

SEGMENTASI PELANGGAN DENGAN ATRIBUT RFM MENGGUNAKAN ALGORITMA K-MEANS DI REDDOORZ SYARIAH NEAR ALUN-ALUN KAJEN

Ragil Nur Hidayat ¹, Risqiati ², Arief Soma Darmawan ³

^{1,2} Teknik Informatika, Institut Widya Pratama

³ Sistem Informasi, Institut Widya Pratama

Jalan Pelita No 25 Kota Pekalongan

hidayatragil972@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk melakukan segmentasi pelanggan pada penginapan RedDoorz Syariah Near Alun-Alun KAJEN untuk mengetahui tingkat loyalitas pelanggan berdasarkan kebiasaan setiap pelanggan dalam melakukan transaksi. Permasalahan utama yang dihadapi adalah belum adanya pengelompokan pelanggan sehingga strategi pemasaran dan pelayanan masih dilakukan secara umum tanpa memperhatikan karakteristik pada setiap pelanggan. Penelitian ini menggunakan metode RFM (*Recency, Frequency, dan Monetary*) untuk menganalisis loyalitas setiap pelanggan dan algoritma *K-Means Clustering* untuk pengelompokan pelanggan. Proses penelitian meliputi pengumpulan data, *Exploratory Data Analysis* (EDA), preprocessing data, pembentukan variabel RFM, normalisasi data, dan penentuan jumlah cluster optimal menggunakan metode *Elbow, Silhouette Score, dan Davies-Bouldin Index*. Dari hasil penelitian yang sudah dilakukan bahwa pelanggan dapat dikelompokkan menjadi tiga *cluster* yaitu *cluster* loyalitas tinggi sebanyak 58 pelanggan, *cluster* loyalitas sedang sebanyak 45 pelanggan, dan *cluster* loyalitas rendah sebanyak 1 pelanggan. Dari hasil segmentasi pelanggan yang sudah dilakukan pada penelitian ini dapat diimplementasikan ke dalam sistem informasi manajemen untuk penyusunan strategi pelayanan dan pemasaran lebih tepat sasaran serta dalam memberikan perilaku khusus kepada pelanggan dengan tingkat loyalitas tertentu.

Kata kunci : RFM, K-Means Clustering, Segmentasi Pelanggan, Loyalitas Pelanggan.

1. PENDAHULUAN

RedDoorz Syariah Near Alun-Alun KAJEN adalah salah satu usaha penginapan yang berada di wilayah KAJEN, Kabupaten Pekalongan, Jawa Tengah. RedDoorz menawarkan program kemitraan kepada hotel-hotel yang ingin bergabung [1]. RedDoorz dikenal sebagai jaringan penginapan yang menyediakan tarif ekonomis dengan fasilitas yang cukup lengkap, sehingga menjadi alternatif favorit bagi wisatawan yang mencari kenyamanan tanpa harus mengeluarkan biaya besar [2]. Hal ini menjadi peluang usaha bagi para pengusaha untuk membuka penginapan maupun hotel yang dapat digunakan sebagai tempat tinggal sementara bagi wisatawan [3]. Banyaknya pengusaha yang mendirikan hotel maupun penginapan menjadikan persaingan bisnis menjadi semakin ketat.

Dalam keberlangsungan bisnis, hotel harus memiliki pelanggan [4]. Untuk menarik minat pelanggan maka memerlukan strategi pemasaran yang tepat. Dalam merancang strategi pemasaran, perusahaan saat ini tidak lagi hanya berfokus pada orientasi produk, tetapi juga lebih fokus pada orientasi pelanggan [5]. Strategi pemasaran berorientasi pelanggan yang efektif dan dapat diterapkan salah satunya adalah dengan segmentasi pelanggan.

Segmentasi pelanggan adalah strategi untuk mengelompokkan pelanggan ke dalam beberapa kategori berdasarkan kesamaan karakteristik atau perilaku [6]. Tujuannya untuk mengidentifikasi kelompok pelanggan yang memiliki kesamaan kebutuhan dan pola perilaku, sehingga perusahaan

dapat memberikan penawaran produk atau layanan yang lebih maksimal dan tepat sasaran [7]. Strategi yang digunakan RedDoorz Syariah Near Alun-Alun KAJEN saat ini masih belum berfokus pada segmentasi pelanggan dan perlakuan pada setiap pelanggannya relatif masih sama. Selain itu data penjualan yang mencakup ratusan bahkan ribuan riwayat transaksi sering kali hanya dimanfaatkan sebagai arsip, padahal dapat digunakan untuk mengetahui nilai pelanggan serta mendukung pengembangan strategi perusahaan [8].

Tanpa adanya proses segmentasi pelanggan, perusahaan akan kesulitan dalam memahami karakteristik dan perilaku pada setiap pelanggan. Hal ini dapat menyebabkan strategi pelayanan dan pemasaran yang diberikan menjadi kurang efektif karena perusahaan memberikan perlakuan yang sama kepada semua pelanggan tanpa mempertimbangkan tingkat loyalitas pelanggan untuk perusahaan. Perusahaan juga akan kesulitan dalam mengidentifikasi pelanggan yang memiliki kontribusi transaksi tertinggi maupun pelanggan yang memiliki peluang besar untuk dipertahankan. Oleh karena itu, dibutuhkan metode yang mampu mengelompokkan pelanggan berdasarkan perilaku dan karakteristiknya sehingga perusahaan dapat menyusun strategi pelayanan dan pemasaran yang lebih tepat sasaran.

Metode RFM merupakan salah satu metode berorientasi pelanggan yang efektif dan banyak digunakan untuk melakukan segmentasi pelanggan berdasarkan data riwayat transaksi [9]. Dengan menggunakan metode RFM, pelanggan akan

dikelompokkan berdasarkan perilaku pelanggan dan tingkat loyalitasnya sehingga perusahaan dapat memahami perilaku masing-masing pelanggan dan melakukan penargetan pemasaran yang lebih efektif sesuai dengan karakteristik pelanggannya. Dengan ini tentunya dapat membantu perusahaan dalam meningkatkan retensi pelanggan dan dapat mendorong peningkatan penjualan. Metode ini menggunakan tiga dimensi yaitu *Recency* (R), *Frequency* (F) dan *Monetary* (M) [10]. *Recency* yaitu waktu sejak pelanggan melakukan pembelian terakhir dalam kurun waktu tertentu, *Frequency* menunjukkan jumlah transaksi yang dilakukan pelanggan dalam kurun waktu tertentu, dan *Monetary* merupakan jumlah uang yang dikeluarkan pelanggan kepada perusahaan dalam periode tertentu [11].

Pengelompokan , Institut Widya Pratama pelanggan menggunakan metode RFM juga dibutuhkan teknik data mining yaitu *clustering*. *Clustering* yaitu metode yang juga digunakan untuk mengelompokkan data menjadi ke dalam beberapa kelompok, di mana data disetiap kelompok yang sama memiliki karakteristik yang sama tinggi, dan data yang karakteristiknya berbeda akan dikelompokkan ke kelompok lainnya [11]. *K-Means clustering* adalah salah satu algoritma pembelajaran mesin tanpa pengawasan yang populer untuk mengelompokkan data dan menggunakan pendekatan berbasis centroid sebagai pemisah sekumpulan data yang tidak memiliki label ke dalam beberapa *cluster* berdasarkan kesamaan karakteristiknya [10]. Dengan menggunakan metode ini dapat membantu perusahaan dalam mengelompokkan pelanggan menjadi beberapa segmen sesuai dengan karakteristiknya.

Penelitian ini melakukan segmentasi pelanggan pada penginapan RedDoorz Syariah Near Alun-Alun KAJEN dengan model RFM dan *K-Means clustering* berdasarkan karakteristik dan perilakunya untuk mengidentifikasi pelanggan yang loyal dan dapat menguntungkan untuk perusahaan. Hasil dari penelitian ini diharapkan dapat membantu dan bermanfaat untuk perusahaan dalam proses pengembangan bisnis khususnya pada strategi pemasaran yang berorientasi pelanggan sehingga perusahaan dapat memberikan penawaran dan layanan yang lebih maksimal dan tepat sasaran.

2. TINJAUAN PUSTAKA

2.1. Segmentasi Pelanggan

Segmentasi pelanggan adalah strategi untuk mengelompokkan pelanggan ke dalam beberapa kategori berdasarkan kesamaan karakteristik atau perilaku [6]. Tujuannya untuk mengidentifikasi kelompok pelanggan yang memiliki kesamaan kebutuhan dan pola perilaku, sehingga perusahaan dapat memberikan penawaran produk atau layanan yang lebih maksimal dan tepat sasaran [7]. Dengan melakukan segmentasi pelanggan perusahaan dapat mengetahui nilai pelanggan serta mendukung pengembangan strategi perusahaan. Selain itu,

segmentasi pelanggan dapat mempermudah perusahaan dalam memahami perilaku pada setiap pelanggan.

2.2. Metode RFM (*Recency, Frequency, Monetary*)

Metode RFM merupakan salah satu metode berorientasi pelanggan yang efektif dan banyak digunakan untuk melakukan segmentasi pelanggan berdasarkan data riwayat transaksi [9]. *Recency* yaitu waktu sejak pelanggan melakukan pembelian terakhir dalam kurun waktu tertentu, *Frequency* menunjukkan jumlah transaksi yang dilakukan pelanggan dalam kurun waktu tertentu, dan *Monetary* merupakan jumlah uang yang dikeluarkan pelanggan kepada perusahaan dalam periode tertentu [11]. Dengan menggunakan metode RFM, pelanggan akan dikelompokkan berdasarkan perilaku pelanggan dan tingkat loyalitasnya sehingga perusahaan dapat memahami perilaku masing-masing pelanggan dan melakukan penargetan pemasaran yang lebih efektif sesuai dengan karakteristik pelanggannya. Dengan ini tentunya dapat membantu perusahaan dalam meningkatkan retensi pelanggan dan dapat mendorong peningkatan penjualan.

2.3. K-Means Clustering

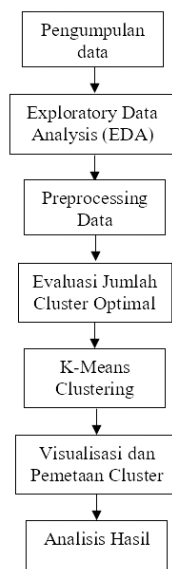
K-Means clustering adalah salah satu algoritma pembelajaran mesin tanpa pengawasan yang populer untuk mengelompokkan data dan menggunakan pendekatan berbasis *centroid* sebagai pemisah sekumpulan data yang tidak memiliki label ke dalam beberapa *cluster* berdasarkan kesamaan karakteristiknya. Dengan menggunakan metode ini dapat membantu perusahaan dalam mengelompokkan pelanggan menjadi beberapa segmen sesuai dengan karakteristiknya

3. METODE PENELITIAN

3.1. Tahap Penelitian

Metode penelitian ini berisi rangkaian dari beberapa langkah yang dirancang secara sistematis dan terstruktur untuk mencapai tujuan yang telah ditentukan dalam penelitian [12]. Penelitian ini dimulai dengan pengumpulan data yang diperoleh dari data internal penginapan RedDoorz Syariah Near Alun-Alun KAJEN. Data yang digunakan mencakup informasi seperti nama pelanggan, tanggal cek in, tanggal cek out, jumlah hari menginap, serta total pembayaran. Data yang sudah terkumpul kemudian dianalisa menggunakan tahap *Exploratory Data Analysis* (EDA) untuk mendapatkan gambaran awal mengenai karakteristik data, distribusi nilai, dan kondisi data secara umum. Selanjutnya dilakukan *preprocessing* data seperti pembentukan variabel *Recency*, *Frequency*, dan *Monetary* (RFM) sebagai representasi perilaku setiap pelanggan. Untuk menentukan jumlah *cluster* yang optimal, dilakukan evaluasi menggunakan metode *Elbow*, *Silhouette Score*, dan *Davies-Bouldin Index*. Dari hasil evaluasi tersebut, ditetapkan jumlah *cluster* yang digunakan

dalam penelitian ini. Tahap berikutnya adalah penerapan algoritma *K-Means* untuk melakukan pengelompokan pelanggan berdasarkan nilai RFM. Hasil *clustering* yang diperoleh kemudian divisualisasikan untuk memudahkan analisis. Selanjutnya, dilakukan analisis terhadap karakteristik pada masing-masing *cluster* untuk memahami pola loyalitas pelanggan, sehingga dapat memberikan informasi yang bermanfaat sebagai dasar pengambilan keputusan bagi pihak perusahaan.



Gambar 1. Alur penelitian

3.2. Pengumpulan Data

Data penelitian ini diambil langsung dari penginapan RedDoorz Syariah Near Alun-Alun KAJEN. Pengumpulan data dilakukan melalui proses wawancara dan observasi langsung ke penginapan RedDoorz Syariah Near Alun-Alun KAJEN. Proses wawancara dilakukan langsung dengan salah satu karyawan yang sudah cukup lama bekerja di RedDoorz Syariah Near Alun-Alun KAJEN. Dari hasil wawancara yang telah dilakukan didapat banyak informasi diantaranya informasi tentang karakteristik pelanggan, sistem reservasi, dan proses pembayaran yang diterapkan. Informasi ini digunakan untuk mendukung proses segmentasi loyalitas pelanggan.

3.3. Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) yaitu proses analisis dan visualisasi data untuk mendapat pemahaman mendalam tentang informasi yang terdapat di dalam data [13]. Proses ini digunakan untuk mendapatkan gambaran awal mengenai karakteristik data, distribusi nilai, dan kondisi data secara umum. EDA juga menggunakan visualisasi untuk mengidentifikasi hubungan antar variabel. Tahap ini sangat penting untuk memastikan kualitas data sebelum melakukan pemrosesan lebih lanjut.

3.4. Preprocessing Data

Preprocessing adalah awal proses dalam data mining, dimana data mentah diubah menjadi data yang siap diproses dan digunakan [14]. Pada proses ini dilakukan *preprocessing* data seperti penghapusan data yang tidak lengkap, penghapusan data duplikat, dan perbaikan format data terutama pada atribut tanggal dan nilai transaksi. Selain itu data transaksi yang tidak valid seperti data pembayaran yang nilainya tidak tepat atau data reservasi yang tidak memiliki informasi yang lengkap juga tidak digunakan dalam proses penelitian. Setelah *preprocessing* data dilakukan selanjutnya adalah pembentukan variabel *Recency*, *Frequency*, dan *Monetary* (RFM) sebagai representasi perilaku setiap pelanggan.

3.5. Evaluasi Jumlah Cluster Optimal

Tahap ini dilakukan untuk pengelompokan jumlah pelanggan yang sesuai sebelum lanjut ke proses *clustering*. Untuk menentukan jumlah *cluster* yang optimal, evaluasi dilakukan menggunakan metode *Elbow*, *Silhouette Score*, dan *Davies-Bouldin Index*. Metode *Elbow* digunakan untuk melihat perubahan nilai WCSS pada setiap jumlah *cluster*. *Silhouette Score* digunakan untuk mengukur tingkat pemisahan pada setiap *cluster*, dan *Davies-Bouldin Index* digunakan untuk menilai kualitas hasil pengelompokan *cluster* berdasarkan tingkat kemiripannya. Jumlah *cluster* optimal ditentukan berdasarkan hasil evaluasi dari ketiga metode yang menunjukkan kualitas *cluster* yang terbaik.

a. Metode Elbow

Metode *Elbow* digunakan untuk melihat perubahan nilai WCSS pada setiap jumlah *cluster*.

$$WCSS = \sum_{k=1}^K \sum_{x \in C_k} \|x - \mu_k\|^2 \quad (1)$$

Keterangan:

K : Jumlah cluster.

x : Data dalam cluster.

C_k : Himpunan data pada cluster ke- k .

μ_k : Centroid cluster ke- k .

b. Silhouette Score

Silhouette Score digunakan untuk mengukur tingkat pemisahan pada setiap *cluster*.

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2)$$

Keterangan:

$a(i)$: Rata-rata jarak data ke- i dengan data lain dalam cluster yang sama.

$b(i)$: Jarak data minimum data ke- i dengan cluster terdekat lainnya.

c. *Davies–Bouldin Index (DBI)*

Davies–Bouldin Index digunakan untuk menilai kualitas hasil pengelompokan *cluster* berdasarkan tingkat kemiripannya.

$$DBI = \frac{1}{K} \sum_{i=1}^K \max_{j \neq i} \left(\frac{S_i + S_j}{M_{ij}} \right) \quad (3)$$

Keterangan:

K : Jumlah cluster

S_i, S_j : Jarak rata-rata data ke centroid cluster ke-i

M_{ij} : Jarak antara centroid cluster ke-i dan ke-j

3.6. *K-Means Clustering*

Penelitian ini menggunakan metode *K-Means clustering* karena dapat mengelompokkan karakteristik data berdasarkan tingkat kemiripannya. Metode ini digunakan untuk mengelompokkan pelanggan berdasarkan nilai RFM yang sudah diproses dan telah melalui tahap normalisasi. Proses *clustering* dilakukan dengan jumlah *cluster* optimal sehingga dapat menghasilkan kelompok pelanggan dengan karakteristik yang sama pada setiap *clusternya*.

a. Jarak Euclidean

K-Means clustering menggunakan perhitungan jarak euclidean untuk mengukur kedekatan data pada pusat cluster [15].

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4)$$

Keterangan:

$d(x, y)$: Jarak antara data x dan centroid y.

n : Jumlah observasi

x_i : Nilai data x pada observasi ke-i.

y_i : Titik pusat y pada observasi ke-i.

b. *Recency*

Recency digunakan untuk menghitung waktu sejak pelanggan melakukan pembelian terakhir dalam kurun waktu tertentu,

$$R_i = T_{acuan} - T_{terakhir,i} \quad (5)$$

Keterangan:

R_i : Nilai *recency* pelanggan.

T_{acuan} : Tanggal acuan atau tanggal terakhir pada data.

$T_{terakhir,i}$: Tanggal terakhir transaksi pelanggan.

c. *Frequency*

Frequency digunakan untuk menghitung jumlah transaksi yang dilakukan pelanggan dalam kurun waktu tertentu.

$$F_i = n_i \quad (6)$$

Keterangan:

F_i : Nilai frequency pelanggan

N_i : Jumlah transaksi yang dilakukan pelanggan

d. *Monetary*

Monetary digunakan untuk menghitung jumlah uang yang dikeluarkan pelanggan kepada perusahaan dalam periode tertentu.

$$M_i = \sum_{j=1}^n x_{ij} \quad (7)$$

Keterangan:

M_i : Nilai monetary pelanggan.

N : Jumlah transaksi pelanggan.

M_{ij} : Nilai transaksi ke-j milik pelanggan ke-i.

3.7. *Visualisasi dan Pemetaan Hasil Cluster*

Tahap visualisasi dilakukan untuk menampilkan hasil dari pengelompokan pelanggan yang sudah diproses secara grafis sehingga perbedaan karakteristik pada setiap *cluster* dapat dilihat dengan lebih jelas.

3.8. *Analisis Hasil*

Setelah semua proses sudah dilakukan selanjutnya melakukan klasifikasi dari hasil *clusterisasi* pelanggan menjadi ke dalam kategori “tingkat loyalitas tinggi”, “tingkat loyalitas sedang”, dan “tingkat loyalitas rendah” berdasarkan nilai rata-rata pada setiap *cluster*. *Cluster* yang mempunyai nilai *recency* yang rendah serta nilai *frequency* dan *monetary* yang lebih tinggi maka dikategorikan sebagai kelompok pelanggan dengan tingkat loyalitas tinggi. Pelanggan dengan nilai *frequency* menengah dan nilai *recency* yang lebih rendah dikategorikan tingkat loyalitas sedang. Sedangkan, *cluster* dengan nilai *recency* yang lebih tinggi serta frekuensi transaksi dan nilai pembelian yang lebih rendah maka dikategorikan sebagai kelompok pelanggan dengan tingkat loyalitas rendah. Analisis dilakukan dengan mengkaji karakteristik dari masing-masing *cluster* berdasarkan metode yang digunakan.

4. *HASIL DAN PEMBAHASAN*4.1. *Deskripsi Data*

Data penelitian ini diambil langsung dari penginapan RedDoorz Syariah Near Alun-Alun Kaje. Pengumpulan data dilakukan melalui proses wawancara dan observasi langsung ke RedDoorz Syariah Near Alun-Alun Kaje. Proses wawancara dilakukan langsung dengan salah satu karyawan yang sudah cukup lama bekerja di RedDoorz Syariah Near Alun-Alun Kaje. Dari hasil wawancara yang telah dilakukan didapat banyak informasi diantaranya informasi tentang karakteristik pelanggan, sistem reservasi, dan proses pembayaran yang diterapkan. Informasi ini digunakan untuk mendukung proses segmentasi loyalitas pelanggan.

4.2. Exploratory Data Analysis (EDA)

Proses ini digunakan untuk mendapatkan gambaran awal mengenai karakteristik data, distribusi nilai, dan kondisi data secara umum. EDA juga menggunakan visualisasi untuk mengidentifikasi hubungan antar variabel. Tahap ini sangat penting untuk memastikan kualitas data sebelum melakukan pemrosesan lebih lanjut.

*** DATA AWAL ***

Unnamed: 0	NO	NAMA	TGL CEK IN	TGL CEK OUT	HARI	YAHU	NO KAMAR	JUMLAH TOTAL	OFFLINE	TRANSFER	CAS TUNAI	CAS TIF	P.A.H TUNAI	P.A.H TIF
0	NAN	NAN	NAN	NAN	168	179	NAN	32314277	12004174.0	7620566.0	3324198.0	3021916.0	2800000.0	2800000.0
1	1.0	PT MAI	2025-10-01	2025-10-09	8	1	109.0	1440000	NAN	1440000.0	NAN	NAN	NAN	NAN
2	2.0	PT MAI	2025-10-01	2025-10-03	2	1	104.0	300000	NAN	300000.0	NAN	NAN	NAN	NAN
3	3.0	NAN	2025-10-01	2025-10-02	1	2	106.0	200000	NAN	200000.0	NAN	NAN	NAN	NAN
4	4.0	NAN	2025-10-01	2025-10-02	1	2	106.0	200000	NAN	200000.0	NAN	NAN	NAN	NAN

Gambar 2. Data pelanggan sebelum diolah

Data mentah yang sudah dikumpulkan lalu dimulai tahap pra pemrosesan untuk membersihkan dan merapikan data agar dapat diolah lebih efektif.

4.3. Preprocessing Data

Data yang menjadi subjek pada penelitian ini diperoleh dari data internal pada penginapan RedDoorz Syariah Near Alun-Alun Kaje.

Tabel 1. Data pelanggan

No	Atribut	Keterangan
1	Nama	Nama pelanggan yang melakukan reservasi
2	Tanggal Cek In	Tanggal awal pelanggan melakukan reservasi atau mulai menginap
3	Tanggal Cek Out	Tanggal akhir reservasi pelanggan selesai menginap
4	Jumlah Total	Total biaya yang dibayarkan pelanggan

Data yang digunakan mencakup empat atribut utama yaitu nama pelanggan, tanggal cek in dan tanggal cek out, serta jumlah total pembayaran pelanggan.

4.4. Metode RFM (Recency, Frequency, Monetary)

*** Hasil RFM:

	NAMA	RECENCY	FREQUENCY	MONETARY
0	adi	27	1	250000
1	ahmad	26	1	200000
2	alamasyah	24	1	250000
3	aldi	20	1	275000
4	alfin	4	1	287500
..	...	...	...	...
99	windari	13	2	180000
100	wisnu	5	2	400000
101	wulan	16	1	200000
102	yusuf	20	2	188066
103	zainudin	30	1	180000

Gambar 3. Output data RFM

Pada tahap selanjutnya, melakukan proses segmentasi pelanggan berdasarkan RFM (*Recency*, *Frequency*, dan *Monetary*) dari data yang sudah dipilih. Nilai *recency* dihitung dari selisih antara tanggal terakhir kunjungan pelanggan dan tanggal

terbaru. Nilai *frequency* dihitung dari jumlah transaksi yang dilakukan pelanggan dalam kurun waktu tertentu. Nilai *monetary* dihitung dari jumlah uang yang dikeluarkan pelanggan kepada perusahaan dalam periode tertentu.

Setelah selesai mencari nilai RFM, tahap selanjutnya yaitu mencari nilai minimum, maksimum, dan rata-rata dari variabel *recency*, *frequency*, *monetary*.

***	Variable	Nilai_Min	Nilai_Max	Nilai_Rata
0	RECENCY	0	30	15.865385
1	FREQUENCY	1	5	1.086538
2	MONETARY	160000	4680000	310714.201923

Gambar 4. Nilai minimum, maksimum, dan rata-rata RFM

Setelah semua proses sudah selesai dilakukan dari mencari nilai RFM sampai mencari nilai minimum, maksimum, dan rata-rata selanjutnya dilakukan proses normalisasi tujuannya untuk menghindari perbedaan data yang signifikan pada nilai RFM setiap pelanggan.

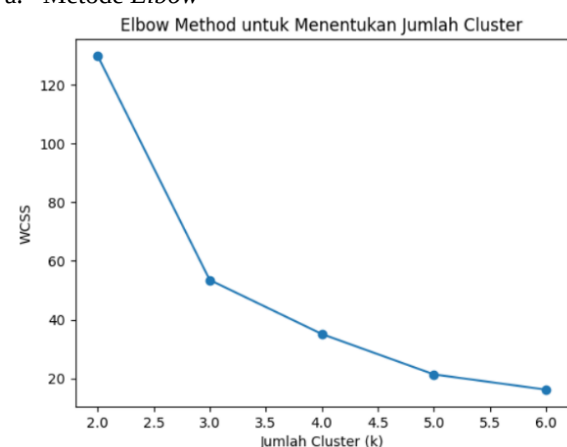
	RECENCY	FREQUENCY	MONETARY
0	1.339461	-0.238806	-0.130423
1	1.222986	-0.238806	-0.237830
2	0.990036	-0.238806	-0.130423
3	0.524137	-0.238806	-0.076719
4	-1.339461	-0.238806	-0.049867

Gambar 5. Data RFM yang sudah dinormalisasi

4.5. Evaluasi Jumlah Cluster Optimal

Tahap ini dilakukan untuk pengelompokan jumlah pelanggan yang sesuai. Untuk menentukan jumlah *cluster* yang optimal, evaluasi dilakukan menggunakan metode *Elbow*, *Silhouette Score*, dan *Davies-Bouldin Index*.

a. Metode Elbow



Gambar 6. Hasil metode elbow

Pada grafik ini menunjukkan penurunan nilai presentase yang sangat signifikan pada jumlah

cluster 2. Setelah terjadi penurunan yang signifikan pada titik tersebut penurunan mulai melambat dan membuat jumlah *cluster* tidak mengalami tanda-tanda peningkatan yang signifikan. Oleh karena itu, menurut metode ini jumlah *cluster* yang optimal adalah 2.

b. Silhouette Score

Silhouette Score memiliki rentang nilai dari -1 hingga 1 dan digunakan untuk mengukur tingkat pemisahan pada setiap *cluster*, dimana nilai yang mendekati 1 merupakan kualitas pengelompokan *cluster* yang terbaik. Dari perhitungan yang sudah didapat, jumlah *cluster* 2 yang memiliki nilai tertinggi 0.892 dan menjadi kualitas *cluster* yang terbaik.

Jumlah Cluster = 2, Silhouette Score = 0.892
 Jumlah Cluster = 3, Silhouette Score = 0.525
 Jumlah Cluster = 4, Silhouette Score = 0.550
 Jumlah Cluster = 5, Silhouette Score = 0.476
 Jumlah Cluster = 6, Silhouette Score = 0.431

Gambar 7. Hasil Silhouette Score

c. Davies–Bouldin Index (DBI)

Dari hasil perhitungan yang didapat, jumlah *cluster* 2 merupakan nilai DBI terendah 0.071, dimana pada perhitungan DBI semakin rendah nilai DBI, maka semakin baik pemisahan antar *cluster* dan kepadatan *cluster* yang tinggi.

Jumlah Cluster = 2, DBI = 0.071
 Jumlah Cluster = 3, DBI = 0.468
 Jumlah Cluster = 4, DBI = 0.455
 Jumlah Cluster = 5, DBI = 0.515
 Jumlah Cluster = 6, DBI = 0.604

Gambar 8. Hasil *davies–bouldin index* (DBI)

4.6. K-Means Clustering

Proses *clustering* dilakukan dengan jumlah *cluster* optimal sehingga dapat menghasilkan kelompok pelanggan dengan karakteristik yang sama pada setiap *cluster*nya.

...	NAMA	REGENCY	FREQUENCY	MONETARY	CLUSTER
0	adi	27	1	250000	0
1	ahmad	26	1	200000	0
2	alamsyah	24	1	250000	0
3	aldi	20	1	275000	0
4	alfin	4	1	287500	2

Gambar 9. Hasil akhir *clustering*

Hasil proses *clustering* menunjukkan bahwa data pelanggan dapat dikelompokkan berdasarkan karakteristik nilai *Recency*, *Frequency*, dan *Monetary*. Proses pengelompokan ini menghasilkan beberapa *cluster* berdasarkan perilaku transaksi pelanggan yang berbeda. Pelanggan yang berada pada *cluster* yang

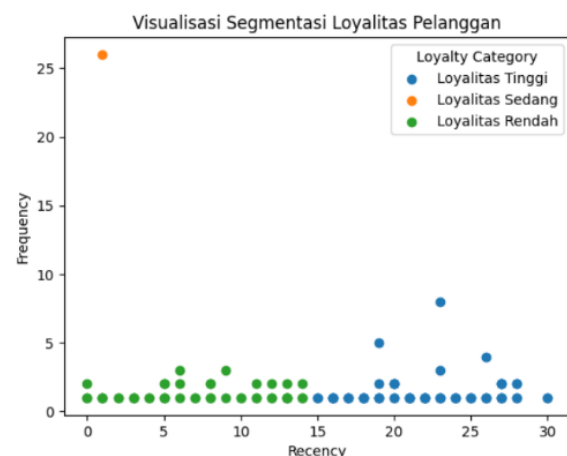
sama memiliki pola transaksi yang mirip, sedangkan pelanggan yang berada pada *cluster* yang berbeda memiliki karakteristik transaksi yang berbeda. Dengan adanya pengelompokan ini perusahaan dapat lebih mudah dalam mengidentifikasi kelompok pelanggan berdasarkan tingkat aktivitas dan kontribusi nilai transaksi terhadap perusahaan.

...	REGENCY	FREQUENCY	MONETARY
CLUSTER			
0	22.137931	1.431034	2.846883e+05
1	1.000000	26.000000	4.680000e+06
2	7.266667	1.311111	2.471635e+05

Gambar 10. Hasil *centroid*

Nilai *centroid* pada setiap *cluster* menggambarkan nilai rata-rata dari atribut *Recency*, *Frequency*, dan *Monetary* pada kelompok pelanggan tersebut. Nilai ini digunakan untuk mengetahui karakteristik dari setiap *cluster*. *Cluster* yang memiliki nilai *recency* lebih rendah serta nilai *frequency* dan *monetary* yang lebih tinggi menunjukkan kelompok pelanggan yang lebih aktif dalam melakukan transaksi dan memiliki kontribusi pendapatan yang lebih besar untuk perusahaan. Sedangkan *cluster* dengan nilai *recency* lebih tinggi serta nilai *frequency* dan *monetary* yang lebih rendah menunjukkan kelompok pelanggan yang jarang melakukan transaksi dan memiliki tingkat loyalitas lebih rendah. Dengan adanya informasi ini pihak manajemen dapat memanfaatkannya untuk menentukan strategi pelayanan dan pemasaran yang berbeda pada setiap pelanggan.

4.7. Visualisasi dan Pemetaan Hasil Cluster



Gambar 11. Visualisasi *scatterplot*

Berdasarkan visualisasi pada segmentasi loyalitas pelanggan menggunakan *scatterplot* nilai *recency* dan nilai *frequency*, segmentasi loyalitas pelanggan dibagi menjadi ke dalam kategori loyalitas tinggi, sedang, dan kategori loyalitas rendah. *Cluster* yang mempunyai nilai *recency* yang rendah serta nilai

frequency dan *monetary* yang lebih tinggi maka dikategorikan sebagai kelompok pelanggan dengan tingkat loyalitas tinggi. pelanggan dengan nilai *frequency* menengah dan nilai *recency* yang lebih rendah dikategorikan tingkat loyalitas sedang. Sedangkan, *cluster* dengan nilai *recency* yang lebih tinggi serta frekuensi transaksi dan nilai pembelian yang lebih rendah maka dikategorikan sebagai kelompok pelanggan dengan tingkat loyalitas rendah.

Berdasarkan visualisasi pada jumlah data pada setiap *cluster*, bahwa *cluster* 0 memiliki jumlah pelanggan yang paling banyak yaitu 58, lalu *cluster* 2 dengan jumlah pelanggan 45 dan *cluster* 1 memiliki jumlah *cluster* paling sedikit yaitu 1.



Gambar 12. Visualisasi jumlah data pada setiap *cluster*

5. KESIMPULAN DAN SARAN

Dari hasil penelitian yang sudah dilakukan, metode RFM (*Recency*, *Frequency*, dan *Monetary*) yang digabungkan dengan algoritma *K-Means Clustering* dapat melakukan segmentasi pelanggan pada data reservasi di RedDoorz Syariah Near Alun-Alun Kajan secara efektif. Proses penelitian dimulai dari tahap *Exploratory Data Analysis* (EDA), *preprocessing* data, pembentukan variabel RFM, normalisasi data, dan menentukan jumlah *cluster* optimal menggunakan metode *Elbow*, *Silhouette Score*, dan *Davies-Bouldin Index* yang menghasilkan tiga *cluster* yaitu kategori loyalitas tinggi sebanyak 58 pelanggan, kategori loyalitas sedang sebanyak 45 pelanggan, dan kategori loyalitas rendah sebanyak 1 pelanggan. Dari hasil segmentasi pelanggan yang sudah dilakukan pada penelitian ini dapat diimplementasikan ke dalam sistem informasi manajemen untuk penyusunan strategi pelayanan dan pemasaran agar lebih tepat sasaran, serta dalam memberikan perilaku khusus kepada pelanggan dengan tingkat loyalitas tertentu.

DAFTAR PUSTAKA

- [1] J. Pariwisata, "PERAN MANAGER DALAM MENJALIN KERJASAMA DENGAN ONLINE TRAVEL AGENT UNTUK MENINGKATKAN TINGKAT," vol. 4, pp. 198–209, 2023.
- [2] U. N. Makassar, "Strategi Inovatif untuk Meningkatkan Kunjungan dan Loyalitas Tamu di Hotel RedDoorz Makassar," pp. 66–74, 2025.
- [3] M. Informatika, F. Sains, and U. Labuhanbatu, "No Title," vol. 1, pp. 7–17, 2021.
- [4] D. Jollyta, "Segmentasi Pelanggan Hotel Labersa Untuk Mengetahui Tingkat Loyalitas Pelanggan Menggunakan Algoritma K-Means Dan RFM (Recency, Frequency, Monetary)," vol. 4, no. 3, 2022.
- [5] G. Shop, G. Shop, G. Shop, G. Shop, K. Kunci, and O. Shop, "STRATEGI PEMASARAN BERORIENTASI PELANGGAN PADA UMKM PENJUALAN ONLINE GULIGULI SHOP CUSTOMER ORIENTED MARKETING STRATEGY FOR UMKM ONLINE SALES GULIGULI SHOP," vol. 2, no. 3, pp. 143–149, 2022.
- [6] R. Nugraha, N. Safriadi, D. Prawira, U. Tanjungpura, P. Studi, and S. Informasi, "Rancang Bangun Sistem CRM (Customer Relationship Management) Pada Usaha Pijat Refleksi Berbasis Web Dengan Pola MVC," vol. 9, no. 1, pp. 70–85, 2022.
- [7] C. Series, "A review of data mining methods in RFM-based customer segmentation A review of data mining methods in RFM-based customer segmentation," 2021, doi: 10.1088/1742-6596/1869/1/012085.
- [8] M. Afdal and A. Marsal, "Business to Business Customer Loyalty Analysis Based on RFM Model Using Fuzzy C-Means Algorithm Analisis Loyalitas Pelanggan Business To Business Berdasarkan Model RFM Menggunakan Algoritma Fuzzy C-Means," vol. 4, no. April, pp. 359–365, 2024.
- [9] L. Rahmadhani, A. Djunaidy, and A. Mukhlason, "Evaluasi Kinerja Pemasok Menggunakan Fuzzy C-Means Clustering dan AHP di CV Delta Raya," vol. 10, no. 2, 2021.
- [10] F. Arifianto, J. Hasudungan, A. Muzaky, and H. T. Y. Achsan, "Segmentasi Pelanggan Berdasarkan Recency, Frequency, dan Monetary dengan K-Means Clustering: Studi Kasus Toko Pakaian Almost Famous komponen penting dari strategi bisnis modern. Dalam era komputer dan internet saat ini, untuk mengumpulkan informasi berharga. Pengumpulan data transaksi dengan teknologi ini faktor kritis, termasuk frekuensi pembelian, jumlah belanja, dan kebaruan. Meskipun metode pengawasan yang populer untuk mengelompokkan data. Algoritma ini menggunakan," vol. 10, no. 1, pp. 122–135, 2024.
- [11] A. Hermawan, N. R. Jayanti, A. Saputra, and C. Tambunan, "Optimalisasi Strategi Pemasaran Melalui Analisis RFM pada Dataset Transaksi Ritel Menggunakan Python," no. 4, 2024.

- [12] R. A. Herdianti, W. Syaifullah, J. Saputra, and D. A. Prasetya, "Pemetaan Karakteristik Wilayah Miskin di Jawa Timur Menggunakan Spectral Clustering untuk Penentuan Target Kebijakan," vol. 10, no. 2, pp. 362–375, 2025.
- [13] J. Prima *et al.*, "ANALISIS BIG DATA DENGAN METODE EXPLORATORY DATA ANALYSIS (EDA) DAN METODE VISUALISASI MENGGUNAKAN JUPYTER NOTEBOOK," vol. 4, no. 2, pp. 2–6, 2021.
- [14] R. Forest and C. Smote, "Optimasi Klasifikasi Sentimen Menggunakan," vol. 10, no. 3, pp. 389–400, 2024.
- [15] B. A. Pradana *et al.*, "Implementasi algoritma k-means++ untuk klasterisasi daerah rawan stunting di kabupaten tasikmalaya," vol. 9, no. 6, pp. 9274–9280, 2025.