

Penjualan Sales Sepeda Motor Di FIF Group Menggunakan Metode Clustering

Rudi Kurniawan¹, Trinanda Syahputra², Moch. Iswan Perangin-angin³

^{1,2}Program Studi Sistem Informasi, STMIK Triguna Dharma

³Program Studi Manajemen Informatika, STMIK Triguna Dharma

Email: ¹rudi0542@gmail.com, ²trinandasyahputra@gmail.com, ³mochammadiswan@gmail.com

Email Penulis Korespondensi: rudi0542@gmail.com

Abstrak

FIF GROUP adalah salah satu perusahaan yang menjual berbagai jenis sepeda motor seperti yamaha, Suzuki, honda dan lain-lain. Adapun yang menimbulkan beberapa permasalahan salah satunya seperti pemilik perusahaan kebingungan dalam proses penambahan stok barang dan terjadinya penumpukan stok sepeda motor yang kurang diminati pembeli, sedangkan stok untuk sepeda motor 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 FIF GROUP dalam hal klasterisasi penjualan sepeda motor yang paling diminati, cukup diminati dan kurang diminati.

Kata Kunci: Data Mining, K-Means, Penjualan

1. PENDAHULUAN

FIF GROUP adalah salah satu perusahaan yang menjual berbagai jenis sepeda motor seperti yamaha, Suzuki, Honda dan lain-lain. Sepeda motor sebagai alat transportasi minimalis yang memiliki fungsi lebih handal dan bebas hambatan dari pada kendaraan beroda empat, akhir - akhir ini diminati oleh pengguna jalan raya khususnya di kota - kota besar yang intensitas kemacetannya tinggi, sehingga mau tidak mau bagi para teknisi kendaraan roda dua ini dituntut untuk bekerja lebih cepat dan akurat. Klasterisasi sepeda motor adalah klasterisasi penjualan sepeda motor yang diambil dari satu sumber data yang dikelompokkan menjadi bagian-bagian yang berbeda. Dalam penelitian ini klasterisasi penjualan sepeda motor 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.

Adapun yang menimbulkan beberapa permasalahan salah satunya seperti pemilik perusahaan kebingungan dalam proses penambahan stok barang dan terjadinya penumpukan stok sepeda motor yang kurang diminati pembeli, sedangkan stok untuk sepeda motor yang sangat diminati dibeli oleh pembeli justru 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 [1]. *Data Mining* adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam *database* [2]. Proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis [3].

Tujuan dilakukannya penelitian ini adalah untuk klasterisasi penjualan sepeda motor 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* dan menerapkan metode *K-Means Clustering*.

Penelitian yang sudah menerapkan *K-Means Clustering* diantaranya digunakan untuk mengklusterisasi lokasi terbaik untuk permintaan barang tambahan [4]. Mengklusterisasi minat nasabah asuransi [5]. Pengelompokan Prestasi Siswa [6]. Dengan menerapkan *K-Means Clustering*, metode ini sangat tepat dalam melakukan pengelompokan jenis data karena akan lebih mudah dalam implementasi dan penerapannya dapat memfilter dengan cepat dan tepat.

2. METODOLOGI PENELITIAN

2.3 Metode K-Means

K-Means merupakan suatu metode paling sederhana dalam pengelompokan dokumen yang awalnya mengambil sebagian dari banyaknya komponen data untuk dijadikan sebagai pusat *cluster* [7]. Prinsip utama *K-Means* adalah menyusun *prototype* atau pusat massa (*centroid*) dari sekumpulan data berdimensi n [8]. Metode *K-Means* dikenal sebagai suatu metode pengelompokan yang paling terkenal diberbagai bidang karena sangat mudah diimplementasi dan mampu klasterisasi data dalam jumlah yang banyak [9]. Algoritma *K-Means* juga sangat akurat dalam pengelompokan data [10]. Tahapan dalam perhitungan algoritme *K-Means* yaitu tentukan banyaknya *cluster* berdasarkan *cluster* data yang akan dibentuk. Tentukan *centroid* atau pusat *cluster* awal secara acak. Kemudian hitung

setiap data dan dikelompokkan ke *cluster* terdekat berdasarkan jarak antara data dengan pusat *cluster* awal. Hitung kembali *centroid* dengan keanggotaan *cluster* yang sekarang. Selanjutnya pada setiap objek memakai pusat *cluster* yang baru. Jika pusat *cluster* tidak berubah lagi maka proses perhitungan selesai [11]. Langkah-langkah melakukan *clustering* dengan metode *K-means* adalah sebagai berikut [12]:

1. Pilih jumlah *cluster* k .
2. Inisialisasi ke pusat *cluster* ini bisa dilakukan dengan berbagai cara. Cara yang paling sering dilakukan adalah dengan *random* atau acak. Pusat-pusat *cluster* diberi dengan nilai awal dengan angka-angka *random*.

$$d(x_i, \mu_j) = \sqrt{\sum (x_i, \mu_j)^2}$$

Dimana:

x_i = data kriteria

μ_j = *centroid* pada *cluster* ke- j s

3. Klasterisasi penjualan setiap data berdasarkan kedekatannya dengan *centroid* atau mencari jarak terkecil.
4. Memperbaharui nilai *centroid* baru, nilai *centroid* baru di peroleh dari rata-rata *cluster* yang bersangkutan dengan menggunakan rumus yaitu :

$$\mu_j^{(t+1)} = \frac{1}{N_{sj}} \sum_{j \in s_j} x_j$$

Keterangan :

$\mu_j^{(t+1)}$ = *centroid* baru pada iterasi $(t+1)$ N_{sj} = Data pada *cluster* S_j

5. Apabila data setiap *cluster* belum berhenti, lakukan perulangan dari langkah 2 hingga 5, sampai anggota tiap *cluster* tidak ada yang berubah.

3. HASIL DAN PEMBAHASAN

3.1 Penerapan Metode *K-Means*

Berikut ini adalah deskripsi data yang diuji setelah dilakukan pengumpulan data melalui wawancara dengan sales gudang FIF GROUP pada September 2020 -September 2021.

Tabel 1. Penjualan September 2020 -September 2021

No	Nama Sales	Pemesanan	Realisasi
1	Adi Pranata Brahmana	330	315
2	Ana Linda Br Ginting	58	53
3	Anisa Br Sembiring	48	45
4	Amri Surbakti	202	202
5	Amika Putri Br Sitepu	172	166
6	Amira Kartika	90	81
7	Apriansyah Putra	102	92
8	Billy Sembiring	81	81
9	Boby Siregar	63	57
10	Bunga Rafelia Br Barus	95	81
11	Elisabet Silalahi	64	58
12	Evan Saputra	151	146
13	Handika P. Tarigan	170	162
14	Handika P. Tarigan	133	128
15	Joel Sion Sembiring	140	132
16	Josua Ginting	330	325
17	Jessica Wati	121	115
18	Kalia Pepayosa Br Ginting	40	36
19	Kristian Saragih	26	25
20	Mita Wulandari	170	156
21	Mailida Nehe	31	22
22	Masita Br Bangun	20	19
23	Maya sartika	151	140
24	Meylinda sari	30	27
25	Rendi pemalemsa sembiring	22	20
26	Reza diansyah	90	88
27	Rere nika br keliat	70	62

28	Rido manta kemit	25	20
29	Rifaldo baru	31	29
30	Revi chania sembiring	21	18
31	Wawan baslius keliat	25	21
32	Sakira venita	20	18
33	Dani marvelanda surbakti	52	50
34	Ananta jaya giting	92	81
35	Angelita clodia	80	77
36	Anike br pandia	36	34
37	Anisa lila br ketaren	43	41
38	Arya giting	80	73
39	Beny andika bukit	35	36
40	Boby persadanta baru	55	52
41	Cristian p. Sitepu	36	31
42	Cristy anastasya br pinem	30	25
43	Cindya br sembiring	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 Menentukan Inisialisasi Jumlah Cluster Dan Titik Awal Cluster

Sebelum melakukan perhitungan terlebih dahulu harus menentukan jumlah *cluster* yang diinginkan. Dalam penelitian ini data-data yang ada akan dikelompokkan menjadi tiga *cluster*. Selanjutnya adalah menentukan titik pusat awal dari setiap *cluster*. Dalam penelitian ini titik pusat *cluster* yang dipilih dapat dilihat pada tabel 3.2 sebagai berikut:

Tabel 2. Contoh Data Titik Pusat *Cluster* pada iterasi-1

Titik Pusat Awal	No Data	Nama Sales	Pemesanan	Penjualan
<i>Cluster 1</i>	45	Friska liana	460	402
<i>Cluster 2</i>	34	Sakira venita	92	81
<i>Cluster 3</i>	32	Ananta jaya giting	20	18

3.1.2 Inisialisasi Ke Pusat *Cluster*

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*

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

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

$$C3 = 429,312$$

3.1.3 Klasterisasi penjualan Setiap Data Centoroid

Berikut tabel dibawah ini dapat dilihat untuk hasil lengkap dari perhitungan pada tahap iterasi pertama:

Tabel 3. Hasil Perhitungan Iterasi ke-1

No	Kode Sales	C1	C2	C3	Jarak Terdekat	Berada Di Cluster
1	A001	156,426	333,766	429,312	156,426	1
2	A002	532,358	44,045	51,662	44,045	2
3	A003	545,154	56,851	38,897	38,897	3
4	A004	326,441	163,527	258,805	163,527	2
5	A005	372,344	116,726	212,151	116,726	2
6	A006	489,838	2,000	94,175	2,000	2
7	A007	473,565	14,866	110,454	14,866	2
8	A008	496,671	11,000	87,693	11,000	2
9	A009	525,960	37,643	58,052	37,643	2
10	A010	486,072	3,000	97,949	3,000	2
11	A011	524,549	36,235	59,464	36,235	2
12	A012	401,269	87,784	183,153	87,784	2
13	A013	376,431	112,450	207,933	112,450	2
14	A014	426,620	62,370	157,699	62,370	2
15	A015	418,688	70,036	165,517	70,036	2
16	A016	151,093	340,852	436,290	151,093	1
17	A017	444,173	44,688	140,036	44,688	2
18	A018	557,096	68,768	26,907	26,907	3
19	A019	574,878	86,556	9,220	9,220	3
20	A020	380,284	108,208	203,823	108,208	2
21	A021	573,098	84,865	11,705	11,705	3
22	A022	583,343	95,016	1,000	1,000	3
23	A023	405,123	83,439	179,011	83,439	2
24	A024	570,548	82,219	13,454	13,454	3
25	A025	581,178	92,849	2,828	2,828	3
26	A026	485,279	7,280	98,995	7,280	2
27	A027	517,397	29,069	66,603	29,069	2

28	A028	578,921	90,609	5,385	5,385	3
29	A029	568,480	80,156	15,556	15,556	3
30	A030	583,247	94,921	1,000	1,000	3
31	A031	578,261	89,939	5,831	5,831	3
32	A032	584,000	95,671	0,000	0,000	3
33	A033	538,858	50,606	45,255	45,255	3
34	A034	488,329	0,000	95,671	0,000	2
35	A035	500,025	12,649	84,149	12,649	2
36	A036	561,427	73,110	22,627	22,627	3
37	A037	551,552	63,253	32,527	32,527	3
38	A038	502,634	14,422	81,394	14,422	2
39	A039	560,875	72,622	23,431	23,431	3
40	A040	535,280	47,011	48,795	47,011	2
41	A041	563,398	75,073	20,616	20,616	3
42	A042	571,864	83,546	12,207	12,207	3
43	A043	583,343	95,016	1,000	1,000	3
44	A044	562,182	73,878	21,932	21,932	3
45	A045	0,000	488,329	584,000	0,000	1
46	A046	521,829	33,601	62,241	33,601	2
47	A047	545,619	57,315	38,419	38,419	3
48	A048	551,357	63,032	32,650	32,650	3
49	A049	528,027	39,699	55,973	39,699	2
50	A050	539,048	50,931	45,222	45,222	3
51	A051	563,494	75,166	20,518	20,518	3
52	A052	552,675	64,382	31,401	31,401	3
53	A053	375,410	113,146	208,741	113,146	2
54	A054	572,524	84,220	11,662	11,662	3
55	A055	550,042	61,717	33,971	33,971	3
56	A056	387,964	101,237	196,581	101,237	2
57	A057	569,795	81,468	14,213	14,213	3
58	A058	557,754	69,426	26,249	26,249	3
59	A059	567,631	79,310	16,401	16,401	3
60	A060	557,754	69,426	26,249	26,249	3
61	A061	551,552	63,253	32,527	32,527	3
62	A062	542,333	54,037	41,725	41,725	3
63	A063	533,671	45,343	50,329	45,343	2
64	A064	551,453	63,127	32,558	32,558	3
65	A065	557,754	69,426	26,249	26,249	3
66	A066	567,539	79,246	16,643	16,643	3
67	A067	537,152	48,836	46,861	46,861	3
68	A068	555,130	66,843	29,000	29,000	3
69	A069	562,837	74,518	21,213	21,213	3
70	A070	581,178	92,849	2,828	2,828	3
71	A071	561,427	73,110	22,627	22,627	3
72	A072	361,544	128,222	223,450	128,222	2
73	A073	404,361	84,149	179,744	84,149	2
74	A074	397,181	92,098	187,385	92,098	2
75	A075	327,844	162,114	257,391	162,114	2
76	A076	563,494	75,166	20,518	20,518	3
77	A077	557,754	69,426	26,249	26,249	3
78	A078	516,551	28,284	67,476	28,284	2

Dari tabel 3.4 diatas dapat hasil iterasi ke-1 sebagai berikut:

- C1= {A001,A016 dan A045}
- C2={ A002, A004, A005, A006, A007, A008, A009, A010, A011, A012, A013, A014, A015, A017, A020, A023, A026, A027, A034, A035, A038, A040, A046, A049, A053, A056, A063, A072, A073, A074, A075, A078}

- C3={ A002, A004, A005, A006, A007, A008, A009, A010, A011, A012, A013, A014, A015, A017, A020, A023, A026, A027, A034, A035, A038, A040, A046, A049, A053, A056, A063, A072, A073, A074, A075, A078}

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

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

$$WCV = 3171,189$$

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

$$a. D(m1,m2) = \frac{\sqrt{(m1-m2)^2}}{\sqrt{(460-92)^2+(402-81)^2}}$$

$$= \frac{488,329}{\dots}$$

$$b. D(m1,m3) = \frac{\sqrt{(m1-m3)^2}}{\sqrt{(460-20)^2+(402-18)^2}}$$

$$= \frac{95,671}{\dots}$$

$$c. D(m2-m3) = \frac{\sqrt{(m2-m3)^2}}{\sqrt{(92-20)^2+(81-18)^2}}$$

$$= \frac{584,000}{\dots}$$

$$\text{Nilai BCV} = d(m1,m2) + d(m1,m3)+d(m2,m3)$$

$$= 488,329 + 95,671 + 584,000$$

$$= 1168,000$$

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

$$BCV/WCV = 1168,000 / 3171,189$$

$$= 0,368$$

3.1.4 Memperbarui Nilai Centroid Baru

Kemudian langkah berikutnya pada iterasi ke-2 adalah Menghitung kembali pusat-pusat baru dari *cluster* yang baru terbentuk dengan menghitung nilai rata-rata pada *cluster*.

1. Pada *cluster* 1 terdapat 3 data maka nilai rata-rata pada *cluster-1* adalah
 $s = (330 + 330 + 460)/3 = 373,333$
 $t = (315 + 325 + 402)/3 = 347,333$
2. Pada *cluster* 2 terdapat 9 data maka nilai rata-rata pada *cluster-2* adalah
 $s = (58 + 202 + 172 + 90 + 102 + 81 + 63 + 95 + 64 + 151 + 170 + 133 + 140 + 121 + 170 + 151 + 90 + 70 + 92 + 80 + 80 + 55 + 65 + 62 + 173 + 160 + 58 + 177 + 152 + 153 + 201 + 72) / 32 = 115,718$
 $t = (53 + 202 + 166 + 81 + 92 + 81 + 57 + 81 + 58 + 146 + 162 + 128 + 132 + 115 + 156 + 140 + 88 + 62 + 81 + 77 + 73 + 52 + 61 + 55 + 160 + 156 + 51 + 177 + 140 + 150 + 201 + 61) / 32 = 109,218$
3. Pada *cluster* 3 terdapat 67 data maka nilai rata-rata pada *cluster-3* adalah
 $s = (48 + 40 + 26 + 31 + 20 + 30 + 22 + 25 + 31 + 21 + 25 + 20 + 52 + 36 + 43 + 35 + 36 + 30 + 20 + 35 + 50 + 45 + 57 + 35 + 45 + 30 + 45 + 31 + 40 + 33 + 40 + 43 + 50 + 44 + 40 + 34 + 56 + 40 + 35 + 22 + 36 + 35 + 40) / 43 = 36,093$
 $t = (45 + 36 + 25 + 22 + 19 + 27 + 20 + 20 + 29 + 18 + 21 + 18 + 50 + 34 + 41 + 36 + 31 + 25 + 19 + 34 + 42 + 39 + 44 + 32 + 37 + 24 + 41 + 27 + 35 + 28 + 35 + 41 + 47 + 40 + 35 + 27 + 48 + 39 + 33 + 20 + 34 + 32 + 35) / 43 = 32,209$

Tabel 4. Iterasi ke-2

TITIK PUSAT AWAL	PENJUALAN	PENJUALAN
<i>Cluster 1</i>	373,333	347,333
<i>Cluster 2</i>	175,333	170,000
<i>Cluster 3</i>	36,093	32,209

3.2 Implementasi Sistem

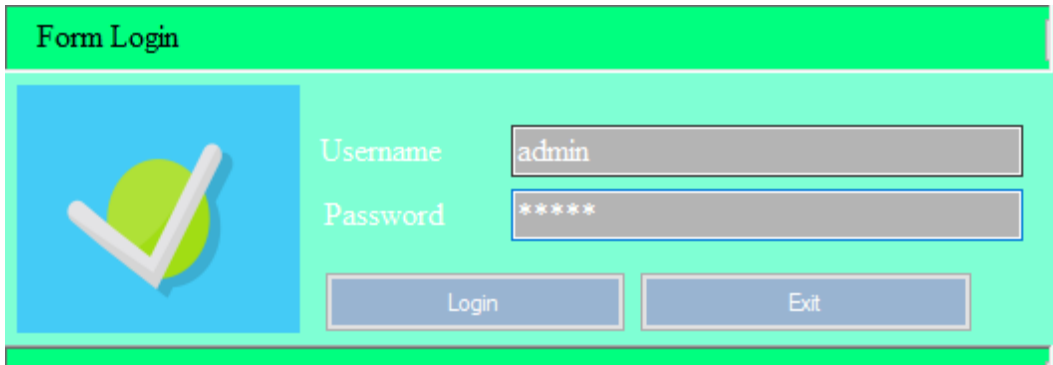
Fungsi dari *interface* (antarmuka) ini adalah untuk memberikan *input* dan menampilkan *output* dari aplikasi. Pada aplikasi ini memiliki *interface* yang terdiri dari *Menu login*, *Data Sales*, *Data Centroid* dan *Menu Proses K-Means*.

3.2.1 Halaman Utama

Dalam halaman utama untuk menampilkan pada tampilan *Menu* pada awal sistem yaitu *Menu login* dan *Menu Utama*. Adapun *Menu* halaman utama sebagai berikut.

1. *Menu Login*

Menu Login digunakan untuk mengamankan sistem dari *user-user* yang tidak bertanggung jawab sebelum masuk ke *Menu Utama*. Berikut adalah tampilan *Menu Login* :



Gambar 1. *Menu Login*

2. *Menu Utama*

Menu Utama digunakan sebagai penghubung untuk *Menu Data Sales*, data *Centroid*, proses dan laporan. Berikut adalah tampilan *Menu Utama*.




Gambar 2. *Menu Utama*

3.2.2 Halaman Administrator

Administrator untuk menampilkan *Menu* pengolahan data pada penyimpanan data kedalam *database* yaitu *Menu Data Sales*, dan *Menu Centroid*. Adapun *Menu* halaman *administrator* utama sebagai berikut.

1. *Menu Data Sales*

Menu data Sales berfungsi untuk pengolahan dalam penginputan data, ubah data dan penghapusan data Sales. Adapun *Menu* data nasbah adalah sebagai berikut.

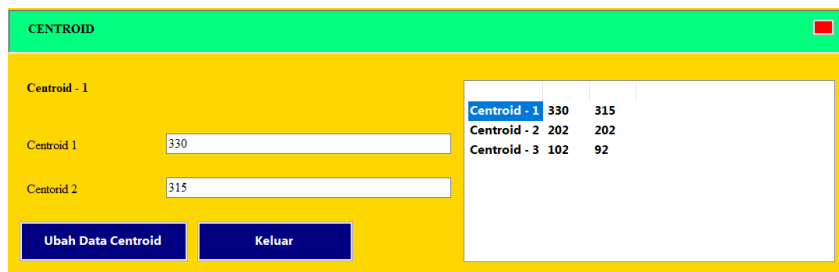


No	Nama Sales	Kode Sales	C1
1	Anisa Br Sembiring	1	48
2	Anni Subakti	2	202
3	Amka Putri Br Sitepu	3	172
4	Amira Kartika	4	90
5	Aptansyah Putra	5	102
6	Billy Sembiring	6	81
7	Boby Siregar	7	63

Gambar 3. *Menu Data Sales*

2. *Menu Data Centroid*

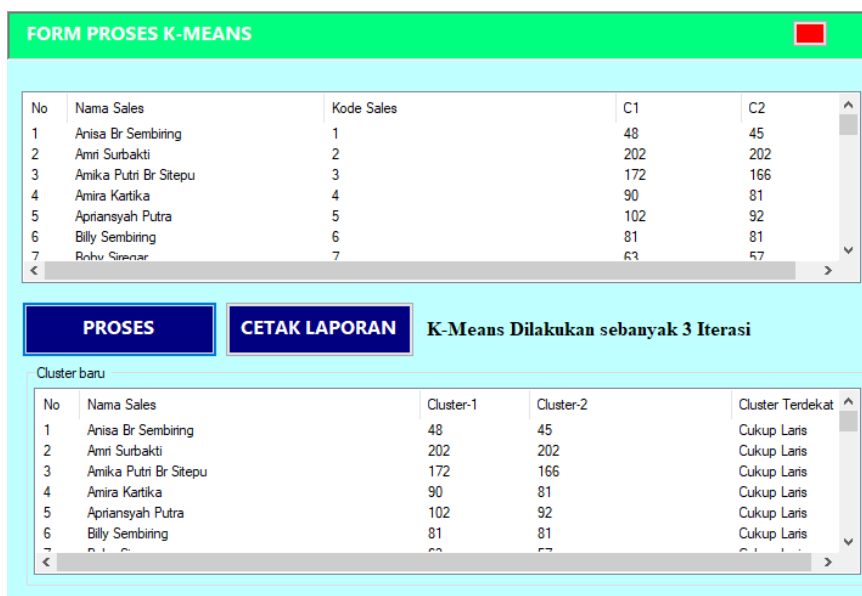
Menu Data Centroid untuk pengolahan data *Centroid* pemilihan produk asuransi. Adapun *Menu Data Centroid* adalah sebagai berikut.



Gambar 4. *Menu Data Centroid*

3.3 Pengujian

Pada bagian ini diminta untuk melakukan pengujian dengan *sampling* data baru untuk dapat menguji keakuratan sistem yang dirancang dengan *tools-tools* yang sudah teruji dan terkalibrasi sebelumnya. Adapun hasil proses program dalam penjualan bahan pokok sebagai berikut.



Gambar 5. Hasil Analisa *K-Means*

4. KESIMPULAN

Berdasarkan hasil analisa dari permasalahan yang terjadi dengan kasus yang dibahas tentang mengelompokkan data perusahaan dengan menerapkan metode terhadap *k-means* sistem yang dirancang dan dibangun maka dapat ditarik kesimpulan sebagai berikut: Dalam menganalisa pengelompokkan segementasi distributor kosmetik dilakukan dengan riset ke tempat perusahaan dan wawancara pihak pengolahan data 78 data sales yang dikelompokkan dalam segementasi penjualan. Dalam merancang ataupun membangun sistem menggunakan bahasa pemodelan *Unified Modeling Language* (UML) yang terdiri dari *use case diagram*, *class diagram*, *activity diagram*. Pembangunan sistem dengan bahasa pemograman *visual basic*. Dalam akurasi metode yang digunakan sangat tepat berdasarkan jumlah data digunakan 78 dan menampilkan hasil pengelompokkan data penjualan.

UCAPAN TERIMAKASIH

Terima Kasih diucapkan kepada Bapak Trinanda Syahputra dan Moch. Iswan Perangin-angin yang memberi motivasi, Doa dan dukungan moral maupun materi, serta pihak-pihak yang telah mendukung dalam proses penyelesaian penelitian yang tidak dapat disebutkan satu persatu. Kiranya bisa memberi manfaat bagi pembaca dan dapat meningkatkan kualitas penelitian selanjutnya.

DAFTAR PUSTAKA

- [1] Juniar Hutagalung, Y. H. Syahputra, “Pemetaan Siswa Kelas Unggulan Menggunakan Algoritma K-Means Clustering,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 1, pp. 606–620, 2022, doi: 10.35957/jatisi.v9i1.1516.
- [2] P. Purwadi, P. S. Ramadhan, and N. Safitri, “Penerapan Data Mining Untuk Mengestimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada BPS Deli Serdang,” *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 18, no. 1, p. 55, 2019, doi: 10.53513/jis.v18i1.104.
- [3] D. Nofriansyah, M. Yetri, K. Erwansyah, and S. Suharsil, “Penerapan Data Mining Dalam Menganalisa Data Penjualan Untuk Mendapatkan Pola Rekomendasi Penjualan Menggunakan Algoritma Apriori Pada K3 Mart,” *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 18, no. 2, p. 176, 2019, doi: 10.53513/jis.v18i2.157.
- [4] 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.
- [5] J. Hutagalung and F. Sonata, “Penerapan Metode K-Means Untuk Menganalisis Minat Nasabah Asuransi,” *J. MEDIA Inform. BUDIDARMA*, vol. 5, no. 3, pp. 1187–1194, 2021, doi: 10.30865/mib.v5i3.3113.
- [6] S. Natalia, B. Sembiring, H. Winata, and S. Kusnasari, “Pengelompokan Prestasi Siswa Menggunakan Algoritma K-Means,” *J. Sist. Inf. TGD*, vol. 1, pp. 31–40, 2022.
- [7] S. U. Tarigan and M. Yetri, “Klasterisasi Data Penanganan Dan Pelayanan Kesehatan Masyarakat,” *J. Sist. Inf. TGD*, vol. 1, no. 3, pp. 193–206, 2022.
- [8] S. Laia, D. H. Pane, and E. Affandi, “Implementasi Metode K-Means Untuk Mengelompokkan Kawasan Potensi Pertanian Karet Produktif,” *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 1, no. 4, p. 282, 2022, doi: 10.53513/jursi.v1i4.5224.
- [9] M. H. Mhd. Gilang Suryanata, Deski Helsa Pane, “Implementasi Algoritma K-Means Untuk Mengukur Tingkat Kepuasan Siswa Terhadap Proses Pembelajaran,” *J. Teknol. Sist. Inf. dan Sist. Komput. TGD*, vol. 2, no. 2, pp. 118–125, 2019.
- [10] D. Nofriansyah and I. Mariami, “Implementasi Data Mining Untuk Pengelempokan Buku Di Perpustakaan Yayasan Nurul Islam Indonesia Baru Dengan Metode K-Means Clustering,” *J. CyberTech*, vol. 1, no. 1, pp. 1–12, 2021.
- [11] J. Hutagalung, N. L. W. S. R. Ginantra, G. W. Bhawika, W. G. S. Parwita, A. Wanto, and P. D. Panjaitan, “COVID-19 Cases and Deaths in Southeast Asia Clustering using K-Means Algorithm,” *J. Phys. Conf. Ser.*, vol. 1783, no. 1, pp. 1–6, 2021, doi: 10.1088/1742-6596/1783/1/012027.
- [12] T. Syahputra, J. Halim, and E. P. Sintho, “Penerapan Data Mining Dalam Menentukan Pilihan Jurusan Bidang Studi SMA Menggunakan Metode,” *Penerapan Data Min. dalam Menentukan Pilihan*