



## ***Customer Segmentation at BC HNI 2 Pekanbaru by Applying the K-Medoids Algorithm and Recency, Frequency, Monetary (RFM) Model***

### **Segmentasi Pelanggan pada BC HNI 2 Pekanbaru dengan Menerapkan Algoritma K-Medoids dan Model Recency, Frequency, Monetary (RFM)**

Alkahfi Madani<sup>1</sup>, Astriana Rahmah<sup>2</sup>, Fadilah Nurunnisa<sup>3</sup>, Andi Elia<sup>4</sup>

<sup>1234</sup>Program Studi Sistem Informasi, Fakultas Sains dan Teknologi,  
UIN Sultan Syarif Kasim Riau, Indonesia  
JL. H.R Soebrantas Km.18 Panam Pekanbaru-Riau

E-Mail: <sup>1</sup>12050310447@students.uin-suska.ac.id, <sup>2</sup>12050320398@students.uin-suska.ac.id,  
<sup>3</sup>12050322891@students.uin-suska.ac.id, <sup>4</sup>11850320503@students.uin-suska.ac.id

*Corresponding Author: Alkahfi Madani*

#### **Abstract**

*PT. Halal Network Indonesia (HNI) is a business company that sells halal and herbal products in Indonesia. One of PT.HNI's branches is located at Jalan Paus no. 69 F, Central Tangerang, Marpoyan Damai Pekanbaru Riau, namely BC 2 HNI Pekanbaru. There is a problem in BC 2 HNI Pekanbaru, namely the company has not been able to classify which customers are profitable for the company. So it is difficult for companies to determine the right marketing strategy for their customers. Researchers utilize data mining model clustering which is very efficient to determine the characteristics of the formed customer groups, using the K-Medoids algorithm which has advantages in classifying data, and applying the RFM model so that the results of data segmentation obtained are more leverage to handle problems in BC 2 HNI Pekanbaru so that The company's marketing is more effective and efficient. By getting the best 2 clusters based on DBI (Davies Bouldin Index). With a DBI value of 0.492, as well as customer segments, the research results obtained were 452 core customers and 351 lost customers.*

**Keyword:** *Business center, Clustering, Data mining, K-medoids, Rapidminer, RFM, Segmentation customer.*

#### **Abstrak**

*PT. Halal Network Indonesia (HNI) merupakan perusahaan bisnis yang melakukan pemasaran pada produk-produk halal dan herbal di Indonesia. Salah satu cabang PT.HNI berada di Jalan Paus no 69 F Tangerang Tengah, Marpoyan Damai Pekanbaru Riau yaitu BC 2 HNI Pekanbaru. Terdapat permasalahan dalam BC 2 HNI Pekanbaru yaitu perusahaan belum mampu mengelompokkan pelanggan mana yang menguntungkan bagi perusahaan. Sehingga perusahaan sulit menentukan strategi pemasaran yang tepat bagi pelanggannya. Peneliti memanfaatkan *data mining* model *clustering* yang sangat efisien untuk menentukan karakteristik kelompok pelanggan yang terbentuk, menggunakan algoritma *K-Medoids* yang memiliki keunggulan dalam mengelompokkan data, serta menerapkan model RFM agar hasil segmentasi data yang didapatkan lebih maksimal untuk menangani permasalahan pada BC 2 HNI Pekanbaru agar pemasaran yang dilakukan perusahaan lebih efektif dan efisien. Dengan mendapatkan 2 *cluster* terbaik berdasarkan DBI (*Davies Bouldin Index*). Dengan nilai DBI 0.492, serta segmen pelanggan hasil penelitian yang didapatkan adalah 452 *core customer* dan 351 *lost customer*.*

**Kata Kunci:** *Business center, Clustering, Data mining, K-medoids, Rapidminer, RFM, Segmentasi pelanggan.*

## 1. PENDAHULUAN

PT. *Halal Network International* (HNI) merupakan perusahaan bisnis yang melakukan pemasaran pada produk-produk halal dan herbal di Indonesia. *Network Marketing* diterapkan dalam perusahaan HNI, sehingga HNI memiliki jaringan pemasaran yang cukup luas [1]. Data tahun 2021 tercatat sebanyak 368 BC tersebar di seluruh Indonesia, salah satunya yaitu BC 2 HNI Pekanbaru yang berlokasi di Jalan Paus no 69 F Tangkerang Tengah, Marpoyan Damai Pekanbaru Riau. Pelanggan pada BC 2 HNI Pekanbaru terdiri dari 2 kategori pelanggan, yaitu member sebesar 92% dan non member sebesar 8% dari seluruh pelanggan yang ada di BC 2 HNI Pekanbaru.

Perusahaan HNI memiliki sebuah sistem informasi penjualan yang mampu mengelola semua transaksi penjualan yang terjadi di BC HNI 2 Pekanbaru. Namun data transaksi tersebut hanya digunakan untuk informasi laba rugi serta pembukuan bagi pelanggan-pelanggan yang sudah melakukan transaksi pembelian. Oleh karena itu, BC HNI 2 Pekanbaru belum mampu untuk mengelompokkan pelanggan yang melakukan transaksi mengenai pelanggan mana yang potensial atau menguntungkan bagi perusahaan. Selain itu, perusahaan tidak pernah mengolah data-data transaksi pembelian secara langsung sehingga BC 2 HNI Pekanbaru kurang memahami karakteristik pelanggan yang ada di perusahaannya, yang menyebabkan perusahaan kesulitan dalam menyusun strategi untuk pemasaran produk yang tepat. Perusahaan harus berfokus pada strategi kebutuhan pelanggan agar dapat menghadapi persaingan yang semakin ketat antar perusahaan. Penanganan data transaksi penjualan dengan perhitungan menggunakan IT yang tepat dapat dijadikan solusi untuk menangani permasalahan yang terdapat pada perusahaan BC 2 HNI Pekanbaru, agar pemasaran lebih efektif dan efisien. Peneliti memanfaatkan *data mining* model *clustering* menggunakan RFM dengan algoritma *K-Medoids* untuk menangani permasalahan pada BC 2 HNI Pekanbaru.

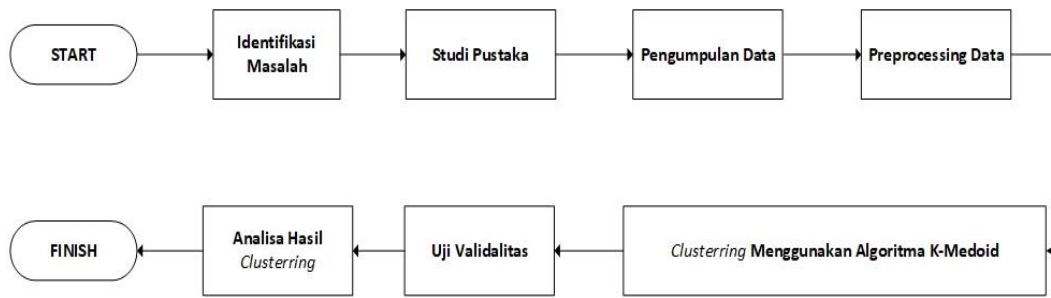
Penggunaan data mining dalam pengolahan data berfungsi agar pola yang ada dapat diketahui serta informasi-informasi yang tersimpan didalamnya dapat diidentifikasi. Pada data mining terdapat suatu model *clustering* yang berfungsi untuk mengelompokkan [2]. Untuk menentukan segmentasi pelanggan digunakan model *clustering* yang sangat efisien dalam melakukan analisis kelompok data yang berfungsi untuk menentukan karakteristik kelompok pelanggan yang terbentuk [3]. Salah satu algoritma teknik *clustering* adalah algoritma *K-Medoids* yang mampu mengelompokkan data-data kedalam *cluster-cluster* dengan objek yang sama [4]. Model RFM diterapkan dalam segmentasi data, agar hasil yang didapatkan lebih tepat [5]. Peneliti menggunakan algoritma *K-Medoids* karena dinilai lebih unggul dibandingkan ketika menggunakan algoritma *K-Means*. Hal ini sesuai dengan hasil penelitian Arora dkk terkait penemuan keunggulan dan keterbatasan menggunakan model *clustering* dengan membandingkan antara algoritma *K-Means* dan *K-Medoids*. Kedua algoritma ini kemudian diuji pada KEEL dataset transaksi. Penelitian tersebut mendapatkan hasil bahwa waktu yang dibutuhkan untuk pemilihan *cluster* dan kompleksitas antar ruang *cluster* jauh lebih baik menggunakan *K-Medoids* dibandingkan *K-Means*. Selain itu, penggunaan algoritma *K-Medoids* juga memiliki keunggulan dalam *running time*, tidak sensitif terhadap outlier dan dapat mereduksi noise dibanding ketika menggunakan algoritma *K-Means*. [6].

Penelitian sebelumnya sudah dilakukan oleh Andini dkk [7] dan Marlina dkk [8] tentang penggunaan algoritma *K-Medoids* dalam proses *clustering*. Kumpulan data yang digunakan dalam penelitian Andini dkk yakni dataset dari RSUD Kota Bandung tentang riwayat penyakit riwayat penyakit pada tahun 2019 [7], sementara itu pada penelitian Marlina dkk yakni kumpulan data yang dipakai yakni dari persebaran anak berkebutuhan khusus pada Provinsi Riau [8]. Dari kedua penelitian tersebut menunjukkan bahwa algoritma *K-Medoids* bekerja lebih baik karena setiap subjek pada setiap cluster memiliki kualitas yang baik, dimana setiap subjek dikelompokkan berdasarkan tingkat kemiripan yang tinggi. Berdasarkan nilai validitasnya algoritma *K-Medoids* dinilai lebih baik dalam mengelompokkan data dibandingkan dengan algoritma *K-Means*.

Metode *clustering* dari berbagai literatur yang telah dilakukan sebelumnya dan telah *direview*, metode *clustering* menggunakan algoritma *K-Medoids* memberikan hasil yang baik dan memuaskan. Dengan dasar tersebut, algoritma *K-Medoids* diterapkan pada penelitian ini, dengan tujuan untuk mengelompokkan transaksi pelanggan pada BC 2 HNI Pekanbaru, terkait pelanggan mana yang potensial atau menguntungkan bagi perusahaan. Dari informasi tersebut perusahaan dapat mengetahui karakteristik pelanggan potensial dan dapat menyusun strategi pemasaran bagi pelanggan potensial tersebut.

## 2. METODOLOGI PENELITIAN

Gambaran metodologi penelitian ditunjukkan pada Gambar 1.



**Gambar 1.** Alur Penelitian

### 2.1 Identifikasi Masalah

Teori yang akan dipelajari adalah metode RFM dan metode algoritma K-medoids, dan metode penelitian penelitian ini dimulai dari identifikasi masalah, dimulai dari identifikasi masalah pada tahap ini, dan dilanjutkan dengan menentukan tujuan penelitian. Metode riset pada penelitian ini dimulai dengan mengidentifikasi masalah dan dilanjutkan dengan menentukan tujuan penelitian. Dengan ditentukannya tujuan penelitian dan permasalahan yang ada penelitian menjadi lebih jelas dan terarah.

### 2.2 Studi Pustaka

Studi pustaka kemudian dilakukan untuk memberikan peneliti dengan dasar referensi yang kuat untuk menerapkan metode yang digunakan dalam penelitian ini.. Teori yang dipertimbangkan adalah metode RFM dan algoritma K-Medoids dari jurnal, artikel, dan sumber lainnya.

### 2.3 Pengumpulan Data

Tahap selanjutnya adalah tahap pengumpulan data. Selama tahap ini, peneliti menggunakan data transaksi pelanggan pada masa pandemi Covid-19, Oktober 2021. Data pelanggan yang digunakan pada BC 2 HNI Pekanbaru yaitu data transaksi pelanggan di BC 2 HNI.

### 2.4 Knowledge Discovery Database (KDD)

*Knowledge Discovery Database (KDD)* adalah suatu keseluruhan proses dalam menganalisa data yang sah, baru dan dapat di mengerti yang nantinya berguna untuk dikelola.[9] Dalam proses KDD ini, Terdapat tahapan *Data Selection* untuk menganalisis dan memilih data yang digunakan. Tahap pemilihan data dilakukan dengan memilih atribut yang akan digunakan. Tiga atribut dipilih dari jumlah atribut (dengan lambang ✓) digunakan dalam pemrosesan data selanjutnya. Tabel 1 menunjukkan pemilihan atribut yang digunakan. Dimensi data yang siap diolah adalah 1.659 baris x 7 kolom.

**Tabel 1.** Pemilihan Atribut

Atribut	Detail Penggunaan Data
No	✓
No_Resi	X
Nama_Pembeli	✓
Tgl.Penjualan	✓
Total VP	X
Item	X
Total	✓

Setelah memilih data, masuk ke tahap pretreatment/pembersihan. Data dibersihkan selama tahap ini. Pada tahap ini, data asli dijamin dengan memeriksa dan menyiapkan integritas data, data kosong atau nol, dan data redundan serta mampu mengolah data sesuai kebutuhan.[10]. Selanjutnya adalah tahap transformasi, yang bertujuan untuk memodifikasi data agar sesuai dengan format yang diolah oleh proses data mining. [11]. Konversi data ini dilakukan dengan mengubah data menjadi data numerik. Hal ini dikarenakan perhitungan clustering hanya dapat melakukan pengolahan data numeric [3]. Selain itu, data numerik dapat dengan mudah dipahami dan diinterpretasikan dengan mentransformasikannya menggunakan model RFM [12][13].

## 2.5 Metode RFM (Recency, Frequency, dan Monetary)

Metode yang dapat digunakan untuk mengidentifikasi kelompok pelanggan yang mungkin mendapatkan perlakuan khusus disebut dengan metode RFM. Metode RFM dapat diterapkan pada pemasaran karena karena perlu mengaktifkan data historis seperti riwayat pembelian dan riwayat penelusuran.[14].

$$\text{Skor RFM} = (\text{Recency} \times 100) + (\text{Nilai Frequency} \times 0) + (\text{Nilai Monetary}) \quad (1)$$

## 2.6 Normalisasi Data

Data yang akan diperoleh nantinya akan sangat bervariasi, sehingga data tersebut bisa saja terlalu besar atau terlalu kecil. Oleh karena itu perlu dilakukan normalisasi data untuk mencegah dimensi data yang terlalu kecil atau dimensi data yang terlalu besar [15]. Normalisasi dilakukan dengan menggunakan Min-Max Normalization, dimana nilai normalisasi merupakan hasil pembagian nilai data aktual dengan nilai maksimum data aktual, dan nilai minimum data aktual. [4].

$$V = \frac{v - \text{Min}}{\text{Max} - \text{Min}} \quad (2)$$

## 2.7 Clustering Menggunakan Algoritma K-Medoid

Algoritma K-Medoids atau Partitioning Around Medoids mengelompokkan sekumpulan objek ke dalam sejumlah grup atau kelompok. Algoritma K-Medoids menggunakan objek dalam koleksi untuk mewakili sebuah cluster. Objek yang dipilih untuk mewakili sebuah cluster disebut medoid. [16].

$$d(x, y) = \|x - y\| = \sum_{i=1}^n (x_i - y_i)^2 \quad (3)$$

## 2.8 Uji Validitas

Validasi cluster merupakan salah satu cara untuk mendapatkan validasi cluster terbaik dari proses eksperimen cluster selanjutnya. Dalam penelitian ini, kami menggunakan Davies Doulbin Index (DBI) sebagai alat untuk memvalidasi hasil analisis cluster. Untuk mengoptimalkan skema cluster serta meminimalkan DBI gunakan uji validasi menggunakan DBI [17].

## 2.9 Analisis Hasil Clustering

Untuk mengetahui tipe pelanggan yang diinginkan, dapat menggunakan matriks loyalitas pelanggan [18]. Matriks mengklasifikasikan perilaku pelanggan berdasarkan segmen menggunakan lambang ( $\uparrow$ ) sementara itu untuk menyatakan nilai rata-rata cluster yang terbentuk yang berada di atas rata-rata nilai total dan lambangnya ( $\downarrow$ ) tujuannya untuk rata-rata nilai cluster yang terbentuk lebih rendah dari rata-rata nilai keseluruhan. [19].

## 3. HASIL DAN PEMBAHASAN

Sesuai dari hasil observasi pada BC 2 HNI didapatkan yaitu data transaksi penjualan yang dilakukan pada bulan Oktober 2021 dengan jumlah sebanyak 1.659 transaksi yang akan digunakan sebagai sampel. Adapun sampel data transaksi terdapat 7 atribut yang dipakai yaitu nomor, no\_invoice, nama\_pembeli, tanggal\_penjualan, total\_VP, item dan total. Berikut ini adalah sampel data transaksi yang digunakan untuk sampel pada bulan Oktober 2021 terdapat pada Tabel 2.

**Tabel 2.** Data Awal Transaksi Penjualan

	A	B	C	D	E	F	G
1	No.	No. Invoice	Nama Pembeli	Tgl. Penjualan	Total VP	Item	Total
2	1	B00822110012	SUMARDIYONO	01-10-2021	813000	48	2509000
3	2	B00822110013	SUMARDIYONO	01-10-2021	571000	25	2080000
4	3	B00822110016	ELYANA	01-10-2021	166500	10	532000
5	4	B00822110018	IYEF.T. SULAEMAN	01-10-2021	0	0	0
6	5	B00822110019	LILY KUSUMAWARDHANI	01-10-2021	90500	7	291000
7	6	B008221100110	MAIDALFI	01-10-2021	25000	1	90000
8	7	B008221100111	MIRI DORINA	01-10-2021	62000	8	204000
9	8	B008221100113	HENDRIYANTO	01-10-2021	40000	2	120000
10	9	B008221100114	ZAINAB	01-10-2021	35500	4	117000
11	10	B008221100116	NURDIN SE.AK.	01-10-2021	13000	1	50000
12	11	B008221100117	ADHELVA	01-10-2021	35000	4	135000
13	12	B008221100118	BUNGA PERMATASARI	01-10-2021	1140000	43	3260000
14	13	B008221100119	ELYANA	01-10-2021	40000	2	120000
15	14	B008221100120	EKA PERMATA SARI	01-10-2021	62000	3	170000
16	15	B008221100122	ROHANUM	01-10-2021	41000	2	150000
17	16	B008221100123	NAILAN NIZAM	01-10-2021	48000	3	150000
18	17	B008221100124	ELFANI ARIS	01-10-2021	40000	2	120000
19	18	B008221100125	SRI INDAR YANTI	01-10-2021	1430000	103	4780000
20	19	B008221100126	NUR RASYIDAH	01-10-2021	335000	24	1068000
21	20	B008221100127	SYEFRA YUDA	01-10-2021	1107000	62	3490000
22	21	B008221100128	ABDULLAH ALI	01-10-2021	943500	67	3005000
23	22	B008221100129	RIAU ANDRIYANI	01-10-2021	2052000	129	6818000

Setelah dilakukannya tahapan *Knowledge Discovery in Database* (KKD), yang terdiri dari tahap *selection* atau seleksi data, tahap *cleaning* atau *data cleaning* dan terakhir tahap transformasi data. Ketika sampai pada tahap seleksi dilakukan dengan seleksi atribut yang ada pada data transaksi pelanggan BC 2 HNI Pekanbaru. Saat proses seleksi harus berdasarkan kriteria RFM yaitu kapan tanggal transaksi terakhir (*Recency*), dalam satu periode berapa kali pelanggan membeli produk (*Frequency*) dan total belanja pelanggan (*Monetary*). Dalam hal ini atribut yang akan dipakai berdasarkan kriteria RFM dan akan digunakan saat proses *clustering* yaitu nomor, nama\_pembeli, tanggal\_penjualan, dan total.

Pada tahap *cleaning data* yang dilakukan yakni menghapus data yang tidak signifikan dengan kriteria model RFM. Adapun data yang sudah signifikan dengan kriteria RFM dapat ditampilkan pada Tabel 3.

**Tabel 3.** Data Cleaning

No	Nama	Tanggal Penjualan	Total
1	SUMARDIYONO	01-10-2021	2509000
2	SUMARDIYONO	01-10-2021	2080000
3	ELYANA	01-10-2021	532000
4	IYEF.T.SULAEMAN	01-10-2021	0
5	LILY KUSUMAWARDHANI	01-10-2021	291000
...	...	...	...
1656	HERO	31-10-2021	30000
1657	NISWATUL HUSNA	31-10-2021	300000
1658	MERI ANDANI	31-10-2021	75000
1659	JANNATUL SOLEHA	31-10-2021	170000

Kemudian data yang lengkap sudah diperoleh dengan atribut yang selaras maka selanjutnya ke tahap mencari nilai transformasi R,F. dan M. Berikut ini merupakan tabel data penjualan berdasarkan kriteria model RFM. Untuk nilai R (*Recency*) ialah jarak waktu antara pembelian terakhir dengan waktu pembelian saat ini atau dengan tanggal ketetapan yang diambil yaitu pada tanggal 31 Oktober 2021, Nilai F (*Frequency*) ialah banyaknya jumlah transaksi pada pelanggan dalam suatu periode, Nilai M (*Monetary*) ialah banyaknya jumlah uang yang telah digunakan saat pembelian dalam suatu periode. Adapun tabel nilai *Recency*, *Frequency* dan *Monetary* dapat dilihat pada Tabel 4.

**Tabel 4.** Data Transformasi *Recency*

No	Nama	Tanggal Transaksi Terakhir	Tanggal Ketetapan	R	F	M
1	ABDUL ROKIP	31/10/2021	31/10/2021	0	6	Rp 54.977.000
2	ABDULLAH ALI	30/10/2021	31/10/2021	1	32	Rp 46.012.000
3	ABDURRAHIM	23/10/2021	31/10/2021	8	2	Rp 5.365.000
4	ABUL BARRO AL AUJAT SITORUS	03/10/2021	31/10/2021	28	1	Rp 90.000
5	ACEP JAMALUDIN	18/10/2021	31/10/2021	13	3	Rp 1.150.000
...	...	...	...	...	...	...
802	ZULKIFLI.M	24/10/2021	31/10/2021	7	2	Rp 442.000
803	ZULMADIAR	31/10/2021	31/10/2021	0	5	Rp 450.000
804	ZULMAYATI ZEIN	25/10/2021	31/10/2021	6	1	Rp 2.017.000
805	ZURRAHMI	22/10/2021	31/10/2021	9	1	Rp 110.000

Setelah mendapatkan hasil dari transformasi data maka selanjutnya dilakukan tahap normalisasi data pada atribut RFM. Selanjutnya gunakan rumus *Min Max Normalization* pada tahap normalisasi data. Nilai atribut M (*Monetary*) pada data RFM tidak seimbang dengan atribut *Recency*, dan *Frequency*, oleh karena itu dilakukanlah tahap normalisasi untuk menghasilkan nilai yang seimbang. Kemudian didapatkan hasil nilai minimum dan maksimum dari data yang ada dapat dilihat pada Tabel 5.

**Tabel 5.** Nilai Minimum dan Maksimum Data RFM

	R	F	M
Mínimum	0	1	Rp 10.000
Maksimum	30	33	Rp 100.822.000

Setelah ke tahap normalisasi data menggunakan *Min Max Normalization* selanjutnya Pada nilai atribut R (*Recency*) dilakukan pengurangan nilai 1- sebab semakin sedikit nilai maka semakin sesuai hasil yang akan

didapatkan. Hasil akhir data akan ditunjukkan melalui hasil normalisasi yang dilakukan dari tahapan *preprocessing* data. Adapun hasil akhir tersebut terdapat pada Tabel 6.

**Tabel 6.** Hasil Normalisasi Data

No	Nama	NR	NF	NM
1	ABDUL ROKIP	1,000	0,156	0,545
2	ABDULLAH ALI	0,967	0,969	0,456
3	ABDURRAHIM	0,733	0,031	0,053
4	ABUL BARRO AL AUJA'I SITORUS	0,067	0,000	0,001
5	ACEP JAMALUDIN	0,567	0,063	0,011
...	...	...	...	...
803	ZULMADIAR	1,000	0,125	0,004
804	ZULMAYATI ZEIN	0,800	0,000	0,020
805	ZURRAHMI	0,700	0,000	0,001

Penerapan algoritma *K-Medoids* akan diterapkan saat tahapan *clustering* di akhir langkah pra-pemrosesan data. Pada penelitian ini menggunakan metode Elbow untuk menentukan berapa jumlah *cluster* serta melakukan beberapa kali percobaan *cluster* sebanyak jumlah *cluster* (*k*) dan metode ini bisa dimulai dari *k*=2 dan seterusnya [20]

### 3.1 Pengelompokkan dengan Algoritma *K-Medoids* Menggunakan *Tools RapidMiner*

Tahap ini akan memproses perhitungan data dengan memakai *Tools RapidMiner*, dengan melakukan percobaan yakni menggunakan 2 *cluster* sampai 6 *cluster* yang bertujuan untuk mendapatkan *cluster* terbaik. Data yang diinputkan adalah hasil dari normalisasi dari model RFM yaitu sebanyak 803 data transaksi pelanggan.

### 3.2 Hasil DBI pada *cluster*

Hasil DBI pada *cluster* 2 sampai dengan *cluster* 7 dapat dilihat pada Tabel 7.

**Tabel 7.** Hasil DBI *Cluster* 2 sampai *Cluster* 7

No	<i>Cluster</i>	Nilai DBI
1.	<i>Cluster</i> 2	0.492
2.	<i>Cluster</i> 3	0.944
3.	<i>Cluster</i> 4	0.802
4.	<i>Cluster</i> 5	0.946
5.	<i>Cluster</i> 6	0.866

Nilai DBI yang terbaik sudah diperoleh dari hasil nilai DBI *cluster* RFM pada data transaksi pelanggan. Hasil *cluster* terbaik yaitu dilihat dari nilai DBI yang paling rendah di setiap *cluster* hal ini bertujuan untuk mencari pelanggan yang potensial dimana ini adalah sebagai tujuan dari penelitian ini. Setelah itu diperoleh data transaksi yaitu pada *cluster* 2 dengan nilai DBI 0.492 dimana nilai tersebut merupakan nilai DBI terbaik pada *cluster*. Berikut adalah nilai DBI terbaik dari *cluster* data transaksi dapat dilihat pada tabel 8.

**Tabel 8.** Nilai DBI 0,492 Terbaik *Cluster* 2

No	<i>Cluster</i>	Pelanggan
1.	<i>Cluster</i> 0	452
2.	<i>Cluster</i> 1	351
Total		803

### 3.3 Analisis dan Hasil *clustering* RFM

Setelah proses *clustering* dilakukan, mendapatkan *cluster* terbaik dari data transaksi selanjutnya ialah analisa hasil *clustering* yang telah dilakukan. Nama-nama pelanggan sesuai hasil *cluster* terbaik yang telah didapatkan dari data transaksi pelanggan BC 2 HNI terdapat pada Tabel 9.

**Tabel 9.** Daftar Nama Pelanggan Sesuai *Cluster* Terbaik

No	Nama	<i>Cluster</i>
1.	ABDUL ROKIP	cluster_0
2.	ABDULLAH ALI	cluster_0
3.	ABDURRAHIM	cluster_0

No	Nama	Cluster
4.	ABUL BARRO AL AUJAI SITORUS	cluster_1
5.	ACEP JAMALUDIN	cluster_0
6.	ADE ELVIRA ROSA	cluster_1
7.	ADE PUTRA DIPRAJA	cluster_0
8.	ADE YUDHISTIRA	cluster_0
9.	ADELINA BATUBARA	cluster_0
10.	ADHELYA	cluster_1
...	...	...
801	ZULHELMI	cluster_1
802	ZULHIDAYATI	cluster_0
803	ZULKIFLI	cluster_0

Lambang peringkat untuk setiap cluster kemudian ditentukan dengan merata-ratakan atribut RFM dari tiap-tiap cluster. Nilai RFM yang dimiliki oleh cluster rata-rata lebih tinggi dari rata-rata RFM sebelum *cluster* diberi lambang ( $\uparrow$ ). Sedangkan cluster yang memiliki nilai rata-rata RFM lebih rendah dari rata-rata RFM sebelum cluster diberi lambang ( $\downarrow$ ), dan nilai rata-ratanya masih dianggap sebagai cluster yang valid. Nilai RFM pelanggan rata-rata sebelum cluster dapat dilihat pada Tabel 10.

**Tabel 10.** Nilai Rata-rata RFM Sebelum *Cluster*

R	F	M
0,585	0,033	0,017

Rata-rata RFM untuk setiap cluster adalah sebagai berikut pada Tabel 11.

**Tabel 11.** Nilai Rata-rata RFM *Cluster* Terbaik

Cluster	Jumlah Pelanggan	R	F	M
0	452	0,822	0,072	0,009
1	351	0,485	0,004	0,002

Lambang peringkat atau *rank* untuk setiap cluster adalah nilai rata-rata awal RFM dibagi dengan nilai rata-rata terbaik dari RFM untuk cluster tersebut, hasilnya dapat dilihat pada Tabel 12.

**Tabel 12.** Lambang peringkat RFM Setiap *Cluster*

Cluster	Jumlah Pelanggan	Lambang Rank RFM
0	452	R $\downarrow$ F $\uparrow$ M $\downarrow$
1	351	R $\uparrow$ F $\downarrow$ M $\downarrow$

#### 4. KESIMPULAN

Berikut merupakan analisa pola RFM per *cluster* berdasarkan variabel RFM untuk menentukan kategori pelanggan, yaitu :

1. *Costumer segment 1: Core costumer group- High frequency buying costumers* (RFM,  $\downarrow\uparrow\downarrow$ ). Merupakan pelanggan dengan kategori pelanggan utama yang *recent transaction time* yang rendah dengan artian bahwa baru-baru ini dan frekuensinya juga tinggi hal ini diartikan bahwa pelanggan sering bertransaksi antara pelanggan dan perusahaan serta rendahnya nilai *monetary* yang berarti total uang yang dipakai tidak terlalu besar. Pelanggan tersebut termasuk dalam pelanggan potensial untuk kedepannya bagi perusahaan, bagaimana cara mempertahankan pelanggan dan meningkatkan nilai rata-rata uang yang dikeluarkan oleh pelanggan.
2. *Costumers segment 2: Lost Costomer group- uncertain lost customers* (RFM,  $\uparrow\downarrow\downarrow$ ). *Recent transaction time* yang tinggi dimana maksudnya adalah para pelanggan hanya bertransaksi hanya pada awal bulan. *Frequency* dan *Monetary* nya rendah berarti merupakan pelanggan yang tergolong bertransaksi ketika membutuhkan sekali barang dari perusahaan, pelanggan seperti ini dapat dilabeli sebagai pelanggan yang tidak pasti sebagai pelanggan potensial.

#### REFERENSI

- [1] K. I. Sari, "Strategi Komunikasi Pemasaran Halal Network International-Herba Penawar Alwahida Indonesia (Hni-Hpai) Dalam Menarik Minat Beli Konsumen Muslim Di Bandar Lampung (Studi Pada Business Center Hni-Hpai Cabang Lampung 3)." UIN Raden Intan Lampung, 2020.
- [2] I. Kamila, U. Khairunnisa, and M. Mustakim, "Perbandingan Algoritma K-Means dan K-Medoids

- untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau,” *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 5, no. 1, p. 119, 2019, doi: 10.24014/rmsi.v5i1.7381.
- [3] R. D. Dana, C. L. Rohmat, and A. R. Rinaldi, “Strategi Marketing Penerimaan Mahasiswa Baru Menggunakan Machine Learning dengan Teknik Clustering,” vol. 04, no. 2, pp. 201–204, 2019, doi: 10.30591/jpit.v4i2-2.1879.
- [4] A. A. D. Sulistyawati and M. Sadikin, “Penerapan Algoritma K-Medoids Untuk Menentukan Segmentasi Pelanggan,” *Sistemasi*, vol. 10, no. 3, p. 516, 2021, doi: 10.32520/stmsi.v10i3.1332.
- [5] T. Hardiani, S. Sulisty, and R. Hartanto, “Segmentasi Nasabah Tabungan Menggunakan Model RFM (Recency, Frequency, Monetary) dan K-Means Pada Lembaga Keuangan Mikro,” *Semin. Nas. Teknol. Inf. dan Komun. Terap.*, no. November, p. 2015, 2015.
- [6] P. Arora, Deepali, and S. Varshney, “Analysis of K-Means and K-Medoids Algorithm for Big Data,” *Phys. Procedia*, vol. 78, no. December 2015, pp. 507–512, 2016, doi: 10.1016/j.procs.2016.02.095.
- [7] A. D. Andini and T. Arifin, “Implementasi Algoritma K-Medoids Untuk Klasterisasi Data Penyakit Pasien Di RSUD Kota Bandung,” *J. Responsif Ris. Sains dan Inform.*, vol. 2, no. 2, pp. 128–138, 2020, doi: 10.51977/jti.v2i2.247.
- [8] D. Marlina, N. Lina, A. Fernando, and A. Ramadhan, “Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak,” *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 2, p. 64, 2018, doi: 10.24014/coreit.v4i2.4498.
- [9] Mustakim, “Pemetaan Digital dan Pengelompokan Lahan Hijau di Wilayah Provinsi Riau Berdasarkan Knowledge Discovery in Databases (KDD) dengan Teknik K-Means Mining,” *Semin. Nas. Teknol. Inf. Komun. dan Ind.* 4, pp. 103–111, 2012.
- [10] A. R. Riszky and M. Sadikin, “Data Mining Menggunakan Algoritma Apriori untuk Rekomendasi Produk bagi Pelanggan,” *J. Teknol. dan Sist. Komput.*, vol. 7, no. 3, pp. 103–108, 2019, doi: 10.14710/jtsiskom.7.3.2019.103-108.
- [11] H. Astuti, “Penerapan Data Mining Menggunakan Metode K-Means Clustering Untuk Pengelompokan Data Pelanggan (Studi Kasus: PT. Pinus Merah Abadi),” *J. Web Inform. Teknol.*, vol. 4, no. 1, pp. 9–16, 2019.
- [12] S. I. Murpratiwi, I. G. Agung Indrawan, and A. Aranta, “Analisis Pemilihan Cluster Optimal Dalam Segmentasi Pelanggan Toko Retail,” *J. Pendidik. Teknol. dan Kejuru.*, vol. 18, no. 2, p. 152, 2021, doi: 10.23887/jptk-undiksha.v18i2.37426.
- [13] M. Aryuni, E. D. Madyatmadja, and E. Miranda, “Customer Segmentation in XYZ Bank Using K-Means and K-Medoids Clustering,” in *2018 International Conference on Information Management and Technology (ICIMTech)*, 2018, pp. 412–416. doi: 10.1109/ICIMTech.2018.8528086.
- [14] Sudriyanto, “CLUSTERING LOYALITAS PELANGGAN DENGAN METODE RFM (RECENCY, FREQUENCY, MONETARY ) DAN FUZZY C-MEANS,” *Pros. SNATIF*, pp. 815–822, 2017.
- [15] S. Wulandari, “Prosiding Seminar Nasional Sains Clustering Microarray Adenoma Menggunakan Spectral Clustering dengan Algoritma Partitioning Around Medoid ( PAM ),” vol. 1, no. 1, pp. 345–351, 2020.
- [16] D. Listiyanti, , Yoanda Alim Syahbana, Silvana Rasio Henim, “Perancangan dan Implementasi Aplikasi Android Penentu Salient Area pada Video dengan Algoritma,” vol. 2, no. 1, pp. 96–101, 2016.
- [17] S. Butsianto *et al.*, “Penerapan Data Mining Terhadap Minat Siswa Dalam Mata Pelajaran Matematika Dengan Metode K-Means,” vol. 3, no. 1, pp. 51–59, 2020.
- [18] D.-C. Li, W.-L. Dai, and W.-T. Tseng, “A two-stage clustering method to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business,” *Expert Syst. Appl.*, vol. 38, no. 6, pp. 7186–7191, 2011, doi: <https://doi.org/10.1016/j.eswa.2010.12.041>.
- [19] S. Monalisa, “Segmentasi Perilaku Pembelian Pelanggan Berdasarkan Model RFM dengan Metode K-Means,” vol. 5341, no. April, pp. 9–15, 2018.
- [20] C. D. Rumiarti and I. Budi, “Customer Segmentation for Customer Relationship Management on Retail Company: Case Study PT Gramedia Asri Media,” *J. Sist. Inf.*, vol. 13, no. 1, pp. 1–10, 2017.