

Integrasi K-Means Clustering dan Rasio Potensi Pasar untuk Prioritisasi Lokasi Toko Bahan Bangunan Berbasis Data

Afreza Wahyu Nugroho-1^{a*}, Budi Santoso-2^a, Bunga Intan-3^a, Sri Tita Faulina-4^a

^aUniversitas Bina Insan, Indonesia

*Email: rezawahyu0410@gmail.com^{1*}, budisantoso@univbinainsan.ac.id²,
bungaintan@univbinainsan.ac.id³, sttabta@gmail.com⁴

Abstrak

Pemilihan lokasi toko bahan bangunan memerlukan pertimbangan terhadap karakteristik wilayah dan tingkat potensi pasar agar keputusan penentuan lokasi tidak hanya didasarkan pada satu indikator. Penelitian ini bertujuan mengelompokkan kecamatan di Kota Lubuklinggau berdasarkan karakteristik demografis dan ketersediaan toko bahan bangunan serta mengintegrasikan hasil pengelompokan dengan rasio potensi pasar untuk menentukan tingkat prioritas lokasi usaha. Penelitian menggunakan pendekatan kuantitatif berbasis *data mining* dengan kerangka CRISP-DM. Data penelitian mencakup delapan kecamatan di Kota Lubuklinggau, dengan sumber data jumlah penduduk, kepadatan penduduk, dan laju pertumbuhan penduduk dari BPS Kota Lubuklinggau serta data jumlah toko bahan bangunan yang diperoleh melalui Google Maps. Pengelompokan dilakukan menggunakan algoritma K-Means setelah proses normalisasi data, sedangkan potensi pasar dihitung melalui rasio jumlah penduduk terhadap jumlah toko bahan bangunan. Hasil penelitian menunjukkan bahwa data delapan kecamatan terbagi menjadi tiga cluster, yaitu Potensi Rendah, Potensi Sedang, dan Potensi Tinggi, dengan nilai *Silhouette Score* sebesar 0,5102. Integrasi hasil *clustering* dan rasio potensi pasar menghasilkan tiga kategori, yaitu Sangat Strategis, Strategis, dan Cukup Strategis. Lubuklinggau Barat I termasuk kategori Sangat Strategis karena berada pada cluster Potensi Tinggi dan memiliki rasio potensi pasar kategori Tinggi. Hasil tersebut menunjukkan bahwa integrasi *clustering* dan rasio potensi pasar dapat digunakan sebagai dasar penyaringan awal dalam menentukan prioritas lokasi toko bahan bangunan.

Kata Kunci: *Data Mining, K-Means Clustering, Lokasi Usaha, Rasio Potensi Pasar, Toko Bahan Bangunan*

1. Latar Belakang

Pemilihan lokasi merupakan salah satu aspek penting dalam pengembangan usaha karena berkaitan dengan jangkauan pasar, tingkat persaingan, dan peluang keberlanjutan usaha [1]. Kondisi demografis suatu wilayah dapat memberikan gambaran mengenai karakteristik pasar yang tersedia, terutama melalui jumlah penduduk, kepadatan penduduk, dan laju pertumbuhan penduduk. Wilayah dengan jumlah penduduk yang besar memiliki basis

konsumen yang lebih luas, sedangkan kepadatan dan pertumbuhan penduduk dapat memberikan informasi tambahan mengenai konsentrasi serta perkembangan kebutuhan masyarakat. Informasi tersebut menjadi semakin penting bagi usaha yang produknya berkaitan dengan kebutuhan pembangunan dan renovasi, seperti toko bahan bangunan. Penentuan lokasi usaha yang hanya mengandalkan pengalaman atau pertimbangan subjektif berpotensi mengabaikan perbedaan karakteristik pasar antarwilayah. Pemanfaatan data

demografis dan data persebaran usaha dapat memberikan dasar yang lebih terukur dalam mengidentifikasi wilayah yang memiliki karakteristik pasar berbeda [2].

Perkembangan data mining memberikan peluang untuk memanfaatkan data wilayah sebagai dasar dalam proses pengambilan Keputusan [3]. Salah satu teknik yang dapat digunakan adalah *clustering*, yaitu proses pengelompokan data berdasarkan kemiripan karakteristik. Algoritma K-Means merupakan salah satu metode *clustering* yang banyak diterapkan untuk mengelompokkan wilayah berdasarkan karakteristik demografis [4]. Penerapannya pada data kepadatan penduduk di Provinsi DKI Jakarta mampu menghasilkan tiga kelompok wilayah berdasarkan tingkat kepadatan penduduk, yaitu kelompok tinggi, sedang, dan rendah, dengan evaluasi menggunakan *Davies-Bouldin Index* [5]. Penggunaan K-Means juga diterapkan untuk mengelompokkan jumlah penduduk berdasarkan wilayah administrasi dan menghasilkan kelompok dengan karakteristik jumlah penduduk yang berbeda [6]. Penerapan serupa pada pemetaan kepadatan penduduk menunjukkan bahwa K-Means dapat digunakan untuk membentuk kelompok wilayah berdasarkan karakteristik jumlah penduduk sehingga pola perbedaan antarwilayah lebih mudah diidentifikasi [7].

Pemanfaatan K-Means tidak terbatas pada pengelompokan karakteristik penduduk, tetapi juga telah digunakan dalam kajian penentuan lokasi usaha. Identifikasi lokasi strategis *coffee shop* menggunakan data spasial berupa koordinat lokasi usaha dan algoritma K-Means mampu mengelompokkan lokasi usaha menjadi beberapa wilayah strategis berdasarkan kedekatan karakteristik spasialnya [8]. Pendekatan pengelompokan juga dapat menggabungkan karakteristik demografis dengan ketersediaan fasilitas untuk menghasilkan tipologi wilayah. Penggunaan variabel kepadatan penduduk dan rasio ketersediaan fasilitas pelayanan

dasar dapat membentuk tipologi kecamatan berdasarkan karakteristik demografi dan kondisi pelayanan wilayah [9]. Temuan tersebut menunjukkan bahwa K-Means dapat digunakan untuk mengidentifikasi pola karakteristik wilayah, tetapi hasil pengelompokan pada dasarnya menggambarkan kesamaan atribut antarwilayah dan belum secara langsung menunjukkan besarnya peluang pasar suatu usaha.

Keterbatasan tersebut menjadi penting ketika K-Means digunakan untuk mendukung penentuan lokasi usaha toko bahan bangunan. Pengelompokan wilayah berdasarkan karakteristik demografis dan jumlah toko dapat menunjukkan kelompok wilayah yang memiliki kondisi relatif serupa, tetapi belum secara langsung menggambarkan hubungan antara jumlah calon konsumen dan jumlah usaha yang telah tersedia [10]. Dua wilayah dapat memiliki jumlah penduduk yang relatif tinggi, tetapi peluang pengembangan usahanya dapat berbeda apabila jumlah toko bahan bangunan yang telah beroperasi juga berbeda. Wilayah dengan jumlah toko yang lebih banyak dapat memiliki tingkat persaingan yang lebih tinggi, sedangkan wilayah dengan jumlah toko yang lebih sedikit relatif terhadap jumlah penduduk dapat menunjukkan adanya ruang pasar yang lebih besar [11]. Oleh karena itu, hasil *clustering* perlu dilengkapi dengan analisis yang dapat menggambarkan perbandingan antara potensi permintaan dan ketersediaan usaha.

Kondisi tersebut relevan dengan karakteristik wilayah Kota Lubuklinggau yang terdiri atas delapan kecamatan dengan kondisi demografis dan persebaran toko bahan bangunan yang berbeda. Data penelitian mencakup jumlah penduduk, kepadatan penduduk, laju pertumbuhan penduduk, serta jumlah toko bahan bangunan pada masing-masing kecamatan. Perbedaan karakteristik tersebut menunjukkan bahwa setiap kecamatan tidak memiliki kondisi pasar yang sama. Beberapa wilayah memiliki jumlah

penduduk dan pertumbuhan penduduk yang relatif tinggi, sementara wilayah lainnya memiliki kepadatan penduduk atau jumlah toko bahan bangunan yang lebih tinggi. Variasi tersebut perlu dianalisis secara bersamaan agar identifikasi wilayah potensial tidak hanya didasarkan pada satu karakteristik demografis [12]. Data jumlah toko juga diperlukan untuk melihat kondisi ketersediaan usaha yang telah ada pada masing-masing wilayah.

Permasalahan utama dalam penentuan lokasi toko bahan bangunan adalah belum tersedianya pendekatan yang secara bersamaan menggambarkan karakteristik wilayah dan keseimbangan antara potensi permintaan dengan ketersediaan usaha. Penggunaan data demografis saja belum cukup untuk menunjukkan apakah suatu wilayah masih memiliki peluang pengembangan usaha, sedangkan penggunaan jumlah toko saja belum menggambarkan besarnya pasar yang dapat dilayani [13]. Kondisi tersebut dapat menyebabkan wilayah dengan pasar potensial belum teridentifikasi secara optimal atau wilayah dengan tingkat persaingan yang relatif tinggi tetap dipertimbangkan sebagai lokasi pengembangan usaha. Pendekatan yang mengintegrasikan beberapa indikator diperlukan agar hasil analisis dapat memberikan informasi yang lebih terstruktur mengenai karakteristik dan potensi setiap kecamatan [14].

Salah satu pendekatan pendukung yang dapat digunakan adalah rasio antara jumlah penduduk dan jumlah toko bahan bangunan. Dalam penelitian ini, rasio tersebut digunakan untuk menggambarkan Rasio Potensi Pasar, yaitu perbandingan jumlah penduduk sebagai indikator permintaan dengan jumlah toko bahan bangunan sebagai indikator ketersediaan usaha. Nilai rasio yang lebih tinggi menunjukkan bahwa jumlah penduduk yang dilayani oleh setiap toko relatif lebih besar, sedangkan nilai rasio yang lebih rendah menunjukkan bahwa jumlah toko relatif lebih banyak dibandingkan jumlah

penduduk. Dengan demikian, rasio tersebut dapat memberikan informasi tambahan mengenai peluang pengembangan usaha yang tidak diperoleh hanya melalui proses *clustering* [15].

Penelitian terdahulu memperlihatkan bahwa K-Means telah digunakan untuk mengelompokkan kepadatan penduduk, jumlah penduduk, karakteristik wilayah, maupun lokasi usaha [16]. Penggunaan K-Means pada kepadatan penduduk mampu membedakan wilayah berdasarkan tingkat kepadatannya, sedangkan penerapan pada data jumlah penduduk menghasilkan kelompok wilayah dengan karakteristik populasi yang berbeda. Pada konteks lokasi usaha, K-Means juga dapat digunakan untuk mengidentifikasi kelompok lokasi strategis berdasarkan data spasial [17]. Penggabungan karakteristik demografis dengan ketersediaan fasilitas menunjukkan bahwa informasi mengenai kondisi penduduk dan ketersediaan suatu layanan dapat digunakan secara bersamaan untuk membentuk tipologi wilayah [18] [19]. Meskipun demikian, pendekatan-pendekatan tersebut belum secara khusus mengintegrasikan karakteristik demografis, jumlah toko bahan bangunan, dan rasio antara jumlah penduduk dengan jumlah toko untuk menentukan prioritas lokasi pengembangan toko bahan bangunan pada tingkat kecamatan.

Kesenjangan tersebut menunjukkan adanya kebutuhan untuk mengembangkan pendekatan yang tidak hanya mengelompokkan wilayah berdasarkan kemiripan karakteristik, tetapi juga mempertimbangkan potensi pasar relatif pada masing-masing wilayah. Penggunaan beberapa indikator secara simultan dalam analisis wilayah memungkinkan karakteristik suatu wilayah direpresentasikan secara lebih komprehensif dibandingkan penggunaan satu indikator saja [20]. K-Means digunakan untuk mengidentifikasi kelompok kecamatan berdasarkan kombinasi jumlah penduduk, kepadatan penduduk, laju pertumbuhan penduduk,

dan jumlah toko bahan bangunan. Penggunaan empat variabel tersebut memungkinkan karakteristik wilayah dianalisis secara lebih menyeluruh dibandingkan penggunaan satu indikator demografis saja. Hasil *clustering* kemudian dilengkapi dengan Rasio Potensi Pasar untuk menggambarkan perbandingan antara jumlah penduduk dan jumlah toko bahan bangunan. Hasil dari kedua analisis tersebut selanjutnya digunakan secara terpadu untuk menentukan prioritas lokasi berdasarkan karakteristik wilayah dan potensi pasar. Penggunaan beberapa variabel secara bersamaan dalam proses *clustering* juga sejalan dengan hasil penelitian yang menunjukkan bahwa karakteristik suatu wilayah dapat berbeda berdasarkan kombinasi jumlah penduduk, kepadatan, pertumbuhan penduduk, dan ketersediaan usaha [21].

Pendekatan yang digunakan dalam penelitian ini memberikan kontribusi pada penerapan *data mining* untuk analisis lokasi usaha, khususnya melalui integrasi K-Means dan Rasio Potensi Pasar pada sektor toko bahan bangunan. K-Means berperan dalam mengidentifikasi karakteristik kelompok kecamatan, sedangkan Rasio Potensi Pasar memberikan informasi mengenai perbandingan potensi permintaan dengan ketersediaan toko. Integrasi keduanya menghasilkan dasar analisis yang lebih terstruktur untuk mengidentifikasi wilayah yang memiliki karakteristik dan peluang pasar yang berbeda. Hasil analisis diharapkan dapat menjadi informasi pendukung bagi proses penentuan lokasi usaha berbasis data, tanpa mengesampingkan faktor lain seperti pendapatan, daya beli masyarakat, aksesibilitas, kondisi geografis, kebijakan wilayah, dan preferensi pelaku usaha yang tidak menjadi variabel dalam penelitian ini.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk (1) mengelompokkan kecamatan di Kota Lubuklinggau berdasarkan karakteristik jumlah penduduk, kepadatan penduduk, laju pertumbuhan penduduk, dan jumlah

toko bahan bangunan menggunakan algoritma K-Means; (2) menganalisis potensi pasar toko bahan bangunan melalui perbandingan jumlah penduduk dan jumlah toko bahan bangunan; serta (3) mengintegrasikan hasil *clustering* dan Rasio Potensi Pasar untuk menentukan prioritas lokasi pengembangan toko bahan bangunan pada tingkat kecamatan. Tujuan tersebut diarahkan untuk menghasilkan analisis lokasi yang berbasis data dan dapat menggambarkan karakteristik wilayah sekaligus kondisi potensi pasar secara lebih terstruktur.

2. Metodologi

Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *data mining* untuk mengidentifikasi karakteristik wilayah dan menentukan prioritas lokasi pengembangan toko bangunan di Kota Lubuklinggau. Proses penelitian mengacu pada kerangka *Cross-Industry Standard Process for Data Mining* (CRISP-DM) yang terdiri atas *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*. Kerangka tersebut digunakan untuk memastikan proses analisis dilakukan secara sistematis, mulai dari pemahaman kebutuhan penelitian hingga interpretasi hasil.

Analisis utama dilakukan menggunakan algoritma K-Means *Clustering* untuk mengelompokkan kecamatan berdasarkan kemiripan karakteristik jumlah penduduk, kepadatan penduduk, pertumbuhan penduduk, dan jumlah toko bangunan. Selanjutnya, Rasio Potensi Pasar digunakan sebagai analisis pendukung untuk menggambarkan perbandingan antara potensi permintaan dan ketersediaan toko bangunan. Hasil kedua analisis tersebut kemudian diintegrasikan untuk memperoleh prioritas lokasi toko bangunan.

Objek dan Sumber Data

Objek penelitian adalah delapan kecamatan di Kota Lubuklinggau. Data

yang digunakan terdiri atas data jumlah penduduk, kepadatan penduduk, pertumbuhan penduduk, dan jumlah toko bangunan pada masing-masing kecamatan. Data kependudukan diperoleh dari Badan Pusat Statistik (BPS) Kota Lubuklinggau, sedangkan data jumlah toko bangunan diperoleh melalui observasi menggunakan Google Maps. Data tersebut kemudian digabungkan ke dalam satu dataset berdasarkan wilayah kecamatan.

Data jumlah penduduk, kepadatan penduduk, dan pertumbuhan penduduk digunakan untuk menggambarkan karakteristik demografis wilayah, sedangkan jumlah toko bangunan digunakan untuk merepresentasikan ketersediaan usaha pada masing-masing kecamatan. Seluruh data selanjutnya dipersiapkan untuk proses *clustering* dan analisis Rasio Potensi Pasar.

Variabel Penelitian

Empat variabel numerik yang digunakan dalam proses K-Means *Clustering*, yaitu (1) Jumlah penduduk (L + P) yang menunjukkan jumlah penduduk pada setiap kecamatan, (2) Kepadatan Penduduk (L + P) yang menunjukkan jumlah penduduk dalam satuan luas wilayah, (3) Pertumbuhan Penduduk yang menunjukkan laju pertumbuhan penduduk pada periode pengamatan, (4) Total Toko Bangunan yang menunjukkan jumlah toko bangunan yang tersedia pada setiap kecamatan.

Atribut Nama Kecamatan tidak digunakan dalam perhitungan karena hanya berfungsi sebagai identitas wilayah. Pemilihan keempat variabel tersebut disesuaikan dengan tujuan penelitian untuk mengelompokkan wilayah berdasarkan karakteristik demografi dan ketersediaan toko bangunan.

Tahapan Pengolahan Data

Pengolahan data dilakukan melalui beberapa tahap, yaitu *data cleaning*, seleksi fitur, normalisasi, *clustering*, analisis Rasio Potensi Pasar, integrasi hasil, dan evaluasi.

Data Cleaning

Tahap *data cleaning* dilakukan untuk memastikan dataset terbebas dari data kosong dan data duplikat. Proses pembersihan dilakukan menggunakan fungsi *dropna()* untuk menangani *missing value* dan *drop_duplicates()* untuk menghapus data duplikat. Hasil pemeriksaan menunjukkan bahwa dataset tetap terdiri atas delapan data setelah proses pembersihan, sehingga tidak terdapat *missing value* maupun data duplikat yang perlu dihapus.

Seleksi Fitur

Setelah data dibersihkan, dilakukan seleksi fitur untuk menentukan atribut yang digunakan dalam proses *clustering*. Empat atribut numerik, yaitu jumlah penduduk, kepadatan penduduk, pertumbuhan penduduk, dan total toko bangunan, digunakan sebagai variabel masukan K-Means. Nama kecamatan tidak disertakan dalam proses perhitungan karena hanya merupakan identitas wilayah.

Normalisasi Data

Normalisasi dilakukan karena setiap variabel memiliki rentang nilai yang berbeda. Jumlah penduduk dan jumlah toko bangunan memiliki skala yang berbeda dengan kepadatan serta pertumbuhan penduduk. Kondisi tersebut dapat menyebabkan variabel dengan rentang nilai lebih besar memberikan pengaruh lebih besar pada perhitungan jarak K-Means.

Normalisasi dilakukan menggunakan metode Min-Max dengan rentang 0–1. Persamaan normalisasi dituliskan sebagai berikut:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

dengan x' merupakan nilai hasil normalisasi, x merupakan nilai asli, sedangkan $x_{\min}$ dan $x_{\max}$ masing-masing merupakan nilai minimum dan maksimum pada atribut yang bersangkutan. Proses ini menghasilkan skala yang seragam untuk

seluruh variabel sebelum digunakan dalam *clustering*.

Metode K-Means dan Rasio Potensi Pasar K-Means Clustering

K-Means digunakan untuk mengelompokkan delapan kecamatan berdasarkan kemiripan empat variabel penelitian. Sebelum proses *clustering*, jumlah cluster optimal ditentukan menggunakan Elbow Method dan Silhouette Score. Berdasarkan hasil pengujian, jumlah cluster yang digunakan adalah $K = 3$.

Proses K-Means dilakukan dengan menginisialisasi centroid, menghitung jarak setiap data terhadap centroid, mengalokasikan data ke centroid dengan jarak terdekat, kemudian memperbarui centroid berdasarkan rata-rata anggota pada setiap cluster. Proses tersebut dilakukan secara iteratif hingga centroid mencapai kondisi konvergen.

Jarak antara data dan centroid dihitung menggunakan Euclidean Distance dengan persamaan:

$$d(x_i, c_j) = \sqrt{\sum_{k=1}^p (x_{ik} - c_{jk})^2}$$

dengan $d(x_i, c_j)$ merupakan jarak data ke- i terhadap centroid ke- j , x_{ik} merupakan nilai atribut ke- k pada data ke- i , c_{jk} merupakan nilai atribut ke- k pada centroid ke- j , dan p merupakan jumlah atribut yang digunakan.

Setiap data kemudian ditempatkan pada cluster dengan jarak Euclidean terkecil. Setelah proses pengelompokan, centroid diperbarui menggunakan nilai rata-rata seluruh data yang berada dalam cluster tersebut dengan persamaan:

$$c_{jk} = \frac{1}{n_j} \sum_{x_i \in C_j} x_{ik}$$

dengan c_{jk} merupakan nilai centroid cluster j pada atribut k , C_j merupakan kumpulan data pada cluster j , dan n_j merupakan jumlah data dalam cluster tersebut. Proses perhitungan jarak, pengelompokan, dan pembaruan centroid dilakukan berulang hingga posisi centroid tidak mengalami perubahan.

Rasio Potensi Pasar

Rasio Potensi Pasar digunakan sebagai analisis pendukung untuk menggambarkan perbandingan antara potensi permintaan dan ketersediaan toko bangunan pada setiap kecamatan. Dalam penelitian ini, jumlah penduduk digunakan sebagai indikator potensi permintaan, sedangkan jumlah toko bangunan digunakan sebagai indikator ketersediaan usaha. Analisis rasio tidak digunakan sebagai variabel pembentuk cluster, tetapi dihitung setelah proses *clustering* sebagai informasi tambahan dalam menentukan prioritas lokasi.

Rasio Potensi Pasar dihitung dengan persamaan:

$$RPM_i = \frac{P_i}{T_i}$$

dengan RPM_i merupakan nilai Rasio Potensi Pasar pada kecamatan i , P_i merupakan jumlah penduduk, dan T_i merupakan jumlah toko bangunan pada kecamatan tersebut.

Nilai rasio menunjukkan jumlah penduduk secara relatif yang terdapat untuk setiap toko bangunan. Nilai rasio yang lebih tinggi menunjukkan jumlah penduduk yang lebih besar dibandingkan ketersediaan toko, sedangkan nilai yang lebih rendah menunjukkan ketersediaan toko yang relatif lebih banyak dibandingkan jumlah penduduk. Nilai rasio selanjutnya dikategorikan menjadi Rendah, Sedang, dan Tinggi menggunakan pendekatan *quantile*.

Proses Implementasi dan Integrasi Hasil

Implementasi analisis dilakukan menggunakan bahasa pemrograman Python pada platform Google Colab. Pengolahan data memanfaatkan pustaka Pandas dan NumPy, sedangkan proses *clustering* dan evaluasi menggunakan Scikit-learn. Visualisasi hasil analisis dibuat menggunakan Matplotlib dan Seaborn.

Tahap implementasi dimulai dengan memasukkan dataset, melakukan pemeriksaan dan pembersihan data, memilih fitur, serta melakukan normalisasi. Data yang telah dinormalisasi kemudian

digunakan dalam proses K-Means dengan tiga cluster. Hasil *clustering* berupa label cluster dan centroid selanjutnya digunakan bersama nilai Rasio Potensi Pasar.

Integrasi dilakukan dengan menggabungkan karakteristik hasil cluster dengan kategori Rasio Potensi Pasar. Hasil kombinasi tersebut digunakan untuk mengidentifikasi tingkat prioritas wilayah yang terdiri atas kategori Sangat Strategis, Strategis, dan Cukup Strategis. Dengan pendekatan tersebut, hasil akhir tidak hanya mempertimbangkan kemiripan karakteristik wilayah dari K-Means, tetapi juga mempertimbangkan perbandingan jumlah penduduk dengan ketersediaan toko bangunan.

Metode Evaluasi

Evaluasi hasil *clustering* dilakukan menggunakan Silhouette Score untuk mengetahui tingkat kedekatan data dalam cluster yang sama dan tingkat pemisahan terhadap cluster lainnya. Nilai Silhouette untuk setiap data dihitung dengan persamaan:

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

dengan $a(i)$ merupakan rata-rata jarak data i terhadap data lain dalam cluster yang sama, sedangkan $b(i)$ merupakan rata-rata jarak minimum data i terhadap data pada cluster terdekat lainnya.

Nilai Silhouette berada pada rentang -1 hingga 1. Nilai yang lebih tinggi menunjukkan bahwa suatu data memiliki kedekatan yang lebih baik dengan cluster-nya dibandingkan dengan cluster lainnya. Selain evaluasi numerik, hasil *clustering*

dibandingkan dengan karakteristik jumlah penduduk dan jumlah toko bangunan pada masing-masing kecamatan untuk membantu interpretasi hasil. Dengan demikian, evaluasi tidak hanya melihat nilai metrik *clustering*, tetapi juga mempertimbangkan kesesuaian pola cluster dengan karakteristik data yang dianalisis.

Lingkungan Penelitian dan Implementasi

Seluruh proses pengolahan dan analisis data dilakukan menggunakan Python pada Google Colab. Dataset disimpan dalam format Excel dan diproses menggunakan beberapa pustaka, yaitu Pandas untuk pengolahan data, NumPy untuk operasi numerik, Scikit-learn untuk normalisasi, K-Means, dan evaluasi *clustering*, serta Matplotlib dan Seaborn untuk visualisasi. Penggunaan lingkungan tersebut memungkinkan seluruh tahapan pengolahan data, mulai dari *preprocessing*, normalisasi, *clustering*, perhitungan Rasio Potensi Pasar, evaluasi, hingga visualisasi dilakukan dalam satu alur analisis yang terintegrasi.

3. Hasil

Data Penelitian

Data penelitian terdiri atas data jumlah penduduk, kepadatan penduduk, laju pertumbuhan penduduk, dan jumlah toko bangunan pada delapan kecamatan di Kota Lubuklinggau. Data demografis diperoleh dari BPS Kota Lubuklinggau, sedangkan jumlah toko bangunan diperoleh melalui observasi Google Maps. Data tersebut kemudian digabungkan menjadi satu dataset sebagai input analisis K-Means *Clustering* dan Rasio Potensi Pasar.

Tabel 1. Data Demografis dan Jumlah Toko Bangunan

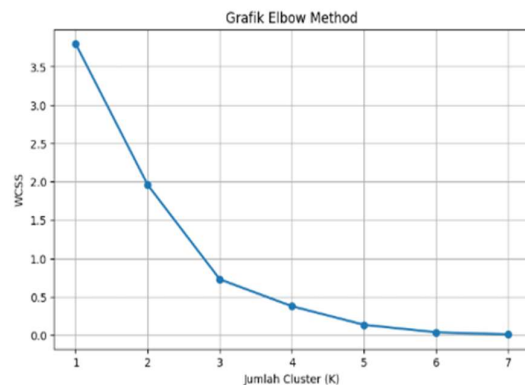
Kecamatan	Jumlah Penduduk	Kepadatan Penduduk (jiwa/km ²)	Pertumbuhan (%)	Jumlah Toko
Lubuklinggau Barat I	41.631	721,01	1,68	15
Lubuklinggau Barat II	21.148	2.731,30	-0,11	27
Lubuklinggau Selatan I	17.896	234,03	2,30	2
Lubuklinggau Selatan II	35.961	894,77	1,57	18
Lubuklinggau Timur I	36.219	2.581,54	1,49	33
Lubuklinggau Timur II	33.583	3.777,62	0,75	34
Lubuk Linnggau Utara I	19.311	138,28	2,01	1
Lubuklinggau Utara II	45.708	1.978,97	2,51	26

Sumber: BPS Kota Lubuklinggau, (2026) dan Google Maps, (2026)

Tabel 1 menunjukkan adanya variasi karakteristik antarwilayah. Jumlah penduduk berkisar antara 17.896 hingga 45.708 jiwa, sedangkan jumlah toko bangunan berkisar antara 1 hingga 34 toko. Perbedaan tersebut menjadi dasar penggunaan K-Means untuk mengelompokkan kecamatan berdasarkan kemiripan karakteristik setelah data dinormalisasi. Data jumlah toko juga digunakan pada tahap berikutnya untuk menghitung Rasio Potensi Pasar.

Penentuan Jumlah Cluster

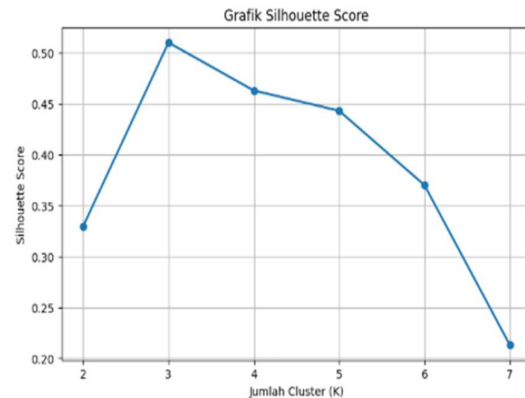
Sebelum proses K-Means dilakukan, jumlah cluster ditentukan menggunakan Elbow Method dan Silhouette Score sebagaimana dijelaskan pada metode penelitian. Hasil Elbow Method menunjukkan bahwa penurunan WCSS terlihat cukup besar hingga $K = 3$, kemudian cenderung melandai pada jumlah cluster berikutnya. Kondisi tersebut menunjukkan bahwa titik siku berada pada $K = 3$.



Gambar 1. Grafik Elbow Method

Hasil Elbow Method kemudian diperkuat dengan evaluasi Silhouette Score. Perhitungan dilakukan pada $K = 2$ sampai

$K = 7$ dan menghasilkan nilai tertinggi sebesar 0,5102 pada $K = 3$. Setelah $K = 3$, nilai Silhouette Score kembali menurun. Dengan demikian, jumlah cluster yang digunakan dalam proses K-Means ditetapkan sebanyak tiga cluster.



Gambar 2. Grafik Silhouette Score untuk Penentuan K

Hasil kedua metode memberikan keputusan yang sama, yaitu $K = 3$. Elbow Method menunjukkan titik siku pada tiga cluster, sedangkan Silhouette Score mencapai nilai tertinggi pada jumlah cluster yang sama. Oleh karena itu, tiga cluster digunakan pada tahap pemodelan K-Means.

Hasil K-Means Clustering

Proses K-Means dilakukan menggunakan data yang telah dinormalisasi dengan MinMaxScaler, kemudian menggunakan tiga cluster dan inisialisasi k-means++. Proses iterasi dilakukan hingga centroid mencapai kondisi konvergen. Hasil implementasi menunjukkan bahwa algoritma mencapai kondisi konvergen setelah 2 iterasi.

Tabel 2. Karakteristik Cluster Berdasarkan Centroid Skala Asli

Cluster	Rata-rata penduduk	Rata-rata kepadatan (jiwa/km ²)	Rata-rata pertumbuhan (%)	Rata-rata toko	Kategori
Cluster 0	18.803,50	186,15	2,15	1,50	Potensi Rendah
Cluster 1	30.316,67	3.030,48	0,71	31,67	Potensi Sedang
Cluster 2	41.100,00	1.198,25	1,92	19,67	Potensi Tinggi

Sumber: Hasil pengolahan data menggunakan K-Means

Hasil *clustering* menunjukkan bahwa Cluster 0 memiliki rata-rata jumlah penduduk dan kepadatan terendah serta jumlah toko bangunan yang sangat sedikit. Cluster ini terdiri atas dua kecamatan dan dikategorikan sebagai Potensi Rendah. Cluster 1 memiliki kepadatan penduduk dan jumlah toko bangunan paling tinggi, tetapi laju pertumbuhannya paling rendah, sehingga dikategorikan sebagai Potensi Sedang. Sementara itu, Cluster 2 memiliki rata-rata jumlah penduduk tertinggi, pertumbuhan penduduk relatif tinggi, dan jumlah toko bangunan lebih rendah dibandingkan Cluster 1, sehingga dikategorikan sebagai Potensi Tinggi.

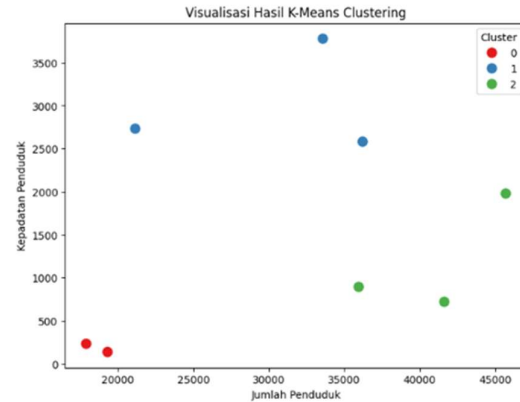
Hasil pengelompokan kecamatan dapat diringkas sebagai berikut:

Tabel 3. Hasil Pengelompokan Kecamatan

Kecamatan	Cluster	Kategori Potensi
Lubuklinggau Selatan I	0	Potensi Rendah
Lubuklinggau Utara I	0	Potensi Rendah
Lubuklinggau Barat II	1	Potensi Sedang
Lubuklinggau Timur I	1	Potensi Sedang
Lubuklinggau Timur II	1	Potensi Sedang
Lubuklinggau Barat I	2	Potensi Tinggi
Lubuklinggau Selatan II	2	Potensi Tinggi
Lubuklinggau Utara II	3	Potensi Tinggi

Sumber: Hasil pengolahan data

Hasil tersebut menunjukkan bahwa delapan kecamatan terbagi ke dalam tiga karakteristik wilayah. Cluster 0 terdiri atas dua kecamatan, sedangkan Cluster 1 dan Cluster 2 masing-masing terdiri atas tiga kecamatan.



Gambar 3. Visualisasi Hasil K-Means Clustering

Visualisasi hasil *clustering* memperlihatkan persebaran kecamatan berdasarkan karakteristik jumlah penduduk dan kepadatan penduduk. Masing-masing kelompok menunjukkan pola persebaran yang berbeda, yang menggambarkan hasil pengelompokan berdasarkan kemiripan karakteristik data.

Evaluasi Hasil Clustering

Kualitas pengelompokan dievaluasi menggunakan Silhouette Score. Hasil pengujian menghasilkan nilai 0,5102. Nilai tersebut menunjukkan adanya pemisahan antarcluster yang cukup jelas serta kemiripan data dalam cluster yang terbentuk.

```
# =====
# SILHOUETTE SCORE
# =====

from sklearn.metrics import silhouette_score

silhouette = silhouette_score(X_normal, df["Cluster"])

print("====60")
print("HASIL EVALUASI CLUSTER")
print("====60")
print(f"Silhouette Score : {silhouette:.4f}")

=====
HASIL EVALUASI CLUSTER
Silhouette Score : 0.5102
```

Gambar 4. Hasil Evaluasi Silhouette Score

Nilai 0,5102 digunakan sebagai indikator kualitas struktur cluster yang terbentuk dan menjadi dasar bahwa hasil K-Means dapat dilanjutkan ke tahap analisis Rasio Potensi Pasar.

Hasil Analisis Rasio Potensi Pasar

Setelah proses *clustering*, dilakukan perhitungan Rasio Potensi Pasar dengan membandingkan jumlah penduduk terhadap jumlah toko bangunan. Rasio ini digunakan sebagai analisis pendukung untuk menggambarkan jumlah penduduk yang secara relatif dilayani oleh setiap toko bangunan. Nilai rasio yang lebih tinggi menunjukkan jumlah penduduk yang lebih besar dibandingkan jumlah toko yang tersedia.

Hasil analisis menunjukkan bahwa nilai rasio memiliki variasi yang cukup besar. Nilai rata-rata rasio adalah 4.703,31, dengan nilai minimum 783,26 dan maksimum 19.311,00.

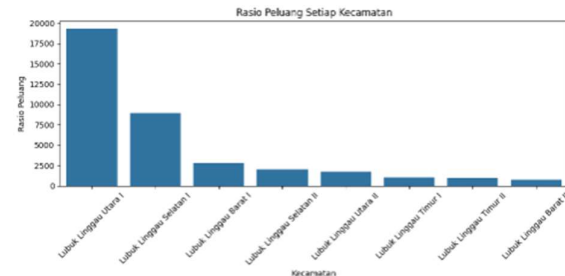
Tabel 4. Hasil Rasio Potensi Pasar

Kecamatan	Rasio Potensi Pasar	Kategori
Lubuklinggau Barat I	2.775,40	Tinggi
Lubuklinggau Barat II	783,26	Rendah
Lubuklinggau Selatan I	8.948,00	Tinggi
Lubuklinggau Selatan II	1.997,83	Sedang
Lubuklinggau Timur I	1.097,55	Sedang
Lubuklinggau Timur II	987,74	Rendah
Lubuklinggau Utara I	19.311,00	Tinggi
Lubuklinggau Utara II	1.758,00	Sedang

Sumber: Hasil pengolahan data

Nilai rasio dihitung dari jumlah penduduk dibagi jumlah toko bangunan; kategori menggunakan metode kuantil. Pengelompokan rasio dilakukan ke dalam tiga kategori, yaitu Rendah, Sedang, dan Tinggi, menggunakan metode kuantil. Hasil skripsi menunjukkan bahwa Lubuklinggau Utara I memiliki rasio tertinggi sebesar 19.311,00, diikuti

Lubuklinggau Selatan I sebesar 8.948,00. Rasio terendah terdapat pada Lubuklinggau Barat II sebesar 783,26.



Gambar 5. Visualisasi Rasio Potensi Pasar per Kecamatan

Visualisasi rasio memperlihatkan perbedaan tingkat potensi pasar antarwilayah. Lubuklinggau Utara I memiliki nilai rasio paling tinggi, sedangkan Lubuklinggau Barat II memiliki nilai paling rendah. Hasil tersebut kemudian digunakan bersama hasil *clustering* dalam tahap integrasi.

Integrasi Hasil Clustering dan Rasio Potensi Pasar

Tahap terakhir menggabungkan kategori hasil K-Means dengan kategori Rasio Potensi Pasar. Integrasi dilakukan untuk memperoleh informasi yang mempertimbangkan karakteristik wilayah sekaligus perbandingan antara jumlah penduduk dan ketersediaan toko bangunan.

Hasil integrasi menghasilkan tiga kategori rekomendasi lokasi, yaitu Sangat Strategis, Strategis, dan Cukup Strategis. Wilayah dengan Cluster Potensi Tinggi dan Rasio Tinggi berada pada kategori Sangat Strategis. Kategori Strategis diberikan pada wilayah yang memiliki salah satu indikator utama yang mendukung, sedangkan Cukup Strategis menunjukkan wilayah yang masih memiliki potensi tetapi peluang pasarnya relatif lebih rendah.

Tabel 5. Hasil Integrasi K-Means *Clustering* dan Rasio Potensi Pasar

Kecamatan	Cluster	Potensi Cluster	Rasio Potensi Pasar	Kategori Rasio	Status Lokasi
Lubuklinggau Barat I	2	Tinggi	2.775,40	Tinggi	Sangat Strategis
Lubuklinggau Barat II	1	Sedang	783,26	Rendah	Cukup Strategis
Lubuklinggau Selatan I	0	Rendah	8.948,00	Tinggi	Strategis

Kecamatan	Cluster	Potensi Cluster	Rasio Potensi Pasar	Kategori Rasio	Status Lokasi
Lubuklinggau Selatan II	2	Tinggi	1.997,83	Sedang	Strategis
Lubuklinggau Timur I	1	Sedang	1.097,55	Sedang	Strategis
Lubuklinggau Timur II	1	Sedang	987,74	Rendah	Cukup Strategis
Lubuklinggau Utara I	0	Rendah	19.311,00	Tinggi	Strategis
Lubuklinggau Utara II	2	Tinggi	1.758,00	Sedang	Strategis

Sumber: Hasil pengolahan data

Hasil integrasi menunjukkan bahwa Lubuklinggau Barat I berada pada Cluster Potensi Tinggi sekaligus memiliki kategori rasio Tinggi sehingga memperoleh kategori Sangat Strategis. Kecamatan lainnya berada pada kategori Strategis atau Cukup Strategis sesuai kombinasi karakteristik cluster dan rasio yang diperoleh.

Hasil integrasi memperlihatkan bahwa K-Means memberikan informasi mengenai karakteristik wilayah, sedangkan Rasio Potensi Pasar memberikan informasi tambahan mengenai perbandingan jumlah penduduk dengan ketersediaan toko bangunan. Dengan demikian, kedua hasil analisis dapat digunakan secara bersama untuk menyusun prioritas lokasi pada tingkat kecamatan.

4. Pembahasan

Interpretasi Hasil K-Means Clustering

Hasil pengelompokan menggunakan K-Means menunjukkan bahwa delapan kecamatan di Kota Lubuklinggau memiliki karakteristik wilayah yang berbeda berdasarkan jumlah penduduk, kepadatan penduduk, laju pertumbuhan penduduk, dan jumlah toko bahan bangunan. Penggunaan keempat variabel tersebut menghasilkan tiga kelompok dengan karakteristik yang berbeda, yaitu Cluster 0 sebagai Potensi Rendah, Cluster 1 sebagai Potensi Sedang, dan Cluster 2 sebagai Potensi Tinggi. Perbedaan tersebut menunjukkan bahwa karakteristik suatu wilayah tidak dapat dilihat hanya dari jumlah penduduk, tetapi perlu mempertimbangkan kepadatan, pertumbuhan penduduk, serta kondisi ketersediaan toko bahan bangunan secara bersamaan. Pendekatan K-Means dengan karakteristik tersebut sejalan dengan

penelitian terdahulu yang menunjukkan bahwa *clustering* dapat digunakan untuk mengelompokkan wilayah berdasarkan kemiripan karakteristik kependudukan [21]

Cluster 0 yang terdiri atas Lubuklinggau Selatan I dan Lubuklinggau Utara I memiliki rata-rata jumlah penduduk sebesar 18.803,50 jiwa dan kepadatan penduduk sebesar 186,15 jiwa/km². Kelompok ini memiliki jumlah toko bahan bangunan yang paling sedikit, yaitu rata-rata 1,50 toko. Meskipun kedua kecamatan memiliki laju pertumbuhan penduduk relatif tinggi, kondisi jumlah penduduk dan kepadatan yang rendah menyebabkan kelompok tersebut dikategorikan sebagai Potensi Rendah. Temuan ini menunjukkan bahwa pertumbuhan penduduk yang relatif tinggi belum secara langsung menjadikan suatu wilayah memiliki karakteristik pasar yang tinggi apabila basis jumlah penduduk dan kepadatannya masih rendah.

Cluster 1 terdiri atas Lubuklinggau Barat II, Lubuklinggau Timur I, dan Lubuklinggau Timur II. Kelompok ini memiliki rata-rata kepadatan penduduk paling tinggi, yaitu 3.030,48 jiwa/km², serta jumlah toko bahan bangunan yang juga paling banyak dibandingkan cluster lainnya. Kondisi tersebut menunjukkan bahwa wilayah dalam kelompok ini memiliki konsentrasi penduduk dan aktivitas usaha yang relatif tinggi. Namun, rata-rata laju pertumbuhan penduduknya hanya sebesar 0,71%, sehingga kelompok ini dikategorikan sebagai Potensi Sedang. Banyaknya toko bahan bangunan pada kelompok ini juga menunjukkan bahwa potensi permintaan berada pada kondisi yang disertai dengan ketersediaan usaha yang relatif tinggi.

Cluster 2 yang terdiri atas Lubuklinggau Barat I, Lubuklinggau Selatan II, dan Lubuklinggau Utara II memiliki rata-rata jumlah penduduk tertinggi, yaitu 41.100 jiwa, dengan rata-rata kepadatan penduduk sebesar 1.198,25 jiwa/km² dan pertumbuhan penduduk sebesar 1,92%. Rata-rata jumlah toko bahan bangunan pada kelompok ini sebesar 19,67 toko, lebih rendah dibandingkan Cluster 1. Kombinasi jumlah penduduk yang tinggi, pertumbuhan penduduk yang relatif baik, dan jumlah toko yang tidak sebanyak Cluster 1 menjadi karakteristik utama yang mendasari kategori Potensi Tinggi. Dengan demikian, Cluster 2 menunjukkan kondisi wilayah yang memiliki basis pasar relatif besar dengan ketersediaan toko yang masih lebih rendah dibandingkan kelompok dengan kepadatan tertinggi.

Pola pengelompokan tersebut memiliki kesesuaian dengan penelitian sebelumnya yang menggunakan K-Means untuk mengidentifikasi karakteristik kependudukan suatu wilayah. Penelitian di Provinsi DKI Jakarta menghasilkan tiga kelompok kepadatan penduduk, sedangkan penelitian di Kota Medan dan Kota Pematangsiantar juga menunjukkan bahwa K-Means dapat membentuk kelompok wilayah berdasarkan karakteristik jumlah atau kepadatan penduduk [22] [23]. Perbedaannya, penelitian ini tidak berhenti pada pengelompokan karakteristik wilayah, tetapi menggunakan hasil cluster sebagai salah satu komponen untuk menganalisis potensi lokasi toko bahan bangunan.

Kualitas dan Karakteristik Cluster

Penentuan jumlah cluster dalam penelitian ini dilakukan menggunakan Elbow Method dan Silhouette Score. Kedua pendekatan menghasilkan jumlah cluster yang sama, yaitu tiga cluster, sedangkan nilai Silhouette Score yang diperoleh sebesar 0,5102. Hasil tersebut menunjukkan bahwa pembagian data menjadi tiga kelompok memberikan struktur pengelompokan yang cukup

terpisah untuk digunakan dalam tahap analisis berikutnya. Penggunaan Elbow Method dan Silhouette Score secara bersamaan juga memberikan dasar evaluasi yang lebih terukur dalam menentukan jumlah kelompok dibandingkan menetapkan jumlah cluster tanpa evaluasi.

Nilai Silhouette Score sebesar 0,5102 perlu dipahami sebagai ukuran kualitas pemisahan cluster pada data yang digunakan, bukan sebagai ukuran akurasi rekomendasi lokasi. Nilai tersebut mendukung penggunaan tiga cluster dalam penelitian karena struktur kelompok yang terbentuk masih dapat dibedakan berdasarkan karakteristik variabel penelitian. Hasil ini kemudian digunakan sebagai dasar untuk menginterpretasikan perbedaan karakteristik antarwilayah sebelum digabungkan dengan analisis rasio potensi pasar.

Interpretasi Rasio Potensi Pasar

Analisis rasio potensi pasar memberikan informasi tambahan mengenai hubungan antara jumlah penduduk sebagai gambaran permintaan dan jumlah toko bahan bangunan sebagai gambaran ketersediaan usaha. Nilai rasio diperoleh dari perbandingan jumlah penduduk dengan jumlah toko pada masing-masing kecamatan. Dengan pendekatan tersebut, nilai rasio yang lebih tinggi menunjukkan bahwa setiap toko secara relatif berada dalam cakupan jumlah penduduk yang lebih besar, sedangkan nilai rasio yang lebih rendah menunjukkan bahwa jumlah toko relatif lebih banyak dibandingkan jumlah penduduk.

Hasil analisis menunjukkan bahwa Lubuklinggau Utara I memiliki rasio tertinggi sebesar 19.311,00, diikuti Lubuklinggau Selatan I sebesar 8.948,00 dan Lubuklinggau Barat I sebesar 2.775,40. Ketiga wilayah tersebut berada pada kategori rasio tinggi. Kondisi tersebut memberikan indikasi adanya jumlah penduduk yang relatif besar dibandingkan jumlah toko bahan bangunan yang teridentifikasi. Sebaliknya, Lubuklinggau

Barat II dan Lubuklinggau Timur II memiliki rasio masing-masing sebesar 783,26 dan 987,74 yang termasuk kategori rendah. Nilai tersebut menunjukkan bahwa jumlah toko relatif lebih banyak dibandingkan jumlah penduduk apabila dilihat melalui indikator rasio yang digunakan.

Rasio tersebut tidak dapat diartikan secara langsung sebagai ukuran permintaan aktual atau jaminan keberhasilan usaha. Nilai rasio hanya menggambarkan hubungan kuantitatif antara jumlah penduduk dan jumlah toko berdasarkan data yang digunakan. Faktor seperti daya beli masyarakat, aksesibilitas, aktivitas pembangunan, lokasi pesaing, dan kondisi geografis belum dimasukkan dalam perhitungan. Oleh karena itu, rasio lebih tepat digunakan sebagai indikator awal untuk melihat peluang pasar yang kemudian dikombinasikan dengan karakteristik wilayah hasil *clustering*.

Integrasi Clustering dan Rasio Potensi Pasar

Integrasi kedua hasil analisis memberikan informasi yang lebih lengkap dibandingkan penggunaan *clustering* atau rasio secara terpisah. *Clustering* memberikan gambaran mengenai karakteristik wilayah berdasarkan beberapa variabel secara simultan, sedangkan rasio memberikan gambaran mengenai hubungan jumlah penduduk dan jumlah toko bahan bangunan. Kombinasi kedua informasi tersebut memungkinkan identifikasi wilayah dengan mempertimbangkan karakteristik internal wilayah sekaligus kondisi relatif antara permintaan dan ketersediaan usaha.

Lubuklinggau Barat I memperoleh kategori Sangat Strategis karena termasuk dalam Cluster 2 yang dikategorikan Potensi Tinggi dan memiliki rasio potensi pasar sebesar 2.775,40 yang termasuk kategori Tinggi. Kondisi tersebut memperlihatkan adanya kombinasi antara karakteristik wilayah dengan jumlah penduduk yang tinggi dan pertumbuhan yang relatif baik

serta rasio penduduk terhadap toko yang tinggi. Posisi ini berbeda dengan Lubuklinggau Selatan I dan Lubuklinggau Utara I. Kedua kecamatan tersebut memiliki rasio tinggi, tetapi berada pada Cluster 0 yang dikategorikan Potensi Rendah. Hal ini menunjukkan bahwa rasio tinggi saja belum cukup untuk menempatkan suatu wilayah pada kategori Sangat Strategis.

Lubuklinggau Selatan II dan Lubuklinggau Utara II berada pada Cluster 2 atau Potensi Tinggi, tetapi memiliki rasio kategori Sedang dengan nilai masing-masing 1.997,83 dan 1.758,00. Kedua wilayah tersebut berada pada kategori Strategis. Kondisi ini memperlihatkan bahwa karakteristik demografis yang mendukung tetap memberikan indikasi potensi meskipun rasio pasar tidak berada pada kategori tinggi. Lubuklinggau Timur I juga termasuk kategori Strategis karena berada pada Cluster 1 dengan rasio kategori Sedang.

Sementara itu, Lubuklinggau Barat II dan Lubuklinggau Timur II berada pada kategori Cukup Strategis. Kedua wilayah tersebut termasuk Cluster 1 atau Potensi Sedang dan memiliki rasio masing-masing dalam kategori Rendah. Hasil ini menunjukkan bahwa wilayah dengan jumlah toko relatif tinggi perlu dipertimbangkan secara berbeda dibandingkan wilayah yang memiliki jumlah toko lebih sedikit. Dalam konteks penelitian ini, kombinasi karakteristik wilayah dan rasio memberikan informasi yang lebih rinci mengenai posisi relatif masing-masing kecamatan.

Temuan tersebut memperkuat penelitian yang menerapkan K-Means untuk mengidentifikasi lokasi strategis usaha coffee shop. Penelitian tersebut menunjukkan bahwa *clustering* dapat digunakan untuk membedakan wilayah berdasarkan karakteristik lokasi usaha. Penelitian lain juga menggunakan K-Means untuk membentuk tipologi kecamatan berdasarkan karakteristik demografi dan ketersediaan fasilitas pelayanan dasar [24].

Berbeda dari kedua pendekatan tersebut, penelitian ini menambahkan rasio jumlah penduduk terhadap jumlah toko bahan bangunan sebagai informasi pendukung dalam menentukan tingkat strategis wilayah. Dengan demikian, kontribusi pendekatan yang digunakan terletak pada penggabungan karakteristik wilayah hasil *clustering* dengan indikator hubungan antara potensi permintaan dan ketersediaan usaha.

Implikasi Hasil dan Keterbatasan Analisis

Hasil penelitian dapat digunakan sebagai referensi awal dalam penyaringan lokasi untuk pengembangan toko bahan bangunan di tingkat kecamatan. Lubuklinggau Barat I menjadi wilayah yang memperoleh kategori Sangat Strategis berdasarkan kombinasi hasil *clustering* dan rasio potensi pasar. Sementara itu, wilayah dengan kategori Strategis maupun Cukup Strategis dapat menjadi alternatif yang perlu dikaji lebih lanjut sesuai kondisi lapangan. Hasil ini dapat membantu mengarahkan analisis lokasi dari pendekatan yang hanya mempertimbangkan satu indikator menuju pertimbangan beberapa karakteristik wilayah secara bersamaan. Pemanfaatan analisis data dan machine learning juga dapat menghasilkan informasi yang mendukung proses pengambilan keputusan berbasis data dalam konteks organisasi dan bisnis [25].

Interpretasi hasil tetap perlu mempertimbangkan keterbatasan data dan metode. Penelitian hanya menggunakan delapan kecamatan di Kota Lubuklinggau dengan variabel jumlah penduduk, kepadatan penduduk, laju pertumbuhan penduduk, dan jumlah toko bahan bangunan. Faktor ekonomi seperti pendapatan dan daya beli, faktor spasial seperti aksesibilitas dan kondisi geografis, serta faktor usaha seperti harga lahan, tingkat persaingan, dan aktivitas pembangunan belum dimasukkan dalam analisis. Oleh sebab itu, kategori strategis yang dihasilkan merupakan hasil

prioritisasi berbasis data pada indikator yang digunakan, bukan keputusan investasi akhir. Pengembangan penelitian selanjutnya dapat menambahkan variabel tersebut dan membandingkan beberapa metode *clustering* untuk memperoleh gambaran lokasi yang lebih komprehensif.

Kesimpulan

Penelitian ini bertujuan menerapkan algoritma K-Means *Clustering* untuk mengelompokkan kecamatan di Kota Lubuklinggau berdasarkan karakteristik demografis dan jumlah toko bahan bangunan, serta mengintegrasikan hasil *clustering* dengan rasio potensi pasar untuk menentukan tingkat prioritas lokasi usaha. Hasil penelitian menunjukkan bahwa delapan kecamatan dapat dikelompokkan menjadi tiga cluster berdasarkan jumlah penduduk, kepadatan penduduk, laju pertumbuhan penduduk, dan jumlah toko bahan bangunan. Cluster 0 dikategorikan sebagai Potensi Rendah, Cluster 1 sebagai Potensi Sedang, dan Cluster 2 sebagai Potensi Tinggi. Penentuan jumlah cluster menghasilkan tiga cluster dengan nilai Silhouette Score sebesar 0,5102.

Integrasi hasil *clustering* dan rasio potensi pasar menghasilkan tiga kategori tingkat strategis, yaitu Sangat Strategis, Strategis, dan Cukup Strategis. Lubuklinggau Barat I berada pada kategori Sangat Strategis karena termasuk dalam cluster Potensi Tinggi dan memiliki rasio potensi pasar kategori Tinggi. Hasil ini menunjukkan bahwa kombinasi karakteristik demografis dan rasio antara jumlah penduduk dengan ketersediaan toko dapat digunakan sebagai dasar penyaringan awal lokasi usaha toko bahan bangunan. Pengembangan penelitian selanjutnya dapat mempertimbangkan variabel tambahan seperti daya beli masyarakat, aktivitas pembangunan, aksesibilitas, kondisi geografis, harga lahan, dan tingkat persaingan usaha, serta membandingkan algoritma *clustering* lainnya untuk menghasilkan analisis yang lebih komprehensif.

Daftar Pustaka

- [1] A. R. Samani, S. Mishra, M. Golias, and D. J.-H. Lee, "What influences the location choice of establishments? An analysis considering establishment types and activities interactions," *Journal of Transport Geography*, vol. 111, Art. no. 103667, 2023, doi: 10.1016/j.jtrangeo.2023.103667
- [2] L. G. Rohman, C. N. Alam, and B. Subaeki, "Implementasi Algoritma K-Means Clustering untuk Identifikasi Lokasi Strategis Coffee Shop," *Bulletin of Computer Science Research*, vol. 5, no. 4, pp. 797–805, 2025, doi: 10.47065/bulletincsr.v5i4.600.
- [3] R. P. Aryanto, A. Nilogiri, and A. E. Wardoyo, "Klasterisasi Jumlah Penduduk Provinsi Jawa Timur Tahun 2021-2023 Menggunakan Algoritma K-Means," *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 9, no. 2, pp. 134–146, 2024, doi: 10.14421/jiska.2024.9.2.134-146.
- [4] N. Oktaviany, N. Suarna, and W. Prihartono, "Implementasi Algoritma K-Means Clustering dalam Mengelompokkan Kepadatan Penduduk di Provinsi DKI Jakarta," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, 2024, doi: 10.36040/jati.v8i1.8241.
- [5] F. Handayanna dan S. Sunarti, "Penerapan Algoritma K-Means Untuk Mengelompokkan Kepadatan Penduduk Di Provinsi DKI Jakarta," *J. Appl. Comput. Sci. Technol.*, vol. 5, no. 1, hlm. 50–55, Mar 2024, doi: 10.52158/jacost.v5i1.477.
- [6] L. Y. Hutabarat, I. Gunawan, I. Purnamasari, M. Safii, dan W. Saputra, "Penerapan Algoritma K-Means Dalam Pengelompokan Jumlah Penduduk Berdasarkan Kelurahan Di Kota Pematangsiantar," *J. Ilmu Komput. Dan Teknol.*, vol. 2, no. 2, hlm. 20–26, Jan 2022, doi: 10.35960/ikomti.v2i2.704.
- [7] P. Marpaung dan R. F. Siahaan, "Penerapan Algoritma K-Means Clustering Untuk Pemetaan Kepadatan Penduduk Berdasarkan Jumlah Penduduk Kota Medan," vol. 5, 2021.
- [8] R. G. Rohman, Cecep Nurul Alam, dan Beki Subaeki, "Implementasi Algoritma K-Means Clustering untuk Identifikasi Lokasi Strategis Coffee Shop," *Bull. Comput. Sci. Res.*, vol. 5, no. 4, hlm. 797–805, Jun 2025, doi: 10.47065/bulletincsr.v5i4.600.
- [9] Y. Radhinal, A. K. Hadid, T. Wahyuningsih, and N. R. Nur, "Tipologi Kecamatan di Kota Palu Berdasarkan Karakteristik Demografi dan Ketersediaan Fasilitas Pelayanan Dasar Menggunakan Algoritma K-Means Clustering," *Jurnal Kesehatan*, vol. 8, no. 11, 2025, doi: 10.56338/jks.v8i11.92693.
- [10] S. Schaffner, S. Schmid, and J. R. H. W, "Location effects: Geo-spatial and socio-demographic determinants of sales dynamics in brick-and-mortar retail stores," *Journal of Retailing and Consumer Services*, vol. 66, Art. no. 102902, 2022, doi: 10.1016/j.jretconser.2021.102902.
- [11] C. Reed, T. E. Yu, and D. Hughes, "Evaluating the factors influencing the location strategies of specialty grocers versus traditional supermarkets in the United States," *Applied Geography*, vol. 158, Art. no. 103034, 2023, doi: 10.1016/j.apgeog.2023.103034.
- [12] J. Lu, X. Zheng, E. Nervino, Y. Li, Z. Xu, and Y. Xu, "Retail store location screening: A machine learning-based approach," *Journal of Retailing and Consumer Services*, vol. 77, Art. no. 103620, 2024, doi: 10.1016/j.jretconser.2023.103620.
- [13] J. R. Graham and D. P. H. McKenzie, "Empirical methods in business location research," *Journal of Urban*

- Technology, vol. 29, no. 3, pp. 344–361, 2022, doi: 10.1080/21681376.2021.1976261
- [14] S. Datta and K. Sudhir, “The agglomeration-differentiation tradeoff in spatial location choice,” *Customer Needs and Solutions*, vol. 10, no. 1, 2023, doi: 10.1007/s40547-023-00135-w
- [15] T. Formanek and O. Sokol, “Location effects: Geo-spatial and socio-demographic determinants of sales dynamics in brick-and-mortar retail stores,” *Journal of Retailing and Consumer Services*, vol. 66, Art. no. 102902, 2022, doi: 10.1016/j.jretconser.2021.102902
- [16] N. Nurhayah, N. Suarna, and W. Prihartono, “Analisis jumlah penduduk menggunakan algoritma K-Means berdasarkan kabupaten/kota di Indonesia,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, pp. 894–901, 2024, doi: 10.36040/jati.v8i1.8863
- [17] R. P. Aryanto, A. Nilogiri, and A. E. Wardoyo, “Klasterisasi jumlah penduduk Provinsi Jawa Timur tahun 2021–2023 menggunakan algoritma K-Means,” *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 9, no. 2, pp. 134–146, 2024, doi: 10.14421/jiska.2024.9.2.134-146
- [18] D. Gultom, I. Gunawan, I. Purnamasari, S. R. Andani, and Z. A. Siregar, “Penerapan algoritma K-Means dalam pengelompokan kepadatan penduduk menurut kecamatan di Kabupaten Simalungun,” *Terapan Informatika Nusantara*, vol. 2, no. 10, pp. 622–628, 2022, doi: 10.47065/tin.v2i10.1375
- [19] D. Erdiansyah, I. N. Abdullah, and A. J. Tallo, “Determination of potential business locations using data mining clustering,” *Jurnal Pilar Nusa Mandiri*, vol. 21, no. 1, pp. 72–81, 2025, doi: 10.33480/pilar.v21i1.6295
- [20] V. Agustine, I. F. Ashari, and R. C. T. H. Permana, “Clustering of regions in Lampung Province based on social and economic aspects using the K-Means algorithm with PCA optimization,” *Jurnal Ekonomi & Studi Pembangunan*, vol. 26, no. 2, 2025, doi: 10.18196/jesp.v26i2.26212
- [21] A. E. P. Haryanto and M. U. Yanuar, “Metode K-Means Clustering untuk pengelompokan kabupaten/kota dalam upaya pengendalian tingkat inflasi di Pulau Jawa dan Sumatera,” *Government and Statistics*, vol. 1, pp. 29–42, 2022, doi: 10.21787/govstat.1.2022.29-42
- [22] F. S. Darusman and B. W. Trenggono, “Population density cluster analysis in DKI Jakarta Province using K-Means algorithm,” *Journal of Information Systems and Informatics*, vol. 4, no. 4, pp. 777–789, 2022
- [23] N. Napitupulu, S. B. Purba, D. T. Silalahi, A. A. Sinamo, and I. Gunawan, “Clustering data kepadatan penduduk Kota Pematangsiantar (per km²) menurut kecamatan menggunakan metode K-Means clustering,” *Jurnal Riset Sistem Informasi dan Aplikasi Komputer (JRSIKOM)*, vol. 1, no. 2, pp. 51–58, 2025, doi: 10.180997/jrsikom.v1i2.24
- [24] S. Han, X. Jia, X. Chen, S. Gupta, A. Kumar, and Z. Lin, “Search well and be wise: A machine learning approach to search for a profitable location,” *Journal of Business Research*, vol. 144, pp. 416–427, 2022, doi: 10.1016/j.jbusres.2022.01.049
- [25] A. A. E. Purba and M. I. P. Nasution, “Business Intelligence dalam Era Ekonomi Digital: Kajian Sistematis atas Strategi dan Implementasi,” *Journal Artificial*, vol. 3, no. 2, pp. 161–171, Aug. 2025, doi: 10.54065/artificial.753