

PENERAPAN DATA MINING MENGGUNAKAN METODE *K-MEANS* UNTUK *CLUSTER* RUMAH TANGGA DENGAN STATUS KEPEMILIKAN RUMAH SEWA

Muhammad Faisal*¹, Wiranti Sri Utami², Janu Ilham Saputro³

¹Program Studi Teknik Informatika, Fakultas Sains dan Teknologi Universitas Raharja

²Program Studi Teknik Informatika, Fakultas Teknik Universitas Cendekia Abditama

³Program Studi Sistem Komputer, Fakultas Sains dan Teknologi Universitas Raharja

Tinggi E-mail: muhammad.faisal@raharja.info*¹, wirantisutami@uca.ac.id²,

janu@raharja.info³

Abstrak

Memiliki Rumah merupakan Impian setiap orang karena dapat dijadikan tempat tinggal sementara maupun tetap suatu individu atau keluarga. Harga rumah memiliki nilai yang berbeda beda tergantung dari luas bangunan dan tanah maupun letak geografis. Terutama rumah yang berada di jantung Ibu Kota misalnya Jakarta, Jakarta sendiri memiliki harga rumah yang berbeda sesuai dengan kategorinya masing-masing. Memiliki sebuah hunian rumah merupakan impian setiap makhluk hidup khususnya Manusia. Namun dengan pertumbuhan Ekonomi membuat harga Rumah menjadi lebih mahal dari tahun ketahun. Sehingga seseorang yang sudah berkeluarga lebih memilih menyewa rumah atau mengontrak. Dalam penelitian ini penulis akan melakukan cluster menggunakan metode K-Means terhadap dataset Keluarga atau Rumah tangga dengan status kepemilikan rumah sewa atau kontrak. Algoritma K-Means dapat membantu dalam mengkategorikan Provinsi dan Tahun Keluarga atau Rumah tangga dengan status rumah sewa atau kontrak. Dengan dilakukannya penelitian ini menggunakan algoritma K-Means didapatkan 2 cluster dengan jumlah cluster terendah terdapat cluster 29 dan cluster tertinggi terdapat cluster 9 dan hasil validitas menggunakan metode Elbow yaitu 29,9.

Kata Kunci—Rumah Sewa, Data Mining, *Cluster*, K-Means

Abstract

Having a house is everyone's dream because it can be used as a temporary or permanent residence for an individual or family. House prices have different values depending on the size of the building and land as well as geographical location. Especially houses located in the heart of the capital city, for example Jakarta, Jakarta itself has different house prices according to their respective categories. Having a home is the dream of every living thing, especially humans. However, with economic growth, house prices have become more expensive from year to year. So that someone who is married prefers to rent a house or contract. In this study, the author will conduct clusters using the K-Means method on the Family or Household dataset with rental or contract home ownership status. The K-Means algorithm can help in categorizing Provinces and Years of Families or Households with rental or contract home status. By conducting this study using the K-Means algorithm, 2 clusters were obtained with the lowest number of clusters, there were cluster 29 and the highest cluster was cluster 9 and the validity results using the Elbow method were 29.9.

Keywords—Rental House, Data Mining, *Cluster*, K-Means

1. PENDAHULUAN

Rumah merupakan sarana tempat tinggal yang dapat memberikan perlindungan bagi penghuninya atau suatu keluarga yang digunakan sebagai tempat istirahat setelah melakukan kegiatan sehari-hari. Memiliki hunian atau rumah merupakan impian setiap orang khususnya yang telah berumah tangga. Namun sebagian besar warga negara Indonesia yang mengalami kesulitan dalam membeli Hunian atau Rumah, hal ini disebabkan harga rumah selalu naik ditiap tahunnya [1]. Terdapat beberapa faktor lain selain kenaikan harga rumah yang mempengaruhi Gen Z atau seseorang yang telah berkeluarga kesulitan dalam membeli rumah, diantaranya memiliki pendapatan yang tidak memadai dengan kebutuhan pokok menjadi faktor utama bagi seseorang yang telah berkeluarga, Pasar kerja yang kompetitif menjadi salah satu faktor dalam kesulitan membeli rumah yang disebabkan dari kemajuan teknologi dan perubahan ekonomi global menjadikan persaingan kerja semakin ketat, kemudian peraturan Pemerintah disertai dengan kondisi ekonomi global yang memerankan peran penting bagi seseorang untuk membeli rumah, kebijakan tersebut seperti suku bunga Kredit Pemilikan Rumah (KPR) yang sangat mempengaruhi daya beli [2]. Kredit Pemilikan Rumah (KPR) menjadi indikator utama seseorang dalam melakukan pembelian rumah, namun hal ini harus diamati dari segi pendapatan atau penghasilan, semakin tinggi penghasilan akan mempengaruhi daya beli yang semakin tinggi pula namun bila pendapatan rendah maka akan semakin rendah juga daya beli [3].

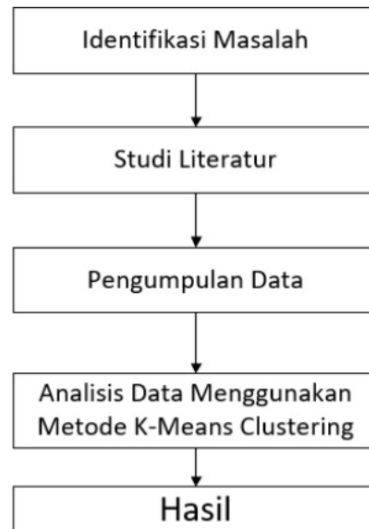
Untuk memenuhi kebutuhan papan atau Rumah seseorang yang sudah berkeluarga atau belum berkeluarga dan belum mampu membeli rumah biasanya akan melakukan Penyewaan Rumah atau mengontrak. Dengan perkembangan zaman maka manusia akan memiliki kebutuhan yang semakin bertambah terutama kebutuhan tempat tinggal, banyak developer property akan berlomba-lomba dalam menyediakan sarana tempat tinggal untuk investasi, Upaya tersebut berdampak terhadap harga rumah yang semakin tinggi [4]. Penelitian terhadap seseorang yang belum memiliki rumah dengan sistem sewa atau mengontrak dapat dilakukan menggunakan *Data mining*, *Data mining* adalah konsep dasar yang digunakan untuk menggali informasi dari sebuah *dataset* dan termasuk dalam *Knowledge Discovery in Database* (KDD) [5]. Dalam *Data mining* terdapat 2 teknik untuk mengolah *database* yaitu *Classification* dan *Clustering*. Pada penelitian ini penulis menggunakan teknik *clustering* dan menggunakan algoritma K-Means. Permasalahan warga negara Indonesia yang belum memiliki rumah menjadi perhatian khusus Pemerintah. Pemerintah Indonesia berupaya memberikan solusi terhadap masyarakat yang berpenghasilan rendah agar mendapatkan rumah yang layak huni dengan menerapkan tabungan perumahan rakyat atau yang disingkat TAPERAN [6]. Namun banyak kalangan yang menolak kebijakan ini salah satunya dari APINDO atau Asosiasi Pengusaha Indonesia, APINDO merasa keberatan dengan kebijakan ini dengan alasan dibebani dengan biaya iuran JAMSOSTEK [7].

Penerapan *Data mining* menggunakan teknik *cluster* dan algoritma K-Means dapat digunakan untuk berbagai penelitian yang berkaitan dengan data, salah satu contohnya adalah *cluster* harga rumah di Jakarta menggunakan algoritma K-Means. Harga rumah pasti berbeda-beda tergantung dari daerah dan kategori masing-masing, sebagai contoh harga rumah di Jakarta yang memiliki perbedaan sesuai dengan kategorinya, maka dibuatlah Penelitian harga rumah di Jakarta menggunakan *clustering* dan algoritma K-Means [8]. K-Means juga dapat digunakan dalam penelitian untuk memprediksi nasabah dengan status kredit macet, pelayanan kredit diberikan oleh lembaga atau badan yang menyalurkan dana kepada seseorang maupun kelompok. Namun pada prakteknya sering terjadi nasabah yang ingin mengajukan *Repeat* order atau pengajuan dana kembali, untuk itu dilakukan penelitian dalam mengklasifikasi kriteria penilaian kelayakan calon nasabah menggunakan Algoritma K-Means [9].

Penerapan *Cluster* K-Means pada penelitian ini bertujuan untuk mencari informasi *cluster* dari *dataset* Penduduk atau keluarga yang kepemilikan rumah dengan status sewa atau kontrak yang diperoleh dari Badan Pusat Statistik (BPS). Dengan adanya penelitian ini diharapkan informasi yang dihasilkan dapat dimanfaatkan oleh Pemerintah atau Pihak terkait.

2. METODE PENELITIAN

Langkah-langkah dalam melakukan penelitian ini akan dirumuskan dalam bentuk kerangka penelitian. Kerangka penelitian merupakan urutan atau langkah-langkah penelitian yang akan menjelaskan tahap demi tahap dalam proses penelitian ini. Berikut ini merupakan kerangka penelitian yang digunakan dalam penelitian ini.



Gambar 1. Kerangka Penelitian

Proses selanjutnya dalam penelitian ini adalah melakukan tahapan analisa, proses tersebut diantaranya sebagai berikut:

2.1. Data Mining

Data mining merupakan upaya untuk mendapatkan informasi dalam himpunan data dalam suatu data yang memiliki *entry* dalam jumlah besar, Tahapan ini antara lain pengumpulan data, ekstrak, analisis, dan statistika [10]. *Data mining* dikaitkan dengan *Knowledge discovery*, *Knowledge extraction*, *data/pattern analysis*, *information harvesting*. *Data mining* dapat diimplementasikan kedalam beberapa penelitian salah satunya dalam melakukan analisa tingkat penjualan barang, *Cluster* atau pengelompokkan data dapat digunakan dalam berbagai penelitian salah satunya digunakan dalam menganalisa Tingkat penjualan barang menggunakan *Data mining* dan algoritma *Cluster K-Means* [11].

2.2. Clustering

Dalam *Data mining* proses *clustering* merupakan suatu proses yang dilakukan untun mendapatkan informasi dan pengetahuan mengenai distribusi data dan mengamati karakteristik setiap *cluster* untuk dianalisa lebih lanjut [12]. *Clustering* merupakan bagian dalam *Data mining*, Pada teknik *Cluster* berbeda dengan Klasifikasi, pada teknik *clustering* tidak memiliki variabel target dalam pengelompokkannya, dalam analisa *clustering* memilih kumpulan data menjadi beberapa kategori dengan kesamaan keunikan atau perhitungan jarak [13].

2.3. K-Means

K-Means merupakan salah satu metode yang digunakan untuk pengelompokkan data non-hierarki dan memiliki kemampuan mempartisi data kedalam dua kelompok maupun lebih [14]. Algoritma K-Means banyak digunakan dalam penelitian untuk menganalisa data yang jumlahnya besar. Salah satu penerapan algoritma K-Means pada penelitian adalah untuk

pengelompokkan UMKM, dalam melakukan analisa pada sektor UMKM maka dapat diperhatikan beberapa hal diantaranya tingkat produktivitas dan inovasi pelaku usaha, kemudahan usaha, akses permodalan, dll. Informasi pengelompokkan UMKM menjadi informasi yang sangat penting untuk mendapatkan informasi menggunakan teknik *Data mining* [15].

Algoritma K-Means digunakan sebagai pengelompokkan data atau *cluster*. *Cluster* yang terbentuk akan meningkatkan kriteria partisi, contohnya adalah perbedaan jarak [16]. Untuk mengukur jarak setiap data yang terbentuk dalam *dataset* dapat menggunakan rumus *Euclidean distance*, rumus tersebut dapat dilihat dibawah ini.

$$d = (b_i, a_t = \sqrt{\sum_j^l (b_{ij} - a_{tj})^2}$$

Keterangan:

d = Jarak nilai data & nilai *centroid*

b_i = nilai data, $i = 1, 2, 3, \dots, n$ = Jumlah data

a_t = nilai *centroid cluster*, $t = 1, 2, 3, \dots, k$ = jumlah *cluster*

l = jumlah atribut

Selain menggunakan rumus *Euclidean distance*, penulis juga menggunakan rumus *Elbow*, penggunaan metode *elbow* untuk memperhatikan nilai perbandingan antara jumlah *cluster* yang akan membentuk siku pada suatu titik, semakin besar jumlah *cluster* k maka nilai (*Sum of Square Error*) akan semakin kecil [17]. Rumus metode *elbow* dapat dilihat dibawah ini.

$$SSE = \sum_k^k = 1 \sum_{x1 \in sk} ||x_1 - c_k||_2^2$$

3. HASIL DAN PEMBAHASAN

Untuk mendukung proses penelitian ini penulis menggunakan Bahasa pemrograman R dengan *tool* atau *software Posit Cloud. Tools* ini digunakan untuk mengelompokkan data menggunakan Algoritma K-Means dan menggunakan *dataset* yang diperoleh dari BPS. Data yang akan dianalisa adalah data Rumah tangga dengan status kepemilikan rumah sewa atau kontrak.

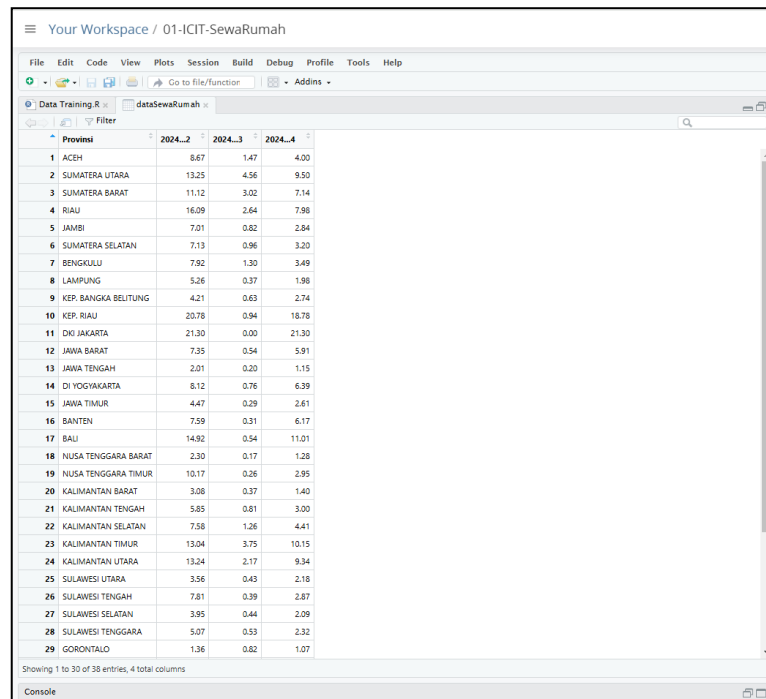
Tabel dibawah ini merupakan *dataset* dari jumlah Rumah tangga disetiap Provinsi yang status kepemilikan rumahnya sewa atau kontrak.

Tabel 1. *Dataset* Rumah tangga status sewa atau kontrak

Provinsi	2024	2024	2024
Aceh	8,67	1,47	4
Sumatera Utara	13,25	4,56	9,5
Sumatera Barat	11,12	3,02	7,14
Riau	16,09	2,64	7,98
Jambi	7,01	0,82	2,84
Sumatera Selatan	7,13	0,96	3,2
Bengkulu	7,92	1,3	3,49
Lampung	5,26	0,37	1,98
Kep. Bangka Belitung	4,21	0,63	2,74

Kep. Riau	20,78	0,94	18,78
DKI Jakarta	21,3	0	21,3
.....
Papua	16,08	1,89	10,44
Papua Selatan	12,01	1,08	5
Papua Tengah	33,62	0,38	7,8
Papua Pegunungan	18,65	0,17	1,53

Selanjutnya *dataset* tersebut akan diimport kedalam *tools Posit Cloud* untuk dilakukan analisa. Gambar tersebut dapat dilihat dibawah ini.



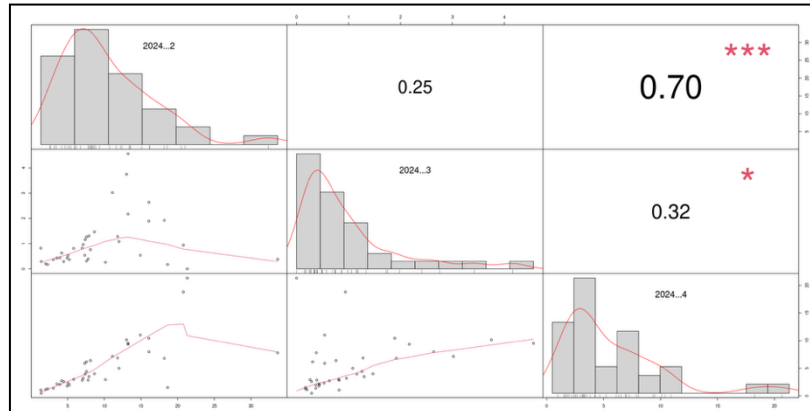
Gambar 2. Import dataset ke Posit Cloud

Selanjutnya data yang akan dianalisa adalah data yang mengandung nilai numerik, maka *syntax* yang digunakan pada Bahasa R dapat dilihat pada gambar dibawah ini.

```
16 index <- dataSewaRumah[,2:4]
17 str(index)
```

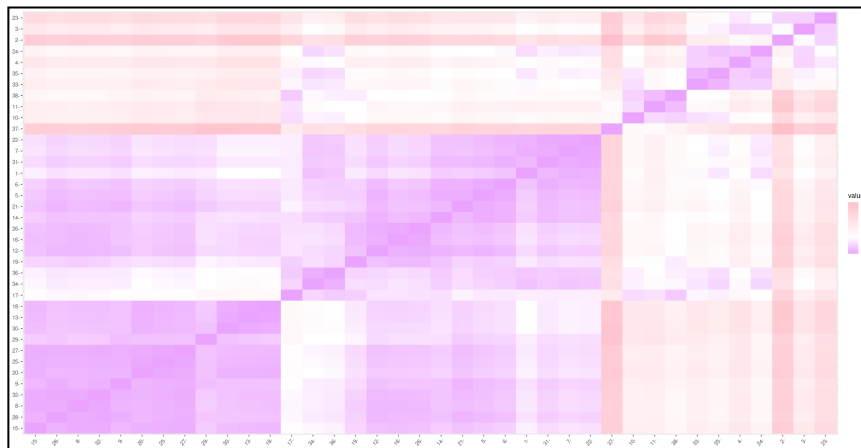
Gambar 3. Mengambil data numerik

Kemudian penulis melakukan korelasi data untuk mencari hubungan antar data yang akan dianalisa, data yang memiliki hubungan akan ditandai dengan tanda **, maka kita akan menggunakan satu variabel saja.



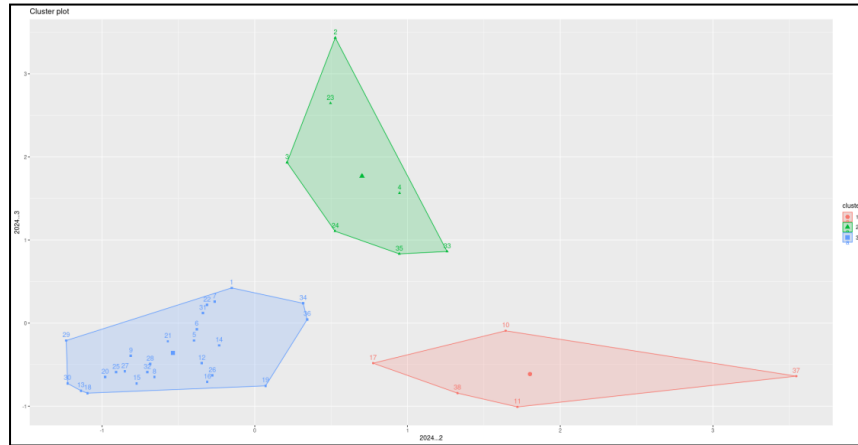
Gambar 4. Diagram korelasi

Langkah berikutnya menentukan ukuran jarak merupakan langkah penting dalam *cluster*. Dalam menentukan jarak penulis menggunakan *Euclidean distance*, Hal ini mendefinisikan bagaimana kesamaan dua elemen antara x,y akan dihitung dan hal ini akan mempengaruhi bentuk *cluster*. Tampilan visualisasi penulis rumuskan kedalam 3 warna agar lebih mudah dimengerti warna magenta menandakan *cluster* sedikit, warna pink menandakan *cluster* tinggi dan warna putih menandakan *cluster* sedang. Matrik *dataset* dapat dilihat pada gambar dibawah ini.



Gambar 5. Tampilan visualisasi matriks jarak

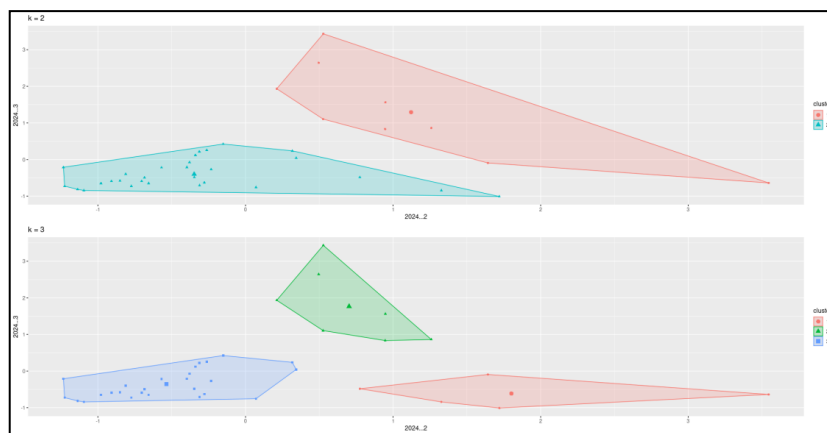
Proses selanjutnya mengelompokkan data menjadi 3 *cluster*, *cluster plot* data Rumah tangga dengan status kepemilikan sewa dapat dilihat pada gambar dibawah ini.



Gambar 6. Cluster plot

Dari data diatas menjelaskan bahwa *cluster* berwarna pink merupakan *cluster* dengan jumlah kepemilikan rumah sewa atau kontrak paling sedikit yaitu Kep. Riau, DKI Jakarta, Bali, Papua Tengah, dan Papua Pegunungan. *Cluster* berwarna Hijau dengan jumlah kepemilikan rumah sewa atau kontrak sedang yaitu Sumatera Utara, Sumatera Barat, Riau, Kalimantan Utara, Papua Barat, dan Papua. *Cluster* terakhir berwarna biru dengan jumlah kepemilikan rumah sewa atau kontrak tinggi.

Penulis mencoba melakukan komparasi bila ingin menggunakan 2 *cluster* dalam penelitian ini. Penentuan *cluster* subyektif tergantung pada penelitian, maka penulis mencoba menggunakan 2 *cluster*.

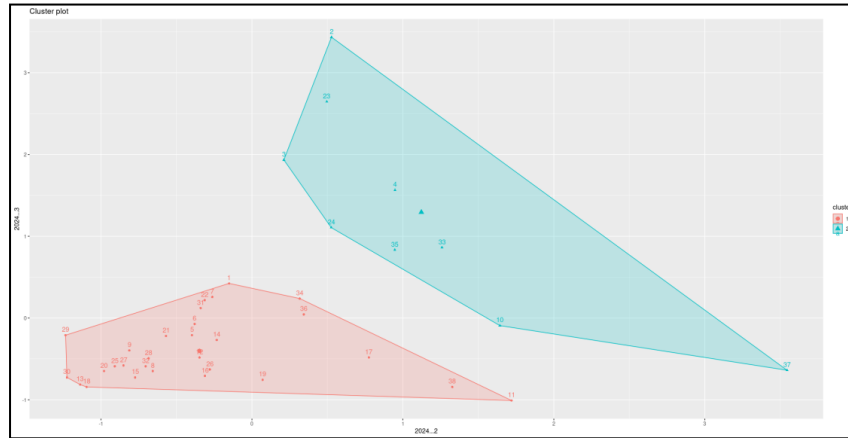


Gambar 7. Cluster plot 2 cluster

Meskipun tampilan nilai visualisasi ini dapat memberikan informasi dimana dilineasi yang sebenarnya terjadi atau tidak terjadi antar *cluster*, hal tersebut belum memberikan informasi berapa jumlah *cluster* yang optimal.

3.2. Cluster K-Means

Dalam menentukan analisis *cluster* menggunakan metode K-Means untuk menentukan *cluster* optimal dari dataset yang akan diteliti penulis akan menggunakan Metode *elbow*, metode *elbow* bekerja dengan memvisualisasikan grafik yang dikenal dengan grafis siku yang akan memetakan nilai dari SSE (*Sum of Square Errors*). Metode akan diproses menggunakan



Gambar 11. Diagram Cluster Plot

Dari total Provinsi 38 Provinsi yang dianalisa maka diperoleh 2 cluster, *cluster 1* terdapat 28 Provinsi dengan Rumah tangga yang status kepemilikan sewa atau kontrak, dan *cluster 2* terdapat 9 Provinsi dengan Rumah tangga yang status kepemilikan sewa atau kontrak.

Dari hasil *cluster plot* diatas kemudian akan ditampilkan dalam bentuk daftar keseluruhan hasil *cluster* Rumah tangga yang status kepemilikan sewa atau kontrak menggunakan metode K-Means.

	Provinsi	X2024...2	X2024...3	X2024...4	final.cluster
1	ACEH	8.67	1.47	4.00	1
2	SUMATERA UTARA	13.25	4.56	9.50	2
3	SUMATERA BARAT	11.12	3.02	7.14	2
4	RIAU	16.09	2.64	7.98	2
5	JAMBI	7.01	0.82	2.84	1
6	SUMATERA SELATAN	7.13	0.96	3.20	1
7	BENGKULU	7.92	1.30	3.49	1
8	LAMPUNG	5.26	0.37	1.98	1
9	KEP. BANGKA BELITUNG	4.21	0.63	2.74	1
10	KEP. RIAU	20.78	0.94	18.78	2
11	DKI JAKARTA	21.30	0.00	21.30	1
12	JAWA BARAT	7.35	0.54	5.91	1
13	JAWA TENGAH	2.01	0.20	1.15	1
14	DI YOGYAKARTA	8.12	0.76	6.39	1
15	JAWA TIMUR	4.47	0.29	2.61	1
16	BANTEN	7.59	0.31	6.17	1
17	BALI	14.92	0.54	11.01	1
18	NUSA TENGGARA BARAT	2.30	0.17	1.28	1
19	NUSA TENGGARA TIMUR	10.17	0.26	2.95	1
20	KALIMANTAN BARAT	3.08	0.37	1.40	1
21	KALIMANTAN TENGAH	5.85	0.81	3.00	1
22	KALIMANTAN SELATAN	7.58	1.26	4.41	1
23	KALIMANTAN TIMUR	13.04	3.75	10.15	2
24	KALIMANTAN UTARA	13.24	2.17	9.34	2
25	SULAWESI UTARA	3.56	0.43	2.18	1
26	SULAWESI TENGAH	7.81	0.39	2.87	1
27	SULAWESI SELATAN	3.95	0.44	2.09	1
28	SULAWESI TENGGARA	5.07	0.53	2.32	1
29	GORONTALO	1.36	0.82	1.07	1
30	SULAWESI BARAT	1.42	0.29	0.51	1
31	MALUKU	7.40	1.16	3.96	1
32	MALUKU UTARA	4.94	0.43	1.80	1
33	PAPUA BARAT	18.19	1.92	6.80	2
34	PAPUA BARAT DAYA	11.83	1.28	6.86	1
35	PAPUA	16.08	1.89	10.44	2
36	PAPUA SELATAN	12.01	1.08	5.00	1
37	PAPUA TENGAH	33.62	0.38	7.80	2
38	PAPUA PEGUNUNGAN	18.65	0.17	1.53	1

Gambar 12. Hasil akhir cluster

4. KESIMPULAN

Penerapan Algoritma *cluster* K-Means dalam melakukan analisa Rumah tangga dengan status kepemilikan rumah sewa atau kontrak diharapkan dapat menjadi informasi untuk pihak terkait dalam mendata Penduduk yang kepemilikan rumah dengan status sewa atau kontrak. Dalam penggunaan algoritma *cluster* K-Means memiliki pengaruh besar terhadap temuan jumlah *cluster* dengan menggunakan metode *elbow*.

5. SARAN

Penulis berhadapan penelitian ini digunakan sebagai rujukan untuk refrensi penelitan yang saling berkaitan. Harapan penulis Pemerintah dapat memberikan program KPR yang sesuai dengan kemampuan masyarakat demi terwujudnya harapan memiliki rumah impian.

DAFTAR PUSTAKA

- [1] Y. Angelica, "Kenapa Gen Z Kesulitan Membeli Rumah? Ini Faktor Penyebabnya," *btnproperti*, 24 September 2024. [Online]. Available: <https://www.btnproperti.co.id/blog/gen-z-tidak-bisa-beli-rumah>. [Accessed 25 Februari 2025].
- [2] T. Siagian, "4 Penyebab Gen Z Sulit untuk Membeli Rumah, Apakah Benar?," *IDN TIMES*, 02 Oktober 2024. [Online]. Available: <https://www.idntimes.com/life/inspiration/theodore-siagian/penyebab-gen-z-sulit-beli-rumah-c1c2?page=all>. [Accessed 25 Februari 2025].
- [3] I. M. Roeskamto and A. S. Atmadja, "PERGERAKAN HARGA RUMAH DI INDONESIA: ANALISIS SISI PERMINTAAN," *Dimensi Utama Teknik Sipil*, vol. 7, no. 2, pp. 53-59, 2020.
- [4] E. F. Rahayuningtyas, F. N. Rahayu and Y. Azhar, "Prediksi Harga Rumah Menggunakan General Regression Neural Network," *JURNAL INFORMATIKA*, vol. 8, no. 1, pp. 59-66, 2021.
- [5] M. Faisal, W. S. Utami, D. Prastyo and J. I. Saputro, "Penerapan Data Mining Menggunakan Metode K-Means Untuk Clustering Tindak Pidana Daerah (Studi Kasus: Data BPS 2018, 2019, 2020)," *ICIT Journal*, vol. 11, no. 1, pp. 79-88, 2025.
- [6] R. Abidin and A. Herawati, "ANALISIS SENTIMEN PUBLIK TERHADAP KEBIJAKAN PROGRAM TABUNGAN PERUMAHAN RAKYAT (TAPERAN)," *Journal of Information System and Computer*, vol. 4, no. 1, pp. 13-19, 2024.
- [7] H. Retnaningsih, "PENOLAKAN IURAN TAPERAN OLEH PEKERJA DAN PENGUSAHA," 11 Juni 2024. [Online]. Available: https://berkas.dpr.go.id/pusaka/files/isu_sepekan/Isu%20Sepekan---II-PUSLIT-Juni-2024-175.pdf. [Accessed 26 Februari 2025].
- [8] Z. F. Daldiri, M. Rafly and I. Veritawati, "Clustering Daftar Harga Rumah di Jakarta Dengan Algoritma K-Means," *Journal of Informatics and Advanced Computing (JIAC)*, vol. 3, no. 2, pp. 155-160, 2022.
- [9] M. N. R. Fitriani, B. Priyatna, B. Huda, A. L. Hananto and T. , "Implementasi Metode K-Means Untuk Memprediksi Status Kredit Macet," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 3, pp. 554-561, 2023.
- [10] A. W. S, I. G. I. Sudipa, T. A. E. Putra, A. J. Wahidin, W. A. Syukrilla, A. K. Wardhani, N. Heryana, T. Indriyani and L. W. Santoso, "Data Mining," in *DATA MINING DAN KNOWLEDGE DISCOVERY PROCESS*, Padang Sumatera Barat, PT GLOBAL EKSEKUTIF TEKNOLOGI, 2023, pp. 1-15.
- [11] D. N. Yoliadi, "Data mining Dalam Analisis Tingkat Penjualan Barang Elektronik Menggunakan Algoritma K-means," *Insearch*, vol. 3, no. 1, pp. 43-53, 2023.

- [12] R. R. Muhima, M. Kurniawan, S. R. Wardhana, A. Yudhana, S. W. M. Rahmawati and G. E. Yuliasuti, "KUPASTUNTAS ALGORITMA CLUSTERING KONSEP PERHITUNGAN MANUAL, DAN PROGRAM," in *PENGANTAR ANALISIS CLUSTERING*, Surabaya, ANDI, 2021, pp. 1-4.
- [13] T. Hardiani, "ANALISIS CLUSTERING KASUS COVID 19 DI INDONESIA MENGGUNAKAN ALGORITMA K-MEANS," *JANAPATI*, vol. 11, no. 2, pp. 156-165, 2022.
- [14] A. Nugraha, O. Nurdiawan and G. Dwilestari, "PENERAPAN DATA MINING METODE K-MEANS CLUSTERING UNTUK ANALISA PENJUALAN PADA TOKO YANA SPORT," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 6, no. 2, pp. 849-855, 2022.
- [15] W. Sudrajat, I. Cholid and J. Petrus, "Penerapan Algoritma K-Means Clustering untuk Pengelompokan UMKM Menggunakan Rapidminer," *Jurnal Jupiter*, vol. 14, no. 1, pp. 27-36, 2022.
- [16] M. Faisal, W. S. Utami, D. Prastyo and J. I. Saputro, "Penerapan Data Mining Menggunakan Metode K-Means Untuk ClusteringTindak Pidana Daerah (Studi Kasus:Data BPS 2018, 2019, 2020)," *ICITJournal*, vol. 11, no. 1, pp. 79-88, 2025.
- [17] N. T. Hartanti, "Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 6, no. 2, pp. 82-89, 2020.