

IMPLEMENTASI ALGORITMA *K-MEANS* UNTUK PEMILIHAN KERAMIK DAN PELANGGAN POTENSIAL PADA CV. JAYA TUNGGAL KERAMIK

Implementation of K-Means Algorithm For Ceramic Selection and Potential Customer at CV. Jaya Tunggal Keramik

Yuliana¹⁾, Mario Richie^{2)*} dan Halim Agung, hagung@bundamulia.ac.id³⁾

¹⁾²⁾ Program Studi Teknik Informatika, Universitas Bunda Mulia

ABSTRACT

CV. Jaya Tunggal Keramik is a company that sale of ceramics. CV. Jaya Tunggal Keramik experienced some problems regarding ceramics and customers such as difficulties in sale ceramics to customers so that some ceramic products accumulate in the warehouse, such as being damaged and ceramic display becomes less good because it is stored too long and difficulty retaining customers because some customers do not want to order ceramic products. Lack of precise decision taken by the management CV. Jaya Tunggal Keramik in determining the strategy to supply ceramic and how to make it CV. Jaya Tunggal Keramik is difficult to estimate the stock of ceramic products to be provided and it is difficult to determine which potential customers can be maintained as a regular customer. This research uses K-Means algorithm. K-Means algorithm is a partitioning clustering method that separates data into different groups with iterative partitioning. By using this application, users can find out the estimated stock and price of ceramics as well as information about potential customers. Based on ceramic data test results, there are some ceramics that are not in accordance with the predicted results so it can be concluded that the K-Means algorithm on the test inventory data inventory in this study is not fully can provide accurate estimates, this is because the use of the K-Means algorithm is strongly influenced by the cluster center results and the attributes used.

Keywords: *customer, inventory, k-means, clustering*

ABSTRAK

CV. Jaya Tunggal Keramik adalah perusahaan yang bergerak dibidang penjualan keramik. CV. Jaya Tunggal Keramik mengalami beberapa permasalahan mengenai keramik dan pelanggan seperti kesulitan memasarkan keramik kepada pelanggan sehingga beberapa produk keramik menumpuk digudang, diantaranya menjadi rusak dan tampilan keramik menjadi kurang bagus karena disimpan terlalu lama serta kesulitan mempertahankan pelanggan dikarenakan beberapa pelanggan tidak ingin melakukan pemesanan produk keramik. Kurangnya tepatnya keputusan yang diambil oleh pihak manajemen CV. Jaya Tunggal Keramik dalam menentukan strategi terhadap persediaan keramik dan cara memasarkannya membuat CV. Jaya Tunggal Keramik kesulitan untuk memperkirakan stok produk keramik yang harus disediakan dan sulit untuk menentukan mana pelanggan yang potensial yang dapat dipertahankan sebagai pelanggan tetap. Penelitian ini menggunakan algoritma *K-Means*. Algoritma *K-Means* adalah metode clustering secara partitioning yang memisahkan data ke dalam kelompok yang berbeda dengan partitioning secara iteratif. Dengan menggunakan aplikasi ini, pengguna dapat mengetahui perkiraan stok dan harga keramik serta informasi mengenai pelanggan potensial. Berdasarkan hasil pengujian data keramik, terdapat beberapa keramik yang tidak sesuai dengan hasil prediksi sehingga dapat disimpulkan bahwa algoritma *K-Means* pada pengujian data persediaan barang pada penelitian ini tidak sepenuhnya bisa memberikan hasil perkiraan yang tepat, hal ini dikarenakan penggunaan algoritma *K-Means* sangat dipengaruhi oleh hasil pusat *cluster* dan atribut yang digunakan.

Kata Kunci: *pelanggan, persediaan barang, k-means, clustering*

PENDAHULUAN

CV. Jaya Tunggal Keramik merupakan salah satu perusahaan yang bergerak dalam bidang penjualan keramik. Dalam menjalankan usahanya, CV. Jaya Tunggal Keramik mengalami beberapa permasalahan mengenai produk keramik dan pelanggan seperti kesulitan memasarkan keramik kepada pelanggan yang mengakibatkan beberapa produk keramik menumpuk digudang, diantaranya menjadi rusak dan tampilan keramik menjadi kurang bagus karena disimpan terlalu lama serta kesulitan mempertahankan pelanggan karena beberapa pelanggan tidak ingin melakukan pemesanan produk keramik. Permasalahan yang terjadi mengakibatkan CV. Jaya Tunggal Keramik mengalami kerugian antara lain penjualan yang mengalami penurunan, target penjualan tidak tercapai, pelanggan kemungkinan beralih kepada perusahaan lain, banyaknya biaya pengeluaran terhadap produk keramik yang kurang diminati pelanggan sehingga mengurangi laba perusahaan.

Kurang tepatnya keputusan yang diambil oleh pihak manajemen CV. Jaya Tunggal Keramik dalam menentukan strategi terhadap persediaan keramik dan cara memasarkannya serta masih menggunakan pencatatan secara manual membuat CV. Jaya Tunggal Keramik kesulitan untuk memperkirakan stok produk keramik yang harus disediakan dan sulit untuk menentukan mana pelanggan yang potensial yang dapat dipertahankan sebagai pelanggan tetap. Data – data manual yang ada tidak dimanfaatkan oleh CV. Jaya Tunggal Keramik secara optimal untuk memperoleh informasi dikarenakan kesulitan mencari data – data yang ada dan memakan waktu yang lama karena harus dicari satu per satu.

Berdasarkan permasalahan yang terjadi di CV. Jaya Tunggal Keramik maka dalam penelitian ini, peneliti mengimplementasikan algoritma *K-Means* untuk dapat memperkirakan stok keramik yang harus disiapkan untuk penjualan

keramik serta dapat diperoleh informasi mengenai data pelanggan yang potensial.

Algoritma *K-Means* dapat digunakan untuk menyelesaikan beberapa permasalahan yang ada, seperti pada penelitian terdahulu yang dilakukan oleh peneliti sebelumnya dengan mengimplementasikan algoritma *K-Means clustering* untuk menentukan strategi *marketing President University* [4]. Pada penelitiannya, algoritma *K-Means* sukses diimplementasikan karena dapat diperoleh strategi promosi yang dapat dilakukan oleh pihak *marketing President University* untuk memperoleh calon mahasiswa baru.

Algoritma *K-Means* juga dapat digunakan untuk *clustering* produk *online shop* dalam penentuan stok barang [3] seperti yang dijelaskan dalam penelitian terdahulu mengenai algoritma *K-Means* dimana dihasilkan beberapa produk yang paling diminati untuk jumlah stok terbanyak dan produk yang kurang diminati untuk jumlah stok sedikit.

Algoritma *K-Means* juga digunakan untuk memberikan rekomendasi dosen tetap berdasarkan penilaian dosen [4] seperti yang dijelaskan dalam penelitian terdahulu mengenai algoritma *K-Means* dimana keberhasilan algoritma dalam memberikan rekomendasi dosen tetap sebesar 55.75%.

TINJAUAN PUSTAKA

Clustering

Pada dasarnya *clustering* merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*similarity*) antara satu data dengan data yang lain. *Clustering* merupakan salah satu metode data *mining* yang bersifat tanpa arahan (*unsupervised*), maksudnya metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*) serta tidak memerlukan target *output*. Dalam data *mining* ada dua jenis metode *clustering* yang digunakan dalam pengelompokan data, yaitu *hierarchical clustering* dan *non-hierarchical clustering* [6].

Clustering lebih ke arah pengelompokan *record*, pengamatan, atau kasus dalam kelas yang memiliki kemiripan. Sebuah *cluster* adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lain dan memiliki ketidakmiripan dengan *record-record* dalam *cluster* yang lain [2].

Hierarchical clustering adalah suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan kedua. Demikian seterusnya sehingga *cluster* akan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah *cluster*. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut [7]. Berbeda dengan metode *hierarchical clustering*, metode *non-hierarchical clustering* justru dimulai dengan menentukan terlebih dahulu jumlah *cluster* yang diinginkan (dua *cluster*, tiga *cluster*, atau lain sebagainya). Setelah jumlah *cluster* diketahui, baru proses *cluster* dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan *K-Means Clustering* [7].

Algoritma *K-Means*

K-Means merupakan metode *clustering* secara *partitioning* yang memisahkan data ke dalam kelompok yang berbeda dengan *partitioning* secara iteratif, *K-Means* mampu meminimalkan rata-rata jarak setiap data ke *cluster* [1].

Proses pengelompokan data secara umum ke dalam *cluster* tertuang dalam algoritma *clustering*, dalam hal ini digunakan algoritma *K-Means*. Algoritma *K-Means* mendasari metode operasinya berdasarkan namanya yaitu 'mean' atau

rerata. Dimana akan dilakukan pengamatan ke dalam kelompok *k*, dimana *k* diberikan sebagai parameter masukan. Kemudian dilakukan pengamatan pola untuk setiap kelompok berdasarkan kedekatan pada rerata jarak *cluster* yang adalah titik tengah, selanjutnya titik pusat atau *centroid* akan dikalkulasi ulang dan proses dimulai kembali. *K-Means* adalah teknik yang dikategorikan *greedy*, komputasi dilakukan dengan teknik yang efisien, menjadi salah satu algoritma *clustering* yang paling banyak digunakan untuk klusterisasi. Pada algoritma *clustering*, fokus ditujukan pada penentuan jumlah *cluster* atau *k*. Jumlah *cluster* yang ingin dibentuk ini akan digunakan sebagai masukan bagi algoritma. Pada dasarnya algoritma tidak mampu menentukan jumlah *cluster* dan ini bergantung sepenuhnya pada pengguna untuk mengidentifikasi terlebih dahulu jumlah *cluster* yang diinginkan. Tidak mudah menentukan banyak *cluster*, dan ini adalah strategi yang dipilih dengan asumsi bahkan pertimbangan yang sifatnya intuitif [8].

Langkah-langkah melakukan *clustering* dengan algoritma *K-Means* adalah [6] Langkah 1, Pilih jumlah *cluster* *k*. Langkah 2, Inisialisasi *k* pusat *cluster* ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara *random*. Pusat-pusat *cluster* diberi nilai awal dengan angka-angka *random*. Langkah 3, Alokasikan semua data / objek ke *cluster* terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke *cluster* tertentu ditentukan jarak antara data dengan pusat *cluster*. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat *cluster*. Jarak paling dekat antara satu data dengan satu *cluster* tertentu akan menentukan suatu data masuk dalam *cluster* mana. Untuk menghitung jarak semua data ke setiap titik pusat *cluster* dapat menggunakan teori jarak *Euclidean* yang dirumuskan pada rumus (1).

$$D(i,j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{in} - x_{jn})^2} \dots\dots\dots (1)$$

.....

dimana $D(i,j)$ = Jarak data ke i ke pusat $cluster$ j , X_{ki} = Data ke i pada atribut data ke k dan X_{kj} = Titik pusat ke j pada atribut ke k

Langkah 4, Hitung kembali pusat $cluster$ dengan keanggotaan $cluster$ yang sekarang. Pusat $cluster$ adalah rata-rata dari semua data / objek dalam $cluster$ tertentu. Jika dikehendaki bisa juga menggunakan *median* dari $cluster$ tersebut. Jadi rata-rata (*mean*) bukan satu-satunya ukuran yang bisa dipakai. Langkah 5, Tugaskan lagi setiap objek memaknai pusat $cluster$ yang baru. Jika pusat $cluster$ tidak berubah lagi maka proses *clustering* selesai. Atau, kembali ke langkah c sampai pusat $cluster$ tidak berubah lagi.

METODE PENELITIAN

Data – data yang digunakan dalam penelitian ini adalah data bulan November 2017. Dalam penelitian ini, hanya digunakan beberapa atribut data seperti nama keramik, *quantity* penjualan, harga untuk data keramik dan nama pelanggan, jumlah pembelian, total transaksi untuk data pelanggan. Pengolahan data dilakukan dengan mengelompokkan berdasarkan kemiripan karakteristik dari setiap data. Berikut ini adalah beberapa data keramik pada bulan November 2017.

Tabel 1 Data Keramik

No.	Nama Keramik	Quantity Penjualan	Harga
1	Maluku Wood	1000	52.000
2	Gianyar	500	52.000
3	Singaraja	650	52.000
4	Arish Hijau	300	45.000
5	Rosetta Beige	150	45.000
6	Ketty Yellowish	300	50.000
7	Matrix Beige	200	47.000

Untuk melakukan pengelompokan data-data keramik menjadi beberapa $cluster$ dilakukan dengan langkah – langkah yaitu Langkah 1, Mengelompokkan data keramik menjadi dua $cluster$.

Langkah 2, Menentukan titik pusat awal dari setiap $cluster$ berdasarkan

8	Fuzy Grey	600	50.000
9	Orion Grey	350	43.000
10	Bawean Grey	200	43.000

Untuk melakukan pengelompokan data-data keramik menjadi beberapa $cluster$ dilakukan dengan langkah – langkah yaitu Langkah 1, Mengelompokkan data keramik menjadi dua $cluster$.

Langkah 2, Menentukan titik pusat awal dari setiap $cluster$ berdasarkan *quantity* penjualan keramik tertinggi, *quantity* penjualan keramik terendah, harga tertinggi dan harga terendah

Tabel 2 Tabel Awal Clustering

	Quantity Penjualan	Harga
C1	1000	52000
C2	150	43000

Langkah 3, Menempatkan setiap data pada $cluster$. Dalam penelitian ini digunakan metode *hard K-Means* untuk mengalokasikan setiap data ke dalam suatu $cluster$, sehingga data akan dimasukkan dalam suatu $cluster$ yang memiliki jarak paling dekat dengan titik pusat dari setiap $cluster$. Untuk mengetahui $cluster$ mana yang paling dekat dengan data, maka perlu dihitung jarak setiap data dengan titik pusat setiap $cluster$. Hasil perhitungan jarak dari data keramik pertama ke pusat $cluster$ pertama adalah $D(1,1) = 0$

Jarak data keramik pertama dengan pusat $cluster$ kedua adalah $D(1,2) = 9040$

Berdasarkan hasil kedua perhitungan di atas, terlihat bahwa jarak data keramik pertama yang paling dekat adalah dengan $cluster$ 1, sehingga data keramik pertama dimasukkan ke dalam $cluster$ 1 yang ditandai dengan angka 1 pada C1. Hasil perhitungan selengkapnya setelah diimplementasikan pada program terlihat pada Tabel 3. *quantity* penjualan keramik tertinggi, *quantity* penjualan keramik terendah, harga tertinggi dan harga terendah

Tabel 2 Tabel Awal Clustering

	Quantity Penjualan	Harga
C1	1000	52000
C2	150	43000

Langkah 3, Menempatkan setiap data pada *cluster*. Dalam penelitian ini digunakan metode *hard K-Means* untuk mengalokasikan setiap data ke dalam suatu *cluster*, sehingga data akan dimasukkan dalam suatu *cluster* yang memiliki jarak paling dekat dengan titik pusat dari setiap *cluster*. Untuk mengetahui *cluster* mana yang paling dekat dengan data, maka perlu dihitung jarak setiap data dengan titik pusat setiap *cluster*. Hasil perhitungan jarak dari data keramik pertama ke pusat *cluster* pertama adalah $D(1,1) = 0$

Jarak data keramik pertama dengan pusat *cluster* kedua adalah $D(1,2) = 9040$

Berdasarkan hasil kedua perhitungan di atas, terlihat bahwa jarak data keramik pertama yang paling dekat adalah dengan *cluster* 1, sehingga data keramik pertama dimasukkan ke dalam *cluster* 1 yang ditandai dengan angka 1 pada C1. Hasil perhitungan selengkapnya setelah diimplementasikan pada program terlihat pada Tabel 3.

Tabel 3 Perhitungan Jarak Antara Pusat Cluster Dengan Data Keramik

No	Nama Keramik	Quantity Penjualan	Harga	Centroid 1 (1000, 52000)	Centroid 2 (150,4 3000)	C 1	C 2
1	MALU KU WOOD	1000	52000	0	9040.049778624	1	0
2	GIAN YAR	500	52000	500	9006.8029844113	1	0
3	SING ARAJ A	650	52000	350	9013.87818866	1	0
4	ARIS H HIJAU	300	45000	7034.9129347846	2005.6171120132	0	1
5	ROSE TTA BEIGE	150	45000	7051.418297052	2000	0	1
6	KETT Y YELL OWIS H	300	50000	2118.9620100417	7001.606958406	1	0
7	MATR IX BEIGE	200	47000	5063.5955604689	4000.3124877939	0	1
8	FUZY GREY	600	50000	2039.6078054371	7014.4493725452	1	0
9	ORIO N GREY	350	43000	9023.441693722	200	0	1
10	BAWEAN GREY	200	43000	9035.4855984612	50	0	1

Langkah 4, Setelah semua data ditempatkan ke dalam *cluster* yang terdekat, kemudian dihitung kembali pusat *cluster*

yang baru berdasarkan rata-rata anggota yang ada pada *cluster* dengan cara membagi jumlah *quantity* penjualan dan harga dengan jumlah data pada masing – masing *cluster*.

Jumlah *quantity* penjualan yang ada pada C1 adalah 3.050, banyaknya data pada C1 adalah 5 sehingga didapatkan hasil pusat *cluster* baru sebesar 610 untuk *quantity* penjualan dan 51.200 untuk harga. Hasil perhitungan selengkapnya untuk pusat *cluster* baru terlihat pada Tabel 4.

Tabel 4 Tabel Pusat Cluster Baru

	Quantity Penjualan	Harga
C1	610	51200
C2	240	44600

Langkah 5, Setelah didapatkan titik pusat yang baru dari setiap *cluster*, dilakukan perhitungan kembali jarak setiap data dengan titik pusat setiap *cluster* hingga titik pusat dari setiap *cluster* tidak berubah lagi dan tidak ada lagi data yang berpindah dari satu *cluster* ke *cluster* yang lain.

Dalam penelitian ini, iterasi *clustering* data keramik terjadi sebanyak 2 kali iterasi. Pada iterasi ke-2, titik pusat dari setiap *cluster* sudah tidak berubah dan tidak ada lagi data yang berpindah dari satu *cluster* ke *cluster* yang lain sehingga didapatkan hasil perkiraan stok keramik dan harga keramik pada program ditunjukkan pada Tabel 5.

Tabel 5 Hasil Perkiraan Stok dan Harga Keramik

No	Nama Keramik	Qty Penjualan	Harga	Qty Stok Prediksi	Harga Prediksi
1	MALUKU WOOD	1000	52000	610	51200
2	GIAN YAR	500	52000	610	51200
3	SINGARAJA	650	52000	610	51200
4	ARISH HIJAU	300	45000	240	44600
5	ROSETTA BEIGE	150	45000	240	44600
6	KETTY YELLOWISH	300	50000	610	51200
7	MATRIX BEIGE	200	47000	240	44600
8	FUZY GREY	600	50000	610	51200
9	ORION GREY	350	43000	240	44600
10	BAWEAN GREY	200	43000	240	44600

Langkah 6, Hasil perkiraan stok dan harga diperoleh dari pusat *cluster* terakhir. Apabila data keramik berada pada C1 maka stok dan harga prediksi akan diambil dari pusat *cluster* C1, sebaliknya apabila data keramik berada pada C2 maka

stok dan harga prediksi akan diambil dari pusat *cluster* C2.

Selanjutnya, untuk melakukan pengelompokan data-data pelanggan menjadi beberapa *cluster* dilakukan langkah – langkah yang sama seperti pengelompokan data keramik. Untuk data pelanggan akan menghasilkan informasi pelanggan yang potensial. Syarat pelanggan potensial di CV. Jaya Tunggal Keramik adalah pelanggan yang mempunyai transaksi sepuluh juta rupiah dan di atas sepuluh juta rupiah serta sering melakukan pembelian keramik. Hasil *output* data pelanggan pada bulan November 2017 dapat dilihat pada Tabel 6.

Tabel 6 Hasil Output Data Pelanggan

No	Nama Pelanggan	Jumlah Pembelian	Total Transaksi	Keterangan
1	Chandra Agung Perkasa	1	4.500.000,00	tidak potensial
2	Cemerlang	1	6.525.000,00	tidak potensial
3	Guna Bangunan	1	3.525.000,00	tidak potensial
4	Klasit Abadi	2	4.700.000,00	tidak potensial
5	Sinar Baru	1	4.350.000,00	tidak potensial
6	Adika Sentosa	1	10.400.000,00	tidak potensial
7	Sinar Remaja	1	4.420.000,00	tidak potensial
8	Omega Ceramic	1	5.000.000,00	tidak potensial
9	Bintang Jaya	1	870.000,00	tidak potensial
10	Laguna	3	15.000.000,00	potensial

HASIL DAN PEMBAHASAN

Berdasarkan hasil pengujian 34 data keramik yang dilakukan antara data prediksi bulan November dengan bulan Desember 2017 terlihat bahwa hanya 6 keramik yang sesuai atau mendekati perkiraan dengan stok bulan Desember yang sedang berjalan. Sedangkan 23 data keramik lainnya tidak sesuai dengan hasil prediksi dan 5 buah keramik tidak distok dengan alasan barang tersebut kurang laku. Untuk hasil data pelanggan seperti yang ditampilkan pada Tabel 6, terlihat bahwa pelanggan yang potensial adalah Laguna karena sering melakukan pembelian sebanyak tiga kali dengan total transaksi

terbesar yaitu lima belas juta rupiah. Hasil data *output* sesuai dengan syarat ketentuan di CV. Jaya Tunggal keramik. Hasil pengujian dengan menggunakan 34 data keramik setelah implementasi pada program dapat dilihat pada Tabel 7

Tabel 7 Hasil Pengujian Data Keramik Pada Program

No	Nama Keramik	Qty Penjualan	Harga	Qty Prediksi	Harga Prediksi	Stok Desember	Keterangan
1	MALUKU WOOD	1000	52000	297	51.735	1500	Tidak sesuai
2	BROMO BROWN	250	52000	297	51.735	150	Tidak sesuai
3	BROMO GREY	150	52000	297	51.735	290	sesuai
4	KRAKA TAU BLACK	200	52000	297	51.735	100	Tidak sesuai
5	KRAKA TAU GREY	200	52000	297	51.735	100	Tidak sesuai
6	MATARAM WOOD	150	52000	297	51.735	100	Tidak sesuai
7	MAHAKAM	300	52000	297	51.735	400	Tidak sesuai
8	JAVA WOOD	150	52000	297	51.735	75	Tidak sesuai
9	GIANYAR	500	52000	297	51.735	50	Tidak sesuai
10	SINGARAJA	650	52000	297	51.735	300	sesuai
11	HOUSTON BROWN	100	52000	297	51.735	50	Tidak sesuai
12	ARISH HIJAU	300	45000	175	44.941	180	sesuai
13	ARISH BEIGE	300	45000	175	44.941	700	Tidak sesuai
14	ARISH ABU	125	45000	175	44.941	800	Tidak sesuai
15	ROSETTA BEIGE	150	45000	175	44.941	0	Tidak sesuai
16	ROSETTA ABU	150	45000	175	44.941	125	Tidak sesuai
17	KETTY YELLOW WISH	300	50000	297	51.735	0	Tidak sesuai
18	KETTY BLUISH	125	50000	297	51.735	295	sesuai
19	JENIFER	150	50000	297	51.735	500	Tidak sesuai
20	BRIAN GREEN	100	50000	297	51.735	100	Tidak sesuai
21	MATRIX BEIGE	200	47000	175	44.941	100	Tidak sesuai
22	MATRIX BLUE	200	47000	175	44.941	100	Tidak sesuai
23	MATRIX GREEN	100	47000	175	44.941	50	Tidak sesuai
24	VENUS BEIGE	100	47000	175	44.941	0	Tidak sesuai
25	VENUS BLUE	125	47000	175	44.941	250	Tidak sesuai
26	FUZY GREY	600	50000	297	51.735	150	Tidak sesuai
27	SOPHIE BEIGE	120	57500	297	51.735	175	Tidak sesuai
28	ORION GREY	350	43000	175	44.941	0	Tidak sesuai
29	ORION BEIGE	150	43000	175	44.941	0	Tidak sesuai
30	BAWEAN GREY	200	43000	175	44.941	50	Tidak sesuai

3 1	AGACI A GREY	100	430 00	175	44.9 41	174	sesuai
3 2	SAHAR A BEIGE	120	450 00	175	44.9 41	300	Tidak sesuai
3 3	AKASH A BEIGE	150	435 00	175	44.9 41	170	sesuai
3 4	AKASH A GREY	150	435 00	175	44.9 41	100	Tidak sesuai

Pembahasan

Berdasarkan hasil pengujian data keramik diatas, maka implementasi algoritma *K-Means* untuk pengujian data bulan November dengan Desember untuk perkiraan stok barang yang akan datang dan harga untuk konsumen tidak sepenuhnya bisa memberikan hasil perkiraan yang tepat. Sedangkan untuk data pelanggan, dapat diperoleh informasi mengenai pelanggan yang potensial untuk dapat dipertahankan sebagai pelanggan tetap.

SIMPULAN

Dari hasil penelitian, dapat disimpulkan bahwa untuk memperoleh hasil pengujian dengan mengimplementasikan algoritma *K-Means* sebaiknya menggunakan pengujian data lebih banyak dengan membandingkan data setiap periode bulan serta dapat menggunakan atribut data lainnya sehingga hasil pengujian akan memberikan hasil yang lebih baik. Penggunaan algoritma *K-Means*, sangat tergantung pada hasil pusat *cluster* dimana hasil *cluster* juga dipengaruhi dari nilai *centroid* awal yang dimasukkan, atribut data dan jumlah data yang digunakan, perbedaan pengambilan data pusat *centroid* awal juga mempengaruhi hasil akhir *centroid*.

DAFTAR PUSTAKA

- [1] Astuti, W. 2015. *Clustering Program Keahlian Pada Pendaftaran Siswa Baru (PSB) Dengan Menggunakan Algoritma K-Means: Studi Kasus Di SMKN 1 Nguling*. ISSN: 2085 – 3092
- [2] Kusriani, Lutfi dan Emha Taufiq. 2009. *Algoritma Data Mining*. Yogyakarta:Penerbit Andi.
- [3] Muningsih, Elly dan Sri Kiswati. 2015. *Penerapan Metode K-Means Untuk Clustering Produk Online Shop Dalam Penentuan Stok Barang*. Yogyakarta: AMIK BSI. ISSN: 2338-9761
- [4] Santoso, Yessica P. Marlina. Agung, H. 2018. *Implementasi Metode K-Means Clustering pada Sistem Rekomendasi Dosen Tetap Berdasarkan Penilaian Dosen*. Banten: Universitas Pamulang. ISSN: 2541-1004
- [5] Ong, Oscar J. 2013. *Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University*. Bekasi: Universitas Presiden. ISSN: 1412-6869
- [6] Santosa, B. 2007. *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*, Graha Ilmu, Yogyakarta.
- [7] Santoso, S. 2010. *Statistik Multivariat*. Jakarta: Elex Media Komputindo.
- [8] Sulianta, Feri. 2014. *Customer Profiling Pada Supermarket Menggunakan Algoritma K-Means Dalam Memilih Produk Berdasarkan Selera Konsumen Dengan Daya Beli Maksimum*. Universitas Widyatama. ISSN : 2407-3911.