

ANALISIS TINGKAT KERAMAIAN PELANGGAN DIGIDREAMS MENGUNAKAN METODE K-MEANS CLUSTERING DAN APRIORI

Muhammad Saifullah, Dwi Hartanti, Ridwan Dwi Irawan

Teknik Informatika, Universitas Duta Bangsa Surakarta
Jl. Bhayangkara No 55-57 Tipes, Serengan, Kota Surakarta, Jawa Tengah
220103199@mhs.udb.ac.id

ABSTRAK

Digidreams Space merupakan *co-working space* yang menyediakan ruang kerja bersama dengan pola kunjungan pelanggan yang berbeda setiap hari. Dalam operasionalnya, Digidreams menghadapi permasalahan berupa fluktuasi tingkat keramaian yang belum teridentifikasi secara sistematis, sehingga berpotensi menyebabkan ketidakseimbangan pemanfaatan fasilitas dan inefisiensi sumber daya operasional. Penelitian ini bertujuan menganalisis pola keramaian dan karakteristik operasional harian Digidreams melalui integrasi metode K-Means Clustering dan Apriori. Penelitian menggunakan kerangka CRISP-DM dengan data historis 110 hari periode September–Desember 2025 yang mencakup jumlah pelanggan harian, *income*, *outcome*, dan *profit*. K-Means diterapkan dengan Elbow Method untuk menentukan jumlah *cluster* optimal, sedangkan Apriori menggunakan minimum *support* 5% dan minimum *confidence* 50%. Hasil K-Means dengan K=3 menghasilkan *Silhouette score* 0,3057 yang termasuk kategori *fair clustering*. Tiga *cluster* yang terbentuk adalah Ramai sebanyak 40 hari (36,4%), Sedang 57 hari (51,8%), dan Inefisien 13 hari (11,8%). *Cluster* Inefisien menunjukkan *profit* negatif Rp280.000 akibat *outcome* rata-rata Rp825.192. Apriori menghasilkan 77 *frequent itemsets* dan 51 *association rules*, dengan rule “Day_Saturday, Income_Tinggi → Pelanggan_Banyak” sebagai pola asosiasi dominan.

Kata kunci : K-Means Clustering, Apriori, Data Mining, Co-working Space, Prediksi Keramaian, CRISP-DM

1. PENDAHULUAN

Kemajuan teknologi digital telah mengubah kebiasaan kerja dan belajar masyarakat secara signifikan, sehingga meningkatkan permintaan akan ruang kerja yang fleksibel, nyaman, dan produktif [1]. *Co-working space* merupakan konsep lingkungan kerja yang menyediakan ruang bagi berbagai individu dengan latar belakang beragam seperti entrepreneur, freelancer, startup, konsultan, investor, dan pelajar [2]. Dinamika kunjungan pada *co-working space* dipengaruhi oleh berbagai faktor seperti hari, waktu, dan kebutuhan pengguna, sehingga pemahaman terhadap pola keramaian menjadi sangat penting untuk memastikan kualitas layanan [3].

Digidreams Space adalah *co-working space* yang melayani berbagai macam pengguna dengan fasilitas seperti internet cepat, monitor tambahan, proyektor, tempat istirahat, ruang sholat, dan tempat parkir [4]. Operasional harian Digidreams sangat fluktuatif karena fleksibilitas layanannya dan keberagaman penggunaannya. Berdasarkan data historis selama periode September hingga Desember 2025 yang terdiri dari 110 hari pengamatan dengan jumlah pelanggan harian, *income*, *outcome*, dan *profit*, terdapat fluktuasi yang menunjukkan adanya perbedaan tingkat keramaian pada periode tertentu. Data menunjukkan nilai *income* tertinggi mencapai Rp1.061.000 dengan 60 pelanggan, sedangkan rata-rata *income* berada pada kisaran Rp568.515 dengan rata-rata 26,8 pelanggan per hari. Terdapat juga kondisi ekstrem dimana *profit* mencapai nilai negatif hingga -Rp615.500 akibat *outcome* yang sangat tinggi pada hari dengan kunjungan rendah.

Ketidakpastian tingkat keramaian menyebabkan beberapa permasalahan operasional yang berdampak langsung pada kinerja bisnis Digidreams. Pertama, ketidakseimbangan pemanfaatan fasilitas terjadi karena manajemen tidak dapat mengantisipasi lonjakan atau penurunan jumlah pengunjung secara akurat. Kedua, perubahan beban kerja operasional yang tidak terprediksi mengakibatkan alokasi sumber daya manusia menjadi tidak optimal. Ketiga, potensi penurunan kualitas layanan pada periode dengan aktivitas tinggi akibat persiapan yang tidak memadai [5]. Selama ini, manajemen operasional masih bergantung pada perkiraan berdasarkan pengalaman tanpa didukung analisis data historis yang terstruktur, sehingga berpotensi menimbulkan ketidaktepatan dalam pengambilan keputusan dan inefisiensi dalam pengelolaan sumber daya.

Beberapa penelitian sebelumnya telah menerapkan kombinasi metode K-Means dan Apriori untuk menganalisis pola operasional bisnis. Penelitian [6] menunjukkan bahwa integrasi K-Means dan Apriori mampu mengidentifikasi pola pembelian dan mendukung penyusunan strategi pemasaran berbasis data melalui pembentukan *cluster* dan *association rules* yang bermakna. Penelitian [7] menerapkan metode yang sama pada aplikasi MPStore dan menghasilkan tiga klaster terbaik dengan *Silhouette Coefficient* sebesar 0,91 serta aturan asosiasi yang signifikan untuk mendukung strategi pemasaran. Namun, masih terbatas penelitian yang mengintegrasikan kedua metode tersebut untuk menganalisis tingkat keramaian pada *co-working space* berdasarkan data historis yang mencakup jumlah

pelanggan harian, *income*, *outcome*, dan *profit* secara bersamaan.

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan penerapan integrasi metode K-Means *Clustering* dan Apriori untuk menganalisis tingkat keramaian pelanggan Digidreams. Penelitian ini bertujuan mengidentifikasi pola kunjungan pelanggan, mengelompokkan tingkat keramaian berdasarkan karakteristik operasional harian, dan menemukan pola asosiasi antar variabel yang memengaruhi keramaian. K-Means digunakan untuk mengelompokkan data harian berdasarkan kesamaan karakteristik jumlah pelanggan, *income*, *outcome*, dan *profit*, sedangkan Apriori menganalisis pola asosiasi antar variabel untuk menemukan hubungan yang sering muncul. Hasil penelitian diharapkan dapat membantu Digidreams merancang strategi operasional yang lebih efektif, mulai dari alokasi sumber daya, pengendalian biaya, hingga perencanaan promosi yang responsif terhadap fluktuasi keramaian pelanggan.

2. TINJAUAN PUSTAKA

Penelitian [6] menerapkan integrasi algoritma K-Means dan Apriori dalam analisis pola pembelian untuk mendukung strategi pemasaran pada bisnis kuliner Premium Salad.co. K-Means digunakan untuk membentuk klaster pelanggan berdasarkan karakteristik transaksi, sedangkan Apriori menghasilkan aturan asosiasi sebagai dasar rekomendasi paket menu dengan nilai *support* dan *confidence* yang telah ditentukan. Temuan penelitian menunjukkan bahwa integrasi K-Means dan Apriori mampu menghasilkan wawasan strategis untuk mendukung pengambilan keputusan bisnis.

Penelitian [7] mengimplementasikan metode K-Means dan Apriori dalam klasterisasi pengguna serta analisis pola pembelian pada aplikasi MPStore. K-Means menghasilkan tiga klaster terbaik dengan nilai *Silhouette Coefficient* sebesar 0,91, sedangkan Apriori menghasilkan aturan asosiasi yang signifikan pada setiap klaster untuk mendukung strategi pemasaran dan peningkatan layanan. Hasil penelitian menunjukkan bahwa integrasi kedua metode tersebut efektif dalam mengidentifikasi pola perilaku pengguna guna mendukung keputusan strategis perusahaan.

2.1. Data Mining

Data mining adalah proses mengekstraksi informasi dan pola yang bernilai dari kumpulan data yang besar sehingga dapat dimanfaatkan dalam pengambilan keputusan bisnis [8]. Proses data mining mencakup beberapa tahapan sistematis mulai dari pemilihan data, *preprocessing*, transformasi, penerapan algoritma, hingga interpretasi hasil. Teknik ini digunakan untuk menemukan pengetahuan baru yang sebelumnya tersembunyi dalam data transaksi atau operasional pelanggan. Pemanfaatan data mining memungkinkan organisasi mengubah data historis menjadi informasi strategis, sehingga keputusan pemasaran, pelayanan, dan operasional dapat

dilakukan secara lebih efektif berdasarkan pola nyata daripada intuisi [9]. *Data mining* mencakup berbagai teknik seperti *classification*, *clustering*, *association rule mining*, dan *regression* yang dapat disesuaikan dengan kebutuhan analisis bisnis [10].

2.2. Clustering

Clustering merupakan salah satu teknik dalam data mining yang bertujuan mengelompokkan data ke dalam beberapa kelompok berdasarkan kemiripan karakteristiknya tanpa menggunakan label yang ditentukan sebelumnya (*unsupervised learning*) [11]. Objek dalam satu *cluster* memiliki tingkat kemiripan yang tinggi satu sama lain, sedangkan objek antar *cluster* memiliki perbedaan yang signifikan. Metode *clustering* membantu pelaku bisnis memahami struktur data sehingga strategi pemasaran dan pengelolaan layanan dapat dilakukan secara tepat sasaran. Dengan melakukan pengelompokan, pelaku usaha dapat mengetahui segmen atau kategori tertentu seperti kelompok tingkat keramaian pelanggan yang berbeda-beda, yang dapat dimanfaatkan untuk merancang strategi operasional yang lebih spesifik dan efektif sesuai karakteristik masing-masing segmen.

2.3. Algoritma K-Means

Algoritma K-Means adalah metode *clustering* yang membagi data ke dalam k kelompok berdasarkan jarak antar data terhadap pusat *cluster* (*centroid*). Proses dimulai dengan inisialisasi k *centroid* secara acak, kemudian setiap data *point* diassign ke *centroid* terdekat menggunakan jarak Euclidean. Setelah itu, *centroid* diupdate berdasarkan rata-rata dari semua data *point* dalam *cluster* tersebut. Proses ini diulang hingga konvergen, yaitu ketika tidak ada lagi perubahan *assignment* data atau perubahan posisi *centroid* sudah sangat kecil [12]. Kelebihan K-Means adalah waktu komputasi yang cepat dan hasil pengelompokan yang jelas, sehingga sangat cocok digunakan untuk mengelompokkan tingkat keramaian pada data kunjungan pelanggan. Kualitas hasil *clustering* K-Means dapat dievaluasi menggunakan *Silhouette Score* yang mengukur seberapa baik pemisahan antar *cluster* [13].

2.4. Association Rule Mining

Association Rule Mining adalah teknik untuk menemukan hubungan antar item dalam suatu dataset sehingga pola transaksi atau kebiasaan pelanggan dapat diketahui. Analisis ini mampu mengungkap kombinasi variabel yang muncul secara bersamaan dalam data sehingga menghasilkan *rule* yang bermanfaat bagi pengembangan strategi bisnis [6]. *Rule* yang dihasilkan memiliki bentuk IF (*antecedent*) THEN (*consequent*), dimana *antecedent* adalah kondisi awal dan *consequent* adalah hasil yang terjadi. *Association rule mining* sangat relevan digunakan untuk mengetahui pola keterkaitan faktor-faktor penyebab keramaian pelanggan seperti hari

kunjungan, *event* promosi, atau momen tertentu yang dapat mempengaruhi tingkat kunjungan.

2.5. Algoritma Apriori

Algoritma Apriori merupakan salah satu algoritma association rule mining yang paling banyak digunakan dan bekerja berdasarkan prinsip bahwa setiap subset dari *frequent itemset* juga harus *frequent* [6]. Proses Apriori dimulai dengan menghitung *support* setiap item tunggal, kemudian secara iteratif membentuk kandidat *itemset* yang lebih besar dengan menggabungkan *frequent itemset* sebelumnya dan mengeliminasi kandidat yang tidak memenuhi minimum *support* (pruning). Setelah semua *frequent itemsets* ditemukan, aturan asosiasi dibentuk berdasarkan nilai minimum *confidence* yang telah ditentukan. Parameter utama dalam Apriori adalah *support* yang mengukur frekuensi kemunculan *itemset* dalam data, *confidence* yang mengukur probabilitas *consequent* benar jika *antecedent* terjadi, dan lift yang mengukur kekuatan asosiasi relatif terhadap kejadian acak [14]. Nilai lift lebih besar dari 1 menunjukkan asosiasi positif, sama dengan 1 menunjukkan independensi, dan kurang dari 1 menunjukkan asosiasi negatif. Pada penelitian ini, algoritma Apriori digunakan untuk menemukan pola yang memengaruhi tingkat keramaian di Digidreams berdasarkan kombinasi hari, kategori pelanggan, kategori *income*, dan kategori *profit*.

2.6. CRISP-DM

CRISP-DM (*Cross Industry Standard Process for Data Mining*) merupakan suatu standar proses yang digunakan dalam penelitian *data mining* untuk mengolah dan menganalisis data secara sistematis. Metode ini memiliki enam tahapan utama yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment* yang digunakan untuk membantu menyelesaikan permasalahan bisnis melalui pemanfaatan data yang tersedia [15]. CRISP-DM telah terbukti efektif sebagai kerangka metodologi dalam berbagai penelitian *data mining* karena strukturnya yang sistematis dan iteratif, memungkinkan peneliti untuk kembali ke tahap sebelumnya jika diperlukan perbaikan atau penyesuaian. Tahapan CRISP-DM memastikan bahwa proyek *data mining* dilakukan dengan terstruktur mulai dari pemahaman masalah bisnis hingga implementasi solusi dalam bentuk sistem yang dapat digunakan.

3. METODE PENELITIAN

3.1. Metode Pengumpulan Data

Penelitian ini menggunakan data historis operasional harian Digidreams selama periode September hingga Desember 2025. Data yang dikumpulkan mencakup 110 hari pengamatan dengan atribut tanggal kunjungan, hari dalam seminggu, bulan, jumlah pelanggan harian, *income*, *outcome*, dan

profit. Metode pengumpulan data yang digunakan adalah:

- Data sekunder: mengumpulkan data penjualan historis dari pencatatan data yang ada pada Digidreams.
- Observasi: pengamatan langsung terhadap sistem pencatatan kedatangan pelanggan untuk memperoleh data aktual terkait jumlah pelanggan setiap hari serta pola keramaian.

3.2. Metode Pengembangan Sistem

Penelitian ini menggunakan metode CRISP-DM sebagai kerangka pengembangan [15], yang terdiri dari enam tahapan:

- Business Understanding*
Mengidentifikasi permasalahan fluktuasi keramaian pelanggan Digidreams dan menetapkan tujuan analisis pola kunjungan untuk mendukung keputusan operasional.
- Data Understanding*
Mengumpulkan dan mengeksplorasi data historis selama 110 hari yang meliputi tanggal, hari, bulan, jumlah pelanggan, *income*, *outcome*, dan *profit*.
- Data Preparation*
Melakukan pembersihan data, penanganan missing values, seleksi atribut, dan transformasi data. Data K-Means dinormalisasi menggunakan StandardScaler, sedangkan data Apriori didiskretisasi menjadi kategori pelanggan, *income*, *outcome*, dan *profit*.
- Modeling*
Pemodelan dilakukan menggunakan K-Means untuk mengelompokkan data berdasarkan jumlah pelanggan, *income*, *outcome*, dan *profit*, dengan penentuan jumlah *Cluster* melalui Elbow Method dan *Silhouette score*. Selanjutnya, Apriori digunakan untuk menemukan pola asosiasi dengan minimum *support* 5% dan minimum *confidence* 50%.
- Evaluation*
Hasil K-Means dievaluasi menggunakan *Silhouette score*, sedangkan aturan Apriori dinilai berdasarkan *support*, *confidence*, dan lift. Fungsionalitas sistem diuji menggunakan Black Box Testing.
- Deployment*
Hasil analisis disajikan melalui aplikasi berbasis Streamlit yang dijalankan secara lokal dalam bentuk informasi dan visualisasi untuk mendukung pengambilan keputusan operasional.

4. HASIL DAN PEMBAHASAN

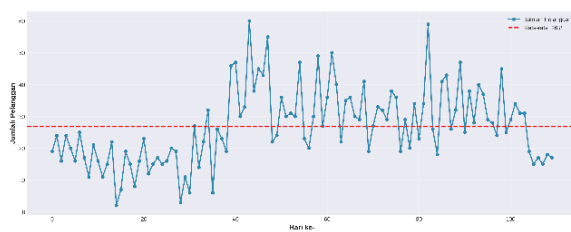
4.1. Analisis Statistik Deskriptif

Tahap awal analisis dilakukan dengan mengeksplorasi karakteristik data operasional DigiDreams selama 110 hari (1 September 2025 – 19 Desember 2025). Data mencakup empat variabel utama: jumlah pelanggan harian, *income*, *outcome*, dan *profit*. Ringkasan statistik deskriptif disajikan pada Tabel 1.

Tabel 1. Statistik Deskriptif Data Historis Digidreams

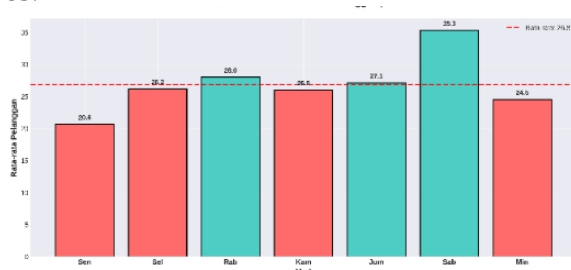
Variabel	Count	Mean	Std	Min	25%	50%	75%	Max
Jumlah Pelanggan	110	26.8	11.94	2	18.25	25.5	34.0	60
Income (Rp)	110	568,515	169,884	230,000	442,250	536,000	679,500	1,061,000
Outcome (Rp)	110	368,634	247,332	8,000	204,750	329,000	452,750	1,413,000
Profit (Rp)	110	199,881	254,856	-615,500	76,262.5	218,500	341,125	794,000

Tabel 1 menunjukkan jumlah pelanggan berkisar antara 2–60 orang dengan rata-rata 26,8 dan standar deviasi 11,94, yang mengindikasikan adanya fluktuasi kunjungan harian yang cukup tinggi. Variabel *Outcome* memiliki standar deviasi sebesar Rp247.332 dengan nilai maksimum mencapai Rp1.413.000 atau sekitar 3,8 kali lipat dari nilai rata-ratanya. Sementara itu, variabel *profit* menunjukkan nilai minimum sebesar -Rp615.500, yang menandakan adanya hari-hari dengan kerugian operasional.



Gambar 1. Tren Pelanggan Harian

Gambar 1 menunjukkan fluktuasi jumlah pelanggan harian dengan garis putus-putus merah menandakan rata-rata 26,8 pelanggan. Terlihat pola naik-turun yang signifikan dengan beberapa puncak mencapai 60 pelanggan pada hari ke-43 dan hari ke-83.



Gambar 2. Rata-rata Jumlah Pelanggan per Hari

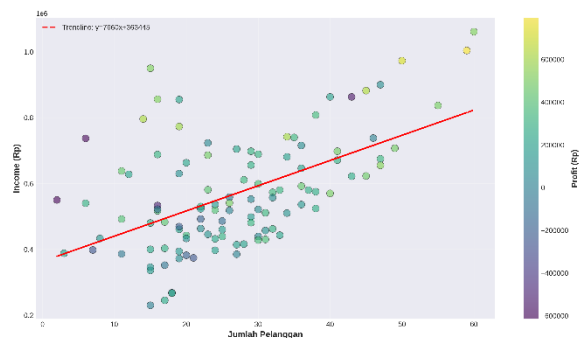
Gambar 2 menunjukkan Sabtu memiliki rata-rata tertinggi (35,3 pelanggan), diikuti Rabu (28,0) dan Jumat (27,1), sedangkan Senin terendah (20,6), mengindikasikan keramaian tidak hanya terjadi di akhir pekan.

Gambar 3 menunjukkan matriks korelasi antar variabel. Jumlah pelanggan memiliki korelasi positif sedang terhadap *income* sebesar 0,538 dan korelasi positif rendah hingga sedang terhadap *profit* sebesar 0,392. Hal ini menunjukkan bahwa peningkatan jumlah pelanggan cenderung diikuti oleh kenaikan *income* dan *profit*. Sebaliknya, *outcome* memiliki korelasi negatif kuat terhadap *profit* sebesar -0,772, yang mengindikasikan bahwa semakin tinggi pengeluaran operasional, maka *profit* cenderung menurun. Sementara itu, korelasi antara jumlah

pelanggan dan *outcome* sangat lemah, yaitu sebesar -0,035, sehingga menunjukkan bahwa perubahan jumlah pelanggan tidak memiliki hubungan yang berarti dengan besarnya *outcome*.



Gambar 3. Correlation Matrix Antar Variabel



Gambar 4. Hubungan Jumlah Pelanggan vs Income

Gambar 4 menampilkan scatter plot antara jumlah pelanggan dan *income* dengan trendline $y = 7660x + 363445$. Garis tren menunjukkan bahwa setiap penambahan 1 pelanggan berkaitan dengan peningkatan *income* rata-rata sebesar Rp7.660. Temuan ini sejalan dengan korelasi positif antara jumlah pelanggan dan *income* ($r = 0,538$). Selain itu, selisih *outcome* antara hari dengan pelanggan rendah (<20 orang, Rp380.952) dan tinggi (≥ 30 orang, Rp366.762) hanya sebesar 3,7%, meskipun jumlah pelanggan berbeda hingga 174%. Hal ini mengindikasikan bahwa *outcome* relatif stabil dan dipengaruhi oleh komponen biaya tetap.

4.2. Preprocessing dan Diskretisasi Data

Sebelum diterapkan algoritma Apriori, data numerik perlu diubah menjadi data kategorikal melalui

proses diskretisasi. Aturan diskretisasi yang digunakan disajikan pada Tabel 2.

Tabel 2. Aturan Diskretisasi Variabel

Variabel	Rentang	Kategori
Jumlah Pelanggan	0 – 15	Sedikit
Jumlah Pelanggan	16 – 30	Sedang
Jumlah Pelanggan	> 30	Banyak
Income	< Persentil 33%	Rendah
Income	Persentil 33% – 66%	Sedang
Income	> Persentil 66%	Tinggi
Outcome	< Median	Rendah
Outcome	≥ Median	Tinggi
Profit	< Persentil 33%	Rendah
Profit	Persentil 33% – 66%	Sedang
Profit	> Persentil 66%	Tinggi

Tabel 2 menunjukkan bahwa jumlah pelanggan didiskretisasi menggunakan batasan tetap, sedangkan variabel finansial menggunakan pendekatan berbasis persentil agar kategori lebih seimbang. Variabel hari dalam seminggu juga ditambahkan untuk menangkap pola temporal dalam *association rules*.

4.3. Hasil K-Means Clustering

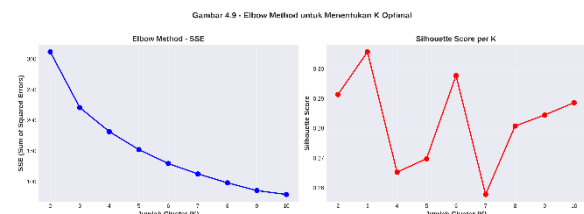
Penentuan jumlah *cluster* optimal dilakukan menggunakan *Elbow Method* dan analisis *Silhouette Score* untuk nilai K dari 2 hingga 10.

Berdasarkan Tabel 3 dan Gambar 5, penurunan SSE terbesar terjadi dari K=2 ke K=3 sebesar 29,17%, sedangkan penurunan pada K berikutnya relatif lebih kecil. Selain itu, *Silhouette Score* tertinggi diperoleh

pada K=3 sebesar 0,3057. Dengan demikian, K=3 dipilih sebagai jumlah *cluster* paling sesuai. Nilai *Silhouette Score* tersebut termasuk kategori *fair clustering*, sehingga hasil pengelompokan cukup layak digunakan untuk menggambarkan pola operasional Digidreams meskipun masih terdapat kemungkinan overlap antar *cluster* [13].

Tabel 3. Hasil Evaluasi Jumlah Cluster Menggunakan Elbow Method

K	SSE	Penurunan SSE (%)	Silhouette Score
2	311.58	0.00	0.2914
3	220.68	29.17	0.3057
4	181.45	17.78	0.2652
5	152.10	16.17	0.2698
6	129.28	15.01	0.2977
7	112.48	12.99	0.2577
8	97.85	13.01	0.2808
9	85.16	12.96	0.2845
10	78.82	7.44	0.2886



Gambar 5. Elbow Method untuk Menentukan K Optimal

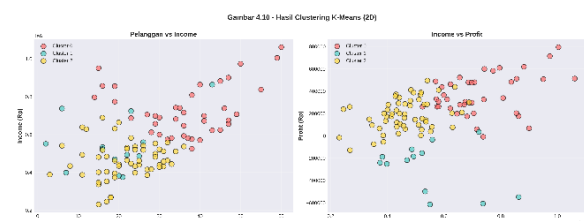
Tabel 4. Karakteristik Detail Setiap Cluster

Cluster	Label	Jml Data	Persentase	Avg Pelanggan	Avg Income	Avg Outcome	Avg Profit	Min-Max Pelanggan
0	Ramai	40	36.4%	36.5	Rp727.729	Rp357.320	+Rp370.409	14-60
1	Inefisien	13	11.8%	18.9	Rp545.192	Rp825.192	-Rp280.000	2-43
2	Sedang	57	51.8%	21.7	Rp462.105	Rp272.446	+Rp189.659	3-36

Berdasarkan Tabel 4, hasil *clustering* membentuk tiga segmen operasional, yaitu Ramai, Sedang, dan Inefisien. Cluster Ramai (36,4%) menunjukkan performa terbaik dengan rata-rata 36,5 pelanggan dan *profit* Rp370.409, sedangkan Cluster Sedang (51,8%) merupakan segmen terbesar dengan *profit* rata-rata Rp189.659. Sementara itu, Cluster Inefisien (11,8%) memiliki *profit* negatif sebesar Rp280.000 akibat *outcome* yang lebih tinggi dibandingkan *income*. Temuan ini menunjukkan bahwa kerugian lebih dipengaruhi oleh tingginya pengeluaran operasional daripada rendahnya jumlah pelanggan.

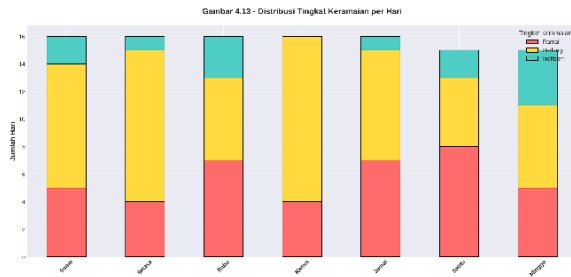
Gambar 6 menunjukkan bahwa Cluster Ramai terkonsentrasi pada tingkat pelanggan dan *income* yang tinggi, sedangkan Cluster Sedang mendominasi area pelanggan dan *income* rendah hingga sedang. Di sisi lain, Cluster Inefisien tersebar pada berbagai tingkat pelanggan, tetapi cenderung memiliki *profit* negatif meskipun *income* relatif tinggi. Temuan ini

mengindikasikan bahwa kerugian pada Cluster Inefisien lebih dipengaruhi oleh tingginya pengeluaran dibandingkan rendahnya pendapatan.



Gambar 6. Hasil Clustering K-Means (2D)

Gambar 7 menunjukkan bahwa Cluster Ramai paling banyak terjadi pada hari Sabtu, sedangkan Cluster Inefisien paling sering muncul pada hari Minggu. Sementara itu, beberapa hari kerja cenderung didominasi oleh Cluster Sedang sebagai kondisi operasional normal.



Gambar 7. Distribusi Tingkat Keramaian per Hari

Untuk memastikan bahwa hasil *clustering* tidak bersifat kebetulan dan dapat digeneralisasi, dilakukan validasi menggunakan *5-Fold Cross Validation*. Data dibagi menjadi 5 subset yang berbeda, dan algoritma K-Means dijalankan ulang pada setiap subset [16]. Hasil validasi disajikan pada Tabel 5.

Tabel 5. Hasil 5-Fold Cross Validation Clustering

Fold	Silhouette Score	Davies-Bouldin Index	Calinski-Harabasz Index
1	0,3032	1,0977	40,65
2	0,3290	0,9682	44,90
3	0,3088	1,0565	46,49
4	0,3071	1,0740	43,09
5	0,3076	1,0678	43,02
Rata-rata	0,3112 ± 0,0091	1,0528	43,63

Berdasarkan Tabel 5, hasil *5-fold cross validation* menunjukkan bahwa pembentukan tiga *cluster* relatif stabil. Hal ini ditunjukkan oleh nilai *Silhouette Score* yang konsisten pada setiap *fold* dengan rata-rata $0,3112 \pm 0,0091$. Selain itu, nilai *Davies-Bouldin Index* sebesar 1,0528 dan *Calinski-Harabasz Index* sebesar 43,63 mendukung kualitas pemisahan antar *cluster*. Dengan demikian, hasil *clustering* dinilai cukup stabil dan dapat diandalkan.

Tabel 7. Top 10 Association rules Berdasarkan Lift

IF (Antecedent)	THEN (Consequent)	Support (%)	Confidence (%)	Lift
Day_Saturday, Income_Tinggi	Profit_Tinggi, Pelanggan_Banyak	5,45	75,00	4,34
Day_Saturday, Profit_Tinggi	Pelanggan_Banyak, Income_Tinggi	5,45	75,00	4,12
Day_Saturday, Pelanggan_Banyak	Profit_Tinggi, Income_Tinggi	5,45	66,67	3,86
Profit_Tinggi, Pelanggan_Banyak, Income_Tinggi	Day_Saturday	5,45	50,00	3,67
Day_Saturday, Income_Tinggi	Pelanggan_Banyak	7,27	100,00	2,97
Day_Saturday, Income_Tinggi, Profit_Tinggi	Pelanggan_Banyak	5,45	100,00	2,97
Day_Saturday	Pelanggan_Banyak, Income_Tinggi	7,27	53,33	2,93
Day_Saturday, Pelanggan_Banyak	Income_Tinggi	7,27	88,89	2,64
Day_Saturday, Profit_Tinggi	Pelanggan_Banyak	6,36	87,50	2,60
Day_Saturday, Pelanggan_Banyak, Profit_Tinggi	Income_Tinggi	5,45	85,71	2,55

Dari Tabel 7 dapat dilihat bahwa seluruh *rules* terkuat melibatkan variabel Day_Saturday sebagai *antecedent* utama. Rule dengan *lift* tertinggi (4,34) menyatakan bahwa jika hari Sabtu dan *income* tinggi, maka 75% kemungkinan *profit* juga tinggi dan

4.4. Hasil Algoritma Apriori

Analisis Apriori diterapkan dengan minimum *support* 5% dan minimum *confidence* 50%. Proses mining menghasilkan 77 *frequent itemsets*. Tabel 6 menyajikan 15 *frequent itemsets* teratas berdasarkan nilai *support*.

Tabel 6. Top 15 Frequent itemsets

Itemsets	Support (%)
Pelanggan_Sedang	50,91
Income_Rendah	33,64
Income_Tinggi	33,64
Pelanggan_Banyak	33,64
Profit_Rendah	33,64
Profit_Tinggi	33,64
Income_Sedang	32,73
Profit_Sedang	32,73
Income_Rendah, Pelanggan_Sedang	21,82
Profit_Rendah, Pelanggan_Sedang	20,00
Income_Sedang, Pelanggan_Sedang	18,18
Pelanggan_Banyak, Income_Tinggi	18,18
Profit_Sedang, Pelanggan_Sedang	18,18
Profit_Tinggi, Income_Tinggi	17,27

Tabel 6 menunjukkan Pelanggan_Sedang sebagai *itemset* paling sering muncul (50,91%), diikuti pasangan *Income_Rendah* + Pelanggan_Sedang (21,82%) dan *Profit_Rendah* + Pelanggan_Sedang (20,00%), mengindikasikan hari sedang cenderung disertai pendapatan dan laba yang rendah.

Dari proses mining Apriori, total dihasilkan 51 *association rules* yang memenuhi threshold minimum *support* 5% dan *confidence* 50%. Dari 51 *rules* tersebut, dipilih 10 *rules* dengan nilai *lift* tertinggi untuk disajikan sebagai *rules* yang paling bermakna secara manajerial. Tabel 7 menyajikan 10 *association rules* terkuat berdasarkan nilai *lift*.

pelanggan banyak. Rule dengan *confidence* 100% menunjukkan bahwa jika hari Sabtu dan *income* tinggi, pelanggan dapat dipastikan dalam kondisi banyak (*lift* 2,97). Secara keseluruhan, hari Sabtu terbukti menjadi

prediktor keramaian yang paling dominan dalam data operasional DigiDreams.

Selain *rules* berbasis hari Sabtu, terdapat pula *rules actionable* terkait efisiensi biaya yang relevan

untuk manajemen operasional, sebagaimana disajikan pada Tabel 8.

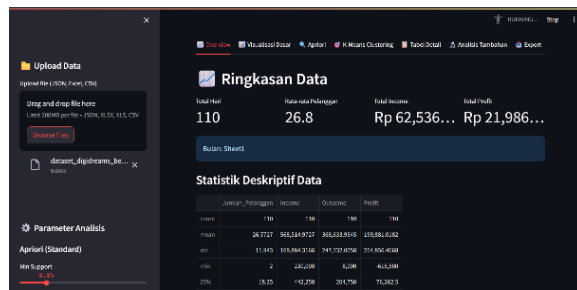
Tabel 8. *Association rules Actionable* (Efisiensi Biaya)

IF (Antecedent)	THEN (Consequent)	Support (%)	Confidence (%)	Lift	Kategori Aksi
<i>Income_Rendah, Outcome_Tinggi</i>	<i>Profit_Rendah</i>	11,8	100,00	3,06	Monitoring
<i>Income_Rendah, Outcome_Tinggi, Pelanggan_Sedang</i>	<i>Profit_Rendah</i>	9,09	100,00	3,06	Cost Efficiency
<i>Income_Tinggi, Outcome_Rendah</i>	<i>Profit_Tinggi</i>	10,91	100,00	2,89	Optimasi Revenue
<i>Outcome_Tinggi, Profit_Tinggi</i>	<i>Income_Tinggi</i>	8,18	100,00	2,89	Monitoring
<i>Income_Rendah, Outcome_Tinggi</i>	<i>Profit_Rendah, Pelanggan_Sedang</i>	9,09	76,92	4,03	Monitoring
<i>Profit_Sedang, Outcome_Rendah, Pelanggan_Sedang</i>	<i>Income_Rendah</i>	10,00	91,67	2,80	Monitoring

Berdasarkan Tabel 8, kombinasi *Income_Rendah* dan *Outcome_Tinggi* selalu berasosiasi dengan *Profit_Rendah* (*confidence* 100%; *lift* 3,06). Sebaliknya, kombinasi *Income_Tinggi* dan *Outcome_Rendah* selalu menghasilkan *Profit_Tinggi* (*confidence* 100%; *lift* 2,89). Temuan ini menunjukkan bahwa pengendalian biaya merupakan faktor penting untuk menjaga *profitabilitas*, terutama pada periode dengan pendapatan rendah.

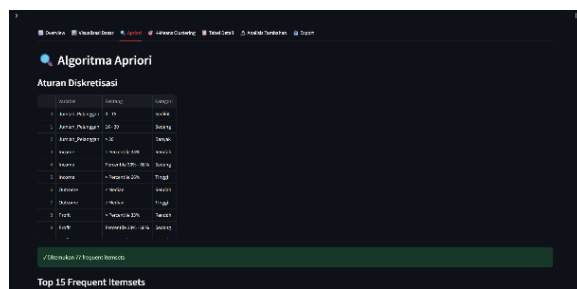
4.5. Implementasi Sistem

Sistem diimplementasikan menggunakan framework Streamlit yang dijalankan secara lokal melalui Google Colab dengan tunneling pyngrok.



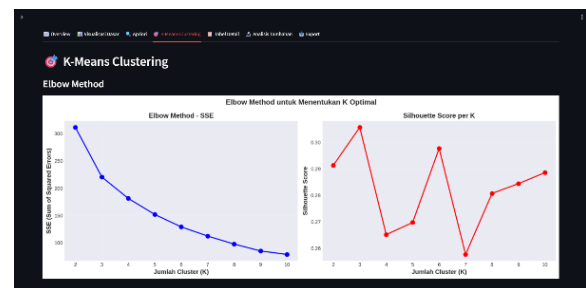
Gambar 8. Tampilan Awal / Halaman Upload

Gambar 8 menunjukkan Tab *Overview* yang memuat empat *metric card* (Total Hari, Rata-rata Pelanggan, Total *Income*, Total *Profit*) dan tabel statistik deskriptif.



Gambar 9. Tab Apriori

Gambar 9 menyajikan Top 15 *Frequent Itemsets* dan Top 10 *Association Rules* beserta interpretasi bisnis.



Gambar 10. Tab K-Means (Elbow)

Gambar 10 menunjukkan tampilan Tab *K-Means Clustering* yang menyajikan hasil analisis pengelompokan data. Pada tampilan ini ditampilkan grafik *Elbow Method*.

4.6. Pembahasan

Integrasi hasil *K-Means Clustering* dan *Apriori* menghasilkan pemahaman yang komprehensif tentang pola operasional DigiDreams [6], [7]. Dari sisi *K-Means*, teridentifikasi tiga segmen hari operasional yang berbeda secara signifikan: *Cluster* Ramai yang merupakan kondisi ideal bisnis, *Cluster* Sedang yang menjadi kondisi dominan dengan 57 hari atau 51,8% dari total data, dan *Cluster* Inefisien yang menjadi prioritas perhatian meskipun hanya mencakup 13 hari atau 11,8% dari keseluruhan periode pengamatan. Sebagaimana diidentifikasi pada sub-bab 4.3, *Cluster* Inefisien dicirikan oleh *outcome* ekstrem yang mengindikasikan pencampuran CAPEX dan OPEX dalam pencatatan, bukan sekadar rendahnya jumlah pelanggan.

Dari sisi *Apriori*, hari Sabtu terbukti secara statistik sebagai hari dengan probabilitas keramaian tertinggi, di mana 8 dari 10 *rules* terkuat melibatkan *Day_Saturday* sebagai *antecedent* utama. Rule "*Day_Saturday* + *Income_Tinggi* → *Pelanggan_Banyak*" dengan *confidence* 100% dan *lift*

2,97 memberikan landasan yang kuat untuk strategi operasional berbasis hari. Selain itu, *rules* efisiensi biaya yang teridentifikasi memberikan panduan bahwa pengendalian *outcome* pada hari berpendapatan rendah merupakan kunci peningkatan *profitabilitas* secara keseluruhan. Konvergensi temuan dari dua metode yang berbeda ini semakin diperkuat oleh konsistensi hasil *clustering* yang menunjukkan bahwa 53,3% hari Sabtu masuk ke dalam *Cluster* Ramai, sehingga memperkuat validitas kesimpulan penelitian.

Berdasarkan gap analysis, kondisi aktual menghasilkan total *profit* Rp21.986.912 selama 110 hari dengan rata-rata Rp199.881 per hari. Apabila seluruh hari dapat beroperasi seperti *Cluster* Ramai, proyeksi total *profit* dapat mencapai Rp40.744.954 dengan rata-rata Rp370.409 per hari. Gap sebesar Rp18.758.042 atau 85,3% ini dapat dipersempit melalui strategi konkret, yaitu meningkatkan rata-rata pelanggan harian dari 26,8 menjadi 36,5 orang, mengendalikan pengeluaran pada hari-hari inefisien, serta mengoptimalkan promosi pada hari-hari strategis berdasarkan temuan *association rules*.

5. KESIMPULAN DAN SARAN

Penelitian ini menganalisis pola keramaian dan karakteristik operasional harian Digidreams melalui integrasi algoritma *K-Means Clustering* dan Apriori berdasarkan 110 hari data historis periode September–Desember 2025. Hasil *K-Means Clustering* membentuk tiga klaster, yaitu Ramai sebesar 36,4%, Sedang sebesar 51,8%, dan Inefisien sebesar 11,8%, dengan nilai *Silhouette Score* 0,3057 yang menunjukkan kategori *fair clustering*. Klaster Inefisien ditandai oleh pengeluaran yang melebihi pendapatan sebesar 51%. Sementara itu, Apriori menghasilkan 77 *frequent itemsets* dan 51 *association rules*, dengan hari Sabtu sebagai prediktor keramaian paling dominan (*confidence* 100%; *lift* 2,97), yang menunjukkan bahwa pendapatan tinggi pada hari Sabtu selalu berkaitan dengan banyaknya pelanggan. Setiap penambahan satu pelanggan juga diperkirakan berkaitan dengan peningkatan pendapatan rata-rata sebesar Rp7.660,00. Temuan ini dapat mendukung pengelola Digidreams dalam alokasi sumber daya, pengendalian pengeluaran, strategi promosi, penjadwalan *maintenance*, dan perencanaan kapasitas. Penelitian selanjutnya disarankan untuk menambahkan variabel eksternal, memperpanjang periode data, mengembangkan model prediktif, membandingkan metode *clustering* lain, serta membangun sistem *monitoring* berbasis *cloud* dengan fitur *real-time alert*.

DAFTAR PUSTAKA

- [1] V. F. Nahuway, “Manajemen perkantoran modern di era digitalisasi: Suatu tinjauan literatur,” *J. Adm. Terap.*, vol. 3, no. 1, pp. 303–315, 2024.
- [2] M. Faisal and M. I. Sa’ud, “Perancangan *co-working space* dengan penerapan fleksibilitas

ruang di Kota Yogyakarta,” *J. Tugas Akhir Mhs. Lanting*, vol. 12, pp. 104–118, 2023.

- [3] I. W. A. Wikantara, I. G. A. N. P. Utami, I. M. D. J. M. Saputra, I. M. W. S. Pranata, and K. B. C. Diputra, *The Spirit of Bali Business: Ketika Budaya, Kreativitas, dan Digital Berkolaborasi*, 1st ed. Badung: Nilacakra Publishing House, 2026.
- [4] “Digidreams Space,” *Digidreams Space*. [Online]. Available: <https://digidreams.space/>. Accessed: Dec. 11, 2025.
- [5] M. M. Nashir and C. S. Widagdo, “Optimalisasi operasional dapur hotel: Analisis hambatan dan strategi efisiensi,” *PUBLIK: J. Manaj. Sumber Daya Manusia, Administrasi dan Pelayanan Publik*, vol. 13, pp. 220–230, 2026, doi: 10.37606/publik.v13i1.2199.
- [6] V. E. Putri and H. D. Purnomo, “Integrasi algoritma Apriori dan K-Means dalam analisis pola pembelian untuk meningkatkan strategi pemasaran,” *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 10, no. 1, pp. 409–423, 2025, doi: 10.29100/jupi.v10i1.5768.
- [7] K. Anam, W. Setiawan, E. Mala, and S. Rochman, “Klusterisasi pengguna aplikasi MPStore dan analisis pola pembelian menggunakan K-Means dan Apriori (studi kasus PT Mitra Pedagang Indonesia),” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 2, pp. 2361–2367, 2025, doi: 10.36040/jati.v9i2.13074.
- [8] I. Syafii, A. A. Ribhi, L. Y. Astutik, G. K. S. Budiono, and A. S. Pamela, “Analisis prediksi customer repeat order menggunakan algoritma Decision Tree pada perusahaan transportasi,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, pp. 1372–1378, 2024, doi: 10.57152/malcom.v4i4.1538.
- [9] S. Syam et al., *Data Mining: Teori dan Penerapannya dalam Berbagai Bidang*, 1st ed. Jambi, Indonesia: PT Sonpedia Publishing Indonesia, 2024.
- [10] R. Swastika, S. Mukodimah, F. Susanto, M. Muslihudin, and S. Ipnuwati, *Implementasi Data Mining (Clustering, Association, Prediction, Estimation, Classification)*, cet. 1. Indramayu: Penerbit Adab, 2023.
- [11] F. Handayani, “Aplikasi data mining menggunakan algoritma K-Means Clustering untuk mengelompokkan mahasiswa berdasarkan gaya belajar,” *J. Teknol. dan Inf.*, vol. 12, 2022, doi: 10.34010/jati.v12i1.
- [12] D. Anggraini and M. S. Hasibuan, “Initial centroid pada algoritma K-Means dan K-Medoids,” *Jatilima: J. Multimedia dan Teknologi Informasi*, vol. 7, no. 3, pp. 487–500, 2025, doi: 10.54209/jatilima.v7i03.1615.
- [13] Abdullah and K. Setiawan, “Implementasi algoritma Apriori terhadap data penjualan untuk mengetahui pola pembelian konsumen pada

- Koperasi IDN Boarding School,” *J. Indonesia: Manajemen Informatika dan Komunikasi*, vol. 4, no. 3, pp. 1646–1653, 2023, doi: 10.35870/jimik.v4i3.390.
- [14] H. Saputra, R. A. Uneputty, L. A. Simyapen, M. F. B. Mustamir, F. A. Ikawanti, and S. P. Kusumawati, “Analisis pola kegiatan belajar mahasiswa terhadap keberhasilan akademik menggunakan algoritma Apriori,” *J. Sains Informatika Terapan (JSIT)*, vol. 4, no. 2, pp. 143–154, 2025, doi: 10.62357/jsit.v4i2.567.
- [15] I. Fitrianti, A. Voutama, and Y. Umaidah, “Clustering film populer pada aplikasi Netflix dengan menggunakan algoritma K-Means dan metode CRISP-DM,” *JTSI*, vol. 4, no. 2, pp. 301–311, 2023.
- [16] A. Fauzan, N. Suarna, I. Ali, and H. Susana, “Penerapan algoritma K-Means Clustering untuk meningkatkan model pengelompokan dan kinerja jaringan Wi-Fi secara optimal,” *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, vol. 13, no. 2, 2025, doi: 10.23960/jitet.v13i2.6272.