

Penerapan *Data Mining* Dalam *Clustering* Pencapaian Target Penjualan Menggunakan Algoritma *K-Means*

Suci Nurani¹, Yohanni Syahra², Ahmad Calam³

^{1,2,3} Sistem Informasi, STMIK Triguna Dharma

Email: ¹sucinurani79576@gmail.com, ²yohanni.syahra@gmail.com, ³calamahmad72@gmail.com

Email Penulis Korespondensi: ¹sucinurani79576@gmail.com

Abstrak

PT. Batam Cahaya Sukses adalah salah satu perusahaan yang menjual berbagai jenis perumahan subsidi. Adapun yang menimbulkan beberapa permasalahan salah satunya seperti pemilik perusahaan kebingungan dalam proses penambahan stok barang dan terjadinya penumpukan stok perumahan subsidi yang kurang diminati pembeli, sedangkan stok untuk perumahan subsidi yang sangat diminati dibeli oleh pembeli justru sering tidak tersedia. Permasalahan tersebut perusahaan membutuhkan bidang keilmuan *data mining* dengan menggunakan metode *K-Means Clustering*. Dengan menggunakan metode ini, data-data yang telah didapatkan dapat dikelompokkan kedalam beberapa *cluster* berdasarkan kemiripan dari data-data tersebut, sehingga data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster* dan yang memiliki karakteristik yang berbeda dikelompokkan dalam *cluster* yang lain yang memiliki karakteristik yang sama. Hasil penelitian ini dapat membantu PT. Batam Cahaya Sukses dalam hal klusterisasi penjualan perumahan yang paling diminati, cukup diminati dan kurang diminati.

Kata Kunci: Data Mining, *K Means*, Perumahan Subsidi, *Clustering*, Penjualan

Abstract

PT. Batam Cahaya Sukses is one of the companies that sells various types of subsidized housing. As for the causes of several problems, one of them is the company owner's confusion in the process of adding stock and the buildup of subsidized housing stocks that are less attractive to buyers, while stocks for subsidized housing that buyers highly value are often not available. The problem is that companies need the scientific field of data mining using the *K-Means Clustering* method. By using this method, the data that has been obtained can be grouped into several clusters based on the similarity of the data, so that data that has the same characteristics are grouped in one cluster and those with different characteristics are grouped in other clusters that have different characteristics. The same characteristics. The results of this study can help PT. Batam Cahaya Success in terms of housing sales clustering that is most desirable, quite desirable and less desirable.

Keywords: Data Mining, *K Means*, Subsidized Housing, *Clustering*, Sales

1. PENDAHULUAN .

Penjualan merupakan salah satu kegiatan yang dilakukan perusahaan untuk mempertahankan bisnisnya untuk berkembang dan untuk mendapatkan laba atau keuntungan yang diinginkan. Kegiatan penjualan merupakan kegiatan pelengkap atau suplemen dari pembelian, untuk memungkinkan terjadinya transaksi. Jadi kegiatan pembelian dan penjualan merupakan satu kesatuan untuk dapat terlaksananya transfer hak atau transaksi. Oleh karena itu, kegiatan penjualan seperti halnya kegiatan pembelian, terdiri dari serangkaian kegiatan yang meliputi penciptaan permintaan, menemukan si pembeli, negosiasi harga, dan syarat-syarat pembayaran. dalam hal ini, penjualan ini, seperti penjual harus menentukan kebijaksanaan dan prosedur yang akan diikuti memungkinkan dilaksanakannya rencana penjualan yang ditetapkan. Subsidi KPR oleh pemerintah merupakan upaya untuk meningkatkan kepemilikan rumah terhadap masyarakat berpenghasilan rendah, di tengah tingginya harga rumah. Undang-Undang Nomor 4 Tahun 1992 Tentang Perumahan dan Permukiman sudah tidak sesuai dengan perkembangan dan kebutuhan perumahan dan permukiman yang layak dan terjangkau dalam lingkungan yang sehat, aman, serasi, dan teratur sehingga diganti dengan Undang-Undang Nomor 1 Tahun 2011 Tentang Perumahan dan Kawasan Permukiman (Jawa Pos, 3 Januari 2012). Teknologi yang semakin pesat pada era globalisasi ini merupakan aspek yang dapat dimanfaatkan untuk mencapai kemudahan dalam kehidupan sehari-hari [1].

Adapun yang menimbulkan beberapa permasalahan salah satunya seperti pemilik perusahaan kebingungan dalam proses penambahan stok barang dan terjadinya penumpukan stok perumahan subsidi yang kurang diminati pembeli, sedangkan stok untuk perumahan subsidi yang sangat diminati dibeli oleh pembelijustru sering tidak tersedia. Oleh sebab itu, permasalahan itu dapat diatasi dengan adalah penerapan *Data Mining*. *Data Mining* adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar. *Data Mining* juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. Pada penelitian ini analisa data mining dilakukan dengan menggunakan metode *K-Means*. Dengan menggunakan metode ini, data-data yang telah didapatkan dapat dikelompokkan kedalam beberapa *cluster* berdasarkan kemiripan dari data-data tersebut dan dikelompokkan dalam *cluster* yang lain yang memiliki karakteristik yang sama.

Data Mining adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar. *Data Mining* juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. Pada penelitian ini analisa data mining dilakukan dengan menggunakan metode *K-Means*. Dengan menggunakan metode ini, data-data yang telah

didapatkan dapat dikelompokkan kedalam beberapa *cluster* berdasarkan kemiripan dari data-data tersebut dan dikelompokkan dalam cluster yang lain yang memiliki karakteristik yang sama [2].

Tujuan dilakukannya penelitian ini adalah untuk klasterisasi penjualan perumahan subsidi yang sangat diminati dan kurang diminati sehingga dapat menjadi informasi dalam penyediaan stok barang yang tepat sesuai dengan tingkat penjualan. Penyediaan stok yang tepat juga dapat mengurangi penumpukan barang yang kurang diminati. Dalam penelitian ini akan dirancang sebuah perangkat lunak berbasis *Desktop Programming*. Aplikasi *desktop* yang digunakan adalah *Visual Basic*. *Visual basic* adalah sebuah bahasa pemrograman yang berpusat pada *object* (*Object Oriented Programming*) digunakan dalam pembuatan aplikasi *windows* yang berbasis *Graphical User Interface* [3].

Data mining berkembang menjadi alat bantu untuk mencari pola-pola yang berharga dalam suatu basis data yang sangat besar jumlahnya [4], sehingga tidak memungkinkan dicari secara manual. Beberapa teknik data mining dapat diklasifikasikan ke dalam kategori berikut, meliputi klasifikasi, *Clustering*, penggalian kaidah asosiasi, analisa pola sekuensial, prediksi, visualisasi data dan lain sebagainya. Teknik *Clustering* adalah teknik yang digunakan untuk menangani data yang besar dengan banyak atribut ke dalam sejumlah kelompok kecil. *Clustering* dilakukan dengan terlebih dahulu menganalisis bagian kecil dari data untuk menentukan klaster. Dalam data mining banyak metode yang sesuai untuk digunakan dalam menganalisa suatu pola penjualan salah satunya adalah algoritma *K-Means* [5]. Algoritma *K-Means* merupakan algoritma pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah *K cluster* yang sudah ditetapkan di awal. Algoritma *K-Means* sederhana untuk diimplementasikan dan dijalankan, relatif cepat, mudah beradaptasi, umum Penggunaannya dalam praktek. *K-Means* dapat diterapkan pada data yang direpresentasikan dalam *r-dimensi ruang tempat* [6].

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan penelitian dalam kasus klasterisasi perumahan subsidi terdapat beberapa bagian penting, yaitu sebagai berikut ini:

- a. Teknik Pengumpulan Data (*Data Collecting*)
Data Collecting adalah suatu teknik pengumpulan data yang digunakan untuk mencari informasi yang dibutuhkan dalam penelitian.
 1. Pengamatan Langsung (*Observasi*)
 2. Wawancara (*Interview*)
- b. Studi Kepustakaan (*Study of Literature*)
- c. Penerapan Metode *K-Means* dalam pengolahan data.

2.2 Klasterisasi

Klasterisasi adalah metode klasterisasi penjualan data dimana satu set objek data dikelompokkan ke dalam himpunan bagian yang disebut dengan cluster. Objek yang didalam cluster memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan cluster yang lain [7]. Dalam penelitian ini klasterisasi penjualan perumahan subsidi akan dikelompokkan menjadi tiga bagian-bagian yaitu paling diminati, cukup diminati dan kurang diminati dan perusahaan bisa dengan bijak mengatur stok barang dengan tepat sesuai dengan tingkat minat pembeli.

2.3 Data Mining

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terikat dari berbagai *database* besar. Berdasarkan pengertian data mining yang telah dijelaskan di atas, maka data mining merupakan pengetahuan yang tersembunyi di dalam *database* yang di proses untuk menemukan pola dan teknik statistik matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi pengetahuan dari *database* tersebut [8].

Knowledge Discovery In Database (KDD) adalah metode yang digunakan untuk mencari pengetahuan atau informasi yang belum diketahui dari sebuah database. KDD merupakan nama lain dari Data Mining walaupun sesungguhnya kedua istilah tersebut memiliki konsep yang tidak sama, namun berkaitan satu sama lain, dan salah satu dari tahapan proses keseluruhan KDD merupakan *Data Mining* yang menjadi inti dari proses KDD. Data Mining merupakan salah satu teknik untuk menemukan, mencari, atau menggali informasi atau pengetahuan baru dari sekumpulan data yang sangat besar, dengan integrasi atau penggabungan dengan disiplin ilmu lain seperti statistika, kecerdasan buatan, serta *machine learning*, menjadikan *Data Mining* sebagai salah satu alat bantu untuk menganalisa data yang kemudian menghasilkan informasi yang berguna [9].

2.4 K-Means Clustering

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 kedalam kelompok yang lain [10]:

K-Means ialah algoritma yang digunakan kedalam suatu pengelompokan secara membagi yang memisahkan data ke dalam kelompok yang berbeda – beda. Algoritma ini mampu memperpendek jarak antara data ke *cluster* nya. Pada dasar penggunaan algoritma ini tergantung dalam proses *clustering* pada data yang dihasilkan dan konklusi yang ingin dicapai di akhir proses. Sehingga dalam penggunaan algoritma *k-means* memiliki beberapa aturan sebagai berikut:

1. Berapa banyak jumlah *cluster* yang perlukan
2. Hanya mempunyai atribut bertipe *numeric*

Algoritma *k-means* hanya mengambil beberapa sampel dari seluruh populasi komponen yang didapatkan agar kemudian dijadikan pusat *cluster* awal, pada penentuan pusat *cluster* ini dipilih dengan cara acak dari populasi data. Kemudian algoritma *k-means* akan menguji setiap komponennya ke dalam jumlah populasi data tersebut dan menandai komponen-komponen tersebut ke dalam salah satu pusat *cluster* yang telah dideskripsikan sebelumnya tergantung dari jarak minimum antar komponen dengan setiap pusat *cluster* yang ada. Selanjutnya posisi pusat *clusternya* akan dihitung kembali sampai semua komponen data dikelompokkan ke dalam setiap *cluster* dan terakhir akan terbentuk *cluster* baru [11].

3. HASIL DAN PEMBAHASAN

3.1 Penerapan Metode K-Means

Penerapan Metode *K-Means* merupakan langkah penyelesaian terkait klasterisasi perumahan subsidi secara berurutan sesuai dengan referensi yang telah digunakan:

3.1.1 Persiapan Data

Berikut ini merupakan data sampel penjualan yang nantinya akan dilakukan proses perhitungan:

Tabel 1. Data Penjualan

No	Nama Sales Promotion	Booking	Penjualan
1	Anton	330	315
2	Jordan Ananda Sinambela	58	53
3	Jenni Nadya	48	45
4	Yenni Olan Br Napitupulu	202	202
5	Novitasari Panjaitan	172	166
6	Putra Pratama Nasution	90	81
7	Eduard	102	92
8	Jandri Ginting	81	81
9	Muhammad Maulana Subhan	63	57
10	Miranda Talita Tesalonika	95	81
11	Johanis Surbakti	64	58
12	Lagu Parningotan L	151	146
13	Sadari Ginting Manik	170	162
14	Dame Laurenta Simamora	133	128
15	Benta Nika Br Sebayang Amd	140	132
16	Ardiansyah Putra Bukit	330	325
17	Jaren Ahgriva	121	115
18	Enrfano Simanungkalit	40	36
19	Rediana Marsauli Napitupulu	26	25
20	Hamdan Irwansyah Siregar	170	156
21	Dameria Sagala	31	22
22	Kiki Esterlita Br Bangun	20	19
23	Josua Alexander Sinambela	151	140
24	Rodes Purba	30	27
25	Reza Fahlevi	22	20
26	Junianti Simanjuntak	90	88
27	Binsar M Hutapea	70	62
28	Matius Satriawan Surbakti	25	20
29	Adiel Ginting	31	29
30	Nani Sri Wahyuni Br Tambunan	21	18
31	Sri Yunita Anovia Br Bangun	25	21

No	Nama Sales Promotion	Booking	Penjualan
32	Sri Defi Sinaga	20	18
33	Elfrida Siregar	52	50
34	Yorena Mianta Marpaung	92	81
35	Elvi Susana Br Sembiring	80	77
36	M Riduan S	36	34
37	Jonathan Pardamean Gabe Tua Simarmata	43	41
38	Maria Magdalena Saragih	80	73
39	Saut Maruli Tua Sinaga	35	36
40	Sarwedi Hotma Tua Saragih	55	52
41	Rickhard Clifffen Purba	36	31
42	Chendra Martua Panjaitan	30	25
43	Dzaki Irfaan Rahmadan	20	19
44	Fransiscus da diva	35	34
45	Friska liana	460	402
46	Hosea oktavianus	65	61
47	Henri paldo	50	42
48	Heru sinulinnga	45	39
49	Iwanta sahputra bukit	62	55
50	Icha wandita br giting	57	44
51	Jesica Anjani	35	32
52	Liska sari br giting	45	37
53	Nindy natasya br baru	173	160
54	Nelsi br purba	30	24
55	Revaldo tarigan	45	41
56	Risa ninta br sitepu	160	156
57	Rima br tarigan	31	27
58	Silvani br keliat	40	35
59	Oska gurusinga	33	28
60	Saskia laura	40	35
61	Paula ninta	43	41
62	Pitra maulana	50	47
63	Widya mayangsari	58	51
64	Wira kesuma	44	40
65	Nurul Silvani	40	35
66	Putri Wahyuni	34	27
67	Ragil Suhardina	56	48
68	Rama Eka Satria	40	39
69	Ramadani	35	33
70	Rizka Amelia Azis	22	20
71	Silvy Dwi Tamara	36	34
72	Siswanto	177	177
73	Sri Agustina Br Tarigan	152	140
74	Tasya Awalia Sitepu	153	150
75	Tri Wulandari	201	201
76	Yunita Sari Sembiring	35	32
77	Zahilul Muhibah	40	35
78	Afnida Aulia Husna Sigalingging	72	61

3.1.1 Proses Perhitungan Metode K-Means

Berikut ini merupakan proses perhitungan Metode *K-Means*:

$$d(x,y) = \|x-y\| = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}; i = 1, 2, 3 \dots n$$

Penerapan jumlah *cluster* (K) yaitu 3 *cluster*, Setelah menetapkan jumlah *cluster*, Tentukan titik pusat awal *cluster* (*Centroid*), Berikut ini titik *Centroid* yang telah dipilih:

Tabel 2. Data Centroid

Titik Pusat Awal	No Data	Nama Sales Promotion	Pemesanan	Penjualan
Cluster 1	45	Friska liana	460	402
Cluster 2	34	Yorena Mianta Marpaung	92	81
Cluster 3	32	Sri Defi Sinaga	20	18

Inisialisasi ke pusat cluster dengan menghitung jarak data ke Centroid menggunakan rumus *Euclidean*, data tersebut dari Cluster terdekatnya.

- a. Menghitung data pertama dengan titik cluster 1

$$C1 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$C1 = \sqrt{(330-460)^2 + (315-402)^2}$$

$$C1 = 156,426$$

- b. Menghitung data pertama dengan titik cluster 2

$$C2 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$C2 = \sqrt{(330-92)^2 + (315-81)^2}$$

$$C2 = 333,766$$

- c. Menghitung data pertama dengan titik cluster 3

$$C3 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$C3 = \sqrt{(330-20)^2 + (315-18)^2}$$

$$C3 = 429,312$$

Hasil Perhitungan Iterasi I

Lakukan proses perhitungan yang sama dengan objek ke 40, Adapun hasil dari perhitungan literasi 1 dapat dilihat pada table dibawah ini. Dimana jarak terdekat dilihat dari perhitungan yang paling dekat ke pusat cluster. Sementara WCV (*Within Cluster Variatio*) adalah keadaan dari hasil pangkat dari perhitungan jarak terdekat ke pusat cluster.

Tabel 3. Tabel literasi 1

No	Kode Sales	C1	C2	C3	Jarak Terdekat	Berada Di Cluster
1	S001	156,426	333,766	429,312	156,426	1
2	S002	532,358	44,045	51,662	44,045	2
3	S003	545,154	56,851	38,897	38,897	3
4	S004	326,441	163,527	258,805	163,527	2
5	S005	372,344	116,726	212,151	116,726	2
6	S006	489,838	2,000	94,175	2,000	2
7	S007	473,565	14,866	110,454	14,866	2
8	S008	496,671	11,000	87,693	11,000	2
9	S009	525,960	37,643	58,052	37,643	2
10	S010	486,072	3,000	97,949	3,000	2
11	S011	524,549	36,235	59,464	36,235	2
12	S012	401,269	87,784	183,153	87,784	2
13	S013	376,431	112,450	207,933	112,450	2
14	S014	426,620	62,370	157,699	62,370	2
15	S015	418,688	70,036	165,517	70,036	2
16	S016	151,093	340,852	436,290	151,093	1
17	S017	444,173	44,688	140,036	44,688	2
18	S018	557,096	68,768	26,907	26,907	3
19	S019	574,878	86,556	9,220	9,220	3

Tabel 3. Tabel literasi 1 (Lanjutan)

No	Kode Sales	C1	C2	C3	Jarak Terdekat	Berada Di Cluster
20	S020	380,284	108,208	203,823	108,208	2
21	S021	573,098	84,865	11,705	11,705	3
22	S022	583,343	95,016	1,000	1,000	3
23	S023	405,123	83,439	179,011	83,439	2
24	S024	570,548	82,219	13,454	13,454	3
25	S025	581,178	92,849	2,828	2,828	3
26	S026	485,279	7,280	98,995	7,280	2
27	S027	517,397	29,069	66,603	29,069	2
28	S028	578,921	90,609	5,385	5,385	3
29	S029	568,480	80,156	15,556	15,556	3
30	S030	583,247	94,921	1,000	1,000	3
31	S031	578,261	89,939	5,831	5,831	3
32	S032	584,000	95,671	0,000	0,000	3
33	S033	538,858	50,606	45,255	45,255	3
34	S034	488,329	0,000	95,671	0,000	2
35	S035	500,025	12,649	84,149	12,649	2
36	S036	561,427	73,110	22,627	22,627	3
37	S037	551,552	63,253	32,527	32,527	3
38	S038	502,634	14,422	81,394	14,422	2
39	S039	560,875	72,622	23,431	23,431	3
40	S040	535,280	47,011	48,795	47,011	2
41	S041	563,398	75,073	20,616	20,616	3
42	S042	571,864	83,546	12,207	12,207	3
43	S043	583,343	95,016	1,000	1,000	3
44	S044	562,182	73,878	21,932	21,932	3
45	S045	0,000	488,329	584,000	0,000	1
46	S046	521,829	33,601	62,241	33,601	2
47	S047	545,619	57,315	38,419	38,419	3
48	S048	551,357	63,032	32,650	32,650	3
49	S049	528,027	39,699	55,973	39,699	2
50	S050	539,048	50,931	45,222	45,222	3
51	S051	563,494	75,166	20,518	20,518	3
52	S052	552,675	64,382	31,401	31,401	3
53	S053	375,410	113,146	208,741	113,146	2
54	S054	572,524	84,220	11,662	11,662	3
55	S055	550,042	61,717	33,971	33,971	3
56	S056	387,964	101,237	196,581	101,237	2
57	S057	569,795	81,468	14,213	14,213	3
58	S058	557,754	69,426	26,249	26,249	3
59	S059	567,631	79,310	16,401	16,401	3
60	S060	557,754	69,426	26,249	26,249	3
61	S061	551,552	63,253	32,527	32,527	3
62	S062	542,333	54,037	41,725	41,725	3
63	S063	533,671	45,343	50,329	45,343	2
64	S064	551,453	63,127	32,558	32,558	3
65	S065	557,754	69,426	26,249	26,249	3
66	S066	567,539	79,246	16,643	16,643	3
67	S067	537,152	48,836	46,861	46,861	3
68	S068	555,130	66,843	29,000	29,000	3
69	S069	562,837	74,518	21,213	21,213	3
70	S070	581,178	92,849	2,828	2,828	3
71	S071	561,427	73,110	22,627	22,627	3
72	S072	361,544	128,222	223,450	128,222	2

Tabel 3. Tabel literasi 1 (Lanjutan)

No	Kode Sales	C1	C2	C3	Jarak Terdekat	Berada Di Cluster
73	S073	404,361	84,149	179,744	84,149	2
74	S074	397,181	92,098	187,385	92,098	2
75	S075	327,844	162,114	257,391	162,114	2
76	S076	563,494	75,166	20,518	20,518	3
77	S077	557,754	69,426	26,249	26,249	3
78	S078	516,551	28,284	67,476	28,284	2

C1= {S001,S016 dan S045}

C2={ S002, S004, S005, S006, S007, S008, S009, S010, S011, S012, S013, S014, S015, S017, S020, S023, S026, S027, S034, S035, S038, S040, S046, S049, S053, S056, S063, S072, S073, S074, S075, S078}

C3={ S002, S004, S005, S006, S007, S008, S009, S010, S011, S012, S013, S014, S015, S017, S020, S023, S026, S027, S034, S035, S038, S040, S046, S049, S053, S056, S063, S072, S073, S074, S075, S078}

1. Hitung nilai WCV (*Within Cluster Variation*) dengan cara memangkatkan jarak terdekat *cluster* dan menjumlahkan setiap nilai WCV.

$$WCV = 156,426 + 44,045 + 38,897 + 163,527 + \dots + 28,284$$

$$WCV = 2671,189$$

2. Hitung nilai BCV (*Between Cluster Variation*) dengan cara menjumlahkan hasil dari jarak diantara setiap *centroid*.

$$\begin{aligned} \text{a. } D(m1,m2) &= \sqrt{(m1-m2)^2} \\ &= \sqrt{(460-92)^2+(402-81)^2} \\ &= 488,329 \end{aligned}$$

$$\begin{aligned} \text{b. } D(m1,m3) &= \sqrt{(m1-m3)^2} \\ &= \sqrt{(460-20)^2+(402-18)^2} \\ &= 95,671 \end{aligned}$$

$$\begin{aligned} \text{c. } D(m2-m3) &= \sqrt{(m2-m3)^2} \\ &= \sqrt{(92-20)^2+(81-18)^2} \\ &= 584,000 \end{aligned}$$

$$\begin{aligned} \text{Nilai BCV} &= d(m1,m2) + d(m1,m3)+d(m2,m3) \\ &= 488,329 + 95,671 + 584,000 \\ &= 1168,000 \end{aligned}$$

3. Menghitung nilai besar rasio dengan membandingkan nilai BCV dan WCV.

$$BCV/WCV = 1168,000 / 2671,189$$

$$= 0,437$$

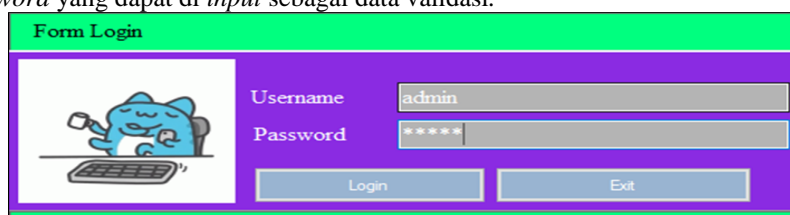
Selanjutnya, dilakukan perhitungan dengan rumus yang sama hingga iterasi selanjutnya.

3.2 Implementasi Sistem

Berikut ini merupakan hasil implementasi sistem yang telah dibangun dengan berbasis *Desktop* menggunakan *Microsoft Visual Studio 2008* dan *database Microsoft Acces*.

1. *Form Login*

Form login berfungsi sebagai validasi akses dari admin untuk masuk kedalam sistem, pada *form login* terdapat *username* dan *password* yang dapat di *input* sebagai data validasi.



Gambar 1. Tampilan *Form Login*

2. *Form Menu Utama*

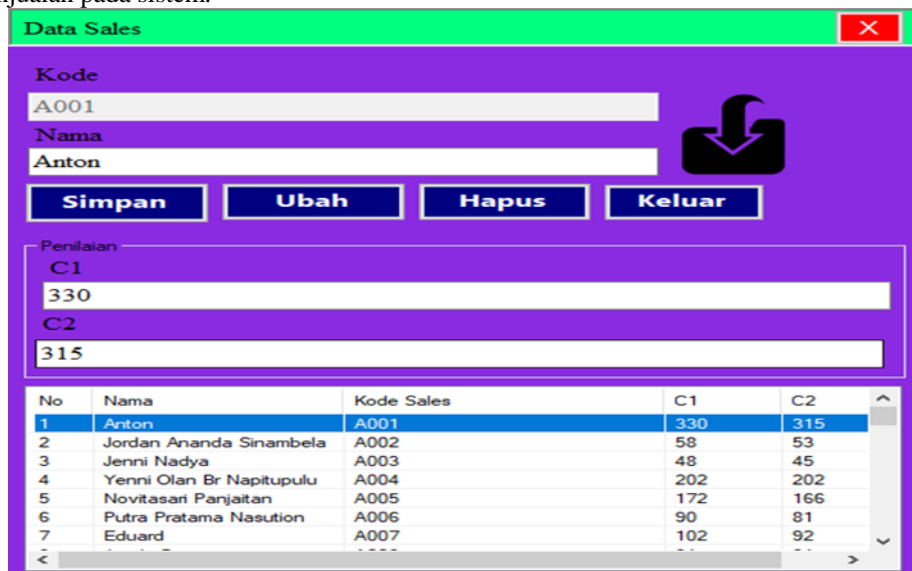
Form Menu Utama berfungsi sebagai halaman navigasi untuk membuka menu-menu yang lainnya.



Gambar 2. Tampilan *Form Menu Utama*

3. *Form Data Penjualan*

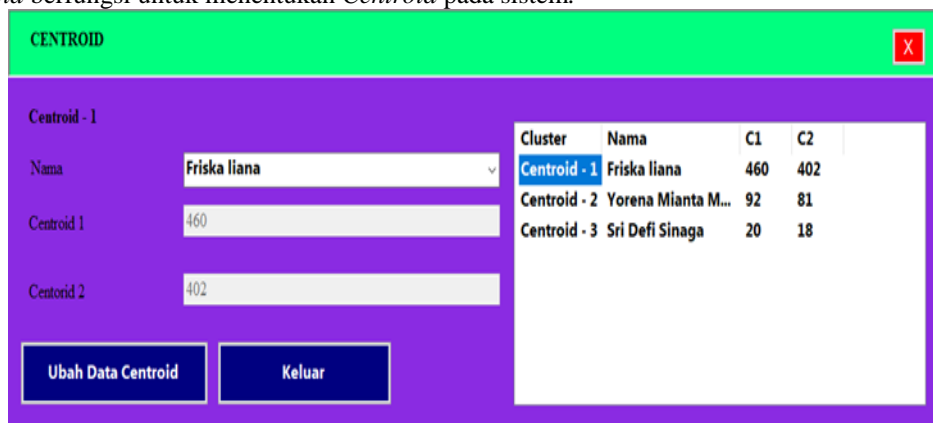
Form Data Penjualan berfungsi untuk mengelola data Penjualan seperti menampilkan, menyimpan, menghapus dan mengubah Penjualan pada sistem.



Gambar 3. Tampilan *Form Data Penjualan*

4. *Form Centroid*

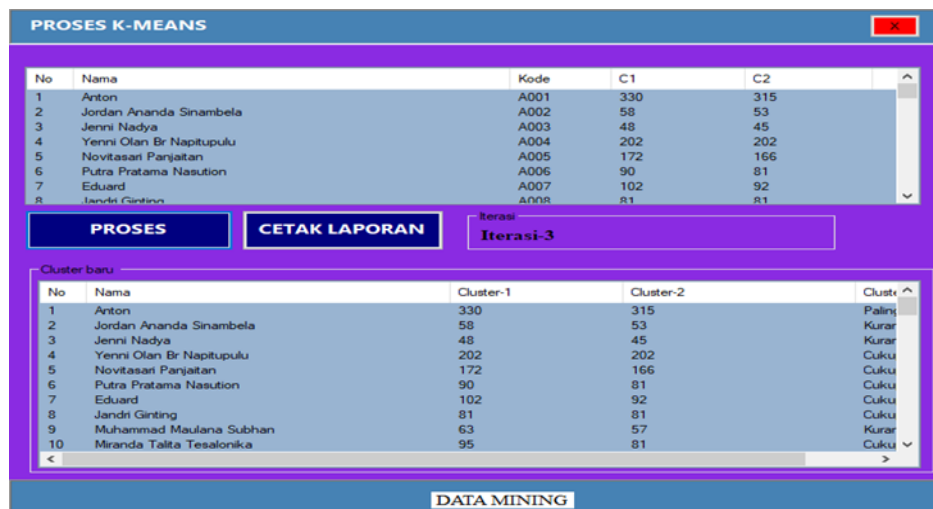
Form Centroid berfungsi untuk menentukan *Centroid* pada sistem.



Gambar 4. Tampilan *Form Centroid*

5. *Form Proses K-Means*

Form Proses K-Means berfungsi untuk melakukan proses *Clustering* data menggunakan algoritma *K-Means Clustering*.



Gambar 5. Tampilan *Form Proses*

4. KESIMPULAN

Dalam mengamati pengelompokkan dilakukan dengan riset ke tempat perusahaan dan wawancara pihak pengolahan data 78 data target penjualan sales yang dikelompokkan dalam segementasi penjualan. Dalam menguji aplikasi sistem penjualan menghasilkan akurasi metode yang digunakan sangat tepat berdasarkan jumlah data digunakan 78 dan menampilkan hasil pengelompokkan data penjualan.

UCAPAN TERIMAKASIH

Puji dan Syukur dipanjatkan kepada Tuhan Yang Maha Esa yang memberikan karunia sehingga mampu menyelesaikan jurnal ini. Kemudian kepada Ibu Yohanni Syahra dan Bapak Ahmad Calam atas segala waktu dan ilmunya yang telah memberikan bimbingan selama masa pengerjaan hingga menyelesaikan jurnal ini dan kepada seluruh dosen serta pegawai kampus STMIK Triguna Dharma yang telah banyak membantu baik dari segi informasi ataupun dukungan lainnya.

DAFTAR PUSTAKA

[1] T. Pradiani, "Pengaruh Sistem Pemasaran Digital Marketing Terhadap Peningkatan Volume Penjualan Hasil Industri Rumahan," *J. Ilm. Bisnis dan Ekon. Asia*, vol. 11, no. 2, pp. 46–53, 2018, doi: 10.32812/jibeka.v11i2.45.

[2] Y. Syahra, "Penerapan Data Mining Dalam Pengelompokkan Data Nilai Siswa Untuk Penentuan Jurusan Siswa Pada SMA Tamora Menggunakan Algoritma K-Means Clustering," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 17, no. 2, p. 228, 2018, doi: 10.53513/jis.v17i2.70.

[3] A. Fitri, Y. Syahra, and R. Kustini, "Penerapan Data Mining Dalam Mengklusterisasi Location Best Pb Tambahan Pada Regional IV PT Indomarco Prismatama Cab.Medan Dengan Menggunakan Metode K-Means," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 19, no. 2, p. 11, 2020, doi: 10.53513/jis.v19i2.2330.

[4] M. Ramadhan, J. Hutagalung, M. Dahria, I. Zulkarnain, and H. Jaya, "Prediksi Penjualan Spare Part Mobil Daihatsu Menggunakan Algoritma Apriori," *Techno.Com*, vol. 22, no. 1, pp. 156–166, 2023, doi: 10.33633/tc.v22i1.7192.

[5] A. Azanuddin and A. Calam, "Data Mining Untuk Memprediksi Penjualan Buah Sawit Menggunakan Metode Multi Regresi Pada PT.Karya Hevea Indonesia," no. x, 2020.

[6] A. F. Siregar, "Peminatan Menu Makanan Dengan Algoritma Analytical," vol. 9, no. 1, pp. 52–57, 2021.

[7] S. D. Prasetiani and N. Rochmawati, "Penerapan Data Mining Untuk Clustering Menu Favorit Menggunakan Algoritma K-Means (Studi Kasus Kedai Expo)," vol. 03, pp. 278–286, 2022.

[8] D. P. Utomo, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," vol. 4, no. April, pp. 437–444, 2020, doi: 10.30865/mib.v4i2.2080.

[9] R. Aditya, S. Defit, and G. W. Nurcahyo, "Prediksi Tingkat Ketersediaan Stock Sembako Menggunakan Algoritma FP-Growth dalam Meningkatkan Penjualan," *J. Inform. Ekon. Bisnis*, vol. 2, pp. 67–73, 2020, doi: 10.37034/infv2i3.44.

[10] F. Yunita, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru," *Sistemasi*, vol. 7, no. 3, p. 238, 2018, doi: 10.32520/stmsi.v7i3.388.

[11] M. Syukri Mustafa, M. Rizky Ramadhan, and A. P. Thenata, "Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Citec J.*, vol. 4, no. 2, pp. 151–162, 2017.