GATOT KACA JOURNAL
Jurnal Keilmuan Teknik: Aplikasi, Perencanaan, Rancang Bangun, Sistem, Simulasi, & Rekayasa
e-ISSN : 2775-4561 (Online)
ISSN : 2775-2992(Print)

Penerapan Data Mining Untuk Menentukan Jumlah Produksi Menggunakan Metode K-Means (Studi Kasus Pada Pt. Coca Cola Distribution Indonesia Bengkulu)

Application Of Data Mining To Determine The Amount Of Production Using The K-Means Method (Case Study At Pt. Coca Cola Distribution Indonesia Bengkulu)

Yessi Mardiana[1)]; Toibah Umi Kalsum [2)];
[1)]Study of Computer System Engineerings Faculty of Computer Science, Universitas Dehasen Bengkulu
Email: yessimrd@gmail.com

How to Cite :
Mardiana, Y.; Kalsum, T. U. (2022). *Application Of Data Mining To Determine The Amount Of Production Using The K-Means Method (Case Study At Pt. Coca Cola Distribution Indonesia Bengkulu)*. Gatotkaca Journal, 3(1) page: 1-10. DOI: https://doi.org/10.37638/gatotkaca.3.1.1-10

ARTICLE HISTORY
Submited [26 Januari 2022]
Received [28 Januari 2022]
Revised [29 Januari 2022]
Accepted [31 Januari 2022]

KEYWORDS
K-Means Clustering, PT. Coca Cola Distribution Indonesia

This is an open access article under the CC–BY-SA license

ABSTRAK

Penelitian ini menerapkan Data Mining dengan menggunakan metode *Clustering* untuk menentukan jumlah produksi berdasarkan tingkat penjualan produk di PT. Coca Cola Distribution Indonesia cabang Bengkulu. Algoritma yang digunakan yaitu *K-Means Clustering,* di mana data dikelompokkan berdasarkan karakteristik yang sama akan dimasukkan ke dalam kelompok yang sama dan set data yang dimasukkan ke dalam kelompok tidak tumpang tindih. Informasi yang ditampilkan berupa kelompok – kelompok data produk berdasarkan tingkat penjualannya, sehingga diketahui jumlah produksi pada PT.Coca Cola Distribution Indonesia yang disesuaikan dengan *cluster* produk dengan tingkat penjualan tinggi, sedang dan rendah, sehingga menghasilkan *cluster-cluster* produk yang dapat dimanfaatkan oleh PT. Coca Cola Distribution Indonesia dalam menentukan jumlah produksi produk berikutnya.

ABSTRACT

This study applies Data Mining using the Clustering method to determine the amount of production based on the level of product sales at PT. Coca Cola Distribution Indonesia Bengkulu branch. The algorithm used is K-Means Clustering, where data grouped based on the same characteristics will be included in the same group and the data sets entered into the groups do not overlap. The information displayed is in the form of product data groups

GATOT KACA JOURNAL
Jurnal Keilmuan Teknik: Aplikasi, Perencanaan, Rancang Bangun, Sistem, Simulasi, & Rekayasa
e-ISSN : 2775-4561 (Online)
ISSN : 2775-2992(Print)

based on the level of sales, so that it is known the amount of production at PT. Coca Cola Distribution Indonesia which is adjusted to product clusters with high, medium and low sales levels, resulting in product clusters that could be utilized by PT. Coca Cola Distribution Indonesia in determining the number of production of the next product

PENDAHULUAN

Seiring dengan pertumbuhan bisnis di era globalisasi dan kemajuan di bidang teknologi informasi yang cepat memberikan pengaruh yang cukup besar baik dalam bidang industri maupun jasa. Hal ini juga membawa suatu perubahan besar dalam tingkat persaingan antara perusahaan, sehingga pelaku-pelaku perusahaan tersebut harus selalu menciptakan berbagai teknik untuk terus berkembang.

Dalam rangka menghadapi persaingan bisnis dan meningkatkan pendapatan perusahaan, pimpinan perusahaan maupun manajemen dalam suatu perusahaan tersebut dituntut untuk dapat mengambil keputusan yang tepat dalam menentukan strategi penjualan. Untuk dapat melakukan hal tersebut, perusahaan membutuhkan sumber informasi yang cukup banyak untuk dapat dianalisis lebih lanjut. Pihak eksekutif perusahaan mengharapkan adanya teknologi yang mampu menghasilkan suatu informasi yang siap digunakan untuk membantu mereka dalam mengambil keputusan strategis perusahaan. Mereka ingin mengetahui produk apa yang harus ditingkatkan produksi berikutnya, seberapa besar pencapaian hasil yang diperoleh oleh perusahaan. Untuk memenuhi kebutuhan-kebutuhan di atas, banyak cara yang dapat ditempuh. Salah satunya adalah dengan melakukan pemanfaatan data perusahaan (*Data Mining*).

PT. Coca Cola Distribution Indonesia, terdapat beberapa permasalahan yang kerap muncul mengenai penjualan dan produksi produk. Perusahaan sulit mendapatkan informasi-informasi strategis seperti tingkat penjualan per periode. Ketersediaan data penjualan yang besar di PT. Coca Cola Distribution Indonesia tidak digunakan semaksimal mungkin, sehingga data penjualan tersebut tidak dimanfaatkan secara optimal dan belum adanya sistem pendukung keputusan dan metode yang dapat digunakan untuk merancang sebuah strategi bisnis untuk menentukan jumlah produksi produk berdasarkan tingkat penjualan yang dihasilkan.

Berdasarkan latar belakang di atas, maka penulis tertarik untuk meneliti bidang ini dengan mengambil judul **"**Penerapan Data Mining untuk Menentukan Jumlah Produksi Menggunakan Metode K-Means**"**.

LANDASAN TEORI

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 terakit dari berbagai database besar. Dan *Data Mining* adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Di mana hasil dari proses penggalian tersebut akan membentuk pola-pola dari kumpulan data, yang sering disebut dengan pengenalan pola (Deka, *et al*, 2017).

GATOT KACA JOURNAL
Jurnal Keilmuan Teknik: Aplikasi, Perencanaan, Rancang Bangun, Sistem, Simulasi, & Rekayasa
e-ISSN : 2775-4561 (Online)
ISSN : 2775-2992(Print)

Algoritma K-Means Clustering

Menurut Prianto (2020:18), Algoritma K-Means Clustering merupakan salah satu algoritma dengan partitional, karena K-Means Clustering didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai centroid awalnya. Dibutuhkan jumlah cluster awal yang diinginkan sebagai masukan dan menghasilkan titik centroid akhir sebagai output. Metode K-Means clustering akan memilih pola k sebagai titik awal centroid secara acak atau random. Jumlah iterasi untuk mencapai cluster centroid akan dipengaruhi oleh calon cluster centroid awal secara random. Sehingga didapat cara dalam pengembangan algoritma dengan menentukan centroid cluster yang dilihat dari kepadatan data awal yang tinggi agar mendapatkan kinerja yang lebih tinggi.

Menurut Tutik (2014), tahapan melakukan *clustering* atau pengelompokan dengan metode *K-Means* adalah sebagai berikut :

1. Menentukan berapa banyak *cluster* yang ingin yang ingin dibentuk, di mana nilai K adalah banyaknya cluster/ jumlah *cluster*.
2. Menentukan pusat *cluster* (*centroid*) awal. *Centroid* awal ditentukan secara acak dari data yang ada dan jumlah *centroid* awal sama dengan jumlah *cluster*.
3. Setelah menentukan *centroid* awal, maka setiap data akan menemukan *centroid* terdekatnya yaitu dengan menghitung jarak setiap data ke masing-masing *centroid* menggunakan rumus korelasi antar dua obyek yaitu *Euclidean Distance.*
4. Setelah menghitung jarak data ke *centroidnya*, maka langkah berikutnya adalah mengelompokkan data berdasarkan jarak minimumnya. Suatu data akan menjadi anggota dari suatu *cluster* yang memiliki jarak terdekat (terkecil) dari pusat *cluster*-nya.
5. Berdasarkan pengelompokkan tersebut, selanjutnya adalah mencari *centroid* baru berdasarkan membership dari masing-masing *cluster* yaitu dengan menghitung rata-rata dari data masing-masing *cluster*.
6. Kembali ke tahap 3.
7. Perulangan berhenti apabila tidak ada data yang berpindah

Menurut Tutik (2014) untuk menentukan nilai pusat (*centroid*) pada tahap *iterasi* digunakan rumus sebagai berikut :

$$v_{ij} = \frac{1}{N_i} = \sum_{k=0}^{N_i} x_{ki} \quad \text{..... (1)}$$

Di mana :

Vij = *centroid* rata-rata *cluster* ke *i* untuk variable ke *j*
Ni = jumlah anggota *cluster* ke *i*
i, k = *indeks* dari *cluster*
j = *indeks* dari variable
*X*kj = nilai data ke *k* variable ke *j* dalam *cluster* tersebut

Menurut Afrisawati (2013) untuk menentukan korelasi antar dua obyek yaitu dengan menggunakan rumus *Euclidean Distance* berikut:

$$d_{Euclidean}(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad \text{..... (2)}$$

Di mana :

d(x,y) = jarak data ke *x* ke pusat cluster *y*
x_i = data ke *i* pada atribut data ke *n*
y_i = data ke *j* pada atribut data ke *n.*

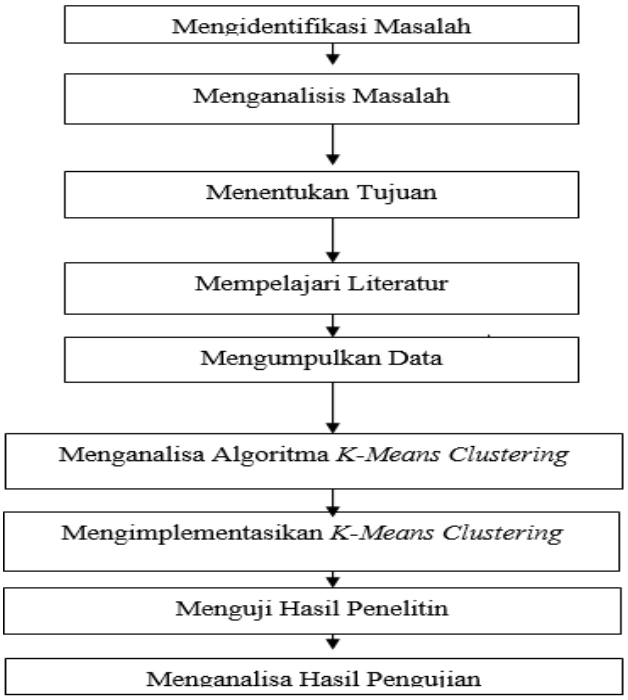
GATOT KACA JOURNAL
Jurnal Keilmuan Teknik: Aplikasi, Perencanaan, Rancang Bangun, Sistem, Simulasi, & Rekayasa
e-ISSN : 2775-4561 (Online)
ISSN : 2775-2992(Print)

METODE PENELITIAN

Metode Analisis

Penelitian ini memiliki beberapa tahapan dalam pelaksanaan kegiatan yang tertuang pada kerangka kerja penelitian yaitu identifikasi masalah, analisa masalah, menentukan tujuan, mempelajari literatur, mengumpulkan data, analisa Algoritma *K-Means Clustering*, pengolahan data, implementasi, dan pengujian. Adapun kerangka kerja yang digunakan dalam penulisan ini adalah seperti terlihat pada gambar berikut :

Gambar 1. Kerangka Kerja



HASIL DAN PEMBAHASAN

Pembahasan

Data yang digunakan pada penelitian ini adalah data penjualan produk di PT. Coca Cola Distribution Indonesia Cabang Bengkulu. *Cluster* atau pengelompokan yang ingin dibentuk yaitu tingkat penjualan produk berdasarkan stok dan total penjualan produk yang diambil penjualan selama satu bulan. Di mana *cluster* 0 mewakili tingkat penjualan rendah, *cluster* 1 mewakili tingkat penjualan sedang, dan *cluster* 2 mewakili tingkat penjualan yang tinggi. Sehingga hasil *cluster* tersebut dapat menentukan jumlah produksi produk. Adapun variabel yang digunakan dalam pengelompokan atau *clustering* adalah stok dan total penjualan produk dalam satu bulan Januari di tahun 2014, sehingga nilai masing-masing data dapat dilihat dalam tabel 1.

GATOT KACA JOURNAL
Jurnal Keilmuan Teknik: Aplikasi, Perencanaan, Rancang Bangun, Sistem, Simulasi, & Rekayasa
e-ISSN : 2775-4561 (Online)
ISSN : 2775-2992(Print)

Tabel 1. Data Penjualan PT. Coca Cola Distribution Indonesia Cabang Bengkulu

No	Nama Produk	Stok	Total
1	Frestea 220 ml	75.000,00	65.087,00
2	Frestea Apel 200 ml	55.000,00	15.520,00
3	Green Tea 250 ml	75.000,00	38.768,00
4	Coke PET 1.25 ml	65.000,00	33.296,00
5	Fanta Strawbery PET 1.25 ml	75.000,00	71.407,00
6	Fanta Strawbery PET 1 Liter	75.000,00	65.342,00
7	Coke Zero PET 1 Liter	65.000,00	30.464,00
8	Coke PET 500 ml	65.000,00	37.011,00
9	Fanta Strawbery PET 500 ml	75.000,00	65.573,00
10	Fanta VIT C Lemon PET 500 ml	55.000,00	12.419,00
11	Coke PET 1.5 ltr	65.000,00	32.077,00
12	Fanta Strawbery PET 1.5 ltr	75.000,00	66.196,00
13	Diet Coke PET 1.5 ltr	45.000,00	9.021,00
14	Coke Can 250 ml 36	75.000,00	72.987,00
15	Soda Water Can 250 ml 36	65.000,00	35.987,00
16	Fanta Cream Can 250 ml 36	45.000,00	16.839,00
17	Sprite Can 250 ml 9x4(36	65.000,00	34.765,00
18	Sprite Ice Can 250 ml 9x4(36	50.000,00	15.442,00
19	Fanta Strawbery Can 250 ml 9x4(36	75.000,00	70.980,00
20	Sprite Can 250 ml X24	75.000,00	70.371,00
21	Fanta Fruit Punch Can 250 ml X24	45.000,00	15.463,00
22	F. Soda Water Can 250 ml X24	60.000,00	28.640,00
23	Coke Scanpack 4 x 6	60.000,00	29.595,00
24	Fanta Strawbery Scanpack 4 x 6	75.000,00	69.300,00
25	A & W Scanpack 4 x 6	45.000,00	8.471,00
26	Ades 600 ml	65.000,00	31.831,00

Untuk dapat melakukan pengelompokan data-data tersebut menjadi beberapa *cluster* perlu dilakukan beberapa langkah, yaitu :

1. Tentukan jumlah *cluster* yang diinginkan. Dalam penelitian ini data-data yang ada akan dikelompokkan mejadi tiga *cluster*. Yaitu *cluster* 0 merupakan kelompok produk dengan penjualan rendah, *cluster* 1 kelompok produk dengan penjualan sedang, dan *cluster* 2 kelompok produk dengan penjualan tinggi.
2. Tentukan titik pusat awal *cluster* (*centroid*). Dalam penelitian ini titik pusat awal ditentukan secara random atau acak.
3. Setelah menentukan *centroid* awal, maka setiap data akan menemukan *centroid* terdekatnya yaitu dengan menghitung jarak setiap data ke masing-masing *centroid* menggunakan rumus korelasi antar dua obyek yaitu *Euclidean Distance.* Adapun penghitungan *centroid* awal secara manual hanya menghitung 10 data *sample* saja ada. Perhitungannya adalah sebagai berikut:

GATOT KACA JOURNAL
Jurnal Keilmuan Teknik: Aplikasi, Perencanaan, Rancang Bangun, Sistem, Simulasi, & Rekayasa
e-ISSN : 2775-4561 (Online)
ISSN : 2775-2992(Print)

Perhitungan untuk *Cluster* 0:

$$D(C0,1) = \sqrt{(45.000 - 75.000)^2 + (8471 - 65.087)^2} = 64.073,17$$
$$D(C0,2) = \sqrt{(45.000 - 55.000)^2 + (8471 - 15.520)^2} = 12.234,72$$
$$D(C0,3) = \sqrt{(45.000 - 75.000)^2 + (8471 - 38.768)^2} = 42636,93$$
$$D(C0,4) = \sqrt{(45.000 - 65.000)^2 + (8471 - 33.296)^2} = 31.879,16$$
$$D(C0,5) = \sqrt{(45.000 - 75.000)^2 + (8471 - 71.407)^2} = 69.720,44$$
$$D(C0,6) = \sqrt{(45.000 - 75.000)^2 + (8471 - 65.342)^2} = 64.298,61$$
$$D(C0,7) = \sqrt{(45.000 - 65.000)^2 + (8471 - 30.464)^2} = 29.729,96$$
$$D(C0,8) = \sqrt{(45.000 - 65.000)^2 + (8471 - 37.011)^2} = 34.850,13$$
$$D(C0,9) = \sqrt{(45.000 - 75.000)^2 + (8471 - 65.573)^2} = 64.503,01$$
$$D(C0,10) = \sqrt{(45.000 - 55.000)^2 + (8471 - 12.419)^2} = 10,751,00$$

Perhitungan untuk *Cluster* 1:

$$D(C1,1) = \sqrt{(65.000 - 75.000)^2 + (30.464,00 - 65.087)^2} = 36.038,20$$
$$D(C1,2) = \sqrt{(65.000 - 55.000)^2 + (30.464,00 - 15.520)^2} = 17.981,19$$
$$D(C1,3) = \sqrt{(65.000 - 75.000)^2 + (30.464,00 - 38.768)^2} = 12,998,32$$
$$D(C1,4) = \sqrt{(65.000 - 65.000)^2 + (30.464,00 - 33.296)^2} = 2.832,00$$
$$D(C1,5) = \sqrt{(65.000 - 75.000)^2 + (30.464,00 - 71.407)^2} = 42.146,52$$
$$D(C1,6) = \sqrt{(65.000 - 75.000)^2 + (30.464,00 - 65.342)^2} = 36.283,26$$
$$D(C1,7) = \sqrt{(65.000 - 65.000)^2 + (30.464,00 - 30.464)^2} = 0$$
$$D(C1,8) = \sqrt{(65.000 - 65.000)^2 + (30.464,00 - 37.011)^2} = 6.547,00$$
$$D(C1,9) = \sqrt{(65.000 - 75.000)^2 + (30.464,00 - 65.573)^2} = 36.505,37$$
$$D(C1,10) = \sqrt{(65.000 - 55.000)^2 + (30.464,00 - 12.419)^2} = 20,630,61$$

Perhitungan untuk *Cluster* 2:

$$D(C2,1) = \sqrt{(75.000 - 75.000)^2 + (65.087,00 - 65.087)^2} = 0$$
$$D(C2,2) = \sqrt{(75.000 - 55.000)^2 + (65.087,00 - 15.520)^2} = 53.449,86$$
$$D(C2,3) = \sqrt{(75.000 - 75.000)^2 + (65.087,00 - 38.768)^2} = 26.319,00$$
$$D(C2,4) = \sqrt{(75.000 - 65.000)^2 + (65.087,00 - 33.296)^2} = 33.326,68$$
$$D(C2,5) = \sqrt{(75.000 - 75.000)^2 + (65.087,00 - 71.407)^2} = 6.320,00$$
$$D(C2,6) = \sqrt{(75.000 - 75.000)^2 + (65.087,00 - 65.342)^2} = 255,00$$
$$D(C2,7) = \sqrt{(75.000 - 65.000)^2 + (65.087,00 - 30.464)^2} = 36.038,20$$
$$D(C2,8) = \sqrt{(75.000 - 65.000)^2 + (65.087,00 - 37.011)^2} = 29.803,72$$
$$D(C2,9) = \sqrt{(75.000 - 75.000)^2 + (65.087,00 - 65.573)^2} = 486,00$$
$$D(C2,10) = \sqrt{(75.000 - 55.000)^2 + (65.087,00 - 12.419)^2} = 56.337,54$$

GATOT KACA JOURNAL
Jurnal Keilmuan Teknik: Aplikasi, Perencanaan, Rancang Bangun, Sistem, Simulasi, & Rekayasa
e-ISSN : 2775-4561 (Online)
ISSN : 2775-2992(Print)

4. Setelah menghitung jarak data ke *centroidnya*, maka langkah berikutnya adalah mengelompokkan data berdasarkan jarak minimumnya.
5. Kemudian menentukan keanggotaan objek ke dalam matrik, dengan elemen matriks bernilai 1 apabila objek menjadi anggota group.

Berdasarkan nilai minimum yang telah dihasilkan pada Tabel 4.6. tersebut di atas pada penentuan nilai *centroid* maka diperoleh hasil pengelompokan seperti terlihat pada tabel 2.

Tabel 2. Hasil Pengelompokan

Kelompok (cluster)	Anggota Kelompok	Jumlah Anggota
0	[2, 10, 13, 16, 18, 21, 25]	7 Anggota
1	[3, 4, 7, 8, 11, 15, 17, 22, 23,26]	10 Anggota
2	[1, 5, 6, 9, 12, 14, 19, 20, 24]	9 Anggota

6. Tahap Selanjutnya dihitung *centroid* yang baru untuk setiap *cluster* berdasarkan data yang bergabung pada setiap *cluster*-nya. Untuk *cluster* 0, ada 7 anggota yang tergabung ke dalamnya yakni 2, 10, 13, 16, 18, 21 dan 25 sehingga :

Centroid 1: $C(1,1)= \frac{55.000+55.000+45.000+45.000+50.000+45.000+45.000}{7} = 48.571,43$

Centroid 2:$C(1,2) = \frac{15.520+12.419,00+9.021,00+16.839,00+15.442,00+15.463+8471}{7} = 13.243,40$

Untuk *cluster* 1, ada 10 anggota yang tergabung ke dalamnya yakni 3, 4, 7, 8, 11, 15, 17, 22, 23 dan 26 sehingga :

Centroid 1 :$C(2,1)=\frac{75.000+65.000+65.000+65.000+65.000+65.000+65.000+60.000+60.000+65.000}{10} =$ 65.000,00

Centroid 2 :C(2,2)=

$$\frac{38.768,00+33.296,00+30.464,00+37.011,00+32.077,00+35.987,00+34.765,00+28640,00 +29.595,00+31.831,00}{10} = 33.243,30$$

Untuk *cluster* 2, ada 9 anggota yang tergabung ke dalamnya yakni 1, 5, 6, 9, 12, 14, 19, 20 dan 24 sehingga :

Centroid 1: $C(3,1)= \frac{75.000+75.000+75.000+75.000+75.000+75.000+75.000+75.000+75.000}{9} = 75.000,00$

Centroid 2 :$C(3,2)= \frac{65.087,00+71.407,00+65.342,00+65.573,00+66.196,00+72.987,00 +70.980,00+70.371,00+69.300,00}{9} = 68.582,56$

Setelah proses perhitungan di atas, maka akan diperoleh *centroid* baru dengan nilai sebagai berikut:

C1 = [48.571,00; 13.310, 71]
C2 = [65.000,00 ; 33.243,40]
C3 = [75.000,00 ; 68.582,56]

7. Setelah didapatkan *centroid* baru langkah selanjutnya kembali lagi ke langkah 3, yakni menghitung jarak setiap data ke masing-masing *centroid* menggunakan

GATOT KACA JOURNAL
Jurnal Keilmuan Teknik: Aplikasi, Perencanaan, Rancang Bangun, Sistem, Simulasi, & Rekayasa
e-ISSN : 2775-4561 (Online)
ISSN : 2775-2992(Print)

rumus korelasi antar dua obyek yaitu *Euclidean Distance* berdasarkan *centroid* baru.

8. Setelah menghitung jarak data ke *centroidnya*, maka langkah berikutnya adalah mengelompokkan data berdasarkan jarak minimumnya.

Karena pada iterasi ke-1 dan ke-2 posisi *cluster* tidak berubah lagi dan tidak ada data lagi yang berpindah dari satu *cluster* ke *cluster* yang lain, maka iterasi dihentikan dan hasil akhir yang diperoleh sebanyak 3 *cluster* dengan 2 iterasi.

Dari proses *clustering* atau pengelompokan di atas maka dihasilkan pengelompokan berdasarkan kedekatan jarak antar titik pusat dengan data produk PT. Coca Cola dari setiap atributnya. Adapun hasil dari proses ini dapat dilihat pada tabel 3 yaitu :

Tabel 3. Hasil Proses

Cluster	**Hasil Proses *Centroid* akhir**		**Anggota**
Cluster 0	48.571,43 13.310,71	;	Jumlah anggota = 7 anggota, yaitu: 1. Frestea Apel 200 ml 2. Fanta VIT C Lemon PET 500 ml 3. Diet Coke PET 1.5 ltr 4. Fanta Cream Can 250 ml 36 5. Sprite Ice Can 250 ml 9x4(36 6. Fanta Fruit Punch Can 250 ml X24 7. A & W Scanpack 4 x 6
Cluster 1	65.000,00 33.243,40	;	Jumlah anggota = 10 anggota, yaitu: 1. Green Tea 250 ml 2. Coke PET 1.25 ml 3. Coke Zero PET 1 Liter 4. Coke PET 500 ml 5. Coke PET 1.5 ltr 6. Soda Water Can 250 ml 36 7. Sprite Can 250 ml 9x4(36 8. F. Soda Water Can 250 ml X24 9. Coke Scanpack 4 x 6 10. Ades 600 ml
Cluster 2	75.000,00; 68.582,56		Jumlah anggota = 9 anggota, yaitu: 1. Frestea 220 ml 2. Fanta Strawbery PET 1.25 ml 3. Fanta Strawbery PET 1 Liter 4. Fanta Strawbery PET 500 ml 5. Fanta Strawbery PET 1.5 ltr 6. Coke Can 250 ml 36 7. Fanta Strawbery Can 250 ml 9x4(36 8. Sprite Can 250 ml X24 9. Fanta Strawbery Scanpack 4 x 6

GATOT KACA JOURNAL
Jurnal Keilmuan Teknik: Aplikasi, Perencanaan, Rancang Bangun, Sistem, Simulasi, & Rekayasa
e-ISSN : 2775-4561 (Online)
ISSN : 2775-2992(Print)

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil penelitian yang dilakukan, maka dapat ditarik kesimpulan bahwa berdasarkan hasil pengujian, maka didapatkan pengelompokan tingkat penjualan produk di PT. Coca Cola Distribution Indonesia sebanyak 3 *cluster.* Yaitu *cluster* 0 yaitu kelompok dengan tingkat penjualan produk rendah, *cluster* 1 *cluster* produk dengan tingkat penjualan sedang dan *cluster* 2 dengan tingkat penjualan produk tinggi. Dan produk yang masuk ke dalam masing-masing *cluster* tersebutlah yang menjadi acuan untuk jumlah produksi produk berikutnya.

Clustering yang dihasilkan digunakan untuk menentukan jumlah produksi produk. C*luster* produk yang memiliki tingkat penjualan tinggi memiliki jumlah produksi yang tinggi pula atau stabil seperti sebelumnya. Kemudian *cluster* produk dengan tingkat penjualan rendah, maka jumlah produksi produk untuk berikutnya dikurangi agar tidak terjadi penumpukan produk di gudang dan mengalami kadarluarsa. Begitupun *cluster* produk dengan tingkat penjualan sedang, jumlah produksinya disesuaikan dengan tingkat penjualan yang telah dicapai.

Saran

Pada penelitian berikutnya agar diadakan penelitian lebih mendalam terhadap pengelompokan data penjualan dengan metode lain supaya didapatkan hasil yang optimal dengan waktu penelitian yang lebih singkat.

Jumlah *cluster* dalam penelitian selanjutnya dapat dikembangkan menjadi beberapa *cluster*, sehingga lebih variatif.

e-ISSN : 2775-4561 (Online)
ISSN : 2775-2992(Print)

DAFTAR PUSTAKA

Adi Budiono, dkk. (2012). *"Analisis Faktor-faktor yang Mempengaruhi Produksi Jagung di Kecamatan Batu Ampar Kabupaten Tanah Laut."* *Jurnal Agribisnis Perdesaan. Volume 02 Nomor 02.*

Afrisawati. (2013). *"Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means."* Ed. *Pelita Informatika Budi Darma, Volume : V,No. 3.*

Bondu Venkateswarlu and Prof G.S.V.Prasad Raju. (2013). "*Mine Blood Donors Information through Improved K-Means Clustering*." *International Journal of Computational Science and Information Technology (IJCSITY) Vol.1, No.3.*

Budanis Dwi Meilani dan Nofi Susanti, (2014), *"Aplikasi Data Mining Untuk Menghasilkan Pola Kelulusan Siswa Dengan Metode Naïve Bayes."* Ed. *Jurnal LINK Vol 21/No.2.*

Deka Dwinavinta,et.al. (2014). *"Klasterisasi Judul Buku dengan Menggunakan Metode K-Means."* Ed. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI).*

Dr. M.P.S Bhatia and Deepika Khurana. (2013). "*Experimental study of Data clustering using k-Means and modified algorithms." International Journal of Data Mining & Knowlegde Management Process (IJDKP). Vol. 3 No.3.*

Er. Nikhil Chaturvedi and Er. Anand Rajavat. (2013). "*An Improvement in K-mean Clustering Algorithm Using Better Time and Accuracy." International Journal of Programming Language and Aplications (IJPLA) Vol. 3 No.4.*

Fadlina. (2014). *"Data Mining untuk Analisa Tingkat Kejahatan Jalanan Dengan Algoritma Association Rule Metode Apriori (Studi Kasus Di Polsekta Medan Sunggal)." Voumel.III No.1.*

Goldie Gunadi dan Dana Indra Sensuse, (2012). "*Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan MenggunakanAlgoritma Apriori dan Frequent Pattern Growth (FP-Grouth) : tudi Kasus Percetakan* PT. Gramedia." Ed. Jurnal Telematika MKom Vol.4 No.1.

Johan Oscar Ong. (2013). "*Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing Presiden University*." Ed. Jurnal Ilmiah Teknik Industri, Vol.12, No.1.

Mujib Ridwan, et.al. (2013). "*Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naïve Bayes Classifier*." Ed. Jurnal EECCIS Vol.7, No.1.

Nova Tumoka. (2013). "*Analisis Pendapatan Usaha Tani Tomat di Kecamatan Kawangkoan Barat Kabupaten Minahasa*. Jurnal EMBA. Vol.1 No.3

Ramesh Singh Yadava and P.K.Mishra. (2012). "*Performance Analysis of High Performance k-Mean Data Mining Algorithm for Multicore Heterogeneous Compute Cluster*." Vol. 2 No.4.

Ramzi A. Haraty, dkk. (2015). "*An Enhancedk-Means Clustering Algorithm for Pattern Discovery in Healthcare Data*." Hindawi Publishing Corporation International Journal of Distributed Sensor Networks Volume 2015, Article ID 615740.

Tutik Khotimah. (2014). "*Pengelompokan Surat Dalam Al Qur'an Menggunakan Algoritma K-Means*." Ed. Jurnal Simetris, Vol 5 No 1