

**CLUSTERING DENGAN ALGORITMA K-MEANS
TERHADAP CALON PENERIMA BANTUAN
RENOVASI RUMAH**

(Studi Kasus Desa Huidu Melito)

Oleh

FAISAL B. ABAY

T3117391

SKRIPSI

Untuk memenuhi salah satu syarat ujian

Guna memperoleh gelar sarjana



**PROGRAM SARJANA
TEKNIK INFORMATIKA
UNIVERSITAS ICHSAN GORONTALO
GORONTALO
2022**

PERSETUJUAN SKRIPSI
CLUSTERING DENGAN ALGORITMA K-MEANS
TERHADAP CALON PENERIMA BANTUAN RENOVASI
RUMAH

(Studi Kasus Desa Huidu Melito)

Oleh
FAISAL B. ABAY
T3117391

SKRIPSI

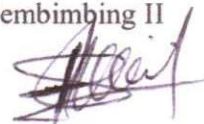
Untuk memenuhi salah satu syarat ujian
Guna memperoleh gelar sarjana
Program study teknik informatika,
Ini telah di setujui oleh pembimbing
Gorontalo, 30 Mei 2022

Pembimbing I



Suhardi Rustam, M.Kom
NIDN.0915088403

Pembimbing II



Sarlis Mooduto, M.Kom
NIDN.0920078803

PENGESAHAN SKRIPSI

CLUSTERING DENGAN ALGORITMA K-MEANS
TERHADAP CALON PENERIMA BANTUAN
RENOVASI RUMAH
(Study Kasus Desa Huidu Melito)

OLEH

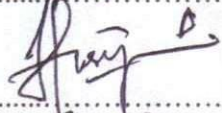
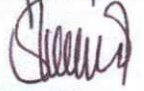
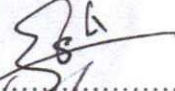
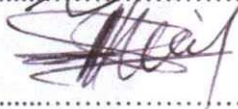
FAISAL B. ABAY

T3117391

Diperiksa Oleh Panitia Ujian Strata Satu (SI)
Universitas Ichsan Gorontalo

Gorontalo, Mei 2022

1. Ketua Penguji
Amiruddin, S.Kom, M.Kom
2. Anggota
Haditsah Annur, S.Kom, M.Kom
3. Anggota
Sumarni, S.Kom, M.Kom
4. Anggota
Suhardi Rustam, S.Kom, M.Kom
5. Anggota
Sarlis Mooduto, S.Kom, M.Kom

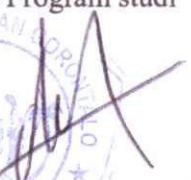

.....

.....

.....

.....

.....

Mengetahui

Dekan Fakultas Ilmu Komputer


Jorry Karim M. Kom
NIDN.0918077302

Ketua Program studi


Sudirman S. Panna M. Kom
NIDN.0924038205

PERNYATAAN SKRIPSI

Dengan ini saya menyatakan bahwa :

1. Karya tulis (Skripsi) saya ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (Sarjana) baik di Universitas Ichsan Gorontalo maupun diperguruan tinggi lainnya.
- 2.
3. Karya tulis (Skripsi) saya ini adalah murni gagasan, rumusan, dan penelitian saya sendiri, tanpa bantuan pihak lain, kecuali arahan dari tim pembimbing.
- 4.
5. Dalam karya tulis (Skripsi) saya ini tidak terdapat karya atau pendapat yang telah dipublikasikan orang lain, kecuali secara tertulis dicantumkan sebagai acuan/sitasi dalam naskah dan dicantumkan pula dalam daftar pustaka.
- 6.
7. Pernyataan ini saya buat dengan sesungguhnya dan apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya tulis ini, serta sanksi lainnya sesuai dengan norma-norma yang berlaku di Universitas Ichsan Gorontalo.

Gorontalo, 30 Juni 2022

Yang membuat pernyataan



Faisal B. Abay

ABSTRACT

FAISAL B. ABAY. T3117391. CLUSTERING USING K-MEANS ALGORITHM FOR PROSPECTIVE RECIPIENTS OF HOUSE RENOVATION ASSISTANCE

This study aims to 1) obtain the results of the performance of the k-means clustering method in grouping population data on prospective recipients of house renovation assistance for precise accuracy, and 2) find out the results of grouping the population data of prospective recipients of house renovation assistance. Based on the level of application, this study is applied research. It can also be seen from the type of information processing as a quantitative research method. The method used in this study to solve the problem of population data clustering is a descriptive method. The method aims to reveal objective truth. In other words, it describes a situation that is currently occurring in society by using scientific procedures to answer a problem accurately. The method used in this study is the K-Means algorithm method. This method can group large amounts of data based on predetermined characteristics. The use of this method is very effective. The data used in this study are up to 83 data. The data has been grouped into the decent population, namely (C1) with 24 data, poor population (C2) has 55 data, and unfit population data (C3) with 4 data. It indicates that this study can properly be completed.

Keywords: *Clustering, K-Means, house renovation assistance, population data*



ABSTRAK

FAISAL B. ABAY. T3117391. CLUSTERING DENGAN ALGORITMA K-MEANS TERHADAP CALON PENERIMA BANTUAN RENOVASI RUMAH

Penelitian ini bertujuan untuk 1). Untuk memperoleh hasil kinerja metode k-meansclustering pada pengelompokan data penduduk terhadap calon penerima bantuan renovasi rumah untuk memperoleh akurasi dengan tepat 2). Untuk mengetahui hasil pengelompokan data penduduk calon penerima bantuan renovasi rumah. Dilihat dari tingkat penerapan maka penelitian ini merupakan penelitian terapan. Dapat dilihat pula dari jenis informasi yang diolah maka, penelitian ini merupakan penelitian kuantitatif. Metode yang digunakan dalam penelitian ini untuk masalah *Clustering* data penduduk yaitu metode deskriptif, dimana metode ini bertujuan untuk mengungkapkan kebenaran yang objektif. Dengan kata lain menjabarkan suatu keadaan atau situasi yang terjadi saat ini dalam suatu masyarakat dengan menggunakan prosedur ilmiah untuk menjawab sebuah masalah secara akurat. Metode yang digunakan dalam penelitian ini adalah metode algoritma k-means, metode ini mampu mengelompokkan data yang banyak berdasarkan karakteristik yang telah di tentukan sebelumnya. Penggunaan metode ini sangat efektif dimana data yang digunakan dalam penelitian ini data yang digunakan berjumlah 83 data. Data tersebut telah dikelompokkan kedalam penduduk layak (C1) 24 data, pendudukkurang layak (C2) 55 data, data penduduk tidak layak (C3) 4 data. Dengan hasil tersebut penelitian ini dapat di selesaikan dengan baik.

Kata kunci: *Clustering*, K-Means, bantuan renovasi rumah, data penduduk



KATA PENGANTAR

Assalamualaikum Wr. Wb.

Alhamdulillah, puji syukur kehadiran Allah SWT. Karena hanya dengan izin dan kuasanya jualah penulis dapat menyelesaikan skripsi ini yang berjudul **“CLUSTERING DENGAN ALGORITMA K-MEANS TERHADAP CALON PENERIMA BANTUAN RENOVASI RUMAH”** tepat pada waktunya. Adapun Tujuan skripsi ini adalah untuk menyelesaikan penelitian yang telah di lakukan sebelumnya guna memperoleh gelar Sarjana SI Program Studi Teknik Informatika Fakultas Ilmu Komputer di Universitas Ichsan Gorontalo.

Ucapan terimakasih penulis sampaikan kepada semua pihak yang telah membantu baik dalam bentuk dukungan moril maupun materil sehingga skripsi ini dapat selesai tepat pada waktunya. Rasa hormat dan terimakasih penulis sampaikan kepada :

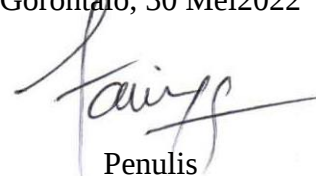
1. Ibu Dr. Hj Jurikose Abdusamad M.Si laku ketua Yayasan Pengembangan Ilmu Pengetahuan dan Teknologi (YPIPT) Ichsan Gorontalo;
2. Bapak Dr. Abdul Gaffar Latjokke, M.Si, selaku Rektor Universitas Ichsan Gorontalo;
3. Bapak Jorri Karim, M.Kom, selaku dekan Fakultas Ilmu Komputer;
4. Bapak Sudirman Melangi, M.Kom, selaku wakil dekan I Bidang Akademik;
5. Ibu Irma Surya Kumala Idris, M.Kom, selaku wakil dekan II bidang Kemahasiswaan;
6. Bapak Sudirman S. Panna, M.Kom, selaku Ketua Program Studi Teknik Informatika Fakultas Ilmu Komputer;
7. Bapak Suhardi Rustam, M.Kom, sebagai Pembimbing Utama yang telah membimbing penulis selama mengerjakan proposal ini;
8. Bapak Sarlis Mooduto, M.Kom sebagai Pembimbing Pendamping yang telah membimbing penulis selama mengerkan proposal ini
9. Kedua orang tua, Ayah dan Ibu, serta Kakak dan Adik, yang juga telah memberikan dukungan, motivasi, dan doa terbaik.
10. Segenap Dosen Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.

11. Rekan-rekan Program Study Teknik Informatika, Kampus 3 Universitas Ichsan Gorontalo.
12. Teman-teman Angkatan 2017 dan Kakak-kakak Fakultas Ilmu Komputer yang banyak memberikan motivasi serta kiritikan kepada penulis sehingga usulan penelitian ini bisa selesai.
13. Dan semua pihak yang tidak dapat di sebut satu persatu yang telah membantu penulis untuk menyelesaikan usulan penelitian ini.

Penulis tentunya menyadari bahwa penyusunan usulan penelitian ini masih terdapat banyak kekurangan dan kelemahan baik dalam penyusunan maupun penulisan. Oleh karena itu penulis berharap kepada semua pihak untuk dapat menyampaikan saran dan kritikan demi membangun atau menambahkan kesempurnaan usulan penelitian ini.

Wassalamu'alaikum wr.wb.

Gorontalo, 30 Mei2022



Penulis

DAFTAR ISI

HALAMAN SAMPUL	
HALAMAN JUDUL	i
PERSETUJUAN SKRIPSI	ii
PENGESAHAN SKRIPSI	iii
PERNYATAAN SKRIPSI	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
BAB II LANDASAN TEORI	6
2.1 Tujuan Studi	6
2.2 Tinjauan Pustaka	7
2.2.1 Rumah Tidak Layak Huni	7
2.2.2 Rehabilitas Sosial Rumah Tidak Layak Huni	8
2.2.3 Data Mining	8
2.2.4 <i>Clustering</i>	10
2.2.5 Algoritma K-Means	11
2.2.6 Penerapan Algoritma K-Means	13
2.4 Kerangka Pikir	19
BAB III METODE PENELITIAN	20
3.1 Objek Penelitian	20
3.2 Metode Penelitian	21

3.3 Tahap Pengumpulan Data	21
3.3.1 Metode Observasi	21
3.3.2 Metode Wawancara.....	21
3.4 Pemodelan.....	22
3.4.1 Pra Pengolahan Data	22
3.4.2 Hasil Clustering	22
3.4.3 Evaluasi.....	22
BAB IV HASIL PENELITIAN.....	23
4.1 Hasil Pengumpulan Data.....	23
4.2 Hasil Pemodelan	26
4.2.1 Pra Pengolahan Data	26
4.2.2 Normalisasi Data.....	26
4.2.3 Hasil Hitung Algoritma K-Means.....	26
4.2.4 Hasil <i>Clustering</i>	33
BAB V PEMBAHASAN	34
5.1 Pembahasan Model	34
5.2 Pembahasan Tools	34
5.3 Pembahasan Hasil Algoritma.....	38
BAB VI KESIMPULAN DAN SARAN	42
6.1 Kesimpulan	42
6.2 Saran	42
 DAFTAR PUSTAKA.....	
LAMPIRAN.....	

DAFTAR TABEL

Tabel 1.1 Dataset Calon Penerima Rumah Mahyani	3
Tabel 2.1 Peneliti Terdahulu	6
Tabel 2.2 Data Penduduk Kel. Bungusari	14
Tabel 2.3 Data Penduduk Kel Bungusari setelah dilakukan perubahan.....	14
Tabel 2.4 Pusat Awal Cluster	15
Tabel 2.5 Hasil Perhitungan dan Pengelompokkan Iterasi	16
Tabel 2.6 Nilai Pusat <i>Cluster</i> yang Baru	17
Tabel 2.7 Hasil Perhitungan dan Pengelompokkan Iterasi	17
Tabel 4.1 Hasil Pengumpulan Data	23
Tabel 4.2 penentuan awal <i>cluster</i>	27
Tabel 4.3 Mencari Jarak Terdekat	31
Tabel 4.4 Jumlah Perhitungan iterasi pertama	33
Tabel 4.5 penentuan <i>cluster</i> baru	34
Tabel 4.6 Hasil iterasi ke-2.....	35
Tabel 4.7 Jumlah Perhitungan iterasi kedua.....	37
Tabel 4.8 Penentuan <i>cluster</i> iterasi ke-3	37
Tabel 4.9 hasil <i>cluster</i> iterasi ke-3.....	38
Tabel 4.10 Jumlah Perhitungan iterasi ketiga	40

DAFTAR GAMBAR

Gambar 2.1 Tahapan Data Mining	9
Gambar 2.2 Flowchart	13
Gambar 3.1 Sistem yang di Usulkan	19
Gambar 3.2 Pemodelan K-Means.....	22
Gambar 5.1 Tampilan Utama Rapidminer	34
Gambar 5.2 tampilan import data	35
Gambar 5.3 Tampilan Lokasi Penyimpanan Data.....	35
Gambar 5.4 Tampilan Data Yang Akan Digunakan	36
Gambar 5.5 Tampilan Desain.....	36
Gambar 5.6 Tampilan Result Data	37
Gambar 5.7 Tampilan <i>cluster</i> model	37
Gambar 5.8 Tampilan Devies Bouldin.....	38

BAB I

PENDAHULUAN

1.1 Latar Belakang

Rumah atau tempat tinggal merupakan salah satu kebutuhan primer masyarakat selain sandang dan pangan. Sandang berarti pakaian dan pangan berarti makanan. Pentingnya rumah atau tempat tinggal yaitu sebagai tempat untuk berlindung, tempat beristirahat, dan sebagai tempat untuk berkumpul bersama keluarga dengan aman dan nyaman. Saat ini di Indonesia sendiri masih banyak terdapat rumah yang masuk dalam kategori dengan keadaan tak layak huni. Dalam hal ini pemerintah memberikan bantuan Rumah Tidak Layak Huni (RTLH) dalam bentuk uang untuk pembelian bahan bangunan guna memperbaiki rumah tidak layak huni dimana rumah tersebut tidak memenuhi persyaratan kecukupan minimal bangunan dan keselamatan suatu bangunan [1].

Bantuan renovasi rumah merupakan bentuk kegiatan pemerintah dalam hal membantu masyarakat kurang mampu. Penyaluran bantuan ini ditangani oleh pemerintah daerah yang kemudian di salurkan kepada beberapa desa yang masuk dalam kategori desa miskin [1]. Desa Huidu Melito merupakan salah satu desa yang masuk dalam kategori desa miskin di Kecamatan Tomilito Kabupaten Gorontalo Utara. Desa Huidu Melito adalah desa mekaran dari desa Molantadu bersamaan dengan terbentuknya Kecamatan Tomilito pada tahun 2011 dengan jumlah penduduk 979 jiwa dan jumlah KK 117. Namun, dalam hal ini Pemerintah Desa Huidu Melito masih memiliki beberapa kendala dalam menentukan sasaran penengangan bantuan Rumah Tidak Layak Huni. Kendala tersebut dikarenakan banyaknya data-data yang berubah sehingga memungkinkan proses penentuan penerima bantuan Renovasi Rumah tidak tepat sasaran. Untuk mengatasi masalah tersebut maka dibutuhkan teknik data mining untuk membantu pengelompokan data dalam penelitian ini.

Program bantuan Rumah Tidak Layak Huni ini masih tergolong baru di desa Huidu Melito sehingga tidak adanya label pada dataset. Maka dibutuhkan metode klasterisasi (*Clustering*) yang merupakan suatu metode untuk mencari data dan mengelompokkan data yang memiliki kemiripan karakteristik yang sama

(*similarity*) antara satu data dengan data yang lain [2]. Metode *Clustering* juga tidak menggunakan label dalam setiap prosesnya jadi, label dapat diberikan saat *cluster* telah terbentuk [1].

Metode klasterisasi banyak diusulkan oleh para ahli, salah satunya ialah *k-Means Method* atau metode *k-Means* termasuk dalam metode berbasis partisi, dimana metode ini bekerja dengan cara membagi data kedalam sejumlah kelompok. Dalam penelitian ini penggunaan metode *K-Means* merupakan metode clustering yang paling sederhana dan umum. Hal ini dikarenakan *k-means* mempunyai kemampuan dalam mengelompokkan/clustering data dalam jumlah yang cukup besar dengan waktu komputasi yang cepat dan efisien [3]. Metode *K-Means* berusaha mengclusterkan data yang ada kedalam beberapa kelompok, dimana data dalam satu kelompok yang memiliki karakteristik yang sama dengan yang lainnya dan memiliki karakteristik yang berbeda dengan data yang ada dalam kelompok yang lain [2].

Bangunan yang baik tidak harus besar, indah dan mewah, akan tetapi harus memenuhi syarat dan kesehatan rumah agar para penghuninya dapat beraktifitas dengan aman dan nyaman. Adapun beberapa kriteria yang dapat di kategorikan sebagai rumah layak huni, seperti adanya ventilasi udara, atap dan langit-langit aman dari panas dan hujan, pembuangan limbah yang lancar, air bersih yang cukup, terhindar dari polusi udara dan pencemaran lingkungan. Dari beberapa kriteria di atas maka peneliti akan mengelompokkan data penduduk yang masuk dalam kategori rumah tak layak huni menggunakan algoritma *k-means*.

Data yang akan digunakan dalam penelitian ini merupakan data yang berada di desa huido melito, yang akan dikelompokkan kedalam penduduk layak (C1), kurang layak (C2), dan tidak layak menerima bantuan (C3). Penentuan *cluster* tersebut terdiri dari 3 variable yaitu umur, jumlah tanggungan dan jumlah pendapatan untuk mengetahui mana saja rumah yang mendekati syarat kriteria rumah tak layak huni akan menjadi bahan pertimbangan bagi pemerintah desa untuk memberikan bantuan atau merenovasi rumah tersebut.

Table 1.1 Dataset Calon Penerima Renovasi Rumah

No	NAMA	L/P	Umur	Jumlah Tanggungan	Pendapatan	Kepemilikan Rumah
1	NOPRI ABJUL	L	38	2	500.000	Milik Sendiri
2	YASMAN HARUN	L	42	3	1.000.000	Milik Sendiri
3	GANI LAMATO	L	41	3	700.000	Milik Orang Tua
4	RINTON ISMAIL	L	36	2	500.000	Milik Orang Tua
5	RIDWAN KADIR	L	37	5	750.000	Milik Sendiri
6	WIRSON KADIR	L	43	3	1.500.000	Milik Orang Tua
7	JON RISKIN HULOPI	L	33	3	500.000	Milik Sendiri
8	SIPRIT RUMOPA	L	31	2	500.000	Milik Sendiri
....						
83	ARIYANTO IKILO	L	42	2	500.000	Kontrak

(Sumber data : Desa Huidu Melito)

Data di atas merupakan data yang digunakan dalam penelitian ini merupakan data yang di ambil dari kantor desa Huidu Melito. Data tersebut berjumlah 83 data. Dalam data tersebut peneliti menemukan masalah tentang calon penerima bantuan renovasi rumah, dimana masih terdapat penerima bantuan renovasi rumah yang memiliki pendapatan tergolong mampu dan juga terdapat rumah dengan kategori layak huni namun mendapatkan bantuan renovasi rumah. Sehingga hal inilah yang menjadi permasalahan bagi pemerintah desa dalam menentukan calon penerima bantuan renovasi rumah yang tidak tepat sasaran dalam menentukan calon penerima bantuan renovasi rumah. Maka dari itu salah satu teknik untuk mengelompokkan data penduduk terhadap calon penerima bantuan renovasi rumah ini menggunakan teknik data mining.

Data mining sebenarnya bukanlah topik yang baru dalam suatu penelitian. *Data mining* digunakan sebagai objek penelitian untuk memperbaiki teknik-teknik sebelumnya sehingga dapat menangani berbagai masalah yang sering di jumpai [4]. Data mining merupakan salah satu disiplin ilmu yang digunakan dalam menentukan sebuah informasi tertentu dalam sekumpulan data sebagai faktor pendukung pengambilan keputusan. Data mining juga merupakan salah satu proses dalam memperoleh informasi yang memiliki nilai guna dari sekumpulan data [5]. Penggunaan data mining menggunakan algoritma k-means dalam mengelompokkan data penduduk calon penerima bantuan renovasi rumah di desa Huidu Melito diharapkan dapat membantu menentukan calon penerima bantuan yang layak mendapatkan dan membutuhkan bantuan renovasi rumah. Dengan adanya metode K-Means *Clustering* akan mempermudah peneliti dalam melakukan pengelompokkan data yang ada secara akurat dan dapat membantu pemerintah desa untuk mengelompokkan data dalam jumlah yang besar.

Berdasarkan uraian atau pemaparan diatas untuk kesesuaian metode dan masalah, maka dianggap perlu untuk melakukan penelitian menggunakan metode k-means clustering, dengan judul “***Clustering dengan Algoritma K-Means Terhadap Calon Penerima Bantuan Renovasi Rumah (Studi Kasus Desa Huidu Melito)***”, maka peneliti tertarik dalam melakukan penelitian terkait membantu pemerintah desa Huidu Melito dalam menentukan sasaran penerima bantuan Rumah Tidak Layak Huni secara tepat menggunakan algoritma *K-Means method*. Hasil dari *cluster* bisa dijadikan bahan masukan bagi pemerintah desa Huidu Melito, agar dapat memberikan kontribusi terhadap sistem pemerintahan desa untuk lebih mempermudah pengolahan data di masa kini dan yang akan datang.

1.2 Identifikasi Masalah

1. Pihak pemerintah desa Huidu Melito masih sulit untuk mengelompokkan data penduduk terhadap calon penerima bantuan renovasi rumah.
2. Belum adanya sistem yang mengatur dalam mengelompokkan data penduduk calon penerima bantuan dengan tingkat akurasi yang tepat dan akurat.

1.3 Rumusan Masalah

Berdasarkan latar belakang diatas dapat dirumuskan masalah pokok yang berkaitan yaitu :

1. Bagaimana kinerja metode k-means *clustering* dalam mengelompokkan data penduduk calon penerima bantuan renovasi rumah?
2. Bagaimana hasil pengelompokkan data penduduk terhadap calon penerima bantuan renovasi rumah menggunakan metode algoritma k-means *clustering*?

1.4 Tujuan Penelitian

1. Memperoleh hasil kinerja metode k-means clustering pada pengelompokkan data penduduk terhadap calon penerima bantuan renovasi rumah untuk memperoleh akurasi dengan tepat.
2. Mengetahui hasil pengelompokkan data penduduk calon penerima bantuan renovasi rumah.

1.5 Manfaat Penelitian

1. Manfaat Teoritis

Memberikan masukan bagi perkembangan ilmu pengetahuan dan teknologi. Khususnya pada bidang ilmu komputer, yaitu dengan penggunaan metode *k-means clustering*.

2. Manfaat Praktis

Memudahkan pemerintah desa untuk memberikan informasi data pengelompokkan penduduk terhadap calon penerima bantuan renovasi rumah di harapkan hasil dari pengelompokkan dapat dijadikan bahan pengambil kebijakan terkait peningkatan kecerdasan dimasa mendatang. Sehingga dapat membantu untuk mengoptimalkan program-program pemerintah di desa huidu melito.

BAB II

LANDASAN TEORI

2.1 Tujuan Studi

Penelitian terdahulu merupakan bagian dari panduan atau contoh dalam penelitian yang akan dilakukan nantinya, yang menjadi acuan dan perbandingan dalam membuat penelitian ini. Terdapat tiga jurnal nasional yang di angkat sebagai panduan yaitu:

Tabel 2.1 Peneliti Terdahulu

No	Nama Peneliti	Judul Penelitian	Tahun	Metode	Hasil
1	Risa Rohima	Aplikasi Menentukan Tingkat Sosial Ekonomi Penduduk Menggunakan Algoritma K-Means di Paberesan Sumenep Institut Teknologi Nasional Malang	2017	K-Means	Kesimpulan yang di dapat ialah dari 20 data yang diuji menghasilkan 15 data yang sesuai dengan hasil dari paberesan, sehingga keakuratan sistem bernilai 75%.
2	Lina Listiani, Yoga Handoko Agustin, Mochhammad Zaenal Ramdhani	Implementasi Algoritma K-Means Cluster Untuk Rekomendasi Pekerjaan Berdasarkan	2019	K-Means Clustering	Setelah melakukan analisis, perancangan, implementasi dan pengujian yang telah dilakukan, maka diperoleh hasil 4 kelompok

		Pengelompokkan Data Penduduk			jenis pekerjaan berdasarkan jenis kelamin, pendidikan, umur dan jenis pekerjaan dengan nilai akurasi yang diperoleh yaitu sebesar 79% dan diklasifikasikan sebagai fair atau cukup pada klasifikasi nilai UAC.
--	--	---------------------------------	--	--	--

2.2 Tinjauan Pustaka

2.2.1 Rumah Layak Huni

Menurut Peraturan Menteri Negara Perumahan Rakyat Republik Indonesia Nomor 22/Permen/M/2008 tentang standar pelayanan bidang perumahan rakyat daerah provinsi dan daerah kabupaten mendefinisikan bahwa rumah layak huni adalah rumah yang memenuhi persyaratan keselamatan bangunan dan kecukupan minimum luas bangunan serta kesehatan penghuninya. Dengan terpenuhinya persyaratan rumah layak huni tersebut dapat dikatakan pemerintah telah berhasil mensejahterakan rakyatnya dengan pembangunan dan juga renovasi rumah sebagai upaya dari peningkatan taraf kesejahteraan masyarakat.

2.2.2 Rumah Tidak Layak Huni

Rumah tidak layak huni merupakan suatu hunian atau tempat tinggal yang tak layak untuk di huni karena bangunan yang tidak memenuhi standar persyaratan untuk hunian baik secara teknis maupun non teknis. Pada umumnya rumah tidak

layak huni erat kaitanya dengan pemukiman kumuh karena di daerah pemukiman kumuh pada dasarnya menggambarkan kemiskinan masyarakat.

Adapun kriteria rumah tidak layak huni apabila [6]:

a. Kondisi Rumah

1. Luas lantai perkapita kota $< 4 \text{ m}^2$, desa $< 10 \text{ m}^2$.
2. Sumber air tidak sehat, akses memperoleh air bersih terbatas.
3. Tidak mempunyai akses mandi, cuci dan kakus.
4. Bahan bangunan tidak permanen atau atap/dinding dari bambu, rumbia.
5. Tidak memiliki pencahayaan matahari dan ventilasi udara.
6. Tidak memiliki pembagian ruangan.
7. Lantai dari tanah atau papan dan rumah lembab.
8. Letak rumah tidak teratur dan berdempetan.

b. Kondisi Lingkungan

1. Lingkungan kumuh dan becek.
2. Saluran pembuangan air tidak memenuhi standar.
3. Jalan setapak tidak teratur.

2.2.3 Rehabilitas Sosial Rumah Tidak Layak Huni

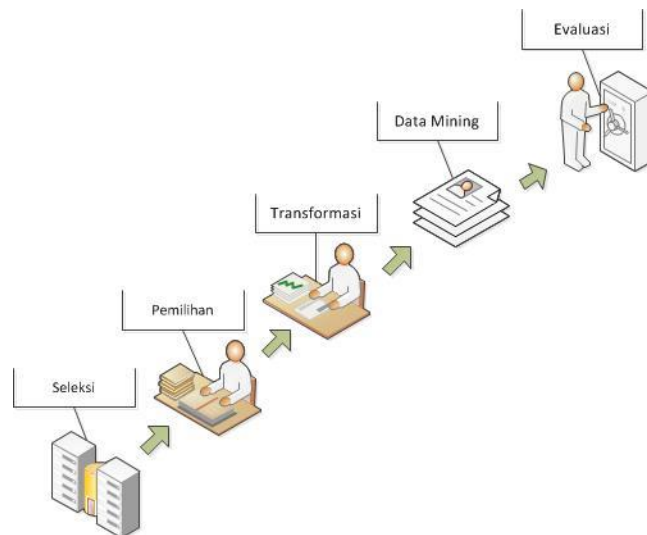
Dalam Peraturan Menteri Sosial Republik Indonesia No.20 tahun 2017 tentang Rehabilitasi Rumah Tidak Layak Huni dan Sarana Prasarana Lingkungan disebutkan bahwa Fakir Miskin adalah orang yang sama sekali tidak memiliki sumber mata pencaharian atau mempunyai sumber mata pencaharian tetapi tidak mempunyai kemampuan untuk memenuhi kebutuhan dasar yang layak bagi kehidupan diri dan keluarganya.

Kegiatan RS-RTLH tersebut dilaksanakan dengan tujuan dalam mengatasi sebagian masalah kemiskinan yang ada, tersedianya rumah layak huni, adanya kenyamanan bertempat tinggal, meningkatkan kemampuan keluarga untuk melaksanakan peran dan fungsi masing-masing keluarga dalam memberikan perlindungan bimbingan dan pendidikan, meningkatnya kualitas kesehatan lingkungan pemukiman dan meningkatnya harkat dan martabat [6].

2.2.4 Data Mining

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. *Data Mining* adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [7].

Data Mining adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode ini dapat digunakan untuk mengambil keputusan di masa mendatang. *Data Mining* ini biasa juga disebut dengan istilah *pattern recognition*.



Gambar 2.1 Tahapan data mining

Berikut adalah tahapan data mining yaitu :

1. Pembersihan Data

Proses ini digunakan untuk membuang data yang tidak konsisten dan bersifat noise dari data awal yang ada.

2. Integrasi Data

Menyatukan data yang terdapat diberbagai basis data yang mungkin berbeda format maupun platform yang kemudian di integrasikan dalam satu database *data warehouse*.

3. Seleksi Data

Data yang terdapat dalam database *data warehouse* kemudian direduksi dengan berbagai teknik. Proses reduksi diperlukan untuk mendapatkan hasil yang lebih akurat dan mengurangi waktu komputasi terutama untuk masalah dengan skala besar (*large scale problem*).

4. Transformasi Data

Transformasi data diperlukan sebagai tahap *pre-procecing*, dimana data yang diolah siap untuk ditambang.

5. Data Mining

Data-data yang telah diseleksi dan ditransformasi ditambang dengan berbagai teknik. Proses *data mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan fungsi-fungsi tertentu. Fungsi atau algoritma dalam *data mining* sangat bervariasi. Pemilihan fungsi atau algoritma yang tepat sangat bergantung pada tujuan dan proses pencarian pengetahuan secara keseluruhan.

6. Evaluasi Pola

Tahap ini merupakan bagian dari proses pencarian pengetahuan yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

7. Representasi Pengetahuan

Langkah terakhir KDD adalah mempresentasikan pengetahuan dalam bentuk yang mudah dipahami oleh pengguna.

2.2.5 *Clustering*

Analisis pengelompokan/*Clustering* merupakan proses membagi data dalam satu himpunan kedalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar dari pada kesamaan data tersebut dengan data dalam kelompok lain [7].

Pada proses analisis *cluster* metode yang digunakan dalam membagi data berdasarkan kesamaan ataupun kemiripan yang telah di tentukan sebelumnya.

Analisis *cluster* secara umum bisa dikatakan bahwa :

- a. Data yang terdapat dalam satu cluster memiliki tingkat kesamaan yang tinggi, dan

- b. Data yang terdapat dalam suatu cluster berbeda yang berbeda memiliki tingkat kesamaan yang rendah.

Clustering adalah metode yang digunakan untuk membagi rangkaian data menjadi beberapa grup berdasarkan kemiripan karakteristik yang telah ditentukan sebelumnya. *Cluster* adalah sekelompok atau sekumpulan objek data yang similar satu dengan yang lain dalam cluster yang sama dan dissimilar terhadap objek-objek data yang berbeda cluster. Objek akan dikelompokkan dalam satu atau lebih cluster sehingga objek yang ada dalam satu cluster akan memiliki kesamaan yang tinggi antara satu sama lain. Objek-objek dikelompokkan berdasarkan prinsip memaksimalkan kesamaan objek untuk cluster yang sama dan memaksimalkan ketidaksamaan untuk cluster yang berbeda. Kesamaan objek biasanya diperoleh dari nilai-nilai atribut yang menjelaskan objek data, sehingga objek data biasanya dipresentasikan sebagai titik dalam ruang multidimensi [8].

2.2.6 Algoritma K-Means

Algoritma K-Means termasuk algoritma *clustering* berulang untuk menetapkan nilai (K) secara acak sebagai nilai pusat dari *cluster/centroid/mean/means*. Dari data yang ada dicari jarak yang paling dekat dengan masing-masing *centroid* menggunakan rumus *euclidian* [9]. Objek akan dikelompokkan ke dalam satu atau lebih *cluster* sehingga objek-objek yang berbeda dalam satu *cluster* akan mempunyai kesamaan yang tinggi antara satu dengan yang lainnya. Objek-objek dikelompokkan berdasarkan prinsip memaksimalkan kesamaan objek pada *cluster* yang berbeda. Kesamaan objek biasanya diperoleh dari nilai-nilai atribut yang menjelaskan objek data, sehingga objek-objek data biasanya dipresentasikan sebagai sebuah titik dalam ruang multidimensi.

K-Means merupakan salah satu metode non-hierarki yang mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang ditetapkan di setiap kelompok dan

memaksimalkan variasi antar kelompok [10]. Metode k-means berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok lain.

Dasar algoritma k-means adalah sebagai berikut:

1. Tentukan nilai k sebagai jumlah *cluster* yang ingin dibentuk
2. Inisialisasi k sebagai *centroid* yang dapat dibangkitkan secara random
3. Hitung jarak setiap data ke masing-masing *centroid* menggunakan persamaan *euclidean distance*
4. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan *centroid*nya
5. Tentukan posisi *centroid* baru
6. Kembali ke langkah 3 jika posisi *centroid* baru dengan *centroid* lama tidak sama.

Hitung jarak setiap data yang ada terhadap masing-masing *centroid* menggunakan rumus *euclidean* hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*, klasifikasikan setiap data berdasarkan kedekatannya dengan *centroid*. Lakukan langkah tersebut hingga nilai *centroid* tidak berubah. Jarak *euclidean* yang dirumuskan sebagai berikut:

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

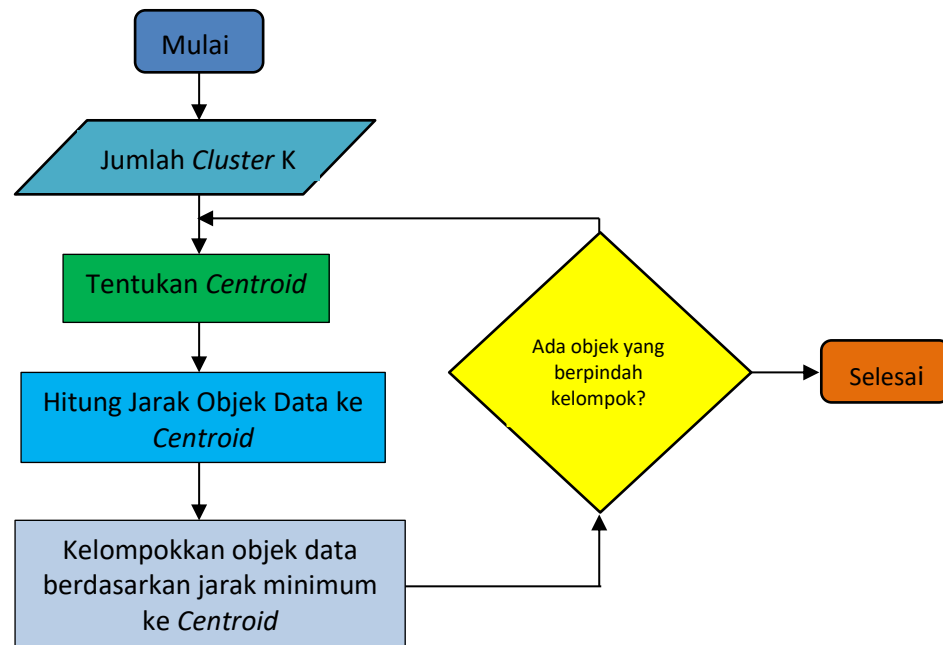
Dimana:

De = *Euclidean distance*

i = banyaknya objek

(x,y) = koordinat objek dan

(s,t) = koordinat *centroid*



Gambar 2.2 Diagram Alur Implementasi Algoritma K-Means

2.2.7 Penerapan Algoritma K-Means

Pada penelitian yang dilakukan oleh (Lina Listiani, Yoga Handoko Agustin, Mochammad Zaenal Ramdhani. 2019) dengan judul Implementasi Algoritma K-Means Cluster untuk Rekomendasi Pekerjaan Berdasarkan Pengelompokan Data Penduduk adalah salah satu contoh kasus penerapan metode K-Means Clustering [9].

Berikut adalah tahapan proses clustering untuk rekomendasi jenis pekerjaan warga bungasari yaitu :

1. Menentukan jumlah cluster untuk rekomendasi jenis pekerjaan warga bungasari yaitu buruh harian lepas, wiraswasta, karyawan swasta, karyawan honorer. Jadi jumlah cluster yang ditentukan adalah sebanyak 4 cluster. Selanjutnya dilakukan perhitungan nilai teloransi, menggunakan parameter jumlah cluster = 4 (buruh harian lepas, jumlah data = 285,) jumlah atribut = 4 (jenis kelamin, umur, pendidikan terakhir, pekerjaan), delta = 0,01, sehingga didapatkan nilai teloransi sebagai berikut :

$$\text{Teloransi error} = \text{delta} * (\text{max-min}) = (0,01*(2-1)) + (0,01*(4-1)) + (0,01*(4-1))$$

$$= 0,01+0,03+0,03$$

$$= 0,07$$

Sehingga, nilai teloransi atau nilai ambang atau *thresthold* pada pengelompokkan pekerjaan berdasarkan data penduduk adalah 0,07.

2. Melakukan inisialisasi pusat cluster (*centroid*) pada data penduduk diubah kedalam bentuk angka untuk mempermudah proses perhitungan dengan algoritma k-means, untuk jenis kelamin L = 1 dan P = 2, untuk meningkatkan hasil maka umur dilakukan pengelompokkan berdasarkan Depkes tahun 2009, sehingga didapat umur 20-25 = 1, 26-35 = 2, 36-45 = 3, dan umur 46-49 = 4, dan pada jenis pekerjaan, wiraswasta = 1, buruh = 2, karyawan swasta = 3 dan karyawan honorer = 4.

Tabel 2.2 Data Penduduk Kel. Bungusari

NIK	JK	Umur	Pendidikan	Pekerjaan
3278091508940002	L	24	SLTA/Sederajat	Wiraswasta
3278092808930001	L	25	SLTA/Sederajat	Wiraswasta
3278091304840001	L	36	SLTA/Sederajat	Buruh Harian Lepas
3206220412810002	L	36	SLTA/Sederajat	Karyawan Swasta
3278090907900005	L	28	SLTA/Sederajat	Buruh Harian Lepas

Tabel 2.3 Data Penduduk Kel. Bungusari setelah dilakukan perubahan

NIK	JK	Umur	Pekerjaan
3278091508940002	L	1	1
3278092808930001	L	1	1
3278091304840001	L	3	2
3206220412810002	L	3	3
3278090907900005	L	2	2

3. Melakukan perhitungan jarak data penduduk terhadap pusat cluster, untuk penentuan awal diasumsikan dengan menggunakan persamaan :

$$ci = \min + \frac{(i-1) * (\max - \min)}{n} + \frac{(\max - \min)}{2 * n}$$

Dimana :

C_i = *centroid* dari kelas ke-i

Min = Nilai minimum pada data

Max = Nilai maksimum pada data

N = Jumlah kelas

Contoh perhitungan pusat awal cluster untuk C1 dan hasilnya pada tabel 2 adalah sebagai berikut :

C1 :

$$\begin{aligned} \text{Jenis kelamin} &= 1 + \frac{(1+1)*(2-1)}{4} + \frac{(2-1)}{2*4} \\ &= 1 + 0 + 0,125 \\ &= 1,125 \end{aligned}$$

$$\begin{aligned} \text{Umur} &= 1 + \frac{(1-1)*(4-1)}{4} + \frac{(4-1)}{2*4} \\ &= 1 + 0 + 0,375 \\ &= 1,375 \end{aligned}$$

$$\begin{aligned} \text{Pekerjaan} &= 1 + \frac{(1-1)*(4-1)}{4} + \frac{(4-1)}{2*4} \\ &= 1 + 0 + 0,375 \\ &= 1,375 \end{aligned}$$

Tabel 2.4 Pusat Awal Cluster

Cluster	JK	Umur	Pekerjaan
C1	1,125	1,375	1,375
C2	1,375	2,125	2,125
C3	1,625	2,875	2,875
C4	1,875	3,625	3,625

Untuk mengukur jarak antara data dengan pusat cluster digunakan *Euclidian distance*, kemudian akan didapatkan matriks jarak sebagai berikut :

$$d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2}$$

C_{ij} = Pusat Cluster

C_{kj} = Data

Σ = Jumlah dari Atribut

Sebagai contoh, perhitungan jarak dari data ke-1 dan ke-3 terhadap pusat cluster adalah :

Perhitungan data ke-1

Jarak data penduduk 1 terhadap pusat *cluster* pertama :

$$\begin{aligned} C1 &= \sqrt{(1,125 - 1)^2 + (1,375 - 1)^2 + (1,375 - 1)^2} \\ &= \sqrt{0,015625 + 0,140625 + 0,140625} \\ &= 0,296875 \end{aligned}$$

Jarak data penduduk 2 terhadap pusat *cluster* kedua :

$$\begin{aligned} C2 &= \sqrt{(1,375 - 1)^2 + (2,125 - 1)^2 + (2,125 - 1)^2} \\ &= \sqrt{0,140625 + 1,265625 + 1,265625} \\ &= 2,671875 \end{aligned}$$

Jarak data penduduk 3 terhadap pusat *cluster* ketiga :

$$\begin{aligned} C3 &= \sqrt{(1,625 - 1)^2 + 2,875 - 1)^2 + (2,875 - 1)^2} \\ &= \sqrt{0,390625 + 3,515625 + 3,515625} \\ &= 7,42188 \end{aligned}$$

Jarak data penduduk 4 terhadap pusat *cluster* keempat :

$$\begin{aligned} C4 &= \sqrt{(1,625 - 1)^2 + (3,625 - 1)^2 + (3,625 - 1)^2} \\ &= \sqrt{0,765625 + 6,890625 + 6,890625} \\ &= 14,54688 \end{aligned}$$

4. Pengelompokkan setiap data penduduk berdasarkan kedekatannya dengan pusat *centroid* untuk data ke-2, data ke-4, ...n. kemudian didapatkan matriks jarak dan kelompok data pada tabel 4.

Tabel 2.5 Hasil Perhitungan dan Pengelompokkan Iterasi 1

NIK	C1	C2	C3	C4	Cluster
3278091508940002	0,296875	2,671875	7,42188	14,54688	1
3278092808930001	0,296875	2,671875	7,42188	14,54688	1
3278091304840001	3,046875	0,921875	1,17188	3,796875	2

3206220412810002	5,296875	1,671875	0,42188	1,546875	3
3278090907900005	0,796875	0,796875	1,92188	6,046875	2

5. Menentukan pusat *cluster* baru, diambil dari hasil rerata iterasi 1 yang sebelumnya sudah dilakukan pada langkah 3. Berikut adalah nilai untuk pusat *cluster* baru :

Tabel 2.6 Nilai Pusat *Cluster* yang Baru

Cluster	JK	Umur	Pekerjaan
C1	1,12	1,7	1,26
C2	1,092683	2,4487	1,873171
C3	1,444444	2,777778	2,944444
C4	1,426667	3,333333	3,916667

6. Ulangi langkah ke 3 (ketiga) hingga hasil rerata pada iterasi lebih kecil dari nilai teloransi error. Iterasi akan terus dilakukan hingga hasil selisih rata-rata dengan nilai *centroid* lebih kecil daripada nilai teloransi error yang telah ditentukan sebelumnya. pada saat nilai hasil rerata pada iterasi lebih kecil dari teloransi error, maka iterasi tidak dilanjutkan. Pada proses penghitungan rerata pada iterasi yang nilainya lebih kecil dari teloransi error adalah pada iterasi ke-5, dimana selisih rerata dan *centroid* iterasi 5 adalah 0. Pada tabel 5 adalah data hasil perhitungan dan pengelompokkan pada iterasi ke 5 atau iterasi akhir.

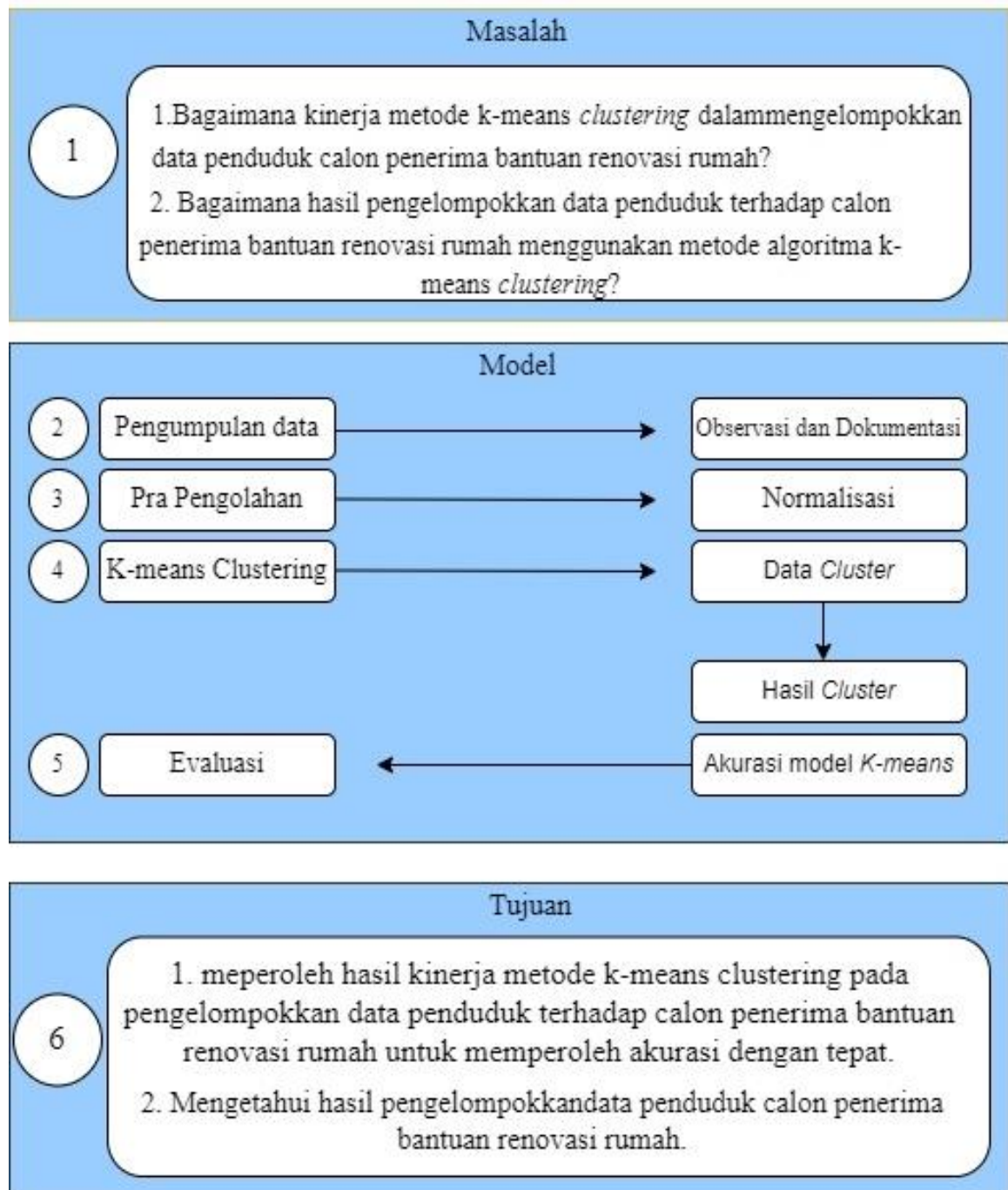
Tabel 2.7 Hasil Perhitungan dan Pengelompokkan Iterasi 5

NIK	C1	C2	C3	C4	<i>Cluster</i>
3278091508940002	1,118662	5,003957	4,76319	11,4325	1
3278092808930001	1,118662	5,003957	4,76219	11,4325	1
3278091304840001	1,472196	0,881507	2,99849	2,5325	2
3206220412810002	3,21967	3,697833	1,88084	0,4325	4
3278090907900005	0,169166	2,350895	1,32202	3,9325	1

Dari hasil pengelompokkan tabel 4.9, diketahui hasil untuk C1 adalah jenis pekerjaan Buruh Harian Lepas dengan rerata kelompok usia adalah 1,848485 atau

pada usia 30 tahun, dengan usia antara 21 sampai 35 tahun. C2 adalah jenis pekerjaan Wiraswasta dengan rerata kelompok umur adalah 3,234694 atau usia 42 tahun, dengan usia antara 36 sampai 49 tahun. C3 adalah jenis pekerjaan Karyawan Swasta dengan rerata kelompok umur adalah 1,661765, atau usia 42 tahun, dengan usia antara 20 sampai 35 tahun. Dan C4 adalah jenis pekerjaan Karyawan Honorer dengan rerata kelompok umur adalah 3,2 atau usia 41 tahun, dengan usia antara 36 sampai 49 tahun.

2.4 Kerangka Pikir



BAB III

METODE PENELITIAN

3.1 Objek Penelitian

Berdasarkan latar belakang dan kerangka pikir seperti yang telah diuraikan pada Bab I dan Bab II, maka yang menjadi objek penelitian adalah Implementasi K-Means Clustering Data Penduduk terhadap Calon Penerima Bantuan Renovasi Rumah di Desa Huidu Melito.

Berikut ini adalah gambaran sistem yang diusulkan menggunakan *Flowchart* sebagai berikut :

3.2 Metode Penelitian

Dilihat dari tingkat penerapan maka penelitian ini merupakan penelitian terapan. Dapat dilihat pula dari jenis informasi yang diolah maka, penelitian ini merupakan penelitian kuantitatif. Metode yang digunakan dalam penelitian ini untuk masalah *Clustering* data penduduk yaitu metode deskriptif, dimana metode ini bertujuan untuk mengungkapkan kebenaran yang objektif. Dengan kata lain menjabarkan suatu keadaan atau situasi yang terjadi saat ini dalam suatu masyarakat dengan menggunakan prosedur ilmiah untuk menjawab sebuah masalah secara akurat.

3.3 Tahap Pengumpulan Data

Tahap pengumpulan data yang digunakan untuk mendapatkan informasi atau data digunakan ada 2 (dua) jenis data yaitu data *Primer* dan data *Sekunder* sebagai berikut :

1. Penelitian Data Primer (Lapangan)

- a. Metode Observasi

Dalam metode ini yang akan dilakukan oleh seorang peneliti adalah melihat serta mempelajari permasalahan yang ada dilapangan yang erat kaitannya dengan objek yang akan diteliti.

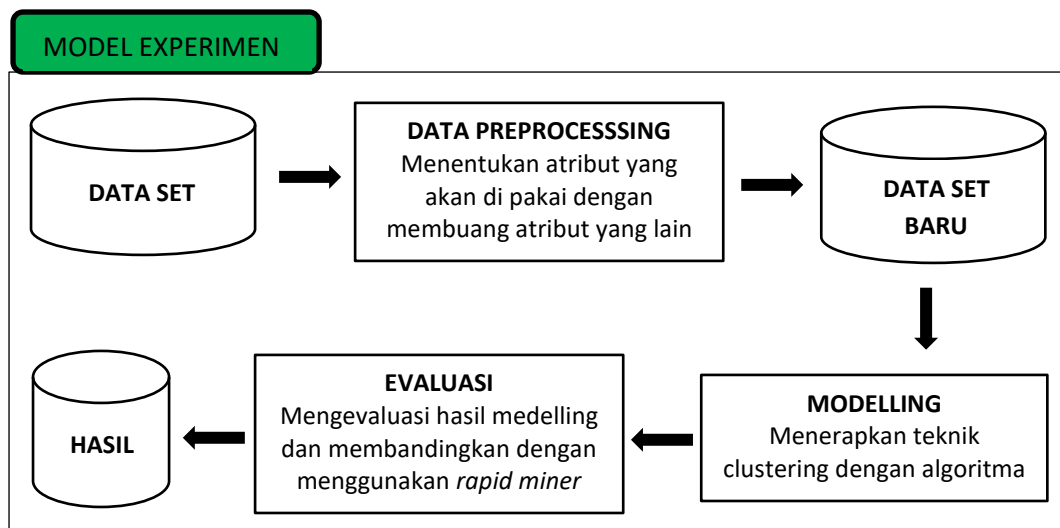
- b. Wawancara

Metode ini digunakan dengan mengajukan beberapa pertanyaan kepada karyawan kantor desa huidu melito untuk mengetahui objek yang menjadi permasalahan yang ada dilapangan.

2. Penelitian Data Sekunder (Kepustakaan)

Metode sekunder ini diperlukan untuk mendapatkan data kepustakaan dengan tujuan melengkapi data primer. Data sekunder didapatkan dari pengkajian kepustakaan yang berisi dasar-dasar teori. Metode sekunder digunakan oleh analisis sistem dengan cara mengambil contoh dokumen-dokumen yang berhubungan dengan materi penelitian. Selain itu, analisis sistem mencari data mengenai hal-hal dan variable yang berupa catatan, buku, majalah dan lain sebagainya yang menjadi penghubung dengan objek penelitian.

3.4 Pemodelan



Gambar 3.2 Pemodelan K-Means

3.4.1 Pra Pengolahan Data

Tahap ini merupakan langkah awal dalam melakukan proses data mining. Dimana sebelum data diolah maka dilakukan proses pembersihan dan normalisasi pada data yang menjadi fokus KDD. Proses pembersihan ini mencakup antara lain membuang data yang sama, membuang data yang inkonsisten dan memperbaiki kesalahan pada data, seperti kesalahan dalam mencetak. hal ini dilakukan karena data yang diolah akan di kelompokkan.

3.4.2 Hasil Clustering

Hasil clustering di dapat dari mengidentifikasi data yang paling dekat dengan pusat *cluter*, kemudian menggabungkan data tersebut kedalam suatu cluster dengan menghitung jarak setiap cluster, lalu mengulangi dari awal sampai semua cluster benar-benar terhubung satu sama lain.

3.4.3 Evaluasi

Setelah mendapatkan hasil kemudian dievaluasi untuk mengetahui hasil kinerja dari metode yang digunakan. Cluster dan output yang telah dihasilkan akan dievaluasi lagi dan setelah itu dimasukkan kedalam data centroid untuk menghitung nilai akurasinya.

BAB IV

HASIL PENELITIAN

4.1 Hasil Pengumpulan Data

Penelitian ini menggunakan data yang berasal dari lokasi penelitian, lokasi penelitian ini berada di Desa Huidu Melito. Penelitian dimulai dengan melakukan observasi langsung ke kantor desa huidu melito, tujuannya adalah untuk mendapatkan data penduduk penerima bantuan renovasi rumah. Sebelum datang ke lokasi penelitian peneliti sudah terlebih dahulu mengantarkan surat izin penelitian untuk mendapatkan izin dari pihak desa untuk meneliti di lokasi tersebut.

Dibawah ini merupakan data yang berasal dari lokasi penelitian, data status kepemilikan rumah akan di konversi kedalam nilai yaitu, status rumah milik sendiri (1), milik orang tua (2), dan kontrak (3).

Tabel 4.1 Hasil Pengumpulan Data

No	NAMA	L/P	Umur	Jumlah tanggungan	Pendapatan	Kepemilikan rumah
1	NOPRI ABJUL	L	38	2	Rp 500.000	1
2	YASMAN HARUN	L	42	3	Rp 1.000.000	1
3	GANI LAMATO	L	41	3	Rp 700.000	2
4	RINTON ISMAIL	L	36	2	Rp 500.000	2
5	RIDWAN KADIR	L	37	5	Rp 750.000	1
6	WIRSON KADIR	L	43	3	Rp 1.500.000	2
7	JON RISKIN HULOPI	L	33	3	Rp 500.000	1
8	SIPRIT RUMOPA	L	31	2	Rp 500.000	1
9	ISWAN HARUN	L	38	4	Rp 1.000.000	3
10	ARIYANTO IKILO	L	42	2	Rp 500.000	3
11	PIDI LATIF	L	32	3	Rp 500.000	3
12	YAMIN HARUN	L	48	3	Rp 500.000	1
13	WINI AJULA	P	45	3	Rp 500.000	1
14	MISRON BANTINGI	L	36	3	Rp 1.500.000	1
15	KASIM MOOWANGGA	L	29	4	Rp 500.000	1
16	HALIM AJULA	L	38	6	Rp 500.000	1
17	HAMJATI ABDULLAH	L	51	2	Rp 1.750.000	1
18	MEI LATIF	P	32	7	Rp 500.000	3
19	OPAN LAMATO	L	27	5	Rp 700.000	1

No	NAMA	L/P	Umur	Jumlah tanggungan	Pendapatan	Kepemilikan rumah
20	SARWIN SUNGE	L	33	2	Rp 1.000.000	1
21	ARSAD LAMATO	L	31	4	Rp 500.000	1
22	UWAN KARIM	L	36	4	Rp 500.000	2
23	NINI MANTALI	L	36	3	Rp 500.000	1
24	ALI NTOROI	L	43	3	Rp 750.000	2
25	JASMIN HARUN	L	28	4	Rp 500.000	1
26	ASMA HARUM	P	37	2	Rp 2.500.000	2
27	KISMAN KASIM	L	31	3	Rp 1.000.000	2
28	RISTO TAPATE	L	27	3	Rp 1.000.000	3
29	HARIS ISINI	L	37	9	Rp 500.000	1
30	YULIANA LAMATO	P	39	4	Rp 750.000	3
31	WIRTIA HARUN	P	30	5	Rp 500.000	1
32	YUSUF SADU	L	29	4	Rp 700.000	2
33	YANTO HABIBI	L	35	4	Rp 500.000	2
34	RAHMAN PATILA	L	32	3	Rp 500.000	1
35	HERSON KANTU	L	36	5	Rp 750.000	2
36	SULEMAN PANGULU	L	38	2	Rp 750.000	2
37	PEKI POLULI	L	36	3	Rp 500.000	1
38	HAMID BOBIHU	L	38	2	Rp 1.500.000	2
39	AIS BOBIHU	L	43	2	Rp 1.000.000	1
40	ERIYANTI ISIINI	P	26	2	Rp 500.000	2
41	UTEN MAHAJANI	L	29	2	Rp 500.000	1
42	NIKO INGO	P	40	6	Rp 500.000	2
43	USMAN TAPATE	L	34	2	Rp 500.000	2
44	PARMAN POKI	L	53	3	Rp 750.000	1
45	IWAN TAHIDINA	L	51	3	Rp 750.000	1
46	RADZAK POU	L	43	4	Rp 500.000	1
47	IWAN LIPUTO	L	45	3	Rp 1.000.000	2
48	RAMDIN LIPUTO	L	34	5	Rp 500.000	1
49	HASAN POMANTO	L	32	6	Rp 500.000	3
50	HARJON LAMATO	L	37	3	Rp 750.000	1
51	ALPIN LAKORO	L	50	4	Rp 500.000	2
52	IRAWATI PULUBUHU	P	35	4	Rp 1.000.000	1
53	ROY OTOLUWA	L	42	3	Rp 2.500.000	1
54	IRPAN SAMIUN	L	33	2	Rp 2.000.000	2
55	FANDI MOHAMAD	L	57	4	Rp 500.000	2
56	ENDANG WAROKA	L	55	3	Rp 1.000.000	2

No	NAMA	L/P	Umur	Jumlah tanggungan	Pendapatan	Kepemilikan rumah
57	RUSTAM KANTU	L	41	3	Rp 500.000	2
58	RAHAMAN J AKULU	L	40	2	Rp 1.000.000	3
59	YUSRIN ABAS	L	38	4	Rp 750.000	2
60	ELIN SADU	P	32	4	Rp 500.000	1
61	YULIANA RACHMAN	P	37	2	Rp 700.000	2
62	UDIN ISMAIL	L	34	3	Rp 1.000.000	3
63	RIYANTO MANTALI	L	43	2	Rp 700.000	1
64	SUMITRO DUDEPO	L	38	3	Rp 1.000.000	1
65	IRFAN SAMIUN	L	32	1	Rp 1.500.000	2
66	YASIN LAUJE	L	28	1	Rp 700.000	2
67	DJAFAR BOBIHU	L	45	2	Rp 1.500.000	1
68	HASMIA ABAY	P	39	2	Rp 500.000	2
69	AMIDIN ISINI	L	41	3	Rp 1.000.000	1
70	JEFRI PULUBUHU	L	37	2	Rp 750.000	1
71	ALPIN ISHAK MONOARFA	L	27	1	Rp 1.000.000	2
72	RAHMAN BOBIHU	L	47	3	Rp 750.000	2
73	SAMSUDIN ALINGGAHE	L	35	1	Rp 750.000	3
74	WARMAN POKI	L	44	2	Rp 1.000.000	2
75	RIYON LAMATO	L	34	2	Rp 1.500.000	1
76	SULEMAN ABDULAH	L	38	3	Rp 700.000	1
77	NOPAN LATIF	L	42	3	Rp 500.000	2
78	YULAN HAJI	P	32	1	Rp 1.000.000	1
79	MARLAN LATIF	L	30	1	Rp 750.000	3
80	YURNI AHMAD	P	29	2	Rp 500.000	2
81	ABDULRAHMAN PULUBUHU	L	27	2	Rp 1.000.000	1
82	DANRESA TANAIYO	L	26	1	Rp 750.000	2
83	AHIM USIRA	L	33	2	Rp 1.500.000	2

Data di atas merupakan data penduduk calon penerima bantuan renovasi rumah di desa huidu melito data tersebut berjumlah 83 data.

4.2 Hasil Pemodelan

4.2.1 Pra Pengolahan Data

Data yang digunakan dalam penelitian ini merupakan data yang berasal dari lokasi penelitian yaitu Desa Huidu Melito, pra pengolahan data sendiri merupakan proses untuk mengenali data yang telah di kumpulkan, data yang sebelumnya tidak dikenal harus di olah untuk bisa dikenali. Hal tersebut bertujuan untuk dapat mengetahui jumlah data yang akan dikelompokkan kedalam *cluster*.

4.2.2 Normalisasi Data

Dalam penelitian ini terdapat 6 variabel yang gunakan untuk menentukan *cluster* yang di inginkan. 6 variabel tersebut adalah (Nama, Jenis Kelamin, Umur, Jumlah Tanggungan, Pendapatan, Kepemilikan Rumah). Dari jumlah variabel tersebut terdapat 3 variabel yang dapat dihitung yaitu Jumlah Tanggungan, Pendapatan, dan kepemilikan rumah yang telah di konversi.

4.2.3 Hasil Hitung Algoritma K-Means

Diketahui :
 Data penduduk : Calon Penerima Bantuan Renovasi Rumah
 Jumlah *cluster* : 3
 Jumlah data : 83
 Jumlah atribut : 3

Data yang digunakan dalam penelitian ini berjumlah 83 data, data tersebut akan di kelompokkan kedalam 3 *cluster* (*Cluster* 1 = Layak, *Cluster* 2 = Kurang layak, *Cluster* 3 = Tidak Layak, dengan menggunakan 3 atribut.

Ket : C1 = Layak
 C2 = Kurang Layak
 C3 = Tidak Layak

Iterasi ke-1

1. Menentukan nilai *centroid*

Dalam menentukan nilai *centroid* data yang digunakan di ambil secara acak. Dalam penelitian ini peneliti mengambil data ke-3, data ke-9, dan data ke-15 sebagai pusat *Cluster*.

Tabel 4.2 penentuan awal *cluster*

NO	NAMA	JUMLAH TANGGUNGAN	PENDAPATAN	KEPEMILIKAN RUMAH
1	NOPRI ABJUL	2	500.000	1
2	YASMAN HARUN	3	1.000.000	1
3	GANI LAMATO	3	700.000	2
4	RINTON ISMAIL	2	500.000	2
5	RIDWAN KADIR	5	750.000	1
6	WIRSON KADIR	3	1.500.000	2
7	JON RISKIN HULOPI	3	500.000	1
8	SIPRIT RUMOPA	2	500.000	1
9	ISWAN HARUN	4	1.000.000	3
10	ARIYANTO IKILO	2	500.000	3
11	PIDI LATIF	3	500.000	3
12	YAMIN HARUN	3	500.000	1
13	WINI AJULA	3	500.000	1
14	MISRON BANTINGI	3	1.500.000	1
15	KASIM MOOWANGGA	4	500.000	1
16	HALIM AJULA	6	500.000	1
17	HAMJATI ABDULLAH	2	1.750.000	1
18	MEI LATIF	7	500.000	3
19	OPAN LAMATO	5	700.000	1
20	SARWIN SUNGE	2	1.000.000	1
21	ARSAD LAMATO	4	500.000	1
22	UWAN KARIM	4	500.000	2
23	NINI MANTALI	3	500.000	1
24	ALI NTOROI	3	750.000	2
25	JASMIN HARUN	4	500.000	1
26	ASMA HARUM	2	2.500.000	2
27	KISMAN KASIM	3	1.000.000	2
28	RISTO TAPATE	3	1.000.000	3
29	HARIS ISINI	9	500.000	1
30	YULIANA LAMATO	4	750.000	3
31	WIRTIA HARUN	5	500.000	1
32	YUSUF SADU	4	700.000	2
33	YANTO HABIBI	4	500.000	2

NO	NAMA	JUMLAH TANGGUNGAN	PENDAPATAN	KEPEMILIKAN RUMAH
34	RAHMAN PATILA	3	500.000	1
35	HERSON KANTU	5	750.000	2
36	SULEMAN PANGULU	2	750.000	2
37	PEKI POLULI	3	500.000	1
38	HAMID BOBIHU	2	1.500.000	2
39	AIS BOBIHU	2	1.000.000	1
40	ERİYANTI ISIINI	2	500.000	2
41	UTEN MAHAJANI	2	500.000	1
42	NIKO INGO	6	500.000	2
43	USMAN TAPATE	2	500.000	2
44	PARMAN POKI	3	750.000	1
45	IWAN TAHIDINA	3	750.000	1
46	RADZAK POU	4	500.000	1
47	IWAN LIPUTO	3	1.000.000	2
48	RAMDIN LIPUTO	5	500.000	1
49	HASAN POMANTO	6	500.000	3
50	HARJON LAMATO	3	750.000	1
51	ALPIN LAKORO	4	500.000	2
52	IRAWATI PULUBUHU	4	1.000.000	1
53	ROY OTOLUWA	3	2.500.000	1
54	IRPAN SAMIUN	2	2.000.000	2
55	FANDI MOHAMAD	4	500.000	2
56	ENDANG WAROKA	3	1.000.000	2
57	RUSTAM KANTU	3	500.000	2
58	RAHAMAN J AKULU	2	1.000.000	3
59	YUSRIN ABAS	4	750.000	2
60	ELIN SADU	4	500.000	1
61	YULIANA RACHMAN	2	700.000	2
62	UDIN ISMAIL	3	1.000.000	3
63	RIYANTO MANTALI	2	700.000	1
64	SUMITRO DUDEPO	3	1.000.000	1
65	IRFAN SAMIUN	1	1.500.000	2
66	YASIN LAUJE	1	700.000	2
67	DJAFAR BOBIHU	2	1.500.000	1

NO	NAMA	JUMLAH TANGGUNGAN	PENDAPATAN	KEPEMILIKAN RUMAH
68	HASMIA ABAY	2	500.000	2
69	AMIDIN ISINI	3	1.000.000	1
70	JEFRI PULUBUHU	2	750.000	1
71	ALPIN ISHAK MONOARFA	1	1.000.000	2
72	RAHMAN BOBIHU	3	750.000	2
73	SAMSUDIN ALINGGAHE	1	750.000	3
74	WARMAN POKI	2	1.000.000	2
75	RIYON LAMATO	2	1.500.000	1
76	SULEMAN ABDULAH	3	700.000	1
77	NOPAN LATIF	3	500.000	2
78	YULAN HAJI	1	1.000.000	1
79	MARLAN LATIF	1	750.000	3
80	YURNI AHMAD	2	500.000	2
81	ABDULRAHMAN PULUBUHU	2	1.000.000	1
82	DANRESA TANAIYO	1	750.000	2
83	AHIM USIRA	2	1.500.000	2

2. Perhitungan jarak pada *Cluster*

Dibawah ini adalah rumus yang akan digunakan dalam menentukan persamaan *euclidean distance space*.

$$C1 = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2}$$

$$C2 = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2}$$

$$C3 = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2}$$

Keterangan

x = data *record*

y = data *cluster*

di bawah ini merupakan cara perhitungan manual.

$$\begin{aligned}
 C1 (1) &= \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2} \\
 &= \sqrt{(2 - 3)^2 + (500.000 - 700.000)^2 + (1 - 2)^2}
 \end{aligned}$$

$$\begin{aligned}
&= \sqrt{(-1)^2 + (-200.000)^2 + (-1)^2} \\
&= \sqrt{1 + 40.000.000.000 + 1} \\
&= \sqrt{40.000.000.002} \\
&= 200.000
\end{aligned}$$

$$\begin{aligned}
C1 (2) &= \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2} \\
&= \sqrt{(3 - 3)^2 + (1.000.000 - 700.000)^2 + (1 - 2)^2} \\
&= \sqrt{(0)^2 + (300.000)^2 + (-1)^2} \\
&= \sqrt{0 + 90.000.000.000 + 1} \\
&= \sqrt{90.000.000.001} \\
&= 300.000
\end{aligned}$$

$$\begin{aligned}
C2 (1) &= \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2} \\
&= \sqrt{(2 - 4)^2 + (500.000 - 500.000)^2 + (1 - 1)^2} \\
&= \sqrt{(-2)^2 + (0)^2 + (0)^2} \\
&= \sqrt{4 + 0 + 0} \\
&= \sqrt{4} \\
&= 2
\end{aligned}$$

$$\begin{aligned}
C2 (2) &= \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2} \\
&= \sqrt{(3 - 4)^2 + (1.000.000 - 500.000)^2 + (1 - 1)^2} \\
&= \sqrt{(-1)^2 + (500.000)^2 + (0)^2} \\
&= \sqrt{1 + 250.000.000.000 + 0} \\
&= \sqrt{250.000.000.001} \\
&= 500.000
\end{aligned}$$

$$\begin{aligned}
C3 (1) &= \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2} \\
&= \sqrt{(2 - 5)^2 + (500.000 - 750.000)^2 + (1 - 1)^2} \\
&= \sqrt{(-3)^2 + (-250.000)^2 + (0)^2} \\
&= \sqrt{9 + 62.500.000.000 + 0} \\
&= \sqrt{62.500.000.009} \\
&= 250.000
\end{aligned}$$

$$\begin{aligned}
 C3 (2) &= \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2} \\
 &= \sqrt{(3 - 5)^2 + (1.000.000 - 750.000)^2 + (1 - 1)^2} \\
 &= \sqrt{(-2)^2 + (250.000)^2 + (0)^2} \\
 &= \sqrt{4 + 62.500.000.000 + 0} \\
 &= \sqrt{62.500.000.004} \\
 &= 250.000
 \end{aligned}$$

Hasil yang di dapatkan dari perhitungan jarak terdekat dengan pusat *cluster*. Jarak yang didapatkan menunjukan bahwa jarak tersebut berada berada satu kelompok dengan *cluster* terdekatnya.

Tabel 4.3 Mencari Jarak Terdekat

JUMLAH TANGGUNGAN	PENDAPATAN	KEPEMILIKAN RUMAH	C1	C2	C3	JARAK TERDEKAT	CLUSTER
2	500.000	1	200000	500000	2	2	C1
3	1.000.000	1	300000	2,236067977	500000	2,236067977	C3
3	700.000	2	0	300000	200000	0	C1
2	500.000	2	200000	500000	2,236067977	2,236067977	C1
5	750.000	1	50000,00005	250000	250000	50000,00005	C2
3	1.500.000	2	800000	500000	1000000	500000	C3
3	500.000	1	200000	500000	1	1	C1
2	500.000	1	200000	500000	2	2	C1
4	1.000.000	3	300000	0	500000	0	C3
2	500.000	3	200000	500000	2,828427125	2,828427125	C1
3	500.000	3	200000	500000	2,236067977	2,236067977	C1
3	500.000	1	200000	500000	1	1	C1
3	500.000	1	200000	500000	1	1	C1
3	1.500.000	1	800000	500000	1000000	500000	C3
4	500.000	1	200000	500000	0	0	C1
6	500.000	1	200000	500000	2	2	C1
2	1.750.000	1	1050000	750000	1250000	750000	C3
7	500.000	3	200000	500000	3,605551275	3,605551275	C1
5	700.000	1	2,236067977	300000	200000	2,236067977	C1
2	1.000.000	1	300000	2,828427125	500000	2,828427125	C3
4	500.000	1	200000	500000	0	0	C1
4	500.000	2	200000	500000	1	1	C1
3	500.000	1	200000	500000	1	1	C1

JUMLAH TANGGUNGAN	PENDAPATAN	KEPEMILIKAN RUMAH	C1	C2	C3	JARAK TERDEKAT	CLUSTER
3	750.000	2	50000	250000	250000	50000	C2
4	500.000	1	200000	500000	0	0	C1
2	2.500.000	2	1800000	1500000	2000000	1500000	C3
3	1.000.000	2	300000	1,414213562	500000	1,414213562	C3
3	1.000.000	3	300000	1	500000	1	C3
9	500.000	1	200000,0001	500000	5	5	C1
4	750.000	3	50000,00002	250000	250000	50000,00002	C2
5	500.000	1	200000	500000	1	1	C1
4	700.000	2	1	300000	200000	1	C1
4	500.000	2	200000	500000	1	1	C1
3	500.000	1	200000	500000	1	1	C1
5	750.000	2	50000,00004	250000	250000	50000,00004	C2
2	750.000	2	50000,00001	250000	250000	50000,00001	C2
3	500.000	1	200000	500000	1	1	C1
2	1.500.000	2	800000	500000	1000000	500000	C3
2	1.000.000	1	300000	2,828427125	500000	2,828427125	C3
2	500.000	2	200000	500000	2,236067977	2,236067977	C1
2	500.000	1	200000	500000	2	2	C1
6	500.000	2	200000	500000	2,236067977	2,236067977	C1
2	500.000	2	200000	500000	2,236067977	2,236067977	C1
3	750.000	1	50000,00001	250000	250000	50000,00001	C2
3	750.000	1	50000,00001	250000	250000	50000,00001	C2
4	500.000	1	200000	500000	0	0	C1
3	1.000.000	2	300000	1,414213562	500000	1,414213562	C3
5	500.000	1	200000	500000	1	1	C1
6	500.000	3	200000	500000	2,828427125	2,828427125	C1
3	750.000	1	50000,00001	250000	250000	50000,00001	C2
4	500.000	2	200000	500000	1	1	C1
4	1.000.000	1	300000	2	500000	2	C3
3	2.500.000	1	1800000	1500000	2000000	1500000	C3
2	2.000.000	2	1300000	1000000	1500000	1000000	C3
4	500.000	2	200000	500000	1	1	C1
3	1.000.000	2	300000	1,414213562	500000	1,414213562	C3
3	500.000	2	200000	500000	1,414213562	1,414213562	C1
2	1.000.000	3	300000	2	500000	2	C3
4	750.000	2	50000,00001	250000	250000	50000,00001	C2
4	500.000	1	200000	500000	0	0	C1
2	700.000	2	1	300000	200000	1	C1
3	1.000.000	3	300000	1	500000	1	C3

JUMLAH TANGGUNGAN	PENDAPATAN	KEPEMILIKAN RUMAH	C1	C2	C3	JARAK TERDEKAT	CLUSTER
2	700.000	1	1,414213562	300000	200000	1,414213562	C1
3	1.000.000	1	300000	2,236067977	500000	2,236067977	C3
1	1.500.000	2	800000	500000	1000000	500000	C3
1	700.000	2	2	300000	200000	2	C1
2	1.500.000	1	800000	500000	1000000	500000	C3
2	500.000	2	200000	500000	2,236067977	2,236067977	C1
3	1.000.000	1	300000	2,236067977	500000	2,236067977	C3
2	750.000	1	50000,00002	250000	250000	50000,00002	C2
1	1.000.000	2	300000	3,16227766	500000	3,16227766	C3
3	750.000	2	50000	250000	250000	50000	C2
1	750.000	3	50000,00005	250000	250000	50000,00005	C2
2	1.000.000	2	300000	2,236067977	500000	2,236067977	C3
2	1.500.000	1	800000	500000	1000000	500000	C3
3	700.000	1	1	300000	200000	1	C1
3	500.000	2	200000	500000	1,414213562	1,414213562	C1
1	1.000.000	1	300000	3,605551275	500000	3,605551275	C3
1	750.000	3	50000,00005	250000	250000	50000,00005	C2
2	500.000	2	200000	500000	2,236067977	2,236067977	C1
2	1.000.000	1	300000	2,828427125	500000	2,828427125	C3
1	750.000	2	50000,00004	250000	250000	50000,00004	C2
2	1.500.000	2	800000	500000	1000000	500000	C3

3. Menentukan pusat *cluster* baru

Pada *cluster* yang sudah terbentuk sebelumnya dapat diketahui bahwa *Cluster* 1 (C1) terdiri dari 41 data, *Cluster* 2 (C2) terdiri dari 14 data, dan *Cluster* 3 (C3), terdiri dari 28 data. Berikut ini merupakan tabel hasil perhitungan *cluster* sebelumnya.

Tabel 4.4 Jumlah Perhitungan iterasi pertama

<i>Cluster</i>	Jumlah Data
C1	41 Data
C2	14 Data
C3	28 Data

Dari data di atas kita dapat menghitung kembali centroidnya untuk dapat menentukan *centroid* baru.

Untuk menghitung centroid baru C1, jumlahkan seluruh data *Cluster* 1 untuk jumlah tanggungan, pendapatan dan juga kepemilikan rumah lalu bagi data tersebut dengan jumlah keseluruhan *Cluster* 1 maka hasil yang di dapatkan merupakan hasil *centroid* baru *Cluster* 1.

Untuk menghitung centroid baru C2, jumlahkan seluruh data *Cluster* 2 untuk jumlah tanggungan, pendapatan dan juga kepemilikan rumah lalu bagi data tersebut dengan jumlah keseluruhan *Cluster* 1 maka hasil yang di dapatkan merupakan hasil *centroid* baru *Cluster* 2.

Untuk menghitung centroid baru C3, jumlahkan seluruh data *Cluster* 3 untuk jumlah tanggungan, pendapatan dan juga kepemilikan rumah lalu bagi data tersebut dengan jumlah keseluruhan *Cluster* 1 maka hasil yang di dapatkan merupakan hasil *centroid* baru *Cluster* 3.

Iterasi ke-2

Ulangi data pada iterasi ke-2 hingga data tidak mengalami perubahan.

Tabel 4.5 penentuan *cluster* baru

NO	<i>Cluster</i>	JumlahTanggungan	Pendapatan	Kepemilikan Rumah
1	C1	3,536585366	534146,3415	1,585365854
2	C2	2,857142857	750000	1,857142857
3	C3	2,428571429	1294642,857	1,678571429

Sama seperti halnya pada perhitunagan iterasi pertama, untuk iterasi kedua juga menggunakan rumus *euclidean distance*.

$$C1 = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2}$$

$$C2 = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2}$$

$$C3 = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2}$$

Berikut ini merupakan Hasil iterasi ke-2

Tabel 4.6 Hasil iterasi ke-2

C1	C2	C3	JARAK TERDEKAT	CLUSTER
34146,3415	250000	794642,8571	34146,3415	C1
465853,6585	250000	294642,8571	250000	C2
165853,6585	50000	594642,8571	50000	C2
34146,3415	250000	794642,8571	34146,3415	C1
215853,6585	2,307927774	544642,8571	2,307927774	C2
965853,6585	750000	205357,1429	205357,1429	C3
34146,34147	250000	794642,8571	34146,34147	C1
34146,3415	250000	794642,8571	34146,3415	C1
465853,6585	250000	294642,8572	250000	C2
34146,34153	250000	794642,8571	34146,34153	C1
34146,3415	250000	794642,8571	34146,3415	C1
34146,34147	250000	794642,8571	34146,34147	C1
34146,34147	250000	794642,8571	34146,34147	C1
965853,6585	750000	205357,1429	205357,1429	C3
34146,34147	250000	794642,8571	34146,34147	C1
34146,34156	250000	794642,8572	34146,34156	C1
1215853,659	1000000	455357,1429	455357,1429	C3
34146,34167	250000	794642,8572	34146,34167	C1
165853,6585	50000,00005	594642,8571	50000,00005	C2
465853,6585	250000	294642,8571	250000	C2
34146,34147	250000	794642,8571	34146,34147	C1
34146,34147	250000	794642,8571	34146,34147	C1
34146,34147	250000	794642,8571	34146,34147	C1
215853,6585	0,202030509	544642,8571	0,202030509	C2
34146,34147	250000	794642,8571	34146,34147	C1
1965853,659	1750000	1205357,143	1205357,143	C3
465853,6585	250000	294642,8571	250000	C2
465853,6585	250000	294642,8571	250000	C2
34146,34191	250000,0001	794642,8572	34146,34191	C1
215853,6585	1,616244071	544642,8571	1,616244071	C2
34146,3415	250000	794642,8571	34146,3415	C1
165853,6585	50000,00001	594642,8571	50000,00001	C2
34146,34147	250000	794642,8571	34146,34147	C1
34146,34147	250000	794642,8571	34146,34147	C1
215853,6585	2,147613768	544642,8571	2,147613768	C2
215853,6585	0,868966076	544642,8571	0,868966076	C2

C1	C2	C3	JARAK TERDEKAT	CLUSTER
34146,34147	250000	794642,8571	34146,34147	C1
965853,6585	750000	205357,1429	205357,1429	C3
465853,6585	250000	294642,8571	250000	C2
34146,3415	250000	794642,8571	34146,3415	C1
34146,3415	250000	794642,8571	34146,3415	C1
34146,34155	250000	794642,8572	34146,34155	C1
34146,3415	250000	794642,8571	34146,3415	C1
215853,6585	0,868966076	544642,8571	0,868966076	C2
215853,6585	0,868966076	544642,8571	0,868966076	C2
34146,34147	250000	794642,8571	34146,34147	C1
465853,6585	250000	294642,8571	250000	C2
34146,3415	250000	794642,8571	34146,3415	C1
34146,34158	250000	794642,8572	34146,34158	C1
215853,6585	0,868966076	544642,8571	0,868966076	C2
34146,34147	250000	794642,8571	34146,34147	C1
465853,6585	250000	294642,8571	250000	C2
1965853,659	1750000	1205357,143	1205357,143	C3
1465853,659	1250000	705357,1429	705357,1429	C3
34146,34147	250000	794642,8571	34146,34147	C1
465853,6585	250000	294642,8571	250000	C2
34146,34147	250000	794642,8571	34146,34147	C1
465853,6585	250000	294642,8571	250000	C2
215853,6585	1,151751107	544642,8571	1,151751107	C2
34146,34147	250000	794642,8571	34146,34147	C1
165853,6585	50000,00001	594642,8571	50000,00001	C2
465853,6585	250000	294642,8571	250000	C2
165853,6585	50000,00001	594642,8571	50000,00001	C2
465853,6585	250000	294642,8571	250000	C2
965853,6585	750000	205357,1429	205357,1429	C3
165853,6586	50000,00003	594642,8571	50000,00003	C2
965853,6585	750000	205357,1429	205357,1429	C3
34146,3415	250000	794642,8571	34146,3415	C1
465853,6585	250000	294642,8571	250000	C2
215853,6585	1,212183053	544642,8571	1,212183053	C2
465853,6585	250000	294642,8571	250000	C2
215853,6585	0,202030509	544642,8571	0,202030509	C2
215853,6586	2,180619646	544642,8571	2,180619646	C2
465853,6585	250000	294642,8571	250000	C2
965853,6585	750000	205357,1429	205357,1429	C3

C1	C2	C3	JARAK TERDEKAT	CLUSTER
165853,6585	50000,00001	594642,8571	50000,00001	C2
34146,34147	250000	794642,8571	34146,34147	C1
465853,6585	250000	294642,8571	250000	C2
215853,6586	2,180619646	544642,8571	2,180619646	C2
34146,3415	250000	794642,8571	34146,3415	C1
465853,6585	250000	294642,8571	250000	C2
215853,6586	1,862629259	544642,8571	1,862629259	C2
965853,6585	750000	205357,1429	205357,1429	C3

Pada perhitungan iterasi ke-2 terdapat perbedaan hasil *cluster* dengan iterasi pertama. Pada iterasi pertama C1 terdapat 41 data, C2 14 data dan C3 terdapat 28 data. Hal tersebut berbeda dengan hasil iterasi ke 2 dimana pada iterasi ke-2 C1 terdapat 34 data, C2 38 data dan C3 terdapat 11 data, dari hasil tersebut harus di lakukan perhitungan iterasi ke-3.

Tabel 4.7 Jumlah Perhitungan iterasi kedua

<i>Cluster</i>	Jumlah Data
C1	41
C2	14
C3	28

Iterasi Ke-3

Dikarenakan terdapat adanya perbedaan antara iterasi pertama dan kedua maka akan dilakukan perhitungan pada iterasi ke-3.

Tabel 4.8 Penentuan *cluster* iterasi ke-3

NO	Cluster	Jumlah Tanggungan	Pendapatan	Kepemilikan Rumah
1	C1	767647,0588	1,545454545	245589,3838
2	C2	809279,7784	1,640826137	339474,1576
3	C3	813636,3636	1,545454545	359091,3941

Untuk Penentuan *cluster* pada iterasi ke tiga maka dilakukan perhitungan yang sama untuk menentukan *cluster* iterasi pertama dan *cluster* iterasi kedua,

rumus yang digunakanpun masih sama yaitu dengan menggunakan rumus *euclidean distance*.

Penggunaan rumus *euclidean distance* pada iterasi ketiga.

$$C1 = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2}$$

$$C2 = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2}$$

$$C3 = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2}$$

Hasil *cluster* iterasi ke-3

Tabel 4.9 hasil *cluster* iterasi ke-3

C1	C2	C3	JARAK TERDEKAT	CLUSTER
1,776675095	352631,5789	1250000	1,776675095	C1
500000	147368,4211	750000	147368,4211	C2
200000	152631,5789	1050000	152631,5789	C2
2,609703744	352631,5789	1250000	2,609703744	C1
250000	102631,579	1000000	102631,579	C2
1000000	647368,4211	250000	250000	C3
1,207316221	352631,5789	1250000	1,207316221	C1
1,952063942	352631,5789	1250000	1,952063942	C1
500000	147368,4211	750000	147368,4211	C2
3,436648605	352631,579	1250000	3,436648605	C1
3,075323147	352631,5789	1250000	3,075323147	C1
1,207316221	352631,5789	1250000	1,207316221	C1
1,207316221	352631,5789	1250000	1,207316221	C1
1000000	647368,4211	250000	250000	C3
1,051033434	352631,579	1250000	1,051033434	C1
2,529582758	352631,579	1250000	2,529582758	C1
1250000	897368,4211	1,016394535	1,016394535	C3
4,477258955	352631,579	1250000	4,477258955	C1
200000	152631,579	1050000	152631,579	C2
500000	147368,4211	750000	147368,4211	C2
1,051033434	352631,579	1250000	1,051033434	C1
2,025998835	352631,5789	1250000	2,025998835	C1
1,207316221	352631,5789	1250000	1,207316221	C1
250000	102631,5789	1000000	102631,5789	C2
1,051033434	352631,579	1250000	1,051033434	C1
2000000	1647368,421	750000	750000	C3

C1	C2	C3	JARAK TERDEKAT	CLUSTER
500000	147368,4211	750000	147368,4211	C2
500000	147368,4211	750000	147368,4211	C2
5,416637832	352631,579	1250000	5,416637832	C1
250000	102631,579	1000000	102631,579	C2
1,658833959	352631,579	1250000	1,658833959	C1
200000	152631,579	1050000	152631,579	C2
2,025998835	352631,5789	1250000	2,025998835	C1
1,207316221	352631,5789	1250000	1,207316221	C1
250000	102631,579	1000000	102631,579	C2
250000	102631,579	1000000	102631,579	C2
1,207316221	352631,5789	1250000	1,207316221	C1
1000000	647368,4211	250000	250000	C3
500000	147368,4211	750000	147368,4211	C2
2,609703744	352631,5789	1250000	2,609703744	C1
1,952063942	352631,5789	1250000	1,952063942	C1
3,065744433	352631,579	1250000	3,065744433	C1
2,609703744	352631,5789	1250000	2,609703744	C1
250000	102631,579	1000000	102631,579	C2
250000	102631,579	1000000	102631,579	C2
1,051033434	352631,579	1250000	1,051033434	C1
500000	147368,4211	750000	147368,4211	C2
1,658833959	352631,579	1250000	1,658833959	C1
3,794573616	352631,579	1250000	3,794573616	C1
250000	102631,579	1000000	102631,579	C2
2,025998835	352631,5789	1250000	2,025998835	C1
500000	147368,4211	750000	147368,4211	C2
2000000	1647368,421	750000	750000	C3
1500000	1147368,421	250000	250000	C3
2,025998835	352631,5789	1250000	2,025998835	C1
500000	147368,4211	750000	147368,4211	C2
2,111305865	352631,5789	1250000	2,111305865	C1
500000	147368,4211	750000	147368,4211	C2
250000	102631,579	1000000	102631,579	C2
1,051033434	352631,579	1250000	1,051033434	C1
200000	152631,5789	1050000	152631,5789	C2
500000	147368,4211	750000	147368,4211	C2
200000	152631,579	1050000	152631,579	C2
500000	147368,4211	750000	147368,4211	C2
1000000	647368,4211	250000	250000	C3

C1	C2	C3	JARAK TERDEKAT	CLUSTER
200000	152631,579	1050000	152631,579	C2
1000000	647368,4211	250000	250000	C3
2,609703744	352631,5789	1250000	2,609703744	C1
500000	147368,4211	750000	147368,4211	C2
250000	102631,579	1000000	102631,579	C2
500000	147368,4211	750000	147368,4211	C2
250000	102631,5789	1000000	102631,5789	C2
250000	102631,579	1000000	102631,579	C2
500000	147368,4211	750000	147368,4211	C2
1000000	647368,4211	250000	250000	C3
200000	152631,5789	1050000	152631,5789	C2
2,111305865	352631,5789	1250000	2,111305865	C1
500000	147368,4211	750000	147368,4211	C2
250000	102631,579	1000000	102631,579	C2
2,609703744	352631,5789	1250000	2,609703744	C1
500000	147368,4211	750000	147368,4211	C2
250000	102631,579	1000000	102631,579	C2
1000000	647368,4211	250000	250000	C3

Pada perhitungan yang telah di lakukan tadi, perhitungan tersebut telah berhenti pada iterasi ke-3. Karena hasil perhitungan pada iterasi ke-2 dan iterasi ke-3 sama. Pada iterasi ke-2 data yang di dapatkan yaitu C1 = 41 data, C2 = 14 data, dan C3 = 28 data.

Tabel 4.10 Jumlah Perhitungan iterasi ketiga

Cluster	Jumlah Data
C1	41
C2	14
C3	28

4.2.4 Hasil Clustering

Data yang telah di olah sebelumnya maka dapat di simpulkan hasil clusternya, yaitu *cluster* 1 (C1) penduduk layak terdapat sebanyak 41 data, dikatakan layak karena jumlah tanggungan dalam keluarga lebih dari 1 dan untuk pendapatan perbulan berada di kisaran Rp 500.000 atau kurang sedangkan status

kepemilikan rumah masih tinggal bersama orang tua ataupun masih mengontrak. *Cluster 2 (C2)* penduduk kurang layak terdapat 38 data, di cluster 2 terdapat 14 data penduduk kurang layak hal tersebut dikarenakan jumlah tanggungan dalam keluarga ada yang hanya memiliki 1 tanggungan dengan pendapatan perbulan di kisaran Rp 750.000 dan ada penduduk yang sudah memiliki rumah sendiri hal tersebutlah yang menyebabkan 14 penduduk masuk dalam kategori penduduk kurang layak. Dan *cluster 3 (C3)* terdapat 28 data, 28 data yang merupakan penduduk kurang layak tersebut di karenakan penduduknya memiliki pendapatan di atas Rp 1.000.000 dan memiliki rumah sendiri, dengan pendapatan dan kepemilikan rumah tersebut maka penduduk itu bisa memenuhi kebutuhan sehari hari dan juga dapat merenovasi rumahnya dengan penghasilan yang ia dapatkan setiap bulannya.

BAB V

PEMBAHASAN

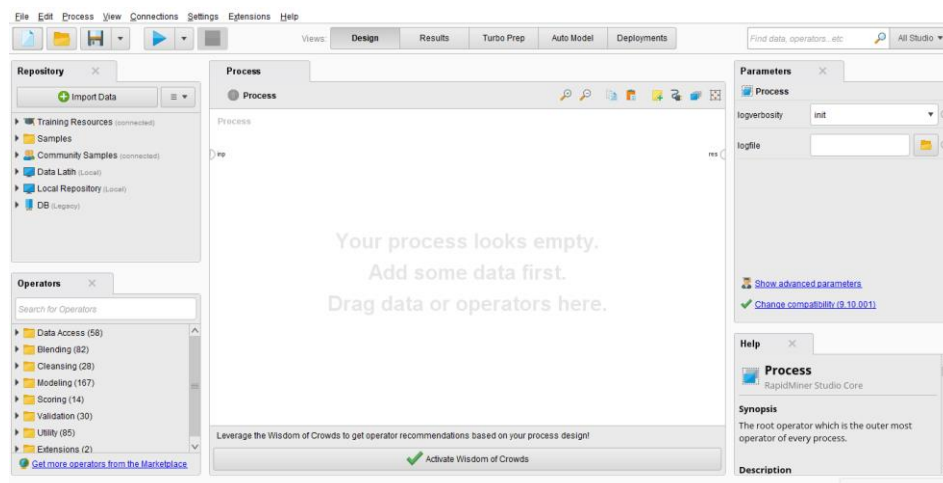
5.1 Pembahasan Model

Pemodelan yang digunakan dalam penelitian ini merupakan pemodelan dari k-means. Pemodelan k-means sendiri di mulai dari tahap pengumpulan data set di lokasi penelitian, kemudian dilanjutkan dengan *preprocessing* data yaitu menentukan atribut yang akan di pakai dengan membuang atribut yang lain, setelah melakukan *preprocessing* data kemudian di lakukan pembuatan data set baru, data set baru merupakan data yang sudah di kelola sebelumnya. Setelah itu tahap *modelling* di mana tahap ini merupakan tahap untuk menerapkan teknik *clustering* dengan menggunakan algoritma yang telah di pilih sebelumnya, setelah tahap modelling kemudian dilakukan evaluasi, evaluasi hasil modeling dan membandingkan dengan menggunakan tools rapidminer sehingga mendapatkan hasil yang di capai.

5.2 Pembahasan Tools

Tools yang akan digunakan dalam penelitan ini adalah tools rapidminer, berikut ini merupakan penjelasann tentang tools rapidminer.

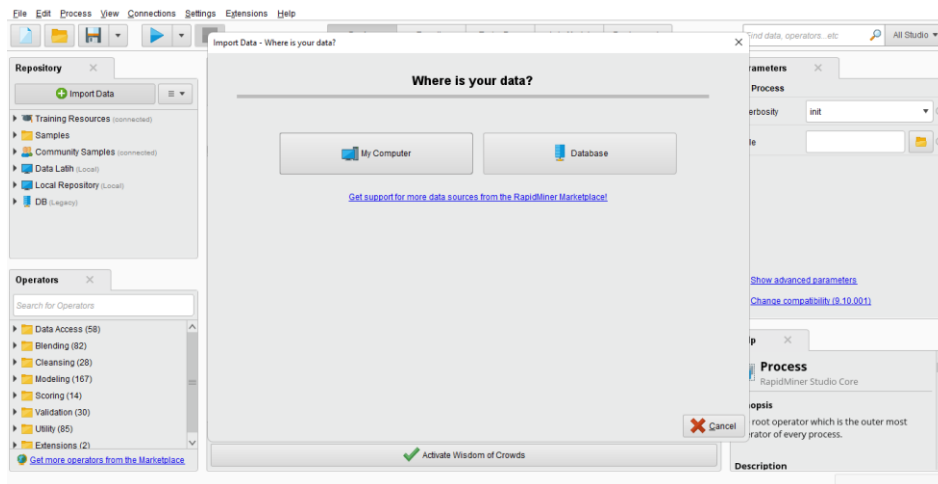
1. Tampilan Utama Rapidminer



Gambar 5.1 Tampilan Utama Rapidminer

Tampilan pada gambar di atas merupakan tampilan utama pada tools rapidminer.

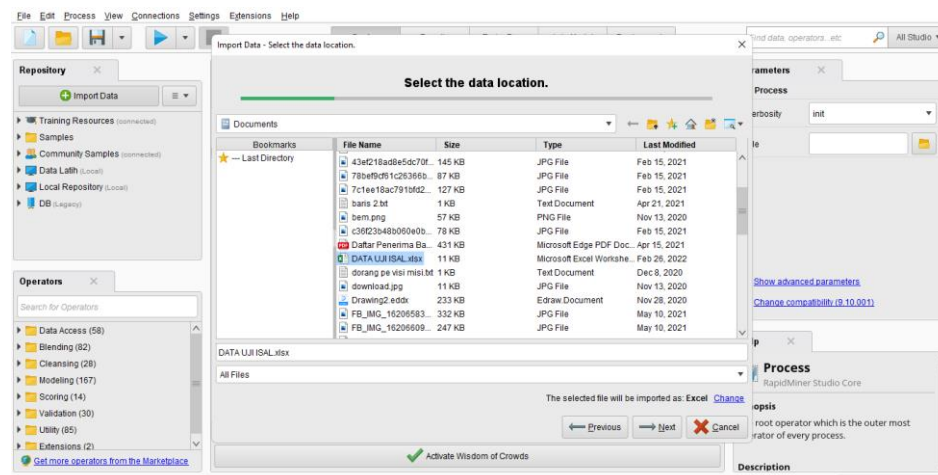
2. Tampilan Import Data



Gambar 5.2 tampilan import data

Tampilan import data merupakan tampilan yang akan mengarahkan pengguna untuk mencari data yang akan digunakan.

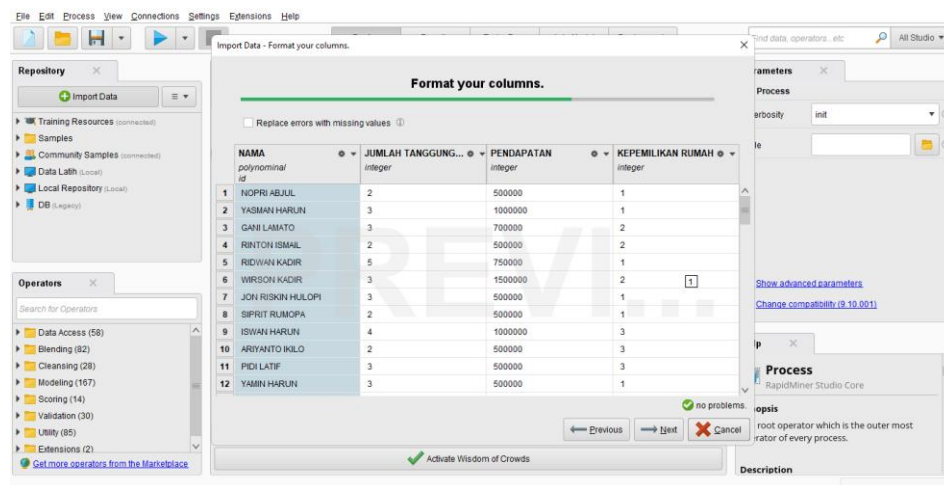
3. Tampilan Lokasi Penyimpanan Data



Gambar 5.3 Tampilan Lokasi Penyimpanan Data

Pada tampilan ini terdapat berbagai lokasi penyimpanan data, cari tempat dimana data disimpan.

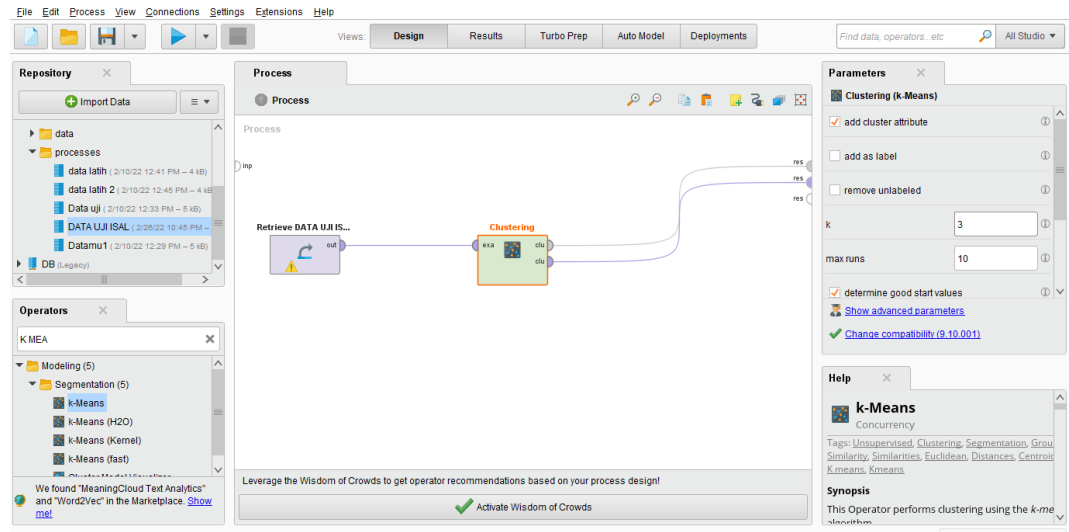
4. Tampilan Data Yang Akan Digunakan



Gambar 5.4 Tampilan Data Yang Akan Digunakan

Tampilan ini merupakan data yang akan digunakan, data tersebut berformat excel.

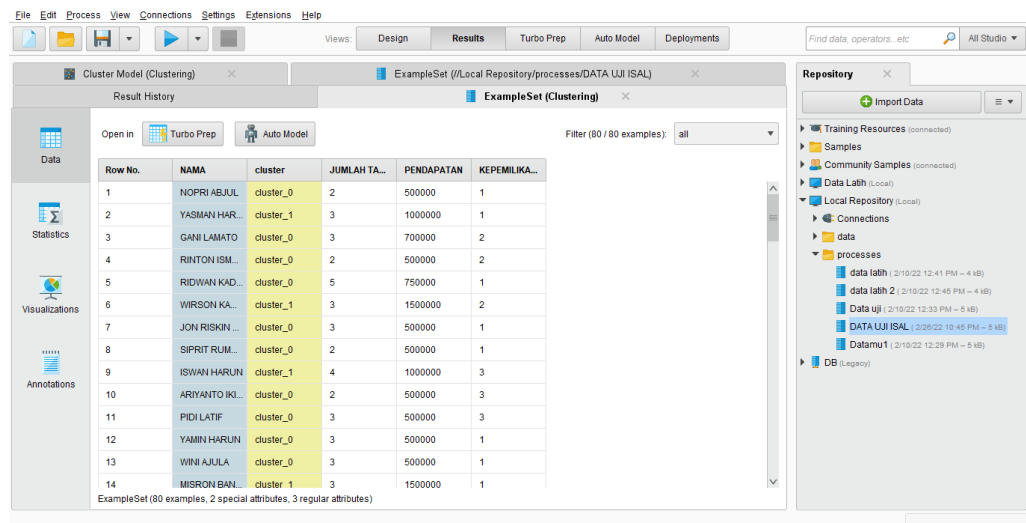
5. Tampilan Desain



Gambar 5.5 Tampilan Desain

Gambar di atas merupakan tampilan desain yang digunakan, tampilan tersebut terdiri dari yang telah di import sebelumnya dan juga sistem operasi k-means yang digunakan. Setelah keduanya di hubungkan maka selanjutnya klik ikon segitiga biru (run) yang berada di bagian kiri atas.

6. Tampilan result data

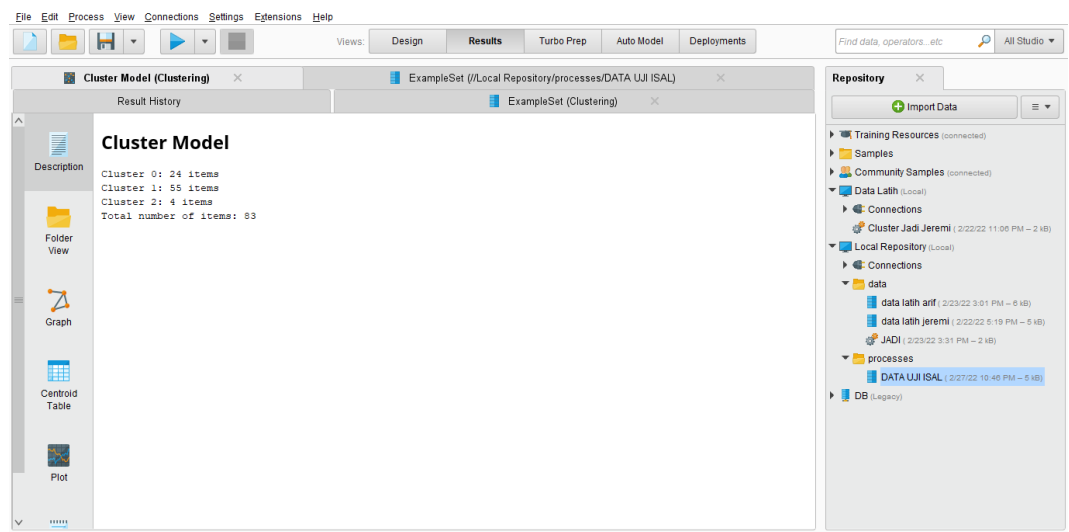


Row No.	NAMA	cluster	JUMLAH TA...	PENDAPATAN	KEPEMILIKAN
1	NOPRI ABUJUL	cluster_0	2	500000	1
2	YASMAN HAR...	cluster_1	3	1000000	1
3	GANI LAMATO	cluster_0	3	700000	2
4	RINTON ISM...	cluster_0	2	500000	2
5	RIDWAN KAD...	cluster_0	5	750000	1
6	WIRSON KA...	cluster_1	3	1500000	2
7	JON RISKIN ...	cluster_0	3	500000	1
8	SIPRIT RUM...	cluster_0	2	500000	1
9	ISWAN HARUN	cluster_1	4	1000000	3
10	ARIYANTO IKI...	cluster_0	2	500000	3
11	PIDI LATIF	cluster_0	3	500000	3
12	YAMIN HARUN	cluster_0	3	500000	1
13	WINI AJULA	cluster_0	3	500000	1
14	MISRON BAN...	cluster_1	3	1500000	1

Gambar 5.6 Tampilan Result Data

Tampilan result data merupakan semua data yang sudah di olah lengkap dengan clusternya.

7. Tampilan Cluster Model



Cluster Model

Cluster 0: 24 items
 Cluster 1: 55 items
 Cluster 2: 4 items
 Total number of items: 83

Gambar 5.7 tampilan cluster model

Tampilan cluster model merupakan tampilan yang menunjukkan clusterisasi keseluruhan data yang digunakan dimana dalam tampilan ini jumlah data yang digunakan akan terlihat beserta ketiga cluster yang sudah ditentukan sebelumnya.

5.3 Pembahasan Hasil Algoritma

Pada pembahasan hasil algoritma telah diketahui bahwa penggunaan algoritma k-means sangat membantu dalam penelitian ini, efektifitas penggunaan algoritma k-means sangat berpengaruh dalam penelitian ini. Data yang digunakan dalam penelitian ini berjumlah 83 data dan data tersebut akan dikelompokkan kedalam 3 *cluster*. *Cluster* pertama (C1) penduduk layak terdapat 41 data. *Cluster* kedua (C2) penduduk kurang layak terdapat 14 data. Dan *cluster* ketiga (C3) penduduk tidak layak terdapat 28 data. Dengan hasil tersebut penggunaan algoritma k-means dirasa sangat membantu pengelompokan data pada penelitian yang dilakukan.

Dibawah ini merupakan tabel hasil *cluster* yang sudah dilakukan :

Tabel 5.1 hasil *clustering*

JUMLAH TANGGUNGAN	PENDAPATAN	KEPEMILIKAN RUMAH	NAMA	Cluster
1	1500000	2	IRFAN SAMIUN	<i>cluster_0</i>
1	1000000	2	ALPIN ISHAK MONOARFA	<i>cluster_0</i>
1	1000000	1	YULAN HAJI	<i>cluster_0</i>
2	1500000	2	HAMID BOBIHU	<i>cluster_0</i>
2	1500000	1	DJAFAR BOBIHU	<i>cluster_0</i>
2	1500000	1	RIYON LAMATO	<i>cluster_0</i>
2	1500000	2	AHIM USIRA	<i>cluster_0</i>
2	1000000	1	SARWIN SUNGE	<i>cluster_0</i>
2	1000000	1	AIS BOBIHU	<i>cluster_0</i>
2	1000000	3	RAHAMAN J AKULU	<i>cluster_0</i>
2	1000000	2	WARMAN POKI	<i>cluster_0</i>
2	1000000	1	ABDULRAHMAN PULUBUHU	<i>cluster_0</i>
3	1500000	2	WIRSON KADIR	<i>cluster_0</i>
3	1500000	1	MISRON BANTINGI	<i>cluster_0</i>
3	1000000	1	YASMAN HARUN	<i>cluster_0</i>
3	1000000	2	KISMAN KASIM	<i>cluster_0</i>
3	1000000	3	RISTO TAPATE	<i>cluster_0</i>
3	1000000	2	IWAN LIPUTO	<i>cluster_0</i>
3	1000000	2	ENDANG WAROKA	<i>cluster_0</i>
3	1000000	3	UDIN ISMAIL	<i>cluster_0</i>

JUMLAH TANGGUNGAN	PENDAPATAN	KEPEMILIKAN RUMAH	NAMA	Cluster
3	1000000	1	SUMITRO DUDEPO	<i>cluster_0</i>
3	1000000	1	AMIDIN ISINI	<i>cluster_0</i>
4	1000000	3	ISWAN HARUN	<i>cluster_0</i>
4	1000000	1	IRAWATI PULUBUHU	<i>cluster_0</i>
1	750000	3	SAMSUDIN ALINGGAHE	<i>cluster_1</i>
1	750000	3	MARLAN LATIF	<i>cluster_1</i>
1	750000	2	DANRESA TANAIYO	<i>cluster_1</i>
1	700000	2	YASIN LAUJE	<i>cluster_1</i>
2	750000	2	SULEMAN PANGULU	<i>cluster_1</i>
2	750000	1	JEFRI PULUBUHU	<i>cluster_1</i>
2	700000	2	YULIANA RACHMAN	<i>cluster_1</i>
2	700000	1	RIYANTO MANTALI	<i>cluster_1</i>
2	500000	1	NOPRI ABJUL	<i>cluster_1</i>
2	500000	2	RINTON ISMAIL	<i>cluster_1</i>
2	500000	1	SIPRIT RUMOPA	<i>cluster_1</i>
2	500000	3	ARIYANTO IKILO	<i>cluster_1</i>
2	500000	2	ERIYANTI ISIINI	<i>cluster_1</i>
2	500000	1	UTEN MAHAJANI	<i>cluster_1</i>
2	500000	2	USMAN TAPATE	<i>cluster_1</i>
2	500000	2	HASMIA ABAY	<i>cluster_1</i>
2	500000	2	YURNI AHMAD	<i>cluster_1</i>
3	750000	2	ALI NTOROI	<i>cluster_1</i>
3	750000	1	PARMAN POKI	<i>cluster_1</i>
3	750000	1	IWAN TAHIDINA	<i>cluster_1</i>
3	750000	1	HARJON LAMATO	<i>cluster_1</i>
3	750000	2	RAHMAN BOBIHU	<i>cluster_1</i>
3	700000	2	GANI LAMATO	<i>cluster_1</i>
3	700000	1	SULEMAN ABDULAH	<i>cluster_1</i>
3	500000	1	JON RISKIN HULOPI	<i>cluster_1</i>
3	500000	3	PIDI LATIF	<i>cluster_1</i>
3	500000	1	YAMIN HARUN	<i>cluster_1</i>
3	500000	1	WINI AJULA	<i>cluster_1</i>
3	500000	1	NINI MANTALI	<i>cluster_1</i>
3	500000	1	RAHMAN PATILA	<i>cluster_1</i>
3	500000	1	PEKI POLULI	<i>cluster_1</i>
3	500000	2	RUSTAM KANTU	<i>cluster_1</i>

JUMLAH TANGGUNGAN	PENDAPATAN	KEPEMILIKAN RUMAH	NAMA	Cluster
3	500000	2	NOPAN LATIF	<i>cluster_1</i>
4	750000	3	YULIANA LAMATO	<i>cluster_1</i>
4	750000	2	YUSRIN ABAS	<i>cluster_1</i>
4	700000	2	YUSUF SADU	<i>cluster_1</i>
4	500000	1	KASIM MOOWANGGA	<i>cluster_1</i>
4	500000	1	ARSAD LAMATO	<i>cluster_1</i>
4	500000	2	UWAN KARIM	<i>cluster_1</i>
4	500000	1	JASMIN HARUN	<i>cluster_1</i>
4	500000	2	YANTO HABIBI	<i>cluster_1</i>
4	500000	1	RADZAK POU	<i>cluster_1</i>
4	500000	2	ALPIN LAKORO	<i>cluster_1</i>
4	500000	2	FANDI MOHAMAD	<i>cluster_1</i>
4	500000	1	ELIN SADU	<i>cluster_1</i>
5	750000	1	RIDWAN KADIR	<i>cluster_1</i>
5	750000	2	HERSON KANTU	<i>cluster_1</i>
5	700000	1	OPAN LAMATO	<i>cluster_1</i>
5	500000	1	WIRTIA HARUN	<i>cluster_1</i>
5	500000	1	RAMDIN LIPUTO	<i>cluster_1</i>
6	500000	1	HALIM AJULA	<i>cluster_1</i>
6	500000	2	NIKO INGO	<i>cluster_1</i>
6	500000	3	HASAN POMANTO	<i>cluster_1</i>
7	500000	3	MEI LATIF	<i>cluster_1</i>
9	500000	1	HARIS ISINI	<i>cluster_1</i>
2	2500000	2	ASMA HARUM	<i>cluster_2</i>
2	2000000	2	IRPAN SAMIUN	<i>cluster_2</i>
2	1750000	1	HAMJATI ABDULLAH	<i>cluster_2</i>
3	2500000	1	ROY OTOLUWA	<i>cluster_2</i>

BAB VI

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Berdasarkan penelitian yang dilakukan oleh peneliti maka dapat di tarik kesimpulan sebagai berikut :

1. Metode yang digunakan dalam penelitian ini adalah metode algoritma k-means, metode ini mampu mengelompokan data yang banyak berdasarkan karakteristik yang telah di tentukan sebelumnya. Penggunaan metode ini sangat efektif dimana data yang digunakan dalam penelitian ini data yang digunakan berjumlah 83 data. Data tersebut telah dikelompokan kedalam penduduk layak (C1) 41 data dengan rata-rata pendapatan Rp 500.000 perbulannnya dan penduduk ini rata-rata masih mengontrak ataupun tinggal bersama orang tua. Penduduk kurang layak (C2) 14 data pendapatan rata-rata perbulan berada di kisaran Rp 750.000 pendapatan ini dinilai sudah bisa mencukupi kebutuhan sehari-hari dan juga ada beberapa penduduk yang sudah memiliki rumah sendiri. Data penduduk tidak layak (C3) 28 data dengan pendapatan berada di atas 1 jt rupiah dan juga memiliki rumah sendiri maka penduduk ini dikelompokan kedalam penduduk tidak layak.
2. *Clustering* dengan algoritma k-means terhadap calon penerima bantuan renovasi rumah menghasilkan *cluster* calon penerima bantuan berdasarkan variabel dan kriteria yang telah di tentukan.

6.2 Saran

Berdasarkan penelitian yang telah dilakukan maka perlu dilakukan percobaan dengan menggunakan metode komputasi lain agar hasil yang didapatkan dapat di bandingkan dengan hasil penggunaan dari metode k-means.

DAFTAR PUSTAKA

- [1] Muhidin, A. , & Baragigiratri, I. (2019). Pemetaan Penduduk Calon Penerima Bantuan Renovasi Rumah Desa Pasangkalan Menggunakan Algoritma Clustering K-Means. *Jurnal Sigma*, 9(3), 75-52.
- [2] Zahrotun, L., & Linarti, U. (2016). Analisis Data Mining Pengelompokan Data Warga Menggunakan Metode Statistik K-Means Untuk Merekomendasikan Pekerjaan Sampingan. *Jurnal Teknologi*, 9(1), 18-23.
- [3] Setiawan, M. Rixco. “Analisis Data Mining Untuk Menentukan Kelompok Prioritas Penerimaan Bantuan Bedah Rumah Di Desa Semen Menggunakan Metode Clustering K-Means.” *Universitas Nusantara Pgri Kediri* (2018).
- [4] Sudibyo, N. A., Iswardani, A., Sari, K., & Suprihatingsih, S. (2020). Penerapan Data Mining Pada Jumlah Pendudk Miskin Di Indonesia. *Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 1(3), 199-207.
- [5] Sibuea, M. L., & Safta, A. (2017). Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering. *Jurteksi (Jurnal Teknologi Dan Sistem Informasi)*, 4(1),85-92.
- [6] Huni, Layak. “Implementasi Kebijakan Tentang Rehabilitas Rumah Tidak Layak Huni Di Desa Perjiwa Kecamatan Tenggarong Seberang Kutai Kartanegara.” (2018).
- [7] Maulida, Linda. “Penerapan Data Mining Dalam Mengelopokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. Dki Jakarta Dengan K-Means.” *Jiska (Jurnalinformatika Sunan Kalijaga)* 2.3 (2018): 167-174.
- [8] Darmi, Y. D., & Setiawan, A. (2016). Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk. *Jurnal Media Infotama*, 12(2).
- [9] Setiawan, Hari Harjanto. “ penanggungan kemiskinan melalui pusat kesejahteraan sosial.” *Sosio informa* 3.3 (2017)
- [10] Zahrotum, Lisna, and Utaminingsih Linarti. “ Analisis Dara Mining Pengelompokan Data Warga Menggunakan Metode Statiskik K-Means

Untuk Merekomendasikan Pekerjaan Sampingan” *Jurnal Teknologi* 9.1 (2016): 18-23.

PAPER NAME

**SKRIPSI_T3117391_FAISAL B. ABAY.doc
x**

AUTHOR

**T3117391_FAISAL B. ABAY faisalabaay
@gmail.com**

WORD COUNT

7501 Words

CHARACTER COUNT

43018 Characters

PAGE COUNT

46 Pages

FILE SIZE

1.1MB

SUBMISSION DATE

May 24, 2022 1:48 AM GMT+8

REPORT DATE

May 24, 2022 1:51 AM GMT+8**● 23% Overall Similarity**

The combined total of all matches, including overlapping sources, for each database.

- 23% Internet database
- 7% Publications database
- Crossref database
- Crossref Posted Content database
- 0% Submitted Works database

● Excluded from Similarity Report

- Bibliographic material
- Small Matches (Less than 25 words)

● 23% Overall Similarity

Top sources found in the following databases:

- 23% Internet database
- 7% Publications database
- Crossref database
- Crossref Posted Content database
- 0% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	ejurnal.dipaneegara.ac.id	Internet	10%
2	ejournal.akprind.ac.id	Internet	2%
3	ejournal.uin-suka.ac.id	Internet	2%
4	eprints.untirta.ac.id	Internet	2%
5	ejurnal.stmik-budidarma.ac.id	Internet	1%
6	repository.nusamandiri.ac.id	Internet	1%
7	123dok.com	Internet	1%
8	ejournal.ap.fisip-unmul.ac.id	Internet	<1%

9	journal.upgris.ac.id Internet	<1%
10	ejournal.pnc.ac.id Internet	<1%
11	Lili Rusdiana, Sam'ani Sam'ani. "Pemodelan K-Means Pada Penentuan ... Crossref	<1%
12	text-id.123dok.com Internet	<1%
13	es.scribd.com Internet	<1%
14	Suliman Suliman. "Implementasi Data Mining Terhadap Prestasi Belaja... Crossref	<1%
15	jurnal.dharmawangsa.ac.id Internet	<1%