
Penerapan Metode Clustering Dengan K-Means untuk Memetakan Presentase Rumah Tangga dengan Akses Terhadap Hunian yang Layak dan Terjangkau Menurut Provinsi di Indonesia Tahun 2021-2023

Application of the Clustering Method Using K-Means to Map the Percentage of Households with Access to Decent and Affordable Housing According to Provinces in Indonesia in 2021-2023

Reni Triyaningsih¹

¹Program Studi Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Jambi
renitriyaningsih72@gmail.com

Abstrak

Selama periode 2021-2023, dinamika sosial, ekonomi, dan demografis di Indonesia mengalami perubahan yang signifikan, yang dapat berdampak pada akses terhadap hunian. Jurnal ini membahas penerapan metode clustering menggunakan algoritma K-Means untuk memetakan presentase rumah tangga dengan akses terhadap hunian yang layak dan terjangkau di berbagai provinsi di Indonesia dalam rentang waktu 2021-2023. Penelitian ini bertujuan untuk mengidentifikasi pola distribusi akses perumahan di tingkat provinsi dan menyusun kelompok rumah tangga berdasarkan karakteristik tertentu. Metode K-Means digunakan untuk mengelompokkan data presentasi rumah tangga berdasarkan variabel yang relevan. Hasil analisis clustering diharapkan dapat memberikan pemahaman yang lebih mendalam mengenai disparitas akses terhadap hunian yang layak dan terjangkau di Indonesia, memberikan dasar untuk perumusan kebijakan perumahan yang lebih efektif dan berkelanjutan.

Kata Kunci: Clustering, K-means, Akses Hunian Layak, Hunian Terjangkau.

Abstract

During the 2021-2023 period, social, economic and demographic dynamics in Indonesia will experience significant changes, which could impact access to housing. This journal discusses the application of the clustering method using the K-Means algorithm to map the percentage of households with access to decent and affordable housing in various provinces in Indonesia in the 2021-2023 time period. This research aims to identify distribution patterns of housing access at the provincial level and organize household groups based on certain characteristics. The K-Means method is used to group household presentation data based on relevant variables. It is hoped that the results of the clustering analysis will provide a deeper understanding of disparities in access to adequate and affordable housing in Indonesia, providing a basis for formulating more effective and sustainable housing policies.

Keywords: Clustering, K-means, Access to Decent Housing, Affordable Housing.

Pendahuluan

Perumahan yang layak dan terjangkau merupakan salah satu aspek kritis dalam meningkatkan kualitas hidup masyarakat, terutama di negara berkembang seperti Indonesia. Meskipun upaya telah dilakukan untuk memperbaiki akses terhadap perumahan, disparitas

yang masih signifikan antarprovinsi mengindikasikan adanya kompleksitas dalam dinamika perumahan di berbagai wilayah. Dalam konteks ini, penelitian ini bertujuan untuk menyelidiki dan memetakan presentasi rumah tangga di Indonesia selama periode 2021-2023 dengan fokus pada akses terhadap hunian yang layak dan terjangkau.

Periode waktu yang dipilih, yaitu 2021-2023, menjadi krusial karena mencakup dinamika sosial, ekonomi, dan demografis yang signifikan, yang dapat memberikan gambaran yang lebih akurat terhadap kondisi perumahan di berbagai wilayah Indonesia. Metode clustering K-Means dipilih karena kemampuannya dalam mengelompokkan data menjadi kelompok yang homogen berdasarkan karakteristik yang relevan, seperti pendapatan, lokasi geografis, dan ketersediaan infrastruktur pendukung. Penerapan metode ini diharapkan dapat memberikan pemahaman yang lebih mendalam tentang disparitas akses perumahan, memberikan landasan bagi penyusunan kebijakan perumahan yang lebih terarah dan efektif.

Dalam konteks dinamika sosial, ekonomi, dan demografis yang terus berubah di Indonesia, penelitian ini juga memberikan kontribusi dalam menangkap perubahan-perubahan tersebut yang dapat memengaruhi akses terhadap perumahan. Dengan demikian, diharapkan hasil penelitian ini dapat menjadi landasan informasi yang berharga bagi pihak-pihak terkait dalam merancang dan melaksanakan kebijakan perumahan yang efektif dan berkelanjutan di tingkat nasional serta provinsi.

Metode Penelitian

Tahap Pengumpulan Data

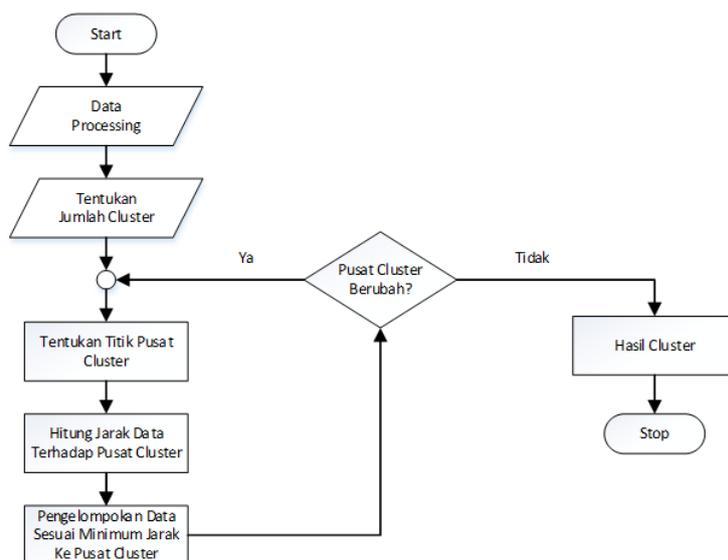
Sumber data penelitian diperoleh dari data yang dikumpulkan berdasarkan dokumen-dokumen yang dihasilkan oleh Badan Pusat Statistik Indonesia melalui situs <https://www.bps.go.id/> Data yang digunakan dalam penelitian ini adalah data Persentase Rumah Tangga yang Memiliki Akses Terhadap Hunian Yang Layak Dan Terjangkau Menurut Provinsi (Persen), 2021-2023.

Tahap Data Preparation

Data preparation adalah proses mengumpulkan, menggabungkan, menyusun, dan mengatur data sehingga bisa dipakai dalam visualisasi data. Pada tahap ini, dilakukan pengecekan data terhadap *missing value*, mengubah bentuk data menjadi data frame, mengubah nama objek pada data menjadi label, serta menghapus data yang sekiranya tidak terpakai untuk masuk ke dalam tahap clustering.

Tahap Data Mining

Pada tahap data mining adalah dimana identifikasi dimulai. Pada penelitian ini menggunakan algoritma K-Means Clustering. Dalam menentukan jumlah cluster optimal, penelitian ini menggunakan metode Elbow dan Average Silhouette. Berikut adalah alur dari algoritma K-Means seperti pada Gambar 1.



Gambar 1. Alur Algoritma Kmeans

Hasil dan Pembahasan

Pengumpulan Data

Data yang digunakan dalam penelitian ini ialah data resmi dari Badan Pusat Statistik, dengan judul Persentase Rumah Tangga yang Memiliki Akses Terhadap Hunian Yang Layak Dan Terjangkau Menurut Provinsi (Persen), 2021-2023 yang dapat diakses melalui <https://www.bps.go.id/id/statistics-table/2/MTI0MSMy/persentase-rumah-tangga-yang-memiliki-akses-terhadap-hunian-yang-layak-dan-terjangkau-menurut-provinsi--persen-.html> Data yang akan digunakan dapat dilihat pada Tabel 1.

Tabel 1. Data Persentase Rumah Tangga yang Memiliki Akses Terhadap Hunian yang Layak dan Terjangkau Menurut Provinsi

Provinsi	2021	2022	2023
Aceh	64,43	64,18	65,91
Sumatera Utara	69,48	67,26	70,95
Sumatera Barat	56,67	58,18	59,85
Riau	70,63	69,43	71,53
Jambi	62,54	60,85	64,12
Sumatera Selatan	57,86	59,96	61,82
Bengkulu	51,36	54,98	54,74
Lampung	61,56	61,24	63,15
Kep. Bangka Belitung	27,6	30,79	32,57
Kep. Riau	50,23	46,69	54,21
DKI Jakarta	40	36,23	38,8
Jawa Barat	53,14	53,37	54,17
Jawa Tengah	66,47	67,02	68,85
DI Yogyakarta	85,15	84,94	85,79
Jawa Timur	66,93	66,28	70,74
Banten	60,78	60,98	63,06
Bali	78,47	81,65	84,26
Nusa Tenggara Barat	62,9	62,3	66,31

Provinsi	2021	2022	2023
Nusa Tenggara Timur	40,41	41,8	42,7
Kalimantan Barat	61,17	60,74	64,16
Kalimantan Tengah	55,34	54,07	56,49
Kalimantan Selatan	57,5	55,96	57,5
Kalimantan Timur	70,7	73,18	75,82
Kalimantan Utara	65,65	67,21	71,43
Sulawesi Utara	69,5	69,57	72,39
Sulawesi Tengah	62,7	58,61	59,6
Sulawesi Selatan	69,11	69,9	71,88
Sulawesi Tenggara	70,45	72,8	74,01
Gorontalo	67,28	69,48	70,75
Sulawesi Barat	57,26	55,37	59,28
Maluku	60,69	60,66	63,61
Maluku Utara	63,85	65,42	68,17
Papua Barat	57,9	53,81	56,92
Papua	28,92	27,28	29,01

Data Preparation

Import Dataset

Sebelum memasuki proses Clustering kita perlu menyiapkan dataset terlebih dahulu, pada Rstudio dibutuhkan Library Cluster, Tidyverse, Factoextra, ggplot2, dan readxl untuk membaca data yang telah kita import melalui file excel. Library-library tersebut perlu dipanggil sebelum masuk ke tahap clustering seperti yang bisa dilihat pada Gambar 2.

```

> library(readxl)
> AksesHunianIndonesia <- read_excel("AksesHunianIndonesia.xlsx")
> View(AksesHunianIndonesia)
> library(cluster)
> library(tidyverse)
— Attaching core tidyverse packages — tidyverse 2.0.0 —
✓ dplyr 1.1.3 ✓ readr 2.1.4
✓ forcats 1.0.0 ✓ stringr 1.5.0
✓ ggplot2 3.4.3 ✓ tibble 3.2.1
✓ lubridate 1.9.3 ✓ tidyr 1.3.0
✓ purrr 1.0.2
— Conflicts — tidyverse_conflicts() —
✗ dplyr::filter() masks stats::filter()
✗ dplyr::lag() masks stats::lag()
i Use the conflicted package to force all conflicts to become errors
> library(factoextra)
Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3wBa
> library(ggplot2)

```

Gambar 2. Memanggil Library pada Rstudio

Cek Missing Value

Untuk mendapatkan iterasi dengan hasil terbaik, diperlukanlah pengecekan Missing Value untuk memastikan apakah ada data yang Null. Pengecekan Missing Value dapat dilihat pada Gambar 3.

```
> AksesHunianIndonesia <- is.na(AksesHunianIndonesia)
> AksesHunianIndonesia
  Provinsi 2021 2022 2023
[1,] FALSE FALSE FALSE FALSE
[2,] FALSE FALSE FALSE FALSE
[3,] FALSE FALSE FALSE FALSE
[4,] FALSE FALSE FALSE FALSE
[5,] FALSE FALSE FALSE FALSE
[6,] FALSE FALSE FALSE FALSE
[7,] FALSE FALSE FALSE FALSE
[8,] FALSE FALSE FALSE FALSE
[9,] FALSE FALSE FALSE FALSE
[10,] FALSE FALSE FALSE FALSE
[11,] FALSE FALSE FALSE FALSE
[12,] FALSE FALSE FALSE FALSE
[13,] FALSE FALSE FALSE FALSE
[14,] FALSE FALSE FALSE FALSE
[15,] FALSE FALSE FALSE FALSE
[16,] FALSE FALSE FALSE FALSE
[17,] FALSE FALSE FALSE FALSE
[18,] FALSE FALSE FALSE FALSE
[19,] FALSE FALSE FALSE FALSE
[20,] FALSE FALSE FALSE FALSE
[21,] FALSE FALSE FALSE FALSE
[22,] FALSE FALSE FALSE FALSE
[23,] FALSE FALSE FALSE FALSE
[24,] FALSE FALSE FALSE FALSE
[25,] FALSE FALSE FALSE FALSE
[26,] FALSE FALSE FALSE FALSE
[27,] FALSE FALSE FALSE FALSE
[28,] FALSE FALSE FALSE FALSE
[29,] FALSE FALSE FALSE FALSE
[30,] FALSE FALSE FALSE FALSE
[31,] FALSE FALSE FALSE FALSE
[32,] FALSE FALSE FALSE FALSE
[33,] FALSE FALSE FALSE FALSE
[34,] FALSE FALSE FALSE FALSE
```

Gambar 3. Pengecekan Missing Value

3.2.4 Mengubah Data Menjadi Data Frame

Sebelum mengeksekusi data, data excel perlu diubah menjadi dataframe agar dapat dibaca oleh Rstudio seperti pada gambar 4.

```
> DataAksesHunianIndonesia <- as.data.frame(AksesHunianIndonesia)
> DataAksesHunianIndonesia
  Provinsi 2021 2022 2023
1 Aceh 64.43 64.18 65.91
2 Sumatera Utara 69.48 67.26 70.95
3 Sumatera Barat 56.67 58.18 59.85
4 Riau 70.63 69.43 71.53
5 Jambi 62.54 60.85 64.12
6 Sumatera Selatan 57.86 59.96 61.82
7 Bengkulu 51.36 54.98 54.74
8 Lampung 61.56 61.24 63.15
9 Kep. Bangka Belitung 27.60 30.79 32.57
10 Kep. Riau 50.23 46.69 54.21
11 DKI Jakarta 40.00 36.23 38.80
12 Jawa Barat 53.14 53.37 54.17
13 Jawa Tengah 66.47 67.02 68.85
14 DI Yogyakarta 85.15 84.94 85.79
15 Jawa Timur 66.93 66.28 70.74
16 Banten 60.78 60.98 63.06
17 Bali 78.47 81.65 84.26
18 Nusa Tenggara Barat 62.90 62.30 66.31
19 Nusa Tenggara Timur 40.41 41.80 42.70
20 Kalimantan Barat 61.17 60.74 64.16
21 Kalimantan Tengah 55.34 54.07 56.49
22 Kalimantan Selatan 57.50 55.96 57.50
23 Kalimantan Timur 70.70 73.18 75.82
24 Kalimantan Utara 65.65 67.21 71.43
25 Sulawesi Utara 69.50 69.57 72.39
26 Sulawesi Tengah 62.70 58.61 59.60
27 Sulawesi Selatan 69.11 69.90 71.88
28 Sulawesi Tenggara 70.45 72.80 74.01
29 Gorontalo 67.28 69.48 70.75
30 Sulawesi Barat 57.26 55.37 59.28
31 Maluku 60.69 60.66 63.61
32 Maluku Utara 63.85 65.42 68.17
33 Papua Barat 57.90 53.81 56.92
34 Papua 28.92 27.28 29.01
```

Gambar 4. Mengubah Data ke Data Frame

Cek Struktur Data

Pengecekan Struktur data bertujuan untuk kembali memastikan apakah data yang digunakan sudah benar menjadi dataframe seperti pada Gambar 5.

```
> str(DataAksesHunianIndonesia)
'data.frame': 34 obs. of 4 variables:
 $ Provinsi: chr "Aceh" "Sumatera Utara" "Sumatera Barat" "Riau" ...
 $ 2021 : num 64.4 69.5 56.7 70.6 62.5 ...
 $ 2022 : num 64.2 67.3 58.2 69.4 60.9 ...
 $ 2023 : num 65.9 71 59.9 71.5 64.1 ...
```

Gambar 5. Cek Struktur Data

Mengubah Nama Objek Menjadi Label

Untuk menerapkan teknik data mining, nama objek pada data tidak ikut dihitung. Maka dari itu diperlukan tahap mengubah nama objek menjadi label agar nama objek tidak ikut dihitung saat proses clustering, seperti pada Gambar 6.

```
> row.names(DataAksesHunianIndonesia) <- c(DataAksesHunianIndonesia$Provinsi)
> DataAksesHunianIndonesia
```

	Provinsi	2021	2022	2023
Aceh	Aceh	64.43	64.18	65.91
Sumatera Utara	Sumatera Utara	69.48	67.26	70.95
Sumatera Barat	Sumatera Barat	56.67	58.18	59.85
Riau	Riau	70.63	69.43	71.53
Jambi	Jambi	62.54	60.85	64.12
Sumatera Selatan	Sumatera Selatan	57.86	59.96	61.82
Bengkulu	Bengkulu	51.36	54.98	54.74
Lampung	Lampung	61.56	61.24	63.15
Kep. Bangka Belitung	Kep. Bangka Belitung	27.60	30.79	32.57
Kep. Riau	Kep. Riau	50.23	46.69	54.21
DKI Jakarta	DKI Jakarta	40.00	36.23	38.80
Jawa Barat	Jawa Barat	53.14	53.37	54.17
Jawa Tengah	Jawa Tengah	66.47	67.02	68.85
DI Yogyakarta	DI Yogyakarta	85.15	84.94	85.79
Jawa Timur	Jawa Timur	66.93	66.28	70.74
Banten	Banten	60.78	60.98	63.06
Bali	Bali	78.47	81.65	84.26
Nusa Tenggara Barat	Nusa Tenggara Barat	62.90	62.30	66.31
Nusa Tenggara Timur	Nusa Tenggara Timur	40.41	41.80	42.70
Kalimantan Barat	Kalimantan Barat	61.17	60.74	64.16
Kalimantan Tengah	Kalimantan Tengah	55.34	54.07	56.49
Kalimantan Selatan	Kalimantan Selatan	57.50	55.96	57.50
Kalimantan Timur	Kalimantan Timur	70.70	73.18	75.82
Kalimantan Utara	Kalimantan Utara	65.65	67.21	71.43
Sulawesi Utara	Sulawesi Utara	69.50	69.57	72.39
Sulawesi Tengah	Sulawesi Tengah	62.70	58.61	59.60
Sulawesi Selatan	Sulawesi Selatan	69.11	69.90	71.88
Sulawesi Tenggara	Sulawesi Tenggara	70.45	72.80	74.01
Gorontalo	Gorontalo	67.28	69.48	70.75
Sulawesi Barat	Sulawesi Barat	57.26	55.37	59.28
Maluku	Maluku	60.69	60.66	63.61
Maluku Utara	Maluku Utara	63.85	65.42	68.17
Papua Barat	Papua Barat	57.90	53.81	56.92
Papua	Papua	28.92	27.28	29.01

Gambar 6. Mengubah Kolom Objek ke Label

Menghapus Kolom Objek

Setelah mengubah kolom objek menjadi label, maka kolom objek tidak diperlukan lagi dan tidak masuk ke dalam perhitungan. Maka, kolom objek kita buang seperti pada Gambar 7.

```
> DataAksesHunianIndonesia1 <- DataAksesHunianIndonesia[-1]
> DataAksesHunianIndonesia1
      2021 2022 2023
Aceh      64.43 64.18 65.91
Sumatera Utara 69.48 67.26 70.95
Sumatera Barat 56.67 58.18 59.85
Riau      70.63 69.43 71.53
Jambi     62.54 60.85 64.12
Sumatera Selatan 57.86 59.96 61.82
Bengkulu  51.36 54.98 54.74
Lampung   61.56 61.24 63.15
Kep. Bangka Belitung 27.60 30.79 32.57
Kep. Riau 50.23 46.69 54.21
DKI Jakarta 40.00 36.23 38.80
Jawa Barat 53.14 53.37 54.17
Jawa Tengah 66.47 67.02 68.85
DI Yogyakarta 85.15 84.94 85.79
Jawa Timur 66.93 66.28 70.74
Banten    60.78 60.98 63.06
Bali      78.47 81.65 84.26
Nusa Tenggara Barat 62.90 62.30 66.31
Nusa Tenggara Timur 40.41 41.80 42.70
Kalimantan Barat 61.17 60.74 64.16
Kalimantan Tengah 55.34 54.07 56.49
Kalimantan Selatan 57.50 55.96 57.50
Kalimantan Timur 70.70 73.18 75.82
Kalimantan Utara 65.65 67.21 71.43
Sulawesi Utara 69.50 69.57 72.39
Sulawesi Tengah 62.70 58.61 59.60
Sulawesi Selatan 69.11 69.90 71.88
Sulawesi Tenggara 70.45 72.80 74.01
Gorontalo 67.28 69.48 70.75
Sulawesi Barat 57.26 55.37 59.28
Maluku    60.69 60.66 63.61
Maluku Utara 63.85 65.42 68.17
Papua Barat 57.90 53.81 56.92
Papua     28.92 27.28 29.01
```

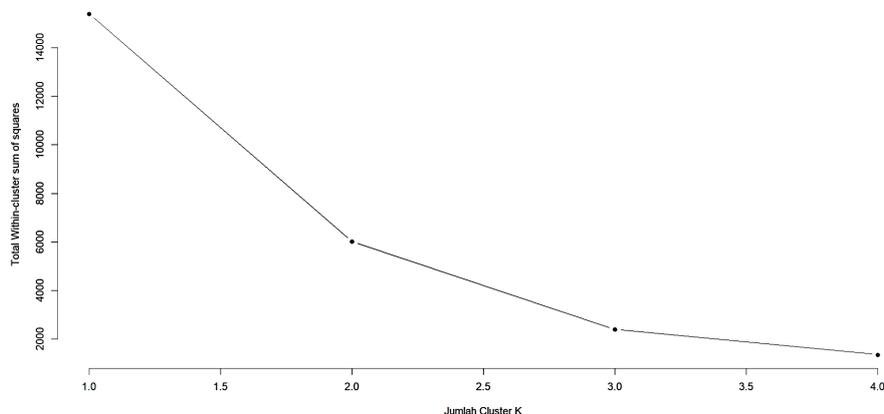
Gambar 7. Hapus Kolom Objek

Data Mining

Menentukan Jumlah Cluster Optimal Menggunakan Metode Elbow

Metode Elbow merupakan salah satu metode untuk menentukan jumlah cluster yang tepat melalui persentase hasil perbandingan antara jumlah cluster yang akan membentuk siku pada suatu titik seperti pada Gambar 8.

```
> wss <- function(k){kmeans(DataAksesHunianIndonesia1, k, nstart = 10)$tot.withinss}
> k.values <- 1:4
> wss_values <- map_dbl(k.values, wss)
> plot(k.values, wss_values,
+      type = "b", pch = 19, frame = FALSE,
+      xlab = "Jumlah Cluster K",
+      ylab = "Total Within-cluster sum of squares")
```

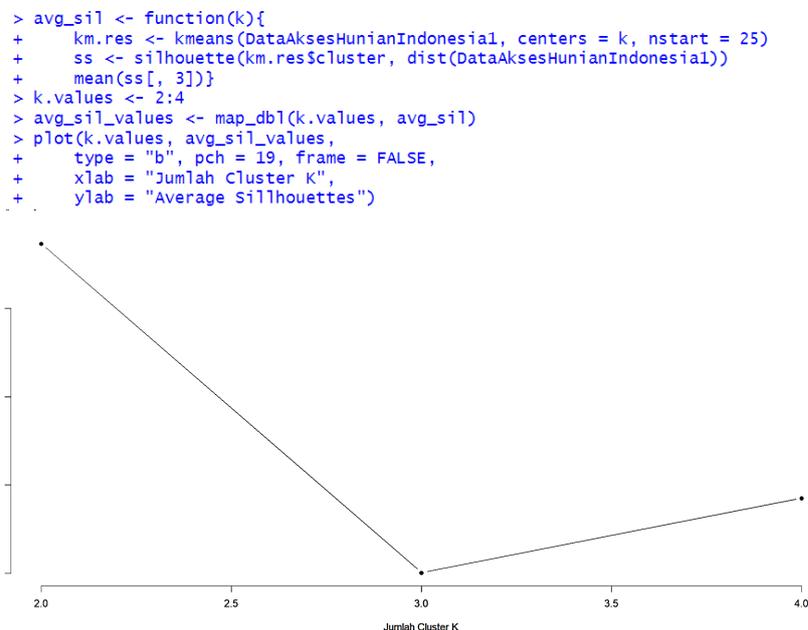


Gambar 8. Visualisasi Metode Elbow

Dengan melihat sisi siku pada grafik elbow, cluster optimal adalah pada Cluster 2.

Menentukan Jumlah Cluster Optimal Menggunakan Metode Average Silhoutte

Silhouette mengacu pada metode interpretasi dan validasi konsistensi dalam kelompok data. Teknik ini memberikan representasi grafis ringkas tentang seberapa baik setiap objek telah diklasifikasikan seperti pada Gambar 9.



Gambar 9. Visualisasi Metode Average Silhouette

Pada grafik Silhouettes, titik optimum berada pada cluster 2. Jadi, berdasarkan grafik Elbow dan grafik Silhouettes jumlah cluster yang optimal untuk data penelitian ini adalah 2 Cluster.

Proses Clustering

Setelah melalui data preparation dan penentuan jumlah cluster, barulah data yang dimiliki kita eksekusi menjadi 2 Cluster seperti pada Gambar 10.

```

> k2 <- kmeans(DataAksesHunianIndonesia1, centers = 2, nstart = 25)
> k2
K-means clustering with 2 clusters of sizes 29, 5

Cluster means:
      2021      2022      2023
1 64.05069 64.11724 66.45724
2 37.43200 36.55800 39.45800

Clustering vector:
Aceh Sumatera Utara Sumatera Barat Riau
1 1 1 1
Jambi Sumatera Selatan Bengkulu Lampung
1 1 1 1
Kep. Bangka Belitung Kep. Riau DKI Jakarta Jawa Barat
2 2 2 1
Jawa Tengah DI Yogyakarta Jawa Timur Banten
1 1 1 1
Bali Nusa Tenggara Barat Nusa Tenggara Timur Kalimantan Barat
1 1 2 1
Kalimantan Tengah Kalimantan Selatan Kalimantan Timur Kalimantan Utara
1 1 1 1
Sulawesi Utara Sulawesi Tengah Sulawesi Selatan Sulawesi Tenggara
1 1 1 1
Gorontalo Sulawesi Barat Maluku Maluku Utara
1 1 1 1
Papua Barat Papua
1 2

Within cluster sum of squares by cluster:
[1] 5031.8765 983.1392
(between_SS / total_SS = 60.9 %)

Available components:
[1] "cluster" "centers" "totss" "withinss" "tot.withinss" "betweenss"
[7] "size" "iter" "ifault"

```

Gambar 10. Hasil Clustering

Visualisasi Hasil Clustering

Peta hasil clustering dapat dilihat dengan jelas melalui grafik pada gambar 11.



Gambar 11. Visualisasi Hasil Clustering

Pengelompokkan Kmeans Clustering

Centroid akhir dari K-Means dapat dilihat pada Tabel 2.

Tabel 2. Centroid Akhir Kmeans

Cluster	2021	2022	2023
C1	64,05069	64,11724	66,45724
C2	37,432	36,558	39,458

Setelah mengetahui hasil centroid akhir K-Means, maka pemodelan sudah terbentuk dan data sudah masuk ke dalam cluster. Pada data penelitian ini, ada 2 Cluster yaitu Cluster 1 jumlah persen terbesar dengan jumlah persentase rumah tangga yang memiliki akses terhadap hunian yang layak dan terjangkau dan cluster 2 yaitu provinsi dengan jumlah persen terkecil dengan jumlah persentase rumah tangga yang memiliki akses terhadap hunian yang layak dan terjangkau. Hasil pengelompokkan K-Means dapat dilihat pada tabel 3.

Tabel 3. Hasil Pengelompokkan Kmeans

Cluster	Provinsi	Kategori
C1	Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Bali, Nusa Tenggara Barat, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat.	Baik
C2	Kep. Bangka Belitung, Kep. Riau, DKI Jakarta, Nusa Tenggara Timur, Papua.	Buruk

Berdasarkan tabel diatas terdapat 29 Provinsi dengan kategori baik pada presentase akses terhadap hunian yang layak dan terjangkau, dan 5 provinsi masuk kedalam kategori buruk terhadap akses hunian yang layak dan terjangkau.

Kesimpulan

Berdasarkan hasil penelitian clustering dengan K-Means yang telah dilakukan, data provinsi-provinsi di Indonesia dibagi menjadi 2 Cluster yaitu provinsi dengan jumlah persen terbesar dengan jumlah persentase rumah tangga yang memiliki akses terhadap hunian yang layak dan terjangkau (C1) dan cluster provinsi dengan jumlah persen terkecil dengan jumlah persentase rumah tangga yang memiliki akses terhadap hunian yang layak dan terjangkau (C2).

Berdasarkan perhitungan K-Means, terdapat provinsi anggota Cluster 1 sebanyak 29 provinsi diantaranya Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Bali, Nusa Tenggara Barat, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, dan Papua Barat. Dimana provinsi yang termasuk bagian dari cluster 1 adalah provinsi yang memiliki setidaknya 66% penduduk yang dapat dengan mudah mengakses hunian yang layak untuk ditinggali dengan harga yang terjangkau.

Kemudian pada Cluster 2, terdapat 5 provinsi diantaranya Kep. Bangka Belitung, Kep. Riau, DKI Jakarta, Nusa Tenggara Timur, dan Papua. Dimana provinsi yang termasuk bagian dari cluster 2 adalah provinsi yang memiliki setidaknya maksimal 38% saja penduduk yang mampu mendapatkan akses terhadap hunian yang layak ditempati dengan harga yang terjangkau. Maka dari itu, pemerintah perlu mengambil tindakan untuk masalah ini agar masyarakat tiap provinsi di Indonesia dapat hidup dengan tempat tinggal yang layak dan terjangkau.

Daftar Rujukan

- [1] Boehmke, B. (t.thn.). *Panduan Pemrograman UC Business Analytics R*. Diambil kembali dari GitHub: https://uc-r.github.io/kmeans_clustering
- [2] Dinda Nabila Batubara, A. P. (2020). Penerapan Datamining Klastering Pada Perusahaan Industri Mikro di Indonesia. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, 331-333.
- [3] Statistik, B. P. (2022). *Persentase Rumah Tangga yang Memiliki Akses Terhadap Hunian Yang Layak Dan Terjangkau Menurut Provinsi (Persen), 2021-2023* yang dapat diakses melalui Diambil kembali dari bps.go.id: <https://www.bps.go.id/id/statistics-table/2/MTI0MSMy/persentase-rumah-tangga-yang-memiliki-akses-terhadap-hunian-yang-layak-dan-terjangkau-menurut-provinsi--persen-.html>.
- [4] Mardhotillah, B., Asyhar, R., & Elisa, E. 2022. Filosofi Keilmuan Statistika Terapan pada Era Smart Society 5.0, *Multi Proximity: Jurnal Statistika. Vol 1. No 2*.
- [5] Ulfa Khaira, S. M. (2023, Oktober 7). *Clustering*. Diambil kembali dari E-Learning Universitas Jambi: <https://elearning.unja.ac.id/course/view.php?id=57678>
- [6] Mardhotillah, B, Elisa, E., Rozi, S. 2022. Implementasi Metode Faktor Ekstraksi dalam Manajemen Anggaran Pemerintah Daerah Dimasa Pandemi Covid 19. *Multi Proximity: Jurnal Statistika. vol 1. No 1*.

- [7] Mardhotillah, B., Fadli, A., Elisa E., & Zurweni. 2023. Indeks Calinski–Harabasz Analisis Fuzzy C–Means dan K–Means Cluster Kabupaten/Kota di Provinsi Jambi Menurut Potensi Pertambangan, Penggalan, Pengadaan Listrik, dan Gas. *Multi Proximity: Jurnal Statistika. vol 2. No 1.*