

Penerapan Data Mining Dalam Mengklusterisasi Location Best Pb Tambahan Pada Regional IV PT Indomarco Prismatama Cab.Medan Dengan Menggunakan Metode K-Means

Annisa Fitri *, Asyahri Hadi Nasyuha, S.Kom.,M.Kom**, Fifi Sonata, S.Kom., M.Kom.**

*Sistem Informasi, STMIK TRIGUNA DHARMA

**Sistem Informasi, STMIK TRIGUNA DHARMA

Article Info

Article history:

Keyword:

Area
K-Means
Cluster
Pengelompokan
Produk

ABSTRACT

Permintaan barang merupakan hal yang sangat penting bagi PT Indomarco Prismatama dalam pemenuhan kebutuhan stok toko. Pemenuhan stok yang tepat sangat berpengaruh besar untuk pendapatan laba yang besar bagi perusahaan. Salah satu hal penting lainnya yang perlu diperhatikan adalah area atau daerah persebaran toko indomaret dalam menjual produknya. Apalagi Indomaret merupakan perusahaan yang bergerak di bidang ritel dalam penjualan kebutuhan pokok, sehingga perusahaan tersebut juga harus memperhatikan daerah yang memiliki potensi jual yang paling besar.

K-Means Clustering merupakan salah satu metode dari data mining untuk pengelompokan suatu kasus ke dalam tiap kelas yang dihitung dengan konsep sebuah iterasi secara berulang untuk mendapatkan sebuah nilai rasio dimana akan dihentikannya perhitungan iterasi ketika nilai rasio tidak lagi lebih besar dari nilai rasio sebelumnya. Metode ini menggunakan rumus Euclidian sehingga ditemukan jarak yang paling dekat dari setiap data dengan centroid. Apabila salah satu cluster atau kelompok memiliki banyak data area, maka area di salah satu cluster itu merupakan kelompok area permintaan paling besar yang terpilih berdasarkan perhitungan K-Means Clustering.

Dengan membangun sebuah sistem dalam penelitian ini, yang dimana dalam penelitian ini menghasilkan sebuah ilmu pengetahuan baru yaitu mengetahui daerah atau area mana yang paling besar dalam pemenuhan pb tambahan yang diharapkan nantinya berguna bagi perusahaan untuk dapat memaksimalkan sistem pemenuhan pb tambahan.

Copyright © 2020 STMIK Triguna Dharma.

All rights reserved.

First Author :

Nama : Annisa Fitri
Program Studi : Sistem Informasi
E-Mail : annisafitri032@gmail.com

1. PENDAHULUAN

Data mining adalah suatu proses pengolahan data yang memiliki tujuan untuk menggali, mendapatkan dan menemukan sebuah pengetahuan yang tersembunyi dari sekumpulan data dalam ukuran besar disebut *data mining* dan dalam artian lain *data mining* merupakan sebuah teknik baru dalam menganalisis sejumlah data yang besar untuk mendapat sesuatu hal yang baru [1].

Clustering atau yang lebih dikenal dengan sebutan analisis *cluster* merupakan suatu proses pengelompokan satu set benda terlihat ataupun tidak ke dalam satu kelas objek yang sama dan di dalam *clustering*, terdapat banyak teknik-teknik yang digunakan, salah satunya adalah dengan menggunakan Algoritma *K-Means Clustering* [2].

PT Indomarco Prismatama atau yang lebih dikenal dengan sebutan Indomaret merupakan sebuah perusahaan ritel nasional waralaba terbesar di Indonesia dengan tujuan mendapatkan target pasar yang besar mulai dari kalangan masyarakat yang ada di Indonesia. Masalah yang terjadi pada perusahaan adalah dalam pemenuhan stok toko khususnya pada pb tambahan atas produk tertentu sering terjadi *over* stok ataupun stok kurang dikarenakan stok yang dipenuhi terkadang tidak sesuai dengan penjualan yang terjadi di daerah toko yang meminta stok pb tambahan.

Pb (Permintaan Barang) merupakan sebuah sistem yang digunakan oleh PT Indomarco Prismatama khususnya departemen *Distribution Center* dalam memenuhi kebutuhan toko. Hal ini digunakan untuk

untuk menunjang stok toko atas item-item atau produk tertentu yang memiliki nilai jual tinggi sehingga apabila pemenuhan sudah sesuai maka dapat mempengaruhi pendapat seles toko yang besar dan juga provit laba yang besar pula bagi perusahaan.

K-Means adalah sebuah metode dalam *clustering* yang memiliki sifat non hierarki dan berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster* dalam arti lain kelompok sehingga data yang memiliki karakteristik yang sama akan dikelompokkan kedalam *cluster* yang sama dan data yang berbeda dikelompokkan ke *cluster* yg lainnya [3].

Berdasarkan uraian latar belakang dan uraian permasalahan diatas maka diambil judul skripsi “Penerapan Data Mining Dalam Mengklusterisasi Location Best Regional IV PT Indomarco Prismatama Cab.Medan Dengan Menggunakan Metode K-Means”.

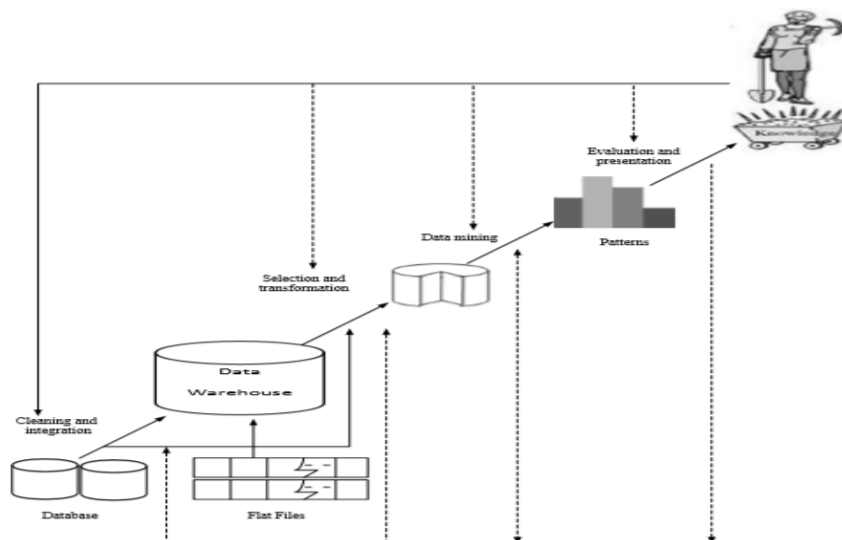
2 KAJIAN PUSTAKA

2.1 Data Mining

Suatu pola yang digunakan untuk menemukan hubungan antar item sehingga membentuk pola pengetahuan yang baru dalam data yang besar dan dapat diselesaikan menggunakan teknik tertentu seperti ilmu statistik, dan matematika disebut juga sebagai pengertian *data mining* menurut [4].

2.1.1 Tahap-Tahap Knowledge Discovery In Database

Proses KDD (*Knowledge Discovery Databases*) secara garis besar dapat dikemukakan sebagai berikut menurut [5] :



Gambar 1.1 Tahap-tahap Data Mining

1. *Data Selection*
Pemilihan (seleksi) sebuah data di dalam sekumpulan data operasional adalah tahap penggalian informasi KDD sebelum dimulai. Proses *data mining* menggunakan data hasil seleksi yang nantinya akan disimpan ke dalam suatu berkas, terpisah dari basis data operasional.
2. *Pre-processing / Cleaning*
Sebelum terjadinya proses *data mining* diperlukannya proses pembersihan pada data yang akan menjadi fokus KDD. Proses pembersihan data mencakup hal-hal seperti membuang duplikasi data, memeriksa data yang inkonsisten dan memperbaiki kesalahan pada data, seperti kesalahan cetak.
3. *Transformation Coding*
Merupakan proses transformasi pada data yang sudah terpilih, yang nantinya data tersebut sudah sesuai untuk proses *data mining*. Proses *coding* di dalam KDD adalah suatu proses yang mempunyai karakteristik dan tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.
4. *Data mining*
Data mining merupakan sebuah proses mencari suatu pola atau sebuah informasi yang menarik dalam data yang terpilih dengan menggunakan teknik atau metode tertentu. Sebuah teknik, metode ataupun algoritma dalam *data mining* sangatlah bervariasi sehingga dalam pemilihan algoritma atau metode yang tepat tergantung pada tujuan dan proses KDD secara keseluruhan.
5. *Interpretation / Evaluation*
Proses dari *data mining* menghasilkan sebuah pola informasi dan penampilan pola tersebut harus dalam bentuk yang mudah untuk dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan tahap

terakhir dari KDD, tujuan dari tahap ini adalah untuk memastikan apakah pola atau informasi yang telah ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

2.2 Clustering

Makna lain yang mengenai *clustering* adalah sebuah mekanisme klasifikasi dan banyak digunakan untuk dirancang dalam mengkategorikan data juga dipaparkan oleh [6]. Penggunaan teknik seperti ini sering digunakan dalam mode eksplorasi data, dan *clustering* dapat dilakukan dengan menggunakan berbagai algoritma yang berbeda salah satunya adalah *K-Means*.

2.3 Algoritma Metode K-Means

K-Means adalah sebuah metode dalam *clustering* yang memiliki sifat non hierarki dan berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster* dalam arti lain kelompok sehingga data yang memiliki karakteristik yang sama akan dikelompokkan kedalam *cluster* yang sama dan data yang berbeda dikelompokkan ke *cluster* yg lainnya [3]. Oleh karena itu, berikut ini merupakan cara kerja dari algoritma *K-Means*:

1. Menentukan k sebagai jumlah *cluster* yang ingin dibentuk.
2. Kemudian membangkitkan k centroid (titik pusat *cluster*) awal secara *random*.
3. Menghitung jarak dari setiap data ke masing-masing *centroid*. Pada tahap ini perlu dihitung antara jarak tiap data ke setiap pusat *cluster*. Jarak antara satu data dengan satu *cluster* tertentu nantinya akan menentukan sebuah data yang masuk kedalam *cluster* mana. Oleh karena itu, untuk menghitung antara jarak ke semua data ke setiap titik pusat *cluster* dapat menggunakan teori jarak *Euclidian* yang dirumuskan sebagai berikut :

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

Dimana :

$D(i, j)$ = Jarak data ke i ke pusat *cluster* j

X_{ki} = Data ke i pada atribut data ke k

X_{kj} = Titik pusat ke j pada atribut ke k

4. Lalu setiap data akan dilakukan pemilihan jarak centroid yang terdekat.
5. Dan tentukan posisi centroid yang baru dengan cara menghitung nilai rata-rata dari data yang terletak pada centroid yang sama.
6. Kemudian kembali ke langkah-3 jika pada posisi centroid yang baru dengan yang lama tidak sama dan begitu seterusnya.

3. METODOLOGI PENELITIAN

3.1 Metode Penelitian

Dalam melakukan suatu penelitian memerlukan langkah-langkah atau cara tertentu yang menjadi pedoman selama proses penelitian, agar hasil penelitian sesuai dengan tujuan yang telah ditetapkan. Berikut adalah metodologi dalam penelitian ini yaitu:

1. Teknik Pengumpulan Data (*Data Collecting*)

Adapun beberapa teknik yang digunakan dalam pengumpulan data dari penelitian yaitu :

- a. Pengamatan (*Observasi*)

Observasi merupakan teknik pengumpulan data yang dilakukan dengan cara melakukan tinjauan langsung ke tempat study kasus dimana akan dilakukan penelitian.

- b. Wawancara (*Interview*)

Teknik wawancara dilakukan untuk mendapatkan informasi tambahan tentang data dari pihak-pihak yang memiliki wewenang agar memperoleh data yang dibutuhkan untuk menunjang penelitian ini.

2. Studi Kepustakaan (*Library Search*)

Untuk mendapatkan hasil teori yang valid untuk dijadikan sebuah landasan dapat mempelajari beberapa buku referensi.

3.2 Metode Perancangan Sistem

Di dalam metode perancangan sistem khususnya mengenai *software* atau perangkat lunak kita dapat mengadopsi beberapa metode diantaranya algoritma *waterfall*. Yang dimulai dari menganalisis masalah, Desain sistem, Pembangunan sistem, Uji coba, dan Implementasi sistem.

3.3 Algoritma Sistem

Algoritma sistem merupakan penjelasan dari langkah-langkah penyelesaian masalah dalam perancangan sistem *data mining* dalam klusterisasi dalam menganalisis area atau daerah yang memiliki permintaan barang paling besar dengan menggunakan metode *K-Means Clustering*.

Analisis yang dilakukan pada PT Indomarco Prismatama Cab.Medan yang bergerak dalam bidang ritel waralaba terbesar di Indonesia, sehingga dalam pemenuhan stok kebutuhan toko harus memiliki berbagai aspek yang dapat menguntungkan. Salah satunya adalah untuk pemenuhan stok tambahan yang yang mempengaruhi dengan jumlah penjualan disetiap masing-masing daerah toko tersebut yang perlu diperhatikan. Hal ini dapat dibantu dengan analisa menggunakan sebuah ilmu pengetahuan yaitu berupa *data mining*.

Berikut merupakan tabel area dari masing-masing *regional IV* PT Indomarco Prismatama Cab.Medan sampel yang diambil dari perusahaan dengan sampel produk rokok sebagai perhitungan pada metode.

Tabel 1.1 Data Permintaan Produk Rokok Setiap Kabupaten

No	Area	Jlh Toko	Jlh Item
1	Aceh	2	15
2	Aceh Barat	8	45
3	Aceh Besar	7	41
4	Aceh Jaya	1	11
5	Aceh Selatan	4	38
6	Aceh Singkil	1	8
7	Aceh Tamiang	12	48
8	Aceh Tengah	4	13
9	Aceh Tenggara	8	42
10	Aceh Utara	4	10
11	Asahan	21	54
12	Banda Aceh	22	14
13	Batu Bara	12	41
14	Belawan	1	10
15	Bener Meriah	2	5
16	Binjai	14	48
17	Bireuen	7	24
18	Dairi	11	49
19	Deli Serdang	131	84
20	Gayo Lues	1	17
21	Humbang Hasundutan	2	24
22	Karo	34	66
23	Kota Medan	250	85
24	Labuhan Batu	2	15
25	Labuhan Batu Utara	9	33
26	Langkat	25	60
27	Langsa	4	26
28	Lhokseumawe	1	12
29	Nagan Raya	4	20
30	Pakpak Barat	1	18
31	Pidie	3	19
32	Pidie Jaya	3	6
33	Samosir	6	52
34	Serdang Bedagai	51	24
35	Siantar	23	76
36	Sibolga	5	39
37	Sigli	1	7
38	Simalungun	28	54
39	Subulussalam	4	20
40	Tanjung Balai	15	10

Dalam proses pengelompokan data kategori permintaan barang menggunakan 2 kriteria sebagai berikut :

Tabel 1.2 Kriteria

Kode Kriteria	Kriteria
K1	Jumlah Toko
K2	Jumlah Item

Dan kemudian dilakukan penyelesaian masalah dengan mengadopsi metode Sesuai dengan referensi yang telah dipaparkan pada bab sebelumnya, berikut ini adalah langkah-langkah penyelesaiannya yaitu:

1. Iterasi ke – 1

a. Penentuan pusat (*centroid*) awal *cluster*

Untuk menentukan pusat (*centroid*) awal ditentukan dengan acak (*random*) dari data yang sudah ada. Pada kasus ini pusat *centroid* awal adalah data Aceh, Batu Bara, dan Tanjung Balai.

Tabel 1.3 Titik Pusat (*Centroid*) Awal *Cluster*

Centroid	Data	K1	K2
Centroid 1	Aceh	2	15
Centroid 2	Batu Bara	12	41
Centroid 3	Tanjung Balai	15	10

b. Perhitungan jarak dengan pusat *cluster*

Perhitungan jarak dari data area ke-1 terhadap titik pusat *cluster* adalah:

$$D(1,1) = \sqrt{(2 - 2)^2 + (15 - 15)^2} = 0,00$$

$$D(1,2) = \sqrt{(2 - 12)^2 + (15 - 41)^2} = 27,86$$

$$D(1,3) = \sqrt{(2 - 15)^2 + (15 - 10)^2} = 13,93$$

Dan seterusnya dilakukan perhitungan jarak untuk data kategori area ke-2 sampai dengan data ke-40 kemudian akan didapatkan hasil perhitungan jarak setiap data terhadap pusat *cluster* baru sebagai berikut

Tabel 1.4 Hasil Perhitungan pada Iterasi ke-1

Area	C1	C2	C3	Jarak Terpendek	Cluster	Jarak Terpendek ²
Aceh	0,00	27,86	13,93	0,00	C1	0,00
Aceh Barat	30,59	5,66	35,69	5,66	C2	32,00
Aceh Besar	26,48	5,00	32,02	5,00	C2	25,00
Aceh Jaya	4,12	31,95	14,04	4,12	C1	17,00
Aceh Selatan	23,09	8,54	30,08	8,54	C2	73,00
Aceh Singkil	7,07	34,79	14,14	7,07	C1	50,00
Aceh Tamiang	34,48	7,00	38,12	7,00	C2	49,00
Aceh Tengah	2,83	29,12	11,40	2,83	C1	8,00
Aceh Tenggara	27,66	4,12	32,76	4,12	C2	17,00
Aceh Utara	5,39	32,02	11,00	5,39	C1	29,00
Asahan	43,38	15,81	44,41	15,81	C2	250,00
Banda Aceh	20,02	28,79	8,06	8,06	C3	65,00
Belawan	5,10	32,89	14,00	5,10	C1	26,00
Bener Meriah	10,00	37,36	13,93	10,00	C1	100,00
Binjai	35,11	7,28	38,01	7,28	C2	53,00
Bireuen	10,30	17,72	16,12	10,30	C1	106,00
Dairi	35,17	8,06	39,20	8,06	C2	65,00
Deli Serdang	146,29	126,53	137,59	126,53	C2	16.010,00
Gayo Lues	2,24	26,40	15,65	2,24	C1	5,00
Humbang Hasundutan	9,00	19,72	19,10	9,00	C1	81,00
Karo	60,21	33,30	59,14	33,30	C2	1.109,00
Kota Medan	257,69	242,03	246,68	242,03	C2	58.580,00
Labuhan Batu	0,00	27,86	13,93	0,00	C1	0,00
Labuhan Batu Utara	19,31	8,54	23,77	8,54	C2	73,00
Langkat	50,54	23,02	50,99	23,02	C2	530,00
Langsa	11,18	17,00	19,42	11,18	C1	125,00
Lhokseumawe	3,16	31,02	14,14	3,16	C1	10,00
Nagan Raya	5,39	22,47	14,87	5,39	C1	29,00
Pakpak Barat	3,16	25,50	16,12	3,16	C1	10,00

Tabel 1.4 Hasil Perhitungan pada Iterasi ke-1 (Lanjutan)

Area	C1	C2	C3	Jarak Terpendek	Cluster	Jarak Terpendek ²
Pidie	4,12	23,77	15,00	4,12	C1	17,00
Pidie Jaya	9,06	36,14	12,65	9,06	C1	82,00
Samosir	37,22	12,53	42,95	12,53	C2	157,00
Serdang Bedagai	49,82	42,54	38,63	38,63	C3	1.492,00
Siantar	64,51	36,69	66,48	36,69	C2	1.346,00
Sibolga	24,19	7,28	30,68	7,28	C2	53,00
Sigli	8,06	35,74	14,32	8,06	C1	65,00
Simalungun	46,87	20,62	45,88	20,62	C2	425,00
Subulussalam	5,39	22,47	14,87	5,39	C1	29,00
Tanjung Balai	13,93	31,14	0,00	0,00	C3	0,00
					WCV	81.193,00

c. Menghitung Besar Rasio

Rasio besaran antara BCV (*Between Cluster Variation*) dengan WCV (*Within Cluster Variation*) :

Karena centroid m1 (2;15), m2 (12;41), m3 (15;10)

$$d(m1,m2) = \sqrt{(2 - 12)^2 + (15 - 41)^2} = 27,587$$

$$d(m1,m3) = \sqrt{(2 - 15)^2 + (15 - 10)^2} = 13,928$$

$$d(m2,m3) = \sqrt{(12 - 15)^2 + (41 - 10)^2} = 31,145$$

$$BCV = d(m1,m2) + d(m1,m3) + d(m2,m3)$$

$$= 27,587 + 13,928 + 31,145$$

$$= 72,930$$

$$\text{Besar Rasio} = 72,930/81.193,00 = 0,001$$

Karena perhitungan masih perhitungan pada iterasi 1 maka lanjutkan ke iterasi 2.

d. Perhitungan pusat *cluster* baru

Untuk mendapatkan pusat *cluster* yang baru diperlukan untuk mengelompokkan data berdasarkan jarak terdekat dengan pusat *cluster*. Kemudian pusat *cluster* baru dihitung berdasarkan data anggota tiap-tiap *cluster* dan membagikan dengan jumlah anggota masing-masing *cluster*.

Tabel 1.5 Kelompok *Cluster* 1

C1			
Area	K1	K2	Cluster
Aceh	2	15	C1
Aceh Jaya	1	11	C1
Aceh Singkil	1	8	C1
Aceh Tengah	4	13	C1
Aceh Utara	4	10	C1
Belawan	1	10	C1
Bener Meriah	2	5	C1
Bireuen	7	24	C1
Gayo Lues	1	17	C1
Humbang Hasundutan	2	24	C1
Labuhan Batu	2	15	C1
Langsa	4	26	C1
Lhokseumawe	1	12	C1
Nagan Raya	4	20	C1
Pakpak Barat	1	18	C1
Pidie	3	19	C1
Pidie Jaya	3	6	C1
Sigli	1	7	C1
Subulussalam	4	20	C1
Mean C1	2,53	14,74	19

Tabel 1.6 Kelompok *Cluster 2*

C2			
Area	K1	K2	Cluster
Aceh Barat	8	45	C2

Tabel 1.6 Kelompok *Cluster 2* (Lanjutan)

C2			
Area	K1	K2	Cluster
Aceh Besar	7	41	C2
Aceh Selatan	4	38	C2
Aceh Tamiang	12	48	C2
Aceh Tenggara	8	42	C2
Asahan	21	54	C2
Batu Bara	12	41	C2
Binjai	14	48	C2
Dairi	11	49	C2
Deli Serdang	131	84	C2
Karo	34	66	C2
Kota Medan	250	85	C2
Labuhan Batu Utara	9	33	C2
Langkat	25	60	C2
Samosir	6	52	C2
Siantar	23	76	C2
Sibolga	5	39	C2
Simalungun	28	54	C2
Mean C2	33,78	53,06	18

Tabel 1.7 Kelompok *Cluster 3*

C3			
Area	K1	K2	Cluster
Banda Aceh	22	14	C3
Serdang Bedagai	51	24	C3
Tanjung Balai	15	10	C3
Mean C3	29,33	16,00	3

Dari hasil tabel *cluster* tersebut maka diperoleh *centroid* baru yaitu :

$$m_1 = (2,53;14,74)$$

$$m_2 = (33,78;53,06)$$

$$m_3 = (29,33;16,00)$$

Tabel 1.8 Pusat (*centroid*) *cluster* baru

Centroid	K1	K2
Centroid 1	2,53	14,74
Centroid 2	33,78	53,06
Centroid 3	29,33	16,00

2. Iterasi ke - 2

a. Perhitungan jarak dengan pusat *cluster*

Perhitungan jarak dari data pelanggan ke-1 terhadap titik pusat *cluster* adalah:

$$D(1,1) = \sqrt{(2 - 2,53)^2 + (15 - 14,74)^2} = 0,59$$

$$D(1,2) = \sqrt{(2 - 33,78)^2 + (15 - 53,06)^2} = 49,58$$

$$D(1,3) = \sqrt{(2 - 29,33)^2 + (15 - 16,00)^2} = 27,35$$

Dan seterusnya dilakukan perhitungan jarak untuk data kategori area ke-2 sampai data ke-40 kemudian akan didapatkan hasil perhitungan jarak setiap data terhadap pusat *cluster* baru sebagai berikut :

Tabel 1.9 Hasil Perhitungan pada iterasi ke-2

Area	C1	C2	C3	Jarak Terpendek	Cluster	Jarak Terpendek ²
Aceh	0,59	49,58	27,35	0,59	C1	0,35
Aceh Barat	30,75	27,01	36,00	27,01	C2	729,39
Aceh Besar	26,64	29,37	33,52	26,64	C1	709,77
Aceh Jaya	4,04	53,32	28,77	4,04	C1	16,29
Aceh Selatan	23,31	33,37	33,55	23,31	C1	543,35
Aceh Singkil	6,91	55,72	29,44	6,91	C1	47,71
Aceh Tamiang	34,59	22,36	36,39	22,36	C2	499,83
Aceh Tengah	2,28	49,91	25,51	2,28	C1	5,19
Aceh Tenggara	27,81	28,05	33,63	27,81	C1	773,24
Aceh Utara	4,96	52,35	26,03	4,96	C1	24,61
Asahan	43,39	12,81	38,90	12,81	C2	164,16
Banda Aceh	19,49	40,79	7,60	7,60	C3	57,78
Batu Bara	27,92	24,89	30,42	24,89	C2	619,61
Belawan	4,98	54,11	28,96	4,98	C1	24,77
Bener Meriah	9,75	57,61	29,46	9,75	C1	95,08
Binjai	35,19	20,41	35,48	20,41	C2	416,72
Bireuen	10,29	39,51	23,72	10,29	C1	105,82
Dairi	35,30	23,14	37,75	23,14	C2	535,27
Deli Serdang	145,96	102,03	122,31	102,03	C2	10.409,72
Gayo Lues	2,73	48,73	28,35	2,73	C1	7,45
Humbang Hasundutan	9,28	43,06	28,48	9,28	C1	86,08
Karo	60,15	12,95	50,22	12,95	C2	167,61
Kota Medan	257,26	218,57	231,20	218,57	C2	47.772,50
Labuhan Batu	0,59	49,58	27,35	0,59	C1	0,35
Labuhan Batu Utara	19,38	31,88	26,50	19,38	C1	375,45
Langkat	50,54	11,19	44,21	11,19	C2	125,27
Langsa	11,36	40,23	27,24	11,36	C1	129,03
Lhokseumawe	3,13	52,54	28,61	3,13	C1	9,82
Nagan Raya	5,47	44,49	25,65	5,47	C1	29,87
Pakpak Barat	3,60	47,99	28,40	3,60	C1	12,98
Pidie	4,29	45,90	26,50	4,29	C1	18,40
Pidie Jaya	8,75	56,23	28,17	8,75	C1	76,56
Samosir	37,42	27,80	42,90	27,80	C2	772,72
Serdang Bedagai	49,35	33,78	23,10	23,10	C3	533,44
Siantar	64,59	25,35	60,33	25,35	C2	642,61
Sibolga	24,39	32,03	33,48	24,39	C1	594,82
Sigli	7,89	56,53	29,73	7,89	C1	62,19
Simalungun	46,80	5,85	38,02	5,85	C2	34,27
Subulussalam	5,47	44,49	25,65	5,47	C1	29,87
Tanjung Balai	13,34	46,97	15,54	13,34	C1	178,03
					WCV	67.437,98

b. Menghitung Besar Rasio

Rasio besaran antara BCV (*Between Cluster Variation*) dengan WCV (*Within Cluster Variation*) :

Karena centroidnya adalah sebagai berikut :

$$m1 = (2,53;14,74)$$

$$m2 = (33,78;53,06)$$

$$m3 = (29,33;16,00)$$

$$d(m1,m2) = \sqrt{(2,53 - 33,78)^2 + (14,74 - 53,06)^2} = 49,447$$

$$d(m1,m3) = \sqrt{(2,53 - 29,33)^2 + (14,74 - 16,00)^2} = 26,837$$

$$d(m2,m3) = \sqrt{(33,78 - 29,33)^2 + (53,06 - 16,00)^2} = 37,321$$

$$BCV = d(m1,m2) + d(m1,m3) + d(m2,m3)$$

$$= 49,447 + 26,837 + 37,321 = 113,605$$

$$\text{Besar Rasio} = 113,605/67.437,98 = 0,002$$

Karena rasio yang baru lebih besar dari rasio sebelumnya maka diperlukan untuk melanjutkan perhitungan ke iterasi selanjutnya.

c. Perhitungan pusat *cluster* baru

Untuk mendapatkan pusat *cluster* yang baru diperlukan untuk klusterisasi data berdasarkan jarak terdekat dengan pusat *cluster*. Kemudian pusat *cluster* baru dihitung berdasarkan data anggota tiap-tiap *cluster* dan membagikan dengan jumlah anggota masing-masing *cluster*.

Tabel .10 Kelompok *Cluster* 1

C1			
Area	K1	K2	Cluster
Aceh	2	15	C1
Aceh Besar	7	41	C1
Aceh Jaya	1	11	C1
Aceh Selatan	4	38	C1
Aceh Singkil	1	8	C1
Aceh Tengah	4	13	C1
Aceh Tenggara	8	42	C1
Aceh Utara	4	10	C1
Belawan	1	10	C1
Bener Meriah	2	5	C1
Bireuen	7	24	C1
Gayo Lues	1	17	C1
Humbang Hasundutan	2	24	C1
Labuhan Batu	2	15	C1
Labuhan Batu Utara	9	33	C1
Langsa	4	26	C1
Lhokseumawe	1	12	C1
Nagan Raya	4	20	C1
Pakpak Barat	1	18	C1
Pidie	3	19	C1
Pidie Jaya	3	6	C1
Sibolga	5	39	C1
Sigli	1	7	C1
Subulussalam	4	20	C1
Tanjung Balai	15	10	C1
Mean C1	3,84	19,32	25

Tabel 1.11 Kelompok *Cluster* 2

C2			
Area	K1	K2	Cluster
Aceh Barat	8	45	C2
Aceh Tamiang	12	48	C2
Asahan	21	54	C2
Batu Bara	12	41	C2
Binjai	14	48	C2
Dairi	11	49	C2
Deli Serdang	131	84	C2
Karo	34	66	C2
Kota Medan	250	85	C2
Langkat	25	60	C2
Samosir	6	52	C2

Tabel 1.12 Kelompok *Cluster 2* (Lanjutan)

C2			
Area	K1	K2	Cluster
Siantar	23	76	C2
Simalungun	28	54	C2
Mean C2	44,23	58,62	13

Tabel 1.12 Kelompok *Cluster 3*

C3			
Area	K1	K2	Cluster
Banda Aceh	22	14	C3
Serdang Bedagai	51	24	C3
Mean C3	36,50	19,00	2

Dari tabel diatas diperoleh *centroid* baru yaitu :

$$m1 = (3,84;19,32)$$

$$m2 = (44,32;58,62)$$

$$m3 = (36,50,19,00)$$

Tabel 1.13 Pusat (*centroid*) *cluster* baru

Centroid	K1	K2
Centroid 1	3,84	19,32
Centroid 2	44,23	58,62
Centroid 3	36,50	19,00

3. Iterasi ke – 3

a. Perhitungan jarak dengan pusat *cluster*

Perhitungan jarak dari data area ke-1 terhadap titik pusat *cluster* adalah sebagai berikut :

$$D(1,1) = \sqrt{(2 - 3,84)^2 + (15 - 19,32)^2} = 4,70$$

$$D(1,2) = \sqrt{(2 - 44,23)^2 + (15 - 58,62)^2} = 60,71$$

$$D(1,3) = \sqrt{(2 - 36,50)^2 + (15 - 19,00)^2} = 34,73$$

Dan seterusnya dilakukan perhitungan jarak untuk data kategori area ke-2

sampai dengan data area ke-40 kemudian akan didapatkan hasil perhitungan jarak setiap data terhadap pusat *cluster* baru sebagai berikut :

Tabel 1.14 Hasil Perhitungan pada iterasi ke-3

Area	C1	C2	C3	Jarak Terpendek	Cluster	Jarak Terpendek ²
Aceh	4,70	60,71	34,73	4,70	C1	22,048
Aceh Barat	26,01	38,70	38,58	26,01	C1	676,768
Aceh Besar	21,91	41,19	36,8	21,91	C1	480,008
Aceh Jaya	8,79	64,31	36,39	8,79	C1	77,288
Aceh Selatan	18,68	45,21	37,65	18,68	C1	348,968
Aceh Singkil	11,67	66,56	37,17	11,67	C1	136,208
Aceh Tamiang	29,82	33,93	37,96	29,82	C1	889,128
Aceh Tengah	6,32	60,82	33,05	6,32	C1	39,968
Aceh Tenggara	23,06	39,86	36,62	23,06	C1	531,688
Aceh Utara	9,32	63,10	33,72	9,32	C1	86,888
Asahan	38,69	23,68	38,28	23,68	C2	560,97041
Banda Aceh	18,92	49,85	15,34	15,34	C3	235,25
Batu Bara	23,16	36,73	32,93	23,16	C1	536,608
Belawan	9,74	65,06	36,62	9,74	C1	94,928
Bener Meriah	14,44	68,25	37,23	14,44	C1	208,448

Tabel 1.14 Hasil Perhitungan pada iterasi ke-3 (Lanjutan)

Area	C1	C2	C3	Jarak Terpendek	Cluster	Jarak Terpendek ²
Binjai	30,43	32,04	36,7	30,43	C1	925,768
Bireuen	5,65	50,84	29,92	5,65	C1	31,888
Dairi	30,53	34,59	39,37	30,53	C1	932,168
Deli Serdang	142,66	90,41	114,7	90,41	C2	8173,2781
Gayo Lues	3,67	60,01	35,56	3,67	C1	13,448
Humbang Hasundutan	5,03	54,60	34,86	5,03	C1	25,288
Karo	55,58	12,62	47,07	12,62	C2	159,20118
Kota Medan	254,77	207,45	223,5	207,45	C2	43037,124
Labuhan Batu	4,70	60,71	34,73	4,70	C1	22,048
Labuhan Batu Utara	14,62	43,56	30,86	14,62	C1	213,768
Langkat	45,85	19,28	42,58	19,28	C2	371,73964
Langsa	6,68	51,79	33,25	6,68	C1	44,648
Lhokseumawe	7,85	63,58	36,18	7,85	C1	61,648
Nagan Raya	0,70	55,76	32,52	0,70	C1	0,488
Pakpak Barat	3,13	59,32	35,51	3,13	C1	9,808
Pidie	0,90	57,18	33,5	0,90	C1	0,808
Pidie Jaya	13,35	66,85	35,93	13,35	C1	178,128
Samosir	32,75	38,80	44,94	32,75	C1	1072,648
Serdang Bedagai	47,39	35,27	15,34	15,34	C3	235,25
Siantar	59,83	27,44	58,58	27,44	C2	752,97041
Sibolga	19,71	43,86	37,31	19,71	C1	388,648
Sigli	12,64	67,33	37,47	12,64	C1	159,848
Simalungun	42,27	16,87	36,02	16,87	C2	284,73964
Subulussalam	0,70	55,76	32,52	0,70	C1	0,488
Tanjung Balai	14,54	56,73	23,31	14,54	C1	211,408
					WCV	62.232,41

b. Menghitung Besar Rasio

Rasio besaran antara BCV (*Between Cluster Variation*) dengan WCV (*Within Cluster Variation*):

Karena centroidnya adalah sebagai berikut :

$$m1 = (3,84;19,32)$$

$$m2 = (44,32;58,62)$$

$$m3 = (36,50,19,00)$$

$$d(m1,m2) = \sqrt{(3,84 - 44,32)^2 + (19,32 - 58,62)^2} = 56,352$$

$$d(m1,m3) = \sqrt{(3,84 - 36,50)^2 + (19,32 - 19,00)^2} = 32,662$$

$$d(m2,m3) = \sqrt{(44,32 - 36,50)^2 + (58,62 - 19,00)^2} = 40,363$$

$$BCV = d(m1,m2) + d(m1,m3) + d(m2,m3)$$

$$= 56,352 + 32,662 + 40,363 = 129,376$$

$$\text{Besar Rasio} = 129,376/62.232,41 = 0,002$$

Karena rasio yang baru tidak lagi lebih besar dari rasio sebelumnya maka proses perhitungan iterasi dihentikan.

4. Kesimpulan

Karena iterasi telah mencapai akhir atau selesai maka dari perhitungan tersebut diperoleh kesimpulan sebagai berikut:

Tabel 1.15 Hasil Proses *K-Means*

Area	C1	C2	C3
Aceh	Area Permintaan Besar		
Aceh Barat	Area Permintaan Besar		
Aceh Besar	Area Permintaan Besar		
Aceh Jaya	Area Permintaan Besar		

Tabel 1.15 Hasil Proses *K-Means* (Lanjutan)

Area	C1	C2	C3
Aceh Selatan	Area Permintaan Besar		
Aceh Singkil	Area Permintaan Besar		
Aceh Tamiang	Area Permintaan Besar		
Aceh Tengah	Area Permintaan Besar		
Aceh Tenggara	Area Permintaan Besar		
Aceh Utara	Area Permintaan Besar		
Asahan		Area Permintaan Sedang	
Banda Aceh			Area Permintaan Kecil
Batu Bara	Area Permintaan Besar		
Belawan	Area Permintaan Besar		
Bener Meriah	Area Permintaan Besar		
Binjai	Area Permintaan Besar		
Bireuen	Area Permintaan Besar		
Dairi	Area Permintaan Besar		
Deli Serdang		Area Permintaan Sedang	
Gayo Lues	Area Permintaan Besar		
Humbang Hasundutan	Area Permintaan Besar		
Karo		Area Permintaan Sedang	
Kota Medan		Area Permintaan Sedang	
Labuhan Batu	Area Permintaan Besar		
Labuhan Batu Utara	Area Permintaan Besar		
Langkat		Area Permintaan Sedang	
Langsa	Area Permintaan Besar		
Lhokseumawe	Area Permintaan Besar		
Nagan Raya	Area Permintaan Besar		
Pakpak Barat	Area Permintaan Besar		
Pidie	Area Permintaan Besar		
Pidie Jaya	Area Permintaan Besar		
Samosir	Area Permintaan Besar		
Serdang Bedagai			Area Permintaan Kecil
Siantar		Area Permintaan Sedang	
Sibolga	Area Permintaan Besar		
Sigli	Area Permintaan Besar		
Simalungun		Area Permintaan Sedang	
Subulussalam	Area Permintaan Besar		
Tanjung Balai	Area Permintaan Besar		

Keterangan :

Klusterisasi untuk area permintaan barang atas sampel item rokok berdasarkan data yang diperoleh dari PT Indomarco Prismatama Cab.Medan untuk pusat *cluster* pertama dinyatakan ada 31 kategori area permintaan barang dengan kondisi area permintaan besar. Sedangkan untuk pusat *cluster* kedua dinyatakan ada 7 kategori area permintaan barang yang sedang dan pada pusat *cluster* ketiga dinyatakan ada 2 kategori area permintaan barang yang kecil.

4. BENTUK INTERFACE APLIKASI

1. *Interface Form* Menu Utama

Menu utama merupakan tampilan awal setelah berhasil login ke sistem. Menu pada menu utama masih belum dapat diakses apabila pengguna belum melakukan *login* ke sistem. Dalam *form* menu utama terdapat menu data, menu proses, menu menampilkan laporan, serta menu keluar dari aplikasi. Berikut ini adalah tampilan *form* utama beranda :

Gambar 1.2 Tampilan *Interface Form* Menu Utama2. *Interface Form* Permintaan Barang

Form permintaan barang merupakan *form* yang digunakan untuk menampilkan data permintaan barang, yang meliputi id, nama area, jumlah toko, dan jumlah item. Pada *form* permintaan barang terdiri 5 (lima) buah tombol *button* yaitu tambah, edit, bersih, hapus, dan batal. Adapun tampilan *form* permintaan barang adalah sebagai berikut :

Id Area	Area	Jlh Toko	Jlh Item
A001	Aceh	2	15
A012	Banda Aceh	22	14
A023	Kota Medan	250	85
A034	Serdang Bedagai	51	24
A040	Tanjung Balai	15	10
A002	Aceh Barat	8	45
A003	Aceh Besar	7	41
A004	Aceh Jaya	1	11
A005	Aceh Selatan	4	38

Gambar 1.3 Tampilan *Interface Form* Permintaan Barang3. *Interface Form* Centroid

Berikut ini terdapat data permintaan barang yang digunakan untuk menentukan titik *cluster* awal yang telah diinisialisasi, yang nantinya digunakan dalam perhitungan *k-means*. Kemudian terdapat tiga *button* yang digunakan sebagai perintah untuk melakukan pengolahan data seperti simpan, batal, dan keluar *form*.

Id Area	Area	Jlh Toko	Jlh Item
A001	Aceh	2	15
A012	Banda Aceh	22	14
A023	Kota Medan	250	85
A034	Serdang Bedagai	51	24
A040	Tanjung Balai	15	10
A002	Aceh Barat	8	45
A003	Aceh Besar	7	41
A004	Aceh Jaya	1	11
A005	Aceh Selatan	4	38
A006	Aceh Singkil	1	8
A007	Aceh Tamiang	12	48
A008	Aceh Tengah	4	13
A009	Aceh Tenggara	8	42
A010	Aceh Utara	4	10
A011	Asahan	21	54

Id Area	Area	Jlh Toko	Jlh Item
A001	Aceh	2	15
A013	Batu Bara	12	41
A040	Tanjung Balai	15	10

Gambar 1.4 Tampilan *Interface Form* Centroid4. Tampilan *Interface Form* K-Means

Pada *form* tersebut terdapat *listview* titik *cluster* yang telah dipilih pada *form* centroid sebelumnya, *listview* anggota *cluster* 1, *listview* anggota *cluster* 2, dan *listview* anggota *cluster* 3 dari masing-masing iterasi dan kesimpulan dari hasil perhitungan *K-Means*. Kemudian terdapat tiga buah *button* yaitu proses, batal dan keluar *form* yang masing-masing digunakan untuk melakukan proses perhitungan metode *K-Means*, membatalkan proses perhitungan, serta menutup *form*. Berikut tampilan *interface form* *K-Means*.

Centroid	Id Area	Area	K1	K2
1	A001	Aceh	2	15
2	A013	Banda Aceh	12	41
3	A040	Tanjung Balai	15	10

KESIMPULAN :
 Dari hasil proses clustering dapat diambil kesimpulan bahwa:
 1. Area Permintaan Cukup Besar, dengan jumlah 31 Area.
 2. Area Permintaan Besar, dengan jumlah 7 Area.
 3. Area Permintaan Kecil, dengan jumlah 2 Area.

Gambar 1.4 Tampilan Proses K-Means

5. Tampilan Form Laporan

Laporan ini berfungsi untuk menampilkan hasil dari pengelompokan data permintaan barang yang telah dihitung melalui form K-Means. Adapun tampilan dari form laporan adalah sebagai berikut :

No	Area	Jlh Toko	Jlh Item	Keterangan
1	Aceh	2	15	Pemintaan Besar
2	Banda Aceh	22	14	Pemintaan Kecil
3	Kota Medan	250	85	Pemintaan Sedang
4	Serdang Bedagai	51	24	Pemintaan Kecil
5	Tanjung Balai	15	10	Pemintaan Besar
6	Aceh Barat	8	45	Pemintaan Besar
7	Aceh Besar	7	41	Pemintaan Besar
8	Aceh Jaya	1	11	Pemintaan Besar

Gambar 1.6 Laporan Hasil Pengelompokan Data Permintaan Barang

5. KESIMPULAN

Berdasarkan hasil uraian analisa dan pengujian perancangan pada bab sebelumnya, kesimpulan yang dapat diambil adalah sebagai berikut :

- Berdasarkan analisis masalah yang terjadi dalam klusterisasi area mana yang memiliki permintaan produk yang besar dapat dianalisis menggunakan sebuah ilmu pengetahuan yaitu berupa penerapan *data mining* dalam membantu pemecahan masalah yang terjadi pada penelitian ini dan berdasarkan hasil penelitian, dalam penerapan *data mining* yang mengadopsi metode *K-Means Clustering* dapat digunakan dalam penyelesaian masalah PT Indomarco Prismaatama Cab.Medan.
- Berdasarkan hasil analisa, penerapan *K-Means Clustering* diawali dengan menentukan titik pusat awal centroid yang kemudian digunakan untuk mencari jarak terdekat dari setiap masing-masing data yang selanjutnya digunakan dalam mencari iterasi dan *cluster*, yang dihitung dengan konsep sebuah iterasi secara berulang untuk mendapatkan sebuah nilai rasio dimana akan dihentikannya perhitungan iterasi ketika nilai rasio tidak lagi lebih besar dari nilai rasio sebelumnya. Maka dilakukan pembentukan kesimpulan pengetahuan baru bagi perusahaan dalam bentuk tabel dengan cara mengurutkan data dari area permintaan paling besar sampai dengan area yang paling kecil dan dalam penelitian ini iterasi dihentikan pada iterasi ke tiga.

3. Berdasarkan hasil implementasi sistem dalam penerapan *data mining* menggunakan metode *K-Means Clustering*, metode ini mampu dalam pengimplementasiannya kedalam aplikasi dan dapat berjalan dengan baik dan mampu juga menyimpulkan hasil dari penelitian dalam menentukan area mana yang memiliki permintaan dari paling besar hingga paling kecil.




UCAPAN TERIMA KASIH

Mengucapkan rasa terima kasih yang sebesar-besarnya kepada pihak-pihak yang telah membantu menyelesaikan artikel ilmiah ini, khususnya kepada Asyahri Hadi Nasyuha, S.Kom., M.Kom selaku Dosen Pembimbing I yang telah banyak memberikan bimbingan serta arahan, Ibu Fifin Sonata, S.Kom., M.Kom selaku Dosen Pembimbing II yang senantiasa membimbing dan memberi masukan, kepada kedua orang tua tercinta yang selalu mendoakan dan memberikan dukungan penuh, seluruh teman-teman kelas 8SIC1 yang telah mengisi hari-hari menjadi menyenangkan.

REFERENSI

- [1] M. Hasanah, S. Defit, and G. W. Nurcahyo, "Implementasi Algoritma K-Means untuk Klasterisasi Peserta Olimpiade Sains Nasional Tingkat SMA," *J. Sistim Inf. dan Teknol.*, vol. 1, no. 3, pp. 31–36, 2019.
- [2] E. Irfiani and S. S. Rani, "Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita," *J. Sist. dan Teknol. Inf.*, vol. 6, no. 4, p. 161, 2018.
- [3] Y. Darmi and A. Setiawan, "Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk," vol. 12, no. 2, pp. 148–157, 2016.
- [4] F. M. Firman Nurdiansyah, Samsul Arifin, "Clustering algorithm untuk pengelompokan pelanggan dalam bidang usaha server reload," *Semin. Nas. Sist. Inf.*, pp. 1043–1047, 2018.
- [5] L. Bruno, "Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2019.
- [6] S. Pai, F. Di Troia, C. A. Visaggio, T. H. Austin, and M. Stamp, "Clustering for malware classification," *J. Comput. Virol. Hacking Tech.*, vol. 13, no. 2, pp. 95–107, 2017.

BIOGRAFI PENULIS

	<p>Annisa Fitri, wanita kelahiran Tanjung Morawa, 6 Maret 1997 ini merupakan mahasiswa STMIK Triguna Dharma jurusan Sistem Informasi. Saat ini tengah melakukan riset ilmiah yang berkaitan tentang bidang keilmuan <i>data mining</i> yang bertujuan untuk memenuhi salah satu syarat mendapatkan gelar strata 1 jurusan Sistem Informasi.</p>
	<p>Asyahri Hadi Nasyuha, S.Kom., M.Kom., merupakan dosen tetap STMIK Triguna Dharma yang aktif sebagai pengajar di STMIK Triguna Dharma.</p>
	<p>Fifin Sonata, S.Kom., M.Kom., merupakan dosen STMIK Triguna Dharma yang aktif sebagai pengajar di STMIK Triguna Dharma.</p>