

# Penerapan K-Means Pada Nilai Input Produksi Industri Mikro Dan Kecil Menurut Provinsi

Heny Sumarno  
AMIK BSI Bekasi  
Jl. Cut Mutiah No. 88 Bekasi  
Heny.hnm@bsi.ac.id

**Abstract**— 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. Nilai Produksi Industri berkaitan dengan nilai tambah, nilai input dan nilai output. Penelitian ini mengambil data tentang nilai input produksi industri mikro dan kecil. Data yang digunakan dalam penelitian ini adalah data sekunder. Data sekunder yang digunakan adalah penggabungan dari deret waktu (*time series*) dari tahun 2013- 2015 dan deret lintang (*cross section*) sebanyak data provinsi di Indonesia yang menghasilkan 34 observasi. Sumber data dalam penelitian ini diperoleh dari publikasi Badan Pusat Statistik (BPS). Penelitian ini menggunakan datamining dengan metode K-Means. Variabel yang digunakan (1) jumlah nilai input produksi mikro dan (2) jumlah nilai input produksi mikro. Data akan diolah dengan melakukan clustering dalam 3 cluster yaitu cluster produksi nilai input tinggi, cluster produksi nilai input tinggi sedang dan produksi nilai input tinggi rendah. Centroid data untuk cluster produksi nilai input tinggi 87899515,0 dan 121258942,3, Centroid data untuk cluster produksi nilai input sedang 10634113,8 dan 15356152,8 serta Centroid data untuk cluster produksi nilai input rendah 3044143,6 dan 1507129,8. Sehingga diperoleh 3 provinsi cluster produksi nilai input tinggi, 9 provinsi cluster produksi nilai input sedang dan 22 cluster produksi nilai input rendah.

**Keywords**— Clustering, K-Means, Data Mining, Produksi, Nilai Input

## I. PENDAHULUAN

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. Peran sektor industri masih cukup penting di Indonesia. Dalam penghitungan Produk Domestik Bruto (PDB), sektor Industri masih memberikan kontribusi yang cukup besar terhadap pertumbuhan ekonomi nasional. Sektor industri tidak saja memberikan kontribusi dalam pertumbuhan ekonomi melalui peningkatan nilai tambah produksi tetapi juga berkontribusi dalam penyerapan tenaga kerja. Pada tahun 2013, sektor industri menyerap tenaga kerja sebesar 14 juta orang yang hampir mencapai 70 persen terletak di industri mikro dan kecil. Nilai produksi tidak sama dengan nilai tambah dikarenakan dalam nilai produksi terdapat nilai antara (*intermediate cost*), yaitu biaya pembelian atau biaya perolehan dari sektor lain yang

telah dihitung sebagai produksi di sektor lain atau berasal dari nilai impor (dihitung sebagai nilai produksi di Negara pengekspor).

Nilai tambah sangat berpengaruh terhadap nilai input atau nilai antara. Semakin kecil nilai input maka nilai tambah akan semakin besar. Penelitian ini bertujuan mengelompokkan nilai input produksi industri mikro dan kecil berdasarkan provinsi dimana hasil pengelompokan ini dapat menjadi masukan kepada pemerintah untuk menekan harga nilai input yang berdampak pada nilai tambah produksi industri mikro dan kecil.

Data yang digunakan dalam penelitian ini adalah data sekunder. Data sekunder yang digunakan adalah penggabungan dari deret waktu (*time series*) dari tahun 2013- 2015 dan deret lintang (*cross section*) sebanyak data provinsi di Indonesia yang menghasilkan 34 observasi. Sumber data dalam penelitian ini diperoleh dari publikasi Badan Pusat Statistik (BPS) dan sumber informasi lain berupa

jurnal ilmiah dan buku-buku teks serta literatur-literatur lain yang mendukung penelitian ini. Penelitian ini menggunakan pengolahan Datamining dengan metode K-Means.

This template, modified in MS Word 2007 and saved as a "Word 97-2003 Document" for the PC, provides authors with most of the formatting specifications needed for preparing electronic versions of their papers. All standard paper components have been specified for three reasons: (1) ease of use when formatting individual papers, (2) automatic compliance to electronic requirements that facilitate the concurrent or later production of electronic products, and (3) conformity, incorporating the applicable criteria that follow.

## II. TINJAUAN PUSTAKA

### A. Data Mining

Data mining merupakan istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database[1]. 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].

### B. Clustering

Analisis clustering merupakan 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].

Teknik *cluster* mempunyai dua metode dalam pengelompokannya yaitu *hierarchical clustering* dan *non-hierarchical clustering*. *hierarchical clustering* merupakan suatu metode pengelompokan data yang cara kerjanya dengan mengelompokkan dua data atau lebih yang mempunyai kesamaan atau kemiripan, kemudian proses dilanjutkan ke objek lain yang memiliki kedekatan dua, proses ini terus berlangsung hingga *cluster* membentuk semacam *tree* dimana ada hirarki atau tingkatan yang jelas antar objek dari yang paling mirip hingga yang paling tidak mirip. Namun

secara logika semua objek pada akhirnya hanya akan membentuk sebuah *cluster*[4].

Pada proses analisis *cluster* metode yang digunakan untuk membagi data menjadi subset data berdasarkan kesamaan atau kemiripan yang telah ditentukan sebelumnya. Jadi analisis *cluster* secara umum dapat dikatakan bahwa [5]:

- Data yang terdapat dalam satu *cluster* memiliki tingkat kesamaan yang tinggi, dan
- Dan yang terdapat dalam suatu *cluster* yang berbeda memiliki tingkat kesamaan yang rendah

### C. Algoritma K-Means

K-Means merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok [6].

Algoritma K-Means pada dasarnya melakukan 2 proses yakni proses pendeteksian lokasi pusat cluster dan proses pencarian anggota dari tiap-tiap cluster. Proses clustering dimulai dengan mengidentifikasi data yang akan dikluster,  $X_{ij}$  ( $i=1, \dots, n; j=1, \dots, m$ ) dengan  $n$  adalah jumlah data yang akan dikluster dan  $m$  adalah jumlah variabel. Pada awal iterasi, pusat setiap kluster ditetapkan secara bebas (sembarang),  $C_{kj}$  ( $k=1, \dots, k; j=1, \dots, m$ ). Kemudian dihitung jarak antara setiap data dengan setiap pusat cluster [7][8][9][10]. Proses dasar algoritma k-means dapat dilihat di bawah ini :

- Tentukan jumlah kluster yang ingin dibentuk dan tetapkan pusat cluster  $k$ .
- Menggunakan jarak *euclidean* kemudian hitung setiap data ke pusat cluster

$$d_{ik} = \sqrt{\sum_j^m (c_{ij} - c_{kj})^2} \quad (1)$$

- Kelompokkan data ke dalam cluster dengan jarak yang paling pendek dengan persamaan

$$\text{Min} \sum_k^k = d_{ik} = \sqrt{\sum_j^m (c_{ij} - c_{kj})^2} \quad (2)$$

- Hitung pusat cluster yang baru menggunakan persamaan

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \quad (3)$$

Dengan :

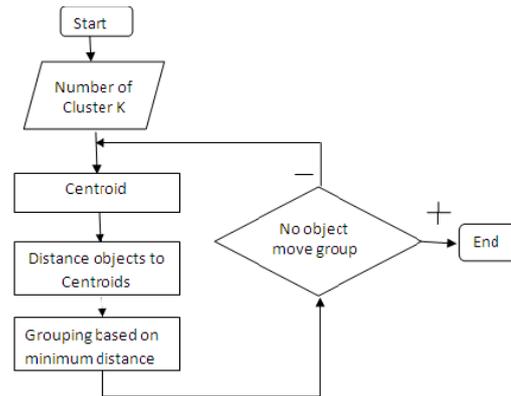
$x_{ij}$  = Kluster ke - k

$p$  = Banyaknya anggota kluster ke - k

- 5) Ulangi langkah dua sampai dengan empat sehingga sudah tidak ada lagi data yang berpindah ke kluster yang lain.

#### D. Kerangka Pemikiran

Kerangka pemikiran merupakan garis besar dari langkah - langkah penelitian yang sedang dilakukan, kerangka pemikiran dijadikan acuan untuk melakukan tahap - tahap yang sedang dilakukan dalam penelitian.



Gambar 1. Kerangka Kerja K-Means

### III. HASIL PENELITIAN

#### A. Contoh Perhitungan Algoritma K-Means Nilai Input Produksi Industri Mikro dan Kecil

Pada bagian ini akan dijelaskan mengenai penggunaan algoritma *k-means* dalam membentuk *clustering* (kelompok). Hal ini bertujuan untuk membuktikan bahwa algoritma *k-means* mampu memberikan informasi yang dibutuhkan. Sample data sebanyak 34 record dari laporan nilai input produksi industri mikro dan kecil. Berikut ini adalah data yang penulis dapatkan yaitu data transaksi 2013-2015, diantaranya:

Tabel 1. Tabel nilai input produksi industri mikro dan kecil (juta rupiah)

Provinsi	2013		2014		2015	
	Nilai Input Menurut Provinsi (Juta Rupiah)		Nilai Input Menurut Provinsi (Juta Rupiah)		Nilai Input Menurut Provinsi (Juta Rupiah)	
	Mikro	Kecil	Mikro	Kecil	Mikro	Kecil
Aceh	1660535	1422148	1708352	648313	1769763	454008
Sumatera Utara	3598386	5926268	3875966	3772030	3932706	2569288
Sumatera Barat	2269048	3406330	3767134	2800022	3358934	2906184
Riau	676468	425482	806405	438869	824055	168435
Jambi	815793	758758	1493541	614258	1292106	730387
Sumatera Selatan	1809627	3112260	4457604	1693424	1393963	716859
Bengkulu	472168	499122	599673	189921	663409	253901
Lampung	3659662	3148364	7037201	3848618	3904366	2670838
Kep. Bangka Belitung	481689	516875	396128	163394	288816	119966
Kep. Riau	399869	420044	475182	757875	192667	104768
Dki Jakarta	1841376	6853446	1583756	7381046	3187459	2969633
Jawa Barat	15490021	48991120	26513715	35030596	31862434	54126757
Jawa Tengah	19415712	56803139	27888134	23416297	43320506	45207851
Di Yogyakarta	2232368	5979784	2662509	3113901	2458585	1273039
Jawa Timur	23577557	37520755	30211811	34716279	45418655	27964033
Banten	1434678	1704669	3134765	2711970	6268138	8863988
Bali	2259106	5359528	3414329	2622635	4979387	2958525
Nusa Tenggara Barat	1456461	2026764	3896431	18307873	5340870	4829042
Nusa Tenggara Timur	941956	347619	2595071	642576	850009	446104
Kalimantan Barat	1091173	1022858	1634863	1006134	2666150	653151
Kalimantan Tengah	724696	407162	985235	279621	632744	219993

Provinsi	2013		2014		2015	
	Nilai Input Menurut Provinsi (Juta Rupiah)		Nilai Input Menurut Provinsi (Juta Rupiah)		Nilai Input Menurut Provinsi (Juta Rupiah)	
	Mikro	Kecil	Mikro	Kecil	Mikro	Kecil
Kalimantan Selatan	2788980	1327232	2243487	1690006	2473472	675924
Kalimantan Timur	1092726	1328463	1034498	728162	563531	412458
Kalimantan Utara	0	0	0	0	60880	52340
Sulawesi Utara	841959	499901	1434177	17271	2058170	20004
Sulawesi Tengah	1020860	531301	1882066	412922	988599	381129
Sulawesi Selatan	3571416	8257088	5022735	15028031	5559252	6916471
Sulawesi Tenggara	1216243	1303657	2428978	336944	2055294	223153
Gorontalo	419528	229068	566275	229129	409854	78357
Sulawesi Barat	339259	872453	743098	993855	315714	96910
Maluku	609792	314426	816402	45027	416277	24277
Maluku Utara	185746	22632	237790	31037	150222	24138
Papua Barat	149182	28516	169018	38978	89011	26942
Papua	483942	382735	1216065	472073	670355	71151

(Sumber : Badan Pusat Statistik)

Untuk membuat cluster pada table di atas, maka terlebih dahulu ditentukan adalah sebagai berikut:

- 1) Menentukan jumlah cluster, dimana jumlah cluster (kelompok) adalah 3.
- 2) Menentukan pusat cluster secara acak. Pada contoh kasus ini nilai pusat cluster 1 = (99.208.023 dan 138.148.473 ), nilai pusat cluster 2 = (12.540.492 dan 15.739.384 ), dan nilai pusat cluster 3 = (60.880 dan 52.340 ).
- 3) Menentukan nilai cluster dari tiap data. Dalam hal ini harus menentukan nilai cluster mana yang paling dekat dengan data, maka perlu dihitung jarak setiap data dengan titik pusat setiap cluster.

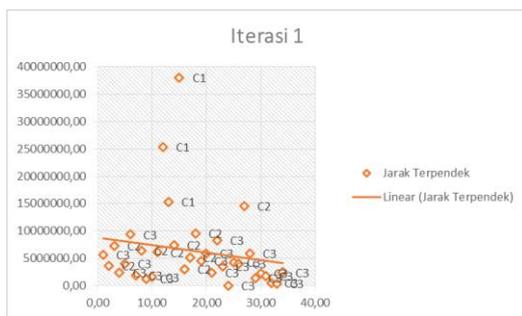
Pada tahap ini *Distance Space* digunakan untuk menghitung jarak antara data dan centroid. Adapun persamaan yang dapat digunakan salah satunya yaitu *Euclidean Distance Space*. *Euclidean distance space* digunakan dalam perhitungan jarak, hal ini dikarenakan hasil yang dapat diperoleh merupakan jarak terpendek antara dua titik yang diperhitungkan. Rumus yang digunakan adalah (1)

Dari hasil perhitungan di atas maka dihasilkan nilai cluster 1, cluster 2 dan cluster ke 3 pada iterasi 1. Hasil perhitungan dari semua perhitungan iterasi 1 dapat dilihat pada table di bawah ini:

**Tabel 2.** Tabel Perhitungan Iterasi ke-1

No	Provinsi	Cluster 1	Cluster 2	Cluster 3	Jarak Terpendek	Hasil
1	Aceh	165054286,21	15146657,88	5647580,89	5647580,89	C3
2	Sumatera Utara	153476405,90	3652130,12	16671772,25	3652130,12	C2
3	Sumatera Barat	157215238,77	7335428,10	13008270,96	7335428,10	C2
4	Riau	167900368,77	17916747,99	2450715,40	2450715,40	C3
5	Jambi	166279523,05	16304803,90	4091750,79	4091750,79	C3
6	Sumatera Selatan	161153526,83	11322163,67	9364181,42	9364181,42	C3
7	Bengkulu	168304184,93	18321787,39	1896494,23	1896494,23	C3
8	Lampung	153836236,91	6411749,15	17432131,39	6411749,15	C2
9	Kep. Bangka Belitung	168750267,07	18776124,29	1334929,45	1334929,45	C3
10	Kep. Riau	168415447,16	18455910,52	1589803,92	1589803,92	C3
11	Dki Jakarta	152320219,74	6106183,45	18360518,67	6106183,45	C2
12	Jawa Barat	25341853,00	136911737,40	156581489,27	25341853,00	C1
13	Jawa Tengah	15346269,29	134642211,97	154662923,15	15346269,29	C1
14	Di Yogyakarta	157370377,63	7467982,08	12632033,47	7467982,08	C2
15	Jawa Timur	37947406,00	121016679,90	140925240,77	37947406,00	C1
16	Banten	152974880,24	2990884,84	17062381,47	2990884,84	C2
17	Bali	154996271,53	5156625,05	15190304,72	5156625,05	C2
18	Nusa Tenggara Barat	143528178,68	9603527,83	27269718,11	9603527,83	C2
19	Nusa Tenggara Timur	166376795,54	16463811,44	4542132,57	4542132,57	C3
20	Kalimantan Barat	164780271,38	14885893,42	5944635,19	5944635,19	C3
21	Kalimantan Tengah	167982674,82	18000048,15	2436524,02	2436524,02	C3

No	Provinsi	Cluster 1	Cluster 2	Cluster 3	Jarak Terpendek	Hasil
22	Kalimantan Selatan	162749816,79	13055963,76	8287610,53	8287610,53	C3
23	Kalimantan Timur	166506696,24	16526288,43	3571678,77	3571678,77	C3
24	Kalimantan Utara	170002052,68	20045549,81	0,00	0,00	C3
25	Sulawesi Utara	167146316,86	17275665,50	4300841,28	4300841,28	C3
26	Sulawesi Tengah	166750715,83	16809787,34	4036632,34	4036632,34	C3
27	Sulawesi Selatan	137429319,77	14551868,70	33280271,63	14551868,70	C2
28	Sulawesi Tenggara	165279093,31	15469919,03	5923403,04	5923403,04	C3
29	Gorontalo	168832162,79	18850288,83	1419891,84	1419891,84	C3
30	Sulawesi Barat	167669944,80	17718247,57	2332280,97	2332280,97	C3
31	Maluku	168698474,00	18714800,69	1812149,50	1812149,50	C3
32	Maluku Utara	169682724,64	19710091,80	513509,89	513509,89	C3
33	Papua Barat	169766067,23	19798507,71	348879,97	348879,97	C3
34	Papua	167951037,32	17968558,82	2469193,65	2469193,65	C3



Gambar 2. Pemetaan Iterasi 1 Pada K-Means

Proses *K-Means* akan terus beriterasi sampai pengelompokan data sama dengan pengelompokan data *iterasi* sebelumnya. Dengan kata lain, proses akan terus melakukan *iterasi* sampai data pada *iterasi* terakhir sama dengan *iterasi* sebelumnya. Pada *iterasi* 1 diperoleh cluster data nilai input proses industri mikro dan kecil, yang dapat dilihat pada tabel 2. Pada *iterasi* 2, akan dilakukan pencarian nilai titik tengah

atau *centroid* untuk iterasi 2 dengan hasil pengelompokan pada *iterasi* 1. Hasil dapat diketahui pada Tabel 3 berikut :

Tabel 3. Centroid Data iterasi 2

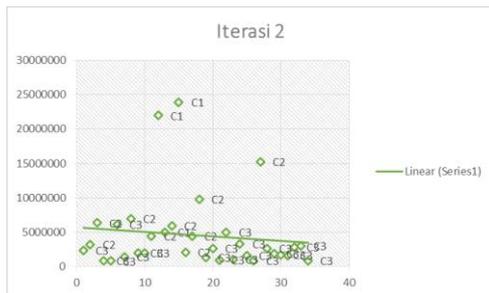
Atribut	Cluster 1	
Nilai	Mikro	Kecil
	87899515,0	121258942,3
Atribut	Cluster 2	
Nilai	Mikro	Kecil
	10634113,8	15356152,8
Atribut	Cluster 3	
Nilai	Mikro	Kecil
	3044143,6	1507129,8

Setelah mendapatkan nilai titik tengah atau *centroid*, proses sama dilakukan dengan mencari jarak terdekat. Proses pencarian jarak terdekat, pengelompokan data pada *iterasi* 2 dan *Clustering* data dapat digambarkan pada tabel dan gambar berikut:

Tabel 4. Tabel Perhitungan Iterasi ke 2

No	Provinsi	Cluster 1	Cluster 2	Cluster 3	Jarak Terpendek	Hasil
1	Aceh	144731599,6	13958948,05	2328505,082	2328505,1	C3
2	Sumatera Utara	133154841,2	3183816,501	13628123,64	3183816,5	C2
3	Sumatera Barat	136893232,5	6365364,558	9908433,436	6365364,6	C2
4	Riau	147581908,2	16568067,44	876635,0166	876635,0	C3
5	Jambi	145959610,9	15003128,91	816162,3272	816162,3	C3
6	Sumatera Selatan	140830047,5	10273175,42	6118880,378	6118880,4	C3
7	Bengkulu	147987229,2	16939018,97	1425309,857	1425309,9	C3
8	Lampung	133511113,1	6935065,463	14147900,42	6935065,5	C2
9	Kep. Bangka Belitung	148434810,6	17363983,84	2006177,079	2006177,1	C3
10	Kep. Riau	148101528,7	17017002,3	1989128,672	1989128,7	C3
11	Dki Jakarta	132041542,8	4425793,351	16097499,01	4425793,4	C2
12	Jawa Barat	21958848,29	138116787	153904568,1	21958848,3	C1
13	Jawa Tengah	4979943,166	136066501,4	151744845,9	4979943,2	C1
14	Di Yogyakarta	137057472,4	5971354,588	9852037,039	5971354,6	C2
15	Jawa Timur	23902227,23	122653972,1	137796897,4	23902227,2	C1
16	Banten	132656919,4	2085475,045	14119238,71	2085475,0	C2
17	Bali	134674306,5	4415504,411	12119571,21	4415504,4	C2
18	Nusa Tenggara Barat	123268114	9807707,607	24862602,04	9807707,6	C2
19	Nusa Tenggara Timur	146054099,6	15257401,81	1344759,051	1344759,1	C3

No	Provinsi	Cluster 1	Cluster 2	Cluster 3	Jarak Terpendek	Hasil
20	Kalimantan Barat	144457317,8	13715259,05	2625634,955	2625635,0	C3
21	Kalimantan Tengah	147663864,3	16659305,1	923300,0351	923300,0	C3
22	Kalimantan Selatan	142424856,6	12075215,58	4968536,461	4968536,5	C3
23	Kalimantan Timur	146190162	15138478	1024810,935	1024810,9	C3
24	Kalimantan Utara	149688564,2	18601074,11	3319077,483	3319077,5	C3
25	Sulawesi Utara	146822644,8	16102473,44	1614103,26	1614103,3	C3
26	Sulawesi Tengah	146428851,2	15566819,65	866659,3048	866659,3	C3
27	Sulawesi Selatan	117174785,9	15256880,51	30769915,32	15256880,5	C2
28	Sulawesi Tenggara	144954774,3	14366113,66	2680203,281	2680203,3	C3
29	Gorontalo	148515361,1	17463378,58	1912988,659	1912988,7	C3
30	Sulawesi Barat	147356607,1	16268841,02	1708090,031	1708090,0	C3
31	Maluku	148380024,9	17362788,56	1645005,859	1645005,9	C3
32	Maluku Utara	149367517,9	18293130,08	2854079,346	2854079,3	C3
33	Papua Barat	149451466,6	18371432,7	2991507,572	2991507,6	C3
34	Papua	147632187,9	16628893,11	889798,4117	889798,4	C3



Gambar 3. Pemetaan Iterasi 2 Pada K-Means

Class centroids:

Class	Mikro	Kecil	Sum of weights	Within-class variance
1	3044143.636	1507129.818	22.000	6587023952364.590
2	10634113.778	15356152.778	9.000	63350478209526.600
3	87899515.000	121258942.333	3.000	539153659953134.000

Gambar 4. Nilai Centroid Akhir Pada Aplikasi XLSTAT

Results by class:

Class	1 (Produksi Rendah)	2 (Produksi Sedang)	3 (Produksi Tinggi)
Objects	22	9	3
Sum of wei	22	9	3
Within-clas	6587023952364.590	63350478209526.600	539153659953134.000
Minimum d	816162.327	2085475.045	4979943.166
Average di	2133152.412	6494106.393	16947006.228
Maximum c	6118880.378	15256880.512	23902227.234
	ACEH RIAU JAMBI SUMATERA SELATAN BENGKULU KEP. BANGKA BELITUNG KEP. RIAU NUSA TENGGARA TIMUR KALIMANTAN BARAT KALIMANTAN TENGAH KALIMANTAN SELATAN KALIMANTAN TIMUR KALIMANTAN UTARA SULAWESI UTARA SULAWESI TENGAH SULAWESI TENGGARA GORONTALO SULAWESI BARAT MALUKU MALUKU UTARA PAPUA BARAT PAPIA	SUMATERA UTARA SUMATERA BARAT LAMPUNG DKI JAKARTA DI YOGYAKARTA BANTEN BALI NUSA TENGGARA BARAT SULAWESI SELATAN	JAWA BARAT JAWA TENGAH JAWA TIMUR

Gambar 5. Hasil Akhir Pengklasteran Pada Aplikasi XLSTAT

### B. Analisa Data

Pada iterasi 2, pengelompokan data yang dilakukan terhadap 3 cluster dengan iterasi 1 didapatkan hasil yang sama. Dari 34 data nilai input produksi industri mikro dan kecil dapat diketahui, 3 provinsi pada cluster 1, 9 provinsi pada cluster 2 dan 22 provinsi pada cluster 3. Pada cluster 3 (cluster rendah) pada nilai input produksi industri mikro dan kecil adalah Aceh, Riau, Jambi, Sumatera Selatan, Bengkulu, Kep. Bangka Belitung, Kep. Riau, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat Dan Papua.

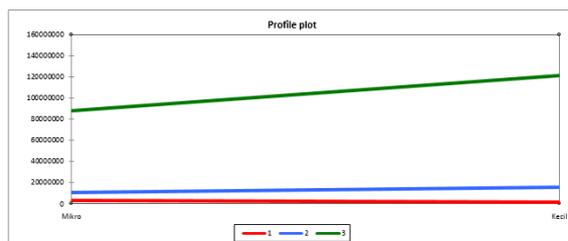
### C. Pengujian dan Hasil

Dari analisa dan perancangan yang telah dilakukan, hasil dari implementasi program yang telah dibangun dengan aplikasi XLSTAT diperoleh hasil cluster yang sama dengan perhitungan manual yang dilakukan. Adapun hasil implementasi dari software XLSTAT dapat terlihat seperti pada tampilan dibawah ini:

Results by object:

Observation	Class	Distance to centroid
ACEH	1	2328505,082
SUMATERA UTARA	2	3183816,501
SUMATERA BARAT	2	6365364,558
RIAU	1	876635,017
JAMBI	1	816162,327
SUMATERA SELATAN	1	6118880,378
BENGKULU	1	1425309,857
LAMPUNG	2	6935065,463
KEP. BANGKA BELITUNG	1	2006177,079
KEP. RIAU	1	1989128,672
DKI JAKARTA	2	4425793,351
JAWA BARAT	3	21958848,286
JAWA TENGAH	3	4979943,166
DI YOGYAKARTA	2	5971354,588
JAWA TIMUR	3	23902227,234
BANTEN	2	2085475,045
BALI	2	4415504,411
NUSA TENGGARA BARAT	2	9807707,607
NUSA TENGGARA TIMUR	1	1344759,051
KALIMANTAN BARAT	1	2625634,955
KALIMANTAN TENGAH	1	923300,035
KALIMANTAN SELATAN	1	4968536,461
KALIMANTAN TIMUR	1	1024810,935
KALIMANTAN UTARA	1	3319077,483
SULAWESI UTARA	1	1614103,260
SULAWESI TENGAH	1	866659,305
SULAWESI SELATAN	2	15256880,512
SULAWESI TENGGARA	1	2680203,281
GORONTALO	1	1912988,659
SULAWESI BARAT	1	1708090,031
MALUKU	1	1645005,859
MALUKU UTARA	1	2854079,346
PAPUA BARAT	1	2991507,572

Gambar 6. Hasil Akhir Detail Pengklasteran Pada Aplikasi XLSTAT



Gambar 7. Profile Plot Klaster dengan XLSTAT

#### IV. KESIMPULAN

Untuk melakukan penilaian terhadap hasil nilai input produksi industri mikro dan kecil dapat menerapkan metode clustering K-Means. Data diolah untuk memperoleh nilai input produksi industri mikro dan kecil. Data tersebut diolah menggunakan Aplikasi XLSTAT. Nilai centroid dalam 3 cluster yaitu cluster tingkat tinggi (C1), cluster tingkat sedang (C2) dan cluster rendah (C3). Centroid data untuk cluster produksi nilai input tinggi 87899515,0 dan 121258942,3, Centroid data untuk cluster produksi nilai input sedang 10634113,8 dan 15356152,8 serta Centroid data untuk cluster produksi nilai input rendah 3044143,6 dan 1507129,8. Dari 34 data penelitian diperoleh hasil C1 (3 Provinsi), C2 (9 Provinsi) dan C3 (22 Provinsi). Hasil yang dari penelitian dapat

digunakan sebagai masukan bagi pemerintah tentang penilaian nilai input produksi industri mikro dan kecil dari seluruh provinsi. Dimana hal ini bisa dijadikan bahan pertimbangan bagi pemerintah dalam mengambil kebijakan karena sangat berpengaruh terhadap nilai tambah suatu produksi.

#### REFERENCES

- [1] Afrisawati. 2013. "Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means". Jurnal Pelita Informatika Budi Karma, Vol 5, No.3, ISSN: 2301-9425
- [2] N. Athina dan L. Iswari, "Klasterisasi Data Kesehatan Penduduk untuk Menentukan Rentang Derajat kesehatan Daerah dengan Metode K-Means," Seminar Nasional Aplikasi Teknologi Informasi (SNATI), Vol. %1 dari %2ISSN 1907 - 5022, pp. B52 - B59, 2014.
- [3] Ong, Johan Oscar. 2003. "Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University". Jurnal Ilmiah Teknik Industri, Vol 12, No. 1, ISSN: 1412-6869
- [4] Anindya Khrisna Wardhani, "Implementasi Algoritma K-Means Untuk Pengelompokan Penyakit Pasien Pada Puskesmas Kajen Pekalongan", Jurnal Transformatika, Volume 14, Nomor 1, Juli 2016, Pp 30-37
- [5] S. Agustina, D. Yhudo, H. Santoso, N. Marnasusanto, A. Tirtana dan F. Khusnu, "Clustering Kualitas Beras Berdasarkan Ciri Fisik Menggunakan Metode K-Means," Universitas Brawijaya Malang, Malang, 2012.
- [6] Suprihatin, "Klastering K-Means Untuk Penentuan Nilai Ujian," JUSI, vol. 1, no. 1, pp. ISSN: 2087-8737, Februari 2011
- [7] Soni, N., & Ganatra, A., 2012, *Categorization of several Clustering algorithms from different perspective: a review*, *International Journal of Advanced Research in Computer Science and Software Engineering*, 2(8), 63-68.
- [8] Windarto AP. Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method. *Int J Artif Intell Res.* 2017;1(2):26-33. <http://ijair.id/index.php/ijair/article/view/17>.
- [9] Windarto AP. Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering. *TechnoCOM.* 2017;16(4):348-357.
- [10] Sadewo MG, Windarto AP, Hartama D. Penerapan Datamining Pada Populasi Daging Ayam Ras Pedaging Di Indonesia Berdasarkan Provinsi Menggunakan K-Means Clustering. *InfoTekJar (Jurnal Nas Inform dan Teknol Jaringan).* 2017;2(1):60-67. <http://jurnal.uisu.ac.id/index.php/infotekjar/article/view/164>.