



Segmentasi Jenis Hunian di Wilayah Bandung Menggunakan Algoritma Clustering K-Means

Ridhaka Gina Amalia¹, Risma Auliya Salsabilla², Alustina Suci Manah³,
Alma Alifya Zafira⁴, Siti Zulhi Nirma Saidah⁵, Ratna Mufidah⁶

¹Universitas Singaperbangsa Karawang

Received: 19 Januari 2026
Revised: 29 Januari 2026
Accepted: 9 Februari 2026

Abstract

The research on property segmentation in Bandung City is motivated by the rapid growth of the property sector, accompanied by market complexity due to variations in price, specifications, and location, which makes it difficult for both buyers and developers to comprehensively understand market characteristics and make informed decisions. To address this issue, this study employs a Knowledge Discovery in Database (KDD) approach, which includes data selection, cleaning, transformation, processing, and evaluation stages, by applying the K-Means Clustering algorithm to a dataset of 57,190 property records from August 2024 to January 2025. The dataset covers variables such as price, installment, land and building area, number of rooms, and location. The determination of the optimal number of clusters was validated using the Elbow Method and Silhouette Score, resulting in two main segments with a Silhouette Score of 0.335: Cluster 0 represents the regular housing segment with 32,424 units and an average price of IDR 1.23 billion, reflecting the middle and more affordable market, while Cluster 1 represents the exclusive housing segment with 24,766 units and an average price of IDR 2.97 billion, representing the upper segment with higher specifications. The results indicate that the K-Means algorithm is effective in mapping large and complex property data, providing practical benefits for buyers in selecting homes according to their preferences and purchasing power, and for developers as a basis for formulating more targeted marketing strategies based on market segmentation.

Keywords: Housing Segmentation, Clustering, K-Means, Property Data Analysis, Bandung

(*) Corresponding Author:

¹2210631170099@student.unsika.ac.id,

²2210631170100@student.unsika.ac.id,

³2210631170006@student.unsika.ac.id,

⁴2210631170005@student.unsika.ac.id,

⁵2210631170103@student.unsika.ac.id, ⁶ratna.mufidah@cs.unsika.ac.id

How to Cite: Amalia, R., Salsabilla, R., Manah, A., Zafira, A., Saidah, S., & Mufidah, R. (2026).

Segmentasi Jenis Hunian di Wilayah Bandung Menggunakan Algoritma Clustering K-Means. *Jurnal*

Ilmiah Wahana Pendidikan, 12(7.D), 217-228. Retrieved from

<https://jurnal.peneliti.net/index.php/JIWP/article/view/14767>.

PENDAHULUAN

Perkembangan teknologi internet memberikan dampak yang sangat signifikan bagi dunia bisnis, termasuk dalam sektor properti. Internet memungkinkan pengelolaan dan analisis data dilakukan secara lebih cepat dan luas, sehingga mendukung pengambilan keputusan yang lebih akurat. Di Kota Bandung, yang dikenal sebagai pusat pendidikan dan pariwisata, pertumbuhan penduduk serta mobilitas masyarakat mendorong peningkatan kebutuhan hunian. Kondisi ini membuat pasar properti di Bandung menjadi sangat dinamis dan beragam.

Kompleksitas pasar tersebut dipengaruhi oleh berbagai faktor seperti lokasi, luas tanah dan bangunan, serta jumlah kamar, sehingga harga properti tidak dapat diprediksi secara sederhana hanya berdasarkan rata-rata pasar. Oleh karena itu, dibutuhkan analisis data yang lebih mendalam untuk memahami pola dan karakteristik properti secara komprehensif.

Namun, proses analisis dan segmentasi data properti secara manual memiliki banyak keterbatasan. Cara ini cenderung memakan waktu, membutuhkan sumber daya yang besar, serta berisiko tinggi terhadap kesalahan. Segmentasi data menjadi sangat penting karena berfungsi untuk mengelompokkan data ke dalam kategori-kategori kecil yang homogen, sehingga mempermudah proses analisis dan pengambilan keputusan. Dalam konteks ini, metode clustering sebagai bagian dari teknik data mining menjadi solusi yang efektif, karena mampu membagi data ke dalam beberapa kelompok (cluster) berdasarkan kemiripan karakteristik yang dimiliki.

Salah satu algoritma clustering yang banyak digunakan adalah K-Means. Algoritma ini bekerja dengan cara meminimalkan perbedaan data di dalam satu kelompok dan memaksimalkan perbedaan antar kelompok yang lain. K-Means memiliki sejumlah keunggulan, antara lain efisiensinya dalam menangani dataset berukuran besar, kemudahan implementasi, serta fleksibilitasnya dalam mengolah data numerik. Karakteristik tersebut menjadikan K-Means sangat cocok digunakan untuk melakukan segmentasi pasar properti yang kompleks dan beragam.

Penelitian ini menggunakan dataset publik yang diperoleh dari Kaggle mengenai data rumah di wilayah Bandung. Variabel yang dianalisis meliputi harga rumah, luas tanah dan bangunan, jumlah kamar tidur dan kamar mandi, ketersediaan garasi, hingga skema cicilan. Berbeda dengan penelitian terdahulu yang umumnya hanya berfokus pada pengelompokan harga rumah, penelitian ini memperluas cakupan analisis dengan memasukkan jenis hunian dan skema cicilan. Dengan demikian, hasil segmentasi yang dihasilkan menjadi lebih kaya informasi dan relevan dengan kondisi pasar yang sebenarnya.

Hasil dari penelitian ini diharapkan memberikan manfaat bagi berbagai pihak. Bagi calon pembeli, segmentasi yang dihasilkan dapat membantu memahami karakteristik rumah yang tersedia sesuai dengan kebutuhan dan kemampuan finansial. Bagi pengembang properti, penelitian ini dapat menjadi acuan dalam melihat segmentasi pasar secara lebih akurat untuk menentukan strategi pengembangan dan pemasaran. Sementara itu, bagi pemerintah, hasil penelitian ini dapat digunakan sebagai referensi dalam merumuskan kebijakan perumahan yang lebih tepat sasaran dan efektif.

METODE

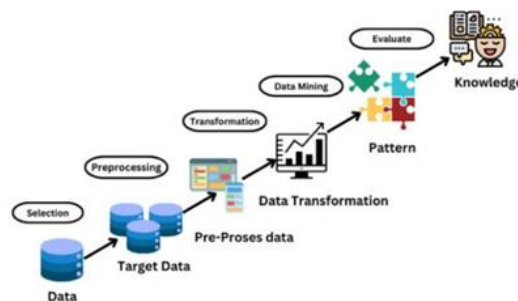
1.1 Metode Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini dilakukan dengan menggunakan metode pencarian data sekunder yang bersumber dari platform terbuka (open source), yaitu Kaggle. Data yang digunakan merupakan hasil ekstraksi (web scraping) dari situs web properti yang memuat informasi penjualan rumah di wilayah Bandung. Secara geografis, data berfokus pada transaksi properti di Kota Bandung, dengan cakupan waktu pengamatan mulai dari Agustus 2024

hingga Januari 2025, sehingga dapat memberikan gambaran kondisi pasar properti dalam periode tersebut secara komprehensif.

1.2 Tahap Penelitian

Metode Knowledge Discovery in Database (KDD) dalam analisis segmentasi hunian di Bandung dimulai dengan data selection, yaitu pemilihan data yang relevan dari basis data yang tersedia. Selanjutnya dilakukan data preprocessing untuk membersihkan data dari nilai yang hilang, data tidak konsisten, dan noise agar kualitas data lebih baik. Data yang telah bersih kemudian melalui tahap data transformation untuk disesuaikan ke dalam format yang tepat, seperti normalisasi atau agregasi. Tahap utama adalah data mining, di mana metode atau algoritma diterapkan untuk menemukan pola dan menentukan segmentasi jenis hunian. Tahap akhir adalah evaluation, yaitu menafsirkan dan menilai hasil segmentasi agar sesuai dengan tujuan penelitian. Alur dari tahapan KDD dapat dilihat pada Gambar 1.



Gambar 1. Tahapan KDD

2.2.1 Data Selection

Tahapan seleksi data dalam penelitian ini diawali dengan pemilihan sumber data yang terpercaya, yaitu dataset dari Kaggle yang diperoleh melalui proses web scraping dari situs penjualan rumah di wilayah Bandung. Data tersebut kemudian diseleksi berdasarkan tiga kriteria utama, meliputi kelengkapan atribut, keterbaruan (recency) data, serta relevansinya terhadap karakteristik hunian di Kota Bandung. Selanjutnya, dilakukan proses eliminasi dengan menghapus data yang tidak relevan maupun yang mengandung informasi redundan atau berulang. Seluruh tahapan ini bertujuan untuk memastikan bahwa data yang digunakan memiliki kualitas tinggi sehingga mampu menunjang akurasi dalam proses segmentasi hunian.

2.2.2 Data Preprocessing

Tahap pra-pemrosesan dan analisis data meliputi pembersihan data untuk menangani missing values, outlier, dan menyeragamkan format atribut, dilanjutkan dengan integrasi berbagai sumber data menjadi satu dataset yang koheren. Selanjutnya, dilakukan analisis data eksploratif (EDA) untuk memahami karakteristik, pola, dan anomali dalam data sebelum analisis lanjutan.

2.2.3 Data Transformation

Tahap transformasi data dilakukan dengan menormalisasi atribut numerik agar memiliki skala yang sebanding sehingga tidak mendominasi proses klusterisasi, serta melakukan encoding pada atribut kategorikal agar dapat diproses secara numerik oleh algoritma. Selain itu, jika diperlukan, dilakukan reduksi

dimensi menggunakan metode seperti Principal Component Analysis (PCA) untuk menyederhanakan komputasi tanpa menghilangkan informasi penting.

2.2.4 Data Mining

Alur pemodelan dilakukan dengan menerapkan algoritma K-Means untuk membentuk segmen hunian yang homogen berdasarkan kedekatan jarak Euclidean terhadap pusat kluster (centroid). Penentuan jumlah kluster optimal (k) dilakukan secara objektif menggunakan Elbow Method, dengan mengamati titik penurunan varians (WCSS) yang mulai melambat, serta Silhouette Score untuk mengevaluasi konsistensi dalam kluster dan tingkat pemisahan antar kluster. K-Means dipilih sebagai model karena memiliki skalabilitas tinggi pada dataset berukuran besar dan mampu menghasilkan segmentasi yang jelas serta mudah diinterpretasikan untuk mendukung pengambilan kebijakan maupun strategi pemasaran.

2.2.5 Evaluation

Tahap analisis dan evaluasi dilakukan dengan mengidentifikasi karakteristik unik pada setiap kluster hunian untuk memahami perbedaan fitur antar segmen, kemudian menilai validitas internal model menggunakan Silhouette Coefficient guna memastikan kualitas pengelompokan dan minimnya tumpang tindih antar kluster. Hasil analisis selanjutnya diinterpretasikan menjadi rekomendasi strategis yang dapat dimanfaatkan oleh pemerintah daerah sebagai dasar perencanaan tata ruang dan zonasi, serta oleh pengembang sebagai acuan dalam merumuskan strategi pembangunan hunian yang sesuai dengan kebutuhan pasar.

PEMBAHASAN

3.1 Karakteristik Dataset

Statistik deskriptif dataset disajikan untuk memberikan gambaran sebaran setiap fitur utama melalui parameter minimum, maksimum, rata-rata (mean), dan standar deviasi. Dataset penelitian ini mencakup 57.190 properti hunian di wilayah Bandung yang dikumpulkan pada periode Agustus 2024 hingga Januari 2025. Setelah melalui tahap pra-pemrosesan, seluruh 57.190 data dinyatakan valid tanpa adanya data yang terbuang, dengan delapan fitur utama yang siap digunakan pada tahap klusterisasi disajikan pada Gambar 2.

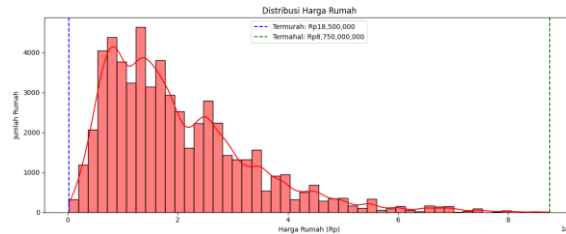
---STATISTICS PARAMETER---					
	harga	kamar_tidur	kamar_mandi	garasi	luas_tanah \
count	5.719000e+04	57190.000000	57190.000000	57190.000000	57190.000000
mean	1.982342e+09	3.179227	2.258699	1.013027	143.615020
min	1.850000e+07	2.000000	1.000000	0.000000	1.000000
25%	9.800000e+08	3.000000	2.000000	1.000000	96.000000
50%	1.650000e+09	3.000000	2.000000	1.000000	126.000000
75%	2.650000e+09	4.000000	3.000000	1.000000	180.000000
max	8.750000e+09	5.000000	4.000000	3.000000	350.000000
std	1.314392e+09	0.903992	0.881830	0.739369	65.410443

	luas_bangunan	waktu_penjualan
count	57190.000000	57190
mean	143.087218	2024-10-15 17:06:16.807134208
min	1.000000	2024-08-01 00:00:00
25%	86.000000	2024-09-01 00:00:00
50%	130.000000	2024-10-01 00:00:00
75%	191.750000	2024-12-01 00:00:00
max	367.000000	2025-01-01 00:00:00
std	74.388380	NaN

Gambar 2. Statistik Deskriptif Dataset

Distribusi harga properti dalam penelitian ini dapat dilihat pada Gambar 3 distribusi harga properti dalam dataset menunjukkan rentang yang sangat lebar, mulai dari Rp 18,5 juta hingga Rp 8,75 miliar, dengan nilai rata-rata sebesar Rp 1,98 miliar dan standar deviasi Rp 1,31 miliar. Besarnya standar deviasi mengindikasikan tingkat heterogenitas yang tinggi, yang mencerminkan

keberagaman tipe hunian dengan karakteristik yang sangat kontras. Selain itu, dataset ini memiliki cakupan geografis yang luas dengan sebaran properti di 166 wilayah berbeda di area Bandung dan sekitarnya.

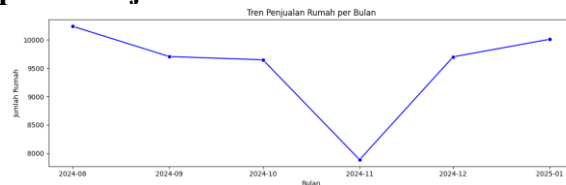


Gambar 3. Histogram Distribusi Harga Rumah

3.2 Hasil *Exploratory Data Analysis* (EDA)

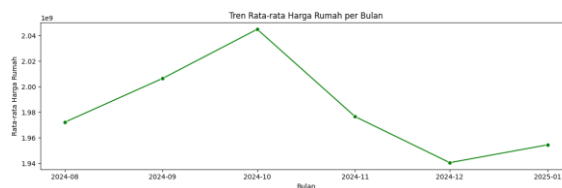
Analisis eksploratori data mengungkap beberapa pola menarik dalam dataset properti Bandung:

3.2.1 Tren Temporal Penjualan



Gambar 4. *Line Chart* Tren Penjualan Bulanan

Gambar 4 menunjukkan tren penjualan rumah yang berfluktuasi, analisis tren penjualan menunjukkan adanya fluktuasi yang signifikan, dengan penurunan tajam pada November 2024 yang mengindikasikan pengaruh musiman serta faktor eksternal seperti kebijakan pemerintah, kondisi ekonomi makro, dan musim liburan. Meskipun demikian, pasar mengalami pemulihan pada Januari 2025, yang mencerminkan tingkat resiliensi serta potensi pertumbuhan positif ke depan. Sementara itu, tren harga properti juga bersifat fluktuatif, dengan puncak tertinggi terjadi pada Oktober 2024, diikuti penurunan tajam pada periode November hingga Desember 2024, sebelum kembali menunjukkan penguatan pada Januari 2025.



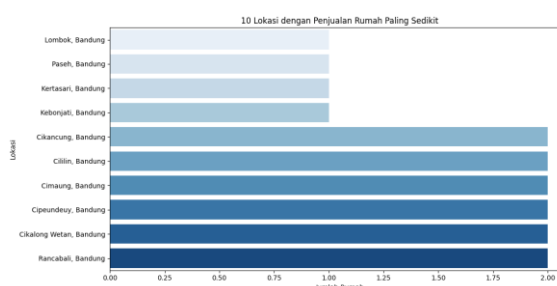
Gambar 5. *Line Chart* Tren Rata-rata Harga Bulanan

3.2.2 Analisis Geografis

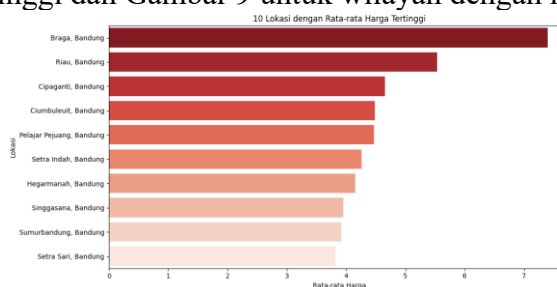
Analisis distribusi geografis properti ditampilkan dalam serangkaian visualisasi. Gambar 6 menunjukkan 10 wilayah dengan volume penjualan tertinggi didominasi oleh Kota Baru Parahyangan.



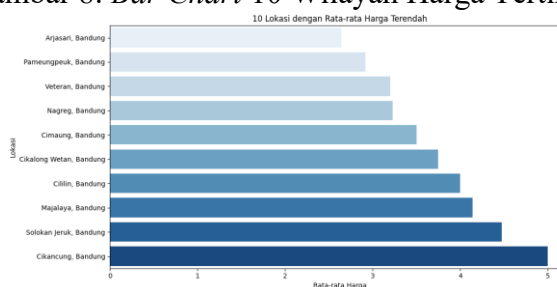
Gambar 6: *Bar Chart* 10 Wilayah Penjualan Tertinggi
Sebaliknya, Gambar 7 mempresentasikan wilayah dengan *volume* penjualan terendah selama periode penelitian.



Gambar 7: *Bar Chart* 10 Wilayah Penjualan Terendah
Variasi harga berdasarkan lokasi dapat dilihat pada Gambar 8 untuk wilayah dengan harga rata-rata tertinggi dan Gambar 9 untuk wilayah dengan harga terendah.



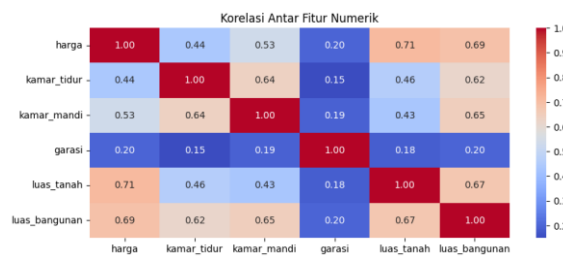
Gambar 8: *Bar Chart* 10 Wilayah Harga Tertinggi



Gambar 9: *Bar Chart* 10 Wilayah Harga Terendah

3.2.3 Analisis Korelasi

Gambar 10 menunjukkan *heatmap* korelasi antar variabel dalam dataset yang mengungkap hubungan signifikan antara berbagai atribut properti.



Gambar 10: *Heatmap* Korelasi Atribut Properti

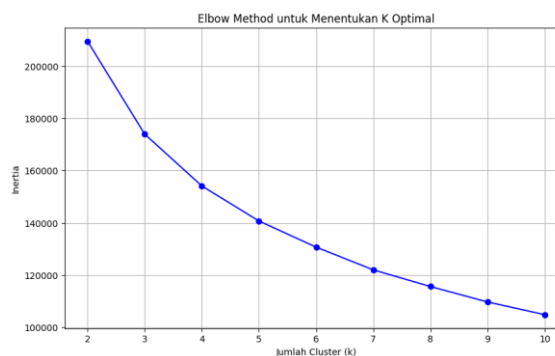
Analisis korelasi menunjukkan bahwa luas tanah ($r = 0,71$) dan luas bangunan ($r = 0,69$) merupakan faktor determinan utama harga properti, yang menegaskan bahwa dimensi fisik menjadi penentu paling dominan dalam pembentukan nilai hunian. Temuan ini dapat dimanfaatkan pengembang untuk mengoptimalkan rasio luas lahan dan bangunan terhadap harga jual agar produk tetap kompetitif. Selain itu, jumlah kamar mandi memiliki korelasi moderat terhadap harga ($r = 0,53$), yang mengindikasikan perannya sebagai pembeda penting pada segmen hunian mewah dan dapat digunakan sebagai fitur kunci dalam strategi positioning produk untuk menyasar pasar premium.

3.3 Penentuan Jumlah *Cluster* Optimal

Penentuan jumlah *cluster* optimal dilakukan menggunakan dua metode evaluasi:

3.3.1 *Elbow Method*

Hasil analisis Elbow Method menunjukkan penurunan nilai Within-Cluster Sum of Squares (WCSS) yang sangat signifikan hingga $k = 2$, kemudian grafik cenderung melandai setelah titik tersebut. Pola ini mengindikasikan bahwa penambahan jumlah kluster di atas $k = 2$ tidak lagi memberikan penurunan varians internal yang berarti. Oleh karena itu, berdasarkan karakteristik grafik pada Gambar 11, jumlah kluster yang paling optimal untuk dataset ini ditetapkan sebanyak dua kluster.



Gambar 11: *Elbow Method Plot (WCSS vs Number of Clusters)*

3.3.2 *Silhouette Score*

Hasil evaluasi menggunakan Silhouette Score menunjukkan bahwa jumlah kluster $k = 2$ menghasilkan nilai tertinggi, yaitu sebesar 0,335, dibandingkan dengan konfigurasi kluster lainnya. Nilai ini mengonfirmasi bahwa pembagian data ke dalam dua kluster memberikan struktur yang paling solid dan konsisten. Selain itu, pada $k = 2$ terlihat bahwa objek dalam setiap kluster memiliki tingkat kemiripan yang kuat, sementara pemisahan antar kluster berlangsung paling jelas dengan tingkat tumpang tindih yang minimal.

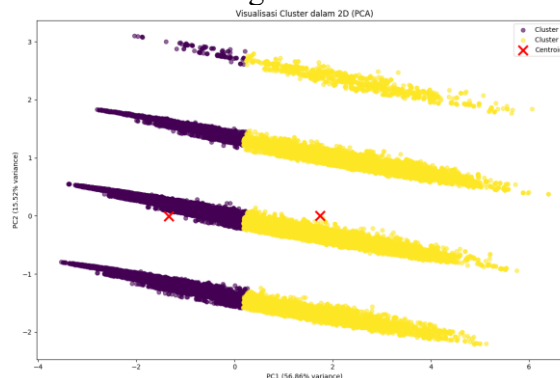
Silhouette Scores:

k	silhouette_score
0	2 0.335182
1	3 0.241134
2	4 0.244508
3	5 0.218587
4	6 0.216572
5	7 0.218488
6	8 0.220327
7	9 0.220167
8	10 0.218065

Gambar 11: *Silhouette Score* untuk Berbagai Nilai K

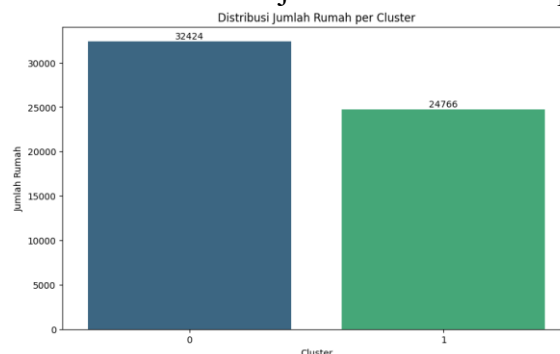
3.4 Hasil *Clustering K-Means*

Berdasarkan hasil analisis, algoritma *K-Means* diterapkan dengan $k=2$ dan divisualisasikan pada Gambar 13. *Plot* menunjukkan dua *cluster* yang terpisah jelas berdasarkan proyeksi data dalam ruang 2D hasil reduksi dimensi *PCA*.



Gambar 13: *Plot Hasil Clustering*

Distribusi jumlah rumah hasil klasterisasi menunjukkan pembagian data ke dalam dua kelompok dengan proporsi yang relatif seimbang. Klaster 0 mendominasi dengan 32.424 rumah atau sekitar 56,7% dari total data, sementara Klaster 1 mencakup 24.766 rumah atau 43,3%. Meskipun Klaster 0 memiliki jumlah lebih besar, kedua klaster tetap bersifat representatif dan memiliki ukuran yang memadai untuk dianalisis lebih lanjut karakteristik serta perbedaannya.

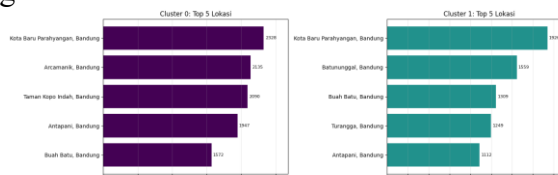


Gambar 14: *Bar Chart Distribusi Cluster*

3.4.1 Karakteristik Lokasi

Lokasi teratas per *cluster* ditunjukkan pada Gambar 15. Distribusi lokasi hasil klasterisasi menunjukkan bahwa Kota Baru Parahyangan mendominasi pada

kedua klaster, yang mengindikasikan kawasan tersebut sebagai area mixed-development yang mampu menjangkau berbagai segmen pasar properti. Klaster 0 cenderung tersebar di wilayah pengembangan baru seperti Arcamanik dan Taman Kopo Indah, sementara Klaster 1 lebih terkonsentrasi di kawasan yang telah mapan seperti Batununggal dan Buah Batu. Perbedaan pola lokasi sekunder ini mencerminkan adanya segregasi spasial antara area pertumbuhan baru dan kawasan pusat kegiatan yang relatif stabil.

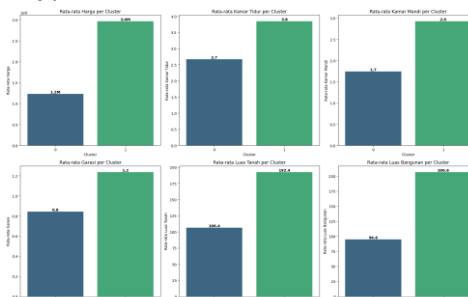


Gambar 15: Bar Chart 5 Lokasi Teratas per Cluster

3.4.2 Profil Cluster

Berdasarkan perbandingan spesifikasi pada Gambar 16, Cluster 1 secara konsisten memiliki nilai rata-rata yang lebih tinggi pada seluruh parameter utama dibandingkan Cluster 0. Perbedaan yang paling signifikan terlihat pada harga properti, jumlah kamar tidur dan kamar mandi, luas tanah, serta luas bangunan. Dengan karakteristik tersebut, Cluster 1 dapat diidentifikasi sebagai segmen hunian kelas atas atau premium, sementara Cluster 0 merepresentasikan segmen hunian menengah hingga lebih terjangkau. Maka dapat disimpulkan profil masing-masing cluster yaitu:

- Cluster 0 (Rumah Reguler):** Jumlah properti 32.424 unit, rata-rata harga Rp 1,2 M. Rata-rata 2,7 kamar tidur, 1,7 kamar mandi, 0,8 garasi, luas tanah 106 m², dan luas bangunan 95 m².
- Cluster 1 (Rumah Eksklusif):** Jumlah properti 24.766 unit, rata-rata harga Rp 3 M. Rata-rata 3,8 kamar tidur, 2,9 kamar mandi, 1,2 garasi, luas tanah 192 m², dan luas bangunan 207 m².



Gambar 16: Bar Chart Karakteristik Cluster

Analisis kesenjangan harga antar klaster menunjukkan perbedaan yang sangat signifikan, di mana rata-rata harga properti pada Cluster 1 mencapai sekitar 2,4 kali lebih tinggi dibandingkan Cluster 0. Secara nilai, Cluster 0 memiliki rata-rata harga sebesar Rp 1,23 miliar, sedangkan Cluster 1 mencapai Rp 2,97 miliar. Kesenjangan yang mencolok ini menegaskan bahwa pasar properti di Bandung telah tersegmentasi secara jelas, sekaligus mencerminkan pembagian target konsumen berdasarkan perbedaan kemampuan ekonomi atau daya beli.

3.5 Evaluasi Model

Evaluasi model K-Means menunjukkan bahwa penggunaan dua klaster ($k = 2$) menghasilkan Silhouette Score sebesar 0,335, yang mengindikasikan kualitas pemisahan klaster yang cukup baik dan valid. Nilai inertia sebesar 209.517,63 dinilai proporsional mengingat besarnya jumlah data (57.190 record) serta skala fitur yang digunakan. Selain itu, uji stabilitas melalui sepuluh kali pengulangan menghasilkan rata-rata skor yang tetap konsisten pada nilai 0,335 dengan deviasi yang sangat kecil, sehingga membuktikan bahwa model memiliki tingkat presisi dan stabilitas yang tinggi, serta memperkuat validitas temuan dalam penelitian ini.

3.6 Rekomendasi Strategis

3.6.1 Untuk Cluster 0 (Rumah Reguler)

Strategi pemasaran untuk Cluster 0 yang merepresentasikan segmen hunian terjangkau difokuskan pada pasangan muda, keluarga kecil, dan pembeli rumah pertama. Pendekatan pemasaran dilakukan melalui optimalisasi media sosial seperti Instagram, TikTok, dan Facebook Ads guna menjangkau audiens secara efektif, didukung oleh strategi penjualan berupa penawaran KPR dengan DP rendah serta insentif tambahan seperti paket furnish ready atau bonus pembangunan garasi. Proposisi nilai yang ditawarkan menekankan pesan “rumah impian dengan harga terjangkau”, sementara pengembangan produk diarahkan pada desain bangunan yang kompak dan efisien serta penyediaan fasilitas komunal yang mendukung interaksi sosial dan gaya hidup penghuni.

3.6.2 Untuk Cluster 1 (Rumah Eksklusif)

Strategi pemasaran untuk Cluster 1 yang merepresentasikan segmen hunian premium difokuskan pada konsumen dengan daya beli tinggi, seperti eksekutif, profesional senior, dan investor properti. Pendekatan penjualan dilakukan secara eksklusif melalui personal selling dan layanan private showing untuk memberikan pengalaman yang lebih personal. Kanal pemasaran diarahkan ke media premium seperti majalah gaya hidup serta partisipasi dalam acara-acara eksklusif, didukung kolaborasi strategis dengan penyedia layanan wealth management dan penasihat keuangan. Proposisi nilai yang ditawarkan menekankan keunggulan fitur mewah dan lokasi strategis, dengan pengembangan produk berupa fasilitas kelas atas seperti teknologi smart home, sistem keamanan terpadu, dan club house.

3.7 Implikasi untuk Stakeholder

3.7.1 Bagi Developer

Hasil penelitian ini memberikan implikasi strategis yang signifikan bagi pengembang properti, khususnya dalam mendukung pengembangan produk yang lebih tepat sasaran berdasarkan pemahaman dua segmen pasar utama, yaitu Cluster 0 dan Cluster 1. Dengan segmentasi yang jelas, pengembang dapat mengoptimalkan alokasi sumber daya seperti modal, tenaga kerja, dan lahan secara lebih efisien sesuai dengan pola permintaan masing-masing segmen. Selain itu, temuan ini memungkinkan penyusunan strategi penetapan harga yang lebih akurat dan kompetitif dengan mempertimbangkan karakteristik teknis serta daya beli konsumen. Secara keseluruhan, pendekatan berbasis data ini membantu meningkatkan efisiensi bisnis dan meminimalkan risiko ketidaksesuaian produk dengan kebutuhan pasar.

3.7.2 Bagi Calon Pembeli

Hasil penelitian memberikan panduan dalam menentukan pilihan hunian berdasarkan *budget* dan preferensi, serta pemahaman tentang *trade-off* antara harga,

lokasi, dan spesifikasi. Calon pembeli juga mendapatkan *insight* tentang tren pasar untuk *timing* pembelian yang optimal.

3.7.3 Bagi Pemerintah Daerah

Data hasil penelitian dapat digunakan untuk perencanaan zonasi dan pengembangan infrastruktur, penyusunan kebijakan perumahan bersubsidi, serta pemahaman tentang *demand pattern* untuk perencanaan kota.

3.8 Limitasi Penelitian

Beberapa limitasi yang perlu dipertimbangkan dalam interpretasi hasil:

1. *Temporal Scope*: Data hanya mencakup periode 6 bulan (Agustus 2024-Januari 2025), sehingga belum menangkap variasi musiman tahunan secara komprehensif.
2. *Geographic Coverage*: Fokus penelitian hanya pada wilayah Bandung dan sekitarnya, sehingga hasil mungkin tidak dapat digeneralisasi ke kota lain dengan karakteristik berbeda.
3. *Data Source*: Ketergantungan pada satu sumber data (*web scraping*) yang mungkin mengandung bias *sampling* dan tidak mencakup seluruh transaksi properti di wilayah studi.
4. *Dynamic Market*: Pasar properti yang dinamis memerlukan *update* model secara berkala untuk memastikan relevansi hasil analisis.

KESIMPULAN

Penelitian ini berhasil memberikan pemahaman yang komprehensif mengenai dinamika pasar properti di wilayah Bandung, khususnya dalam memetakan heterogenitas harga yang sangat ekstrem, mulai dari Rp 18,5 juta hingga Rp 8,75 miliar, dengan rata-rata harga sebesar Rp 1,98 miliar. Temuan ini menegaskan bahwa pasar properti Bandung memiliki keragaman karakteristik yang tinggi, baik dari sisi harga maupun tipe hunian, sehingga membutuhkan pendekatan analitik yang sistematis untuk memahami strukturnya secara menyeluruh.

Dari sisi spasial, penelitian ini mengidentifikasi Kota Baru Parahyangan sebagai wilayah mixed-development yang mendominasi kedua kluster, sekaligus menunjukkan perbedaan pola sebaran antara kawasan pengembangan baru yang cenderung reguler dan wilayah yang telah mapan serta bersifat eksklusif. Selain itu, analisis korelasi mengonfirmasi secara kuantitatif bahwa luas tanah ($r = 0,71$) dan luas bangunan ($r = 0,69$) merupakan faktor determinan utama dalam pembentukan harga properti, yang menegaskan pentingnya dimensi fisik sebagai penentu nilai hunian di Bandung.

Melalui penerapan algoritma K-Means, penelitian ini menghasilkan dua kluster optimal dengan tingkat validitas yang baik ($k = 2$, Silhouette Score 0,335). Cluster 0 merepresentasikan segmen hunian reguler dengan 32.424 properti dan rata-rata harga Rp 1,23 miliar, sedangkan Cluster 1 mencerminkan segmen hunian eksklusif dengan 24.766 properti dan rata-rata harga Rp 2,97 miliar. Segmentasi ini tidak hanya membuktikan adanya pembagian pasar yang jelas, tetapi juga memberikan dasar analitis yang kuat untuk memahami perbedaan karakteristik dan daya beli antar segmen.

Secara strategis, hasil penelitian ini berkontribusi sebagai kerangka analitik yang dapat dimanfaatkan oleh pengembang dalam merumuskan strategi produk dan pemasaran, oleh konsumen sebagai panduan dalam pengambilan keputusan

pembelian, serta oleh pemerintah daerah sebagai referensi dalam perencanaan tata ruang dan kebijakan perumahan. Ke depan, penelitian ini masih memiliki ruang pengembangan melalui perluasan cakupan temporal data, integrasi variabel eksternal seperti infrastruktur dan indikator makroekonomi, eksplorasi algoritma klustering alternatif, pengembangan model prediktif harga, serta implementasi sistem pemantauan pasar properti secara real-time untuk menangkap dinamika pasar yang terus berubah.

DAFTAR PUSTAKA

- Aji, B. G., Sondawa, D. C. A., Gifari, M. R., & Wijayanto, S. (2023). Penerapan algoritma K-Means untuk clustering harga rumah di Bandung. Bandung: Jurnal Ilmiah Informatika Global.
- Dzikrillah, M. R., Gutama, D. H., Wijaya, D. P., & Danianti, D. (2024). Implementasi Algoritma K-Means dalam Pengelompokkan Tingkat Transaksi Produk PPOB Nusantara. Malang: Jurnal Informatika Polinema.
- Hadivibowo, A. W. T. (2023). Analisis Segmentasi Pengunjung Mall Menggunakan Algoritma K-Means Clustering. Bekasi: Universitas Bhayangkara Jakarta Raya.
- Krisna, A. B., & Putri, D. M. (2025). Analisis Segmentasi Klien Menggunakan K-Means Clustering Pada Agensi Kreatif Anoorapro. Surabaya: JATI.
- Marcelina, D., Kurnia, A., & Terttiaavini. (2023). Cluster Analysis of Small Medium Enterprise Performance with K-Means Clustering Algorithm. Palembang: Universitas Indo Global Mandiri.
- Prayogo, D., Kurniawan, R., & Wijaya, Y. A. (2025). Algoritma K-Means Untuk Peningkatan Model Segmentasi Data Aset Tetap pada PT. XYZ. Jakarta: Jurnal Informatika dan Teknik Elektro Terapan.
- Rafu, M. D., Rahaningsih, N., & Ali, I. (2025). Segmentasi Penjualan Alat Kesehatan Menggunakan Algoritma K-Means di PT. Arnetha. Surabaya: JATI.
- Sabitha, F. A. (2022). Analisis pengaruh tingkat urbanisasi terhadap ketersediaan lahan permukiman perumahan di Kota Surabaya. *Jurnal Lembaga Ketahanan Nasional Republik Indonesia*, 10(1), 19–26.
- Sari, P., Efan, & Syahri, R. (2023). Algoritma K-Means Clustering: Sebuah Studi Literatur. Pagar Alam: Institut Teknologi Pagar Alam.
- Sinaga, K. P., & Yang, M.-S. (2020). Unsupervised K-means clustering algorithm. Piscataway: IEEE.
- Sofyan, Z. A., Kurniawan, R., & Wijaya, Y. A. (2023). Optimasi Parameter Algoritma K-Means pada Data Rekam Medis Klinik Ananda Medika Rancah. Surabaya: JATI.
- Soleh, O., & Jonas, D. (2024). Penggunaan Algoritma K-Means untuk Segmentasi Data Pelanggan pada Sistem Pemasaran Berbasis Data Mining. Tangerang: Universitas Raharja.
- Wardana, H. K., & Igon, S. S. (2025). Sistem Informasi Geografis Pemetaan Wilayah Pelabuhan di Nusa Tenggara Timur Dengan Metode Clustering K-Means. Surabaya: JATI.
- Wijaya, D. D. R., & Daulay, S. R. (2025). Strategi pemasaran berbasis segmentasi: Menentukan, menganalisis, dan meramalkan pasar sasaran untuk peningkatan kinerja bisnis. *Jurnal Ilmiah Ekonomi dan Manajemen*, 3(2), 464–478.

- Yolanda, Rosa, I., & Ramadhanu, A. (2024). Identifikasi cerdas apel Fuji dan apel hijau: Pendekatan K-Means clustering untuk segmentasi buah. *Journal of Education Research*, 5(3), 4232–4239.
- Zahrial, & Siddik, M. (2023). Penerapan Metode K-Means Clustering Dalam Perancangan Sistem Untuk Menentukan Jumlah Penjualan Obat Terbanyak. Makassar: JMApTeKsi.