

PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK MENGANALISIS POLA RESERVASI PELANGGAN DI VILLA SAYONG HOUSE UBUD

Putu Bayu Surya mahesa, I Nyoman Yudi Anggara Wijaya, Made Irma Dwiputranti

Sistem Informasi, Primakara University
Jl. Tukad Badung No.135, Renon, Denpasar, Bali
Putumahesa2004@gmail.com

ABSTRAK

Persaingan industri akomodasi di Ubud yang semakin ketat menuntut pemanfaatan data reservasi secara optimal. Villa Sayong House Ubud memiliki catatan pemesanan yang belum dimanfaatkan untuk strategi pemasaran efektif. Penelitian ini menganalisis pola reservasi menggunakan algoritma K-Means Clustering untuk mengidentifikasi karakteristik pelanggan berdasarkan data historis. Metode CRISP-DM diterapkan dengan menganalisis 1.644 data reservasi periode Januari 2024–Februari 2025 mencakup empat atribut: *Booking Date*, *Source*, *Room Type*, dan *Total Night*. Pengolahan data menggunakan RapidMiner Studio dengan metode *Elbow* menghasilkan tiga kluster optimal. Hasil menunjukkan Cluster 0 merupakan pelanggan loyal dengan frekuensi tinggi dan durasi menginap panjang (8+ malam, 33,3%), Cluster 1 adalah pelanggan musiman dengan pemesanan moderat dan durasi pendek (1–3 malam, 33,2%), dan Cluster 2 adalah pelanggan reguler dengan durasi menengah (4–7 malam, 33,2%). Visualisasi mengungkapkan pola musiman di Januari–Juli dan Desember, dominasi kamar *Double Bed*, preferensi platform Booking.com dan Agoda, serta durasi menginap 2–3 malam. Rekomendasi strategi meliputi program loyalitas untuk Cluster 0, promo musiman untuk Cluster 1, dan paket bundling untuk Cluster 2. Evaluasi model menunjukkan performa baik dengan Silhouette Score 0,58 dan Davies-Bouldin Index 0,95. Penelitian membuktikan K-Means Clustering efektif mengidentifikasi pola perilaku pelanggan sebagai dasar keputusan berbasis data untuk meningkatkan okupansi villa.

Kata kunci : K-Means Clustering, Pola Reservasi, CRISP-DM, Segmentasi Pelanggan, Strategi Promosi

1. PENDAHULUAN

Pariwisata di Bali terus menunjukkan tren pertumbuhan yang positif. Berdasarkan data dari Badan Pusat Statistik (BPS) Provinsi Bali, jumlah kunjungan wisatawan mancanegara ke Bali mencapai lebih dari 5,2 juta orang pada tahun 2023 [1]. Ubud, sebagai salah satu destinasi utama di Bali, menarik sekitar 15% dari total kunjungan tersebut, menjadikannya kawasan yang sangat kompetitif dalam industri perhotelan [2]. Villa Sayong House Ubud yang terletak di Jl. Dewi Sita, Ubud, Kabupaten Gianyar, Bali, merupakan salah satu penginapan di Bali yang bergantung pada platform pemesanan online seperti Agoda, Booking.com, Tiket.com, dan Expedia untuk menarik pelanggan.

Namun, saat ini data reservasi pelanggan belum dimanfaatkan secara optimal untuk membuat strategi promosi yang sesuai. Tanpa strategi promosi yang tepat, villa ini berisiko kehilangan peluang dalam meningkatkan okupansi [3]. Kurangnya pemanfaatan data reservasi menjadi masalah utama dalam merancang strategi promosi yang efektif. Villa Sayong House Ubud memiliki banyak data pemesanan dari berbagai platform online, namun belum dimanfaatkan secara maksimal untuk dianalisis secara mendalam. Tanpa pemahaman secara menyeluruh terhadap data ini, akan sulit untuk mengidentifikasi tren pemesanan, preferensi tamu terhadap tipe kamar, serta periode dengan tingkat hunian rendah. Akibatnya, strategi pemasaran yang dijalankan menjadi kurang tepat sasaran dan tidak mampu meningkatkan okupansi secara optimal [4].

Pendekatan berbasis data menjadi solusi untuk mengoptimalkan strategi promosi. Dengan memahami pola pemesanan pelanggan, pengelola dapat menyesuaikan promosi berdasarkan preferensi tamu, periode sepi, serta kanal pemesanan yang efektif. Penelitian ini akan menggunakan metode data mining dengan algoritma K-Means untuk menganalisis data reservasi. K-Means adalah salah satu algoritma clustering yang sangat efektif dalam menangani data dengan skala besar berdasarkan karakteristik tertentu [5].

Untuk mendukung proses analisis sistematis, penelitian ini menggunakan pendekatan CRISP-DM (*Cross Industry Standard Process for Data Mining*), metodologi standar yang terdiri dari enam tahapan: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*. Penelitian ini bertujuan menganalisis pola reservasi pelanggan di Villa Sayong House Ubud menggunakan algoritma K-Means Clustering guna memperoleh pemahaman mendalam terhadap perilaku pemesanan berdasarkan data historis dan memberikan rekomendasi strategi promosi yang tepat sasaran.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian berjudul “Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering dengan Metode K-Means” menerapkan K-Means untuk clustering data transaksi penjualan dan berhasil mengelompokkan produk berdasarkan tingkat permintaan [6].

Penelitian berjudul Algoritma “K-Means Clustering Untuk Segmentasi Pelanggan (Studi Kasus: Fashion Viral Solo)” menggunakan K-Means untuk segmentasi pelanggan fashion dan menghasilkan strategi pemasaran yang lebih tepat sasaran [7].

Penelitian berjudul “Enhancing Segmentation: A Comparative Study of Clustering Methods” membandingkan berbagai metode clustering dan menemukan K-Means memberikan Silhouette Score terbaik untuk segmentasi pelanggan [8].

Penelitian berjudul “Journal of Artificial Intelligence and Engineering Applications K-Means Clustering Method to Make Credit Payment Grouping Efficient” menerapkan K-Means untuk mengelompokkan pola pembayaran kredit dengan hasil yang efisien. Penelitian-penelitian tersebut membuktikan K-Means efektif untuk berbagai aplikasi segmentasi dan analisis pola [9].

2.2. Data Mining

Data mining merupakan proses mengekstraksi informasi penting dari kumpulan data yang besar dan kompleks dengan tujuan untuk menemukan pola tersembunyi, hubungan antar data, serta mendukung pengambilan Keputusan [10]. Data mining dapat diterapkan untuk menganalisis dan mengelompokkan data dinamis, menunjukkan kemampuannya dalam mengolah data dalam jumlah besar untuk menghasilkan kluster informasi yang terstruktur. Tujuan utama data mining adalah membantu organisasi memahami perilaku dan tren berdasarkan data historis, sehingga dapat meningkatkan efisiensi, produktivitas, dan kualitas layanan.

2.3. Clustering

Clustering merupakan salah satu metode dalam data mining yang digunakan untuk membagi data ke dalam kelompok-kelompok atau kluster berdasarkan tingkat kemiripan antar data, tanpa memerlukan label atau klasifikasi sebelumnya [11]. Karena tidak menggunakan data berlabel, teknik ini digolongkan ke dalam pendekatan unsupervised learning. Clustering bermanfaat dalam mengidentifikasi pola-pola tersembunyi maupun struktur alami dari dataset yang besar dan kompleks. Tujuan utama adalah mengelompokkan data yang memiliki karakteristik serupa ke dalam satu kluster, sekaligus memastikan adanya perbedaan yang jelas antara kluster satu dengan yang lainnya [12].

2.4. Algoritma K-Means Clustering

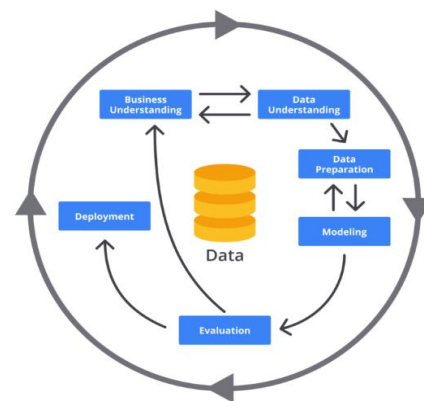
K-Means merupakan algoritma clustering yang bekerja secara iteratif untuk mengelompokkan data ke dalam beberapa cluster berdasarkan karakteristik setiap objek dalam data [13]. Algoritma ini membentuk cluster melalui perhitungan jarak terdekat antara objek dengan pusat cluster (centroid) menggunakan rumus *Euclidean Distance*. K-Means telah terbukti efektif dalam berbagai penelitian seperti segmentasi pelanggan, pengelompokan data

akademik, dan analisis pola pemesanan. Keunggulan K-Means adalah kesederhanaan implementasi, kecepatan komputasi, dan kemampuan menangani dataset besar dengan efisien.

2.5. Metode CRISP-DM

CRISP-DM (Cross Industry Standard Process for Data Mining) merupakan metodologi standar dalam pengembangan proyek data mining yang terdiri dari enam tahapan utama: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*. Metode ini bersifat fleksibel dan dapat diterapkan di berbagai industri termasuk sektor pariwisata dan perhotelan [14]. Penelitian sebelumnya menunjukkan CRISP-DM efektif dalam mengelola proyek data mining dan menghasilkan model yang valid secara statistik untuk mendukung pengambilan keputusan.

3. METODE PENELITIAN



Gambar 1. Tahapan metode CRISP-DM

Penelitian ini menggunakan metode CRISP-DM yang terdiri dari enam tahapan: (1) *Business Understanding* untuk memahami tujuan bisnis yaitu pemanfaatan data reservasi untuk strategi promosi, (2) *Data Understanding* untuk mengumpulkan dan meninjau data reservasi dari berbagai platform, (3) *Data Preparation* untuk pembersihan dan transformasi data, (4) *Modeling* menggunakan algoritma K-Means Clustering, (5) *Evaluation* untuk mengevaluasi hasil kluster, dan (6) *Deployment* untuk menyusun rekomendasi strategis [15].

3.1. Data dan Atribut Penelitian

Data yang digunakan merupakan data primer berupa informasi reservasi pelanggan dari Villa Sayong House Ubud periode Januari 2024 hingga Februari 2025, dengan total 1.644 data reservasi yang diperoleh langsung dari pihak pengelola villa. Dari seluruh atribut yang tersedia, dipilih empat atribut utama berdasarkan relevansi terhadap pola perilaku reservasi: *Booking Date* (tanggal pemesanan), *Source* (sumber pemesanan seperti Agoda, Booking.com, Tiket.com, Expedia), *Room Type* (jenis kamar: Single, Double, Triple), dan *Total Night* (jumlah malam

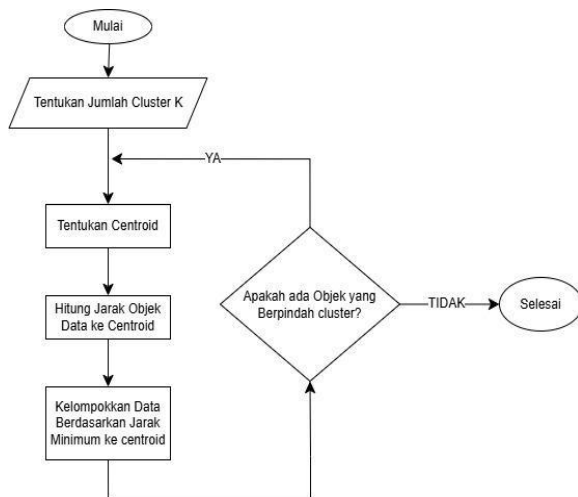
menginap). Berikut dibawah adalah tabel data atribut penelitian yang digunakan.

Tabel 1. Atribut Data Penelitian

No	Atribut	Keterangan
1	Booking Date	Tanggal pelanggan melakukan pemesanan
2	Source	Platform pemesanan (Agoda, booking.com, dll)
3	Room Type	Jenis kamar yang dipesan
4	Total Night	Jumlah malam menginap

3.2. Algoritma K-Means Clustering

Algoritma K-Means digunakan untuk mengelompokkan data reservasi berdasarkan kemiripan karakteristik dengan langkah-langkah: (1) Menentukan jumlah kluster k menggunakan metode Elbow, (2) Inisialisasi centroid secara acak, (3) Menghitung jarak setiap data ke centroid menggunakan *Euclidean Distance*, (4) Mengelompokkan data ke kluster terdekat, (5) Menghitung ulang centroid sebagai rata-rata data dalam kluster, (6) Mengulangi langkah 3-5 hingga konvergen. Alur kerja algoritma K-Means dapat dilihat pada gambar 2.



Gambar 2. Alur kerja algoritma K-Means

Gambar 2. menunjukkan alur kerja algoritma K-Means yang dimulai dari penentuan nilai k, inisialisasi centroid, perhitungan jarak, pengelompokan data, update centroid, hingga proses iterasi sampai mencapai konvergensi.

Rumus Encludian Distance yang digunakan:

$$d(x_i, u_j) = \sqrt{\sum (x_i - u_j)^2}$$

Dimana d adalah jarak antara titik data x_i dan pusat kluster u_j . Proses iterasi berhenti ketika tidak ada lagi perpindahan data antar kluster atau mencapai jumlah iterasi maksimum.

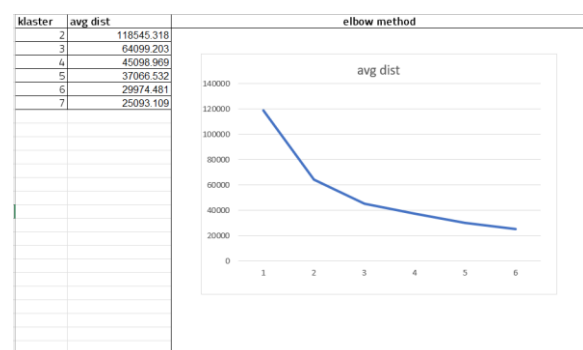
3.3. Evaluasi Model

Evaluasi model clustering dilakukan menggunakan dua metrik: Silhouette Score untuk mengukur kemiripan data dalam kluster dibandingkan dengan kluster lain (nilai -1 hingga 1, semakin mendekati 1 semakin baik), dan Davies-Bouldin Index (DBI) untuk mengukur pemisahan antar kluster (nilai semakin kecil semakin baik).

4. HASIL DAN PEMBAHASAN

4.1. Penentuan Jumlah Kluster Optimal

Penentuan jumlah kluster optimal dilakukan menggunakan metode *Elbow* dengan menghitung nilai *Within-Cluster Sum of Squares* (WCSS) untuk berbagai nilai k (1 hingga 10). Hasil pengujian metode Elbow ditampilkan pada Gambar 3.

Gambar 3. Hasil pengolahan *Elbow Method*

Berdasarkan Gambar 3, nilai average distance mengalami penurunan signifikan dari kluster 2 hingga kluster 3, kemudian mulai melandai setelah kluster ke-3. Titik "elbow" terlihat pada kluster ke-3, menunjukkan bahwa jumlah kluster optimal adalah 3 kluster dengan keseimbangan antara kompleksitas model dan efektivitas pengelompokan.

4.2. Hasil Clustering Data Reservasi

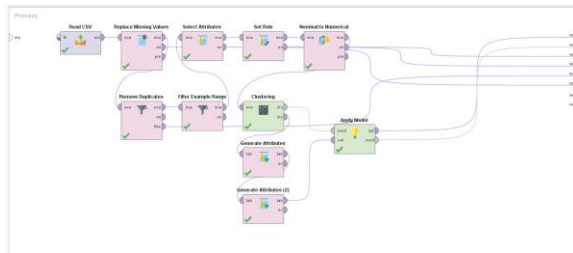
Proses clustering menggunakan RapidMiner Studio menghasilkan pengelompokan data reservasi menjadi 3 kluster dengan distribusi seimbang. Algoritma mencapai konvergensi setelah 7 iterasi dengan nilai WCSS stabil pada 4.567,23. Hasil pembagian data ditampilkan pada Tabel 2.

Tabel 2. Hasil pengelompokan data reservasi

Cluster	Jumlah Data	Karakteristik Utama
Cluster 0	547	Pelanggan loyal, long stay (8+ malam)
Cluster 1	545	Pelanggan musiman, short stay (1-3 malam)
Cluster 2	546	Pelanggan regular, medium stay (4-7 malam)

Dari Tabel 2 dapat dilihat bahwa distribusi data pada ketiga cluster relatif seimbang, menunjukkan bahwa algoritma K-Means berhasil melakukan pengelompokan yang proporsional tanpa adanya dominasi berlebihan pada satu cluster tertentu. Proses

clustering dilakukan menggunakan RapidMiner Studio dengan konfigurasi operator seperti yang ditunjukkan pada Gambar 4.



Gambar 4. Model proses clustering di RapidMiner

Gambar 4 menunjukkan alur proses *clustering* di RapidMiner yang dimulai dari pembacaan data (*Read CSV*), preprocessing dengan *Replace Missing Values* dan *Nominal to Numerical*, hingga proses clustering dengan operator K-Means dan evaluasi hasil.

4.3. Karakteristik Setiap Klaster

Karakteristik setiap cluster dapat dilihat dari nilai centroid yang terbentuk sebagaimana ditampilkan pada Tabel 3.

Tabel 3. Nilai centroid setiap cluster

CLuster	Booking Date	Source	Room Type	Total Night
Cluster 0	273.0	1.978	3.130	4.576
Cluster 1	1365.0	2.305	3.850	4.459
Cluster 2	2145.0	3.200	4.500	6.000

Cluster 0 (Pelanggan Loyal): Merupakan kelompok pelanggan dengan frekuensi reservasi tinggi dan durasi menginap panjang (8+ malam). Karakteristik meliputi pemesanan rutin sepanjang tahun, preferensi kamar Double Bed, dan dominan menggunakan platform Booking.com. Cluster ini merepresentasikan pelanggan setia yang berpotensi untuk program loyalitas.

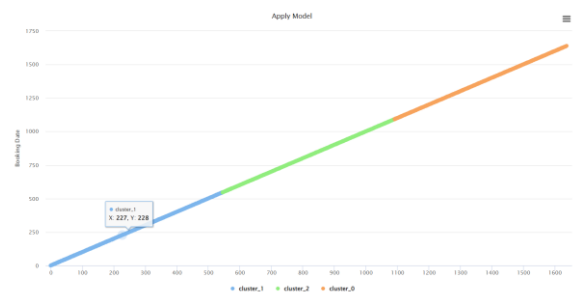
Cluster 1 (Pelanggan Musiman): Mencakup pelanggan dengan aktivitas pemesanan moderat dan durasi menginap pendek (1-3 malam). Karakteristik menunjukkan pemesanan meningkat pada periode Januari-Juli dan Desember, preferensi kamar Double Bed, dan menggunakan platform Agoda dan Expedia. Cluster ini merepresentasikan wisatawan weekend atau liburan singkat.

Cluster 2 (Pelanggan Regular): Berisi pelanggan dengan durasi menginap menengah (4-7 malam). Karakteristik menunjukkan pola pemesanan yang stabil, variasi tipe kamar lebih beragam, dan menggunakan berbagai platform pemesanan. Cluster ini merepresentasikan wisatawan *leisure* yang menginap untuk relaksasi atau eksplorasi.

4.4. Analisis Visualisasi Atribut

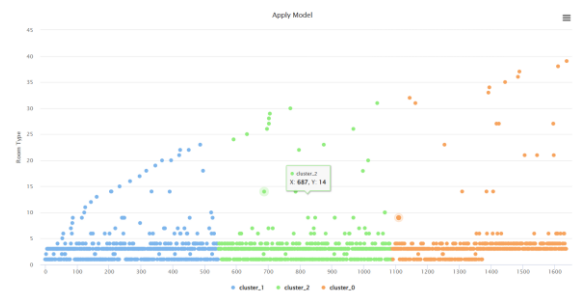
Visualisasi setiap atribut dilakukan untuk memahami pola distribusi data pada masing-masing

cluster. Gambar 5 menunjukkan visualisasi atribut Booking Date.



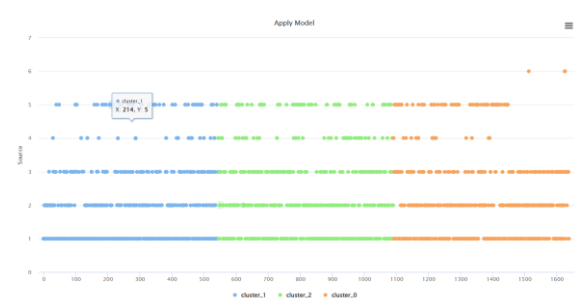
Gambar 5. Visualisasi Booking Date

Dari Gambar 5 dapat dilihat bahwa frekuensi reservasi meningkat pada periode Januari-Juli dan Desember dengan puncak pada Januari (255 reservasi), menandakan pola musiman dalam pemesanan kamar. Visualisasi atribut Room Type ditampilkan pada Gambar 6 yang menunjukkan preferensi pelanggan terhadap jenis kamar.



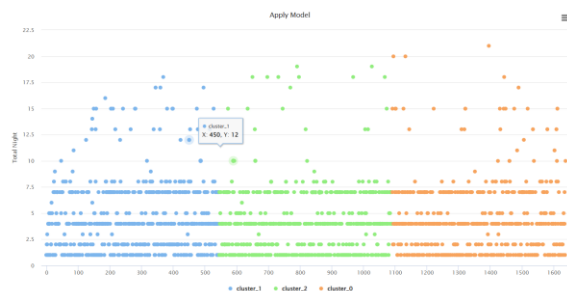
Gambar 6. Visualisasi Room Type

Gambar 6 memperlihatkan dominasi kamar Double Bed sebagai pilihan utama pelanggan di seluruh cluster, diikuti oleh tipe Single Bed dan Triple Bed. Distribusi sumber pemesanan divisualisasikan pada Gambar 7 untuk mengidentifikasi platform yang paling banyak digunakan.



Gambar 7. Visualisasi Source

Dari Gambar 7 terlihat bahwa Booking.com (624 reservasi) dan Expedia (487 reservasi) merupakan platform utama yang digunakan pelanggan, sementara pemesanan langsung relatif lebih sedikit. Visualisasi durasi menginap ditampilkan pada Gambar 8 untuk memahami pola lama tinggal pelanggan.



Gambar 8. Visualisasi Total Night

Gambar 8 mengungkapkan bahwa mayoritas pelanggan menginap selama 2-3 malam, menandakan pola kunjungan jangka pendek yang umum di wilayah wisata seperti Ubud.

4.5. Evaluasi Performa Model

Evaluasi model clustering dilakukan menggunakan metrik Silhouette Score dan Davies-Bouldin Index (DBI). Hasil evaluasi ditampilkan pada Tabel 4.

Tabel 4. Hasil evaluasi model clustering

Matrik Evaluasi	Nilai	Interpretasi
Silhouette Score	0,58	Kualitas clustering baik
Davies-Bouldin Index	0,95	Pemisahan cluster optimal

Dari Tabel 4 dapat dilihat bahwa Silhouette Score sebesar 0,58 mengindikasikan kualitas clustering baik, dimana data dalam setiap cluster memiliki kemiripan tinggi dan terpisah dengan baik dari cluster lain. Davies-Bouldin Index (DBI) sebesar 0,95 menunjukkan nilai rendah yang menandakan pemisahan cluster optimal dengan overlap minimal. Hasil evaluasi mengkonfirmasi bahwa algoritma K-Means dengan 3 cluster berhasil mengidentifikasi pola perilaku pelanggan secara efektif.

4.6. Rekomendasi Strategi Promosi

Berdasarkan hasil clustering dan analisis visualisasi, direkomendasikan strategi promosi yang disesuaikan dengan karakteristik setiap cluster sebagaimana dirangkum dalam Tabel 5.

Tabel 5. Rekomendasi strategi promosi

Cluster	Target Pelanggan	Strategi Promosi
Cluster 0	Pelanggan Loyal	Program loyalitas: member card dengan diskon progresif 10-15%, complimentary upgrade room, prioritas booking
Cluster 1	Pelanggan Musiman	Promo musiman: early booking discount 20% untuk periode peak season, last-minute deals weekend dengan late check-out gratis
Cluster 2	Pelanggan reguler	Paket bundling: aktivitas wisata lokal (rice terrace tour, yoga class, cooking class) dengan harga ekonomis

Dari Tabel 5 dapat dilihat bahwa setiap cluster memiliki strategi promosi yang berbeda sesuai karakteristiknya. Selain itu, disarankan untuk mengoptimasi platform Booking.com dan Agoda sambil mendorong direct booking melalui website resmi dengan benefit eksklusif, serta menyediakan Weekend Package Promo dengan potongan harga untuk pemesanan minimal 2 malam.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil menerapkan algoritma K-Means Clustering untuk menganalisis pola reservasi pelanggan di Villa Sayong House Ubud dengan menggunakan metode CRISP-DM. Dari 1.644 data reservasi periode Januari 2024–Februari 2025, diperoleh 3 klaster optimal yang merepresentasikan segmen pelanggan berbeda: Cluster 0 (pelanggan loyal dengan long stay 8+ malam, 33,3%), Cluster 1 (pelanggan musiman dengan short stay 1-3 malam, 33,2%), dan Cluster 2 (pelanggan reguler dengan medium stay 4-7 malam, 33,2%). Metode Elbow berhasil menentukan jumlah klaster optimal dengan nilai WCSS stabil pada 4.567,23 setelah 7 iterasi. Evaluasi model menunjukkan performa baik dengan Silhouette Score 0,58 dan Davies-Bouldin Index 0,95. Visualisasi mengungkapkan pola musiman pada Januari-Juli dan Desember, dominasi kamar Double Bed, preferensi platform Booking.com dan Agoda, serta durasi menginap rata-rata 2-3 malam. Rekomendasi strategi promosi disesuaikan dengan karakteristik setiap cluster meliputi program loyalitas, promo musiman, dan paket bundling. Penelitian membuktikan K-Means Clustering efektif mengidentifikasi pola perilaku pelanggan sebagai dasar keputusan berbasis data untuk meningkatkan okupansi dan pendapatan villa. Untuk penelitian selanjutnya, disarankan menambahkan atribut lain seperti harga, rating pelanggan, atau demografi, serta membandingkan performa K-Means dengan algoritma clustering lainnya seperti DBSCAN atau Hierarchical Clustering untuk mendapatkan hasil yang lebih komprehensif.

DAFTAR PUSTAKA

- [1] BPS, "Badan Pusat Statistik Provinsi Bali."
- [2] P. Krista, "Ubud Masuk 5 besar Kota Terbaik di Dunia Versi Travel and Leisure." Accessed: May 19, 2025. [Online]. Available: https://www.detik.com/bali/wisata/d-6835550/ubud-masuk-5-besar-kota-terbaik-di-dunia-versi-travel-and-leisure#google_vignette
- [3] A. Pardede Br, A. Sri Sulityawati, and I. N. Tri, "Strategi pemasaran pada masa pandemi di hotel sofitel bali nusa dua beach resort, bali," vol. 8, Nov. 2024.
- [4] Y. Sulistyaningtyas, "Analisis Kontribusi Online Segment Terhadap Tingkat Hunian Di Swiss-Belhotel Balikpapan," vol. 5, pp. 83–86, 2023.
- [5] I. D. Setiawan and A. Triayudi, "Penerapan Data Mining Dengan Menggunakan Algoritma

- Clustering K-Means Untuk Pembagian Jurusan Pada Sekolah Menengah Atas,” *Journal of Computer System and Informatics (JoSYC)*, vol. 5, no. 2, pp. 380–392, Feb. 2024, doi: <https://doi.org/10.47065/josyc.v5i2.4970>.
- [6] N. Afiasari, N. Suarna, and N. Rahaningsi, “Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering dengan Metode K-Means,” *Jurnal SAINTEKOM*, vol. 13, no. 1, pp. 100–110, Mar. 2023, doi: <https://doi.org/10.33020/saintekom.v13i1.40>.
- [7] R. B. Ardi, F. Ely Nastiti, and S. Sumarlinda, “Algoritma K-Means Clustering Untuk Segmentasi Pelanggan (Studi Kasus: Fashion Viral Solo),” *INFOTECH journal*, vol. 9, no. 1, pp. 124–131, May 2023, doi: <https://doi.org/10.31949/infotech.v9i1.5214>.
- [8] L. S. Ling and C. T. Weiling, “Enhancing Segmentation: A Comparative Study of Clustering Methods,” *IEEE Access*, 2025, doi: <https://ieeexplore.ieee.org/document/10921704>.
- [9] S. Nur Illah, N. Suarna, I. Ali, and D. Solihudin, “Journal of Artificial Intelligence and Engineering Applications K-Means Clustering Method to Make Credit Payment Groupinhg Efficient,” 2025. [Online]. Available: <https://ioinformatic.org/>
- [10] K. U. Fernando, J. Titaley, and K. A. Lapu, “Analisis Clustering Perubahan Cuaca dan Iklim di Kota Manado Menggunakan Metode K-Means,” 2022. [Online]. Available: <https://ejournal.unsrat.ac.id/v3/index.php/IJIDS/index>
- [11] S. Sindi *et al.*, “Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia,” *Jurnal Teknologi Informasi*, vol. 4, no. 1, 2020.
- [12] R. H. Zain, Randy Permana, and Sarjon Defit, “Multi-Process Data Mining with Clustering and Support Vector Machine for Corporate Recruitment,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 9, no. 2, pp. 276–282, Mar. 2025, doi: <https://doi.org/10.29207/resti.v9i2.6197>.
- [13] R. Riadi and Mesran, “Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Analisa Penjualan Parfume,” *Journal of Informatics, Electrical and Electronics Engineering*, vol. 2, no. 4, pp. 138–145, Jun. 2023, doi: <https://doi.org/10.47065/jieeee.v2i4.1181>.
- [14] Y. Suhandi, I. Kurniati, and S. Norma, “Penerapan Metode Crisp-DM Dengan Algoritma K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik,” *Jurnal Teknologi Informatika dan Komputer*, vol. 6, no. 2, pp. 12–20, Sep. 2020, doi: <https://doi.org/10.37012/jtik.v6i2.299>.
- [15] M. Nazar Alwi, “CRISP-DM: Tahapan, Studi Kasus, Kelebihan, dan Kekurangan,” Muhammad Nazar Alwi .[16] J. Jabbar, “Sistem Informasi Stok Barang Menggunakan Metode Clustering Kmeans (Studi Kasus Rmd Store),” *INFOTECH journal*, vol. 8, no. 1, pp. 70–75, Apr. 2022