

Segmentasi Pelanggan Berdasarkan Analisis RFM Menggunakan Algoritma K-Means Sebagai Dasar Strategi Pemasaran (Studi Kasus PT Coversuper Indonesia Global)

A T Widiyanto¹, A Witanti²

^{1,2}Program Studi Informatika, Fakultas Teknologi Informasi, Universitas Mercu Buana

E-mail : aguswidi636@gmail.com ¹, arita@mercubuana-yogya.ac.id²

Abstrak. PT. Coversuper Indonesia Global adalah perusahaan yang bergerak dibidang aksesoris otomotif. Persaingan dalam bisnis khususnya perusahaan yang bergerak dibidang aksesoris otomotif semakin banyak. Agar dapat meningkatkan penjualan produk yang dijual, para pelaku bisnis di bidang ini harus mempunyai strategi. Berdasarkan permasalahan tersebut maka digunakan *Data Mining* dengan teknik *cluster* untuk mengetahui potensi dan karakteristik dari setiap pelanggan mereka dalam melakukan pembelian produk. Penelitian ini menggunakan metode kombinasi model RFM dan clustering *K-Means* yang bertujuan untuk melakukan segmentasi pelanggan. Model RFM digunakan sebagai atribut kuantitatif untuk variabel masukan. Selanjutnya menggunakan algoritma K-means untuk melakukan clustering pelanggan. Hasil penelitian ini adalah 4 segmen pelanggan yang dimiliki oleh perusahaan dan karakteristik masing-masing pelanggan. Berdasarkan akurasi yang didapat pada perhitungan menggunakan sistem segmentasi pelanggan dari 29 kali pengujian, terdapat 29 pelanggan (100%) dengan karakteristik yang dihasilkan sistem sesuai dengan pengetahuan user, 0 pelanggan (0%) tidak sesuai dengan pengetahuan user. Dengan nilai akurasi tersebut maka dapat dikategorikan baik. Karakteristik pelanggan ini akan membantu PT Coversuper Indonesia Global untuk mengambil keputusan dalam memprioritaskan tenaga dan sumber dayanya ke pelanggan tertentu (potensial).

Kata kunci: *Algoritma K-means, Data Mining, Toko Aksesoris Otomotif.*

Abstract. PT. Coversuper Indonesia Global is a company engaged in automotive accessories. Competition in business, especially companies engaged in automotive accessories, is increasing. In order to increase sales of products being sold, business people in this field must have a strategy. Based on these problems, Data Mining with cluster techniques is used to determine the potential and characteristics of each of their customers in making product purchases. This study uses a combination method of RFM model and K-Means clustering which aims to perform customer segmentation. The RFM model is used as a quantitative attribute for the input variables. Furthermore, using the K-means algorithm to perform customer clustering. The results of this study are 4 customer segments owned by the company and the characteristics of each customer. Based on the accuracy obtained in the calculation using the customer segmentation system from 29 times the test, there are 29 customers (100%) with the characteristics generated by the system according to user knowledge, 0 customers (0%) do not match the user's knowledge. With this accuracy value, it can be categorized as good. These customer characteristics will help PT Coversuper Indonesia Global to make decisions in prioritizing its manpower and resources to certain (potential) customers.

Keywords: *K-means Algorithm, Data Mining, Automotive Accessory Shop.*

1. Pendahuluan

Cover Super adalah perusahaan yang bergerak di bidang aksesoris otomotif. Cover Super memproduksi produk perlindungan dari bidang otomotif hingga kereta, kapal laut dan pesawat terbang. Persaingan dalam bisnis khususnya perusahaan yang bergerak dibidang aksesoris otomotif semakin banyak. Agar dapat meningkatkan penjualan produk yang dijual, para pelaku bisnis di bidang ini harus mempunyai strategi yang tepat guna mengambil langkah ke depannya. Salah satu cara yang bisa dilakukan oleh perusahaan dengan cara memanfaatkan data dari transaksi penjualan yang ada. Data dari penjualan produk atau data transaksi yang dilakukan oleh perusahaan tersebut lambat laun akan menghasilkan tumpukan data. Sehingga sangat disayangkan jika data transaksi tersebut tidak di analisa kembali. Berdasarkan data penjualan pada 2020, ada sekitar 1000 lebih data transaksi hanya digunakan untuk rekap penjualan dan arsip pada perusahaan tersebut. Data transaksi yang digunakan tersebut, data dapat diolah lebih lanjut sehingga didapatkan informasi baru dari kumpulan data yang ada.

Cover Super sebagai perusahaan dengan kegiatan utamanya sebagai produsen cover atau penutup kendaraan di Indonesia. Setelah berkecimpung di bidang aksesoris otomotif sejak tahun 2012, saat ini Cover Super mampu menangani lebih dari 50 transaksi perhari. Dalam menghadapi persaingan dan agar mampu mempertahankan eksistensinya di era pemasaran *modern* saat ini, yang harus menjadi fokus utama perusahaan adalah pelanggan. Pelanggan adalah berupa perorangan, kelompok, instansi, lembaga, atau organisasi yang membeli produk perusahaan secara rutin karena produk yang dibelinya dirasa sangat bermanfaat. Dapat dikatakan juga pelanggan, yaitu orang-orang yang memberikan perhatian penuh terhadap produk yang dihasilkan oleh perusahaan.

Strategi setiap pelanggan dalam melakukan pembelian berbeda. Ada pelanggan antara yang melakukan pembelian dalam jumlah besar karena memiliki dana lebih atau menyimpan stok barang. Ada juga pelanggan yang melakukan pembelian satu kali dalam tiga bulan langsung dalam jumlah besar dan ada yang melakukan pembelian bertahap dalam setiap enam bulannya. Karena setiap pelanggan memiliki karakteristik yang berbeda-beda itulah yang terkadang menyulitkan perusahaan dalam melakukan *maintenance* pelanggan dan memasarkan produk mereka. Oleh karena itu penting bagi perusahaan untuk mengetahui potensi dan karakteristik dari setiap pelanggan mereka dalam melakukan pembelian produk. Dalam kasus ini segmentasi pelanggan diperlukan perusahaan untuk lebih memahami karakteristik dari pelanggan mereka yang berbeda-beda, dengan demikian perusahaan bisa merencanakan pendekatan yang tepat terhadap setiap segmen pelanggan berbeda.

Beberapa penelitian telah dikembangkan untuk menentukan model dan metode yang terbaik dalam melakukan segmentasi pelanggan. Beberapa penelitian terdahulu yang pernah dikembangkan menyebutkan bahwa model yang banyak digunakan dan memiliki tingkat akurasi yang tinggi adalah *Recency, Frequency, Monetary* (RFM). Model ini merupakan proses penilaian berdasarkan perilaku pelanggan yang akan dilihat dari waktu transaksi terakhir pelanggan (*Recency*), jumlah transaksi (*Frequency*), dan uang yang dikeluarkan (*Monetary*). Dengan mempelajari rekaman interaksi dan transaksi *customers* di masa lalu, perusahaan dapat melakukan penilaian terhadap 3 (tiga) aspek dimensi tersebut. Model RFM diaplikasikan secara luas pada database pemasaran dan merupakan *tool* yang umum digunakan untuk membangun strategi pemasaran. RFM berdasarkan segmentasi pelanggan menghasilkan kemampuan segmentasi antara 75% sampai dengan 85%. Dalam penelitian ini dilakukan segmentasi pelanggan dari dataset retail berdasarkan karakteristik masing masing pelanggan pada perusahaan dan menganalisis segmentasi untuk diketahui pelanggan yang profitable dan sebaliknya. Teknik *data mining* yang digunakan untuk segmentasi pelanggan adalah dengan menggunakan teknik *clustering* dengan algoritma *K-Means*. *K-Means* yaitu metode data *non-hierarchical clustering* yang digunakan untuk mengelompokkan data ke dalam *cluster* berdasarkan kemiripan dari data satu dengan yang lain, jadi data dengan karakteristik yang sama akan dikelompokkan dalam satu *cluster* dan data dengan karakteristik berbeda akan dikelompokkan pada *cluster* lain yang memiliki karakteristik sama dengan data tersebut. Alasan menggunakan algoritma *K-Means* adalah karena metode interaktif yang mudah untuk diterapkan dan bersifat dinamis pada data yang tersebar serta yang terpenting, hasil dari segmentasi yang dihasilkan lebih akurat. Maka penelitian ini akan dikembangkan dengan menggunakan kombinasi analisis RFM, *clustering K-Means* yang dikembangkan oleh [1]. Model ini menggunakan analisa RFM (*Recency, Frequency, Monetary*) sebagai atribut kuantitatif untuk menjadi *input*, Selanjutnya menggunakan algoritma *k-means* untuk *clustering* nilai dari *customer*.

Diharapkan dengan adanya penelitian ini dapat membantu perusahaan untuk mengelompokkan pelanggan serta membantu memprioritaskan tenaga dan sumber dayanya ke pelanggan tertentu (potensial).

2. Tujuan

Tujuan dari penelitian ini adalah menghasilkan sebuah aplikasi segmentasi pelanggan menggunakan kombinasi metode *K-Means* dan model RFM yang dapat membantu perusahaan dalam menemukan karakteristik dari setiap pelanggan agar perusahaan mampu memprioritaskan sumber daya dan tenaga untuk pelanggan tertentu. Menganalisis akurasi dari *cluster* yang sudah terbentuk yang akan digunakan menjadi bahan untuk menarik kesimpulan dari penelitian ini sehingga bisa memberikan solusi kepada perusahaan.

3. Tinjauan Pustaka

Sebagai tinjauan pustaka berikut ini beberapa contoh penelitian yang sudah dilakukan oleh para peneliti yang dapat digunakan sebagai acuan dan pengetahuan.

Penelitian yang mendukung penggunaan Metode Fuzzy C-Means adalah penelitian segmentasi pada tesis dari Yohana Nugrahaeni yang menggunakan Metode Fuzzy pada penelitian yang berjudul “*Data Mining* dengan Metode *Fuzzy* untuk *Customer Relationship Management* pada Perusahaan Retail”. Penelitian tersebut menekankan pada segmentasi pelanggan dengan menggunakan Metode Fuzzy dengan objek yang diteliti adalah sebuah Perusahaan Retail UD. Fenny. Penelitian ini membahas tentang Proses *Data Mining* dari data konsumen dan menggunakan data yang berasal dari data penjualan atau transaksi, tujuan dari penelitian pada jurnal ini adalah untuk mencari konsumen potensial [2].

Penelitian dalam judul “Segmentasi Pelanggan Menggunakan Algoritma K-Means Sebagai Dasar Strategi Pemasaran pada LAROIBA Seluler”. Penelitian ini bertujuan untuk melakukan segmentasi pelanggan pada LAROIBA Seluler dengan memanfaatkan proses data mining dengan teknik Clustering. Metode yang digunakan adalah CRISP-DM dengan melalui proses business understanding, data understanding, data preparation, modeling dan evaluasi. Algoritma yang digunakan untuk pembentukan kluster adalah algoritma K-Means. K-Means menghasilkan model *visual cluster* dengan *tools* Rapidminer 5.2 yang merepresentasikan jumlah pelanggan di setiap kluster dengan menggunakan atribut RFM (*Recency, Frequency dan Monetary*). Setelah dilakukan segmentasi data pelanggan LAROIBA Seluler dengan menggunakan algoritma *K-Means* terbentuk empat *cluster* yang memenuhi kelas pada Piramida Pelanggan yaitu *Most Valuable Customer* ada 4 pelanggan, *Most Growable Customer* ada 76 pelanggan, *Below Zero* 20 pelanggan dan *Migrators* ada 173 pelanggan [3].

Penelitian dalam judul “Segmentasi Pelanggan Menggunakan Model Rfm Dan Teori Rough Set Untuk Memahami Karakteristik Pelanggan (Studi Kasus: Pt.Abbott Indonesia,Tbk Cabang Malang)”. Penelitian ini bertujuan menemukan karakteristik dari setiap pelanggan agar perusahaan mampu memprioritaskan sumber daya dan tenaga untuk pelanggan tertentu. Metode menggunakan kombinasi model RFM, klustering *K-Means* dan Teori *Rough Set* untuk melakukan segmentasi pelanggan. Model RFM digunakan sebagai atribut kuantitatif untuk variabel masukan. Selanjutnya menggunakan algoritma K-means untuk melakukan klustering pelanggan dan terakhir menghasilkan aturan klasifikasi dengan teori *Rough Set*. Setelah dilakukan segmentasi data pelanggan PT Abbott dengan algoritms K-means terbentuk 4 segmen pelanggan yang dimiliki oleh perusahaan dan karakteristik masing-masing pelanggan. Karakteristik pelanggan ini akan membantu PT Abbott Indonesia cabang Malang untuk mengambil keputusan dalam memprioritaskan tenaga dan sumber dayanya ke pelanggan tertentu (potensial) [4].

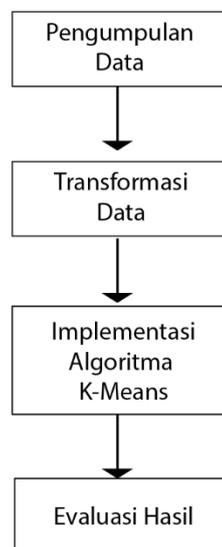
Penelitian dalam judul “Penerapan Clustering K-Means Pada Customer Segmentation Berbasis Recency Frequency Monetary (Rfm) (Studi Kasus : Pt. Sinar Kencana Intermoda Surabaya)”. Penelitian ini bertujuan untuk penerapan algoritma *k-Means* untuk membentuk segmentasi pelanggan, sehingga dapat mengetahui tingkat potensial pelanggan. Analisis data dilakukan dengan dua cara yaitu pembobotan RFM untuk menghasilkan bobot RFM dimana *recency* merupakan transaksi terakhir, *frequency* merupakan jumlah transaksi yang dilakukan dan *monetary* yang merupakan total biaya transaksi yang telah dilakukan. Setelah itu barulah dikelompokkan dengan metode k-Means. Hasil evaluasi sistem didapatkan bahwa pengelompokkan pelanggan pada

katategori BZ memiliki nilai presentase 54,3%, kategori MVC sebesar 21,8% dan MGC sebesar 23,9% [5].

Penelitian dalam judul “Segmentasi Dan Evaluasi Loyalitas Pelanggan Distributor Produk Etikal Farmasi Berdasarkan Nilai Pelanggan”. Penelitian ini bertujuan untuk menganalisa pelanggan yang potensial dengan membuat sebuah sistem aplikasi menggunakan algoritma K-means untuk mengelompokkan data pelanggan potensial berbasis web. Metode K-Means Clustering mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama. Data pelanggan dikelompokkan menjadi dua cluster yaitu tidak potensial, dan potensial. Kemudian setiap *cluster* diklasifikasikan berdasarkan kriteria mana yang lebih diprioritaskan. Hasil dari proses ini membentuk *cluster-cluster* yang digunakan untuk memberi saran pertimbangan dalam menentukan strategi penjualan yakni memberikan reward untuk pelanggan dengan posisi *cluster* teratas [6].

4. Metodologi Penelitian

Penelitian ini dilakukan dengan melewati empat tahapan diantaranya : pengumpulan data, transformasi data, implementasi algoritma *k-means* dan evaluasi hasil. Metodologi yang dilakukan, ditunjukkan pada Gambar 3.1.



Gambar 1. Tahapan Penelitian

5. Pembahasan

Pada penelitian ini, data yang di pergunakan adalah data satu tahun terakhir lebih tepatnya tanggal 01 Desember 2020 – 21 Desember 2020 sejumlah 990 data mentah. Setelah melakukan proses *data pre-processing (data cleaning)* dilakukan perbaikan ditemukan kesalahan penulisan, *duplikasi* data atau inkonsistensi pada data transaksi sehingga data yang di dapatkan adalah 797 data transaksi *clean* setelah di seleksi dan yang dipakai untuk pengujian data adalah 736 data pelanggan, dengan format *upload* data bertipe .CSV agar dapat dibaca oleh sistem yang sudah dirancang. Penelitian ini menghasilkan hasil analisis berupa tabel yang berisi kombinasi dari nilai R, F dan M. Model ini membedakan pelanggan yang penting dari sejumlah data dengan menggunakan tiga atribut yaitu: selang waktu (interval) pemakaian pelanggan, frekuensi dan jumlah uang. Nilai RFM dari semua atribut item dapat kita peroleh dari kriteria atribut RFM dengan 5,4,3,2,1 yang merepresentasikan kontribusi pelanggan terhadap keuntungan perusahaan. Nilai 5 menunjukkan sangat berkontribusi dan 1 menunjukkan sedikit kontribusi. Urutkan data pada ketiga atribut mulai dari yang paling besar hingga yang terkecil. Daftar hasil perhitungan nilai RFM dari tiap atribut seperti terlihat pada Tabel 1

Tabel 1. RFM Scoring

NAMA	Recency	Frequency	Monetary
Al ali	4	3	1
Tsakif	4	3	3
Titin juhari	4	3	2
M. Shalihin	4	3	3
Aryudan Biphy Gunenda	4	3	2
Zulfiandi	4	3	1
Dora	4	3	3
Guntur Widyaswara	4	3	2
Siti Rohmawati	4	3	1
Dewa suryadika	4	3	1
wiwit eka mardiyanti	4	3	1
Yoga DC	4	3	3
Anggu	4	3	1
Mokhamad Carnegie			
Trihandono	4	3	1
David Mahendra	4	3	1
Ahmad Fuad	4	3	3
Haikal Gilang Ramadhan	4	3	1
Bernard Sagala	4	5	5
Cyrillus Manek	4	3	2

5.1. Clustering menggunakan algoritma K-means

Penelitian ini menggunakan metode *k-means* dalam pengolahan data transaksi yang dilakukan oleh pelanggan. penghitungan algoritma ini terdiri dari inialisasi data berupa data keanggotaan, menghitung *centroid*, menghitung jarak data ke *centroid*, menghitung nilai keanggotaan, dan menghitung nilai fungsi objektif. Hasil dari aplikasi ini berupa nilai fungsi objektif dan jumlah iterasi, dari nilai-nilai tersebut *user* dapat melihat kelompok pelanggan menurut jenis-jenis karakteristiknya.

Pertama menentukan jumlah *cluster* yang optimal untuk melakukan proses *clustering* dengan algoritma *k-means*. Kita akan melakukan proses *clustering* dengan algoritma *k-means* dengan nilai $k = 4$ (jumlah *cluster* optimal yang ditentukan). Setelah didapatkan jumlah *cluster*, menentukan *Initial cluster center* merupakan letak pusat *cluster (centroid)* pada tahap awal *clustering*. Untuk hasil *initial cluster center* bisa dilihat pada Tabel 2.

Tabel 2. Hasil *Initial Cluster Center*

Centroid	Recency	Frequency	Monetary
1	4	3	1
2	4	3	2
3	4	3	3
4	4	5	5

Dalam proses *clustering* titik awal tersebut akan terus berganti hingga mencapai titik yang paling optimal yang dinamakan *final cluster center* yang bisa kita lihat pada Tabel 3.

Tabel 3. Hasil *Final Cluster Center*

Centroid	Recency	Frequency	Monetary
1	3.5221	3.0354	1.2522
2	3.0936	2.99	3.4448
3	2.84	5	4.64
4	1.5645	2.9677	1.4516

Proses iterasi yang berlangsung hingga tidak ada lagi kedudukan dari anggota *cluster* yang berubah terhadap titik pusat *cluster* (*centroid*). Sementara itu untuk melihat jarak antara titik pusat satu *cluster* dengan titik pusat *cluster* lainnya bisa dilihat pada Tabel 4 yang menunjukkan jarak antar *final cluster center*. Jika dilihat pada besarnya jarak dalam tabel maka bisa kita simpulkan bahwa kedekatan antara *cluster* yang terbentuk adalah *cluster* 1-4-2-3. informasi jarak ini akan digunakan untuk membantu dalam membuat keputusan tingkat *loyalty* pelanggan yang akan dilakukan pihak perusahaan.

Tabel 4. Hasil Jarak Antara *Final Cluster Center*

Cluster	1	2	3	4
1	0.5415	2.6074	4.3122	2.4772
2	1.8245	0.4547	2.5914	2.1117
3	4.2636	2.5431	0.394	4.3338
4	1.5433	2.6783	4.2374	0.6282

Hasil akhir dari proses data mining menghasilkan total fungsi objektif, perubahan fungsi objektif, jumlah iterasi, dan hasil *clustering* pelanggan.

Tabel 5. Hasil Jumlah Kasus Setiap *Cluster*

Cluster	Jumlah	Valid	Missing
1	226		
2	299		
3	25	736	0
4	186		

Dari tabel tersebut bisa kita lihat bahwa *cluster* 1 terdiri atas 226 pelanggan, *cluster* 2 terdiri atas 299 pelanggan, *cluster* 4 terdiri atas 186 pelanggan dan *cluster* 3 memiliki jumlah anggota *cluster* yang paling sedikit yaitu 25 pelanggan dari total pelanggan sebanyak 736 pelanggan. Kesimpulan dari informasi yang didapat dari tabel ini adalah jika diurutkan dari yang memiliki anggota *cluster* paling banyak hingga paling sedikit adalah *cluster* 2-1-4-3.

5.2. Analisis antar atribut RFM

Analisis hasil RFM juga digunakan untuk menentukan karakteristik dari setiap pelanggan perusahaan. Dalam penelitian ini akan dibagi ke dalam tiap analisis antar atribut. Untuk analisis antar atribut *recency* dan *frequency* bisa dilihat pada Tabel 6.

Tabel 6. Analisis antar Atribut *Recency* dan *Frequency*

<i>Cluster</i>	Hasil Analisis
1	Cluster 1 memiliki rentang <i>recency</i> antara 16-21 hari yang berarti memiliki nilai <i>recency</i> jauh . Sementara itu nilai <i>frequency</i> juga hanya 1 kali sehingga bernilai sangat rendah . Dengan demikian maka pelanggan yang berada pada <i>cluster</i> 1 kemungkinan adalah para pelanggan yang sangat jarang melakukan transaksi.
2	<i>Cluster</i> 2 memiliki nilai <i>recency</i> yang berkisar antara 11-15 hari yang berarti nilainya biasa . Nilai <i>frequency</i> dari <i>cluster</i> ini biasa karena melakukan transaksi antara 1-3 kali. Kemungkinan pelanggan pada <i>cluster</i> 2 melakukan transaksi sekali setiap akhir bulan.
3	Pada <i>cluster</i> 3 nilai <i>recency</i> adalah sangat pendek karena nilainya yang berkisar kurang dari 10 hari. Disamping itu nilai <i>frequency</i> dari <i>cluster</i> 3 adalah tinggi dengan kisaran transaksi sebanyak 2-20 kali. Dari hasil tersebut maka pelanggan yang berada pada <i>cluster</i> 3 merupakan pelanggan yang rutin melakukan transaksi dengan perusahaan karena pada bulan terakhir masih melakukan transaksi dan frekuensinya pun sering.
4	Nilai <i>recency</i> dari <i>Cluster</i> 4 berada pada rentang 6-10 hari yang berarti nilai <i>recency</i> pendek . Namun untuk nilai <i>frequency</i> hanya berkisar 1-2 kali melakukan transaksi yang berarti nilai <i>frequency</i> dari <i>cluster</i> 4 adalah rendah . Hal ini menunjukkan bahwa pelanggan yang berada pada <i>cluster</i> 4 kemungkinan melakukan transaksi hanya pada bulan-bulan terakhir sehingga nilai <i>recency</i> mereka tidak terlalu jauh.

Selanjutnya analisis akan dilanjutkan dengan melihat antara nilai *recency* dan *monetary*. Hasil analisis bisa dilihat pada Tabel 7.

Tabel 7. Analisis antar Atribut *Recency* dan *Monetary*

<i>Cluster</i>	Hasil Analisis
1	<i>Cluster</i> 1 memiliki rentang <i>recency</i> antara 16-21 hari yang berarti memiliki nilai <i>recency</i> jauh . Sementara itu nilai <i>monetary</i> sangat sedikit dengan nominal kurang dari Rp 549.000,-. Kemungkinan para pelanggan yang berada pada <i>cluster</i> 1 ini merupakan pelanggan yang jarang melakukan transaksi serta dalam skala kecil.
2	<i>Cluster</i> 2 memiliki nilai <i>recency</i> yang berkisar antara 11-15 hari yang berarti nilainya biasa . Nilai <i>monetary</i> dari <i>cluster</i> ini normal karena nilai pembeliannya berkisar antara Rp 749.000 – Rp 889.000. Berdasarkan hasil ini kemungkinan para pelanggan yang ada pada <i>cluster</i> 2 melakukan transaksi dalam jumlah besar pada beberapa bulan terakhir.

Cluster	Hasil Analisis
3	Pada <i>cluster</i> 3 nilai <i>recency</i> adalah sangat pendek karena nilainya yang berkisar kurang dari 10 hari. Disamping itu nilai <i>monetary</i> dari <i>cluster</i> 3 adalah sangat banyak dengan kisaran lebih besar dari Rp Rp 1.499.000. Dari hasil tersebut maka pelanggan yang berada pada <i>cluster</i> 3 merupakan pelanggan yang melakukan transaksi dalam jumlah sangat besar pada bulan-bulan terakhir dalam kurun waktu tertentu.
4	Nilai <i>recency</i> dari <i>Cluster</i> 4 berada pada rentang 6-10 hari yang berarti nilai <i>recency</i> pendek . sementara itu <i>monetary</i> berkisar antara Rp 549.000-Rp 749.000 sehingga memiliki nilai <i>monetary</i> yang Sedikit. Hal ini menunjukkan bahwa pelanggan yang berada pada <i>cluster</i> 4 kemungkinan melakukan transaksi pada bulan terakhir dengan jumlah transaksi yang lumayan besar.

Untuk selanjutnya akan dianalisis hasil cluster berdasarkan atribut *frequency* dan *monetary*. Hasil analisisnya bisa dilihat pada Tabel 8.

Tabel 8 Analisis antar Atribut *Frequency* dan *Monetary*

Cluster	Hasil Analisis
1	<i>Cluster</i> 1 memiliki nilai <i>frequency</i> hanya 1 kali sehingga bernilai sangat rendah . Sementara itu nilai <i>monetary</i> sangat sedikit dengan nominal kurang dari Rp 549.000,-. Kemungkinan para para pelanggan yang berada pada <i>cluster</i> 1 ini merupakan pelanggan yang jarang melakukan transaksi bahkan dalam kurun waktu 6 bulan serta dalam skala kecil. Pelanggan yang masuk kedalam <i>cluster</i> ini bisa dikatakan bukan pelanggan setia atau kemungkinan besar pelanggan baru di tahun 2020.
2	<i>Cluster</i> 2 memiliki nilai <i>frequency</i> biasa karena melakukan transaksi antara 1-3 kali. Nilai <i>monetary</i> dari <i>cluster</i> ini normal karena nilai pembeliannya berkisar antara Rp 749.000 – Rp 889.000. Berdasarkan hasil ini kemungkinan para pelanggan yang ada pada <i>cluster</i> 2 tidak terlalu sering melakukan transaksi dalam kurun waktu 6 bulan terakhir dan juga tidak melakukan transaksi dalam skala besar namun juga tidak terlalu kecil.
3	Pada <i>cluster</i> 3 nilai <i>frequency</i> adalah tinggi dengan kisaran transaksi sebanyak 2-20 kali. Didukung oleh nilai <i>monetary</i> yang sangat banyak dengan kisaran nominal transaksi lebih besar dari Rp Rp 1.499.000. Dari hasil tersebut maka pelanggan yang berada pada <i>cluster</i> 3 merupakan pelanggan dengan tingkat loyalitas yang tinggi karena sering melakukan transaksi dengan perusahaan serta dalam jumlah yang cukup besar.

Cluster	Hasil Analisis
4	Nilai <i>frequency cluster</i> 4 berkisar antara 1-2, sehingga nilai <i>frequency</i> dari <i>cluster</i> 4 adalah rendah . Sementara itu <i>monetary</i> berkisar antara Rp 549.000 - Rp 749.000 sehingga memiliki nilai <i>monetary</i> yang Sedikit . Hal ini menunjukkan bahwa pelanggan yang berada pada cluster 4 kemungkinan adalah pelanggan yang melakukan transaksi dalam jumlah lumayan besar dalam satu waktu tertentu yang mengakibatkan frekuensi belanja pelanggan tersebut tidak terlalu banyak.

5.3. Menemukan karakteristik pelanggan

Setelah melakukan analisis antar atribut RFM maka selanjutnya yang perlu dilakukan adalah menunjukkan hasil analisis kepada pihak perusahaan. Hal ini bertujuan untuk menyamakan persepsi dengan perusahaan sehingga bisa meminimalisir adanya kesalahan penentuan karakteristik. Setelah itu jika karakteristik pelanggan sudah ditemukan maka akan diberikan label/nama untuk setiap *cluster* untuk mempermudah pihak perusahaan dalam mengingat segmen pelanggan mereka. Pemberian nama/label disesuaikan dengan karakteristik pelanggan dan juga dengan persetujuan pihak perusahaan. Seluruh hasil analisis dan diskusi dengan pihak perusahaan akan digabung sehingga menghasilkan karakteristik setiap segmen pelanggan yang bisa dilihat pada tabel 4.9.

Tabel 4.9. Karakteristik pelanggan

Cluster	Karateristik	Label
1	<ol style="list-style-type: none"> 1. Para pelanggan yang sangat jarang melakukan transaksi dalam 6 bulan terakhir terutama pada bulan-bulan terakhir. 2. Pelanggan hanya memiliki jumlah transaksi Rp. 549.000 ke bawah. 3. Pelanggan dengan tingkat loyalitas Rendah. 4. Kemungkinan besar pelanggan baru di tahun 2020. 	<i>Consumers</i>
4	<ol style="list-style-type: none"> 1. Pelanggan yang melakukan pembelian 1-2 kali namun dengan nominal transaksi antara Rp 549.000-Rp 749.000 di waktu tertentu. 2. Pelanggan dengan tingkat loyalitas biasa. 	<i>Ordinary</i>
2	<ol style="list-style-type: none"> 1. Merupakan Pelanggan yang jarang melakukan transaksi tetapi membeli dalam jumlah besar pada waktu tertentu pada kurun waktu tertentu. 2. Pelanggan yang melakukan pembelian 1-3 kali. 3. Pelanggan memiliki jumlah transaksi berkisar antara Rp 749.000 – Rp 889.000. 4. Pelanggan dengan tingkat loyalitas Tinggi. 	<i>Big Consumers</i>

Cluster	Karakteristik	Label
3	<ol style="list-style-type: none"> 1. Pelanggan yang berada di cluster 4 adalah pelanggan dengan tingkat sangat tinggi. 2. Tipe pelanggan seperti ini sebagian besar berasal dari Agen. 3. Merupakan pelanggan yang rutin melakukan pembelian dalam jumlah besar pada minggu terakhir setiap bulannya. 4. Pelanggan yang melakukan pembelian 2-20 kali. 5. Pelanggan memiliki jumlah transaksi diatas Rp 1.400.000. 	Top Class

Berdasarkan hasil yang terlihat pada tabel 4.9 maka didapatkan karakteristik dan nama dari keempat segmen pelanggan yang didapatkan dari proses *clustering*.

5.4. Validasi hasil clustering

Pada tahap ini dilakukan wawancara langsung terhadap *user*. Terdapat *user*/perwakilan dari tim *customer representative* perusahaan yang terlibat pada proses UAT ini. Profil dari *user* antara lain, *user* berusia 28 tahun dan telah menjadi pegawai PT.Coversuper selama 7 tahun. Pertanyaan meliputi ketepatan karakteristik pelanggan dari hasil penelitian dengan keadaan di lapangan. Selama proses wawancara *user* akan diberikan sampel pelanggan secara acak dari setiap label untuk dinilai apakah karakter pelanggan tersebut sesuai dengan karakter pada lapangan. setiap jawaban *user* akan dicatat apakah karakter pelanggan tepat dengan kondisi lapangan, tidak tepat ataupun pelanggan yang karakteristiknya sama sekali tidak dikenali. Hasil UAT penelitian ini disajikan dalam Table 10.

Tabel 10. Hasil *User Acceptance Testing*

Label	Kode Pelanggan	Nama Pelanggan	Jawaban
Consumers	011220-RG-E-12365	Al ali	Benar
	011220-RG-E-12370	Zulfiandi	Benar
	011220-RG-E-12372	Guntur Widyaswara	Benar
	011220-RG-E-12376	Yoga DC	Benar
	011220-RG-E-12378	Mokhamad Carnegie Trihandono	Benar
	011220-TP-E-12389	Arif Triman	Benar
	021220-SP-E-12399	Sadiyyah mahardika	Benar
	021220-RG-E-12403	Moh Fauzan	Benar
	021220-RG-E-12410	Mama Putri	Benar
	021220-SP-E-12413	Irawan	Benar
Ordinary	161220-RG-E-12958	David H	Benar
	161220-RG-E-12961	Muhamad Miproh	Benar
	161220-RG-E-12962	Sultan Allaudin Parawansa	Benar
	161220-TP-E-12967	agus setiawan	Benar
	161220-RG-E-12969	Taufiq Miswar	Benar
	161220-SP-E-12976	Yana	Benar
	161220-RG-E-12977	Erni Dwi S	Benar
	171220-SA-E-12979	Fentarie Gumilang	Benar

Label	Kode Pelanggan	Nama Pelanggan	Jawaban
Big Company	211220-RG-E-13108	Sony Erlawan	Benar
	211220-RG-E-13110	Azzam	Benar
	211220-RG-E-13129	Audi Motor	Benar
	211220-SM-E-13075	Ardian mohammad iqbal	Benar
	211220-RG-E-13078	Adhi	Benar
	211220-TP-E-13082	Eriel F N	Benar
Top Class	211220-RG-E-13101	Robinson	Benar
	211220-SP-E-13106	Irwan Jodi Setiawan	Benar
	211220-RG-E-13109	M. Yasin Mustopa	Benar
	011220-RG-E-12368	Bernard Sagala	Benar
	041220-RG-A-12513	Coversuper Store	Benar

Berdasarkan hasil UAT dari 29 kali pengujian, terdapat 29 pelanggan (100%) dengan karakteristik yang dihasilkan sistem sesuai dengan pengetahuan user dan 0 pelanggan (0%) tidak sesuai dengan pengetahuan *user*. Sehingga sistem aplikasi ini sudah memberikan hasil informasi yang akurat.

6. Kesimpulan

Berdasarkan pembahasan yang telah diuraikan dalam penelitian ini, maka dapat diambil kesimpulan yaitu :

Hasil aplikasi segmentasi pelanggan dapat melakukan proses segmentasi menggunakan model RFM untuk mengetahui karakteristik dari pelanggan dan metode K-means yang digunakan untuk melakukan segmentasi kepada pelanggan. Berdasarkan hasil *clustering* dengan metode *K-Means* jumlah segmen pelanggan yang dimiliki oleh PT.Coversuper Indonesia Global adalah 4 *cluster*/segmen pelanggan dari keseluruhan jumlah pelanggan sebanyak 736. Hasil olah data pelanggan menggunakan algoritma *K-Means* dengan metode *clustering* dan dengan menggunakan atribut RFM diimplementasikan ke dalam proses pengelompokan pelanggan potensial pada perusahaan yang menghasilkan kelompok pelanggan *Consumers* 226, *Ordinary* 186, *Big Consumers* 299 dan *Top Class* 25. Berdasarkan hasil *User Acceptance Testing* dari 29 kali pengujian, terdapat 29 pelanggan (100%) dengan karakteristik yang dihasilkan sistem sesuai dengan pengetahuan *user*, 0 pelanggan (0%) tidak sesuai dengan pengetahuan *user*. Dengan nilai akurasi tersebut maka sistem aplikasi ini sudah memberikan hasil informasi yang baik bagi perusahaan.

7. Referensi

- [1] C.-H. Cheng and Y.-S. Chen, "Classifying the segmentation of customer value via RFM model and RS theory," *Expert Systems with Applications*, p. 4176–4184, 2009.
- [2] Y. Nugraheni, "Tesis data mining dengan metode fuzzy untuk Customer relationship management (CRM) pada perusahaan retail," 2011.
- [3] M. I. Istiana, "Segmentasi Pelanggan Menggunakan Algoritma K-Means Sebagai Dasar Strategi Pemasaran pada LAROIBA Seluler," 2013.
- [4] A. Husna, "Segmentasi Pelanggan Menggunakan Model RFM Dan Teori Roughset Untuk Memahami Karakteristik Pelanggan (STUDI KASUS: PT.ABBOTT INDONESIA, Tbk Cabang Malang)," 2015.
- [5] R. D. F. Ruli, "Penerapan CLustering K-means Pada Customer Segmentation Berbasis Recency Frequency Monetary RFM (Studi Kasus : Pt. Sinar Kencana Intermoda Surabaya)," *Seminar Nasional Matematika dan Aplikasinya*, pp. 418-427, 2017.
- [6] S. Rudiarto, "Implementasi Algoritma K-means Clustering Pada Aplikasi Pencari Pelanggan Potensial Pada Restoran XYZ," *Jurnal Ilmu Teknik Dan Komputer*, 2018.
- [7] R. Nainggolan, R. Perangin-angin, E. Simarmata and F. A. Tarigan, "Improved the Performance of the K-Means Cluster Using the," *Journal of Physics: Conference Series*, 2019.

- [8] A. V. Angelie, "Segmentasi Pelanggan Menggunakan Clustering K-Means dan Model RFM (Studi Kasus: PT. Bina Adidaya Surabaya)," 2017.
- [9] E. Ngai, L. Xiu and D. Chau, "Application of data mining techniques in customer relationship management:," *Expert Systems with Applications*, p. 2592–2602, 2009.
- [10] B. Setyobudi, "Application of Segmentation in Determining Policy Analysis Marketing Strategy," vol. 2, pp. 124-132, 2011.
- [11] B. Shim, K. Choi and Y. Suh, "CRM strategies for a small-sized online shopping mall based on association rules," *Expert Systems with Applications*, pp. 7736-7742, 2012.
- [12] D. Peppers and M. Rogers, *Managing Customer relationships*, New Jersey: John Wiley & Sons, Inc, 2011.
- [13] A. M. Hughes, in *Strategic Database marketing*, Cambridge, Probus Publising, 1994, p. 300.
- [14] J. Tabak, *Geometry: The Language of Space and Form*, New York: Facts On File, Inc, 2014.
- [15] B. E. Adiana, i. Soesanti and A. E. Permanasari, "ANALISIS SEGMENTASI PELANGGAN MENGGUNAKAN KOMBINASI RFM MODEL DAN TEKNIK CLUSTERING," *Jurnal Terapan Teknologi Informasi*, vol. 2, pp. 23-32, 2018.