

Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)

Baginda Harahap
profesionalbaginda@gmail.com

Abstract

Building materials have become a necessity that must be met in development, and one of the businesses that is currently growing rapidly is the property business, so that the community's need for shelter is a business opportunity sought by the community. Therefore, UD. Indarung YD Building Stores in determining the grouping of building materials is in demand, in demand and in demand. Software development and support in data processing in decision making, the authors conducted research to apply Data Mining K-Means which is classified as very in demand, in demand and in demand.

Keywords: Clustering, K-Means, Building Materials Very in demand, in demand and in demand, RapidMiner Studio.

Abstrak

Bahan bangunan sudah menjadi kebutuhan yang harus dipenuhi dalam pembangunan, serta salah satu bisnis yang saat ini berkembang pesat yaitu bisnis properti, sehingga kebutuhan masyarakat akan tempat tinggal menjadi peluang usaha yang dicari oleh masyarakat. Oleh sebab itu, UD. Toko Bangunan YD Indarung dalam menentukan pengelompokan bahan bangunan sangat laris, laris dan kurang laris. Pengembangan perangkat lunak dan pendukung dalam pengolahan data dalam pengambilan keputusan, penulis melakukan penelitian untuk menerapkan Data Mining K-Means yang tergolong bahan bangunan sangat laris, laris dan kurang laris.

Keywords: Clustering, K-Means, Bahan Bangunan Sangat Laris, Laris dan Kurang Laris, RapidMiner Studio.

1. Pendahuluan

Bahan bangunan sudah menjadi kebutuhan yang harus dipenuhi dalam pembangunan, serta salah satu bisnis yang saat ini berkembang pesat yaitu bisnis properti, sehingga kebutuhan masyarakat akan tempat tinggal menjadi peluang usaha yang dicari oleh masyarakat. Oleh sebab itu, UD. Toko Bangunan YD Indarung dalam menentukan pengelompokan bahan bangunan sangat laris, laris dan kurang laris. Pengembangan perangkat lunak dan pendukung dalam pengolahan data dalam pengambilan keputusan, penulis melakukan penelitian untuk menerapkan Data Mining K-Means yang tergolong bahan bangunan sangat laris, laris dan kurang laris.

Knowledge Discovery in Database (KDD) didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses *Knowledge Discovery in Database* melibatkan hasil proses data mining (proses pengekstrakan kecenderungan suatu pola data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami. *Knowledge Discovery in Database (KDD)* adalah proses menentukan informasi yang berguna serta pola-pola yang ada dalam data. Informasi ini terkandung dalam basis data yang berukuran besar yang sebelumnya tidak diketahui dan potensial bermanfaat. *Data Mining* merupakan salah satu langkah dari serangkaian proses *iterative KDD*.

K-Means Clustering adalah metode untuk mengkategorikan atau pengelompokan sekelompok objek sesuai dengan atribut yang sama atau karakteristik ke dalam sejumlah groups (jumlah bilangan bulat positif). Ini mendefinisikan sebuah *cluster* oleh massa yang mewakili mean dari *cluster*. Dalam data mining analisis klaster populer adalah *K-means*. Ini adalah metode kuantisasi vector. Hal ini sesuai dengan permasalahan yang saya temukan di lapangan yakni identifikasi bahan bangunan sangat laris, laris dan kurang laris, dimana metode data mining dengan algoritma *K-Means clustering* cocok untuk mengumpulkan dan mengelompokkan bahan bangunan tersebut menjadi beberapa kategori yaitu sangat laris, laris dan kurang laris.

Pada penelitian ini, di kemukakan suatu metode pengolahan data dengan mengelompokkan data menggunakan algoritma *K-Means clustering* dari hasil stok penjualan UD. Toko Bangunan YD Indarung. Hasil observasi dibandingkan dengan identifikasi menggunakan algoritma *K-Means clustering*, sehingga di harapkan untuk penstockan bias ditambah dan bias dikurangi.

Algoritma K-means, mungkin yang pertama dari algoritma pengelompokan yang diusulkan, didasarkan pada ide yang sangat sederhana: Diberi serangkaian kumpulan awal, menetapkan setiap titik ke salah satunya, lalu setiap pusat gugus diganti dengan titik rata-rata pada kluster masing-masing. Dua langkah sederhana ini diulang hingga konvergensi. Suatu titik ditugaskan ke cluster yang dekat dalam jarak Euclidean ke titik.

Algoritma K-Means adalah salah satu algoritma analisis cluster yang paling umum digunakan. Inefisiensi yang digunakan *K-Means* untuk menangani data yang lebih besar harus diperhatikan. Seiring dengan set data yang diproses menjadi lebih besar, pendekatan host berbasis CPU tunggal tidak berdaya. Dalam tulisan ini, penulis ingin mencapai algoritma *K-Means* universal paralel yang mampu menangani kumpulan data yang lebih besar. Hal ini diimplementasikan oleh CUDA. Penulis terutama memperhatikan fleksibilitas dan skalabilitas algoritma, implementasinya mungkin bukan metode yang paling efisien.

Data Mining, sering juga disebut *knowledge discovery in database* (KDD), adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari data mining bisa dipakai untuk memperbaiki pengambilan keputusan dimasa depan.

Data Mining merupakan proses iterative dan interaktif untuk menemukan pola atau model baru yang sempurna, bermanfaat dan dapat di mengerti dalam suatu database yang sangat besar.

Untuk menyelesaikan permasalahan adalah dengan pemanfaatan algoritma *K-Means Clustering*. Algoritma *K-Means* adalah *algoritma Clustering* yang paling sederhana dibanding *algoritma Clustering* yang lain. Algoritma ini mempunyai kelebihan mudah diterapkan dan dijalankan, relatif cepat, mudah untuk diadaptasi, dan paling banyak dipraktekkan dalam tugas *Data Mining*.

2. Metode Penelitian

Pada Metodologi penelitian ini dilakukan secara sistematis menjelaskan proses bagaimana kerangka kerja penelitian yang dilaksanakan. beberapa tahapan yang digambarkan dalam bentuk kerangka Gambar kemudian dapat di pahami mulai dari proses analisa kebutuhan sampai dengan hasil dari penelitian ini. Tahapan-tahapan dalam melakukan penelitian, yakni menganalisa kebutuhan data, proses pengumpulan data, analisa data menggunakan algoritma *K-Means*, kemudian dilakukan pengolahan data dengan RapidMiner, dan hasil dari penelitian ini. Berdasarkan kerangka kerja nantinya akan dijabarkan langkah-langkah pada penelitian ini. yang akan di jelaskan pada Gambar 1.

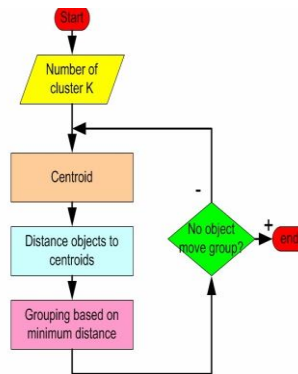
2.1. Kerangka Kerja Penelitian

Tahapan pada poin ini yakni dimana kerangka kerja berfungsi untuk menggambarkan proses yang sistematis dalam penelitian ini, dari tahap menganalisa kebutuhan data hingga menyimpulkan hasil. yang akan di jelaskan pada Gambar 1



Gambar1. Kerangka Kerja Penelitian

Steps of K-Means Processing



Gambar2. Langkah Proses *K-Means*

Pada Gambar 1 dan 2 merupakan kerangka kerja dan langkah proses *K-Means* dengan menggunakan metode *K-Means*. Adapun tahapan-tahapan proses dalam algoritma *clustering* menggunakan metode *K-Means* yakni sebagai berikut (Ong., 2013) :

- Memilih jumlah *cluster* k.
- Inisialisasi k pusat *cluster* pada umum nya banyak cara dalam proses ini, Tetapi pilihan utama pada tahapan *cluster* adalah dengan cara random.
- Alokasikan semua data / objek ke *cluster* paling dekat.

Untuk melakukan proses pengolahan data pada titik tiap titik pusat *cluster* yakni dengan teori jarak Euclidean yang dirumuskan sebagai berikut :

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

$D_{(i,j)}$ = Jarak dari 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

- Proses selanjut nya adalah menghitung ulang pusat *cluster* dengan keanggotaan *cluster* yang terbaru. rata-rata dari semua data/ objek dalam *cluster* merupakan pusat *cluster*. Sehingga mean bukanlah prioritas parameter yang di gunakan.
- Menugaskan kembali setiap objek dengan menggunakan pusat *cluster* baru, apakah pusat *cluster* berubah hingga proses *clustering* selesai, ulang kembali proses “c” sampai di temukan nilai pada pusat *cluster* tidak ada berubah.

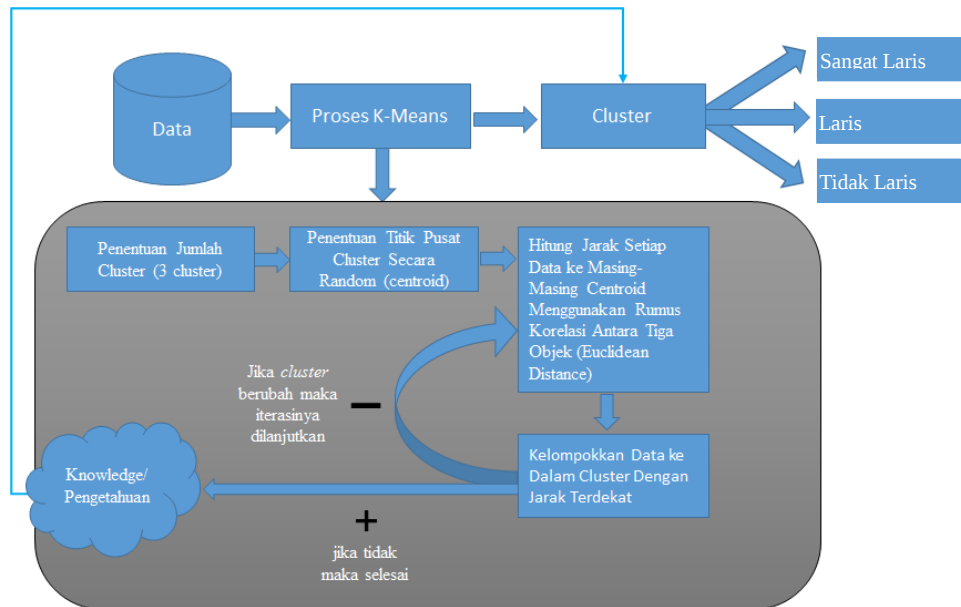
2.2. *K-Means Clustering*

K-Means clustering ialah metode yang termasuk pada *clustering* non-hirarki dimana setiap obyek yang masuk dalam kelompok adalah obyek-obyek yang sama dan berkorelasi. Data yang tergabung dalam kelompok mempunyai tingkat kemiripan yang lebih besar dan memiliki tingkat perbedaan yang besar pula dengan kelompok lainnya

Pada dasarnya *clustering* adalah metode untuk mengkategorikan atau pengelompokan sekelompok objek sesuai dengan atribut yang sama atau karakteristik dengan data-lainnya.

Clustering merupakan suatu metode pada data mining dimana proses kerja pada algoritma ini sifatnya tanpa arahan (unsupervised), artinya metode ini tidak lagi memerlukan lagi suatu training dan tanpa guru bahkan output tidak di perlukan. Pada data mining terdapat dua pembagian jenis metode *clustering* untuk proses pengelompokkan data, yakni *hierarchical clustering* dan *non-hierarchical clustering*.

Adapun langkah-langkah pengolahan data dengan metode *K-Means Cluster* terlihat pada Gambar 3 :



Gambar 3. Alur Kerja *K-Means Clustering*

Pada Gambar 3 merupakan proses algoritma *K-Means*, dimana tahap pertama adalah mengumpulkan data, seleksi data, data cleaning, dan data transformasi. Tahapan selanjutnya adalah proses pengolahan *K-Means* dengan menentukan jumlah *cluster* secara *random*, titik pusat *cluster*, menghitung jarak setiap data dan mengelompokkan data pada *cluster* sehingga menghasilkan *knowledge* yaitu *clustering*.

3. Hasil dan Pembahasan

Tabel 4.21. Nilai Centroid Baru Iterasi 9

Centroid	Tidak Terjual	Terjual
C0	21.125	62.54166667
C1	242.1666667	941.1666667
C2	53	222

Tabel 4.21 merupan nilai centroid pada iterasi kesembilan untuk menghitung pusat *cluster*, tahapan ini akan dihitung jarak setiap data kemasing-masing centroid menggunakan rumus Euclidean Distance (D) iterasi kesembilan. Menghitung jarak masing-masing data ke titik pusat *cluster* (C0).

- a. Pada stage (D1), yaitu (Platon minyak golden Yellow 803), yaitu A₁, memiliki jarak dengan titik pusat *cluster* (C0) adalah :

$$\begin{aligned}
 D_1 &= \sqrt{(A_{1x} - C0_x)^2 + (A_{1y} - C0_y)^2} \\
 &= \sqrt{(30 - 21.125)^2 + (200 - 62.5416)^2} \\
 &= 137.7445426
 \end{aligned}$$

- b. Pada stage (D2), yaitu (Thiner keiji 1 ltr), yaitu A2, memiliki jarak dengan pusat *cluster* (C0) adalah :

$$\begin{aligned}
 D_2 &= \sqrt{(A_{2x} - C0_x)^2 + (A_{2y} - C0_y)^2} \\
 &= \sqrt{(60 - 21.125)^2 + (280 - 62.5416)^2} \\
 &= 220.905845
 \end{aligned}$$

Untuk Dn seterusnya akan di tampilkan pada Tabel 4.22 pengelompokan data dalam *cluster* dengan jarak terdekat. Pada tahap ini akan di hitung jarak setiap data kemasing-masing centroid menggunakan rumus Euclidean Distance iterasi kesembilan. Menghitung jarak masing-masing data ke titik pusat *cluster* (C1).

- a. Pada stage (D1), yaitu (Platon Minyak Golden Yellow 803) yaitu A₁, memiliki jarak dengan titik pusat *cluster* (C₁) adalah :

$$\begin{aligned}
 D_1 &= \sqrt{(A_{1x} - C1_x)^2 + (A_{1y} - C1_y)^2} \\
 &= \sqrt{(30 - 242.1666)^2 + (200 - 941.1666)^2} \\
 &= 770.9362634
 \end{aligned}$$

- b. Pada stage (D2), yaitu (Thinner keiji 1 Ltr) yaitu A₂, memiliki jarak dengan titik pusat *cluster* (C₁) adalah :

$$\begin{aligned} D_2 &= \sqrt{(A_{2x} - C1_{1x})^2 + (A_{2y} - C1_y)^2} \\ &= \sqrt{(60 - 242.1666)^2 + (280 - 941.1666)^2} \\ &= 685.8032193 \end{aligned}$$

Untuk Dn seterusnya akan di tampilkan pada Tabel 4.22 pengelompokan data dalam *cluster* dengan jarak terdekat.

Pada tahapan ini akan di hitung jarak setiap data ke masing-masing centroid menggunakan rumus *Euclidean Distance* iterasi kesembilan. Menghitung jarak masing-masing data ke titik pusat *cluster* (C₂).

- a. Pada stage (D1), yaitu (Platon Minyak Golden 803) yaitu A₁, memiliki jarak dengan titik pusat *cluster* (C₂) adalah :

$$\begin{aligned} D_1 &= \sqrt{(A_{1x} - C2_x)^2 + (A_{1y} - C2_y)^2} \\ &= \sqrt{(30 - 53)^2 + (200 - 222)^2} \\ &= \mathbf{31.82766093} \end{aligned}$$

- b. Pada stage (D2), yaitu (Thinner Keiji 1 Ltr) yaitu A₂, memiliki jarak dengan titik pusat *cluster* (C₂) adalah :

$$\begin{aligned} D_2 &= \sqrt{(A_{2x} - C2_x)^2 + (A_{2y} - C2_y)^2} \\ &= \sqrt{(60 - 53)^2 + (280 - 222)^2} \\ &= \mathbf{58.42088668} \end{aligned}$$

Dari perhitungan *Euclidean Distance*, di dapat perbandingan jarak dari masing-masing data ke C₀, C₁ dan C₂, seperti pada Tabel 4.22 berikut ini :

Kode	Jarak ke C0	Jarak ke C1	Jarak ke C2	Jarak Terdekat ke C0	Jarak Terdekat ke C1	Jarak Terdekat ke C2
A1	137.7445426	770.9362634	31.82766093	0	0	1
A2	220.905845	685.8032193	58.42088668	0	0	1
A3	26.18318216	932.0336487	188.512599	1	0	0
A4	6.362051122	912.1411014	168.6712779	1	0	0
A5	158.4804689	754.6783347	51.89412298	0	0	1
A6	50.45848816	907.6449686	172.8380745	1	0	0
A7	37.47522329	870.011143	126.3843345	1	0	0
A8	26.18318216	932.0336487	188.512599	1	0	0
A9	28.85698692	877.2111427	133.9888055	1	0	0
A10	5.850273026	910.9252013	167.7050983	1	0	0
A11	11.9955142	894.4328122	151.2514463	1	0	0
A12	4.625187687	910.1959802	166.9041641	1	0	0
A13	26.80284241	932.7565182	189.2749323	1	0	0
A14	938.7657104	44.71452657	776.1784589	0	1	0
A15	827.8028503	215.9893259	673.1456009	0	1	0
A16	821.6739372	135.7573407	659.6188293	0	1	0
A17	220.1061691	690.0613902	61.84658438	0	0	1
A18	14.5679681	919.9021989	176.7399219	1	0	0
A19	5.935123797	911.8958579	168.4666139	1	0	0
A20	45.46675373	951.4270277	208.0793118	1	0	0
A21	33.81038836	938.1641944	194.5353438	1	0	0
A22	30.89324135	933.1377474	189.5600169	1	0	0
A23	31.31920116	935.2636289	191.637679	1	0	0
A24	287.1231891	622.9045317	126.0039682	0	0	1
A25	167.6933482	742.1406351	24.35159132	0	0	1
A26	47.29579644	862.8379853	122.0368797	1	0	0
A27	47.29579644	862.8379853	122.0368797	1	0	0
A28	12.59202239	918.4421896	175.1370892	1	0	0

A29	16.45880801	890.0820499	146.5366848	1	0	0
A30	877.6664471	82.16683572	716.8856255	0	1	0
A31	155.7489188	796.5170362	120.8014901	0	0	1
A32	758.8419197	147.3297511	596.4000335	0	1	0
A33	35.46912781	871.4375416	127.8162744	1	0	0
A34	17.16757684	922.346675	178.7652091	1	0	0
A35	138.9324501	767.1115014	24.0208243	0	0	1
A36	96.98827606	815.6118699	75.71657678	0	0	1
A37	111.3577973	803.3575723	67.2086304	0	0	1
A38	79.38970773	833.4466923	92.17917335	1	0	0
A39	1250.319716	361.302351	1087.976562	0	1	0
A40	10.39514444	916.2641844	172.7917822	1	0	0

Tabel 4.22. Pengelompokan Data *Cluster* dengan Jarak Terdekat

Dari Tabel 4.22 diatas dapat kita lihat 40 sampel data dengan hasil perhitungan jarak terdekat C0 adalah 24, C1 adalah 6, dan C2 adalah 10. maka didapatkanlah kelompok dari anggota C0, C1 dan C2 seperti dibawah ini :

- Anggota *cluster* 0 (C0) terdiri dari 24 anggota yang meliputi :
{A3, A4, A6, A7, A8, A9, A10, A11, A12, A13, A18, A19, A20, A21, A22, A23, A26, A27, A28, A29, A33, A34, A38, 140}
- Anggota *cluster* 1 (C1) terdiri dari 6 anggota yang meliputi :
{A14, A15, A16, A30, A32, A39}
- Anggota *cluster* 2 (C2) terdiri dari 10 anggota yang meliputi :
{A1, A2, A5, A17, A24, A25, A31, A35, A36, A37}

Kemudian lakukan iterasi kembali tentukan posisi *centroid* baru dengan cara menghitung rata-rata dari data-data yang berada pada *centroid* yang sama. Dibawah ini untuk perhitungan pusat *cluster* baru :

$$C0(X1) = \frac{(12+18+70+20+12+30+23+27+22+13+22+19+22+12+3+0+3+50+50+20+21+22+12+9+17)}{24}$$

$$C0(X1) = 21.125$$

$$C0(X2) = \frac{(38+57+50+100+38+90+57+73+58+37+48+57+18+34+40+37+100+100+50+79+98+48+141+53)}{24}$$

$$C0(X2) = 62.54166667$$

$$C1(X1) = \frac{(220+403+122+308+200+200)}{6}$$

$$C1(X1) = 242.1666667$$

$$C1(X2) = \frac{(980+797+878+892+800+1300)}{6}$$

$$C1(X2) = 941.1666667$$

$$C2(X1) = \frac{(30+60+100+38+52+30+150+52+11+7)}{10}$$

$$C2(X1) = 53$$

$$C2(X2) = \frac{(200+280+200+282+348+230+150+198+159+173)}{10}$$

$$C2(X2) = 222$$

Dari hasil pengelompokan di atas, dapat dilihat tidak terjadi lagi perubahan anggota untuk masing-masing *cluster*. Oleh sebab itu proses iterasi cukup sampai diiterasi 9, dan dihentikan sampai disini.

Maka dapat disimpulkan bahwa *Cluster* 0 (C₀) ada 24 items bahan bangunan, yang artinya ada 24 bahan bangunan yang termasuk kategori bahan bangunan kurang laris, pada cluster 1 (C₁) ada 6 items bahan bangunan, yang artinya 6 bahan bangunan yang termasuk kategori bahan bangunan sangat laris, dan pada

cluster 2 (C₂) ada 10 items bahan bangunan, yang artinya ada **10** bahan bangunan yang termasuk kategori bahan bangunan laris.

Berikut daftar nama-nama bahan bangunan yang masuk dalam kategori bahan bangunan kurang laris berjumlah 24 Items, dapat dilihat pada tabel 4.23 sebagai berikut :

Tabel 4.23 Kelompok Kategori Bahan Bangunan Kurang Laris

No	Kode Barang	Nama Barang	Kode	Jarak Terdekat
1	CTLC-EXT-05-GS	Catylac Ext Grey Sand	A3	26.18318216
2	CTLC-EXT-05- Putih	Catylac Ext Putih 5 Kg 44855	A4	6.362051122
3	124-401-KR-01	Kran Cuci Piring Kranz KD 01	A6	50.45848816
4	13-07-05	Closed Jongkok Global Cream	A7	37.47522329
5	22-09-P-C	Pintu Paket Cream	A8	26.18318216
6	26-IS-T	Lem Isarplas Tube	A9	28.85698692
7	92-001-V-12AW	Pipa Invilon 1/2" AW	A10	5.850273026
8	92-001-V-1A	Pipa Invilon 1" AW	A11	11.9955142
9	92-001-V-212D	Pipa Invilon 2 1/2 D	A12	4.625187687
10	BCP-001-01	Bcp Crisna Mini	A13	26.80284241
11	LGG-16-3	Knee Soket 3"	A18	14.5679681
12	LGG-17-212	Knee Grest 2 1/2"	A19	5.935123797
13	LGG-35-212	Tee Langgeng 2 1/2" D	A20	45.46675373
14	RCK-14-12	Knee Rucika 1/2"	A21	33.81038836
15	RCK-34-1	Tee Rucika 1"	A22	30.89324135
16	RCK-34-12	Tee Rucika 1/2"	A23	31.31920116
17	160-206-12	Pahat Besi Picak Osiris 12"	A26	47.29579644
18	17-101-01	Thinner Keiji 1 Ltr	A27	47.29579644
19	171-001-316	Rantai 3/16 mm	A28	12.59202239
20	174-101-01-Lsn	Sarung Tangan Kain Bintik Lsn	A29	16.45880801
21	27-03-08	Reng Baut m8	A33	35.46912781
22	27-610-1065	Dinabol 10 x 65	A34	17.16757684
23	BK-408-75	Baut Kuning 8 x 75	A38	79.38970773
24	MBBTNLMSK-12	Mata Bor Beton Masaki 12 mm	A40	10.39514444

Berikut daftar nama-nama bahan bangunan yang masuk dalam kategori bahan bangunan sangat laris berjumlah 6 Items, dapat dilihat pada tabel 4.24 sebagai berikut :

Tabel 4.24. Kelompok Bahan Bangunan Sangat Laris

No	Kode Barang	Nama Barang	Kode	Jarak Terdekat
1	BSKS-08	Besi KS 8	A14	44.71452657
2	BSKSTY-06	Besi KSTY 6	A15	215.9893259
3	KB-00-1-KG	Kawat Beton 1 Kg	A16	135.7573407
4	20-04-1-BNT	Paku Bintang 1 Kg	A30	82.16683572
5	25-01-12-Mtr	Slang Benang Fuso 1/2" Mtr	A32	147.3297511
6	KB-00-1-KG	Kawat Beton 1 Kg	A39	361.302351

Berikut daftar nama-nama bahan bangunan yang masuk dalam kategori bahan bangunan laris berjumlah 10 Items, dapat dilihat pada tabel 4.25 sebagai berikut :

Tabel 4.25 Bahan Bangunan Kategori Laris

No	Kode Barang	Nama Barang	Kode	Jarak Terdekat
1	1000-09-803	Platon Minyak Golden Yellow 803	A1	31.82766093
2	17-101-01	Thinner Keiji 1 Ltr	A2	58.42088668
3	KUASCRO-3	Kuas Crocodile 3"	A5	51.89412298
4	LG-3-D	Pipa Lg 3" D	A17	61.84658438
5	SMN-PCC-MRH	Semen Padang Pcc Merah	A24	126.0039682
6	001-RB-26	Kawat Las RB 2,6 mm	A25	24.35159132
7	201-001-01	Bata	A31	120.8014901
8	28-02-10-11-MTR	Kawat Patri 1 x 1 x 1 Mtr	A35	24.0208243
9	30-01-112	Klem Slang 1 1/2"	A36	75.71657678
10	BJR-B-R	Baut Reng Baja Ringan	A37	67.2086304

4. Implementasi Sistem

Hasil dari implementasi sistem dengan menampilkan centroid, kelompok *cluster*, dan *diagram* terlihat pada Gambar dibawah ini

Gambar 5. *Cluster C₀* bahan bangunan kurang laris

cluster_0		
Catylac Ext Grey Sand	12	38
Catylac Ext Putih 5 Kg 44855	18	57
Kran Cuci Piring Kranz KD 01	70	50
Closed Jongkok Global Cream	20	100
Pintu Paket Cream	12	38
Lem Isarplas Tube	30	90
Pipa Invilon 1/2" AW	23	57
Pipa Invilon 1" AW	27	73
Pipa Invilon 2 1/2 D	22	58
Bcp Crisna Mini	13	37
Knee Soket 3"	22	48
Knee Grest 2 1/2"	19	57
Tee Langgeng 2 1/2" D	12	18
Knee Rucika 1/2"	3	34
Tee Rucika 1"	0	40
Tee Rucika 1/2"	3	37
Pahat Besi Picak Osiris 12"	50	100
Thinner Keiji 1 Ltr	50	100
Rantai 3/16 mm	20	50
Sarung Tangan Kain Bintik Lsn	21	79
Reng Baut m8	22	98

Gambar 6. *Cluster C₁* bahan bangunan sangat laris

cluster_1		
Besi KS 8	220	980
Besi KSTY 6	403	797
Kawat Beton 1 Kg	122	878
Paku Bintang 1 Kg	308	892
Slang Benang Fuso 1/2" Mtr	200	800
Kawat Beton 1 Kg	200	1300

Gambar 6. *Cluster C₂* bahan bangunan laris

cluster_2		
Platon Minyak Golden Yellow 803	30	200
Thinner Keiji 1 Ltr	60	280
Kuas Crocodile 3"	100	200
Pipa Lg 3" D	38	282
Semen Padang Pcc Merah	52	348
Kawat Las RB 2,6 mm	30	230
Bata	150	150
Kawat Patri 1 x 1 x 1 Mtr	52	98
Klem Slang 1 1/2"	11	159
Baut Reng Baja Ringan	7	173

Gambar 7. Hasil dari perhitungan *Text View* dengan menggunakan RapidMiner

Cluster Model

```
Cluster 0: 24 items
Cluster 1: 6 items
Cluster 2: 10 items
Total number of items: 40
```

5. Kesimpulan

Berdasarkan hasil yang diperoleh dari penelitian ini dengan metode *K-Means clustering*. Hasilnya mendapatkan 3 kategori penjualan bahan bangunan, yaitu kurang laris, sangat laris dan laris. Hal ini dibuktikan berdasarkan perhitungan jarak terdekat berdasarkan penentuan nilai centroid secara random dengan menggunakan rumus *Euclidean Distance*, pada jumlah stok barang diambil dari jumlah barang terjual dan jumlah barang tidak terjual. Maka didapat 3 iterasi dengan nilai *centroid* C_0 tidak terjual 21.125, terjual 62.54166667, nilai *centroid* C_1 tidak terjual 242.166667, terjual 941.1666667 dan *centroid* C_2 tidak terjual 53, terjual 222. Maka penelitian ini cocok sekali digunakan dalam pengetahuan penstockan barang, guna untuk meningkatkan aset penjualan.

Daftar Pustaka

- Agustin.W.Erlin., 2016. Implementasi Metode K-measn Cluster Analysis untuk Memilih Strategi Promosi Penerimaan Mahasiswa Baru. Seminar Nasional Ilmu Komputer (SNIK 2016)-Semarang. ISSN : 978-602-1034-40-8.
- Chang.E.C.dkk., 210. *Using K-means method and spectral clustering technique in an outfitter's value analysis*. ©Spiringer Science+Business Media B.V. 2009.
- Chowdhury.T.dkk., 2017. *An Efficient MapReduce-based Adaptive K-Means Clustering for Large Dataset*. 978-1-5386-1356-6/17 \$31.00© IEEE.
- Defit.S., 2013. Penggunaan Algoritma Apriori Dalam Menganalisa Prilaku Mahasiswa Dalam Memilih Mata Kuliah (Studi Kasus : FKIP UPI "YPTK"). Jurnal Media Processor Vol.8, No.3, Oktober 2013.
- Du.W.dkk., 2016. *A New Projection-based K-Means Initialization Algorithm*. *Proceedings of 2016 IEEE Chinese Guidance, Navigation and Controll Conference*. 978-1-4673-8318-9/16/\$31.00©2016 IEEE.
- Gornitz.N.dkk., 2017. *Support Vector Data Descriptions and k-Means Clustering : One Clas?*. 2162-237X© 2017.
- Hu.J.dkk., 2017. *Improved K-Means Algorithm Based on Hybrid Fruit Fly Optimization and Differential Evolution*. 978-1-5090-2508-4/17/\$31.00©2017 IEEE.
- Gu.L., 2016. *A Novel Locality Sensitive k-Means Clustering Algorithm based on Subtractive Clustering*. IEEE. 978-1-4673-9904-3/16/\$31.00 ©2016.
- Ikhwan.A.dkk., 20015. Penerapan Data Mining dengan Algoritma Fp-Growth untuk Mendukung Strategi Promosi Pendidikan. Jurnal Ilmiah SAINTIKOM Sains dan Komputer. ISSN : 1978-6603.
- Ong.J.O., 2013. Implementasi Algoritma K-Means Cluster Untuk menentukan Strategi Marketing President University. Jurnal Ilmiah Teknik Industri. ISSN : 1412-6869.
- Mythili.S., Madhiya.E., 2014. *An Analysis on Clustering Algorithms in Data Mining*. IJCSMC, Vol.3, Issue. A, January 2014, pg.334-340.
- Sharma.G.dkk., 2016. *Analysis of K-Means clustering for Human Capital trends*. IEEE. 978-1-5090-5515-9/16/\$31.00©2016.
- Siska.S.T., 2016. Analisa dan penerapan Data Mining Untuk menentukan Kubikasi Air Terjual Berdasarkan pengelompokan Pelanggan menggunakan Algoritma K-Means Cluster. STMIK-AMIK Riau : Jurnal Teknologi Informasi & Pendidikan. ISSN : 20186-4981.
- Tampubolon.K.dkk., 2013. Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan. Majalah Ilmiah Informasi dan Teknologi Ilmiah (INTI). ISSN : 2339-210X.
- Widiarlina., 2015. Algoritma Cluster Dinamik untuk Optimasi Cluster pada Algoritma K-Means dalam Pemetaan nasabah Potensial. Journal of Intelligent Systems. ISSN : 2356-3982.Vol.1.
- Zhong.S.dkk., 2016. *The expansibility Research of K-Means Algorithm under the GPU*. IEEE. 978-1-4673-9904-3/16/\$31.00©2016.