ZONAsi Jurnal Sistem Informasi
ZONAsi VOL. 8 NO. 3 September 2026
Page: 1266 - 1274 Akreditasi Sinta 4
ISSN: 2656-7407 (Online)
2656-7393 (Print)

SEGMENTASI PELANGGAN PT BIN SEF TRAVEL DENGAN HYBRID RFM DAN ALGORITMA K-MEANS

Ilham Amani Saiful Haq[1], Nurmalitasari [2], Afu Ichsan Pradana [3]
Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Duta Bangsa
(Jl. Bhayangkara No.55, Tipes, Kec. Serengan, Kota Surakarta, Jawa Tengah, Indonesia)
e-mail: ilhamamanish@gmail.com

Abstrak

Pemanfaatan data transaksi pelanggan secara optimal merupakan tantangan utama yang dihadapi PT Bin Sef Travel dalam merancang strategi pemasaran yang tepat sasaran. Penelitian ini bertujuan mengidentifikasi segmen pelanggan menggunakan kombinasi model RFM (Recency, Frequency, Monetary) dan algoritma K-Means Clustering dengan kerangka pengembangan CRISP-DM. Data yang digunakan adalah 448 catatan transaksi dari 68 pelanggan periode September hingga Desember 2025. Tahapan penelitian meliputi transformasi data ke variabel RFM, normalisasi menggunakan StandardScaler, penentuan jumlah klaster optimal melalui Silhouette Score dan Metode Elbow, serta implementasi K-Means Clustering. Hasil pengujian menunjukkan jumlah klaster optimal K=4 dengan Silhouette Score sebesar 0,4143. Keempat segmen yang terbentuk adalah Champions (1 pelanggan, 1,5%), Loyal Customers (29 pelanggan, 42,6%), Potential Loyalist (32 pelanggan, 47,1%), dan At Risk (6 pelanggan, 8,8%). Segmen Loyal Customers menjadi kontributor pendapatan terbesar bagi perusahaan, sementara segmen At Risk memerlukan strategi reaktivasi segera akibat rendahnya frekuensi transaksi dan nilai kebaruan yang tinggi. Hasil penelitian ini memberikan landasan berbasis data bagi manajemen PT Bin Sef Travel untuk beralih dari pemasaran massal menuju strategi pemasaran yang dipersonalisasi sesuai karakteristik masing-masing segmen pelanggan.

Kata kunci*:* *Segmentasi Pelanggan, RFM, K-Means, CRISP-DM*

Abstract

PT Bin Sef Travel faces challenges in optimally utilizing customer transaction data to design targeted marketing strategies. This study aims to identify customer segments using a combination of the RFM (Recency, Frequency, Monetary) model and K-Means Clustering algorithm within the CRISP-DM development framework. The dataset comprises 448 transaction records from 68 customers spanning September to December 2025. The research stages include RFM variable transformation, data normalization using StandardScaler, optimal cluster determination through Silhouette Score and Elbow Method, and K-Means Clustering implementation. Results indicate an optimal cluster number of K=4 with a Silhouette Score of 0.4143. Four segments were identified: Champions (1 customer, 1.5%), Loyal Customers (29 customers, 42.6%), Potential Loyalist (32 customers, 47.1%), and At Risk (6 customers, 8.8%). The Loyal Customers segment contributes the highest revenue to the company, while the At Risk segment requires immediate reactivation strategies due to low transaction frequency and high recency values. These findings provide a data-driven foundation for PT Bin Sef Travel management to shift from mass marketing toward personalized marketing strategies tailored to each customer segment.

Keywords: *Customer Segmentation, RFM Model, K-Means Clustering, CRISP-DM.*

1. PENDAHULUAN

Industri pariwisata dan perjalanan di Indonesia telah berkembang pesat selama sepuluh tahun terakhir. Hal ini disebabkan oleh semakin banyaknya orang yang tertarik pada wisata religi maupun pariwisata umum. Persaingan di sektor ini pun semakin ketat. Karena itu, perusahaan jasa travel mulai menerapkan berbagai metode pemasaran. Mereka tidak lagi menggunakan metode yang sama seperti yang mereka terapkan di masa Industri pariwisata bisnis mengalami kemajuan pesat berkat strategi pemasaran digital yang mampu memenuhi ekspektasi para penggemar wisata yang terus berubah.Analisis data dan personalisasi berbasis kecerdasan buatan mulai mengubah tingkat daya saing perusahaan serta kinerja keuangan mereka. Pergeseran paradigma pemasaran dari pemasaran massal ke pemasaran yang disesuaikan dengan target pasar kini tak terhindarkan lagi, mengingat perilaku dan preferensi pelanggan yang semakin beragam. Dalam konteks ini, pemahaman mendalam terhadap karakteristik pelanggan menjadi landasan utama dalam merancang strategi pemasaran yang efektif, efisien, dan berkelanjutan[1].

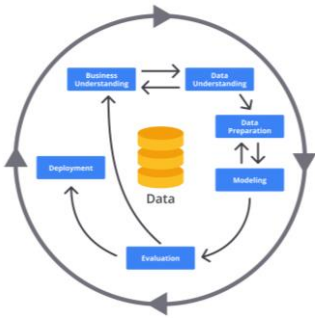
PT Bin Sef Travel adalah biro perjalanan yang bergerak di bidang layanan umrah. Biro perjalanan ini menghadapi tantangan besar dalam mengelola basis pelanggannya yang terus bertambah. Berkurangnya jumlah jemaah umrah belakangan ini menunjukkan bahwa diperlukan strategi pemasaran baru. Strategi ini harus memanfaatkan data untuk mempertahankan pelanggan dan memperluas pasar. Data transaksi pelanggan yang telah dikumpulkan belum dimanfaatkan secara optimal untuk mendukung pengambilan keputusan pemasaran. Akibatnya, sumber daya promosi sering kali digunakan secara tidak tepat. Semua pelanggan diperlakukan sama, tanpa mempertimbangkan nilai masing-masing dan pola pembelian mereka. Hal ini menyebabkan perusahaan memperoleh pendapatan yang lebih sedikit[2][3].

Penelitian terdahulu yang dilakukan oleh Hermawan [4].Dengan judul "Analisis Segmentasi Pelanggan Berbasis RFM dan Evaluasi Efektivitas Kampanye Pemasaran untuk Meningkatkan Retensi" temuan utama studi ini menunjukkan bahwa Kampanye 6 merupakan kampanye yang paling efektif ,Studi ini berpendapat bahwa pendekatan pemasaran yang seragam tidak efektif dan bahwa setiap segmen memerlukan strategi yang disesuaikan secara khusus. Penelitian kedua dilakukan oleh Nugraha dan Priti [5]. Berjudul "Analisis Segmentasi Customer dengan RFM dan K-Means Clustering: Studi E-Commerce PT PQR" Penelitian ini mengaplikasikan analisis RFM dan K-Means pada 56.999 data transaksi e-commerce PT PQR (2018–2020) dan menghasilkan tiga segmen: top, medium, dan low. Analisis menunjukkan adanya perbedaan preferensi produk, metode pembayaran, dan demografi antar klaster. Secara operasional, ditemukan kerugian refund sebesar Rp218,8 juta (6,5% total order) akibat kegagalan seller memproses pesanan yang telah dibayar.

Untuk mengatasi tantangan ini, pendekatan berbasis data dalam segmentasi pelanggan menawarkan solusi dengan membagi pelanggan ke dalam kelompok berdasarkan karakteristik, perilaku, atau kebutuhan yang sama. Hal ini sangat penting dilakukan, karena setiap pelanggan memiliki preferensi yang berbeda-beda [6]. Perusahaan dapat menggunakan model RFM dan algoritma clustering K-Means untuk melakukan segmentasi pelanggan. Hal ini memudahkan untuk mengumpulkan, menganalisis, memproses, dan memvisualisasikan data pelanggan[7]. Model *Recency Frequency Monetary* (RFM) adalah salah satu cara untuk menganalisis potensial pelanggan, Analisis RFM memprediksi perilaku konsumen berdasarkan *Recency* atau jangka waktu(R), *Frequency* atau jumlah transaksi (F), dan *Monetary* atau total jumlah uang (M) [8][9]. Algoritma K-Means adalah metode untuk membagi sekumpulan data pengamatan menjadi beberapa kelompok berdasarkan seberapa dekat data tersebut dengan pusat rata-rata (centroid),Kemudian Algoritma K-Means ini membagi data berdasarkan seberapa jauh jarak antar elemennya dan membentuk kluster berdasarkan pola perilaku pembelian [10][11][12].

2. METODE PENELITIAN

CRISP-DM (Cross Industry Standard Process for Data Mining) CRISP-DM adalah suatu pendekatan untuk merancang dan mengelola proyek penambangan data. Metodologi ini membagi proses analisis menjadi enam fase utama:businessunderstanding ,understanding data,Data preparation, modeling, Evaluation,dan Deployment[13].



Gambar 1. Enam Fase Utama Kerangka Kerja CRISP-DM dalam Proses Data Mining

Penerapan CRISP-DM dalam penelitian ini didasarkan pada kemampuannya untuk menyediakan kerangka kerja terstruktur yang mengarahkan proses penelitian[14].Kerangka kerja ini membantu mengarahkan proses penelitian mulai dari tahap identifikasi masalah bisnis hingga evaluasi hasil akhir implementasi opsional.

2.1 Business Understanding

Pada tahap ini akan membahas tantangan besar yang dihadapi PT Bin Sef Travel. Masalah utama yang teridentifikasi adalah strategi pemasaran perusahaan belum cukup memanfaatkan data transaksi pelanggan historis. Perusahaan masih menerapkan pendekatan yang sama untuk semua pelanggan. Artinya, mereka tidak memperhitungkan seberapa besar pengeluaran masing-masing pelanggan atau seberapa besar transaksi yang mereka lakukan.

2.2 Data Understanding

Penelitian ini menganalisis sekitar 448 catatan transaksi pelanggan dari basis data internal PT Bin Sef Travel. Data ini diambil dari periode September hingga Desember 2025. Untuk memastikan kelayakan data ,analisis dilakukan untuk memeriksa struktur, distribusi, dan kualitas dataset. Tahap ini mencakup penanganan missing values, deteksi outlier, serta pengelompokan pola transaksi pelanggan secara menyeluruh.

Tabel 1. Deskripsi Variabel Data Transaksi

Variabel	Keterangan	
ID Pelanggan	Kode unik identitas pelanggan	Recency, Frequency, Monetary
Tanggal Transaksi	Tanggal pelanggan melakukan pemesanan/pembayaran	Recency, Frequency
Jumlah Transaksi	Banyaknya transaksi yang dilakukan pelanggan	Frequency
Nilai Pembayaran (Rp)	Total nominal yang dibayarkan pelanggan per transaksi	Monetary

2.3 Data Preparation

Tahap persiapan data merupakan fase terpanjang dalam CRISP-DM yang mencakup tiga sub-proses utama :

a) Data Cleaning

Sangat penting untuk membersihkan data sebelum menganalisisnya guna memastikan validitas dan kualitasnya. Prosedur ini berfokus pada empat langkah utama, yakni:(1) Memperbaiki nilai yang hilang dengan menggunakan perkiraan atau menghapusnya,(2) Menghapus duplikat catatan untuk menghindari masalah akibat penghitungan ganda,(3) Mencari dan menangani nilai-nilai yang sangat

menyimpang dengan menggunakan diagram kotak, dan (4) Memastikan format seluruh datanya seragam[15].

b) Transformasi Data ke Model RFM

Setelah data bersih, dilakukan transformasi variabel mentah menjadi tiga atribut RFM. Tanggal referensi analisis ditetapkan sebagai hari setelah tanggal transaksi terakhir dalam dataset.

Tabel 2 Rumus Perhitungan Variabel RFM

Variabel	Rumus / Cara Hitung
Recency (R)	R = Tanggal Referensi - Tanggal Transaksi Terakhir Pelanggan (dalam hari)
Frequency (F)	F = Jumlah total transaksi pelanggan selama periode penelitian
Monetary (M)	M = Total nilai pembayaran pelanggan selama periode penelitian (Rp)

c) Normalisasi Data

Normalisasi memastikan bahwa semua variabel berada pada skala yang sama. Hal ini membantu memastikan bahwa tidak ada satu pun variabel yang mendominasi proses pengelompokan akibat perbedaan satuan dan rentang nilai. Penelitian ini menggunakan metode yang disebut "normalisasi min-max." Ini adalah proses yang mengubah setiap nilai sehingga nilainya berada dalam rentang [0,1]

$$X' = (X - X_min) / (X_max - X_min)$$

nilai X adalah nilai asli variabel, X_min dan X_max setiap nilai minimum dan maksimum variabel tersebut dalam dataset.

2.4 Modeling

Tahap pemodelan terdiri dari dua Langkah :

a) Metode Elbow

Untuk menentukan jumlah kluster (K) yang ideal secara objektif, dilakukan pengujian dengan menganalisis grafik jumlah kuadrat dalam kluster (WCSS). Nilai K yang optimal dapat ditentukan dengan mengamati di mana titik "siku" muncul ketika berbagai nilai K diplot pada kurva.

b) Implementasi K-Means Clustering

Dengan merujuk pada nilai K dari pengujian Elbow, algoritma K-Means dijalankan melalui langkah-langkah berikut:

1) Menetapkan K titik centroid secara acak sebagai inisialisasi awal.
2) Menghitung jarak setiap titik data terhadap seluruh centroid menggunakan Euclidean Distance.
3) Mengalokasikan setiap data ke dalam klaster dengan centroid terdekat.
4) Mengkalkulasi ulang posisi centroid berdasarkan nilai rata-rata data di dalam klaster tersebut.
5) Mengulang proses (langkah 2 hingga 4) sampai centroid konvergen dan keanggotaan klaster tidak lagi mengalami perubahan.

2.5 Evaluation

Untuk menilai kualitas hasil pengelompokan, penelitian ini menggunakan metrik Silhouette Score. Evaluasi ini bertujuan untuk menilai keakuratan penempatan data dalam suatu kelompok jika dibandingkan dengan kelompok-kelompok lainnya[15], dengan menggunakan persamaan berikut

$$S(i) = [b(i) - a(i)] / max\{a(i), b(i)\}$$

Dalam rumus ini, *a(i)* mewakili jarak rata-rata antara titik data ke-i dan semua anggota kluster yang sama, sedangkan *b(i)* adalah jarak rata-rata titik data ke- *(i)* ke kluster terdekat. Rentang hasil evaluasi ini berkisar antara -1 dan 1. Jika skor akhir mendekati 1, hal ini menandakan bahwa data telah dikelompokkan dengan sangat baik dan akurat.

2.6 Perangkat yang digunakan

Penelitian ini menggunakan berbagai perangkat lunak untuk memproses dan menganalisis data pada setiap tahap. Python adalah bahasa pemrograman utama yang menjalankan seluruh proses komputasi[16]. Proses ini memanfaatkan pustaka Pandas dan NumPy untuk memodifikasi, mentransformasi, dan menstandarkan data RFM.Menggunakan pustaka Scikit-learn untuk menjalankan algoritma K-Means dan perhitungan Silhouette Score. Untuk menampilkan hasilnya secara visual, selanjutnya pustaka Matplotlib dan Seaborn untuk membuat grafik Elbow, grafik distribusi klaster, dan diagram kotak RFM. Selain alat berbasis kode ini, file CSV dengan Microsoft Excel juga digunakan pada tahap awal untuk menyiapkan data mentah dan memeriksa keakuratan hasil transformasi model RFM.

3. HASIL DAN PEMBAHASAN

3.1 Hasil

3.1.1 Eksplorasi Data

Dataset yang digunakan adalah data invoice PT Bin Sef Travel periode 17 September 2025 hingga 25 Desember 2025, terdiri dari 448 transaksi yang melibatkan 68 pelanggan unik.Distribusi transaksi per bulan disajikan pada gambar tabel berikut:

Bulan	Jumlah Transaksi	Total Pendapatan
September 2025	36	Rp 1,202
Oktober 2025	9	Rp 0,332
November 2025	220	Rp 28,021
Desember 2025	183	Rp 23,45

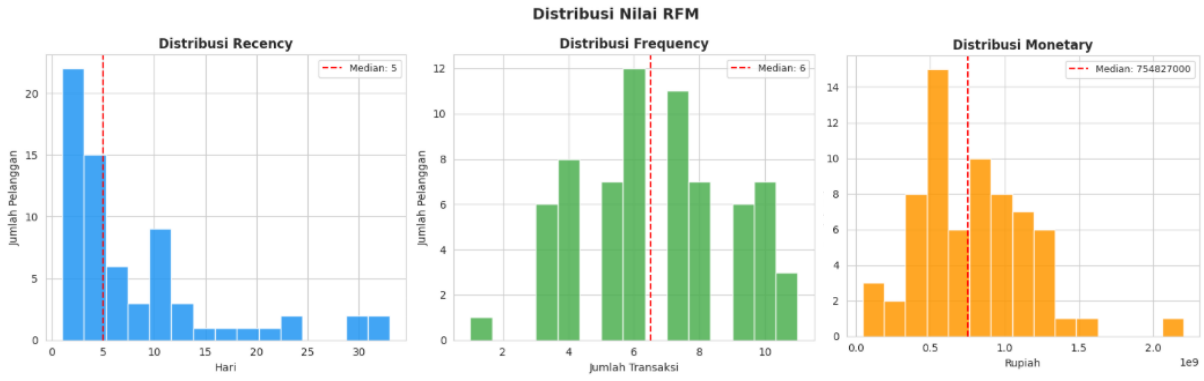
Gambar 2. Distribusi Jumlah Transaksi dan Total Pendapatan per Bulan

3.1.2 Hasil perhitungan RFM

Dari 68 pelanggan, dihitung nilai Recency, Frequency, dan Monetary masing-masing. Statistik deskriptif RFM disajikan pada gambar tabel berikut:

count	68.00	68.00	6.800000e+01
mean	8.13	6.59	7.796232e+08
std	7.90	2.36	3.768056e+08
min	1.00	1.00	4.371300e+07
25%	3.00	5.00	5.237062e+08
50%	5.00	6.50	7.548270e+08
75%	11.00	8.00	1.022342e+09
max	33.00	11.00	2.203160e+09

Gambar 3. Grafik Distribusi Nilai Recency, Frequency, dan Monetary



Gambar 4. Penentuan Jumlah Klaster Optimal dan Proporsi Pelanggan per Segmen

Secara umum, pelanggan tergolong aktif, dengan transaksi terbaru (median Recency 5 hari) dan frekuensi transaksi yang tinggi (Frequency 6 kali). Namun, distribusi nilai moneter (median Rp754,8 juta) condong ke kanan karena adanya segelintir pelanggan dengan transaksi yang sangat besar (outlier alami). Ketidakseimbangan dalam profil ini merupakan alasan kuat mengapa algoritma K-Means diperlukan untuk memisahkan pelanggan bernilai tinggi dari pelanggan biasa secara objektif dan tepat

3.1.3 Normalisasi Data

	Recency	Frequency	Monetary
0	0.1107	-0.2507	-0.2441
1	2.7892	1.4539	3.8060
2	-0.6546	-0.2507	0.5982
3	0.3658	1.0277	1.0654
4	2.0239	-1.5291	-1.5381

Mean: [-0. 0. 0.] (harusnya ~0)
Std : [1.0074 1.0074 1.0074] (ha

Gambar 5. Hasil dari normalisasi

Proses normalisasi berhasil menyeimbangkan skala ketiga variabel RFM, yang dibuktikan dengan nilai rata-rata mendekati 0 dan simpangan baku mendekati 1 (1,0074). Transformasi ini mutlak diperlukan mengingat variabel asal memiliki rentang yang sangat bervariasi (hari, jumlah transaksi, hingga miliaran rupiah). Tanpa standardisasi, besaran variabel Monetary akan mendominasi komputasi K-Means dan merusak validitas klaster. Setelah dinormalisasi, nilai positif dan negatif berfungsi sebagai indikator posisi pelanggan terhadap nilai rata-rata keseluruhan. Contohnya, pelanggan ke-1 dengan Recency positif 2,7892 dan Monetary positif 3,8060 menunjukkan karakteristik pelanggan yang sudah lama tidak bertransaksi, namun memiliki riwayat pengeluaran yang jauh melampaui rata-rata pelanggan lainnya. Melalui pendekatan ini, seluruh variabel dipastikan memberikan kontribusi yang adil dan berimbang dalam penentuan segmen.

3.1.4 Penentuan Jumlah Klaster Optimal

Pengujian dilakukan untuk nilai K menggunakan Silhouette Score dan Metode Elbow secara bersamaan

K	Silhouette Score	Inertia
2	0.3736	119.21
3	0.4115	76.29
4	0.4143	59.73
5	0.3352	45.21
6	0.3303	38.30
7	0.3283	31.48
8	0.3356	26.57
9	0.3222	23.56
10	0.3291	19.94

✅ K Optimal: 4 | Silhouette Score: 0.4143

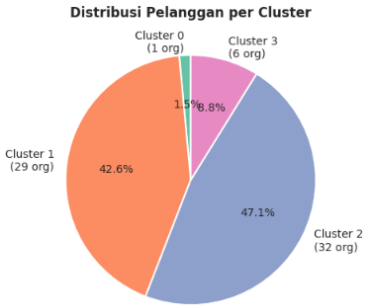
Gambar 6. Nilai silhouette score untuk menentukan Cluster

Nilai Silhouette tertinggi diperoleh pada K = 4, dengan nilai 0,4143, demikian jumlah kluster optimal yang digunakan dalam penelitian ini adalah 4

3.1.5 Hasil K-Means

K-Means dijalankan dengan K = 4,tersebut menghasilkan hasil sebagai berikut :

Cluster	Jumlah_Pelanggan	Recency_Avg	Frequency_Avg	Monetary_Avg	Monetary_Total	Monetary_Avg_Juta	Monetary_Total_Miliar
0	1	30.0	10.0	2.203160e+09	2203160500	2203.2	2.20
1	29	6.6	8.6	1.071221e+09	31065412200	1071.2	31.07
2	32	5.2	5.2	5.497218e+08	17591096550	549.7	17.59
3	6	27.2	3.5	3.591177e+08	2154706000	359.1	2.15



Gambar 7. Hasil clustering mendapatkan 4 segmen

Output yang diluarkan berhasil mengklasifikasikan basis pelanggan ke dalam empat segmen yang secara sistematis menggambarkan keragaman perilaku transaksi. Hasil perhitungan menunjukkan bahwa mayoritas populasi terkonsentrasi pada segmen Pelanggan Loyal Potential (47,1%), yang menunjukkan tingkat aktivitas interaksi terkini tertinggi, disusul oleh segmen Loyal Customers(42,6%), yang menjadi pilar utama profitabilitas berkat frekuensi dan nilai pembelian yang tinggi. Di selain kelompok ini, pengelompkan ini juga secara logis mengidentifikasi segmen At Risk (8,8%), yang rentan terhadap churn akibat ketidakaktifan yang berkepanjangan, serta segmen Champions(1,5%), yang mewakili outlier positif dengan kontribusi finansial maksimum. Pemetaan kuantitatif objektif ini mengonfirmasi relevansi teori segmentasi modern, sekaligus menyediakan landasan strategis spesifik bagi manajemen untuk merancang intervensi pemasaran yang dipersonalisasi guna mengoptimalkan retensi pelanggan dan efisiensi alokasi promosi.

3.2 Pembahasan

Penelitian ini berhasil mengatasi tantangan utama PT Bin Sef Travel dalam memanfaatkan data transaksi pelanggan historis secara optimal. Tujuan penelitian untuk menentukan segmen pelanggan tercapai melalui penerapan kombinasi model RFM (Recency, Frequency, Monetary) dan algoritma K-Means Clustering. Pendekatan ini menunjukkan bahwa basis pelanggan perusahaan, yang sebelumnya diperlakukan secara seragam melalui pemasaran massal, dapat disegmentasikan secara objektif ke dalam kelompok-kelompok tertentu berdasarkan pola pembelian mereka. Hal ini memberikan solusi konkret terhadap masalah alokasi sumber daya promosi, yang sebelumnya kurang tepat sasaran.

3.2.1 Interpretasi Hasil

Nilai Silhouette Score sebesar 0,4143 menunjukkan bahwa data transaksi pelanggan telah dikelompokkan dengan cukup baik menjadi empat segmen dengan makna perilaku sebagai berikut:

1) Champions : Segmen ini mewakili pelanggan tingkat elit atau mitra strategis. Walaupun recency berada di angka 30 hari, rekam jejak frekuensi (10 kali) dan nilai transaksi finansialnya sangat dominan dibandingkan pelanggan lain. Ini menunjukkan kontributor pendapatan yang masif bagi perusahaan.
2) Loyal Customers : Merupakan pilar utama profitabilitas PT Bin Sef Travel. Pelanggan ini sangat aktif berinteraksi dengan layanan (rata-rata recency 6,6 hari) dan memiliki frekuensi pemesanan yang tinggi (8,6 kali)
3) Potential Loyalist : Segmen populasi terbesar yang memiliki tingkat interaksi atau kebaruan paling tinggi (rata-rata recency 5,2 hari). Meskipun nilai transaksi dan frekuensi mereka masih berada di level menengah, antusiasme terbaru mereka menunjukkan potensi besar untuk dikonversi menjadi Loyal Customers jika dirawat dengan baik.
4) At Risk : Kelompok rentan dengan interval transaksi terlama (rata-rata recency 27,2 hari) dan frekuensi terendah (3,5 kali). Data ini mengindikasikan bahwa kelompok ini berada di ambang churn (berisiko beralih ke agen travel kompetitor) sehingga membutuhkan perhatian mendesak.

3.2.2 Perbandingan dengan Literatur

Temuan dalam penelitian ini sejalan dan memperkuat penelitian-penelitian sebelumnya. Serupa dengan temuan Nugraha dan Priti [5] pada data e-commerce, penelitian ini membuktikan bahwa metode RFM dan K-Means sangat efektif untuk menemukan perbedaan perilaku transaksi secara objektif. Namun, jika penelitian Nugraha menghasilkan tiga segmen umum (top, medium, low),

penelitian ini yang berfokus pada industri pariwisata religi berhasil mengidentifikasi granularitas perilaku yang lebih spesifik melalui empat segmen. Lebih lanjut, hasil ini memperkuat argumen pada penelitian Hermawan [4] yang menyatakan bahwa strategi "satu ukuran untuk semua" sudah tidak relevan, karena setiap segmen pelanggan memerlukan pendekatan treatment yang unik untuk mempertahankan loyalitasnya.

3.2.3 Implikasi Penelitian

Secara teoritis, penelitian ini berkontribusi pada pengembangan pengetahuan di bidang analitik pemasaran, khususnya penerapan kombinasi model RFM dan pengelompokan K-Means dalam konteks industri perjalanan umrah, yang memiliki karakteristik unik dibandingkan dengan e-commerce konvensional. Secara praktis, hasil segmentasi ini memberikan landasan kuat bagi manajemen PT Bin Sef Travel untuk merancang strategi pemasaran yang berbeda-beda untuk setiap segmen: (1) mempertahankan pelanggan unggulan melalui layanan VIP dan komunikasi eksklusif yang dipersonalisasi, (2) memperkuat loyalitas pelanggan setia melalui program penghargaan dan penawaran paket prioritas, (3) mendorong Pelanggan Potensial untuk meningkatkan frekuensi transaksi mereka melalui insentif dan promosi yang ditargetkan, dan (4) mengaktifkan kembali pelanggan Berisiko dengan kampanye reaktivasi berbasis data seperti diskon khusus atau pengingat yang dipersonalisasi. Pendekatan berbasis data ini memungkinkan alokasi anggaran promosi yang lebih efisien dan dapat meningkatkan profitabilitas perusahaan secara keseluruhan.

Penelitian ini memiliki dataset private perusahaan travel binsef yang digunakan hanya mencakup 68 pelanggan unik dari 448 transaksi dalam rentang waktu yang relatif singkat (September–Desember 2025), sehingga generalisasi hasil segmentasi perlu dilakukan dengan hati-hati. Kedua, penelitian ini hanya menggunakan tiga variabel RFM tanpa mempertimbangkan faktor demografis, preferensi layanan, atau saluran komunikasi pelanggan yang dapat memperkaya dimensi segmentasi. Ketiga, nilai Silhouette Score sebesar 0,4143 menunjukkan kualitas klaster yang cukup baik ,algoritma K-Means bersifat sensitif terhadap nilai inisialisasi centroid awal dan dapat menghasilkan hasil yang berbeda pada setiap eksekusi tanpa pengaturan random seed yang konsisten. Keterbatasan-keterbatasan ini dapat menjadi arah pengembangan penelitian selanjutnya dengan memperluas dataset, menambahkan variabel segmentasi, serta mengeksplorasi algoritma alternatif lainnya.

4. KESIMPULAN

Berdasarkan analisis empiris terhadap 448 catatan transaksi dari PT Bin Sef Travel dengan pendekatan hybrid model RFM (Recency, Frequency, Monetary) dan algoritma K-Means Clustering, penelitian ini berhasil mengidentifikasi empat segmen pelanggan utama: Champions (1 pelanggan), Loyal customer (29 pelanggan), Potential loyalist (32 pelanggan), dan At risk(6 pelanggan). Pendekatan ini efektif dalam memetakan keberagaman karakteristik pelanggan, sehingga memungkinkan perusahaan untuk beralih dari pemasaran massal ke strategi pemasaran yang dipersonalisasi dan didorong oleh data untuk mengoptimalkan alokasi sumber daya promosi dan retensi pelanggan perusahaan.

Daftar Pustaka

[1] G. S. Isokova, "The Impact of Digital Marketing on the Transformation of the Business Travel Industry," *American Journal of Social and Humanitarian Research*, vol. 7, no. 3, pp. 153–163, 2026.

[2] M. Rahmad and A. Aziz, "Kecerdasan Buatan untuk Personalisasi Layanan: Studi Kasus Chatbot di Industri Travel," *DESTINASI: Journal of Hospitality and Tourism Research*, vol. 1, no. 1, pp. 138–147, 2026.

[3] F. F. Muhimmah, "Karyawan dan Produk (Sebuah Tinjauan MSDM Tour and Travel Haji dan Umroh di Sedati Sidoarjo Perspektif Ekonomi Syariah)," *IQTISADIE: Journal of Islamic Banking and Shariah Economy*, vol. 5, no. 1, pp. 67–80, 2025.

[4] A. Hermawan, F. Aditama, L. R. Ramadhani, N. I. Harahap, N. M. Ilham, A. Saputra, and N. R. Jayanti, "Analisis Segmentasi Pelanggan Berbasis RFM dan Evaluasi Efektivitas Kampanye Pemasaran untuk Meningkatkan Retensi," *Neptunus: Jurnal Ilmu Komputer Dan Teknologi*

Informasi, vol. 2, no. 4, pp. 28–42, 2024.
[5] A. S. Nugraha and Priti, "Analisis Segmentasi Customer dengan RFM dan K-Means Clustering: Studi E-Commerce PT PQR," *JMSI*, vol. 6, no. 2, pp. 187–196, 2025.
[6] A. Kanivia, N. A. Putrie, S. Anugrah, and I. R. Bawono, "Literature Review: Analisis Profitabilitas Pelanggan," *JRAK (Jurnal Riset Akuntansi & Komputerisasi Akuntansi)*, vol. 16, no. 1, pp. 48–65, 2025.
[7] A. Andirwan, V. Asmilita, M. Zhafran, A. Syaiful, and M. Beddus, "Strategi Pemasaran Digital: Inovasi untuk Maksimalkan Penjualan Produk Konsumen di Era Digital," *Jurnal Ilmiah Multidisiplin Amsir*, pp. 155–166, 2023.
[8] E. Purwanto, F. Mohd, Z. A. Long, and S. Purnomo, "Sales Transaction Data Evaluation Information System Using Clustering Model," *4th International Conference of Health, Science and Technology (ICOHETECH)*, pp. 20–24, 2023.
[9] P. A. Wicaksana, I. B. A. Swamardika, and R. S. Hartati, "Literature Review Analisis Perilaku Pelanggan Menggunakan RFM Model," *Majalah Ilmiah Teknologi Elektro*, vol. 21, no. 1, pp. 21–30, 2022.
[10] K. K. Ningrum, J. Maulindar, and A. Farida, "Penerapan Algoritma K-Means untuk Pengelompokkan Penilaian Akhir Semester di SDN Kadokan 01 Sukoharjo," *INFOTECH Journal*, vol. 9, no. 1, pp. 190–197, 2023.
[11] A. S. Kedaton, B. A. Bukhari, NurAfrianto, R. W. Febriyanto, and A. A. Sari, "Pemetaan Wilayah Berdasarkan Minat Baca Masyarakat dengan Algoritma K-Means," *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis (SENATIB)*, pp. 322–327, 2025.
[12] M. F. Fadhillah, A. L. A. Suyoso, and I. Puspitasari, "Customer Segmentation with Clustering Algorithm Based on Recency, Frequency, and Monetary (RFM) Attributes," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 1, pp. 48–56, 2025.
[13] A. Rianti, N. W. A. Majid, and A. Fauzi, "CRISP-DM: Metodologi Proyek Data Science," *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis (SENATIB)*, pp. 107–114, 2023.
[14] D. Ruswanti, D. Susilo, and Riani, "Implementasi CRISP-DM pada Data Mining untuk Melakukan Prediksi Pendapatan dengan Algoritma C.45," *GO INFOTECH: Jurnal Ilmiah STMIK AUB*, vol. 30, no. 1, pp. 111–121, 2024.
[15] D. B. Saputra, V. Atina, and F. E. Nastiti, "Penerapan Model CRISP-DM pada Prediksi Nasabah Kredit Menggunakan Algoritma Random Forest," *IDEALIS: Indonesia Journal Information System*, vol. 7, no. 2, pp. 240–247, 2024.
[16] S. N. Wardani and Nurmalitasari, "Penerapan Regresi Linier Berganda Untuk Memprediksi Harga Laptop Dengan Menggunakan Software Python," *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis (SENATIB)*, pp. 377–382, 2023.

***ZONAsi**: Jurnal Sistem Informasi*
Is licensed under a Creative Commons Attribution International (CC BY-SA 4.0)