, dan Tabel 4.4.

Tabel 4.1 Variabel Data Pertama

Kabupaten /Kota	Jumlah Penduduk Miskin (Ribuan Jiwa)						
	2017	2018	2019	2020	2021	2022	2023
Sumba Barat	36,690	36,300	36,560	36,950	37,780	37,060	37,150
Sumba Timur	78,180	76,800	77,390	77,300	78,330	75,280	75,660
Kupang	84,350	88,670	92,020	94,940	91,250	88,020	90,230
Timor Tengah Selatan	136,450	130,630	130,310	128,980	125,680	120,450	119,510
Timor Tengah Utara	58,590	56,090	56,940	56,980	58,330	55,880	57,190
Belu	33,950	33,910	34,080	34,180	35,410	33,980	33,190
Alor	43,900	44,130	44,320	43,550	43,830	42,300	41,910
Lembata	36,260	36,950	37,460	37,920	38,750	37,880	37,940

Flores Timur	26,970	27,990	27,830	27,890	28,900	28,080	30,930
Sikka	45,010	44,020	43,320	42,180	43,090	40,870	40,810
Ende	65,110	66,010	63,450	65,220	66,380	63,400	63,100
Ngada	20,210	20,790	20,310	20,620	20,980	20,140	20,570
Manggarai	71,860	69,320	69,330	69,520	71,030	69,680	70,000
Rote Ndao	45,570	46,120	47,660	48,770	51,710	52,430	53,550
Manggarai Barat	49,390	48,530	49,230	49,400	51,150	49,950	49,950
Sumba Tengah	25,370	24,930	25,120	25,370	25,480	24,490	24,240
Sumba Barat Daya	99,550	97,280	96,300	97,810	100,420	98,500	101,400
Nagekeo	19,200	18,690	18,700	18,510	19,110	18,010	18,570
Manggarai Timur	74,850	74,880	75,830	76,690	77,170	74,550	74,400
Sabu Raijua	28,220	28,880	29,490	30,140	30,600	30,000	30,430
Malaka	30,910	30,790	30,830	31,120	32,170	30,480	29,190
Kota Kupang	40,220	40,440	39,860	39,720	41,770	40,200	41,200

Sumber: Kantor BPS Provinsi NTT

Tabel 4.2 Variabel Data Kedua

Kabupaten /Kota	Data Rata-rata Lamanya Sekolah (Tahun)						
	2017	2018	2019	2020	2021	2022	2023
Sumba Barat	6.5	6.5	6.5	6.6	6.8	6.9	6.9
Sumba Timur	6.7	6.7	6.9	7.1	7.3	7.3	7.6
Kupang	7.1	7.1	7.4	7.4	7.4	7.4	7.4
Timor Tengah Selatan	6.4	6.5	6.7	6.7	6.7	6.8	7.0
Timor Tengah Utara	7.1	7.3	7.5	7.8	8.0	8.0	8.2
Belu	7.1	7.1	7.1	7.4	7.4	7.4	7.4
Alor	7.8	7.8	8.1	8.4	8.4	8.4	8.5
Lembata	7.6	8.0	8.2	8.2	8.2	8.3	8.3
Flores Timur	7.1	7.4	7.7	7.7	7.7	7.8	8.0
Sikka	6.6	6.7	6.7	6.9	7.0	7.0	7.0
Ende	7.6	7.8	7.8	7.8	8.0	8.1	8.2

Ngada	7.9	8.1	8.4	8.5	8.5	8.5	8.8
Manggarai	7.0	7.3	7.3	7.4	7.6	7.6	7.6
Rote Ndao	7.0	7.2	7.3	7.6	7.7	7.8	7.8
Manggarai Barat	7.1	7.2	7.2	7.3	7.6	7.8	7.9
Sumba Tengah	5.5	5.8	6.0	6.3	6.5	6.7	7.0
Sumba Barat Daya	6.3	6.3	6.3	6.3	6.4	6.4	6.4
Nagekeo	7.5	7.8	7.8	7.9	7.9	7.9	8.1
Manggarai Timur	6.5	6.7	6.9	7.1	7.4	7.4	7.7
Sabu Raijua	6.0	6.1	6.3	6.7	6.7	6.8	7.0
Malaka	6.3	6.6	6.9	6.9	7.1	7.1	7.1
Kota Kupang	11.5	11.5	11.5	11.6	11.6	11.6	11.6

Sumber: Kantor BPS Provinsi NTT

Tabel 4.3 Variabel Data Ketiga

Kabupaten /Kota	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan						
	2017	2018	2019	2020	2021	2022	2023
Sumba Barat	577,851	623,154	617,684	748,104	790,695	802,682	917,095
Sumba Timur	742,557	658,944	787,602	813,884	889,558	977,986	1,168,989
Kupang	676,374	587,138	648,552	745,021	691,687	827,021	843,693
Timor Tengah Selatan	657,523	690,268	711,629	703,941	724,446	804,853	928,636
Timor Tengah Utara	604,747	721,678	687,873	723,831	804,350	798,752	833,651
Belu	647,171	670,364	741,934	772,546	781,774	804,789	880,913
Alor	586,393	580,037	658,032	714,393	777,295	795,984	905,010
Lembata	546,909	588,550	614,899	656,508	674,257	779,226	924,794
Flores Timur	625,467	613,623	658,364	684,513	642,997	817,671	854,811
Sikka	573,477	669,277	677,617	747,178	736,632	839,377	873,304
Ende	700,226	722,880	720,088	863,295	902,309	1,008,449	1,121,834
Ngada	875,225	796,475	909,268	890,564	966,322	999,115	1,111,044
Manggarai	621,129	687,434	787,448	787,670	916,118	895,492	982,219

Rote Ndao	677,533	627,958	643,673	724,752	746,745	636,817	742,828
Manggarai Barat	692,602	839,763	914,129	900,406	1,012,327	1,025,805	1,005,036
Sumba Tengah	523,244	553,076	654,161	634,478	720,463	735,943	981,022
Sumba Barat Daya	447,779	502,582	541,040	575,195	662,124	804,300	811,193
Nagekeo	597,158	704,136	595,880	699,972	766,010	709,168	823,489
Manggarai Timur	628,891	528,087	601,452	569,215	711,029	731,612	773,644
Sabu Raijua	619,913	568,191	678,823	797,369	901,397	935,378	997,011
Malaka	500,761	608,130	652,270	639,105	645,871	764,563	899,354
Kota Kupang	1,313,418	1,359,601	1,407,029	1,503,031	1,519,469	1,378,657	1,447,960

Sumber: Kantor BPS Provinsi NTT

Tabel 4.4 Variabel Data Keempat

Kabupaten /Kota	Persentase Penduduk yang Sulit Mengakses Layanan Kesehatan						
	2017	2018	2019	2020	2021	2022	2023
Sumba Barat	3.81	3.45	1.69	0.75	0.00	0.24	0.28
Sumba Timur	4.90	3.67	8.19	3.59	1.12	1.70	1.76
Kupang	6.05	10.58	11.30	1.68	1.10	2.85	1.98
Timor Tengah Selatan	8.77	5.39	4.18	4.86	1.83	1.34	0.35
Timor Tengah Utara	5.09	6.95	3.32	1.09	1.13	0.42	0.55
Belu	3.28	3.68	3.75	3.68	1.00	0.04	0.00
Alor	3.24	7.43	2.43	6.03	0.87	0.04	0.95
Lembata	2.95	6.21	2.03	0.00	0.19	0.43	0.57
Flores Timur	3.77	1.69	0.37	0.00	0.74	0.00	0.00
Sikka	1.48	2.06	2.04	0.51	1.00	0.18	1.02
Ende	0.68	3.52	4.81	0.24	0.00	2.72	2.89
Ngada	1.66	0.63	2.79	1.54	4.64	0.06	0.68
Manggarai	2.40	3.98	1.27	3.03	1.62	0.54	0.64
Rote Ndao	2.54	1.49	1.20	2.07	3.46	0.58	0.51

Manggarai Barat	2.37	9.36	2.37	2.29	1.71	0.71	2.01
Sumba Tengah	1.97	10.43	3.48	0.84	0.00	1.84	1.38
Sumba Barat Daya	3.91	7.26	2.77	0.77	3.31	1.43	0.98
Nagekeo	2.97	9.81	0.65	1.17	1.61	2.15	2.15
Manggarai Timur	3.04	21.66	3.33	16.04	1.06	1.82	7.69
Sabu Raijua	4.05	1.07	0.00	0.40	0.00	0.00	0.00
Malaka	9.46	4.32	0.45	1.99	2.97	0.45	3.96
Kota Kupang	0.36	0.45	0.00	0.70	0.85	0.23	0.00

Sumber: Kantor BPS Provinsi NTT

4.2 Rata-rata Nilai pada Keempat Variabel Data

Data yang telah dikumpulkan, akan dihitung nilai rata-ratanya dari rentang waktu tahun dari 2017 sampai tahun 2023. Hal ini bertujuan agar mempermudah analisis karakteristik *cluster* yang terbentuk.

Perhitungan rata-rata nilai pada keempat variabel data dari tahun 2017-2023 menggunakan fungsi *average* pada *microsoft excel*. Hasil Perhitungan rata-rata setiap variabel dapat dilihat pada Tabel 4.5.

Tabel 4.5 Hasil Perhitungan Rata-rata pada Variabel Data

Kabupaten/ Kota	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang sulit Mengakses Layanan Kesehatan
Sumba Barat	36,927	6.7	725,324	1.46
Sumba Timur	76,991	7.1	862,789	3.56
Kupang	89,926	7.3	717,069	5.08
Timor Tengah Selatan	127,430	6.7	745,899	3.82
Timor Tengah Utara	57,143	7.7	739,269	2.65
Belu	34,100	7.2	757,070	2.20

Alor	43,420	8.2	716,735	3.00
Lembata	37,594	8.1	683,592	1.77
Flores Timur	28,370	7.6	699,635	0.94
Sikka	42,757	6.8	730,980	1.18
Ende	64,667	7.9	862,726	2.12
Ngada	20,517	8.4	935,430	1.71
Manggarai	70,106	7.4	811,073	1.93
Rote Ndao	49,401	7.5	685,758	1.69
Manggarai Barat	49,657	7.4	912,867	2.97
Sumba Tengah	25,000	6.2	686,055	2.85
Sumba Barat Daya	98,751	6.3	620,602	2.92
Nagekeo	18,684	7.9	699,402	2.93
Manggarai Timur	75,481	7.1	649,133	7.81
Sabu Raijua	29,680	6.5	785,440	0.79
Malaka	30,784	6.9	672,865	3.37
Kota Kupang	40,487	11.5	1,418,452	0.37

Data hasil perhitungan rata-rata pada Tabel 4.5, menunjukkan variasi nilai datanya pada setiap variabel, untuk itu diperlukan tahapan normalisasi data pada *microsoft excel*. Normalisasi ini penting agar nilai-nilai tersebut memiliki rentang atau skala yang tidak beda jauh.

4.3 Normalisasi Data

Normalisasi membantu memastikan bahwa setiap data memiliki nilai yang seimbang terhadap pengukuran jarak. Pada penelitian ini, ada 4 variabel data yang digunakan dengan nilai yang berbeda.

Data jumlah penduduk miskin memiliki nilai data dalam bentuk ribuan, data rata-rata lamanya sekolah dan data persentase penduduk yang sulit mengakses layanan kesehatan memiliki nilai dalam bentuk angka desimal, data rata-rata pengeluaran perkapita sebulan makanan dan *non* makanan memiliki

nilai data berupa angka ratusan hingga jutaan. Oleh Karena itu, diperlukan normalisasi data untuk mencegah variabel data dengan rentang nilai yang besar mendominasi proses *clustering* dan rentang nilai pada setiap variabel data itu tidak berbeda jauh. Normalisasi data menggunakan persamaan dari metode *min-max scaling*. Untuk perhitungan metode *min-max scaling* dapat dilihat pada Persamaan 4.1.

$$NB = \frac{NL - N_{Min \text{ dalam data}}}{N_{Max \text{ dalam data}} - \text{Nilai}_{Min \text{ dalam data}}} \dots\dots\dots(4.1)$$

Keterangan:

NB = Nilai Baru

NL = Nilai Lama

N_{Min} = Nilai minimum dalam data

N_{Max} = Nilai maksimal dalam data

Setelah menghitung rata-rata pada setiap variabel data, hal yang perlu dilakukan adalah mencari nilai paling tinggi dan paling rendah dari masing-masing 4 variabel data tersebut. Tabel nilai paling tinggi (maksimum) dan nilai paling rendah (minimum) pada setiap variabel data dan hasil normalisasinya dapat dilihat pada tabel Tabel 4.6.

Tabel 4.6 Tabel Hasil Normalisasi Data

Kabupaten/ Kota	Jumlah Penduduk Miskin		Rata-rata Lamanya Sekolah		Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan		Persentase Penduduk yang Kesulitan Mengakses Layanan Kesehatan	
	<i>Max</i>	<i>Min</i>	<i>Max</i>	<i>Min</i>	<i>Max</i>	<i>Min</i>	<i>Max</i>	<i>Min</i>
	127,430	18,684	11.5	6.2	1,418,452	620,602	7.81	0.37
Hasil Normalisasi								
Sumba Barat	0.168		0.764		0.083		0.131	
Sumba Timur	0.536		0.818		0.161		0.304	
Kupang	0.655		0.540		0.202		0.121	
Timor Tengah Selatan	1.000		0.723		0.084		0.157	
Timor Tengah Utara	0.354		0.530		0.273		0.149	
Belu	0.142		0.249		0.190		0.171	
Alor	0.227		0.477		0.369		0.120	
Lembata	0.174		0.673		0.351		0.079	
Flores Timur	0.089		0.076		0.265		0.099	
Sikka	0.221		0.167		0.111		0.138	
Ende	0.423		0.579		0.314		0.303	
Ngada	0.017		0.133		0.405		0.395	
Manggarai	0.473		0.456		0.217		0.239	
Rote Ndao	0.282		0.750		0.235		0.082	
Manggarai Barat	0.285		0.347		0.227		0.366	
Sumba Tengah	0.058		1.000		0.000		0.082	
Sumba Barat Daya	0.736		0.767		0.019		0.000	
Nagekeo	0.000		0.144		0.305		0.099	
Manggarai Timur	0.522		0.683		0.157		0.036	
Sabu Raijua	0.101		0.836		0.048		0.207	
Malaka	0.111		0.270		0.116		0.066	
Kota Kupang	0.200		0.000		1.000		1.000	

Pada Tabel 4.6 didapatkan hasil normalisasi data dengan parameter nilai datanya dari angka 0 hingga 1, dengan begitu nilai yang didapatkan sudah tidak

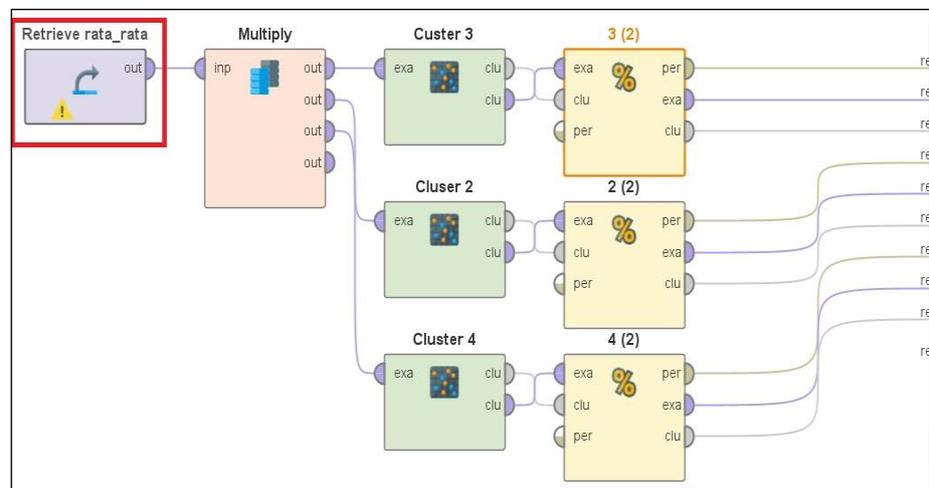
ada perbedaan yang jauh lagi, sehingga tidak mempengaruhi proses klasterisasi data dan *dataset* sudah bisa dilakukan proses *clustering* menggunakan algoritma *k-means* di dalam *tools Rapid Miner* dan *microsoft excel*.

4.4 Pengelompokkan Data pada *Tools Rapid Miner*

Pada tahap ini, pengelompokkan kabupaten/kota dengan algoritma *k-means* dilakukan menggunakan *tools Rapid Miner*. Berikut ini dijelaskan tahapan pengelompokkan data menggunakan *tools Rapid Miner*.

4.4.1 *Import Data ke Tools Rapid Miner*

Pada tahap ini, data dalam format *excel* yang telah dinormalisasikan akan di-*import* ke *tools Rapid Miner* menggunakan operator *retrieve*. *Import* data menggunakan operator *retrieve* dapat dilihat pada Gambar 4.1.



Gambar 4.1 Operator *Retrieve*

Garis merah menunjukkan operator *retrieve*. Pada *Rapid Miner*, *retrieve* adalah operasi yang digunakan untuk mengambil data dari sumber eksternal, seperti *file* lokal, basis data, atau sumber data lainnya

yang dapat diakses. *Dataset* yang di-*import* dalam *retrieve* dapat dilihat pada Gambar 4.2.

Row No.	Wilayah	Jumlah Pen...	Rata-Rata L...	Rata-rata P...	Persentase...
1	Sumba Barat	0.168	0.094	0.131	0.147
2	Sumba Timur	0.536	0.170	0.304	0.429
3	Kupang	0.655	0.208	0.121	0.633
4	Timor Tenga...	1	0.094	0.157	0.464
5	Timor Tenga...	0.354	0.283	0.149	0.306
6	Belu	0.142	0.189	0.171	0.246
7	Alor	0.227	0.377	0.120	0.353
8	Lembata	0.174	0.358	0.079	0.188
9	Flores Timur	0.089	0.264	0.099	0.077
10	Sikka	0.221	0.113	0.138	0.109
11	Ende	0.423	0.321	0.303	0.235
12	Ngada	0.017	0.415	0.395	0.180
13	Manggarai	0.473	0.226	0.239	0.210

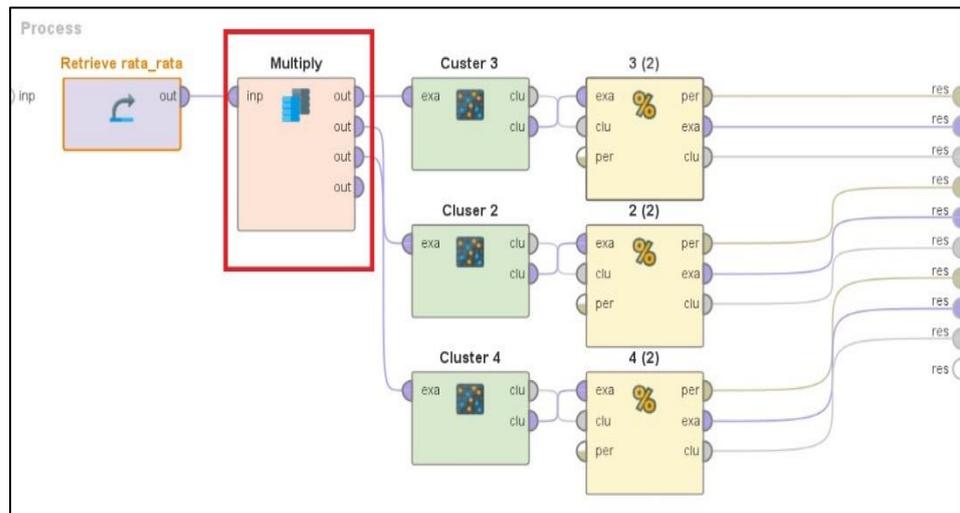
ExampleSet (22 examples, 1 special attribute, 4 regular attributes)

Gambar 4.2 *Dataset* dalam *Retrieve*

Gambar 4.2 merupakan *dataset* pada operator *retrieve* yang telah dinormalisasikan menggunakan metode *min-max scaling* di dalam *microsoft excel*, kemudian di-*import* dalam *retrieve*.

4.4.2 Menyambungkan *Retrieve* dengan Operator *Multiply*

Pada tahap ini, operator *retrieve* disambungkan pada operator *multiply*. Operator *multiply* digunakan untuk menggandakan sebuah *dataset* yang diambil menggunakan operator *retrieve*, sehingga dapat digunakan oleh beberapa operator *k-means* yang masing-masing mewakili jumlah *cluster* yang berbeda, karena akan dilakukan evaluasi jumlah *cluster*. Operator *multiply* dapat dilihat pada Gambar 4.3.



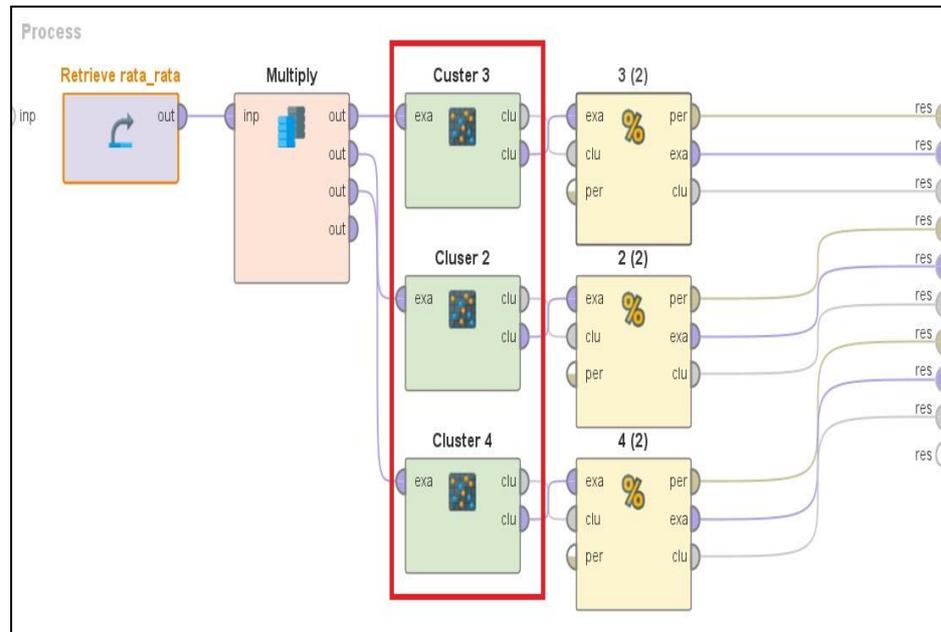
Gambar 4.3 Operator *Multiply*

Garis merah pada Gambar 4.3 merupakan operator *multiply*. Dengan menggandakan *dataset* menggunakan operator *multiply*, setiap salinan dapat digunakan oleh masing-masing operator *k-means*, sehingga dapat dilakukan perbandingan hasil jumlah *cluster* secara cepat dan efisien tanpa harus membuat ulang *dataset* yang sama pada operator *retrieve*.

4.4.3 Proses Pengelompokkan Data

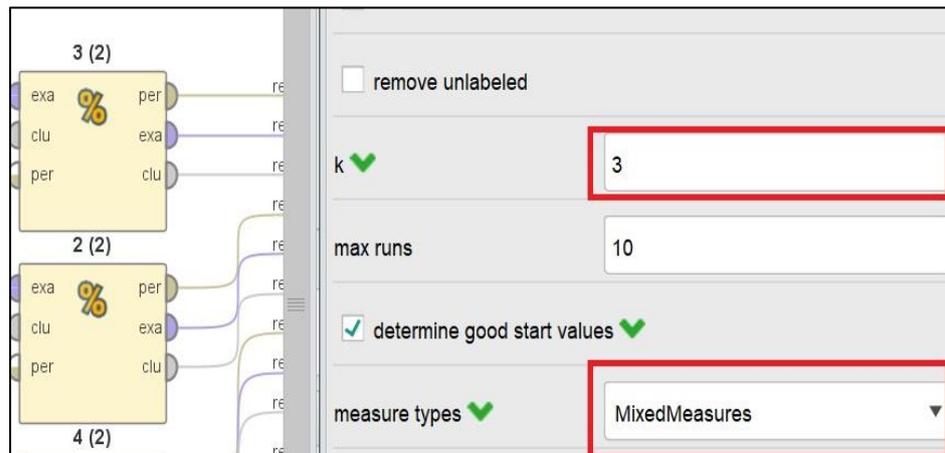
Pada tahap ini data akan dikelompokkan dalam beberapa jumlah *cluster* menggunakan operator *k-means*. Jumlah *cluster* yang digunakan adalah $k=2$, $k=3$, dan $k=4$. Data yang dikelompokkan dihitung jaraknya ke *centroid* menggunakan rumus *eucliden distance* pada operator *k-means* di dalam *Rapid Miner*. Klusterisasi data dilakukan dengan jumlah $k=3$, sesuai dengan tujuan hasil penelitian yang akan dicapai yaitu *cluster* 0, *cluster* 1, *cluster* 2. Jumlah $k=2$ dan jumlah $k=4$ digunakan untuk

evaluasi jumlah *cluster*. Operator *k-means* ($k=3$, $k=2$, $k=4$) dapat dilihat pada Gambar 4.4.



Gambar 4.4 Operator *K-Means*

Di dalam operator *k-means*, dapat diatur jumlah *cluster*-nya dan juga analisis data menggunakan *mixedmeasures*. Fungsi *mixedmeasures* digunakan untuk menganalisis data yang memiliki struktur campuran, seperti data yang mengandung campuran variabel numerik dan kategorikal. Pengaturan jumlah *cluster* dan *mixedmeasures* dapat dilihat pada Gambar 4.5.

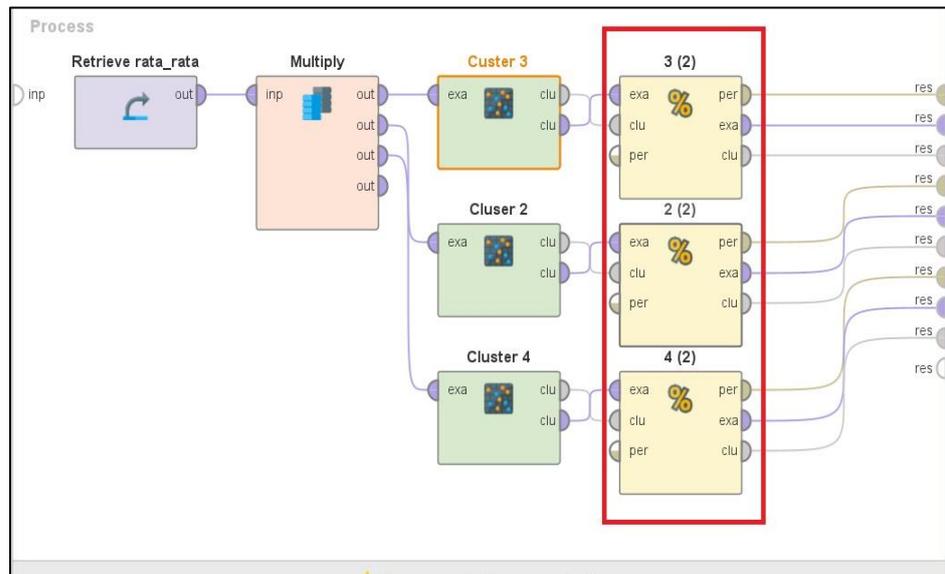


Gambar 4.5 Parameter *Mixedmeasures*

Dataset penelitian yang disimpan dalam operator *retrieve*, memiliki variabel data campuran sehingga membutuhkan fungsi *mixedmeasures* yang dapat dilihat pada Gambar 4.5. Variabel *dataset* mencakup kabupaten/kota yang termasuk ke dalam variabel data kategorikal, sedangkan variabel data jumlah penduduk miskin, rata-rata lamanya sekolah, rata-rata pengeluaran perkapita sebulan makanan *non* makanan, dan persentase penduduk yang sulit mengakses layanan kesehatan termasuk dalam variabel data numerik.

4.4.4 Evaluasi Jumlah *Cluster* Menggunakan Operator *Performance*

Pada perhitungan *k-means* di *Rapid Miner*, operator *performance* digunakan untuk mengevaluasi kualitas jumlah *cluster* yang dihasilkan oleh operator *k-means*. Operator *performance* dapat dilihat pada Gambar 4.6.



Gambar 4.6 Operator *Performance*

Operator *performance* pada Gambar 4.6 memberikan beberapa metrik evaluasi yang dapat digunakan untuk mengukur seberapa baik *cluster* yang dihasilkan oleh algoritma *k-means* saat mengelompokkan data. Metrik yang digunakan yaitu SSW (*Sum of Squares Within-Cluster*), SSB (*Sum of Squares Between-Cluster*), dan metrik rasio, ketiga metrik tersebut akan menghasilkan nilai DBI (*Davies Bouldin Index*). Di dalam *Rapid Miner* perhitungan metrik SSW, SSB, dan rasio tidak ditampilkan secara detail, hanya ditampilkan hasil akhirnya berupa nilai DBI.

4.4.5 Hasil pada *Tools Rapid Miner*

Pada tahap ini hasil yang diperoleh dapat dilihat dalam *example set* setiap *cluster* ($k=3$, $k=2$, $k=4$), *cluster model*, visualisasi data dengan *scatter plot 2D*, dan nilai DBI (*Davies Bouldin Index*).

4.4.5.1 Example Set

Fungsi *example set* dalam *Rapid Miner* untuk menyediakan data hasil *clustering* menggunakan algoritma *k-means*. Operator ini mengelompokkan data ke dalam jumlah *cluster* yang telah ditentukan ($K=3$, $K=2$, $K=4$). Data hasil *cluster* di *example set* dapat dilihat pada Gambar 4.7.

Row No.	id	Wilayah	cluster
1	1	Sumba Barat	cluster_0
2	2	Sumba Timur	cluster_1
3	3	Kupang	cluster_1
4	4	Timor Tenga...	cluster_1
5	5	Timor Tenga...	cluster_0
6	6	Belu	cluster_0
7	7	Alor	cluster_0
8	8	Lembata	cluster_0
9	9	Flores Timur	cluster_0
10	10	Sikka	cluster_0
11	11	Ende	cluster_0

Row No.	id	Wilayah	cluster
12	12	Ngada	cluster_0
13	13	Manggarai	cluster_0
14	14	Rote Ndao	cluster_0
15	15	Manggarai B...	cluster_0
16	16	Sumba Tengah	cluster_0
17	17	Sumba Barat ...	cluster_1
18	18	Nagekeo	cluster_0
19	19	Manggarai TI...	cluster_1
20	20	Sabu Raijua	cluster_0
21	21	Malaka	cluster_0
22	22	Kota Kupang	cluster_2

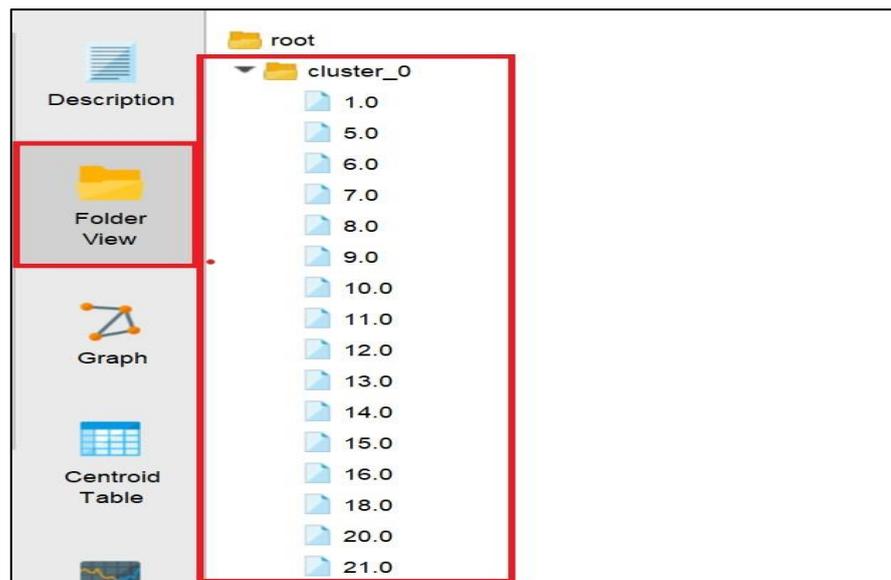
Gambar 4.7 Example Set Cluster 0,1, dan 2

Gambar 4.7 menampilkan hasil *cluster* yang terbentuk dalam *dataset* pada setiap kabupaten/kota. Masing-masing Kabupaten/kota dikelompokkan dalam *cluster* 0, 1, dan 2.

4.4.5.2 Cluster Model ($K=3$)

Operator *cluster model* pada *Rapid Miner* digunakan untuk melakukan analisis klasterisasi atau pengelompokan data. Operator ini membantu dalam mengidentifikasi pola atau struktur dalam data, dengan cara mengelompokkan data ke dalam sejumlah kelompok (*cluster*) berdasarkan kemiripan fitur atau karakteristik yang dimiliki

oleh setiap data. Di dalam *cluster model* terdapat operator *folder view*, operator *description*, dan operator *centroid table*. Operator *folder view* akan menyajikan informasi tentang data ke-n dari setiap *cluster* yang dihasilkan, operator *description* menyajikan informasi berapa banyak data yang terbentuk disetiap *cluster*, dan operator *centroid table* menyajikan informasi mengenai *centroid* yang terbentuk pada iterasi terakhir dalam klasterisasi data. *Folder view*, *description*, *centroid table* yang ada di dalam *cluster model* dapat dilihat pada gambar dan tabel berikut.



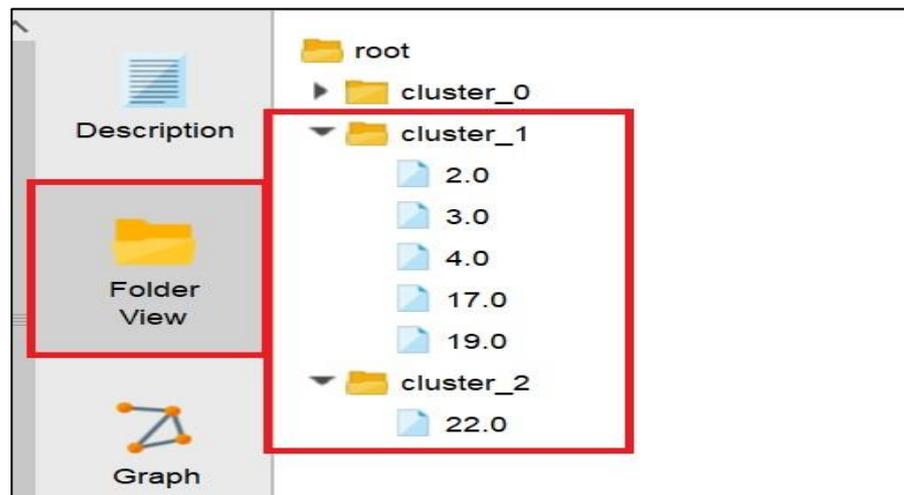
Gambar 4.8 *Folder View Cluster 0*

Nilai-nilai atau angka yang ditampilkan pada Gambar 4.8, merupakan angka yang merepresentasikan titik data atau data ke-n dari masing-masing kabupaten/kota yang terbentuk pada *cluster 0*. Kabupaten/kota yang terdapat pada data ke-n di *cluster 0* dapat dilihat pada Tabel 4.7.

Tabel 4.7 Data Ke-n Cluster 0

Data Ke-n	Kabupaten/Kota
1.0	Sumba Barat
5.0	Timor Tengah Utara
6.0	Belu
7.0	Alor
8.0	Lembata
9.0	Flores Timur
10.0	Sikka
11.0	Ende
12.0	Ngada
13.0	Manggarai
14.0	Rote Ndao
15.0	Manggarai Barat
16.0	Sumba Tengah
18.0	Nagekeo
20.0	Sabu Raijua
21.0	Malaka

Pada Tabel 4.7 sudah dipaparkan kabupaten/kota apa saja yang berada pada titik data atau data ke-n dari *cluster* 0 yang terbentuk. *Folder view cluster* 1 dan 2 dapat dilihat pada Gambar 4.9.



Gambar 4.9 Folder View Cluster 1 dan 2

Nilai-nilai atau angka yang ditampilkan Pada Gambar 4.9, merupakan angka yang merepresentasikan titik data atau data ke-n dari masing-masing kabupaten/kota yang terbentuk pada *cluster* 1 dan

2. Kabupaten/kota yang terdapat pada data ke-n di *cluster* 1 dan 2 dapat dilihat pada Tabel 4.8 dan Tabel 4.9.

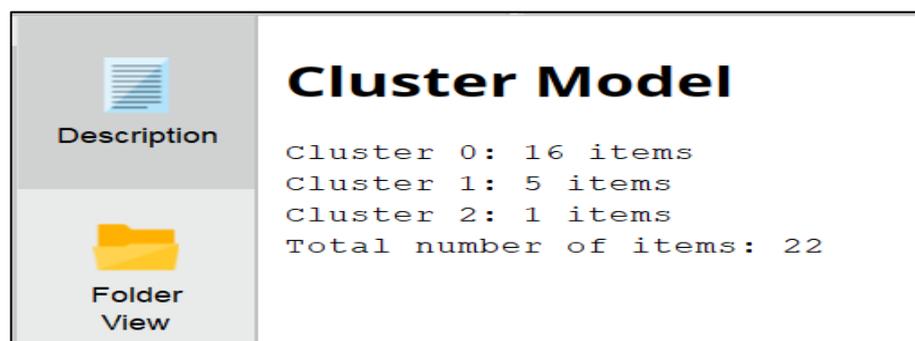
Tabel 4.8 Data Ke-n *Cluster* 1

Data Ke-n	Kabupaten/Kota
2.0	Sumba Barat
3.0	Timor Tengah Utara
4.0	Belu
17.0	Alor
19.0	Lembata

Tabel 4.9 Data Ke-n *Cluster* 2

Data Ke-n	Kabupaten/Kota
22.0	Kota Kupang

Pada Tabel 4.8 dan Tabel 4.9 sudah dipaparkan kabupaten/kota yang berada pada titik data atau data ke-n di setiap *cluster* 1 dan 2 yang terbentuk dari hasil pengelompokkan *dataset* menggunakan algoritma *k-means*. Operator *description* dapat dilihat pada Gambar 4.10



Gambar 4.10 *Description* K=3

Gambar 4.10 menampilkan banyaknya *items* yang terbentuk pada setiap *cluster*-nya. *Items* yang dimaksudkan adalah kabupaten/kota. Jadi *cluster model* memaparkan bahwa ada 16 kabupaten/kota yang masuk ke dalam *cluster* 0, kemudian ada 5

kabupaten/kota yang masuk ke dalam *cluster* 1, dan 1 kabupaten/kota yang masuk ke dalam *cluster* 2. Hasil *centroid* pada *Rapid Miner* dapat dilihat pada Gambar 4.11.

Attribute	cluster_0	cluster_1	cluster_2
Jumlah Penduduk Miskin	0.195	0.690	0.200
Rata-Rata Lamanya Sekolah	0.226	0.132	1
Rata-rata Pengeluaran per Kapita Sebu...	0.170	0.123	1
Persentase Penduduk yang Kesusaha...	0.232	0.574	0

Gambar 4.11 *Centroid Cluster 3*

Gambar 4.11 merupakan *centroid table* untuk jumlah $k=3$. *Centroid* yang dihasilkan merupakan *centroid* pada iterasi terakhir yang terbentuk dalam *tools Rapid Miner*, walaupun setiap iterasi tidak ditampilkan secara detail di dalamnya. Tabel *centroid* dapat terlihat jelas pada Tabel 4.10.

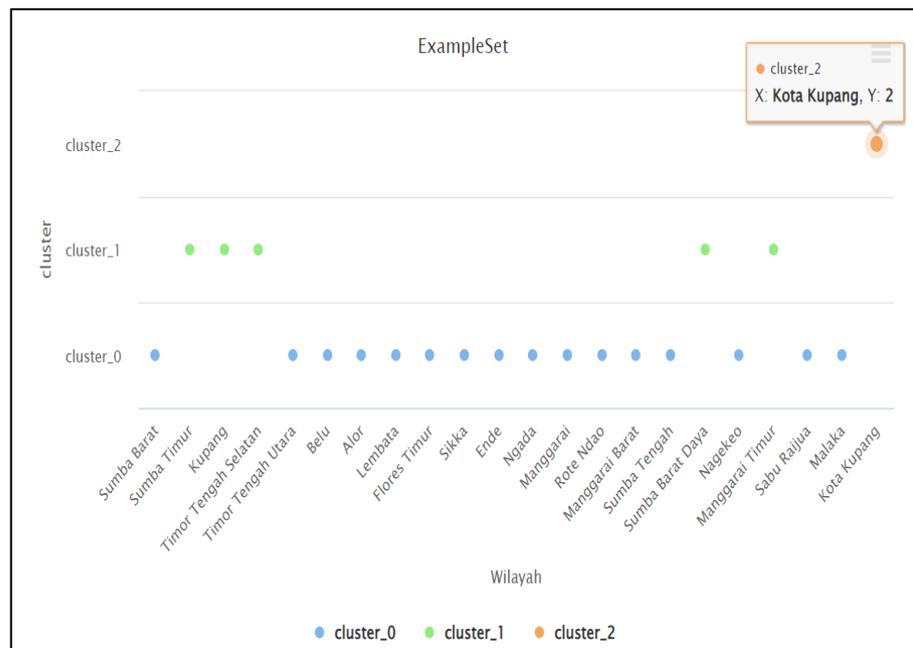
Tabel 4.10 Hasil *Centroid* pada *Rapid Miner*

Atribut	Cluster_0	Cluster_1	Cluster_2
Jumlah Penduduk Miskin	0.195	0.690	0.200
Rata-rata lamanya Sekolah	0.226	0.132	1
Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	0.170	0.123	1
Persentase Penduduk yang Kesulitan Mengakses Layanan Kesehatan	0.232	0.574	0

Tabel 4.10 merupakan hasil *centroid* untuk jumlah $k=3$ pada *tools Rapid Miner* yang telah diperjelas gambarnya menggunakan tabel.

4.4.5.3 Visualisasi Data dengan *Scatter plot* 2D (2 Dimensi)

Scatter plot digunakan untuk memvisualisasikan hasil *clustering* dalam fitur dua dimensi, hal ini membantu dalam memahami sebaran data dan bagaimana *cluster* terbentuk. Hasil visualisasi *clustering* dengan *scatter plot* dapat dilihat pada Gambar 4.12.



Gambar 4.12 *Scatter Plot Cluster 3*

Cluster 0 ditandai dengan warna biru, *cluster 1* ditandai dengan warna hijau, *cluster 2* ditandai dengan warna orange. Hasil *cluster* yang terbentuk pada *scatter plot* dapat dilihat pada Tabel 4.11.

Tabel 4.11 Hasil Pengelompokkan Data *Cluster 3*

<i>Cluster 0</i>	<i>Cluster 1</i>	<i>Cluster 2</i>
Sumba Barat	Sumba Timur	Kota Kupang
Timor Tengah Utara	Kupang	
Belu	Timor Tengah Selatan	
Alor	Sumba Barat Daya	
Lembata	Manggarai Timur	
Flores Timur	Manggarai Barat	
Sikka	Nagekeo	
Ende	Malaka	
Ngada		
Manggarai		
Rote Ndao		
Manggarai Barat		
Sumba Tengah		
Nagekeo		
Sabu Raijua		
Malaka		

Tabel 4.11 memberikan informasi tentang kabupaten/kota yang telah dikelompokkan ke dalam tiga *cluster* berbeda, yaitu *cluster 0*, *cluster 1*, dan *cluster 2*, sebagai hasil dari analisis *cluster* menggunakan algoritma *k-means* dalam *tools Rapid Miner*. Proses pengelompokkan ini dilakukan berdasarkan pada empat variabel data yang telah di-*import* sebelumnya di dalam operator *retrieve*.

4.4.6 Performance Vector

Operator *performance vector* biasanya digunakan untuk mengevaluasi kualitas jumlah *cluster* yang dihasilkan dengan menggunakan nilai DBI (*Davies Bouldin Index*). Semakin kecil nilai DBI, semakin optimal jumlah *cluster* yang digunakan (Kaligis &

Yulianto, 2022). Perbandingan nilai DBI dari masing-masing jumlah *cluster* dapat dilihat pada Gambar 4.13, Gambar 4.14, dan Gambar 4.15.



Gambar 4.13 Nilai DBI *Cluster* 3

Pada Gambar 4.13 ditampilkan hasil nilai *Davies Bouldin Index* untuk jumlah $k=3$ (3 *cluster*). Nilai DBI yang dihasilkan akan digunakan untuk perbandingan nilai DBI dengan jumlah $k=2$ dan $k=4$.



Gambar 4.14 Nilai DBI *Cluster* 2

Pada Gambar 4.13 ditampilkan hasil nilai *Davies Bouldin Index* untuk jumlah $k=2$ (2 *cluster*). Nilai DBI yang dihasilkan akan digunakan untuk perbandingan nilai DBI dengan jumlah $k=3$ dan $k=4$.



Gambar 4.15 Nilai DBI *Cluster* 4

Pada Gambar 4.15 ditampilkan hasil nilai *Davies Bouldin Index* untuk $k=4$ (4 *cluster*). Nilai DBI yang dihasilkan akan digunakan untuk perbandingan nilai DBI dengan jumlah $k=2$ dan $k=3$. Perbandingan nilai DBI untuk $k=2$, $k=3$, dan $k=4$ dapat dilihat pada Tabel 4.12.

Tabel 4.12 Hasil Nilai *Davies Bouldin Index*

Jumlah <i>Cluster</i>	Nilai DBI
2	0.914
3	0.629
4	0.801

Tabel 4.12 menunjukkan bahwa nilai pada jumlah $k=3$ lebih kecil dalam mendekati angka 0 dari jumlah $k=2$ dan $k=4$. Sehingga Jumlah $k=3$ lebih optimal dalam klasterisasi data.

4.5 Perhitungan Manual untuk Jumlah *Cluster* 3

Pada tahap ini, pengelompokkan kabupaten/kota dengan algoritma *k-means* dihitung secara manual menggunakan *microsoft excel* dengan jumlah *cluster* 3. Tahapan perhitungan manual dilakukan seperti berikut:

4.5.1 Pengelompokkan Data pada Iterasi 1

Perhitungan manual untuk pengelompokkan data dilakukan lebih utama dengan jumlah $k=3$. Pada iterasi 1, *centroid* dipilih secara random

pada data yang telah dinormalisasikan. Pemilihan *centroid* dapat dilihat pada Tabel 4.13.

Tabel 4.13 *Centroid* Iterasi Pertama

<i>Cluster</i>	Data Ke-n	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Kesulitan Mengakses Layanan Kesehatan
<i>Centroid 0</i>	6	0.142	0.190	0.171	0.247
<i>Centroid 1</i>	2	0.536	0.161	0.304	0.429
<i>Centroid 2</i>	22	0.200	1.000	1.000	0.000

Nilai pada setiap titik data yang telah dinormalisasikan akan dihitung jaraknya dengan *centroid* menggunakan rumus *eucliden distance*. Setelah itu data dikelompokkan berdasarkan jarak terdekat ke *centroid*. Hasil perhitungan jarak ke *centroid* dan pengelompokkannya dapat dilihat pada Tabel 4.14. Sebelum itu ada *rules* dalam melakukan *cluster* data:

1. Jika jarak pada *centroid 0* < jarak pada *centroid 1*, dan jarak pada *centroid 0* < jarak *centroid 2*, maka masuk dalam *cluster 0*.
2. Jika jarak pada *centroid 1* < jarak pada *centroid 0*, dan jarak pada *centroid 1* < jarak *centroid 2*, maka masuk dalam *cluster 1*.
3. Jika jarak pada *centroid 2* < jarak pada *centroid 0*, dan jarak pada *centroid 2* < jarak *centroid 1*, maka masuk dalam *cluster 2*.

Tabel 4.14 Hasil Perhitungan Jarak dan *Cluster* pada Iterasi 1

Kabupaten/ Kota	Jarak ke <i>Centroid</i> 0	Jarak ke <i>Centroid</i> 1	Jarak ke <i>Centroid</i> 2	Jarak terdekat ke <i>Centroid</i>	<i>Cluster</i>
Sumba Barat	0.154	0.501	1.272	0.154	0
Sumba Timur	0.455	0.000	1.219	0.000	1
Kupang	0.645	0.301	1.420	0.301	1
Timor Tengah Selatan	0.892	0.494	1.551	0.494	1
Timor Tengah Utara	0.236	0.291	1.171	0.236	0
Belu	0.000	0.455	1.186	0.000	0
Alor	0.231	0.422	1.139	0.231	0
Lembata	0.197	0.525	1.143	0.197	0
Flores Timur	0.206	0.614	1.171	0.206	0
Sikka	0.180	0.481	1.243	0.180	0
Ende	0.335	0.271	1.030	0.271	1
Ngada	0.340	0.632	0.887	0.340	0
Manggarai	0.341	0.244	1.145	0.244	1
Rote Ndao	0.186	0.427	1.211	0.186	0
Manggarai Barat	0.266	0.279	1.062	0.266	0
Sumba Tengah	0.242	0.559	1.405	0.242	0
Sumba Barat Daya	0.649	0.400	1.538	0.400	1
Nagekeo	0.219	0.598	1.206	0.219	0
Manggarai Timur	0.855	0.631	1.656	0.631	1
Sabu Raijua	0.244	0.592	1.244	0.244	0
Malaka	0.205	0.490	1.351	0.205	0
Kota Kupang	1.186	1.219	0.000	0.000	2

Pada Tabel 4.14 dapat diketahui jarak setiap data ke *centroid* dari masing-masing kabupaten/kota. Hasil *cluster* pada setiap kabupaten/kota dapat dilihat pada Tabel 4.15.

Tabel 4.15 Hasil *Cluster* pada Iterasi 1

<i>Cluster</i> 0	<i>Cluster</i> 1	<i>Cluster</i> 2
Sumba Barat	Sumba Timur	Kota Kupang
Timor Tengah Utara	Kupang	
Belu	Timor Tengah Selatan	
Alor	Ende	

Lembata	Manggarai
Flores Timur	Sumba Barat Daya
Sikka	Manggarai Timur
Ngada	
Rote Ndao	
Manggarai Barat	
Sumba Tengah	
Nagekeo	
Sabu Raijua	
Malaka	

Tabel 4.15 merupakan tabel hasil *cluster* data pada iterasi 1. Pada iterasi 1 ada 14 kabupaten yang masuk ke dalam *cluster* 0, 7 kabupaten masuk ke dalam *cluster* 1, 1 kabupaten masuk *cluster* 2.

4.5.2 Pengelompokkan Data pada Iterasi 2

Setelah mendapatkan hasil pengelompokkan kabupaten/kota pada iterasi 1, maka dilanjutkan lagi pengelompokkan data pada iterasi 2. Untuk menentukan *centroid* pada iterasi 2, dibutuhkan perhitung rata-rata nilai disetiap *cluster* yang terbentuk pada iterasi 1. Untuk mendapatkan *centroid* baru digunakan perhitungan pada Persamaan 4.2.

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \dots \dots \dots (4.2)$$

Keterangan:

V_{ij} = rata-rata pusat *cluster* ke-i untuk variabel ke-j

N_i = jumlah anggota dalam kelompok pertama (*cluster* ke-i)

i, k = indeks yang merujuk ke *cluster*

j = indeks yang merujuk ke variabel

X_{kj} = nilai variabel ke- j pada data ke- k dalam kelompok tersebut

Perhitungan rata-rata nilai disetiap *cluster*, menggunakan nilai yang telah dinormalisasikan dalam *dataset* sebelumnya. Perhitung rata-rata nilai disetiap *cluster* yang terbentuk dilakukan dengan menjumlahkan setiap nilai titik data yang terbentuk dalam masing-masing *cluster*, lalu dibagi dengan banyaknya nilai titik data pada setiap *cluster* tersebut. Rata-rata *cluster* 0 dapat dilihat pada Tabel 4.16.

Tabel 4.16 *Cluster* 0 untuk *Centroid* Baru pada Iterasi 2

<i>Cluster</i> 0	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster</i> 0			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Kesulitan Mengakses Layanan Kesehatan
Sumba Barat	0.168	0.083	0.131	0.147
Timor Tengah Utara	1.000	0.273	0.149	0.307
Belu	0.142	0.190	0.171	0.247
Alor	0.227	0.369	0.120	0.354
Lembata	0.174	0.351	0.079	0.188
Flores Timur	0.089	0.265	0.099	0.076
Sikka	0.221	0.111	0.138	0.110
Ngada	0.017	0.405	0.395	0.181
Rote Ndao	0.282	0.235	0.082	0.178
Manggarai Barat	0.285	0.227	0.366	0.350
Sumba Tengah	0.058	0.000	0.082	0.333
Nagekeo	0.000	0.305	0.099	0.344
Sabu Raijua	0.101	0.048	0.207	0.056
Malaka	0.111	0.116	0.066	0.404
Rata-rata	0.159	0.213	0.156	0.234

Tabel 4.16 menampilkan *cluster 0* yang terbentuk pada iterasi 1. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 2 untuk perhitungan jarak menggunakan rumus *euclidian distance*. Rata-rata nilai normalisasi pada *cluster 1* dapat dilihat pada Tabel 4.17.

Tabel 4.17 *Cluster 1* untuk *Centroid* Baru pada Iterasi 2

<i>Cluster 1</i>	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster 1</i>			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Kesulitan Mengakses Layanan Kesehatan
Sumba Timur	0.536	0.161	0.304	0.429
Kupang	0.655	0.202	0.121	0.633
Timor Tengah Selatan	1.000	0.084	0.157	0.464
Ende	0.423	0.314	0.303	0.236
Manggarai	0.473	0.217	0.239	0.209
Sumba Barat Daya	0.736	0.019	0.000	0.343
Manggarai Timur	0.522	0.157	0.036	1.000
Rata-Rata	0.544	0.142	0.122	0.412

Tabel 4.17 menampilkan *cluster 1* yang terbentuk pada iterasi 1. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 2 untuk perhitungan jarak menggunakan rumus *euclidian distance*. Rata-rata nilai normalisasi pada *cluster 0* dapat dilihat pada Tabel 4.18.

Tabel 4.18 *Cluster 2* untuk *Centroid* Baru pada Iterasi 2

<i>Cluster 2</i>	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster 2</i>			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Kesulitan Mengakses Layanan Kesehatan
Kota Kupang	0.200	0.000	1.000	1.000
Rata-Rata	0.200	0.000	1.000	1.000

Tabel 4.18 menampilkan *cluster 2* yang terbentuk pada iterasi 1. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 2 untuk perhitungan jarak menggunakan rumus *eucliden distance*.

Centroid baru yang dihasilkan dari perhitungan rata-rata pada setiap nilai normalisasi berdasarkan *cluster*-nya dapat dilihat pada Tabel 4.19.

Tabel 4.19 Hasil *Centroid* Baru pada Iterasi 2

<i>Centroid</i>	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Kesulitan Mengakses Layanan Kesehatan
<i>Centroid 0</i>	0.417	0.722	0.161	0.137
<i>Centroid 1</i>	0.174	0.258	0.245	0.188
<i>Centroid 2</i>	0.200	0.000	1.000	1.000

Hasil *centroid* baru pada Tabel 4.19 yang dihasilkan akan digunakan untuk perhitungan jarak dengan nilai-nilai normalisasi pada setiap titik data menggunakan rumus *eucliden distance*. Hasil perhitungan jarak

data ke *centroid* serta hasil *cluster*-nya pada iterasi 2 dapat dilihat pada

Tabel 4.20.

Tabel 4.20 Hasil Perhitungan Jarak dan *Cluster* pada Iterasi 2

Kabupaten/ Kota	Jarak ke <i>Centroid</i> 0	Jarak ke <i>Centroid</i> 1	Jarak ke <i>Centroid</i> 2	Jarak terdekat ke <i>Centroid</i>	<i>Cluster</i>
Sumba Barat	0.158	0.464	1.272	0.158	0
Sumba Timur	0.452	0.183	1.219	0.183	1
Kupang	0.638	0.254	1.420	0.254	1
Timor Tengah Selatan	0.881	0.464	1.551	0.464	1
Timor Tengah Utara	0.216	0.255	1.171	0.216	0
Belu	0.035	0.440	1.186	0.035	0
Alor	0.211	0.394	1.139	0.211	0
Lembata	0.165	0.482	1.143	0.165	0
Flores Timur	0.189	0.579	1.171	0.189	0
Sikka	0.173	0.444	1.243	0.173	0
Ende	0.319	0.329	1.030	0.319	0
Ngada	0.342	0.689	0.887	0.342	0
Manggarai	0.325	0.256	1.145	0.256	1
Rote Ndao	0.156	0.365	1.211	0.156	0
Manggarai Barat	0.272	0.371	1.062	0.272	0
Sumba Tengah	0.266	0.514	1.405	0.266	0
Sumba Barat Daya	0.638	0.268	1.538	0.268	1
Nagekeo	0.222	0.573	1.206	0.222	0
Manggarai Timur	0.858	0.595	1.656	0.595	1
Sabu Raijua	0.254	0.582	1.244	0.254	0
Malaka	0.220	0.437	1.351	0.220	0
Kota Kupang	1.178	1.340	0.000	0.000	2

Pada Tabel 4.20 dapat diketahui jarak setiap data ke *centroid* dan hasil pengelompokkan kabupaten/kota berdasarkan jarak minimum ke *centroid*. Hasil *cluster* data dapat dilihat pada Tabel 4.21.

Tabel 4.21 Hasil Pengelompokkan Data pada Iterasi 2

<i>Cluster 0</i>	<i>Cluster 1</i>	<i>Cluster 2</i>
Sumba Barat	Sumba Timur	Kota Kupang
Timor Tengah Utara	Kupang	
Belu	Timor Tengah Selatan	
Alor	Manggarai	
Lembata	Sumba Barat Daya	
Flores Timur	Manggarai Timur	
Sikka		
Ende		
Ngada		
Rote Ndao		
Manggarai Barat		
Sumba Tengah		
Nagekeo		
Sabu Raijua		
Malaka		

Tabel 4.21 merupakan tabel hasil *cluster* data pada iterasi 2. Pada iterasi 2 ada 15 kabupaten yang masuk ke dalam *cluster 0*, 6 kabupaten masuk ke dalam *cluster 1*, 1 kabupaten masuk *cluster 2*. Hasil *cluster 1* dapat dilihat pada Tabel 4.22.

Tabel 4.22 Perbandingan *Cluster* dengan Iterasi 1

<i>Cluster 0</i>	<i>Cluster 1</i>	<i>Cluster 2</i>
Sumba Barat	Sumba Timur	Kota Kupang
Timor Tengah Utara	Kupang	
Belu	Timor Tengah Selatan	
Alor	Ende	

Lembata	Manggarai
Flores Timur	Sumba Barat Daya
Sikka	Manggarai Timur
Ngada	
Rote Ndao	
Manggarai Barat	
Sumba Tengah	
Nagekeo	
Sabu Raijua	
Malaka	

Pada Tabel 4.21 dan Tabel 4.22, terdapat perubahan dalam *cluster* baru yang berbeda dengan *cluster* sebelumnya pada iterasi 1 yaitu, Kabupaten Ende yang sebelumnya pada iterasi 1 berada di *cluster* 1 sudah berpindah ke *cluster* 0 pada iterasi 2, maka proses *cluster* data akan diulang sampai tidak ada hasil *cluster* data yang berubah.

4.5.3 Pengelompokan Data pada Iterasi 3

Perhitungan rata-rata nilai pada setiap *cluster* yang terbentuk untuk menentukan *centroid* baru pada iterasi 3 dapat dilihat pada Tabel 4.23, Tabel 4.24, dan Tabel 4.25.

Tabel 4.23 *Cluster* 0 untuk *Centroid* Baru pada Iterasi 3

<i>Cluster</i> 0	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster</i> 0			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang sulit Mengkses Layanan Kesehatan
Sumba Barat	0.168	0.083	0.131	0.147
Timor Tengah Utara	0.354	0.273	0.149	0.307

Belu	0.142	0.190	0.171	0.247
Alor	0.227	0.369	0.120	0.354
Lembata	0.174	0.351	0.079	0.188
Flores Timur	0.089	0.265	0.099	0.076
Sikka	0.221	0.111	0.138	0.110
Ende	0.423	0.314	0.303	0.236
Ngada	0.017	0.405	0.395	0.181
Rote Ndao	0.282	0.235	0.082	0.178
Manggarai Barat	0.285	0.227	0.366	0.350
Sumba Tengah	0.058	0.000	0.082	0.333
Nagekeo	0.000	0.305	0.099	0.344
Sabu Raijua	0.101	0.048	0.207	0.056
Malaka	0.111	0.116	0.066	0.404
Rata-rata	0.177	0.220	0.166	0.234

Tabel 4.23 menampilkan *cluster* 0 yang terbentuk pada iterasi 2. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 3 untuk perhitungan jarak menggunakan rumus *euclidian distance*. Hasil perhitungan rata-rata pada *cluster* 1 dapat dilihat pada Tabel 4.24.

Tabel 4.24 *Cluster* 1 untuk *Centroid* Baru pada Iterasi 3

<i>Cluster</i> 1	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster</i> 1			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Sulit Mengakses Layanan Kesehatan
Sumba Timur	0.536	0.161	0.190	0.171
Kupang	0.655	0.202	0.369	0.120
Timor Tengah Selatan	1.000	0.084	0.265	0.099
Manggarai	0.473	0.167	0.217	0.138

Sumba Barat Daya	0.736	0.133	0.019	0.395
Manggarai Timur	0.522	0.347	0.157	0.366
Rata-Rata	0.654	0.140	0.143	0.513

Tabel 4.24 menampilkan *cluster* 1 yang terbentuk pada iterasi 2. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 3 untuk perhitungan jarak menggunakan rumus *eucliden distance*. Hasil perhitungan rata-rata pada *cluster* 2 dapat dilihat pada Tabel 4.25

Tabel 4.25 *Cluster* 2 untuk *Centroid* Baru pada Iterasi 3

<i>Cluster</i> 2	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster</i> 2			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Sulit Mengakses Layanan Kesehatan
Kota Kupang	0.200	0.000	1.000	1.000
Rata-Rata	0.200	0.000	1.000	1.000

Tabel 4.25 menampilkan *cluster* 2 yang terbentuk pada iterasi 2. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 2 untuk perhitungan jarak menggunakan rumus *eucliden distance*. Hasil *centroid* baru dapat dilihat pada Tabel 4.26.

Tabel 4.26 Hasil *Centroid* Baru pada Iterasi 3

<i>Centroid</i>	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Sulit Mengakses Layanan Kesehatan
<i>Centroid</i> 0	0.177	0.220	0.166	0.234
<i>Centroid</i> 1	0.654	0.140	0.143	0.513
<i>Centroid</i> 2	0.200	1.000	1.000	0.000

Hasil *centroid* baru pada Tabel 4.26 yang didapatkan akan digunakan untuk perhitungan jarak nilai-nilai normalisasi pada setiap titik data ke *centroid* menggunakan rumus *eucliden distance*. Hasil perhitungan jarak data ke *centroid* serta hasil *cluster*-nya pada iterasi 3 dapat dilihat pada Tabel 4.27.

Tabel 4.27 Hasil Perhitungan Jarak dan Hasil *Cluster*-nya pada Iterasi 3

Kabupaten/ Kota	Jarak ke <i>Centroid</i> 0	Jarak ke <i>Centroid</i> 1	Jarak ke <i>Centroid</i> 2	Jarak terdekat ke <i>Centroid</i>	<i>Cluster</i>
Sumba Barat	0.166	0.611	1.272	0.166	0
Sumba Timur	0.435	0.217	1.219	0.217	1
Kupang	0.625	0.137	1.420	0.137	1
Timor Tengah Selatan	0.865	0.355	1.551	0.355	1
Timor Tengah Utara	0.199	0.388	1.171	0.199	0
Belu	0.048	0.580	1.186	0.048	0
Alor	0.203	0.510	1.139	0.203	0
Lembata	0.164	0.620	1.143	0.164	0
Flores Timur	0.197	0.726	1.171	0.197	0
Sikka	0.173	0.592	1.243	0.173	0

Ende	0.297	0.432	1.030	0.297	0
Ngada	0.339	0.806	0.887	0.339	0
Manggarai	0.306	0.374	1.145	0.306	0
Rote Ndao	0.147	0.513	1.211	0.147	0
Manggarai Barat	0.256	0.469	1.062	0.256	0
Sumba Tengah	0.281	0.641	1.405	0.281	0
Sumba Barat Daya	0.626	0.266	1.538	0.266	1
Nagekeo	0.235	0.696	1.206	0.235	0
Manggarai Timur	0.853	0.516	1.656	0.516	1
Sabu Raijua	0.261	0.726	1.244	0.261	0
Malaka	0.232	0.559	1.351	0.232	0
Kota Kupang	1.166	1.394	0.000	0.000	2

Pada Tabel 4.27 dapat diketahui jarak setiap titik data ke *centroid* dan hasil pengelompokkan kabupaten/kota berdasarkan jarak minimum ke *centroid*. Hasil *cluster* pada iterasi 3 dapat dilihat pada Tabel 4.28.

Tabel 4.28 Hasil Pengelompokkan Data pada Iterasi 3

<i>Cluster 0</i>	<i>Cluster 1</i>	<i>Cluster 2</i>
Sumba Barat	Sumba Timur	Kota Kupang
Timor Tengah Utara	Kupang	
Belu	Timor Tengah Selatan	
Alor	Sumba Barat Daya	
Lembata	Manggarai Timur	
Flores Timur		
Sikka		
Ende		
Ngada		
Manggarai		
Rote Ndao		
Manggarai Barat		
Sumba Tengah		

Nagekeo
Sabu Raijua
Malaka

Tabel 4.28 merupakan tabel hasil *cluster* data pada iterasi 3. Pada iterasi 3 ada 16 kabupaten yang masuk ke dalam *cluster* 0, 5 kabupaten masuk ke dalam *cluster* 1, 1 kabupaten masuk *cluster* 2. Perbandingan *cluster* dengan Iterasi 1 dapat dilihat pada Tabel 4.29.

Tabel 4.29 Perbandingan *Cluster* dengan Iterasi 1

<i>Cluster</i> 0	<i>Cluster</i> 1	<i>Cluster</i> 2
Sumba Barat	Sumba Timur	Kota Kupang
Timor Tengah Utara	Kupang	
Belu	Timor Tengah Selatan	
Alor	Manggarai	
Lembata	Sumba Barat Daya	
Flores Timur	Manggarai Timur	
Sikka		
Ende		
Ngada		
Rote Ndao		
Manggarai Barat		
Sumba Tengah		
Nagekeo		
Sabu Raijua		
Malaka		

Pada Tabel 4.28 dan Tabel 4.29, terdapat perubahan dalam *cluster* baru yang berbeda dengan *cluster* sebelumnya pada iterasi 2 yaitu , Kabupaten Manggarai yang sebelumnya pada iterasi 2, berada di *cluster*

1, sudah berpindah ke *cluster* 0 pada iterasi 3 maka proses *cluster* akan diulang lagi.

4.5.4 Pengelompokan Data pada Iterasi 4

Perhitungan rata-rata nilai pada setiap *cluster* yang terbentuk untuk menentukan *centroid* baru pada iterasi 4 dapat dilihat pada Tabel 4.30, Tabel 4.31, dan Tabel 4.32.

Tabel 4.30 *Cluster* 0 untuk *Centroid* Baru pada Iterasi 4

<i>Cluster</i> 0	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster</i> 0			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Sulit Mengakses Layanan Kesehatan
Sumba Barat	0.168	0.083	0.131	0.147
Timor Tengah Utara	0.354	0.273	0.149	0.464
Belu	0.142	0.190	0.171	0.307
Alor	0.227	0.369	0.120	0.247
Lembata	0.174	0.351	0.079	0.354
Flores Timur	0.089	0.265	0.099	0.188
Sikka	0.221	0.111	0.138	0.076
Ende	0.423	0.314	0.303	0.110
Ngada	0.017	0.405	0.395	0.236
Manggarai	0.473	0.217	0.239	0.181
Rote Ndao	0.282	0.235	0.082	0.209
Manggarai Barat	0.285	0.227	0.366	0.178
Sumba Tengah	0.058	0.000	0.082	0.350
Nagekeo	0.000	0.305	0.099	0.333
Sabu Raijua	0.101	0.048	0.207	0.056
Malaka	0.111	0.116	0.066	0.404
Rata-rata	0.195	0.219	0.170	0.232

Tabel 4.30 menampilkan *cluster 0* yang terbentuk pada iterasi 3. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 4 untuk perhitungan jarak menggunakan rumus *eucliden distance*. Rata-rata nilai normalisasi pada *cluster 1* dapat dilihat pada Tabel 4.31.

Tabel 4.31 *Cluster 1* untuk *Centroid* Baru pada Iterasi 4

<i>Cluster 1</i>	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster 1</i>			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk Sulit Akses Layanan Kesehatan
Sumba Timur	0.536	0.161	0.304	0.429
Kupang	0.655	0.202	0.121	0.633
Timor Tengah Selatan	1.000	0.084	0.157	0.464
Sumba Barat Daya	0.736	0.019	0.000	0.343
Manggarai Timur	0.522	0.157	0.036	1.000
Rata-Rata	0.690	0.125	0.123	0.574

Tabel 4.31 menampilkan *cluster 1* yang terbentuk pada iterasi 3. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 4 untuk perhitungan jarak menggunakan rumus *eucliden distance*. Rata-rata nilai normalisasi pada *cluster 2* dapat dilihat pada Tabel 4.32.

Tabel 4.32 *Cluster 2* untuk *Centroid* Baru pada Iterasi 4

<i>Cluster 2</i>	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster 2</i>			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk Sulit Akses Layanan Kesehatan
Kota Kupang	0.200	0.000	1.000	1.000
Rata-Rata	0.200	0.000	1.000	1.000

Tabel 4.32 menampilkan *cluster 2* yang terbentuk pada iterasi 3. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 4 untuk perhitungan jarak menggunakan rumus *eucliden distance*.

Tabel 4.33 Hasil *Centroid* Baru pada Iterasi 4

<i>Centroid</i>	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Sulit Akses Layanan Kesehatan
<i>Centroid 0</i>	0.195	0.219	0.170	0.232
<i>Centroid 1</i>	0.690	0.125	0.123	0.574
<i>Centroid 2</i>	0.200	1.000	1.000	0.000

Hasil *centroid* baru pada Tabel 4.33 yang didapatkan akan digunakan untuk perhitungan jarak dengan nilai-nilai normalisasi pada setiap titik data menggunakan rumus *eucliden distance*. Hasil perhitungan jarak data ke *centroid* serta hasil *cluster*-nya pada iterasi 4 dapat dilihat pada Tabel 4.34.

Tabel 4.34 Hasil Perhitungan Jarak dan Hasil *Cluster*-nya pada Iterasi 4

Kabupaten/ Kota	Jarak ke <i>Centroid</i> 0	Jarak ke <i>Centroid</i> 1	Jarak ke <i>Centroid</i> 2	Jarak terdekat ke <i>Centroid</i>	<i>Cluster</i>
Sumba Barat	0.166	0.611	1.272	0.166	0
Sumba Timur	0.435	0.217	1.219	0.217	1
Kupang	0.625	0.137	1.420	0.137	1
Timor Tengah Selatan	0.865	0.355	1.551	0.355	1
Timor Tengah Utara	0.199	0.388	1.171	0.199	0
Belu	0.048	0.580	1.186	0.048	0
Alor	0.203	0.510	1.139	0.203	0
Lembata	0.164	0.620	1.143	0.164	0
Flores Timur	0.197	0.726	1.171	0.197	0
Sikka	0.173	0.592	1.243	0.173	0
Ende	0.297	0.432	1.030	0.297	0
Ngada	0.339	0.806	0.887	0.339	0
Manggarai	0.306	0.374	1.145	0.306	0
Rote Ndao	0.147	0.513	1.211	0.147	0
Manggarai Barat	0.256	0.469	1.062	0.256	0
Sumba Tengah	0.281	0.641	1.405	0.281	0
Sumba Barat Daya	0.626	0.266	1.538	0.266	1
Nagekeo	0.235	0.696	1.206	0.235	0
Manggarai Timur	0.853	0.516	1.656	0.516	1
Sabu Raijua	0.261	0.726	1.244	0.261	0
Malaka	0.232	0.559	1.351	0.232	0
Kota Kupang	1.166	1.394	0.000	0.000	2

Pada Tabel 4.34 dapat diketahui jarak setiap *centroid* dan hasil *cluster* kabupaten/kota berdasarkan jarak paling minimum ke *centroid*.

Hasil *cluster* pada iterasi 4 dapat dilihat pada Tabel 4.35.

Tabel 4.35 Hasil Pengelompokan Data pada Iterasi 4

<i>Cluster 0</i>	<i>Cluster 1</i>	<i>Cluster 2</i>
Sumba Barat	Sumba Timur	Kota Kupang
Timor Tengah Utara	Kupang	
Belu	Timor Tengah Selatan	
Alor	Sumba Barat Daya	
Lembata	Manggarai Timur	
Flores Timur		
Sikka		
Ende		
Ngada		
Manggarai		
Rote Ndao		
Manggarai Barat		
Sumba Tengah		
Nagekeo		
Sabu Raijua		
Malaka		

Tabel 4.35 merupakan tabel hasil *cluster* data pada iterasi 4. Pada iterasi 4 ada 16 kabupaten yang masuk ke dalam *cluster 0*, 5 kabupaten masuk ke dalam *cluster 1*, 1 kabupaten masuk *cluster 2*. Perbandingan *cluster* dengan iterasi 3 dapat dilihat pada Tabel 4.36.

Tabel 4.36 Perbandingan *Cluster* dengan Iterasi 3

<i>Cluster 0</i>	<i>Cluster 1</i>	<i>Cluster 2</i>
Sumba Barat	Sumba Timur	Kota Kupang
Timor Tengah Utara	Kupang	
Belu	Timor Tengah Selatan	
Alor	Sumba Barat Daya	
Lembata	Manggarai Timur	

Flores Timur
Sikka
Ende
Ngada
Manggarai
Rote Ndao
Manggarai Barat
Sumba Tengah
Nagekeo
Sabu Raijua
Malaka

Pada Tabel 4.35 dan Tabel 4.36, perbandingan hasil *cluster* pada iterasi 3 dan pada iterasi 4 tidak ada perubahan, hasil *clusternya* sama. Namun pada iterasi 3 terdapat perbedaan *cluster* data dengan iterasi 2 yaitu, Kabupaten Manggarai yang sebelumnya pada iterasi 2, berada di *cluster* 1, sudah berpindah ke *cluster* 0 pada iterasi 3 maka proses *cluster* akan diulang lagi

Oleh karena itu, untuk memastikan tidak ada lagi pengelompokan kabupaten/kota yang berbeda pada iterasi 4, maka proses *cluster* data akan dilakukan lagi pada iterasi 5 dan 6 untuk melihat konsistensi dan optimasi hasil *cluster*.

4.5.5 Pengelompokan Data pada Iterasi 5

Perhitungan rata-rata nilai pada setiap *cluster* yang terbentuk untuk menentukan *centroid* baru pada iterasi 5 dan dapat dilihat pada Tabel 4.37, Tabel 4.38, dan Tabel 4.39.

Tabel 4.37 *Cluster 0* untuk *Centroid* Baru pada Iterasi 5

<i>Cluster 0</i>	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster 0</i>			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Sulit Mengakses Layanan Kesehatan
Sumba Barat	0.168	0.083	0.131	0.147
Timor Tengah Utara	0.354	0.273	0.149	0.464
Belu	0.142	0.190	0.171	0.307
Alor	0.227	0.369	0.120	0.247
Lembata	0.174	0.351	0.079	0.354
Flores Timur	0.089	0.265	0.099	0.188
Sikka	0.221	0.111	0.138	0.076
Ende	0.423	0.314	0.303	0.110
Ngada	0.017	0.405	0.395	0.236
Manggarai	0.473	0.217	0.239	0.181
Rote Ndao	0.282	0.235	0.082	0.209
Manggarai Barat	0.285	0.227	0.366	0.178
Sumba Tengah	0.058	0.000	0.082	0.350
Nagekeo	0.000	0.305	0.099	0.333
Sabu Raijua	0.101	0.048	0.207	0.056
Malaka	0.111	0.116	0.066	0.404
Rata-rata	0.195	0.219	0.170	0.232

Tabel 4.37 menampilkan *cluster 0* yang terbentuk pada iterasi 4. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 5 untuk perhitungan jarak menggunakan rumus *euclidian distance*.

Tabel 4.38 *Cluster 1* untuk *Centroid* Baru pada Iterasi 5

<i>Cluster 1</i>	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster 1</i>			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk Sulit Akses Layanan Kesehatan
Sumba Timur	0.536	0.161	0.304	0.429
Kupang	0.655	0.202	0.121	0.633
Timor Tengah Selatan	1.000	0.084	0.157	0.464
Sumba Barat Daya	0.736	0.019	0.000	0.343
Manggarai Timur	0.522	0.157	0.036	1.000
Rata-Rata	0.690	0.125	0.123	0.574

Tabel 4.38 menampilkan *cluster 1* yang terbentuk pada iterasi 4. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 5 untuk perhitungan jarak menggunakan rumus *euclidian distance*.

Tabel 4.39 *Cluster 2* untuk *Centroid* Baru pada Iterasi 5

<i>Cluster 2</i>	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster 2</i>			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk Sulit Akses Layanan Kesehatan
Kota Kupang	0.200	0.000	1.000	1.000
Rata-Rata	0.200	0.000	1.000	1.000

Tabel 4.39 menampilkan *cluster 2* yang terbentuk pada iterasi 4. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data

akan digunakan sebagai *centroid* baru pada iterasi 5 untuk perhitungan jarak menggunakan rumus *eucliden distance*.

Tabel 4.40 Hasil *Centroid* Baru pada Iterasi 5

<i>Centroid</i>	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Sulit Akses Layanan Kesehatan
<i>Centroid</i> 0	0.195	0.219	0.170	0.232
<i>Centroid</i> 1	0.690	0.125	0.123	0.574
<i>Centroid</i> 2	0.200	1.000	1.000	0.000

Hasil *centroid* baru pada Tabel 4.40 yang didapatkan akan digunakan untuk perhitungan jarak dengan nilai-nilai normalisasi pada setiap titik data menggunakan rumus *eucliden distance*. Hasil *centroid* pada iteasi 5 dan iterasi 4 terlihat sama. Hasil perhitungan jarak data ke *centroid* serta hasil *cluster*-nya pada iterasi 5 dapat dilihat pada Tabel 4.41.

Tabel 4.41 Hasil Perhitungan Jarak dan Hasil *Cluster*-nya pada Iterasi 5

Kabupaten/ Kota	Jarak ke <i>Centroid</i> 0	Jarak ke <i>Centroid</i> 1	Jarak ke <i>Centroid</i> 2	Jarak terdekat ke <i>Centroid</i>	<i>Cluster</i>
Sumba Barat	0.168	0.676	1.272	0.166	0
Sumba Timur	0.420	0.280	1.219	0.217	1
Kupang	0.612	0.104	1.420	0.137	1
Timor Tengah Selatan	0.848	0.333	1.551	0.355	1
Timor Tengah Utara	0.184	0.455	1.171	0.199	0
Belu	0.063	0.643	1.186	0.048	0
Alor	0.202	0.568	1.139	0.203	0

Lembata	0.168	0.684	1.143	0.164	0
Flores Timur	0.207	0.793	1.171	0.197	0
Sikka	0.169	0.660	1.243	0.173	0
Ende	0.280	0.504	1.030	0.297	0
Ngada	0.345	0.871	0.887	0.339	0
Manggarai	0.287	0.449	1.145	0.306	0
Rote Ndao	0.137	0.580	1.211	0.147	0
Manggarai Barat	0.246	0.533	1.062	0.256	0
Sumba Tengah	0.291	0.689	1.405	0.281	0
Sumba Barat Daya	0.611	0.286	1.538	0.266	1
Nagekeo	0.251	0.750	1.206	0.235	0
Manggarai Timur	0.847	0.468	1.656	0.516	1
Sabu Raijua	0.266	0.792	1.244	0.261	0
Malaka	0.24s1	0.606	1.351	0.232	0
Kota Kupang	1.163	1.450	0.000	0.000	2

Pada Tabel 4.41 dapat diketahui jarak setiap data ke *centroid* dan hasil *cluster* dari masing-masing kabupaten/kota. hasil *cluster*-nya pada iterasi 5 sama dengan iterasi ke-4. Hasil *cluster* pada iterasi 5 dapat dilihat pada Tabel 4.42.

Tabel 4.42 Hasil Pengelompokkan Data pada Iterasi 5

<i>Cluster 0</i>	<i>Cluster 1</i>	<i>Cluster 2</i>
Sumba Barat	Sumba Timur	Kota Kupang
Timor Tengah Utara	Kupang	
Belu	Timor Tengah Selatan	
Alor	Sumba Barat Daya	
Lembata	Manggarai Timur	
Flores Timur		
Sikka		
Ende		
Ngada		

Manggarai
Rote Ndao
Manggarai Barat
Sumba Tengah
Nagekeo
Sabu Raijua
Malaka

Tabel 4.42 merupakan tabel hasil *cluster* data pada iterasi 5. Pada iterasi 5, ada 16 kabupaten yang masuk ke dalam *cluster* 0, 5 kabupaten masuk ke dalam *cluster* 1, 1 kabupaten masuk *cluster* 2.

Tabel 4.43 Perbandingan *Cluster* dengan Iterasi 4

<i>Cluster</i> 0	<i>Cluster</i> 1	<i>Cluster</i> 2
Sumba Barat	Sumba Timur	Kota Kupang
Timor Tengah Utara	Kupang	
Belu	Timor Tengah Selatan	
Alor	Sumba Barat Daya	
Lembata	Manggarai Timur	
Flores Timur		
Sikka		
Ende		
Ngada		
Manggarai		
Rote Ndao		
Manggarai Barat		
Sumba Tengah		
Nagekeo		
Sabu Raijua		
Malaka		

Perbandingan *centroid* dan hasil *cluster* pada iterasi 4 dan pada iterasi 5 tidak ada perubahan. Proses pengelompokkan data dilanjut pada iterasi 6 untuk lebih memastikan konsistensi hasil *cluster*.

4.5.6 Pengelompokkan Data pada Iterasi 6

Perhitungan rata-rata nilai pada setiap *cluster* yang terbentuk pada iterasi 5 untuk menentukan *centroid* baru pada iterasi 6 dan dapat dilihat pada Tabel 4.44 Tabel 4.45, dan Tabel 4.46.

Tabel 4.44 *Cluster 0* untuk *Centroid* Baru pada Iterasi 6

<i>Cluster 0</i>	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster 0</i>			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Sulit Mengakses Layanan Kesehatan
Sumba Barat	0.168	0.083	0.131	0.147
Timor Tengah Utara	0.354	0.273	0.149	0.464
Belu	0.142	0.190	0.171	0.307
Alor	0.227	0.369	0.120	0.247
Lembata	0.174	0.351	0.079	0.354
Flores Timur	0.089	0.265	0.099	0.188
Sikka	0.221	0.111	0.138	0.076
Ende	0.423	0.314	0.303	0.110
Ngada	0.017	0.405	0.395	0.236
Manggarai	0.473	0.217	0.239	0.181
Rote Ndao	0.282	0.235	0.082	0.209
Manggarai Barat	0.285	0.227	0.366	0.178
Sumba Tengah	0.058	0.000	0.082	0.350
Nagekeo	0.000	0.305	0.099	0.333
Sabu Raijua	0.101	0.048	0.207	0.056

Malaka	0.111	0.116	0.066	0.404
Rata-rata	0.195	0.219	0.170	0.232

Tabel 4.44 menampilkan *cluster* 0 yang terbentuk pada iterasi 5. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 6 untuk perhitungan jarak menggunakan rumus *eucliden distance*.

Tabel 4.45 *Cluster* 1 untuk *Centroid* Baru pada Iterasi 6

<i>Cluster</i> 1	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster</i> 1			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk Sulit Akses Layanan Kesehatan
Sumba Timur	0.536	0.161	0.304	0.429
Kupang	0.655	0.202	0.121	0.633
Timor Tengah Selatan	1.000	0.084	0.157	0.464
Sumba Barat Daya	0.736	0.019	0.000	0.343
Manggarai Timur	0.522	0.157	0.036	1.000
Rata-Rata	0.690	0.125	0.123	0.574

Tabel 4.45 menampilkan *cluster* 1 yang terbentuk pada iterasi 5. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 6 untuk perhitungan jarak menggunakan rumus *eucliden distance*.

Tabel 4.46 *Cluster 2* untuk *Centroid* Baru pada Iterasi 6

<i>Cluster 2</i>	Nilai Data yang Telah Dinormalisasikan pada <i>Cluster 2</i>			
	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk Sulit Akses Layanan Kesehatan
Kota Kupang	0.200	0.000	1.000	1.000
Rata-Rata	0.200	0.000	1.000	1.000

Tabel 4.46 menampilkan *cluster 2* yang terbentuk pada iterasi 5. Hasil perhitungan rata-rata nilai normalisasi pada keempat variabel data akan digunakan sebagai *centroid* baru pada iterasi 6 untuk perhitungan jarak menggunakan rumus *eucliden distance*.

Tabel 4.47 Hasil *Centroid* Baru pada Iterasi 6

<i>Centroid</i>	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Sulit Akses Layanan Kesehatan
<i>Centroid 0</i>	0.195	0.219	0.170	0.232
<i>Centroid 1</i>	0.690	0.125	0.123	0.574
<i>Centroid 2</i>	0.200	1.000	1.000	0.000

Hasil *centroid* baru pada Tabel 4.47 yang didapatkan akan digunakan untuk perhitungan jarak nilai-nilai normalisasi pada setiap titik data ke *centroid* menggunakan rumus *eucliden distance*. hasil *cluster* pada iterasi 6 dapat dilihat pada Tabel 4.48.

Tabel 4.48 Hasil Perhitungan Jarak dan Hasil *Cluster*-nya pada Iterasi 6

Kabupaten/ Kota	Jarak ke <i>Centroid</i> 0	Jarak ke <i>Centroid</i> 1	Jarak ke <i>Centroid</i> 2	Jarak terdekat ke <i>Centroid</i>	<i>Cluster</i>
Sumba Barat	0.168	0.676	1.272	0.166	0
Sumba Timur	0.420	0.280	1.219	0.217	1
Kupang	0.612	0.104	1.420	0.137	1
Timor Tengah Selatan	0.848	0.333	1.551	0.355	1
Timor Tengah Utara	0.184	0.455	1.171	0.199	0
Belu	0.063	0.643	1.186	0.048	0
Alor	0.202	0.568	1.139	0.203	0
Lembata	0.168	0.684	1.143	0.164	0
Flores Timur	0.207	0.793	1.171	0.197	0
Sikka	0.169	0.660	1.243	0.173	0
Ende	0.280	0.504	1.030	0.297	0
Ngada	0.345	0.871	0.887	0.339	0
Manggarai	0.287	0.449	1.145	0.306	0
Rote Ndao	0.137	0.580	1.211	0.147	0
Manggarai Barat	0.246	0.533	1.062	0.256	0
Sumba Tengah	0.291	0.689	1.405	0.281	0
Sumba Barat Daya	0.611	0.286	1.538	0.266	1
Nagekeo	0.251	0.750	1.206	0.235	0
Manggarai Timur	0.847	0.468	1.656	0.516	1
Sabu Raijua	0.266	0.792	1.244	0.261	0
Malaka	0.241	0.606	1.351	0.232	0
Kota Kupang	1.163	1.450	0.000	0.000	2

Pada Tabel 4.48 dapat diketahui jarak setiap data ke *centroid* dan hasil *cluster* dari masing-masing kabupaten/kota. Hasil Jarak data ke *centroid* dan hasil *cluster*-nya pada iterasi 6 dapat dilihat pada Tabel 4.49.

Tabel 4.49 Hasil Pengelompokan Data pada Iterasi 5

<i>Cluster 0</i>	<i>Cluster 1</i>	<i>Cluster 2</i>
Sumba Barat	Sumba Timur	Kota Kupang
Timor Tengah Utara	Kupang	
Belu	Timor Tengah Selatan	
Alor	Sumba Barat Daya	
Lembata	Manggarai Timur	
Flores Timur		
Sikka		
Ende		
Ngada		
Manggarai		
Rote Ndao		
Manggarai Barat		
Sumba Tengah		
Nagekeo		
Sabu Raijua		
Malaka		

Tabel 4.49 merupakan tabel hasil *cluster* data pada iterasi 5. Pada iterasi 5, ada 16 kabupaten yang masuk ke dalam *cluster 0*, 5 kabupaten masuk ke dalam *cluster 1*, 1 kabupaten masuk *cluster 2*. perbandingan *cluster* dengan iterasi 6 dapat dilihat pada Tabel 4.50.

Tabel 4.50 Perbandingan *Cluster* dengan Iterasi 6

<i>Cluster 0</i>	<i>Cluster 1</i>	<i>Cluster 2</i>
Sumba Barat	Sumba Timur	Kota Kupang
Timor Tengah Utara	Kupang	
Belu	Timor Tengah Selatan	
Alor	Sumba Barat Daya	
Lembata	Manggarai Timur	
Flores Timur		
Sikka		

Ende
Ngada
Manggarai
Rote Ndao
Manggarai Barat
Sumba Tengah
Nagekeo
Sabu Raijua
Malaka

Perbandingan hasil optimalisasi pada *cluster* pada iterasi 4, 5, dan 6 tidak ada lagi perubahan dan tetap konsisten. Hasil ini menunjukkan bahwa hasil klasterisasi pada iterasi 6 telah konvergen. Maka proses pengelompok data menggunakan algoritma *k-means* dihentikan.

4.6 Perhitungan Manual Evaluasi Jumlah *Cluster* untuk K=3

Perhitungan manual untuk mencari nilai DBI, dilakukan pada iterasi terakhir (iterasi 6), yang pengelompokan datanya tidak mengalami perubahan lagi. Langkah-langkah perhitungan manual untuk evaluasi jumlah *cluster* dipaparkan sebagai berikut:

4.6.1 Menghitung *Sum of Squares Within-Cluster (SSW)*

Perhitungan SSW membutuhkan jarak data ke *centroid* yang telah terbentuk pada iterasi 6. Tabel jarak data ke *centroid* pada masing-masing *cluster* dapat dilihat pada Tabel 4.51, Tabel 4.52, dan Tabel 4.53.

Tabel 4.51 Jarak *Cluster 0* ke *Centroid* pada Iterasi 6

<i>Cluster 0</i>
0.168
0.184

0.063
0.202
0.168
0.207
0.169
0.280
0.345
0.287
0.137
0.246
0.291
0.251
0.266
0.241

Tabel 4.51 merupakan jarak data ke *centroid* pada *cluster* 0 pada iterasi 6 untuk jumlah k=3. Jarak data tersebut akan digunakan untuk menghitung SSW. Berikut ini adalah perhitungan SSW untuk *cluster* 0.

$$SSW_0 = \frac{1}{16} \times (0.168 + 0.184 + 0.063 + 0.202 + 0.168 + 0.207 + 0.169 + 0.280 + 0.345 + 0.287 + 0.137 + 0.246 + 0.291 + 0.251 + 0.266 + 0.241) = \mathbf{0.219}$$

Tabel 4.52 Jarak *Cluster* 1 ke *Centroid* pada Iterasi 6

<i>Cluster</i> 1
0.280
0.104
0.333
0.286
0.468

Tabel 4.52 merupakan jarak data dengan *centroid* pada *cluster* 1 di iterasi 6 untuk jumlah $k=3$. Jarak data tersebut akan digunakan untuk menghitung SSW. Berikut ini adalah perhitungan SSW untuk *cluster* 1.

$$SSW_1 = \frac{1}{5} \times (0.217 + 0.137 + 0.355 + 0.266 + 0.516) = \mathbf{0.294}$$

Tabel 4.53 Jarak *Cluster* 2 ke *Centroid* pada Iterasi 6

<i>Cluster</i> 2
0.00

Tabel 4.53 merupakan jarak data dengan *centroid* pada *cluster* 0 di iterasi 6. Jarak data tersebut akan digunakan untuk menghitung SSW.

$$SSW_2 = \frac{1}{1} \times (0) = \mathbf{0}$$

4.6.2 Menghitung *Sum of Squares Between-Cluster* (SSB)

Tabel 4.54 *Centroid* Iterasi 5 untuk Perhitungan SSB

<i>Centroid</i>	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang Sulit Mengakses Layanan Kesehatan
<i>Centroid</i> 0	0.195	0.219	0.170	0.232
<i>Centroid</i> 1	0.690	0.125	0.123	0.574
<i>Centroid</i> 2	0.200	1.000	1.000	0.000

Tabel 4.54 merupakan *centroid* dari iterasi 6 yang akan digunakan untuk menghitung *Between-Cluster Sum of Squares* (SSB).

1. Menghitung SSB *Centroid 0 dan Centroid 1*

$$SSB_{(0,1)} = \sqrt{\begin{matrix} (0.195-0.690)^2 + (0.219 - 0.125)^2 + \\ (0.170 + 0.123)^2 + (0.232 + 0.574)^2 \end{matrix}} = \mathbf{0.610}$$

2. Menghitung SSB *Centroid 0 dan Centroid 2*

$$SSB_{(0,2)} = \sqrt{\begin{matrix} (0.195-0.200)^2 + (0.219 - 1.000)^2 + \\ (0.170 + 1.000)^2 + (0.232 + 0.000)^2 \end{matrix}} = \mathbf{1.163}$$

3. Menghitung SSB *Centroid 0 dan Centroid 2*

$$SSB_{(1,2)} = \sqrt{\begin{matrix} (0.690 - 0.200)^2 + (0.125 - 1.000)^2 + \\ (0.123 + 1.000)^2 + (0.574 + 0.000)^2 \end{matrix}} = \mathbf{1.450}$$

4.6.3 Menghitung Rasio

Perhitungan rasio membutuhkan hasil perhitungan dari SSW (*Sum of Squares Within-cluster*) dan SSB (*Sum of Squares Between-cluster*). Hasil perhitungan tersebut disajikan secara terperinci dalam Tabel 4.55 untuk SSW dan Tabel 4.56 untuk SSB.

Tabel 4.55 Hasil Perhitungan SSW

SSW ₀	SSW ₁	SSW ₂
0.219	0.298	0

Tabel 4.56 Hasil Perhitungan SSB

SSB _(0,1)	SSB _(0,2)	SSB _(1,2)
0.610	0.163	1.450

1. Menghitung $R_{0,1} = \frac{SSW_0 + SSW_1}{SSB_{0,1}} = \frac{0.219+0.298}{0.610} = \mathbf{0.848}$

2. Menghitung $R_{0,2} = \frac{SSW_0 + SSW_2}{SSB_{0,2}} = \frac{0.219 + 0}{1.163} = \mathbf{0.188}$

3. Menghitung $R_{1,2} = \frac{SSW_1 + SSW_2}{SSB_{1,2}} = \frac{0.298 + 0}{1.450} = \mathbf{0.203}$

Tabel 4.57 Matriks Rasio

Hasil Matriks Rasio				
Rasio	0	1	2	Nilai <i>Max</i> Rasio
0	0	0.848	0.188	0.848
1	0.848	0	0.206	0.848
2	0.188	0.206	0	0.203

Tabel 4.57 merupakan matriks hasil perhitungan rasio, dari matriks tersebut akan ditentukan nilai maksimum dari masing-masing nilai rasio. Nilai maksimum atau nilai paling tinggi dari matriks rasio adalah 0.848, dan 0.203. Nilai maksimum rasio yang dihasilkan akan digunakan untuk mendapatkan nilai DBI (*Davies Bouldin Index*).

4.6.4 Menghitung Nilai DBI (*Davies Bouldin Index*)

Pada tahap ini, perhitungan nilai DBI akan menggunakan nilai *maximum* dari hasil perhitungan matriks rasio. Berikut ini adalah perhitungan nilai DBI.

$$DBI = \frac{1}{K} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \dots \dots \dots (4.3)$$

Nilai DBI untuk K=3: $\frac{1}{3} \times (0.848 + 0.848 + 0.206) = 0.634$

Nilai DBI untuk jumlah k = 3 telah didapatkan dalam perhitungannya menggunakan *microsoft excel*.

4.7 Analisis Karakteristik *Cluster*

Setelah melakukan pengelompokkan data dan evaluasi jumlah *cluster* menggunakan *tools Rapid Miner* dan *microsoft excel*, hal yang perlu dilakukan adalah menganalisis karakteristik dari masing-masing *cluster* yang terbentuk untuk mengetahui setiap *cluster* merepresentasikan pembangunan sosial-

ekonomi yang rendah, sedang atau tinggi. Oleh karena itu, dibutuhkan nilai dari 4 variabel data yang telah dihitung rata-ratanya untuk dilakukan analisis deskriptif (Utari, 2021). Tabel 4.58 akan menampilkan nilai rata-rata dari 4 variabel data serta hasil *cluster* pada *tools Rapid Miner* yang ditandai dengan 3 warna yang mewakili setiap *cluster* (c0, c1, dan c2).

Tabel 4.58 Hasil Perhitungan Rata-rata untuk Analisis Deskriptif

Kabupaten/ Kota	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang sulit Mengakses Layanan Kesehatan
Sumba Barat	36,927	6.7	725,324	1.46
Sumba Timur	76,991	7.1	862,789	3.56
Kupang	89,926	7.3	717,069	5.08
Timor Tengah Selatan	127,430	6.7	745,899	3.82
Timor Tengah Utara	57,143	7.7	739,269	2.65
Belu	34,100	7.2	757,070	2.20
Alor	43,420	8.2	716,735	3.00
Lembata	37,594	8.1	683,592	1.77
Flores Timur	28,370	7.6	699,635	0.94
Sikka	42,757	6.8	730,980	1.18
Ende	64,667	7.9	862,726	2.12
Ngada	20,517	8.4	935,430	1.71
Manggarai	70,106	7.4	811,073	1.93
Rote Ndao	49,401	7.5	685,758	1.69
Manggarai Barat	49,657	7.4	912,867	2.97
Sumba Tengah	25,000	6.2	686,055	2.85
Sumba Barat Daya	98,751	6.3	620,602	2.92
Nagekeo	18,684	7.9	699,402	2.93
Manggarai Timur	75,481	7.1	649,133	7.81
Sabu Raijua	29,680	6.5	785,440	0.79
Malaka	30,784	6.9	672,865	3.37
Kota Kupang	40,487	11.5	1,418,452	0.37

Tabel 4.58 menampilkan warna yang berbeda pada setiap kabupaten/kota yang masuk ke dalam *cluster* 0, 1, dan 2. Untuk *cluster* 0 ditandai dengan warna biru, *cluster* 1 ditandai dengan warna hijau, dan *cluster* 2 ditandai dengan warna *orange* untuk mempermudah dalam analisis karakteristik *cluster* dan membedakan masing-masing *cluster*. Setelah itu memisahkan setiap kabupaten/kota berdasarkan *cluster*-nya dalam bentuk tabel. Hasil *cluster* 0 dan nilai rata-rata pada setiap variabel datanya dapat dilihat pada Tabel 4.59.

Tabel 4.59 Analisis Karakteristik *Cluster* 0

Kabupaten/ Kota	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang sulit Mengakses Layanan Kesehatan
Sumba Barat	36,927	6.7	725,324	1.46
Timor Tengah Utara	57,143	7.7	739,269	2.65
Belu	34,100	7.2	757,070	2.20
Alor	43,420	8.2	716,735	3.00
Lembata	37,594	8.1	683,592	1.77
Flores Timur	28,370	7.6	699,635	0.94
Sikka	42,757	6.8	730,980	1.18
Ende	64,667	7.9	862,726	2.12
Ngada	20,517	8.4	935,430	1.71
Manggarai	70,106	7.4	811,073	1.93
Rote Ndao	49,401	7.5	685,758	1.69
Manggarai Barat	49,657	7.4	912,867	2.97
Sumba Tengah	25,000	6.2	686,055	2.85
Nagekeo	18,684	7.9	699,402	2.93
Sabu Raijua	29,680	6.5	785,440	0.79
Malaka	30,784	6.9	672,865	3.37

Tabel 4.59 merupakan hasil *cluster* 0. Analisis dilakukan dengan melihat *range* atau parameter (batas minimum dan batas maksimum nilai data) dari

nilai setiap variabel data pada *cluster* 0. Analisis karakteristik *cluster* 1 menggunakan nilai *range* dapat dilihat pada Tabel 4.60.

Tabel 4.60 Nilai *Range Cluster* 0

Variabel Data	<i>Range</i>
Jumlah Penduduk Miskin	18,684 sampai 70,106
Rata-rata Lamanya Sekolah	6.2 sampai 8.4
Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	672,865 sampai 912,867
Persentase Penduduk yang sulit Mengakses Layanan Kesehatan	0.79 sampai 3.37

Penggunaan nilai *range* pada Tabel 4.60 berfungsi untuk menganalisis perbandingan nilai *range* antara *cluster* 0 dengan *cluster* 1 dan 2, sehingga dapat ditemukan karakteristik dari *cluster* 0, dan dapat menentukan *cluster* 0 termasuk pembangunan sosial-ekonomi yang rendah, sedang atau tinggi. Hasil *cluster* 1 dan nilai rata-rata pada setiap variabel datanya dapat dilihat pada Tabel 4.61.

Tabel 4.61 Analisis Karakteristik *Cluster* 1

Kabupaten/ Kota	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang sulit Mengakses Layanan Kesehatan
Sumba Timur	76,991	7.1	862,789	3.56
Kupang	89,926	7.3	717,069	5.08
Timor Tengah Selatan	127,430	6.7	745,899	3.82
Sumba Barat Daya	98,751	6.3	620,602	2.92
Manggarai Timur	75,481	7.1	649,133	7.81

Tabel 4.61 merupakan hasil *cluster* 1. Analisis dilakukan dengan melihat *range* atau parameter nilai pada setiap variabel data pada *cluster* 1. Analisis karakteristik *cluster* 1 menggunakan nilai *range* dapat dilihat pada Tabel 4.62.

Tabel 4.62 Nilai *Range Cluster 1*

Variabel Data	<i>Range</i>
Jumlah Penduduk Miskin	75,481 sampai 127,430
Rata-rata Lamanya Sekolah	6.3 sampai 7.3
Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	620,602 sampai 862,789
Persentase Penduduk yang sulit Mengakses Layanan Kesehatan	2.92 sampai 7.81

Sumber: Primer

Penggunaan nilai *range* pada Tabel 4.62 berfungsi untuk menganalisis perbandingan nilai *range* antara *cluster 1* dengan *cluster 0* dan *2*, sehingga dapat ditemukan karakteristik dari *cluster 1*, dan dapat menentukan *cluster 1* termasuk pembangunan sosial-ekonomi yang rendah, sedang atau tinggi. Hasil *cluster 2* dan nilai rata-rata pada setiap variabel datanya dapat dilihat pada Tabel 4.63.

Tabel 4.63 Analisis Karakteristik *Cluster 2*

Kabupaten/ Kota	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk yang sulit Mengakses Layanan Kesehatan
Kota Kupang	40,487	11.5	1,418,452	0.37

Tabel 4.63 merupakan tabel *cluster 2*. Analisis dilakukan dengan melihat *range* atau parameter nilai pada setiap variabel data pada *cluster 2*. Analisis karakteristik *cluster 2* menggunakan nilai *range* dapat dilihat pada Tabel 4.64.

Tabel 4.64 Nilai *Range Cluster 2*

Variabel Data	<i>Range</i>
Jumlah Penduduk Miskin	40,487
Rata-rata Lamanya Sekolah	11.5
Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	1,418,452
Persentase Penduduk yang sulit Mengakses Layanan Kesehatan	0.37

Penggunaan nilai *range* pada Tabel 4.64 berfungsi untuk menganalisis perbandingan nilai *range* antara *cluster 2* dengan *cluster 0* dan *1*, sehingga dapat ditemukan karakteristik dari *cluster 2*, dan dapat menentukan *cluster 2* termasuk pembangunan sosial-ekonomi yang rendah, sedang atau tinggi.

Nilai *range* atau parameter dari masing-masing *cluster* (*cluster 0*, *1*, dan *2*) yang telah di paparkan pada Tabel 4.60, Tabel 4.62, dan Tabel 4.64 akan digunakan untuk analisis karakteristik pada setiap *cluster* dengan membandingkan nilai *range*-nya. Untuk itu Pada Tabel 4.65 akan di tampilkan perbandingannya.

Tabel 4.65 Perbandingan Nilai *Range* Setiap *Cluster*

Variabel Data	<i>Range Cluster 0</i>	<i>Range Cluster 1</i>	<i>Range Cluster 2</i>
Jumlah Penduduk Miskin	18,684 - 70,106	75,481 - 127,430	40,487
Rata-rata Lamanya Sekolah	6.2 - 8.4	6.3 - 7.3	11.5
Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	672,865 - 912,867	620,602 - 862,789	1,418,452
Persentase Penduduk yang sulit Mengakses Layanan Kesehatan	0.79 - 3.37	2.92 - 7.81	0.37

Tabel 4.65 menunjukkan perbedaan nilai *range* antara *cluster* 0, 1, dan 2. Berikut ini akan dijelaskan perbandingan nilai *range* antara *cluster* 0, 1 dan 2:

1. Perbandingan Nilai *Range* untuk Variabel Data Jumlah Penduduk Miskin

Pada Tabel 4.58, *cluster* 0 memiliki nilai *range* dari angka 18,684 sampai 70,106, pada *cluster* 1 nilai *range*-nya dari angka 75,481 sampai 127,430, dan pada *cluster* 2 nilai *range*-nya 40,487. Hal ini dapat disimpulkan bahwa jumlah penduduk miskin pada *cluster* 1 cenderung lebih banyak dibandingkan dengan *cluster* 0 dan 2.

2. Perbandingan Nilai *Range* untuk Variabel Data Rata-rata Lamanya Sekolah

Pada Tabel 4.58, *cluster* 0 memiliki nilai *range* dari angka 6.2 sampai 8.4, pada *cluster* 1 nilai *range*-nya dari angka 6.3 sampai 7.3, dan pada *cluster* 2 nilai *range*-nya 11.5. Hal ini dapat disimpulkan bahwa rata-rata lamanya sekolah pada *cluster* 1 cenderung lebih rendah dibandingkan dengan *cluster* 0 dan 2.

3. Perbandingan Nilai *Range* untuk Variabel Data Rata-rata Pengeluaran Perkapita Sebulan Makanan dan *Non* Makanan

Pada Tabel 4.58, *cluster* 0 memiliki nilai *range* dari angka 672,865 sampai 912,867, pada *cluster* 1 nilai *range*-nya dari angka 620,602 sampai 862,789, dan pada *cluster* 2 nilai *range*-nya 1,418,452. Hal ini dapat disimpulkan bahwa rata-rata pengeluaran perkapita sebulan

makanan dan *non* makanan pada *cluster* 1 cenderung lebih rendah dibandingkan dengan *cluster* 0 dan 2.

4. Perbandingan Nilai *Range* untuk Variabel Data Persentase Penduduk yang Susah Mengakses Layanan Kesehatan

Pada Tabel 4.58, *cluster* 0 memiliki nilai *range* dari angka 2.92 % sampai 7.81 %, pada *cluster* 1 nilai *range*-nya dari angka 2.92 % sampai 7.81 %, dan pada *cluster* 2 nilai *range*-nya 0.37 %. Hal ini dapat disimpulkan bahwa persentase penduduk yang susah mengakses layanan kesehatan pada *cluster* 1 cenderung lebih tinggi dibandingkan dengan *cluster* 0 dan 2.

Hasil analisis menggunakan nilai *range* dari 4 variabel data dapat disimpulkan bahwa:

1. Kabupaten/kota yang masuk ke dalam *cluster* 1 merupakan wilayah dengan pembangunan sosial-ekonomi yang rendah, karena jumlah penduduk miskin yang lebih banyak dibandingkan dengan *cluster* 0 dan 2, rata-rata lamanya sekolah, dan rata-rata pengeluaran perkapita sebulan makanan dan *non* makanan yang lebih rendah dari *cluster* 0 dan 2, serta persentase penduduk yang susah mengakses layanan kesehatan yang lebih tinggi dibandingkan *cluster* 0 dan 2.
2. Kabupaten/kota yang masuk ke dalam *cluster* 0 merupakan wilayah dengan pembangunan sosial-ekonomi yang sedang, karena jumlah penduduk miskin yang cukup rendah dari *cluster* 1 namun tidak lebih

rendah dari *cluster 2* , rata-rata lamanya sekolah, dan rata-rata pengeluaran perkapita sebulan makanan dan *non* makanan yang cukup lebih tinggi dari *cluster 0* namun tidak lebih tinggi dari *cluster 2*, serta persentase penduduk yang susah mengakses layanan kesehatan yang lebih rendah dibandingkan *cluster 0* namun tidak lebih rendah dari *cluster 2*.

3. Kabupaten/kota yang masuk ke dalam *cluster 2* merupakan wilayah dengan pembangunan sosial-ekonomi yang tinggi, karena jumlah penduduk miskin yang lebih rendah dibandingkan dengan *cluster 0* dan 1, rata-rata lamanya sekolah, dan rata-rata pengeluaran perkapita sebulan makanan dan *non* makanan yang lebih tinggi dari *cluster 0* dan 1, serta persentase penduduk yang susah mengakses layanan kesehatan yang lebih rendah dibandingkan *cluster 0* dan 1.

4.8 Perbandingan Hasil *Cluster* dan *Centroid* untuk Jumlah $K=3$

Perbandingan hasil *cluster* data pada *Rapid Miner* dapat dilihat pada Gambar 4.16, Gambar 4.17, dan Gambar 4.18.

Row No.	id	Wilayah	cluster
1	1	Sumba Barat	cluster_0
2	2	Sumba Timur	cluster_1
3	3	Kupang	cluster_1
4	4	Timor Tenga...	cluster_1
5	5	Timor Tenga...	cluster_0
6	6	Belu	cluster_0
7	7	Alor	cluster_0
8	8	Lembata	cluster_0
9	9	Flores Timur	cluster_0
10	10	Sikka	cluster_0
11	11	Ende	cluster_0

Row No.	id	Wilayah	cluster
12	12	Ngada	cluster_0
13	13	Manggarai	cluster_0
14	14	Rote Ndao	cluster_0
15	15	Manggarai B...	cluster_0
16	16	Sumba Tengah	cluster_0
17	17	Sumba Barat ...	cluster_1
18	18	Nagekeo	cluster_0
19	19	Manggarai Ti...	cluster_1
20	20	Sabu Raijua	cluster_0
21	21	Malaka	cluster_0
22	22	Kota Kupang	cluster_2

Gambar 4.16 *Example Set Rapid Miner*

Gambar 4.16 merupakan *example set* dalam *tools Rapid Miner* yang akan digunakan untuk perbandingan dengan perhitungan manual terkait hasil klasterisasi dalam *dataset*. Hasil *cluster* data perhitungan manual menggunakan *microsoft excel* dalam *dataset* dapat dilihat pada Tabel 4.66.

Tabel 4.66 Hasil *Cluster* Perhitungan Manual

No	Kabupaten/Kota	Cluster
1	Sumba Barat	0
2	Sumba Timur	1
3	Kupang	1
4	Timor Tengah Selatan	1
5	Timor Tengah Utara	0
6	Belu	0
7	Alor	0
8	Lembata	0
9	Flores Timur	0
10	Sikka	0
11	Ende	0
12	Ngada	0
13	Manggarai	0
14	Rote Ndao	0
15	Manggarai Barat	0
16	Sumba Tengah	0
17	Sumba Barat Daya	1
18	Nagekeo	0
19	Manggarai Timur	1
20	Sabu Raijua	0
21	Malaka	0
22	Kota Kupang	2

Pada Tabel 4.66. *Cluster* 0 ditandai dengan warna biru, *cluster* 1 warna hijau dan *cluster* 2 warna orange. Perbandingan *cluster* data pada Gambar 4.16 dan Tabel 4.66 terlihat sama, menunjukkan konsistensi antara pengelompokkan kabupaten/kota dalam *tools Rapid Miner* dan *microsoft excel* dengan menggunakan algoritma *k-means*. Hasil klasterisasi data menggunakan *tools Rapid Miner* dan hasil klasterisasi data menggunakan *microsoft excel* tidak

menunjukkan adanya perbedaan. Selanjutnya perbandingan hasil *centroid* pada *tools Rapid Miner* dengan perhitungan manual di *microsof excel*. Hasil *centroid* pada *tools Rapid Miner* dapat dilihat pada Tabel 4.67.

Tabel 4.67 Hasil *Centroid* pada *Tools Rapid Miner*

Atribut	<i>Centroid</i> ₀	<i>Centroid</i> ₁	<i>Centroid</i> ₂
Jumlah Penduduk Miskin	0.195	0.690	0.200
Rata-rata lamanya Sekolah	0.226	0.132	1
Rata-rata Pengeluaran PerKapiata Sebulan Makanan dan <i>Non</i> Makanan	0.170	0.123	1
Persentase Penduduk Kesulitan Akses Kesehatan	0.232	0.574	0

Tabel 4.60 merupakan hasil *centroid* menggunakan *tools Rapid Miner*.

Centroid tersebut akan dibandingkan dengan *centroid* hasil dari perhitungan dataset menggunakan *microsoft excel*. *Centroid* hasil perhitungan pada *microsoft excel* dapat dilihat pada Tabel 4.68.

Tabel 4.68 Hasil *Centroid* pada Perhitungan Manual

<i>Centroid</i>	Jumlah Penduduk Miskin	Rata-rata Lamanya Sekolah	Rata-rata Pengeluaran Perkapita Sebulan Makanan dan <i>Non</i> Makanan	Persentase Penduduk Kesulitan Akses Kesehatan
<i>Centroid</i> 0	0.195	0.219	0.170	0.232
<i>Centroid</i> 1	0.690	0.125	0.123	0.574
<i>Centroid</i> 2	0.200	1.000	1.000	0.000

Pada Tabel 4.60 dan Tabel 4.61, terlihat bahwa ada perbedaan atau selisih nilai *centroid* 0 dan *centroid* 1 pada variabel data rata-rata lamanya sekolah. Selisih *centroid* 0 dan *centroid* 1 pada *tools Rapid Miner* dan *microsoft excel* yaitu 0.007. Selisih dapat terjadi karena perbedaan implementasi algoritma *k-mans* pada perangkat lunak atau platform yang berbeda.

4.9 Perbandingan Hasil Nilai DBI untuk K=3



Gambar 4.17 Nilai DBI pada *Tools Rapid Miner* untuk K=3

Gambar 4.17 merupakan hasil nilai DBI pada *tools Rapid Miner* dengan jumlah k=3. Nilai DBI tersebut akan digunakan untuk membandingkan hasil nilai DBI pada *microsoft excel*.



Gambar 4.18 Nilai DBI pada *Microsoft Excel* untuk K=3

Gambar 4.17 merupakan hasil nilai DBI pada *microsoft excel* dengan jumlah k=3. Gambar 4.17 dan Gambar 4.18, menampilkan nilai hasil DBI yang tidak sama dengan selisihnya adalah 0.005.

4.10 Perbandingan Hasil Nilai DBI untuk K=2



Gambar 4.19 Nilai DBI K=2 pada *Tools Rapid Miner*

Gambar 4.19 merupakan hasil nilai DBI pada *tools Rapid Miner* dengan jumlah $k=2$. Nilai DBI tersebut akan digunakan untuk membandingkan hasil nilai DBI pada *microsoft excel*.

$DBI = \frac{1}{K} \sum_{i=1}^k \max_{i \neq j} (R_{i,j})$		<p>Ket : K=Cluster yang ada $R_{i,j}$ = Rasio antara cluster i dan j Max = Dicari rasio antar cluster yang terbesar</p>
DAVIES BOULDIN INDEX =		0.913

Gambar 4.20 Nilai DBI K=2 pada *Microsoft Excel*

Gambar 4.20 merupakan hasil nilai DBI pada *microsoft excel* dengan jumlah $k=2$. Gambar 4.19 dan Gambar 4.20, , menampilkan nilai hasil DBI yang tidak sama dengan selisihnya adalah 0.005.

4.11 Perbandingan Hasil Nilai DBI untuk K=4

<p>Criterion</p> <ul style="list-style-type: none"> Avg. within centroid dis... <li style="background-color: #e1f5fe;">Davies Bouldin 	<h2 style="margin: 0;">Davies Bouldin</h2> <div style="border: 2px solid red; padding: 5px; display: inline-block;"> Davies Bouldin: 0.801 </div>
---	---

Gambar 4.21 Nilai DBI K=4 pada *Rapid Miner*

Gambar 4.21 merupakan hasil nilai DBI pada *tools Rapid Miner* dengan jumlah $k=4$. Nilai DBI tersebut akan digunakan untuk membandingkan hasil nilai DBI pada *microsoft excel*.

$DBI = \frac{1}{K} \sum_{i=1}^k \max_{i \neq j} (R_{i,j})$		Ket: K=Cluster yang ada $R_{i,j}$ = Rasio antara cluster i dan j Max = Dicari rasio antar cluster yang terbesar
DAVIES BOULDIN INDEX		0.676

Gambar 4.22 Nilai DBI K=4 di *Microsoft Excel*

Gambar 4.22 merupakan hasil nilai DBI pada *microsoft excel* dengan jumlah k=4. Gambar 4.21 dan Gambar 4.22, menampilkan nilai hasil DBI yang tidak sama dengan selisihnya adalah 0.125.