

Penerapan K-Means Clustering Dalam Menentukan Banyaknya Desa/Kelurahan Menurut Keberadaan dan Jenis Industri Kecil dan Mikro (Desa)

Fira Fania*, Agus Perdana Windarto, Dedy Hartama

Sistem Informasi, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Email: ^{1,*}firafania23@gmail.com, agus.perdana@amiktunasbangsa.ac.id, dedyhartama@amiktunasbangsa.ac.id

Abstrak—Industri Pengolahan adalah suatu kegiatan ekonomi yang melakukan kegiatan mengubah suatu barang dasar secara mekanis, kimia, atau dengan tangan sehingga menjadi barang jadi/setengah jadi, dan atau barang yang kurang nilainya menjadi barang yang lebih tinggi nilainya, dan sifatnya lebih dekat kepada pemakai akhir. Penelitian ini bertujuan melakukan model pengelompokan dalam menentukan Menentukan Banyaknya Desa/Kelurahan Menurut Keberadaan dan Jenis Industri Kecil dan Mikro (Desa). Penelitian ini menjadi referensi khususnya bagian pemerintah agar potensi penyerapan tenaga kerja yang dimiliki kelompok industri ini dapat terus dikembangkan dan dioptimalkan. Kontribusi pemerintah dapat diwujudkan melalui penciptaan kondisi sosial, ekonomi dan politik yang stabil dan melalui kebijakan penentuan arah pengembangan usaha Industri Mikro dan Kecil. Data dari penelitian ini diambil dari sebuah website penyedia data statistik pemerintah yaitu BPS (Badan Pusat Statistik) www.bps.go.id. Penelitian ini menggunakan metode K-Mens dan diuji dengan software RapidMiner untuk membuat 3 cluster yaitu cluster tingkat tinggi, sedang dan rendah dan melihat apa saja isi dari cluster tersebut. Dari hasil penelitian diperoleh centroid data cluster tinggi yaitu ((2151,79), (1494,34), (1135,76) cluster sedang yaitu ((406,64), (525,06), (616,218) dan cluster rendah yaitu ((455,361), (345,523), (1074,09), (176,434), (1410,34), (243,749), (295,151), (463,266), (5868,13), (9344,07), (170,925), (8818,85), (1031,65), (433,61), (5985,505), (1630,75), (367,928), (119,082), (560,907), (172,333), (545,342), (226,174), (776,643), (1880,85), (1482,39), (115,573), (232,734), (187,04), (142,884), (455,674), (441,934). Dengan analisis ini diharapkan menjadi masukan dan informasi bagi pemerintah setiap wilayah untuk lebih memperhatikan daerah-daerah yang industri mikro/kecil yang menduduki posisi di clushter rendah(C1) agar dapat meningkatkan kualitas industri pada wilayah tersebut.

Kata Kunci: Datamining; Clustering; K-Means; RapidMiner

Abstract—Processing Industry is an economic activity that carries out activities to change a basic item mechanically, chemically, or by hand so that it becomes finished / semi-finished goods, and or goods of less value to goods of higher value, and which are closer to the end user. This study aims to model the grouping in determining the Number of Villages / Villages According to the Existence and Types of Small and Micro Industries (Villages). This research is a reference, especially for the government, so that the potential for employment in this industry group can continue to be developed and optimized. Government contributions can be realized through the creation of stable social, economic and political conditions and through the policy of determining the direction of business development of Micro and Small Industries. The data from this study were taken from a government statistical data provider website, BPS (Statistics Indonesia) www.bps.go.id. This research uses the K-Mens method and is tested with RapidMiner software to create 3 clusters, namely high, medium and low level clusters and see what the contents of the cluster are. From the research results obtained by high cluster data centroids namely ((2151.79), (1494.34), (1135.76), moderate clusters ((406.64), (525.06), (616.218), and low clusters ((455,361), (345,523), (1074.09), (176,434), (1410,34), (243,749), (295,151), (463,266), (5868,13), (9344.07), (170,925), (8818,85), (1031,65), (433,61), (5985,505), (1630,75), (367,928), (119,082), (560,907), (172,333), (545,342), (226,174), (776,643), (1880,857), (172,333), (545,342), (226,174), (776,643), (1880,853), (1880,853), (18,80,853), (1880,853), (1880,853), (1880,853), (1880,853), (1880,853), (1880,853), (1880,853)), (1482.39), (115,573), (232,734), (187.04), (142,884), (455,674), (441,934) With this analysis expected to be input and information for the government of each region to pay more attention to regions micro / small industrial areas occupying low clushter (C1) positions in order to improve industrial quality in the region.

Keywords: Data Mining; Clustering; K-Means; Rapid Miner

1. PENDAHULUAN

Industri mikro adalah perusahaan/usaha industri pengolahan yang mempunyai tenaga kerja (termasuk pengusaha) 1-4 orang. Industri kecil adalah perusahaan/usaha industri pengolahan yang mempunyai tenaga kerja (termasuk pengusaha) 5-19 orang. Industri Mikro dan Kecil yang merupakan salah satu komponen dari sektor industri pengolahan secara keseluruhan mempunyai andil yang sangat besar dalam menciptakan lapangan pekerjaan bagi masyarakat. Industri Mikro dan Kecil dengan sifat usahanya yang kebanyakan masih bersifat transisi mempunyai beberapa permasalahan utama yang sering dihadapi antara lain, masalah permodalan dan pemasaran. Permasalahan tersebut sebenarnya merupakan permasalahan klasik dari usaha Industri Mikro dan Kecil yang ditinjau dari sifat usahanya yang pada umumnya masih tergolong tradisional atau usaha keluarga dan biasanya tidak bersifat padat modal. Usaha kecil ini pada umumnya masih menghadapi berbagai kompleksitas masalah baik internal maupun eksternal, seperti penguasaan teknologi yang rendah, kekurangan modal, akses pasar yang terbatas, kelemahan dalam pengelolaan usaha dan lain sebagainya Hal tersebut mengakibatkan akses terhadap sumber pembiayaan dan pasar serta rentannya menghadapi persaingan dengan industri skala menengah dan besar yang merupakan faktor eksternal menjadi rendah. Permasalahan tersebut di atas yang menyebabkan kinerja usaha Industri Mikro dan Kecil ini jauh lebih rendah dibandingkan usaha Industri Besar dan Sedang. Oleh sebab itu, sudah selayaknya mendapat perhatian pemerintah agar potensi penyerapan tenaga kerja yang dimiliki kelompok industri ini dapat terus dikembangkan dan dioptimalkan. Kontribusi pemerintah dapat diwujudkan melalui penciptaan kondisi

sosial, ekonomi dan politik yang stabil dan melalui kebijakan penentuan arah pengembangan usaha Industri Mikro dan Kecil.

Berdasarkan penjelasan tersebut, peneliti menggunakan datamining dapat menyelesaikan masalah diatas. Dalam algoritma *clustering* konsep utama yang ditekankan adalah pencarian pusat *cluster* secara iteratif, dimana pusat *cluster* ditentukan berdasarkan jarak minimum setiap data pada pusat *cluster*. Data yang digunakan dalam penelitian ini berdasarkan dokumen- dokumen yang dihasilkan oleh Badan Pusat Statistik Nasional melalui situs <https://www.bps.go.id>. Dalam hal ini peneliti mengangkat topik mengelompokkan desa/kelurahan Dalam Menentukan Banyaknya Desa/Kelurahan Menurut Keberadaan dan Jenis Industri Kecil dan Mikro (Desa) dimana proses algoritma yang dilakukan adalah *clustering*. Hasil dari *cluster* dapat dijadikan masukan bagi pemerintah agar provinsi yang masuk kedalam *cluster* rendah mendapat perhatian lebih agar kelompok industri ini dapat terus dikembangkan dan dioptimalkan. Proses *cluster* dibagi kedalam 3 (tiga) *cluster* yakni sarana kesehatan tertinggi, sarana kesehatan sedang dan sarana kesehatan rendah[1].

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining merupakan istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining juga merupakan metode yang digunakan dalam pengolahan data berskala besar oleh karena itu data mining memiliki peranan yang sangat penting dalam beberapa bidang kehidupan diantaranya yaitu bidang industri, bidang keuangan, cuaca, ilmu dan teknologi. Dalam data mining juga terdapat metode – metode yang dapat digunakan seperti klasifikasi, *clustering*, regresi, seleksi variabel, dan market basket analisis[2].

2.2 Clustering Analisis Pengelompokan

Proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar daripada kesamaan data tersebut dengan data dalam kelompok lain. Potensi *clustering* adalah dapat digunakan untuk mengetahui struktur dalam data yang dapat dipakai lebih lanjut dalam berbagai aplikasi secara luas seperti klasifikasi, pengolahan gambar, dan pengenalan pola [3].

2.3 Metode K-Means

K-Means merupakan suatu algoritma yang digunakan dalam mengelompokkan secara pertisi yang memisahkan data ke dalam kelompok yang berbeda – berbeda. Algoritma ini mampu meminimalkan jarak antara data ke *clusternya*[4].

a. Tahap Pengumpulan Data

Dalam Pemanfaatan Algoritma *Clustering* Dalam Menentukan Banyaknya Desa/Kelurahan Menurut Keberadaan dan Jenis Industri Kecil dan Mikro (Desa), diperlukan data terkait tentang hal itu. Sumber data penelitian diperoleh dari data yang dikumpulkan berdasarkan dokumen-dokumen keterangan dihasilkan oleh Direktorat Jenderal melalui situs <https://www.bps.go.id>. Data yang digunakan dalam penelitian ini adalah Banyaknya Desa/Kelurahan Menurut Keberadaan dan Jenis Industri Kecil dan Mikro (Desa), yang terdiri dari 34 provinsi.

b. Tahap pengolahan data

Data akan diolah dengan melakukan *clustering* Banyaknya Desa/Kelurahan Menurut Keberadaan dan Jenis Industri Kecil dan Mikro (Desa) dalam 3 *cluster* yakni *cluster* tingkat tinggi, *cluster* tingkat sedang dan *cluster* tingkat rendah. sehingga pada tahapan ini sudah diperoleh perhitungan nilai yang akan diproses pada tahap selanjutnya. Proses pengolahan menggunakan bantuan aplikasi *RapidMiner*.

c. Tahap Clustering

Clustering adalah suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki ke- samaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan kedua. De- mikian seterusnya sehingga *cluster* akan memben- tuk semacam pohon dimana ada hierarki (tingkat- an) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip[5].

3. HASIL DAN PEMBAHASAN

Dalam melakukan *clustering*, data yang diperoleh akan dihitung nilai rata-rata terlebih dahulu agar lebih mudah dalam pengolahannya seperti yang dapat dilihat pada tabel berikut:

Tabel 1. Data Penelitian

No	Provinsi	Banyaknya Desa/Kelurahan Menurut Keberadaan Dan Jenis Industri Kecil Dan Mikro (Desa)					
		Industri Dari Kulit	Industri Dari Kayu	Industri Dari Logam	Industri Dari Anyaman	Industri Dari Gerabah	Industri Dari Makanan

1	Aceh	21	1119,5	288	652,5	341,5	390,5	1536
2	Sumatera Utara	73	1072	326	694,5	353,5	739,5	1623,5
3	Sumatera Barat	108	824	221	282	375	480,5	792,5
4	Riau	17	803	158	297	277,5	255	579,5
5	Jambi	6,5	740,5	166	196,5	178,5	158,5	483,5
6	Sumatera Selatan	18,5	1196,5	227,5	498	425	381,5	1001
7	Bengkulu	7	519	102	227,5	169,5	142	558,5
8	Lampung	23,5	1384,5	221,5	584	767,5	456	1104,5
9	Kep. Bangka Belitung	1	263	80,5	145,5	171	58,5	275,5
10	Kep. Riau	14,5	218	49	113,5	97	73	280
11	DKI Jakarta	95	141	132,5	29,5	10,5	160	174
12	Jawa Barat	562,5	3976	584,5	2060,5	1690,5	2137,5	3830
13	Jawa Tengah	482,5	5921,5	925,5	2711	2437,5	3483,5	6225
14	DI Yogyakarta	146	409	144,5	281	186	235,5	398,5
15	Jawa Timur	806,5	5991,5	1353	2817	2698,5	2175,5	5648
16	Banten	211	931	155	558,5	444,5	526	1081,5
17	Bali	58,5	498,5	314	394,5	333	388	574,5
18	Nusa Tenggara Barat	14,5	765	126	407,5	544,5	337	600
19	Nusa Tenggara Timur	7,5	1374,5	114,5	507	426,5	1299	934
20	Kalimantan Barat	1,5	696,5	140	287	258,5	169,5	628
21	Kalimantan Tengah	7,5	382	77	315	169,5	64,5	441
22	Kalimantan Selatan	11	498	173	406,5	230,5	212	935,5
23	Kalimantan Timur	11	283,5	36	175,5	150,5	99,5	366
24	Kalimantan Utara	1	90,5	32,5	82,5	49,5	15,5	57,5
25	Sulawesi Utara	6	470,5	55,5	146	193,5	110	611,5
26	Sulawesi Tengah	9,5	1032	147,5	466	307,5	146,5	805
27	Sulawesi Selatan	18	1685	363	592	462,5	578,5	1625
28	Sulawesi Tenggara	8,5	1271	164,5	634	506	359,5	1491
29	Gorontalo	1	422	64,5	214,5	261	130,5	494
30	Sulawesi Barat	2,5	312	32,5	134	138	153	255,5
31	Maluku	6	325	23,5	184	94	87	330
32	Maluku Utara	5	426,5	27	253	136	38,5	465
33	Papua Barat	5	128,5	13,5	125,5	83	25	130,5
34	Papua	18	254	28	566	135,5	88,5	191

Setelah data diperoleh selanjutnya adalah menentukan *centroid* data awal dari data yang telah diperoleh. Adapun *centroid* data awal yang diperoleh dapat dilihat pada tabel berikut.

Tabel 2. *centroid* data

	A	B	C	D	E	F	G
C1	806,5	5991,5	1353	2817	2698,5	3483,5	6225
C2	81,92647	1071,324	207,8529	530,5588	444,2059	475,1471	1074,309
C3	1	90,5	13,5	29,5	10,5	15,5	57,5

Keterangan:

- A : Industri Dari Kulit
- B : Industri Dari Kayu
- C : Industri Dari Logam
- D : Industri Dari Anyaman
- E : Industri Dari Gerabah
- F : Industri Dari Kain
- G : Industri Dari Makanan

Selanjutnya adalah melakukan proses pencarian *cluster* tingkat tinggi (C1) tingkat sedang (C2) tingkat rendah (C3) dapat dilihat pada tabel berikut:

Tabel 3. *Cluster* iterasi 1

No	Provinsi	A	B	C	D	E	F	G	C1	C2	C3	Jarak
1	Aceh	21	1119,5	288	652,5	341,5	390,5	1536	8202,53	508,1324	1989,691	508,1324
2	Sumatera Utara	73	1072	326	694,5	353,5	739,5	1623,5	8031,289	648,565	2145,364	648,565
3	Sumatera Barat	108	824	221	282	375	480,5	792,5	8878,475	456,1112	1243,219	456,1112
4	Riau	17	803	158	297	277,5	255	579,5	9136,651	674,047	1000,741	674,047
5	Jambi	6,5	740,5	166	196,5	178,5	158,5	483,5	9319,863	865,1014	838,9425	838,9425
6	Sumatera Selatan	18,5	1196,5	227,5	498	425	381,5	1001	8505,226	188,8146	1638,533	188,8146
7	Bengkulu	7	519	102	227,5	169,5	142	558,5	9408,967	930,7168	723,1658	723,1658
8	Lampung	23,5	1384,5	221,5	584	767,5	456	1104,5	8198,668	458,6203	1972,048	458,6203
9	Kep. Bangka Belitung	1	263	80,5	145,5	171	58,5	275,5	9784,485	1307,959	350,4775	350,4775
10	Kep. Riau	14,5	218	49	113,5	97	73	280	9834,238	1358,337	291,6325	291,6325
11	DKI Jakarta	95	141	132,5	29,5	10,5	160	174	9943,84	1490,096	244,9485	244,9485
12	Jawa Barat	562,5	3976	584,5	2060,5	1690,5	2137,5	3830	3721,353	4802,206	6435,871	3721,353
13	Jawa Tengah	482,5	5921,5	925,5	2711	2437,5	3483,5	6225	609,9092	8276,898	9910,096	609,9092
14	DI Yogyakarta	146	409	144,5	281	186	235,5	398,5	9502,768	1043,978	631,1591	631,1591
15	Jawa Timur	806,5	5991,5	1353	2817	2698,5	2175,5	5648	1429,613	7756,546	9390,404	1429,613
16	Banten	211	931	155	558,5	444,5	526	1081,5	8531,25	206,3071	1596,229	206,3071
17	Bali	58,5	498,5	314	394,5	333	388	574,5	9207,955	792,6065	950,4457	792,6065
18	Nusa Tenggara Barat	14,5	765	126	407,5	544,5	337	600	9024,167	611,8302	1137,32	611,8302

No	Provinsi	A	B	C	D	E	F	G	C1	C2	C3	Jarak
19	Nusa Tenggara Timur	7,5	1374,5	114,5	507	426,5	1299	934	8170,298	897,4736	2115,558	897,4736
20	Kalimantan Barat	1,5	696,5	140	287	258,5	169,5	628	9209,773	733,4742	927,4885	733,4742
21	Kalimantan Tengah	7,5	382	77	315	169,5	64,5	441	9567,295	1090,628	587,6319	587,6319
22	Kalimantan Selatan	11	498	173	406,5	230,5	212	935,5	9098,726	696,0749	1091,614	696,0749
23	Kalimantan Timur	11	283,5	36	175,5	150,5	99,5	366	9705,552	1228,972	425,4404	425,4404
24	Kalimantan Utara	1	90,5	32,5	82,5	49,5	15,5	57,5	10088,32	1612,773	68,49088	68,49088
25	Sulawesi Utara	6	470,5	55,5	146	193,5	110	611,5	9439,645	973,7702	713,51	713,51
26	Sulawesi Tengah	9,5	1032	147,5	466	307,5	146,5	805	8856,7	462,4142	1326,314	462,4142
27	Sulawesi Selatan	18	1685	363	592	462,5	578,5	1625	7726,391	850,1813	2441,23	850,1813
28	Sulawesi Tenggara	8,5	1271	164,5	634	506	359,5	1491	8128,661	498,6642	2049,545	498,6642
29	Gorontalo	1	422	64,5	214,5	261	130,5	494	9493,67	1018,684	642,8264	642,8264
30	Sulawesi Barat	2,5	312	32,5	134	138	153	255,5	9753,393	1280,186	367,0317	367,0317
31	Maluku	6	325	23,5	184	94	87	330	9722,188	1242,639	406,6033	406,6033
32	Maluku Utara	5	426,5	27	253	136	38,5	465	9568,492	1090,124	587,6921	587,6921
33	Papua Barat	5	128,5	13,5	125,5	83	25	130,5	10000,03	1523,337	146,1215	146,1215
34	Papua	18	254	28	566	135,5	88,5	191	9739,99	1315,575	594,8445	594,8445

Setelah diperoleh nilai C1, C2, C3 serta menentukan jarak terpendeknya, selanjutnya hasil akan dikelompokkan kedalam sebuah table. Hasil iterasi I dapat dilihat pada tabel berikut:

Tabel 4. Hasil Iterasi 1

No	Provinsi	C1	C2	C3
1	Aceh		1	
2	Sumatera Utara		1	
3	Sumatera Barat		1	
4	Riau		1	
5	Jambi			1
6	Sumatera Selatan		1	
7	Bengkulu			1
8	Lampung		1	
9	Kep. Bangka Belitung			1
10	Kep. Riau			1
11	DKI Jakarta			1
12	Jawa Barat	1		
13	Jawa Tengah	1		
14	DI Yogyakarta			1
15	Jawa Timur	1		
16	Banten		1	
17	Bali		1	
18	Nusa Tenggara Barat		1	
19	Nusa Tenggara Timur		1	
20	Kalimantan Barat		1	
21	Kalimantan Tengah			1
22	Kalimantan Selatan		1	
23	Kalimantan Timur			1
24	Kalimantan Utara			1
25	Sulawesi Utara			1
26	Sulawesi Tengah		1	
27	Sulawesi Selatan		1	
28	Sulawesi Tenggara		1	
29	Gorontalo			1
30	Sulawesi Barat			1
31	Maluku			1
32	Maluku Utara			1
33	Papua Barat			1
34	Papua			1

Proses terus melakukan iterasi sampai hasil dari iterasi tersebut sama dengan hasil iterasi sebelumnya. Pada kasus ini proses iterasi berhenti pada iterasi ke 5. Pada iterasi ke -5 dilakukan proses pencarian *centroid* data yang dapat dilihat pada tabel 5 berikut:

Tabel 5. Centroid Akhir

	A	B	C	D	E	F	G
C1	617,1667	5296,333	954,3333	2529,5	2275,5	2598,833	5234,333
C2	49,85	1189	222,85	546,85	440,95	535,75	1199,4
C3	20,7381	411,7143	94,07143	237,2381	184,1429	142,9048	420,4524

Berikut adalah hasil dari proses iterasi terakhir pada perhitungan jarak pusat *cluster* yang dapat dilihat pada tabel berikut ini:

Tabel 6. Hasil Perhitungan *Cluster* terakhir

No	Provinsi	A	B	C	D	E	F	G	C1	C2	C3	Jarak
1	Aceh	21	1119,5	288	652,5	341,5	390,5	1536	6638,015	406,6404	1428,819	406,6404
2	Sumatera Utara	73	1072	326	694,5	353,5	739,5	1623,5	6487,87	525,0598	1591,703	525,0598

No	Provinsi	A	B	C	D	E	F	G	C1	C2	C3	Jarak
3	Sumatera Barat	108	824	221	282	375	480,5	792,5	7326,607	616,2183	696,0783	616,2183
4	Riau	17	803	158	297	277,5	255	579,5	7571,095	840,5463	455,3613	455,3613
5	Jambi	6,5	740,5	166	196,5	178,5	158,5	483,5	7753,447	1026,021	345,523	345,523
6	Sumatera Selatan	18,5	1196,5	227,5	498	425	381,5	1001	6934,287	258,5673	1074,09	258,5673
7	Bengkulu	7	519	102	227,5	169,5	142	558,5	7845,736	1098,546	176,4342	176,4342
8	Lampung	23,5	1384,5	221,5	584	767,5	456	1104,5	6625,573	402,8616	1410,339	402,8616
9	Kep. Bangka Belitung	1	263	80,5	145,5	171	58,5	275,5	8223,952	1481,702	243,7492	243,7492
10	Kep. Riau	14,5	218	49	113,5	97	73	280	8275,293	1529,654	295,151	295,151
11	DKI Jakarta	95	141	132,5	29,5	10,5	160	174	8394,083	1659,541	463,2664	463,2664
12	Jawa Barat	562,5	3976	584,5	2060,5	1690,5	2137,5	3830	2151,793	4636,717	5868,133	2151,793
13	Jawa Tengah	482,5	5921,5	925,5	2711	2437,5	3483,5	6225	1494,337	8105,128	9344,073	1494,337
14	DI Yogyakarta	146	409	144,5	281	186	235,5	398,5	7947,09	1221,086	170,9248	170,9248
15	Jawa Timur	806,5	5991,5	1353	2817	2698,5	2175,5	5648	1135,762	7592,279	8818,854	1135,762
16	Banten	211	931	155	558,5	444,5	526	1081,5	6974,246	333,5877	1031,654	333,5877
17	Bali	58,5	498,5	314	394,5	333	388	574,5	7659,941	965,5916	433,6109	433,6109
18	Nusa Tenggara Barat	14,5	765	126	407,5	544,5	337	600	7461,535	786,9743	595,5047	595,5047
19	Nusa Tenggara Timur	7,5	1374,5	114,5	507	426,5	1299	934	6644,511	838,2888	1630,75	838,2888
20	Kalimantan Barat	1,5	696,5	140	287	258,5	169,5	628	7643,148	901,7789	367,928	367,928
21	Kalimantan Tengah	7,5	382	77	315	169,5	64,5	441	8001,888	1264,516	119,0817	119,0817
22	Kalimantan Selatan	11	498	173	406,5	230,5	212	935,5	7541,089	848,481	560,9069	560,9069
23	Kalimantan Timur	11	283,5	36	175,5	150,5	99,5	366	8145,155	1401,244	172,3325	172,3325
24	Kalimantan Utara	1	90,5	32,5	82,5	49,5	15,5	57,5	8529,141	1785,713	545,3423	545,3423
25	Sulawesi Utara	6	470,5	55,5	146	193,5	110	611,5	7877,231	1137,969	226,1738	226,1738
26	Sulawesi Tengah	9,5	1032	147,5	466	307,5	146,5	805	7277,964	602,7995	776,6428	602,7995
27	Sulawesi Selatan	18	1685	363	592	462,5	578,5	1625	6156,323	672,4123	1880,852	672,4123
28	Sulawesi Tenggara	8,5	1271	164,5	634	506	359,5	1491	6553,536	373,8446	1482,39	373,8446
29	Gorontalo	1	422	64,5	214,5	261	130,5	494	7931,674	1191,808	115,5726	115,5726
30	Sulawesi Barat	2,5	312	32,5	134	138	153	255,5	8194,472	1451,645	232,7339	232,7339
31	Maluku	6	325	23,5	184	94	87	330	8159,426	1413,303	187,0399	187,0399
32	Maluku Utara	5	426,5	27	253	136	38,5	465	8000,733	1260,047	142,8842	142,8842
33	Papua Barat	5	128,5	13,5	125,5	83	25	130,5	8439,575	1696,931	455,6742	455,6742
34	Papua	18	254	28	566	135,5	88,5	191	8173,991	1491,233	441,9338	441,9338

Berdasarkan hasil perhitungan jarak pusat *cluster* pada iterasi ke-5 sama dengan pengelompokan pada iterasi ke-6 yang ditunjukkan pada tabel berikut:

Tabel 7. Hasil *Cluster* terakhir

Provinsi	C1	C2	C3
Aceh		1	
Sumatera Utara		1	
Sumatera Barat		1	
Riau			1
Jambi			1
Sumatera Selatan		1	
Bengkulu			1
Lampung		1	
Kep. Bangka Belitung			1
Kep. Riau			1
DKI Jakarta			1
Jawa Barat	1		
Jawa Tengah	1		
DI Yogyakarta			1
Jawa Timur	1		
Banten		1	
Bali			1
Nusa Tenggara Barat			1
Nusa Tenggara Timur		1	
Kalimantan Barat			1
Kalimantan Tengah			1
Kalimantan Selatan			1
Kalimantan Timur			1
Kalimantan Utara			1
Sulawesi Utara			1
Sulawesi Tengah		1	
Sulawesi Selatan		1	
Sulawesi Tenggara		1	
Gorontalo			1
Sulawesi Barat			1
Maluku			1
Maluku Utara			1
Papua Barat			1
Papua			1

Selanjutnya peneliti memanfaatkan *software RapidMiner* untuk menguji keakuratan pada penelitian yang telah dilakukan dan hasilnya dapat dilihat pada gambar berikut :

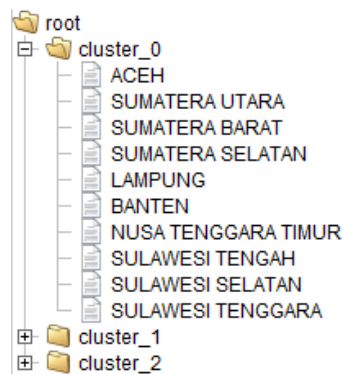
Attribute	cluster_0	cluster_1	cluster_2
industri dari	49.850	20.738	617.167
industri dari	1189	411.714	5296.333
industri dari	222.850	94.071	954.333
industri dari	546.850	237.238	2529.500
industri dari	440.950	184.143	2275.500
industri dari	535.750	142.905	2598.833
industri dari	1199.400	420.452	5234.333

Gambar 1. Hasil *centroid* iterasi -5 menggunakan *RapidMiner*

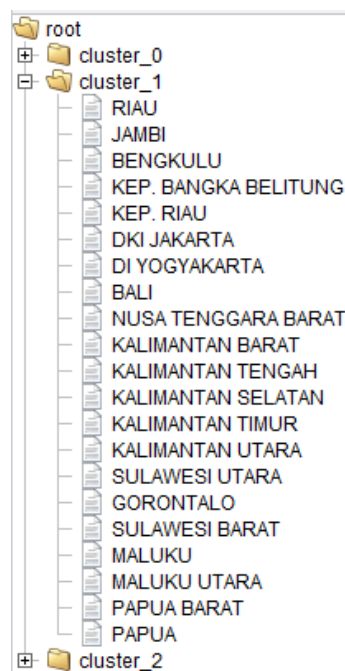
Cluster Model

```
Cluster 0: 10 items
Cluster 1: 21 items
Cluster 2: 3 items
Total number of items: 34
```

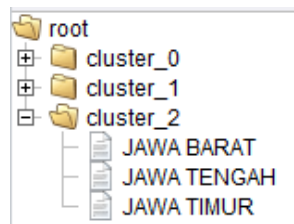
Gambar 2. *Cluster Model*



Gambar 3. Folder view *Cluster C1*



Gambar 4. Folder view *Cluster C2*



Gambar 5. Folder view *Cluster C3*

4. KESIMPULAN

Algoritma *K-Means* clustering dapat diterapkan dalam menentukan banyaknya Desa/Kelurahan Menurut Keberadaan dan Jenis Industri Kecil dan Mikro (Desa). Dengan memanfaatkan software *RapidMiner* dapat menjadi pilihan terbaik dalam menguji algoritma *K-Means* clustering. Penelitian ini diharapkan menjadi masukan kepada pemerintah kota di setiap wilayah agar memberi perhatian lebih kepada wilayah yang berada di posisi cluster tingkat rendah (C3).

REFERENCES

- [1] M. G. Sadewo, A. P. Windarto, and A. Wanto, "Penerapan Algoritma *Clustering* Dalam Mengelompokkan Banyaknya Desa/Kelurahan Menurut Upaya Antisipasi/ Mitigasi Bencana Alam Menurut Provinsi Dengan *K-Means*," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 2, no. 1, pp. 311–319, 2018, doi: 10.30865/komik.v2i1.943.
- [2] Suyanto, *Data mining Untuk Klasifikasi dan Klusterisasi Data*. 2017.
- [3] H. Sumarno, "Penerapan *K-Means* Pada Nilai Input Produksi Industri Mikro Dan Kecil Menurut Provinsi," *Publ. J. Penelit. Tek. Inform.*, vol. 3, no. 1, pp. 279–285, 2018.
- [4] A. K. Wardhani, "Implementasi Algoritma *K-Means* untuk Pengelompokan Penyakit Pasien pada Puskesmas Kaje Pekalongan," *J. Transform.*, vol. 14, no. 1, pp. 30–37, 2016.
- [5] D. N. Batubara, A. P. Windarto, D. Hartama, and H. Satria, "Analisis Metode *K-MEANS* Pada Pengelompokan Keberadaan Area Resapan Air Menurut Provinsi," no. x, pp. 345–349, 2019.