

# Implementation of the K-Means Algorithm for Clustering the KBIH Jabal Nur Hajj Pilgrim Area

## Implementasi Algoritma K-Means untuk Clustering Wilayah Jamaah Haji KBIH Jabal Nur

Deva Martha Shella<sup>1)</sup>, Ade Eviyanti<sup>2)</sup>, Azmuri Wahyu Azinar<sup>3)</sup>, Nuril Lutvi Azizah<sup>4)</sup>

<sup>1,2,3,4)</sup> Program Studi Teknik Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

\*Email Penulis Korespondensi : [adeeviyanti@umsida.ac.id](mailto:adeeviyanti@umsida.ac.id)

**Abstract.** KBIH Jabal Nur is a hajj and umrah counseling institution under the auspices of the Muhammadiyah Sidoarjo Regional Government. KBIH Jabal Nur experiences obstacles in managing pilgrim data due to the increasing number of hajj registrants every year which has an impact on marketing efficiency and development strategies. The method used in this study is K-means which groups regions into two categories, namely the largest area and the lowest area. In this study, it was evaluated using the Elbow Method to find the best k with the result k = 2. The results of this grouping indicate that C1 is the largest area with a total of 1,261 data covering 1 area, namely Sidoarjo, C0 is an area with few registrants covering several areas such as Surabaya, Pasuruan, Mojokerto, Lamongan, Blitar and Jember.

**Keywords** - K-Means, Hajj pilgrims, KBIH, regional promotion, Elbow method

**Abstrak.** KBIH Jabal Nur merupakan lembaga bimbingan haji dan umrah yang berada di bawah naungan Pemerintah Daerah Muhammadiyah Sidoarjo yang berada di wilayah Kabupaten Sidoarjo. KBIH Jabal Nur mengalami kendala dalam pengelolaan data jamaah dikarenakan semakin banyaknya pendaftar haji dan umrah setiap tahunnya yang berdampak pada efisiensi pemasaran dan strategi pengembangan. Metode yang digunakan dalam penelitian ini adalah K-means yang mengelompokkan wilayah menjadi dua kategori yaitu wilayah terluas dan wilayah terendah. Pada penelitian ini dilakukan evaluasi dengan menggunakan Metode Elbow untuk mencari k terbaik. Hasil pengelompokan tersebut menunjukkan bahwa Cluster 1 merupakan area terluas dengan jumlah data sebanyak 1.261 yang meliputi 1 area yaitu Sidoarjo, sedangkan Cluster 0 merupakan area dengan pendaftar sedikit yang meliputi beberapa area seperti Surabaya, Pasuruan, Mojokerto, Lamongan, Blitar dan Jember.

**Kata Kunci** - K-Means, jamaah haji, KBIH, promosi wilayah, Metode Elbow

## I. PENDAHULUAN

Melaksanakan ibadah haji untuk yang mampu baik secara fisik maupun finansial adalah bagian dari rukun islam bagi umat muslim. [1]. Ibadah haji dilakukan di kota Mekkah, Arab Saudi selama bulan Dzulhijjah dalam kalender Hijriyah. Ibadah Haji dan umroh bertujuan mendekatkan diri kepada Allah, memperkuat iman dan membersihkan dosa di Baitullah. Perjalanan religius, terutama haji dan umroh, memiliki banyak peminat di kalangan masyarakat saat ini karena mereka menawarkan berbagai macam fasilitas yang dapat disesuaikan dengan keuangan jamaah[9].

KBIH Jabal Nur merupakan lembaga konseling haji dan umroh dibawah naungan Pemerintahan Daerah Muhammadiyah di kabupaten Sidoarjo. Pendaftar Haji dan Umroh di KBIH Jabal Nur tidak hanya dari Sidoarjo saja, tetapi beberapa terdapat dari luar wilayah Sidoarjo. Pada KBIH Jabal Nur dari 2011 sampai sekarang terdapat 1449 data yang terdaftar.

Meningkatnya jumlah data jamaah pertahunnya menjadikan penumpukan data dalam database[11]. Pengelompokan data yang dilakukan oleh KBIH Jabal Nur masih dilakukan secara manual sehingga menjadi kurang efisien. Tidak hanya itu permasalahan yang dihadapi lainnya adalah sangat sulit mengenali strategi yang dibutuhkan dalam pertumbuhan pengelolaan KBIH Jabal Nur kedepannya.

Oleh karena itu dibutuhkan strategi khusus oleh pegawai KBIH dalam melakukan pemasaran untuk mencari calon pendaftar agar promosi yang dilakukan efektif dan efisien[5]. Untuk dapat melakukan promosi yang efektif dan efisien maka pada penelitian ini akan dilakukan dengan melakukan clustering.. Clustering wilayah merupakan salah satu tujuan untuk mengetahui wilayah mana yang kurang dalam hal promosi yang dilakukan oleh Jabal Nur.

Berdasarkan penelitian yang dilakukan oleh Nurzamana [6] membahas tentang clustering data jamaah umroh menggunakan metode fuzzy c-means. Hasil dari penelitian tersebut menggunakan Principal Component Analysis (PCA) menunjukkan bahwa komponen utama pertama (PC1) memiliki nilai eigen value 21.06%, sedangkan komponen utama kedua (PC2) memiliki nilai eigen value 14.65%. Untuk variasi tiga komponen utama, PC3 memiliki nilai eigen value 8.16%. Pada penelitian ini metode Fuzzy C-means perlu melakukan iterasi sebanyak 6 kali untuk

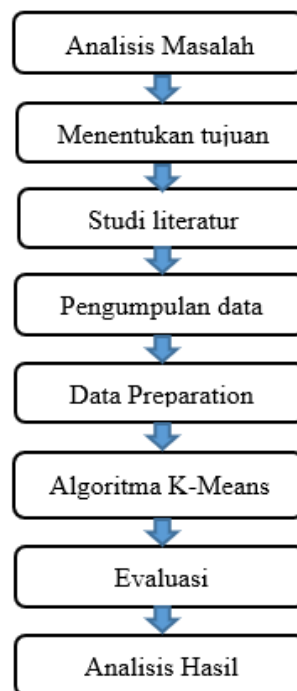
mendapatkan error terkecilnya untuk pengelompokan. Pada penelitian MY. Nurzamana et al (2023) [6] hasil metode K-Means hanya memerlukan 2 iterasi saja untuk mendapatkan rasio yang tepat untuk mengelompokkan dibandingkan dengan metode clustering lain.

Menurut penelitian yang dibuat oleh M.Iqbal [4] K-Means Clustering merupakan algoritma pengelompokan dengan melakukan partisi set data ke dalam k cluster yang sudah ditetapkan di awal. Karena efisiensi dan kesederhanaannya dalam clustering, algoritma K-Means adalah salah satu algoritma yang paling banyak digunakan. Tetapi hasil yang didapatkan pada penelitian ini masih belum efisien dan kurang maksimal dalam menentukan jumlah K cluster. Oleh karena itu, penulis membutuhkan metode yang efektif untuk menentukan kelompok K. Salah satu metode ini adalah metode elbow, yang dapat mengoptimalkan kinerja K-Means.

Dengan penjelasan latar belakang tersebut, penulis tertarik menggunakan K-Means Clustering dalam pengelompokan wilayah karena lebih efisien dalam mengelompokkan data dengan metode Elbow dimana data dikelompokkan berdasarkan karakteristik yang sama sehingga data wilayah terkelompok berdasarkan kabupaten. Maka penulis membuat judul untuk penelitian ini yaitu “Implementasi Algoritma K-Means untuk clustering wilayah jamaah haji KBIH Jabal Nur”.

## II. METODE

Metodologi penelitian adalah alur sistematis atau keseluruhan tahapan yang dilakukan selama penelitian. Pentingnya metodologi ini karena dapat menentukan kualitas penulisan karya ilmiah. [12]. Tujuan dari penelitian ini adalah untuk mengelompokkan data jamaah haji sesuai dengan wilayah tempat tinggal mereka menggunakan algoritma K-Means. Selain itu, mereka juga ingin mengetahui seberapa layak sistem yang dibuat untuk digunakan dalam proses evaluasi pengelompokan data jamaah haji. Langkah langkah pada penelitian yang akan dilakukan dapat dilihat pada Gambar 1.



**Gambar 1.** Langkah-langkah Penelitian

Gambar 1 merupakan bentuk langkah langkah untuk menyusun penelitian ini. Para peneliti akan menjadi lebih memahami alur dari penelitian ini. Tujuannya adalah untuk mengidentifikasi suatu masalah yang ada dalam penelitian.

## 2.1 Analisis Masalah

Pada tahap awal dimulai dengan identifikasi. Pada KBIH Jabal nur terdapat masalah terjadinya penumpukan data karena banyaknya pendaftar setiap tahunnya. Pengelompokan data masih dilakukan secara manual oleh pihak KBIH Jabal Nur. Yang kemudian masalah yang ada dipecahkan dengan metode Algoritma K-Means yang akan digunakan pada penelitian ini.

## 2.2 Menentukan Tujuan

Menentukan tujuan pada penelitian ini digunakan untuk menentukan tujuan penelitian yang jelas. Tujuan penelitian ini adalah sebagai berikut :

- Pengelompokan Jamaah haji berdasarkan variabel usia, gender dan wilayah
- Metode Clustering digunakan untuk menentukan wilayah terbanyak dan sedikit.
- Menggali lebih dalam data pendaftar jamaah haji KBIH Jabal Nur untuk mencari informasi yang belum terpecah.

## 2.3 Studi Literatur

Pencarian pustaka adalah metode penyaringan yang digunakan pada penelitian ini. Literatur bisa ditemukan di buku, prosiding, artikel yang membahas tentang Data Mining, *Clustering*, Algoritma K-Means, KDD (*Knowledge Discovery in Database*) atau pembahasan lainnya yang terkait dengan penelitian. Tujuan dari studi literatur ini adalah supaya penelitian yang dibuat menjadi lebih jelas dan berorientasi.

## 2.4 Pengumpulan Data

Metode pengumpulan data meliputi teknik ataupun alat yang digunakan agar mendapatkan data primer atau sekunder [2]. Dalam mengumpulkan informasi untuk penelitian ini, beberapa metode yang diterapkan sebagai berikut :

- Dokumen

Pada metode ini penulis menggunakan data dari KBIH Jabal Nur dimana data tersebut adalah data jamaah yang mendaftar haji yang berbentuk excel dari tahun 2011-2024.

- Observasi

Observasi dilakukan dengan mendatangi tempat KBIH secara langsung dengan mencari informasi terkait data yang akan diuji pada penelitian ini.

## 2.5 Data Preparation

Pada tahap persiapan data, langkah ini melibatkan mempersiapkan data yang telah dikumpulkan untuk proses data mining. Data diambil dari data pendaftar haji pada KBIH Jabal Nur di Sidoarjo yang diambil dari tahun pemberangkatan 2011-2024. Yang kemudian data tersebut diubah menjadi format yang bersih dan siap untuk diproses dan dianalisis lebih lanjut.

## 2.6 Algoritma K-Means

Langkah berikutnya adalah teknik *clustering*. Menurut IEEE, algoritma K-Means dinobatkan sebagai salah satu dari sepuluh algoritma data mining terbaik karena efisiensi dan kemudahan penggunaan[3].

Metode pada Algoritma K-Means dengan langkah langkah berikut:

- Menentukan nilai K.
- Menentukan titik pusat *cluster* atau *centroid* secara acak.
- Pada langkah ini, pusat cluster atau Centroid awal ditentukan; ini dilakukan dengan mengambil data secara acak dari dataset.
- Setelah menemukan titik pada cluster, rumus *Euclidian Distance* digunakan untuk menghitung jarak antara setiap data dan centroid.

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + \dots (n_i - n_j)^2} \quad \dots(1)$$

Ket :

$D_{ij}$  : Jarak antara i dan j

$X_i X_j$  : Koordinat x objek

$X_i X_j$  : Koordinat x pusat

$X_i X_j$  : Koordinat y objek

$X_i X_j$  : Koordinat y pusat

5. Setelah mengumpulkan hasil perhitungan jarak terdekat, setiap data yang dilihat digabungkan berdasarkan jarak terdekat antara data dan cluster.
6. Menemukan centroid baru dengan menggunakan nilai rata-rata dari data cluster yang sebanding.

## 2.7 Evaluasi

Pada tahap evaluasi, digunakan teknik pengukuran menggunakan *Elbow* selama tahap evaluasi yang dimaksudkan untuk membantu dalam menentukan nilai k terbaik. Dimana metode Elbow ini akan menampilkan nilai k dari 1-9 yang kemudian dimana cluster terbaik akan membentuk siku.

## 2.8 Analisis Hasil

Analisis hasil dilakukan untuk mengetahui tingkat keakuratan hasil yang dikombinasikan dengan hasil data yang sebenarnya. Ini dilakukan untuk memastikan bahwa penelitian yang menggunakan algoritma K-Means mendapatkan hasil yang sama antara penerapan metode secara manual dan perhitungan *google colab*.

# III. HASIL DAN PEMBAHASAN

## 3.1 Data Selection

Tahap pertama mulai dari seleksi data yang diperoleh dari KBIH Jabal Nur, Data Selection merupakan tahapan dimana dilakukan proses pengumpulan dataset awal dengan melakukan pemilihan variable. Berikut pada tabel 1 merupakan data awal pendaftar haji pada KBIH Jabal Nur.

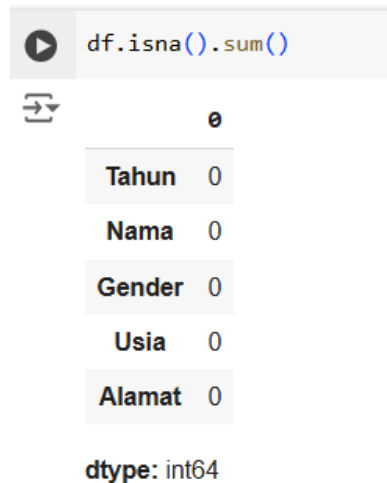
Tabel 1. Hasil data seleksi

Tahun Berangkat	Nama	Jenis Kelamin	Usia	Alamat
2011	Bina Retina	W	64	J. H. P. Blok AC-16 RT. 5/4 Gedangan- Sidoarjo
2011	Nur Madiah Bt Chusairi	W	66	Jl. Wijaya Kusuma 31 RT. 31 RW. 04 Bligo- Candi
2012	Suwito BN Moch Akiyak	P	50	Kepatihan RT. 4/2 Tulangan-Sidoarjo
2012	Nafi'ah BT Kasmun	W	60	Kepatihan RT. 4/2 Tulangan-Sidoarjo
2012	Siti Fatimah BT Soenjoen Fatah	W	70	Penambangan RT. 02 RW. 01 Balongbendo- Sidoarjo
2012	Susanto BN Sumarlan	P	60	Penambangan RT. 02 RW. 01 Balongbendo- Sidoarjo
2012	Innamah BT Mudakir	W	62	Sambisari KAV Utara F-6 RT 35 RW 06 Sambibulu- Taman
2012	Broto Mulyoko BN Sudewo Benowo	P	62	Sambisari KAV Utara F-6 RT 35 RW 06 Sambibulu- Taman
2012	Titi Surtiningsih Sunarja	BT W	67	Simomulyo RT. 15 RW. 04 Kesambi- Porong
2012	Yuli Sudarmanto Munaji	BT W	58	Simomulyo RT. 15 RW. 04 Kesambi- Porong
2012	Siti Muawanah BT Affandi	H. W	70	Trosobo No. 8 RT 3/4 Taman-Sidoarjo

Tabel 1 merupakan data yang telah dianalisis dalam penelitian ini adalah informasi pendaftar haji dari KBIH Jabal Nur selama 2011 sampai sekarang. Dan terdiri dari 9 atribut menjadi 5 atribut dengan total 1449 record.

### 3.2 Data Preprocessing

Pada tahap preprocessing data, biasanya melibatkan proses pembersihan data seperti cleaning data, termasuk penanganan *missing value* [6],[13]. Pengecekan *missing value* dapat dilihat pada gambar berikut.



```
df.isna().sum()
Tahun    0
Nama     0
Gender   0
Usia     0
Alamat   0
dtype: int64
```

**Gambar 2.** Pengecekan *Missing Value*

Gambar 2 merupakan hasil Pengecekan *missing value* dari dataset yang diseleksi dan tidak terdapat *missing value*.

### 3.3 Data Transformation

Setelah itu, akan ditransformasikan sehingga metode K-Means dapat digunakan untuk menghitung data. Namun, sebelum metode ini dapat digunakan untuk perhitungan, data sampel hasil transformasi harus diproses. Berikut pada Tabel 2 merupakan data yang akan menjadi dataset klusterisasi menggunakan algoritma K-Means.

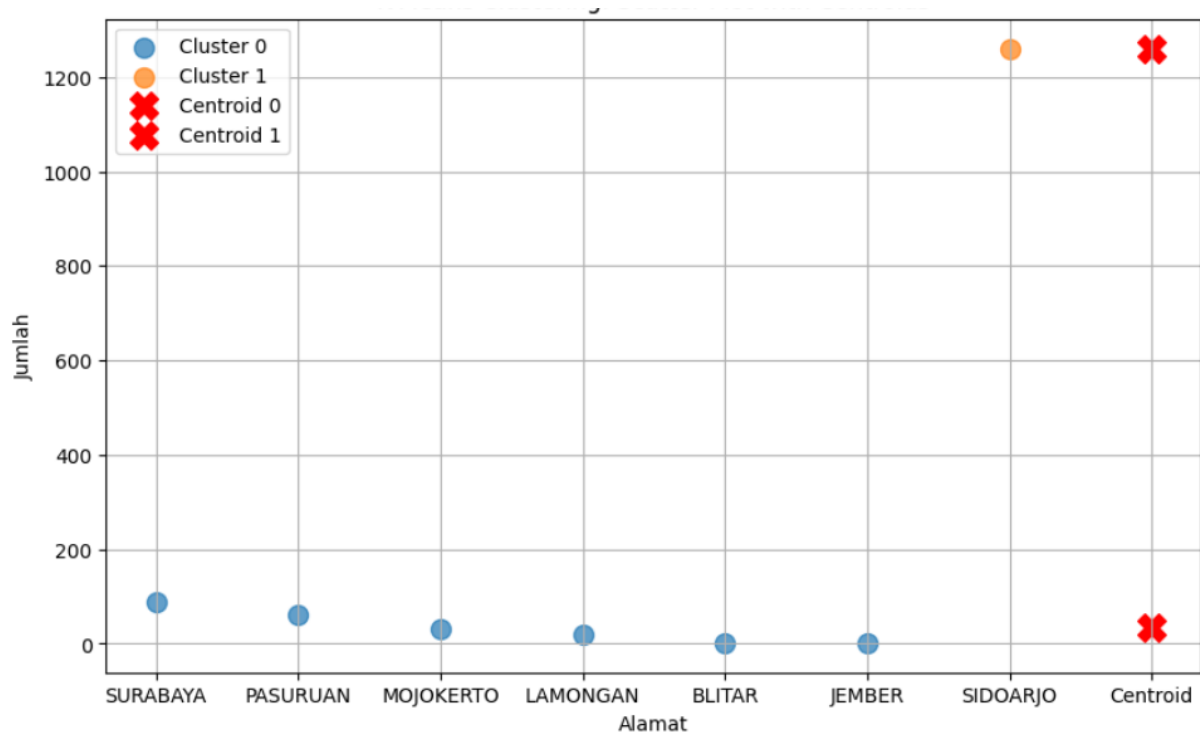
**Tabel 2.** Hasil Transformasi

Index	Tahun	Nama	Gender	Usia	Alamat
0	2011	Heru Santoso BN Madsaridi	0	84	5
1	2011	Bina Retina	1	56	5
2	2012	Nur Madiah BT Chusairi	1	60	5
3	2012	SuWito BN Moch Akiyak	0	61	5
4	2012	Nafi'ah BT Kasmun	1	64	5
5	2012	Siti Fatimah Bt Soenjoen Fatah	1	60	5
6	2012	Susanto Bn Sumartan	0	63	4
7	2012	Innamah Bt Mudakir	1	78	4
8	2012	Broto Mulyono Bn Sadewo	0	59	5
...	...	...	...	...	...
1418	2050	Zainuri Bn Adam SariPah	0	65	5
1419	2050	Mariyam Bt Qasim/Tasrifah	1	58	5
1420	2050	Senifah ( Haji Plus)	1	67	5
1421	2050	Soffyah Bt Kadiran Miun	1	59	5
1422	2050	Suliamah	1	67	5

Tabel 2 merupakan hasil data yang sudah ditranformasi pada kolom gender dan alamat menjadi numerik. Dari 5 variable diatas hanya diambil 3 variabel yang nantinya akan diterapkan pada Metode K-Means.

### 3.4 Pengujian Menggunakan Python pada K-Means

Pada tahapan ini penerapan python menggunakan google Colabs[7],[14]. Hasil pengelompokan data pada KBIH Jabal Nur menggunakan dua kelompok, yang disajikan dalam tampilan bar. Setelah mengumpulkan data tentang masing-masing centroid, peneliti melakukan inisialisasi k pusat cluster (centroid) secara acak. Selanjutnya, mereka mengklasterkan objek berdasarkan jarak ke centroid minimum (terdekat). Untuk hasil pengujian, digunakan dengan nilai  $k=2$ . yang dapat dilihat pada Gambar 3.

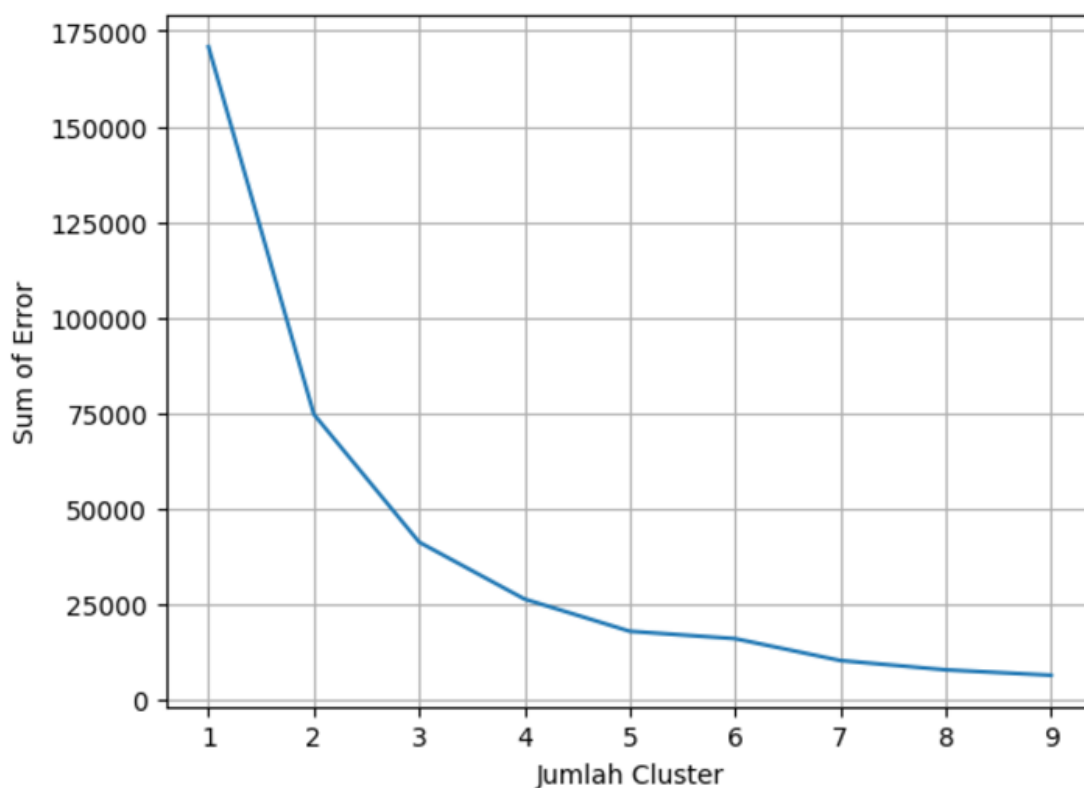


**Gambar 3.** Visualisasi Grafik

Gambar 3 menjelaskan bahwa di KBIH Jabal Nur Cluster 0 yaitu daerah yang paling sedikit jumlah jamaah yang mendaftar keseluruhan 162 data, dan pada Cluster 1 yaitu daerah terbanyak jumlah jamaah yang mendaftar dengan keseluruhan 1260 data dari tahun 2011-2025.

### 3.5 Evaluasi

Pada bagian ini disampaikan mekanisme pengukuran k terbaik [15]. Metode uji dari penelitian ini menggunakan metode Elbow. Pada gambar 4 berikut merupakan gambar yang menunjukkan grafik Elbow.



**Gambar 4.** Grafik Elbow

Grafik 4 adalah visualisasi kluster optimal dengan menggunakan metode Elbow. Berdasarkan gambar 4, menunjukkan bahwa cluster 2 memberikan sudut yang membentuk siku dan mengalami penurunan paling besar. Oleh karena itu dari hasil grafik diatas menunjukkan cluster terbaik adalah cluster 2.

### 3.6 Analisa Hasil

Setelah melakukan pemodelan dengan metode Elbow yang menggunakan 2 cluster pada dataset KBIH Jabal Nur maka diperoleh hasil pada table 3.

**Tabel 3.** Hasil Keseluruhan tiap Cluster

Index	Alamat	Frequency	Cluster	Cluster_Label
0	5	1261	1	Wilayah Terbanyak
1	6	89	0	Wilayah Sedikit
2	4	40	0	Wilayah Sedikit
3	3	26	0	Wilayah Sedikit
4	2	3	0	Wilayah Sedikit
5	0	2	0	Wilayah Sedikit
6	1	2	0	Wilayah Sedikit

Tabel 3 merupakan hasil klasterisasi menggunakan algoritma K-Means yang disajikan, menunjukkan bahwa alamat yang paling terbanyak adalah wilayah 5 atau Sidoarjo yang terletak pada cluster 1, dimana Sidoarjo memiliki 1261 jumlah data pendaftar. Sedangkan wilayah yang paling sedikit dengan jumlah 2 data dimiliki oleh wilayah 0 dan 1 yaitu Blitar dan Jember.

#### IV. SIMPULAN

Dalam penelitian ini berhasil mengelompokkan wilayah jamaah haji pada KBIH Jabal Nur di Sidoarjo dengan mengaplikasikan algoritma K-Means. Studi ini menggunakan 1423 data yang sudah di *cleaning* dengan 5 parameter yaitu tahun, nama, gender, usia dan alamat. Dimana hasil dimana hasil diperoleh dengan bantuan google colab terbentuk 2 cluster dalam kategori wilayah terbanyak di cluster 0 dengan jumlah keseluruhan 874 dan wilayah sedikit di cluster 1 dengan jumlah keseluruhan 549. Hasil evaluasi yang dilakukan dengan metode Elbow untuk menentukan k terbaik menghasilkan 2 kluster, dengan K=2 sebagai kluster terbaik setelah eksperimen dilaksanakan dengan nilai K=1 sampai k=9. Penelitian ini memiliki kelebihan termasuk kemudahan pada preprocessing data dan integrasi dengan machine learning yang membantu peneliti dalam memetakan wilayah terbanyak dan sedikit pada data pendaftar KBIH Jabal Nur di Sidoarjo. Studi ini memberi KBIH di Sidoarjo informasi terbaru yang dapat membantu mereka menerapkan strategi pemasaran yang lebih baik.

Penelitian selanjutnya dapat membandingkan algoritma K-Means dengan metode clustering lainnya, seperti Fuzzy C-Means atau DBSCAN, untuk mengevaluasi keunggulan dan kelemahannya dalam mengelompokkan data jamaah haji. Selain itu, Penelitian ini menggunakan variabel seperti gender, usia, dan alamat. Dapat ditambah variabel tambahan seperti tingkat pendidikan, pekerjaan, atau penghasilan dapat dimasukkan untuk analisis yang lebih mendalam dan akurat. Penelitian juga dapat dikembangkan dengan menerapkan model berbasis machine learning yang lebih kompleks atau mengintegrasikan hasil clustering ke dalam sistem informasi berbasis web untuk mempermudah pengambilan keputusan strategis oleh pihak KBIH.

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh pihak yang telah mendukung penelitian ini, khususnya KBIH Jabal Nur di Sidoarjo yang telah menyediakan data dan informasi berharga untuk keperluan penelitian. Ucapan terima kasih juga disampaikan kepada Universitas Muhammadiyah Sidoarjo, khususnya Fakultas Sains dan Teknologi, atas bimbingan, fasilitas, dan dukungan yang diberikan selama proses penelitian ini. Tidak lupa, apresiasi yang mendalam ditujukan kepada rekan-rekan peneliti dan pihak-pihak lain yang telah memberikan masukan dan kontribusi sehingga penelitian ini dapat terselesaikan dengan baik. Semoga hasil penelitian ini dapat bermanfaat bagi pengembangan strategi pemasaran dan pengelolaan data jamaah haji di masa mendatang.

## REFERENSI

- [1] Ali, A. (2019). Klasterisasi Data Rekam Medis Pasien Menggunakan Metode K-Means *Clustering* di Rumah Sakit Anwar Medika Balong Bendo Sidoarjo. *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 186-195. <https://doi.org/10.30812/matrik.v19i1.529>
- [2] Djaka Permana, M., Lia Hananto, A., Novalia, E., Huda, B., & Paryono, T. (2023). Klasterisasi Data Jamaah Umrah pada Tanurmutmainah Tour Menggunakan Algoritma K-Means. *Jurnal KomtekInfo*, 15–20. <https://doi.org/10.35134/komtekinfo.v10i1.332>
- [3] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, “Penerapan Algoritma K-Means Untuk *Clustering* Data Obat-Obatan,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [4] M. Iqbal, “Klasterisasi Data Jamaah Umroh Pada Auliya Tour & Travel Menggunakan Metode K-Means *Clustering*,” *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 5, no. 2, pp. 97–104, 2019, doi: 10.33330/jurteksi.v5i2.352.
- [5] M.A Albaihaqi, “STRATEGI PEMASARAN KELOMPOK BIMBINGAN IBADAH HAJI (KBIH) AT TAQWA UJUNG HARAPAN BEKASI DALAM MENINGKATKAN JUMLAH JAMA’AH.”
- [6] Muhammad Yamin Nurzaman, & Nurina Sari, B. (2023). Implementasi K-Means *Clustering* Dalam Pengelompokan Banyaknya jumlah petani. *Jurnal Teknik Informatika Dan Sistem Informasi*, 10(3). <http://jurnal.mdp.ac.id>
- [7] Naufal Hayyu Triwardana. (n.d.). IMPLEMENTASI METODE FUZZY C-MEANS UNTUK CLUSTERING TRAVEL UMROH BERDASARKAN PRINCIPAL COMPONENT ANALYSIS
- [8] S. A. Azis, S. Defit, and Y. Yunus, “Jurnal Informatika Ekonomi Bisnis Klasterisasi Dana Bantuan Pada Program Keluarga Harapan ( PKH ),” vol. 3, pp. 53–59, 2021, doi: 10.37034/infeb.v3i2.66.
- [9] T. Hidayat, “Klasifikasi Data Jamaah Umroh Menggunakan Metode K-Means Clustering,” *Jurnal Sistim Informasi dan Teknologi*, pp. 19–24, Feb. 2022, doi: 10.37034/jsisfotek.v4i1.115.
- [10] Yunitasari, Melinda, Tamara Maharani, and Bagus Hikmahwan Akademi Komunitas Negeri Pacitan Jl Walanda Maramis. 2022. “IMPLEMENTASI METODE K-MEANS UNTUK PENGELOMPOKAN DATA JAMAAH” 09 (1).
- [11] I. Manajemen Promosi KBIH Dalam Meningkatkan Minat Bimbingan Jamaah Haji Saepurrahmat, A. Iwan Setiawan, J. Manajemen Dakwah, and U. Sunan Gunung Djati, “Anida (Aktualisasi Nuansa Ilmu Dakwah),” 2017.
- [12] I. Sanela, A. Nazir, F. Syafria, E. Haerani, and L. Oktavia, “Penerapan Metode Clustering Dengan K-Means Untuk Memetakan Potensi Tanaman Padi di Sumatera,” *Journal of Computer System and Informatics (JoSYC)*, vol. 5, no. 1, pp. 82–92, Nov. 2023, doi: 10.47065/josyc.v5i1.4523.
- [13] A. A. Hussein, “Improve The Performance of K-means by using Genetic Algorithm for Classification Heart Attack,” *Int. J. Electr. Comput. Eng.*, vol. 8, no. 2, p. 1256, 2018, doi: 10.11591/ijece.v8i2.pp1256-1261.
- [14] R. K. Dinata, H. Novriando, N. Hasdyna, and S. Retno, “Reduksi Atribut Menggunakan Information Gain,” vol. 6, no. 1, pp. 48–53, 2020.
- [15] I. Kamila, U. Khairunnisa, and M. Mustakim, “Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau,” *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 5, no. 1, p. 119, 2019, doi: 10.24014/rmsi.v5i1.7381.

***Conflict of Interest Statement:***

*The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.*