

SEGMENTASI PELANGGAN PADA MITRA REDDOORZ JM TUPAREV KARAWANG MENGGUNAKAN K-MEANS CLUSTERING

Fisca Shafa Salsabila¹, Hannie², Billy Ibrahim Hasbi³

Sistem Informasi, Fakultas Ilmu Komputer

Universitas Singaperbangsa Karawang

Jl. HS. Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat

salsabilashafafisca@gmail.com¹, hannie@staff.unsika.ac.id², billy.ibrahim@staff.unsika.ac.id³

Abstrak

Penelitian ini bertujuan untuk melakukan segmentasi pelanggan pada Mitra *RedDoorz JM* Tuparev Karawang menggunakan algoritma *K-Means Clustering* dengan pendekatan *CRISP-DM*. Data yang digunakan berupa data transaksi pelanggan yang dianalisis berdasarkan tipe kamar, lama menginap, total harga, dan deposit. Penentuan jumlah cluster optimal dilakukan dengan metode *Elbow* dan *Silhouette Coefficient* serta *Davies-Bouldin Index* dengan nilai *Silhouette* sebesar 0,57 dan *Davies-Bouldin Index* sebesar 0,66 yang menunjukkan kualitas pemisahan cluster cukup baik. Hasil penelitian menghasilkan empat segmen pelanggan, yaitu *Lost Customers* (1.395 pelanggan), *Occasional Customers* (1.784 pelanggan), *Potential Customers* (377 pelanggan), dan *Loyal Customers* (18 pelanggan). Setiap *cluster* memiliki karakteristik transaksi yang berbeda, sehingga menunjukkan bahwa pelanggan bersifat heterogen. Hasil segmentasi divisualisasikan menggunakan Tableau untuk memperjelas distribusi pelanggan dan mendukung penyusunan strategi bisnis yang lebih tepat sasaran, seperti reaktivasi, peningkatan keterlibatan, personalisasi layanan, dan retensi pelanggan berbasis data.

Kata Kunci: Segmentasi Pelanggan, K-Means Clustering, CRISP-DM, Data Mining, RedDoorz.

Abstract

This study aims to perform customer segmentation at RedDoorz Partner JM Tuparev Karawang using the K-Means Clustering algorithm with the CRISP-DM approach. The data used consist of customer transaction records analyzed based on room type, length of stay, total transaction value, and deposit amount. The optimal number of clusters was determined using the Elbow method, Silhouette Coefficient, and Davies-Bouldin Index, resulting in a Silhouette score of 0.57 and a Davies-Bouldin Index of 0.66, indicating a fairly good cluster separation quality. The results produced four customer segments: Lost Customers (1,395 customers), Occasional Customers (1,784 customers), Potential Customers (377 customers), and Loyal Customers (18 customers). Each cluster exhibits distinct transaction characteristics, demonstrating that customers are heterogeneous and cannot be treated as a homogeneous group. The segmentation results were visualized using Tableau to clarify the distribution of customers across clusters and to support the development of more targeted business strategies, such as customer reactivation, engagement enhancement, service personalization, and data-driven retention strategies.

Keywords: Customer Segmentation, K-Means Clustering, CRISP-DM, Data Mining, RedDoorz)

PENDAHULUAN

Industri perhotelan berperan penting dalam mendorong pertumbuhan ekonomi dan sektor pariwisata. Sebagai sebuah perusahaan atau entitas industri, hotel memiliki berbagai bagian dengan fungsi dan tanggung jawab masing-masing [1]. Segmentasi pelanggan memiliki peran krusial dalam keberhasilan strategi pemasaran dan peningkatan pendapatan perusahaan. Dengan membagi konsumen berdasarkan preferensi dan perilaku mereka, perusahaan dapat merancang strategi harga dan promosi yang lebih terarah serta meningkatkan kepuasan pelanggan [2].

Mitra *RedDoorz JM* Tuparev di Karawang merupakan salah satu properti yang sedang mengalami pengembangan bisnis. Berdasarkan hasil wawancara dengan pemilik, properti ini merupakan cabang ketiga, yang secara khusus ingin dijadikan sebagai penyempurnaan dari dua cabang sebelumnya. Upaya penyempurnaan ini tidak hanya berfokus pada peningkatan fasilitas, tetapi juga pada optimalisasi strategi operasional dan pemasaran. Namun, hingga saat ini, properti tersebut belum memiliki analisis segmentasi pelanggan yang terstruktur untuk memahami karakteristik dan

perilaku tamu yang datang, baik dari sisi tujuan perjalanan, frekuensi pemesanan, maupun nilai transaksi.

Selain itu, karakteristik pasar Karawang yang sangat heterogen turut memperkuat urgensi segmentasi pelanggan. Karawang merupakan kawasan industri besar yang menarik banyak tamu pekerja proyek, teknisi, sales, serta mobilitas perusahaan. Berdasarkan Statistik Daerah Kabupaten Karawang 2022, jumlah penduduk Karawang mencapai 2.505.247 jiwa menunjukkan tingginya aktivitas dan mobilitas penduduk di wilayah ini [3].

Adapun algoritma *K-Means* digunakan karena mampu mengelompokkan data numerik secara efektif dengan cara meminimalkan jarak antar data dalam satu kelompok dan memaksimalkan perbedaan antar kelompok. Selain itu, prosesnya relatif sederhana, efisien, dan hasilnya mudah diinterpretasikan, sehingga sesuai dengan kebutuhan penelitian yang tidak hanya menekankan aspek teknis, tetapi juga kemudahan dalam menganalisis karakteristik setiap segmen untuk mendukung strategi bisnis. Penelitian ini menggunakan metodologi *CRISP-DM* karena memberikan tahapan terstruktur mulai dari pemahaman bisnis hingga *deployment*, sehingga proses segmentasi dapat dilakukan secara sistematis dan relevan dengan kebutuhan operasional [4]

METODE PENELITIAN

Penelitian ini menggunakan pendekatan data mining dengan standar proses *CRISP-DM* (*Cross Industry Standard Process for Data Mining*). Pendekatan ini terdiri dari enam tahapan utama yang membentuk alur sistematis dalam pembangunan model analisis data. Terdapat 6 tahapan utama yaitu:

1. *Business Understanding*

Tahap awal ini bertujuan untuk memahami secara menyeluruh kebutuhan dan tujuan bisnis yang menjadi dasar penelitian. Dalam konteks penelitian ini, fokus utama adalah bagaimana melakukan segmentasi pelanggan *RedDoorz* secara efektif untuk meningkatkan strategi pemasaran dan layanan.

2. *Data Understanding*

Tahapan ini berfokus pada eksplorasi data pelanggan yang tersedia. Setelah data dikumpulkan dari sistem internal perusahaan, dilakukan analisis awal terhadap struktur dan kualitas data.

3. *Data Preparation*

Tahap ini merupakan proses transformasi data mentah menjadi format yang siap digunakan dalam proses pemodelan. Tahapan ini penting karena kualitas data sangat memengaruhi hasil analisis.

4. *Modeling*

Setelah data siap, tahap berikutnya adalah membangun model clustering dengan algoritma *K-Means*.

5. *Evaluation*

Setelah model dianggap cukup baik, dilakukan evaluasi secara menyeluruh terhadap hasil segmentasi untuk memastikan model memenuhi tujuan bisnis. Tahapan evaluasi model akhir ini dilakukan dengan mengukur kualitas hasil clustering menggunakan *Silhouette Coefficient* dan *Davies-Bouldin Index* untuk menilai tingkat kesesuaian dan pemisahan antar cluster yang terbentuk [5].

6. *Deployment*

Tahap akhir dari proses penelitian ini adalah menyajikan visualisasi hasil segmentasi dan rekomendasi strategi bisnis berdasarkan karakteristik tiap *cluster* yang terbentuk [6].

HASIL DAN PEMBAHASAN

1. *Business Understanding*

Mitra *RedDoorz* JM Tuparev Karawang merupakan salah satu unit usaha di bidang perhotelan yang melayani berbagai jenis pelanggan dengan karakteristik dan perilaku transaksi yang berbeda-beda. Perbedaan tersebut tercermin dari lama menginap, jenis kamar yang dipilih, serta nilai transaksi yang dihasilkan oleh setiap pelanggan. Namun, hingga saat ini pihak manajemen belum memiliki pemetaan pelanggan yang jelas berdasarkan karakteristik tersebut, sehingga strategi pemasaran dan pelayanan masih bersifat umum. Oleh karena itu, tujuan bisnis dari

penelitian ini adalah untuk membantu pihak manajemen dalam mengelompokkan pelanggan berdasarkan kesamaan karakteristik transaksi dan perilaku menginap. Hasil pengelompokan pelanggan diharapkan dapat digunakan sebagai dasar dalam pengambilan keputusan strategis, seperti penyusunan program promosi, penentuan segmentasi pasar, serta peningkatan kualitas layanan yang lebih tepat sasaran.

2. Data Understanding

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Mitra *RedDoorz JM Tuparev Karawang*. Dataset yang digunakan bernama DATASET, yang berisi data transaksi pelanggan yang melakukan pesanan kamar selama periode bulan Januari sampai Desember tahun 2025 sebanyak 3.731 baris data. Data tersebut mencakup informasi terkait tipe kamar, waktu cek in dan cek out, lama menginap, harga kamar, biaya service, pajak, subtotal dan deposit. Adapun atribut awal yang terdapat pada dataset terdiri dari No., Kamar Tipe, *Cek In*, *Cek Out*, Hari, Nama, Harga, *Service*, Pajak, Subtotal, dan Dep dengan jumlah 11 atribut. Berdasarkan proses seleksi atribut, maka atribut yang digunakan dalam penelitian ini adalah Kamar Tipe, Hari, Harga, dan Dep dengan jumlah 4 atribut yang akan digunakan dalam pengolahan data atau proses analisis segmentasi pelanggan pada tahap modeling. Setiap atribut dievaluasi berdasarkan tingkat keterkaitannya dengan tujuan penelitian, sehingga hanya atribut yang memiliki pengaruh signifikan terhadap proses *clustering* yang dipertahankan. Alasan pemilihan masing-masing atribut dalam penelitian ini disajikan pada Tabel 1.

Tabel 1. Alasan Pemilihan Dataset

No	Atribut	Alasan Pemilihan
1	Kamar Tipe	Merepresentasikan preferensi kelas layanan pelanggan.
2	Hari	Menggambarkan durasi menginap yang mencerminkan pola konsumsi.
3	Harga	Faktor utama dalam segmentasi berdasarkan nilai transaksi
4	Deposit	Menggambarkan komitmen finansial dan pola pembayaran pelanggan.

3. Data Preparation

a. Penyesuaian Tipe Data

Pada tahap penyesuaian tipe data, dilakukan perubahan tipe data pada atribut Hari, Harga, dan Deposit yang sebelumnya bertipe *object* menjadi tipe data numerik dengan tujuan agar data dapat diolah secara matematis pada tahap analisis dan pemodelan. Hasil konversi menunjukkan bahwa ketiga atribut tersebut memiliki tipe data *float*, yang sesuai untuk proses perhitungan statistik dan algoritma clustering.

b. Penanganan Missing Values

Pada tahap penanganan *missing values*, seluruh baris data yang mengandung nilai kosong dihapus dari dataset. Keputusan ini diambil karena baris data yang memiliki *missing values* merupakan data yang tidak lengkap dan tidak merepresentasikan transaksi yang valid, sehingga tidak tepat untuk dilakukan pengisian nilai menggunakan metode imputasi seperti nilai rata-rata atau teknik penanganan lainnya.

c. Penanganan Data Duplikat

Pada tahap penanganan data duplikat, tidak dilakukan proses penghapusan secara terpisah. Hal ini dikarenakan baris data yang terdeteksi sebagai duplikat merupakan baris data yang sama dengan baris data yang mengandung *missing values*.

d. Penanganan *Outlier* (Nilai Ekstrem)

Pada penanganan *outlier*, tidak dilakukan penghapusan terhadap data yang terdeteksi sebagai *outlier*. Keputusan ini diambil karena nilai-nilai ekstrem yang muncul pada data masih merepresentasikan transaksi yang valid dan mencerminkan variasi perilaku pelanggan. Penghapusan *outlier* berpotensi menghilangkan informasi penting yang dibutuhkan dalam proses segmentasi pelanggan.

e. Transformasi Data (Encoding)

Pada tahap transformasi data, dilakukan proses *encoding* pada atribut Kamar Tipe yang

sebelumnya bertipe kategorikal menjadi bentuk numerik. Hasil dari proses *encoding* pada atribut kamar tipe dengan nilai numerik yang dihasilkan dari proses *encoding*, seperti 0, 1, 2, 3, dan 4, tidak merepresentasikan tingkatan kualitas atau urutan nilai dari rendah ke tinggi, melainkan hanya berfungsi sebagai kode numerik untuk merepresentasikan masing-masing kategori kamar, seperti *Deluxe*, *Suite*, dan kategori lainnya.

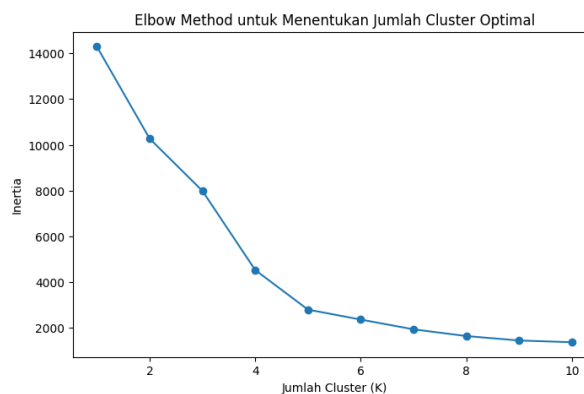
f. Normalisasi Data

Tahapan normalisasi data dilakukan dengan menggunakan metode *StandardScaler*. Normalisasi dilakukan dengan mentransformasikan setiap atribut numerik agar memiliki nilai rata-rata (*mean*) mendekati nol dan standar deviasi mendekati satu. Tahap ini bertujuan untuk menghilangkan perbedaan skala antar atribut, seperti Kamar Tipe, Hari, Harga, dan Deposit, sehingga tidak ada *variabel* yang memiliki pengaruh lebih besar hanya karena perbedaan satuan atau rentang nilai.

4. Modeling

a. Penentuan Jumlah *Cluster* (*Elbow Method*)

Penentuan jumlah cluster merupakan tahap penting dalam penerapan metode *K-Means Clustering* karena jumlah *cluster* yang tidak tepat dapat memengaruhi kualitas hasil pengelompokan. Pada penelitian ini, penentuan jumlah *cluster* optimal dilakukan menggunakan metode *Elbow Method*. Metode ini bekerja dengan menghitung nilai inerti atau *within-cluster sum of squares (WCSS)* pada berbagai jumlah *cluster*.



Gambar 2. Grafik *Elbow Method*

Berdasarkan grafik *Elbow Method* yang ditampilkan pada Gambar 2 terlihat bahwa penurunan nilai inerti terjadi secara signifikan hingga jumlah *cluster* tertentu, kemudian mulai melandai pada penambahan *cluster* berikutnya. Pola tersebut membentuk titik siku (*elbow*) yang menunjukkan jumlah *cluster* optimal. Dari hasil pengamatan grafik, titik siku terlihat pada jumlah *cluster* sebanyak empat, sehingga pada penelitian ini ditetapkan nilai $K = 4$ sebagai jumlah *cluster* yang digunakan dalam proses *K-Means Clustering*.

b. Proses *Clustering*

Kamar	Tipe	Hari	Harga	Dep	Cluster
0	1	250.0	250	0	
0	1	250.0	250	0	
0	1	200.0	200	1	
0	1	235.0	235	0	
0	1	200.0	200	1	
0	1	250.0	250	0	
0	1	225.0	225	0	
0	1	250.0	250	0	
0	1	250.0	250	0	
0	1	200.0	200	1	

Gambar 3. Hasil *Clustering*

K-Means Clustering pada penelitian ini diterapkan dengan parameter $n_clusters = 4$, sehingga menghasilkan empat *cluster* pelanggan yang memiliki karakteristik berbeda. Keempat *cluster* tersebut diberi label *Cluster 0*, *Cluster 1*, *Cluster 2*, dan *Cluster 3* seperti yang ditampilkan pada Gambar 3. Setelah proses *K-Means Clustering* dilakukan dan diperoleh empat *cluster* pelanggan, tahap selanjutnya adalah melakukan interpretasi dan penentuan kategori *cluster*. Penentuan kategori dilakukan dengan menganalisis nilai rata-rata (*mean*) setiap atribut pada masing-masing *cluster*, yaitu Kamar Tipe (jenis kamar), Hari (durasi menginap), Harga, dan Dep (Deposit). Nilai rata-rata tersebut digunakan untuk menggambarkan tingkat pengeluaran dan perilaku transaksi pelanggan pada setiap *cluster*. Berdasarkan perbandingan karakteristik antar *cluster*, masing-masing *cluster* kemudian diberi label kategori yang mencerminkan tingkat nilai pelanggan seperti pada Tabel 2.

Tabel 2. Karakteristik *Cluster*

Cluster	Kategori	Kamar Tipe	Hari	Harga	Deposit
0	Lost Customer	Standar	Sangat Rendah	Sedang	Sedang
1	Occasional Customer	Standar	Rendah	Rendah	Rendah
2	Potential Customer	Standar-Menengah	Tinggi	Sedang	Sedang
3	Loyal Customer	Menengah-Premium	Menengah-Tinggi	Tinggi	Tinggi

Berdasarkan tabel 2, *Cluster 0* dikategorikan sebagai *Lost Customer* karena memiliki karakteristik durasi menginap yang sangat rendah serta kecenderungan memilih kamar dengan tipe standar. Nilai harga dan deposit berada pada tingkat sedang, namun frekuensi dan lama transaksi yang rendah menunjukkan bahwa pelanggan pada klaster ini sudah jarang melakukan pemesanan.

Cluster 1 termasuk dalam kategori *Occasional Customer* yang ditandai dengan durasi menginap rendah dan nilai harga serta deposit yang paling rendah dibandingkan klaster lainnya. Pelanggan pada klaster ini umumnya memilih kamar tipe standar dan melakukan transaksi secara tidak rutin.

Cluster 2 diklasifikasikan sebagai *Potential Customer* karena memiliki durasi menginap yang relatif tinggi dengan nilai harga dan deposit berada pada tingkat sedang. Pelanggan pada klaster ini menunjukkan intensitas penggunaan layanan yang lebih baik dibandingkan klaster sebelumnya dan mulai menunjukkan pola transaksi yang lebih stabil.

Cluster 3 merupakan kategori *Loyal Customer* yang memiliki karakteristik durasi menginap menengah hingga tinggi, kecenderungan memilih kamar dengan tipe menengah hingga premium, serta nilai harga dan deposit tertinggi. Pelanggan pada klaster ini menunjukkan frekuensi dan nilai transaksi yang konsisten, sehingga memberikan kontribusi pendapatan yang signifikan.

5. Evaluation

```
silhouette_avg = silhouette_score(df_scaled, cluster_labels)
print("Nilai Silhouette Coefficient:", silhouette_avg)

Nilai Silhouette Coefficient: 0.5741337089101323
```

Gambar 4. Evaluasi *Silhouette Coefficient*

Berdasarkan hasil evaluasi menggunakan *Silhouette Coefficient*, diperoleh nilai rata-rata sebesar 0,574 seperti pada Gambar 4. Nilai ini menunjukkan bahwa model *K-Means Clustering* yang dibangun memiliki kualitas pengelompokan yang baik, karena berada di atas ambang batas 0,5 yang secara umum mengindikasikan bahwa objek data telah terkelompok dengan cukup jelas dan memiliki tingkat kemiripan yang tinggi di dalam *cluster* serta perbedaan yang memadai antar *cluster*.

```
dbi = davies_bouldin_score(df_scaled, cluster_labels)
print("Nilai Davies-Bouldin Index:", dbi)

Nilai Davies-Bouldin Index: 0.6637871432300958
```

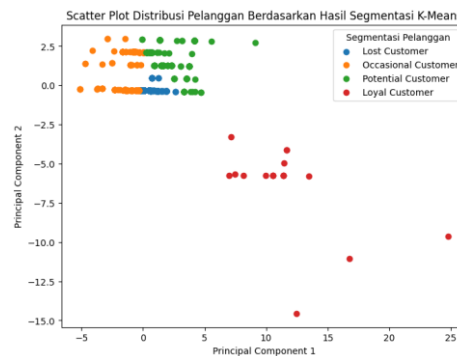
Gambar 5. Evaluasi *Davies-Bouldin Index*

Berdasarkan hasil perhitungan menggunakan *Davies-Bouldin Index (DBI)* pada Gambar 5, diperoleh nilai sebesar 0,6638. Nilai ini menunjukkan bahwa struktur cluster yang terbentuk memiliki tingkat pemisahan antar *cluster* yang cukup baik serta tingkat kekompakan dalam masing-masing cluster yang relatif terjaga. Dengan demikian, nilai DBI sebesar 0,6638 mengindikasikan bahwa hasil clustering yang diperoleh masih berada dalam kategori baik dan mampu merepresentasikan perbedaan karakteristik pelanggan secara cukup jelas.

6. Deployment

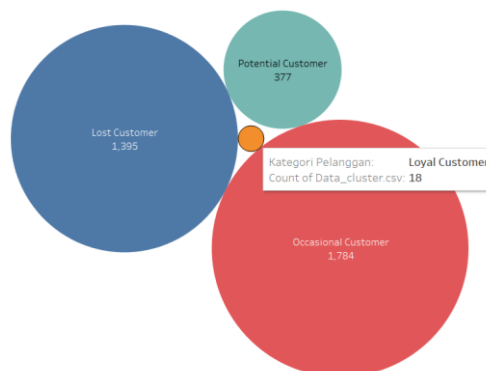
a. Visualisasi Hasil Segmentasi

Visualisasi hasil segmentasi pelanggan ditunjukkan melalui dua jenis grafik, yaitu *scatter plot* berbasis *Principal Component Analysis (PCA)* dan diagram lingkaran (*pie chart*). *Scatter plot* digunakan untuk menggambarkan distribusi dan pemisahan antar segmen pelanggan berdasarkan hasil *K-Means Clustering* dalam ruang dua dimensi, sedangkan *pie chart* digunakan untuk menunjukkan proporsi jumlah pelanggan pada setiap *cluster* yang terbentuk.



Gambar 6. *Scatter Plot* Hasil Segmentasi

Berdasarkan *scatter plot* pada Gambar 6, hasil reduksi dimensi menggunakan PCA, terlihat bahwa masing-masing segmen pelanggan membentuk pola distribusi yang relatif terpisah satu sama lain. *Cluster Lost Customer*, *Occasional Customer*, dan *Potential Customer* cenderung terkonsentrasi pada area tertentu dengan jarak antar titik yang cukup berdekatan, yang menunjukkan adanya kemiripan karakteristik perilaku pelanggan dalam masing-masing *cluster* tersebut. Sementara itu, *Cluster Loyal Customer* tampak terpisah cukup jauh dari *cluster* lainnya, dengan sebaran titik yang lebih jarang.



Gambar 7. *Packed Bubbles* Hasil Segmentasi

Selanjutnya, *packed bubbles* pada Gambar 7 menunjukkan distribusi jumlah pelanggan pada setiap *cluster* hasil *K-Means Clustering*. Berdasarkan visualisasi tersebut, *Occasional*

Customer Cluster merupakan *cluster* dengan proporsi terbesar, yaitu sebesar 49,9%, diikuti oleh *Lost Customer Cluster* sebesar 39,0%. *Potential Customer Cluster* memiliki proporsi sebesar 10,5%, sedangkan *Loyal Customer Cluster* hanya sebesar 0,5% dari total pelanggan. Hasil ini menunjukkan bahwa sebagian besar pelanggan masih berada pada kategori dengan tingkat loyalitas yang belum stabil, sehingga memerlukan strategi bisnis yang berfokus pada peningkatan keterlibatan dan loyalitas pelanggan. Secara keseluruhan, kedua visualisasi tersebut saling melengkapi dalam menjelaskan hasil segmentasi pelanggan.

b. Rekomendasi Strategi Bisnis

Rekomendasi strategi bisnis dalam penelitian ini disusun berdasarkan hasil segmentasi pelanggan menggunakan metode *K-Means Clustering*. Strategi yang diusulkan tidak didasarkan pada intuisi atau asumsi subjektif, melainkan diturunkan secara langsung dari karakteristik masing-masing segmen pelanggan yang terbentuk melalui proses analisis data.

1. *Lost Customer Cluster*

Lost Customer Cluster ditandai dengan frekuensi transaksi rendah dan nilai transaksi relatif kecil. Karakteristik ini menunjukkan rendahnya keterlibatan pelanggan serta potensi terjadinya *customer churn*. Dalam perspektif *CRM (Customer Relationship Management)* strategis, pelanggan yang mengalami penurunan aktivitas tidak serta-merta diabaikan, melainkan perlu dianalisis berdasarkan nilai potensialnya di masa depan [7]. Oleh karena itu, pendekatan yang relevan untuk *cluster* ini adalah strategi reaktivasi pelanggan, yang meliputi:

- a. Pengiriman promo personal melalui *email* atau *WhatsApp* berdasarkan riwayat transaksi terakhir.
- b. Penawaran diskon
- c. Program *win-back campaign* untuk pelanggan yang lama tidak aktif
Strategi ini sejalan dengan *CRM*, di mana perusahaan harus mengidentifikasi pelanggan yang berisiko *churn*, membedakan nilainya, melakukan interaksi ulang, serta menyesuaikan penawaran [8].

2. *Loyal Customer Cluster*

Loyal Customer Cluster memiliki frekuensi transaksi tinggi dan nilai transaksi terbesar dibandingkan *cluster* lain. Secara teoritis, pelanggan dengan *purchase habits* tinggi dapat dikategorikan sebagai *heavy users* yang berkontribusi besar terhadap pendapatan perusahaan. Dalam konsep *CRM*, pelanggan bernilai tinggi perlu dipertahankan karena biaya mempertahankan pelanggan lebih rendah dibandingkan biaya memperoleh pelanggan baru [9]. Strategi yang direkomendasikan untuk *cluster* ini adalah strategi retensi berbasis *loyalty program*, antara lain:

- a. Program *membership* bertingkat (*multi-tier loyalty program*)
- b. Reward eksklusif untuk pelanggan dengan frekuensi tertentu
- c. Prioritas layanan dan benefit khusus
Program loyalitas yang dirancang dengan baik dapat meningkatkan *customer lifetime value (CLV)* dan memperkuat hubungan jangka panjang.

3. *Occasional Customer Cluster*

Occasional Customer Cluster memiliki jumlah pelanggan terbesar dengan frekuensi transaksi tidak stabil dan nilai transaksi menengah. Secara strategis, segmen ini memiliki potensi besar untuk ditingkatkan menjadi pelanggan loyal. Dalam pendekatan *CRM* analitik, pelanggan dengan aktivitas menengah perlu dikelola melalui peningkatan intensitas interaksi dan pengalaman pelanggan [7]

Strategi yang direkomendasikan untuk *cluster* ini meliputi:

- a. Program akumulasi poin sederhana
- b. Promo musiman berbasis perilaku pembelian sebelumnya
- c. Paket bundling untuk meningkatkan frekuensi transaksi
Pendekatan ini bertujuan meningkatkan engagement pelanggan sehingga frekuensi transaksi menjadi lebih konsisten. Secara teoritis, peningkatan frekuensi pembelian merupakan indikator keberhasilan strategi retensi dan penguatan loyalitas.

4. *Potential Customer Cluster*

Potential Customer Cluster memiliki nilai transaksi relatif tinggi namun frekuensi pembelian

masih berada pada tingkat menengah. Karakteristik ini menunjukkan adanya potensi peningkatan loyalitas apabila dikelola dengan pendekatan yang tepat. CRM juga menekankan pentingnya *customization* dalam membangun hubungan jangka panjang [8]

Strategi yang direkomendasikan antara lain:

- a. Rekomendasi tipe kamar berdasarkan riwayat pemesanan
- b. Upselling dan cross-selling terpersonalisasi
- c. Penawaran upgrade layanan pada momen tertentu

Pendekatan ini sejalan dengan konsep *customer-centric marketing* yang menempatkan personalisasi sebagai faktor utama dalam meningkatkan loyalitas dan retensi pelanggan [10].

SIMPULAN

Berdasarkan hasil segmentasi pelanggan pada Mitra *RedDoorz JM Tuparev Karawang* menggunakan algoritma *K-Means Clustering* dengan pendekatan *CRISP-DM*, penelitian ini menghasilkan empat *cluster* utama, yaitu *Lost Customers* sebanyak 1.395 pelanggan, *Occasional Customers* sebanyak 1.784 pelanggan, *Potential Customers* sebanyak 377 pelanggan, dan *Loyal Customers* sebanyak 18 pelanggan, dengan nilai *Silhouette Coefficient* sebesar 0,57 dan nilai *Davies-Bouldin Index* 0,66 yang menunjukkan bahwa kualitas pemisahan *cluster* tergolong cukup baik dan mampu merepresentasikan perbedaan karakteristik perilaku pelanggan berdasarkan tipe kamar, lama menginap, total transaksi, dan deposit. Visualisasi hasil pengelompokan melalui grafik persebaran *cluster* terbukti efektif dalam memperjelas distribusi pelanggan serta batas antar *cluster* sehingga pola dan ciri khas masing-masing *cluster* dapat dipahami secara lebih intuitif. Hasil segmentasi ini memberikan kontribusi praktis bagi perusahaan dalam menyusun strategi bisnis yang lebih terarah dan berbasis data, seperti penerapan strategi reaktivasi pada kelompok *Lost Customers*, penguatan retensi bagi *Loyal Customers*, peningkatan frekuensi transaksi pada *Occasional Customers*, serta personalisasi layanan pada *Potential Customers* untuk meningkatkan nilai pelanggan, sehingga mendukung pengambilan keputusan manajerial yang lebih efektif dan efisien.

DAFTAR PUSTAKA

- [1] H. H. Arrosyid, Z. Pratama, and G. Priambodo, "Penurunan Cancellation Rate Pada City Hotel Menggunakan Metode Issue Tree," *Jukompak (Jurnal Komputasi dan Pengemb. Apl.*, vol. 1, no. 1, pp. 31–39, 2025.
- [2] S. Merdikawati, R. D. Oktaviani, and N. Rahmah, "Analisis Dampak Segmentasi Konsumen terhadap Strategi Pemasaran dalam Tarif Dua Bagian Menggunakan Algoritma Metaheuristics The Impact of Consumer Segmentation on Marketing Strategies in Two-Part Tariff Pricing using Metaheuristics Approach," *J. IPTEK*, vol. 9, no. 1, pp. 12–18, 2025.
- [3] P. Yuyu Ananda, *Statistik daerah*. Karawang: Badan Pusat Statistik Kabupaten/Kota, 2022.
- [4] Y. Suhanda, I. Kurniati, and S. Norma, "Penerapan Metode Crisp-DM Dengan Algoritma K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik," *J. Teknol. Inform. dan Komput. MH Thamrin*, vol. 6, no. 2, pp. 12–20, 2020.
- [5] K. Z. Wijaya, A. Djunaidy, and F. Mahananto, "Segmentasi Pelanggan Menggunakan Algoritma K-Means dan Analisis RFM di Ova Gaming E-Sports Arena Kediri," *J. Tek. ITS*, vol. 10, no. 2, 2021.
- [6] N. H. Harani, C. Prianto, and F. A. Nugraha, "Segmentasi Pelanggan Produk Digital Service Indihome Menggunakan Algoritma K-Means Berbasis Python," *J. Manaj. Inform.*, vol. 10, 2020, doi: 10.34010/jamika.v10i2.
- [7] M. Dr. Agus Wibowo, M.Kom, M.Si, *Teori & Praktek CRM (Customer Relationship Management)*. 2023.
- [8] C. V. Andaryani and H. Alifahmi, "Penerapan Customer Relationship Management Dalam Sosial Media Untuk Mempertahankan Loyalitas Pelanggan," vol. 6, no. 2, pp. 404–417, 2023.
- [9] K. K. N. Kartika Imasari, "PENGARUH CUSTOMER RELATIONSHIP MANAGEMENT TERHADAP LOYALITAS PELANGGAN PADA PT BCA Tbk," vol. 10, no. 3, pp. 183–192, 2011.
- [10] H. D. Wijaya, "Strategi Customer-Centric Dalam Marketing : Dampaknya Pada Loyalitas," vol. 1, 2023.