ZONAsi Jurnal Sistem Informasi
ZONAsi VOL. 8 NO. 1 Jan 2026
Page: 372 - 382 Accreditation Sinta 4
ISSN: 2656-7407 (Online)
2656-7393 (Print)

PENERAPAN *DATA MINING* MENGGUNAKAN ALGORITMA *K-MEANS* UNTUK CLSUTERING PESERTA KELUARGA BERENCANA AKTIF DI KABUPATEN TIMOR TENGAH UTARA

Venansius Apriano Arik[1], Yoseph Pius Kurniawan Kelen[2], Budiman Baso[3]
[1,2,3] Universitas Timor
(Program Studi Teknologi Informasi Fakultas Pertanian Sains dan Kesehatan)
(Jl. Eltari KM. 15 Kefamenanu, Timor Tengah Utara, Nusa Tenggara Timur)
e-mail: [1]arik19042002@gmail.com, [2] yosepkelen@unimor.ac.id, [3]budimanbaso@gmail.com

Abstrak

Penelitian ini dilatarbelakangi oleh kebutuhan pemerintah daerah dalam memahami pola penggunaan kontrasepsi pada peserta program Keluarga Berencana (KB) aktif secara lebih sistematis dan berbasis data. Data peserta KB yang cukup besar serta memiliki karakteristik yang beragam memerlukan metode pengolahan yang mampu mengelompokkan data ke dalam kelompok-kelompok yang memiliki kemiripan. Dengan demikian, hasil pengelompokan tersebut dapat membantu dalam penyusunan kebijakan serta perencanaan program KB yang lebih tepat sasaran. Penelitian ini bertujuan untuk melakukan pengelompokan (clustering) peserta KB aktif di Kabupaten Timor Tengah Utara dengan menerapkan algoritma K-Means sebagai salah satu metode dalam data mining. Metode penelitian yang digunakan meliputi beberapa tahapan, yaitu pengumpulan data, preprocessing data, dan proses clustering menggunakan algoritma K-Means. Tahap preprocessing dilakukan untuk memastikan data siap dianalisis, yang meliputi proses pembersihan data (data cleaning), transformasi data, serta penyesuaian format agar sesuai dengan kebutuhan pemodelan. Selanjutnya, proses clustering dilakukan menggunakan bahasa pemrograman Python yang memiliki berbagai pustaka yang mendukung analisis data secara efisien. Hasil penelitian menunjukkan bahwa algoritma K-Means mampu mengelompokkan peserta KB aktif ke dalam empat cluster utama berdasarkan jenis kontrasepsi yang digunakan. Dari hasil pengelompokan tersebut diketahui bahwa metode kontrasepsi suntikan dan implan merupakan jenis yang paling dominan digunakan oleh peserta KB aktif di Kabupaten Timor Tengah Utara. Setiap cluster memiliki karakteristik penggunaan kontrasepsi yang berbeda, sehingga dapat memberikan informasi penting bagi instansi terkait dalam memahami pola penggunaan KB di masyarakat. Temuan penelitian ini diharapkan dapat mendukung pengambilan keputusan berbasis data serta membantu pemerintah daerah dalam merancang strategi dan kebijakan program KB yang lebih efektif dan tepat sasaran.

Kata Kunci*: Data Mining, Clustering, Peserta KB Aktif, K-Means, Python.*

Abstract

The study is motivated by the need of the local government to better understand patterns of contraceptive use among active Family Planning (KB) program participants in a more systematic and data-driven manner. The large amount of KB participant data with diverse characteristics requires a data processing method capable of grouping participants into clusters with similar characteristics. Such grouping can assist in designing more targeted policies and planning more effective family planning programs. The purpose of this study is to cluster active KB participants in Timor Tengah Utara Regency by applying the K-Means algorithm as one of the methods in data mining. The research method consists of several stages, including data collection, data preprocessing, and clustering using the K-Means algorithm. The preprocessing stage is carried out to ensure that the data is ready for analysis, including data cleaning, data transformation, and format adjustment to meet the requirements of the modeling process. The clustering process is performed using the Python programming language, which provides various libraries that support efficient data processing and data mining analysis. The results of the study show that the K-Means algorithm is able to group active KB participants into four main clusters based on the type of contraceptive methods used. The clustering results indicate that injectable and implant contraceptive methods are the most dominantly used by

active KB participants in Timor Tengah Utara Regency. Each cluster shows different characteristics of contraceptive use, providing valuable information for related institutions in understanding family planning patterns within the community. The findings of this research are expected to support data-based decision making and assist local governments in developing more effective and targeted family planning policies and strategies.

Keywords: *Data Mining, Clustering, Aktiv KB Participants , K-Means, Python*.

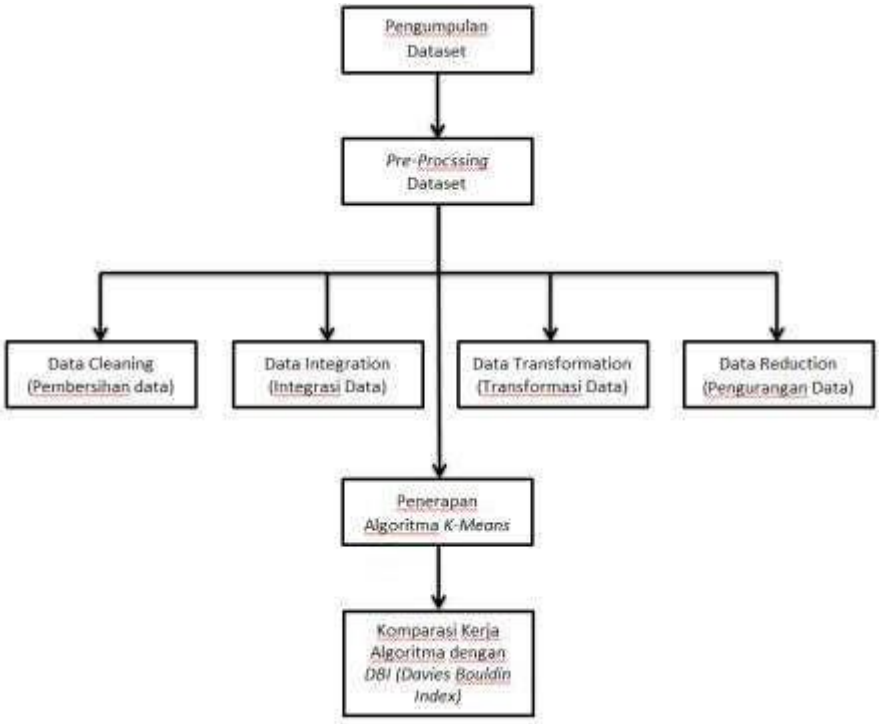
1. PENDAHULUAN

Program keluarga berencana (KB) adalah upaya untuk mewujudkan keluarga berkualitas melalui promosi, perlindungan, dan bantuan dalam mewujudkan hak-hak reproduksi serta penyelenggaraan pelayanan, pengaturan dan dukungan yang diperlukan untuk membentuk keluarga dengan usia kawin yang ideal. Program keluarga berencana juga mengatur jumlah Keluarga Berencana (KB) merupakan program penting yang bertujuan untuk mengendalikan pertumbuhan populasi dan meningkatkan kualitas hidup masyarakat. Dalam konteks ini, pemahaman tentang karakteristik peserta KB aktif sangat krusial untuk merumuskan kebijakan yang lebih efektif dan tepat sasaran. Salah satu pendekatan yang dapat digunakan untuk memahami karakteristik peserta KB adalah melalui teknik clustering, yang memungkinkan pengelompokan individu berdasarkan kesamaan dalam atribut tertentu, seperti jenis kontrasepsi yang digunakan. Clustering merupakan metode analisis data yang mengelompokkan objek berdasarkan ke smiripan karakteristik tanpa memerlukan label sebelumnya. Dalam penelitian ini, teknik clustering diterapkan untuk mengelompokkan peserta KB aktif berdasarkan jenis KB yang mereka pilih. Dengan menggunakan algoritma seperti *K- Means*, data peserta KB dapat dikelompokkan menjadi beberapa cluster, yang masing-masing mencerminkan karakteristik demografis dan preferensi kontrasepsi yang berbeda. Pentingnya penelitian ini terletak pada kemampuannya untuk memberikan wawasan yang lebih mendalam tentang kebutuhan dan preferensi peserta KB aktif. Dengan mengidentifikasi kelompok- kelompok ini, pemerintah dan pemangku kepentingan dapat merancang intervensi yang lebih spesifik dan efektif. Misalnya, peserta dalam kelompok pengguna jenis KB tertentu mungkin memiliki preferensi yang berbeda dengan kelompok pengguna jenis KB lainnya. Dengan demikian, analisis clustering dapat membantu dalam merumuskan kebijakan yang lebih responsif terhadap kebutuhan masyarakat. Selain itu, pemanfaatan data minig dan teknik analisis data yang canggih seperti clustering menjadi semakin relevan dalam konteks kesehatan masyarakat. Dengan kemajuan teknologi dan ketersediaan data yang melimpah, teknik ini dapat digunakan untuk mengidentifikasi unmet needs dalam program KB, sehingga dapat meningkatkan partisipasi dan efektivitas program tersebut. Melalui penelitian ini, diharapkan dapat diperoleh pemahaman yang lebih baik mengenai karakteristik peserta KB aktif, yang pada gilirannya dapat mendukung pengembangan strategi dan kebijakan yang lebih efektif dalam program keluarga berencana (Wardani, 2023).

Pada penelitian ini analisa *data mining* dilakukan dengan menggunakan algoritma *K- Means*. Dengan menggunakan metode ini, data-data yang telah didapatkan dapat dikelompokan ke dalam beberapa cluster berdasarkan kemiripan dari data-data tersebut, sehingga data-data yang memiliki karakteristik yang sama dikelompokan dalam satu cluster dan yang memliki karakteristik yang berbeda dikelompokan dalam cluster yang lain yang memiliki karakteristik yang sama (Aulia, 2021). *K-means* merupakan salah satu metode *clustering* non-hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*. Metode ini mempartisi data ke dalam cluster sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karateristik yang berbeda dikelompokan ke dalam cluster yang lain. *K-Means* adalah suatu metode penganalisaan data atau metode *Data Mining* yang melakukan proses pemodelan tanpa supervisi (unsupervised) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Terdapat dua jenis data clustering yang sering dipergunakan dalam proses pengelompokan data yaitu Hierarchical dan Non- Hierarchical, dan *K-Means* merupakan salah satu metode data clustering nonhierarchical atau Partitional Clustering. Metode *K-Means* Clustering berusaha mengelompokkan data yang ada
ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain (Hasyim, 2023).

2. METODE PENELITIAN

K-means merupakan salah satu algoritma yang bersifat unsupervised learning. Unsupervised learning adalah penggunaan algoritma kecerdasan buatan atau *artificial intelligence* (AI) untuk mengidentifikasi pola dalam kumpulan data yang berisi titik data yang tidak diklasifikasikan atau diberi label. *K-Means* memiliki fungsi untuk mengelompokkan data kedalam data cluster. Algoritma ini dapat menerima data tanpa ada label kategori. Algoritma *Clustering K-Means* juga merupakan metode non-hierarchy. Metode Algoritma *Clustering* adalah mengelompokkan beberapa data ke dalam kelompok yang menjelaskan data dalam satu kelompok memiliki karakteristik yang sama dan memiliki karakteristik yang berbeda dengan data yang ada di kelompok lain. Ada beberapa tahapan yang akan dilakukan dalam penelitian clustering peserta keluarga berencana aktif berdasarkan penggunaan jenis kontrasepsi menggunakan Algoritma K-Means yang dapat dilihat sebagai berikut :

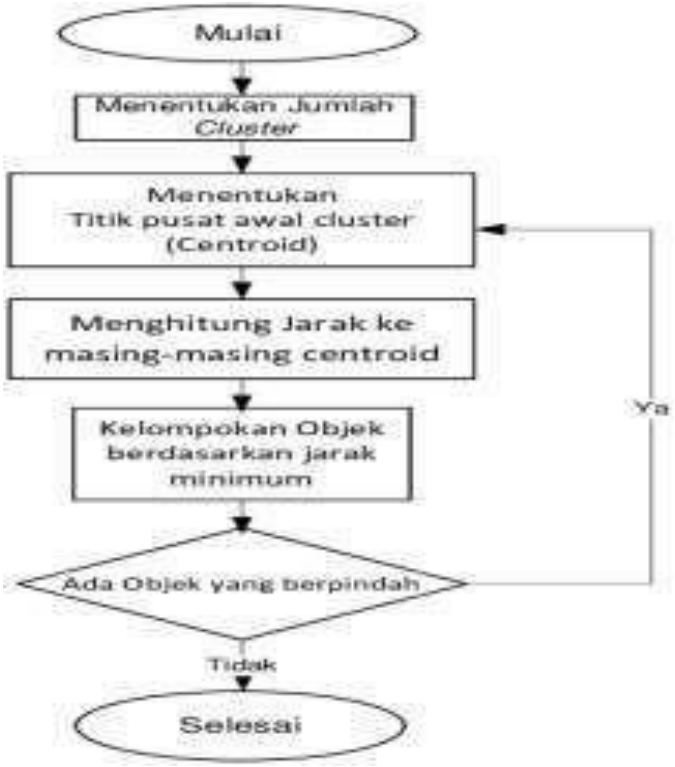


Gambar 1. Tahapan Penelitian Clustering

Tahapan ini diawali dengan tahapan pengumpulan dataset. Dataset yang digunakan adalah dataset yang dikumpulkan di BKKBN Kabupaten Timor Tengah Utara. Masuk pada tahap *Pre-Processing* dataset adalah proses yang mengubah data mentah ke dalam bentuk yang digunakan sebagai masukan Algoritma. Proses ini penting dilakukan karena data mentah sering kali tidak memiliki format yang teratur. Selain itu Algoritma *K-Means Clustering* juga tidak dapat memproses data mentah sehingga proses ini penting dilakukan sehingga mempermudah proses berikutnya. Berikut adalah empat tahap *Pre-Processing* dataset : *Data Cleaning (Pembersihan Data*) Tahap *Pre-Processing.*

Pertama adalah data cleaning. Tahap ini merupakan proses identifikasi dan perbaikan masalah dalam data. Misalnya, kesalahan input, nilai hilang dan duplikasi. Tujuan utama dari tahap ini adalah untuk menciptakan dataset yang konsisten dan akurat. ***Data Integration (Integrasi Data)*** Kedua, tahap *Pre-Processing* data adalah data integration. Hal ini melibatkan penggabungan data dari berbagai sumber menjadi satu dataset yang kohesif. Tahapan ini penting dilakukan ketika informasi dikumpulkan dari berbagai database atau saat bekerja dengan dataset besar yang tersebar. ***Data Transformation (Transformasi Data)*** Tahap *Pre-Processing* data berikutnya adalah data transformation. Tahap ini adalah proses mengubah data ke dalam format atau struktur yang lebih sesuai

untuk analisis. ***Data Reduction (Pengurangan Data)*** Tahap *Pre- Processing* data terakhir adalah data reduction. Ini merupakan proses pengurangan data untuk mengurangi volume data tanpa menhilangkan informasi penting. Hal ini bisa mempercepat pemrosesan dan analisis. Berikut adalah tahap Penerapan Algoritma *K-Means* Langkah-langkah algoritma *K-Means* dalam penelitian ini antara lain :



Gambar 2. Tahapan Algoritma *K-Means*

1. Menentukan nilai (k) sebagai jumlah cluster yang dibentuk.
2. Menentukan centroid sebagai titik pusat cluster awal secara acak pada dataset peserta Keluarga Berencana.
3. Menghitung jarak antara pusat cluster dan seluruh *intances* pada dataset Peserta Keluarga Berencana menggunakan rumus *Eucliedence Distance* sebagai berikut :

$$d(x,y) = \sqrt{(\sum_{i=1}^{n}(xi - yi)^2)}$$

Keterangan :
xi = Pusat titik centroid
yi = Data *Instances*
i = Atribut atau variable data peserta keluarga berencana

4. Mengalokasikan masing-masing objek dalam centroid terdekat.
5. Melakukan iterasi serta mendapatkan centroid baru menggunakan rumus berikut :

$$v = \frac{\sum_{i=1}^{n} x_i}{n} \; ; 1,2,3, \ldots n$$

Keterangan :
xi = Pusat titik centroid
n = Jumlah data peserta Keluarga Berencana
i = Atribut data peserta Keluarga berencana

6. Mengulangi langkah ketiga jika centroid baru tidak sama dengan centroid lama.

Tahap berikut pada *Pre-Processing* data adalah Komparasi Kerja Algoritma Menggunakan DBI *(Davies Bouldin Index).* Komparasi kinerja dari klusterisasi *K-Means Clustering* menggunakan DBI *(Davies Bouldin Index)* sebagai metode evaluasinya. Metode evaluasi DBI *(Davies Bouldin Index)* memiliki kelebihan dalam hal menguukur evaluasi *cluster* pada suatu metode pengelompokan disebabkan nilai kohesi dan separasinya. Kohesi didefenisikan sebagai jumlah dari kedekatan data terhadap centroid dari *cluster* yang diikuti dikenal sebagai *Sum Of Square Within cluster* (SSW). Sedangkan, separasi didasarkan pada jarak antar centroid dari *cluster*nya, dikenal sebagai *Sum Of Square Between cluster* (SSB). Melalui nilai kohesi dan separasi inilah yang menjadikan semakin kecil nilai DBI *(Davies Bouldin Index)* yang didapatkan (Mendekati 0 atau sama dengan 0) maka menunjukan semakin baik *cluster* dari hasil pengelompokan. Perhitungan DBI *(Davies Bouldin Index)* diawali dengan mencari nilai melalui perhitungan nilai *Sum Of Square Wihtin* (SSW), yaitu untuk mengetahui matrik kohesi dalam sebuah *cluster* ke-i yang dirumuskan sebagai berikut :

$$SSW_i = \frac{1}{m_i}\sum_{j=1}^{m_i} d(x_j, c_i)$$

Keterangan :
SSW : *Sum of Square Within cluster*
mi : Jumlah data dalam *cluster* ke-*i*
ci : Centroid *cluster* ke-*i*
d (X*j*, C*i*) : Jarak dari data ke-*i* ke titik *cluster i*
Berikutnya, dilakukan langkah selanjutnya menggunkan rumus *SSB (Sum of Square Between)* untuk mengetahui separasi antar *cluster :*

$$SSB_{ij} = d(i, j)$$

Keterangan :
SSB : *Sum of Square Between cluster*
d(i, j): Jarak *Euclidience Distance* data ke *i* dan data ke *j*
Selanjutnya, yaitu mencari rasio dari hasil perhitugan *SSW* dan *SSB* dengan rumus berikut :

$$R_{ij} = \frac{SSW_i + SSW_j}{SSB_{ij}}$$

Keterangan :
Rij : Rasio antar *cluster*
Untuk proses perhitungan tahap akhir menggunakan rumus berikut untuk memperoleh nilai DBI.

$$DBI = \frac{1}{k}\sum_{i=1}^{k} \max i \neq j \, (R_{i,j})$$

Keterangan :
k : Jumlah *cluster* yang dihitung
Ri,j : Rasio antar *cluster i* dan *j*
max : Dicari rasio antar *cluster* yang terbesar
Semakin kecil nilai yang didapatkan, yakni nilai yang mendekati 0 atau sama dengan 0 maka menunjukan semakin baik *cluster* dari hasil pengelompokan. Nantinya nilai DBI didapatkan dari algoritma *K-Means.* Nilai DBI dikomparasi sebagai evaluasi kinerja dari *Clustering.*

3. HASIL DAN PEMBAHASAN

3.1. Hasil Clustering *K-Means* dengan *python*

Metode yang digunakan dalam permasalahan adalah clustering, proses awal yang dilakukan dalam pembentukan cluster adalah transformasi data kedalam bentuk numerik dengan kode-kode yang telah di tentukan jumlah group (K), hitung centroid, hitung jarak ke centroid dan kemudian groubkan berdasarkan jarak terdekat, jika tidak ada objek yang berpindah group maka itersi selesai.

Tabel 1. Data Peserta Keluraga Berencana Aktif

Nama	Umur	Kecamatan	Jenis KB	Status Peserta	Informed _Consent	Jenis_ Tindakan	Penggunaan _ Asuransi
P1	24	MIOMAFO TIMUR	SUNTIKAN	Peserta KB Ulangan	Tanpa Informed Consent	Operatif/Pemberian/Pemasangan	BPJS Kesehatan
P2	35	MIOMAFO TIMUR	IMPLAN	Peserta KB Ulangan	Tanpa Informed Consent	Operatif/Pemberian/Pemasangan	BPJS Kesehatan
P3	21	MIOMAFO TIMUR	SUNTIKAN	Peserta KB Ulangan	Tanpa Informed Consent	Operatif/Pemberian/Pemasangan	BPJS Kesehatan
P4	30	MIOMAFO TIMUR	SUNTIKAN	Peserta KB Ganti Cara	Informed Consent	Operatif/Pemberian/Pemasangan	BPJS Kesehatan
P5	42	MIOMAFO TIMUR	IMPLAN	Peserta KB Ulangan		Operatif/Pemberian/Pemasangan	BPJS Kesehatan
P6	40	MIOMAFO TIMUR	SUNT I KAN	Peserta KB Baru		Operatif/Pemberian/Pemasangan	BPJS Kesehatan
P7	35	MIOMAFO TIMUR	IMPLAN	Peserta KB Baru		Operatif/Pemberian/Pemasangan	BPJS Kesehatan
…..	…..	….	….	….		…..	…..
P50 0	33	MIOMAFO TIMUR	IMPLAN	Peserta KB Baru	Informed Consent	Operatif/Pemberian/Pemasangan	BPJS Kesehatan

1. Data *Drop*

Dapat dilihat bahwa atribut atau fitur yang digunakan untuk *clustering* terdapat empat kolom yang memiliki tipe data *numerik* dan *teks* yaitu umur, jenis kb, kecamatan dan status peserta.

	UMUR	KECAMATAN	JENIS_KB	STATUS_PESERTA
0	24	MIOMAFO TIMUR	SUNTIKAN	Peserta KB Ulangan
1	35	MIOMAFO TIMUR	IMPLAN	Peserta KB Ulangan
2	21	MIOMAFO TIMUR	SUNTIKAN	Peserta KB Ulangan
3	30	MIOMAFO TIMUR	SUNTIKAN	Peserta KB Ganti Cara
4	42	MIOMAFO TIMUR	IMPLAN	Peserta KB Ulangan
...	...	...	...	...
495	20	BIBOKI FOETLEU	SUNTIKAN	Peserta KB Baru
496	33	BIBOKI FOETLEU	SUNTIKAN	Peserta KB Ulangan
497	32	BIBOKI FOETLEU	SUNTIKAN	Peserta KB Ulangan
498	34	BIBOKI FOETLEU	SUNTIKAN	Peserta KB Baru
499	50	BIBOKI FOETLEU	SUNTIKAN	Peserta KB Baru

500 rows × 4 columns

Gambar 3. Proses Data *Drop*

2. Data *Cleaning*

Tahapan ini menggunakan operator *Replace missing value* di *python* yang mana untuk mengetahui apakah data yang ada terdapat data kosong maupun missing. Hal ini dihasilkan bahwa data yang ada tidak berisi data yang kosong atau missing. Tampilan seperti gambar di bawah menunjukkan posisi nilai kosong dari setiap data dan masing-masing atribut yang digunakan. Terdapat empat kolom atribut yang digunakan yaitu umur, kecamatan, jenis kb dan status peserta dan juga 500 baris data yang dihitung mulai dari index 0 sampai 499 bernilai *false* dimana posisi setiap data bernilai dan tidak kosong. Jika pada kolom atau baris ada yang bernilai *true* maka data pada baris dan kolom tersebut tidak memiliki nilai atau kosong.

	UMUR	KECAMATAN	JENIS_KB	STATUS_PESERTA
0	24	MIOMAFO TIMUR	SUNTIKAN	Peserta KB Ulangan
1	35	MIOMAFO TIMUR	IMPLAN	Peserta KB Ulangan
2	21	MIOMAFO TIMUR	SUNTIKAN	Peserta KB Ulangan
3	30	MIOMAFO TIMUR	SUNTIKAN	Peserta KB Ganti Cara
4	42	MIOMAFO TIMUR	IMPLAN	Peserta KB Ulangan
...	...	...	...	...
495	20	BIBOKI FOETLEU	SUNTIKAN	Peserta KB Baru
496	33	BIBOKI FOETLEU	SUNTIKAN	Peserta KB Ulangan
497	32	BIBOKI FOETLEU	SUNTIKAN	Peserta KB Ulangan
498	34	BIBOKI FOETLEU	SUNTIKAN	Peserta KB Baru
499	50	BIBOKI FOETLEU	SUNTIKAN	Peserta KB Baru

500 rows × 4 columns

Gambar 4. Proses data *Cleaning*

3. Data Transformasi

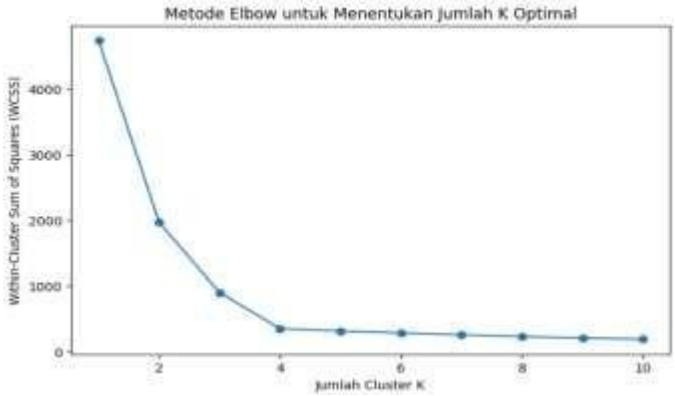
Setelah melewati tahap data *cleaning*, selanjutnya data akan ditransformasikan sesuai dengan ketetapan proses data mining. Yang dimana data tidak dapat diproses jika variabel data bukan numerik. Disini user menggunakan operator *Kategorik To Numeric* untuk mengubah data dari atribut kategorik menjadi atribut numerik.

	UMUR	KECAMATAN	JENIS_KB	STATUS_PESERTA
0	24	18	2	3
1	35	18	1	3
2	21	18	8	3
3	30	18	8	2
4	42	18	1	3
5	40	18	7	1
6	35	18	1	1
7	33	18	1	1
8	32	18	8	1
9	35	18	8	1
10	34	18	8	3
11	37	18	7	3
12	41	18	7	1
13	50	18	8	3
14	35	18	1	1
15	23	18	2	3
16	21	18	1	3
17	22	18	0	1
18	26	18	1	3
19	29	18	1	3
20	27	18	8	1
21	30	18	8	3
22	32	18	8	3
23	43	16	8	3
24	45	16	6	1
25	46	16	7	3
26	50	16	7	3
27	21	16	8	3
28	22	16	8	2
29	28	16	8	3
30	19	16	8	3
31	23	16	0	1
32	30	16	8	3
33	33	16	0	1
34	31	16	8	3
35	32	16	1	3
36	37	16	8	3
37	39	16	7	3
38	50	16	8	1
39	41	16	7	3

Gambar 5. Hasil data Transformaasi

4. Mencari K terbaik dengan metode *elbow*

Pada tahapan ini penentuan jumlah *cluster* (K) terbaik menggunakan metode *elbow*. dimana nilai *inertias* atau *error* yang dihasilkan semakin menurun maka jumlah K yang dipilih akan semakin bagus. dimana nilai *inertias* atau *error* yang dihasilkan semakin menurun maka jumlah K yang dipilih akan semakin bagus. dimana nilai *inertias* atau *error* yang dihasilkan semakin menurun maka jumlah K yang dipilih akan semakin bagus.



Gambar 6. Grafik K terbaik dengan metode *elbow*

5. Menentukan *centroid* awal

Tahapan selanjutnya setelah menentukan jumlah K terbaik dalam proses *K-Means clustering* yaitu menentukan *centroid* awal atau titik pusat *cluster* pada tahap ini centroid awal ditentukana secara acak dari data yang digunakan untuk *clustering*.

6. Menghitung jarak dari masing-masing data

Jarak setiap data ke titik pusat *cluster* (*centroid*) dihitung menggunakan *Euclidean Distance*. Perhitungan dilakukan sebanyak 4 iterasi hingga masing-masing *cluster* bernilai optimal dan tidak berpindah tempat.

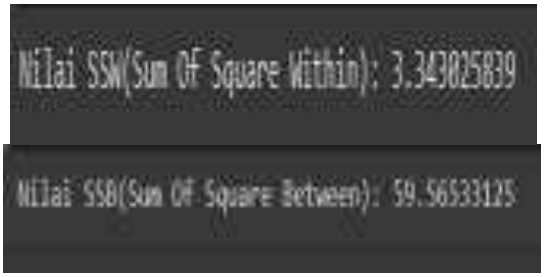
Gambar 7. Hasil iterasi 1 sampai 3

Karena pada iterasi pertama sampai dengan iterasi ketiga masih ada nilai (label) yang berpindah maka dilanjutkan hasil iterasi keempat sampai nilai (label) tidak berpindah lagi.

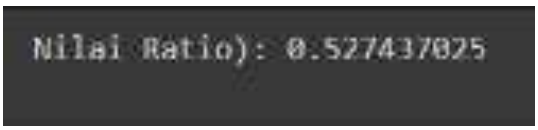
Gambar 8. Hasil iterasi 4 sampai 5

Pada iterasi keempat dan kelima titik pusat cluster atau objeknya tidak ada yang berpindah sehingga tidak diperlukan lagi perhitungan jarak dan mendapatkan bahwa perhitungan *K-Means* clustering peserta KB aktif pada tiap cluster telah selesai. Selanjutnya hasil cluster dari K- Means dievaluasi untuk mengetahui seberapa optimal cluster tersebut. Ada empat langkah dalam menghitung nilai DBI, sebagai berikut:

Gambar 9. Hasil Nilai SSW



Gambar 10. Hasil Nilai SSB



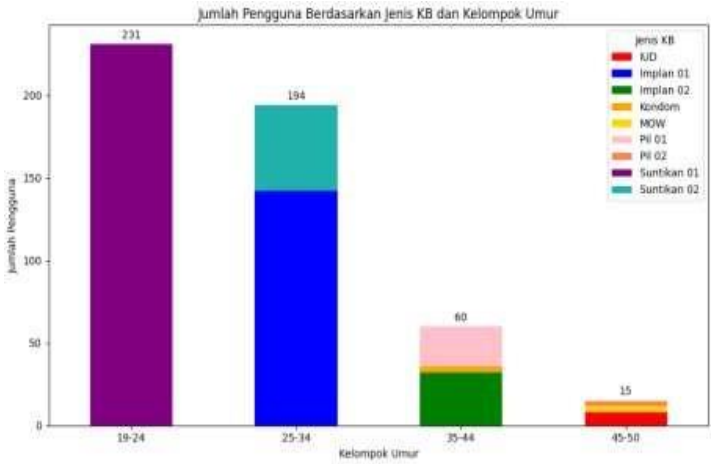
Gambar 11. Hasil Nilai Ratio

Gambar 12. Hasil Nilai DBI

Hasil evaluasi *clustering* peserta keluarga berencana aktif dengan DBI *(Davies Bouldin Index)* dengan jumlah K = 4 memiliki kualitas yang sangat baik karena nilai DBI-nya bernilai hampir mendekati 0 (nol) atau sama dengan 0 (nol).

7. Visualisasi hasil cluster

Hasil *cluster* pada masing-masing kelompok terdapat 4 cluster dimana cluster tersebut ditandai dengan warna dari masing-masing jenis KB.



Gambar 13. Hasil *Clustering* Peserta KB Aktif

1. Suntikan 01 (Suntikan 3 bulan).
2. Suntikan 02 (Suntikan 1 bulan).

3. Implan 01 (Implan 1 batang).
4. Implan 02 (Implan 2 batang).
5. Pil 01 (Pil Kombinasi).
6. Pil 02 (Pil Progestin).
7. Iud (Intrauterin Device).
8. Mow (Menstruation Overweight).
9. Kondom

Berdasarkan analisis tentang penggunaan kontrasepsi (KB) menurut kelompok umur, terlihat variasi yang cukup besar dalam jenis KB yang dipilih. Secara keseluruhan, analisis berdasarkan kelompok umur menunjukkan bahwa setiap kelompok memiliki pola penggunaan kontrasepsi yang berbeda, yang dipengaruhi oleh fase kehidupan, kebutuhan, dan pengetahuan tentang kontrasepsi. Ini menekankan pentingnya edukasi dan penyuluhan yang disesuaikan dengan masing-masing kelompok umur untuk meningkatkan penggunaan kontrasepsi secara efektif.

3.2 PEMBAHASAN

Hasil clustering menunjukan bahwa peserta keluarga berencana aktif dikelompokan dalam empat cluster. Penggunaan algoritma *K-Means* dalam penelitian ini juga menunjukan keunggulan dalam mengelompokan data dengan karakteristik serupa yang dapat membantu pemerintah dalam merumuskan kebijakan yang lebih tepat sasaran. Dengan memahami karakteristik dari masing-masing cluster, intervensi yang lebih spesifik dapat dirancang untuk meningkatkan keberhasilan program keluarga berencana (Hidayat 2023). Dengan Informasi ini pemerintah atau organisasi kesehatan dapat menyesuaikan kebijakan dan program mereka untuk lebih fokus pada jenis kontrasepsi yang paling sering digunakan, sambil mempertimbangkan upaya untuk meningkatkan pemahaman dan akses ke jenis KB yang tidak sering digunakan.

4. KESIMPULAN

Hasil clustering peserta KB aktif berdasarkan jenis kontrasepsi yang digunakan menunjukan beberapa kesimpulan penting:

1. Dalam penerapan dataset peserta keluarga berencana aktif dengan memanfaatkan algoritma K-Means dan dibantu denga python, data peserta KB dapat dibagi kedalam 4 cluster dimana setiap cluster dengan peengguna jenis KB suntikan dan implan paling sering digunakan oleh peserta KB.
2. Pemahaman terhadap karakteristik setiap cluster, seperti distribusi jenis KB, dapat membantu pemerintah dalam merancang intervensi yang lebih spesifik dan tepat sasaran untuk meningkatkan keberhasilan program KB (Keluarga Berencana).
3. Intervensi yang disesuaikan dengan kebutuhan unik setiap cluster, misalnya pendekatan edukasi yang berbeda untuk dapat meningkatkan efektivitas program KB secara keseluruhan

DAFTAR PUSTAKA

[1] Wardani, Syafrina Dyah Kusuma, et al. "Perbandingan Hasil Metode Clustering K- Means, Db Scanner & Hierarchical Untuk Analisa Segmentasi Pasar." *JIKO (Jurnal Informatika dan Komputer)* 7.2 (2023): 191-201.

[2] Aulia, Putri Ulil Fatma, and Sudin Saepudin. "Penerapan Data Mining K-Means Clustering Untuk Mengelompokkan Berbagai Jenis Merk Laptop." *Prosiding Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra*. Vol. 1. 2021.

[3] Hasyim, Fuadz, M. Kom, and M. Muafi. "Implementasi Data Mining Dalam Menentukan Strategi Promosi Program KB Menggunakan Algoritma K-Means Clustering." *Jurnal Kecerdasan Buatan* 3.1 (2022).

[4] Ua, Angelina MTI Sambi, et al. "Penggunaan Bahasa Pemrograman Python Dalam Analisis Faktor Penyebab Kanker Paru-Paru." *Jurnal Publikasi Teknik Informatika* 2.2 (2023): 88-99.

[5] Maesaroh, Maya, Tesa Nur Padilah, and Jajam Haerul Jaman. "PENERAPAN ALGORITMA K-MEANS CLUSTERING PADA PENGELOMPOKAN DAERAH PENYEBARARAN DIARE DI PROVINSI JAWA BARAT." *JATI (Jurnal Mahasiswa Teknik Informatika)* 7.4 (2023): 2783-2787.

[6] Indriana, Ika, Sarah Sambiran, and Neni Kumayas. "Implementasi Program Keluarga Berencana Di Kecamatan Kotamobagu Selatan Kota Kotamobagu." *Jurnal Eksekutif* 1.1 (2021).

[7] Jaya, I. N. T. D., & udytama, i. W. W. W. (2021). Efektifitas undang-undang nomor 52 tahun 2009 tentang perkembangan penduduk dan pembangunan keluarga terhadap keluarga berencana (studi kasus di desa medewi, kecamatan pekutatan kabupaten jembrana). *Jurnal hukum mahasiswa*, *1*(1).

[8] Fajri, Muhammad Bhakti, and Susan Dian Purnamasari. "Klasterisasi Pola Penyebaran Penyakit Pasien Berdasarkan Usia Pasien Menggunakan K-Means Clustering." *Journal of Information Technology Ampera* 3.3 (2022): 317-334.

[9] Hidayat, Andrian, Nurhidayati Nurhidayati, and Amri Muliawan Nur. "IMPLEMENTASI ALGORITMA K-MEANS UNTUK KLASTERISASI PESERTA KELUARGA BERENCANA BERDASARKAN TINGKAT RISIKO KEHAMILAN DI DESA PRINGGASELA SELATAN." *Jurnal PRINTER: Jurnal Pengembangan Rekayasa Informatika dan Komputer* 1.2 (2023): 154-166.

[10] Sari, Herlina Latipa, and Ila Yati Beti. "Penerapan Data Mining Dalam Pengelompokkan Buku Yang Dipinjam Menggunakan Algoritma K-Means." *KLIK: Kajian Ilmiah Informatika dan Komputer* 3.6 (2023): 925-933.

[11] Utami, Wiranti Sri, Nila Pratiwi, and Faisal Muhammad. "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Clustering Perokok Usia Lebih dari 15 Tahun." *Bulletin of Information Technology (BIT)* 4.4 (2023): 501-507.

[12] Fathurrahman, Fathurrahman, Sri Harini, and Ririen Kusumawati. "Evaluasi clustering K-Means dan K-Medoid pada persebaran Covid-19 di Indonesia dengan metode Davies-Bouldin Index (DBI)." *Jurnal Mnemonic* 6.2 (2023): 117-128.

[13] Jollyta, Deny, et al. "Optimasi Cluster Pada Data Stunting: Teknik Evaluasi Cluster Sum of Square Error dan Davies Bouldin Index." *Prosiding Seminar Nasional Riset Information Science (SENARIS)*. Vol. 1. 2021

[14] Kelen, Yoseph PK, et al. "Decision support system for the selection of new prospective students using the simple additive weighted (SAW) method." *AIP Conference Proceedings*. Vol. 2798. No. 1. AIP Publishing, 2023.

[15] Fairuzabadi, Muhammad, et al. *SISTEM PENDUKUNG KEPUTUSAN: KONSEP, METODE DAN IMPLEMENTASI*. Get Press Indonesia, 2024.