

PENERAPAN DATA MINING PADA JUMLAH PELANGGAN PERUSAHAAN AIR BERSIH MENURUT PROVINSI MENGGUNAKAN METODE K-MEANS CLUSTERING

Lestari Sinaga¹, Abdullah Ahmad², M.Safii³

^{1,2,3}STIKOM Tunas Bangsa Pematangsiantar
Jl. W. R. Supratman, Proklamasi, Kec. Siantar Bar., Kota Pematang Siantar, Sumatera Utara 21146 -
Indonesia

e-mail: lestarisinaga445@gmail.com¹, abdul@amiktunasbangsa.ac.id², m.safii@amiktunasbangsa.ac.id³

Received : August, 2019

Accepted : October, 2019

Published : October, 2019

Abstract

Water is one of the primary needs for humans so that everyone has the right to get clean water for their daily needs. Along with increasing population, the need for water will increase. So with that the PDAM must sell clean / decent water to its customers, clean water becomes the focus of attention and has the greatest power compared to other problems. Because water is a basic necessity, most of the companies impose rates that can be reached by the community and prices are adjusted to the growth in demand. The purpose of this research is to get a grouping of the number of customers of clean water companies in all provinces using the K-Means Algorithm, K-Means is a method for grouping data into a cluster by calculating the closest distance from a data to a centroid point. Clusters used are high level clusters (C1), medium level clusters (C2), and for low level clusters (C3). Centroid data obtained is for high-level clusters (C1) which are as many as 7710154, for medium-level clusters as much as 929586, and for low-level clusters as much as 112462. Based on the calculated data obtained high-level results, namely the province of Indonesia, for the medium level namely province North Sumatra, DKI Jakarta, West Java, Central Java and East Java, and other provinces are low levels. So that this result can be a support for the company in order to increase water needs.

Keywords: *maining data, k-means, customers, clean water*

Abstrak

Air merupakan kebutuhan primer bagi manusia sehingga setiap orang memiliki hak mendapatkan air bersih untuk kebutuhan setiap harinya. Seiring dengan bertambahnya penduduk maka kebutuhan air akan semakin meningkat. Maka dengan itu PDAM harus menjual air bersih kepada pelanggannya, air bersih menjadi salah satu focus perhatian pemerintah. Karena air merupakan kebutuhan pokok maka kebanyakan dari perusahaan tersebut mengenakan tarif yang terjangkau oleh masyarakat dan harga disesuaikan dengan pertumbuhan permintaan. Tujuan penelitian ini adalah untuk mendapatkan pengelompokan jumlah pelanggan PDAM diseluruh provinsi menggunakan Algoritma K-Means, K-Means merupakan metode untuk untuk mengelompokkan data ke dalam suatu cluster dengan melakukan cara menghitung jarak terdekat dari suatu data ke sebuah titik centroid. Cluster yang digunakan yaitu cluster tingkat tinggi (C1), cluster tingkat sedang (C2), dan untuk cluster tingkat rendah (C3). Centroid data yang didapat yaitu untuk cluster tingkat tinggi (C1) yaitu sebanyak 7710154, untuk cluster tingkat sedang sebanyak 929586, dan untuk cluster tingkat rendah sebanyak 112462. Berdasarkan data yang telah dihitung diperoleh hasil untuk tingkat tinggi yaitu provinsi Indonesia, untuk

1

tingkat sedang yaitu provinsi Sumatera Utara, DKI Jakarta, Jawa Barat, Jawa Tengah dan Jawa Timur, dan provinsi lainnya merupakan tingkat rendah. Sehingga hasil ini dapat menjadi pendukung bagi perusahaan agar dapat meningkatkan kebutuhan air.

Kata Kunci: data mining, k-means, pelanggan, air bersih

1. PENDAHULUAN

Air senantiasa harus ada dalam kehidupan sehari-hari. Baik dalam hal konsumsi, MCK (Mandi, Cuci, Kakus), atau bahkan sebagai salah satu sarana membersihkan rumah. Bahkan keberadaan serta kondisi air cukup berpengaruh dalam kesehatan masyarakat sendiri. Kondisi air yang layak untuk dipergunakan terlebih untuk dikonsumsi masyarakat tentunya air yang sesuai dengan parameter kualitas air bersih seperti yang tertera dalam PERMENKES 416 / 1990 dan dipertegas oleh PERMENKES No .32 tahun 2017 Tentang Standar Baku Mutu Kesehatan Kesling. Melihat pentingnya air bersih dalam kehidupan masyarakat, pemerintah tentunya ikut ambil bagian dalam pendistribusian air bersih kepada masyarakat guna kesejahteraan rakyat. Melalui Perusahaan Daerah Air Minum (PDAM).

Pemerintah selalu mengupayakan pendistribusian air bersih secara merata di seluruh Indonesia. PDAM di berbagai daerah adalah perusahaan pengoperasiannya di bawah naungan tanggung jawab Pemerintah daerah masing-masing yang diharapkan dapat memberikan pelayanan kepada masyarakat, walaupun pada kenyataannya belum bisa secara maksimal memberikan pelayanan sesuai dengan yang diharapkan oleh masyarakat.

Perusahaan Daerah Air Minum (PDAM) sendiri selain berfungsi sebagai pemberi layanan kepada masyarakat, PDAM juga merupakan perusahaan yang berorientasi pada profit atau keuntungan. Melihat hal tersebut tak heran jika PDAM dituntut harus bisa menjalankan perannya dalam pelayanan air bersih kepada masyarakat dan juga menghasilkan laba yang optimal. Perusahaan Daerah Air Minum (PDAM) sendiri diberi kewajiban untuk menyetorkan beberapa persen pendapatannya ke daerah dalam bentuk pajak pendapatan. Hal tersebut diatur pada UU No. 36 Tahun 2008 Tentang Pajak Penghasilan. Sebagai perusahaan yang terbilang sebagai perusahaan yang

memonopoli penjualan air bersih, PDAM tentunya dapat menjadi salah satu perusahaan penyumbang pendapatan Negara.

Pada penelitian kali ini penulis ingin menentukan pengelompokan jumlah pelanggan perusahaan air bersih dalam konteks ini yang dimaksud adalah PDAM, yang diamati berdasarkan wilayah (provinsi) menggunakan metode K-means clustering. Adapun tujuan penelitian ini yaitu untuk mengetahui wilayah mana dengan jumlah pelanggan terbanyak dan penyumbang pendapatan yang lebih besar bagi daerahnya dan Negara.

2. METODE PENELITIAN

2.1 Data Mining

Konsep data mining semakin banyak dikenal dalam berbagai macam bidang ilmu, misalnya: ilmu pendidikan, ilmu pemerintahan dan ilmu kesehatan. Pada dasarnya, data mining ada sebab dari banyaknya data yang tersimpan pada sebuah *database*. Sehingga dari data secara keseluruhan yang tersimpan di dalam *database* bisa dikelola dan menghasilkan pengetahuan dan informasi yang baru[1]. Istilah data mining memiliki beberapa pandangan, seperti *knowledge discover* ataupun *pattern recognition*. Kedua istilah tersebut sebenarnya memiliki ketepatannya masing-masing, istilah *knowledge discovery* atau penemuan ilmu pengetahuan tepat karena digunakannya manfaat utama dari data mining memang untuk mendapat pengetahuan yang masih tersembunyi di dalam bongkahan data yang akan di olah. Istilah *pattern recognition* akan tetap digunakan sebab ilmu pengetahuan yang akan di pahami lebih mendalam pada hakikatnya memang berupa node-node yang masih perlu dipahami secara mendalam dari berupa data mentah yang tengah dihadapi. Banyak definisi bagi istilah ini dan belum ada yang dibakukan atau disepakati semua pihak. Namun demikian, istilah ini memiliki hakikat sebagai sebuah bidang keilmuan yang bermanfaat bermanfaat untuk memperoleh, mencari, atau menggali pengetahuan dari data

atau informasi yang kita punya, kegiatan tersebut yang menjadi titik perhatian dari disiplin ilmu *data mining*[2]. Penerapan metode-metode data mining untuk menganalisis kesalahan data dapat menjadi solusi permasalahan yang dihadapi. Data mining digunakan pada saat mengidentifikasi dan menganalisa pola sensor yang mengalami kesalahan pada peroperasian berdasarkan pada data yang dihasilkan sehingga data-data yang tidak normal diperoleh pada jaringan sensor tanpa kabel dapat direduksi[3].

Data mining juga bisa diartikan sebagai rangkaian kegiatan untuk menentukan pola yang menarik dari data dalam jumlah besar, lalu data-data tersebut dapat disimpan dalam database, data *warehouse* atau penyimpanan informasi. Memiliki beberapa bidang keilmuan yang mendukung teknik data mining diantaranya ialah data analisis, *signal processing*, *neural network* dan pengenalan node[4].

2.2 Clustering

Clustering termasuk ke dalam *descriptive methods*, dan juga termasuk *unsupervised learning* dimana tidak ada penjelasan kelas dari objek sebelumnya. Sehingga *clustering* bisa dimanfaatkan pada penentuan label kelas bagi masing-masing data yang belum diketahui kelasnya. Konsep dasar *clustering* ialah mengelompokkan beberapa objek ke dalam *cluster* dimana *cluster* yang baik adalah *cluster* yang memiliki tingkat kemiripan yang tinggi antar objek pada suatu *cluster* dan tingkat ketidaksamaan yang tinggi dengan objek *cluster* yang lainnya[5].

Clustering mengacu pada mengelompokkan data, observasi, atau kasus ke dalam kelas yang sama. Sebuah *cluster* adalah kumpulan dari catatan yang mirip satu sama lain dan berbeda dengan data dalam *cluster* lain. *Clustering* tidak sama dengan klasifikasi yang tidak memiliki variabel target untuk *clustering*. Sebaliknya, pada algoritma klasterisasi mencari ke segmen data seluruh set menjadi sub kelompok yang relatif homogen atau kelompok, di mana kesamaan catatan dalam *cluster* dimaksimalkan, dan kesamaan catatan luar klaster ini diminimalkan. Pada analisis *cluster* ada beberapa penjelasan yang harus diperhatikan, sebagai contoh, kita akan perlu menentukan:

1. Bagaimana mengukur kesamaan
2. Bagaimana *recode* variabel kategori
3. Bagaimana standarisasi atau menormalkan variabel numerik
4. Berapa banyak *cluster* yang akan dibentuk[6].

2.3 K-means

K-Means pada dasarnya adalah sebuah metode pengelompokan data *non-hierarki* (sekatan) yang berusaha memisahkan data yang ada ke dalam bentuk dua atau lebih kelompok. Dalam metode ini, pada proses pemisahan data ke dalam kelompok-kelompok sehingga data yang memiliki karakter sama dikelompokkan ke dalam satu kelompok yang sama dan data yang memiliki karakter berbeda dikelompokkan kedalam kelompok yang lain. Adapun manfaat pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diatur pada proses pengelompokan, yang biasanya berupaya meminimalisir variasi data di dalam suatu kelompok dan mengoptimalkan variasi data yang berbeda kelompok[7].

Algoritma *K-means* pada sebelumnya melakukan 2 proses yaitu proses penemuan 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* [Soni and Ganatra, 2012]. Proses penjumlahan landasan algoritma *k-means* dapat dilihat dengan ketentuan di bawah ini :

1. Menentukan jumlah *cluster* yang akan dibuat pada bentuk kelompok dan ditetapkan pusat *cluster* k .
2. Memakai jarak *euclidean* selanjutnya dijumlah setiap data ke pusat *cluster*

$$d(i, k) = \sqrt{\sum_i^m (C_{ij} - C_{kj})^2} \quad (1)$$

3. Mengelompokkan data pada *cluster* dengan jarak terpendek dengan persamaan

$$\min \sum_k^i -a_{ik} = \sqrt{\sum_i^m (C_{ij} - C_{kj})^2} \quad (2)$$

4. Menjumlahkan pusat *cluster* yang baru menggunakan persamaan

$$C_{kj} = \frac{\sum_k^i x_{ij}}{P} \quad (3)$$

Dengan : X_{ij} ? Kluster ke k p = banyaknya anggota kluster ke – k.

Ulangi langkah kedua sampai dengan seterusnya sehingga sudah tidak ada lagi data yang berpindah ke kluster yang lain[8].

3. HASIL DAN PEMBAHASAN

Metode Data *Maining* pada penelitian ini menggunakan Algoritma *K-means*, dimana terlebih dahulu yang dihitung adalah nilai rata-rata dari jumlah pelanggan air bersih menurut provinsi (1995-2017).

Berikut merupakan data jumlah pelanggan air bersih berdasarkan provinsi di Indonesia periode 1995-2017.

Provinsi	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2017	
Aceh	40031	43715	43804	53740	55548	62207	63471	67577	62122	53209	57786	42543	58750	57112	80588	111995	131076	137765	146332	148158	163812	163015	
Sumatera Utara	312047	348571	376786	400005	409685	454917	499671	539086	534166	530621	538100	624618	575540	617721	621169	642851	688320	780765	812716	891429	839212	864318	
Sumatera Barat	105187	113581	120567	128958	133901	138670	136374	130987	152187	150616	154514	188343	117440	183936	198751	206787	220873	229561	237718	256152	274384	329510	
R i a u	46499	53672	81210	79142	80328	93686	87659	92644	206199	47880	50718	195654	72891	60978	59606	58865	64610	64108	66002	73085	80824	98543	
J a m b i	40839	61289	64334	73527	77613	64548	69970	93181	90395	129224	100855	146561	150965	126857	117146	112148	115550	132799	113711	122102	607029	658151	
Sumatera Selatan	114807	125094	125883	134430	144432	126108	141454	55490	144143	153423	158800	163699	242593	247436	181644	210607	299487	338881	379681	447746	693138	756369	
Bengkulu	23992	25296	29093	33408	35862	36523	36121	36848	38569	59419	57723	37987	43784	44728	51429	52968	55184	59339	56610	63893	74956	83334	
Lampung	39501	57765	48200	57144	63775	71709	74281	74675	81285	74603	102855	78366	67765	68938	62029	66060	79554	88511	67538	75649	79477	82889	
Kep Bangka Belitung	-	-	-	-	-	18722	19456	19083	20224	15499	12860	14400	9299	16502	15191	15975	16863	20908	18110	19077	21817	27370	
Kepulauan Riau	-	-	-	-	-	-	-	-	98671	135761	139766	53704	39989	173708	176196	241429	242602	243715	267900	316739	407828		
DKI Jakarta	366730	396707	461241	493428	493948	568054	606841	1282658	1307612	1309675	1658177	1399321	1524362	1676798	1986486	1201557	1611446	809522	820251	837776	851447		
Jawa Barat	550206	580130	660398	672833	721948	665884	700349	741123	778064	759361	866607	723960	639852	813527	1259973	1386562	1140474	1180249	1411853	1364622	1278987	1390628	
Jawa Tengah	421306	485831	568633	592767	641189	690027	721705	943421	754149	807115	869978	829094	1130139	1048992	1285562	1075901	1133719	1200072	1275162	1361421	1448984	1431582	
DI Yogyakarta	60547	66104	73889	76332	84672	89061	94662	87947	101529	105861	105579	120906	84447	122277	125052	118292	122124	129659	133121	140796	147465	161917	
Jawa Timur	58825	648205	714675	775483	826242	867718	920571	963687	996884	1010691	1107175	1105928	1122306	1807759	1699386	1529829	1355320	1432272	1570300	1629463	1176421	2009332	
Banten	166778	142294	158538	130888	148464	177778	165513	168737	74064	213058	215275	200111	203897	251301	238103	246957	275473						
B a l i	120362	144101	177265	197772	201778	205294	213509	216338	252050	290922	284122	250110	329965	621341	322126	334055	342412	353747	372463	390306	406959	436734	
Nusa Tenggara Barat	43483	49191	56550	61440	63142	75761	77453	82261	89618	96889	105131	269761	96744	119062	129081	143465	156683	173677	187959	202145	232662	250091	
Nusa Tenggara Timur	38439	43583	45827	50108	56933	61302	70587	63960	69864	59394	59394	163466	80262	55874	91755	98584	99266	128473	185128	117141	127921	139590	
Kalimantan Barat	58138	68946	77951	83528	90665	107984	111268	113586	113641	126267	131679	135873	134319	113845	155495	141103	160914	158431	158431	221150	187904	137516	
Kalimantan Tengah	26857	27697	35130	34840	36956	43394	33581	67516	81308	60147	80492	72728	70428	77489	83747	99342	98498	103364	110826	111891	118544	137505	
Kalimantan Selatan	65238	71563	81950	87574	100698	106055	121669	121073	143417	153044	182726	193853	189791	221571	374739	373207	398026	429275	324588	360097	370094	421590	
Kalimantan Timur	96910	113100	126485	133138	131868	100803	158388	151390	180066	188044	211091	181914	254663	273684	279413	304485	321208	457000	329692	370007	379732	466534	
Kalimantan Utara	92219	71818	93600	103473	104734	48299	63893	63797	62116	61922	91755	74490	57437	79121	158610	168264	180529	250938	268800	120763	134575	105636	
Sulawesi Tengah	24476	28830	29017	30051	34539	36599	34188	40596	41286	44869	62340	53909	53408	109139	56906	99195	73506	97203	90205	86407	103847	81797	
Sulawesi Selatan	127830	137001	145658	158776	170438	185243	196077	213388	243102	224832	251939	260900	271008	363982	361843	309910	320746	342066	361753	409438	404430	539516	
Sulawesi Tenggara	22910	26665	29990	32252	34298	32637	36400	37095	40915	44162	45771	47066	48428	50967	51761	59275	62609	69480	66344	72625	74327	109748	
Gorontalo	-	-	-	-	-	11967	15231	17702	19009	20080	21608	27005	26318	28920	29826	35620	47114	51738	49593	56141	63098	69362	
Sulawesi Barat	-	-	-	-	-	-	-	-	-	24829	26211	49270	17884	30203	16829	96502	79925	28894	27043	29366	34767	41843	
M a l i u	22969	27288	28458	33691	36725	35677	19068	9815	16038	18343	21271	22193	11024	20201	46217	29609	37583	33888	41695	42465	45521	53062	
Maluku Utara	-	-	-	-	-	-	-	12259	5536	14494	15708	16425	25223	28177	27676	28074	34345	35505	37912	38723	51899	62244	75475
Maluku Barat	-	-	-	-	-	-	-	-	-	-	-	4531	14765	15284	15546	15332	17769	19098	19186	18467	22198	22663	
P a p u a	41341	38258	57824	50634	57732	62107	70363	42550	64354	69795	13885	40019	39802	42547	42477	44675	44591	46233	46946	46946	81721		
Indonesia	3495509	3835459	4350318	4629524	4889647	5272728	5567613	6433478	6825458	6909130	7814216	7818026	7785213	9254555	10370451	9565778	9960010	9836538	10304006	10662081	11146504	1289458	

Gambar 1. Jumlah Pelanggan Perusahaan Air Bersih berdasarkan Provinsi Periode 1995-2017

[Sumber:www.bps.go.id]

Data tersebut selanjutnya terlebih dahulu dicari nilai rata-ratanya, untuk mempermudah pengerjaannya, peneliti dalam melakukan pencarian nilai rata-rata di *Microsoft Excel*. Berikut merupakan data jumlah rata-rata pelanggan air bersih berdasarkan provinsi di Indonesia periode 1995-2017 (Sumber: www.bps.go.id) yang telah diolah ke dalam *Microsoft Excel* untuk mempermudah cara perhitungan.

Tabel 1. Jumlah Rata-rata Pelanggan Perusahaan Air Bersih berdasarkan Provinsi

Provinsi	Rata-rata
Aceh	83021
Sumatera Utara	586285
Sumatera Barat	177686
R i a u	82423
J a m b i	148582

Sumatera Selatan	244769
Bengkulu	46863
Lampung	70525
Kep Bangka Belitung	17722
Kepulauan Riau	195068
DKI Jakarta	1021632
Jawa Barat	922572
Jawa Tengah	941670
DI Yogyakarta	107300
Jawa Timur	1175270
Banten	181072
B a l i	293807
Nusa Tenggara Barat	125056
Nusa Tenggara Timur	86675
Kalimantan Barat	130464
Kalimantan Tengah	74195
Kalimantan Selatan	222356
Kalimantan Timur	236832
Kalimantan Utara	39177

Sulawesi Utara	111620
Sulawesi Tengah	59619
Sulawesi Selatan	272999
Sulawesi Tenggara	49778
Gorontalo	34725
Sulawesi Barat	40120
Maluku	29446
Maluku Utara	32781
Papua Barat	16787
Papua	49931
Indonesia	7710154

3.1 Centroid Data

Centroid adalah titik tengah pada suatu *cluster*. *Centroid* berupa nilai. *Centroid* dimanfaatkan untuk menghitung jarak suatu objek data terhadap *centroid*. Suatu objek data termasuk dalam *cluster* jika mempunyai jarak terpendek terhadap *centroid cluster* tersebut[9]. Menentukan *centroid* awal dapat di peroleh secara acak. *Centroid* awal merupakan titik pusat cluster pertama. Penentuan *cluster* ada tiga bagian yaitu *cluster* tingkat Tinggi (C1), *cluster* tingkat Sedang (C2), dan *cluster* tingkat Rendah (C3). Untuk *cluster* tingkat Tinggi (Max) untuk *cluster* tingkat sedang (Rata-rata), dan untuk *cluster* tingkat rendah (Min). Berikut adalah *centroid* data awal dari penelitian ini:

Tabel 3. *Centroid* data awal

MAX (C1)	7710154
AVERAGE(C2)	446257
MIN(C3)	16787

3.2 Clustering Data

Proses *cluster* atau mengelompokan data bisa dilaksanakan dengan dua metode, yaitu Metode *Hirarki* dan Metode *Non-Hirarki*. Metode yang digunakan dalam penelitian ini adalah metode *Non-Hirarki*. Metode ini dimanfaatkan sebagai alternatif metode *cluster* untuk data yang berukuran lebih besar sebab memiliki kecepatan yang lebih tinggi dibandingkan dengan metode hirarki.[10]. Metode *non-hierarchical clustering* harus dimulai dengan penentuan jumlah cluster yang diinginkan (2 cluster, 3 cluster, atau lain sebagainya). Kemudian, setelah jumlah *cluster* sudah diperoleh, baru proses *cluster* dilakukan tanpa mengikuti proses hirarki. Berikut adalah data *clustering* yang telah dicari :

Iterasi 1			Jarak Terpendek	C1	C2	C3
C1	C2	C3				
7627133	363236	66234	66234			1
7123869	140028	569499	140028		1	
7532468	268570	160900	160900			1
7627732	363834	65636	65636			1
7561573	297675	131795	131795			1
7465386	201488	227982	201488		1	
7663291	399394	30076	30076			1
7639629	375731	53739	53739			1
7692433	428535	935	935			1
7515086	251188	178282	178282			1
6688523	575375	1004845	575375		1	
6787582	476316	905786	476316		1	
6768484	495414	924884	495414		1	
7602854	338956	90514	90514			1
6534885	729013	1158483	729013		1	
7529082	265184	164286	164286			1
7416347	152450	277020	152450		1	
7585099	321201	108269	108269			1
7623479	359582	69889	69889			1
7579690	315792	113678	113678			1
7635960	372062	57408	57408			1
7487798	223900	205570	205570			1
7473323	209425	220045	209425		1	
7670978	407080	22390	22390			1
7598534	334637	94834	94834			1
7650536	386638	42832	42832			1
7437155	173258	256212	173258		1	
7660376	396478	32992	32992			1
7675429	411531	17939	17939			1
7670034	406136	23334	23334			1
7680709	416811	12659	12659			1
7677373	413475	15995	15995			1
7693368	429470	0	0			1
7660223	396325	33145	33145			1
0	7263898	7693368	0		1	
			JUMLAH		1	9
				7710154,36	632870,6	88519,69

Gambar 2. Hasil Perhitungan dari *Centroid* Awal (Iterasi 1)

Proses *K-Means* akan terus beriterasi terus hingga pengelompokan data sama dengan pengelompokan data iterasi sebelumnya atau sama dengan proses akan terus melakukan iterasi sampai data pada iterasi terakhir sama dengan iterasi sebelumnya.

Tabel 3. Data *Centroid*(Iterasi 3)

MAX (C1)	7710154
AVERAGE(C2)	929486
MIN(C3)	112462

Setelah diperoleh nilai data *centroid* maka proses sama dilakukan dengan mencari jarak terdekat. Proses pencarian jarak terpendek. Berikut adalah table dari iterasi 3:

- Tirta Raharja Menggunakan Algoritma K-Means, Universitas Jenderal Achmad Yani, Jawa Barat,” vol. 2016, no. Sentika, pp. 18–19, 2016.
- [6] F. TASLIM, “Penerapan algoritma k-mean untuk clustering data obat pada puskesmas rumbai, Program Studi Teknik Informatika, Universitas Lancang Kuning Pekanbaru,” vol. x, no. x, pp. 108–114, 2016.
- [7] M. G. SADEWO, A. P. WINDARTO, and D. HARTAMA, “Penerapan Datamining Pada Populasi Daging Ayam Ras Pedaging Di Indonesia Berdasarkan Provinsi Menggunakan K-Means Clustering,” pp. 60–67, 2016.
- [8] L. MAULIDA, “Penerapan Data Mining Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. DKI Jakarta Dengan K-Means, Program Studi Manajemen Informatika, AMIK BSI Tangerang,” vol. 2, no. 3, pp. 167–174, 2018.
- [9] A. P. WINDARTO, “Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering, Program Studi Sistem Informasi, Data Mining,” vol. 16, no. 4, pp. 348–357, 2017.
- [10] E. RIANTI, “Data Mining Dalam Menentukan Kacamata Pada Optik Zal Laris Dan Tidak Laris Menggunakan Metode Clustering, Universitas Indonesia,” vol. 4, no. 2, pp. 267–283, 2017.