

# Pengelompokan Populasi Hewan Ternak Menggunakan Metode Clustering

## (Studi Kasus : Dinas Pertanian dan Ketahanan Pangan Kabupaten Langkat)

Putri Fadila Sari<sup>1</sup>, Akim M H Pardede<sup>2</sup>, Yani Maulita<sup>3</sup>  
<sup>1,2,3</sup>STMIK Kaputama

Email : <sup>1</sup>[putripadila1112@gmail.com](mailto:putripadila1112@gmail.com), <sup>2</sup>[akimmhp@live.com](mailto:akimmhp@live.com), <sup>3</sup>[yanimaulita26@gmail.com](mailto:yanimaulita26@gmail.com)

### Abstrak

Data mining berisi pencarian pola yang diinginkan dalam database besar untuk membantu pengambilan keputusan diwaktu akan datang. Secara sederhana data mining adalah kegiatan mengekstrasi atau menambang pengetahuan dari data yang berjumlah besar dan menghasilkan informasi. Populasi merupakan sekumpulan makhluk hidup yang memiliki karakteristik yang sama dan hidup di wilayah geografis yang sama pada waktu tertentu dan bisa berproduksi antara sesama makhluk hidup lainnya. Karena perkembangan populasi hewan ternak yang semakin meningkat setiap tahunnya dan ada pula yang menurun pada tahun tertentu, maka sangat dibutuhkan data yang ril dan aktual. Dan hasil yang diperoleh berdasarkan analisis ini yaitu jumlah kelompok yang terbentuk sebanyak 3 kelompok. *Group 1 Centroid* (5.45;2.35;2.60) dengan data populasi hewan ternak padagroup wilayah adalah Salapian dengan jenis ternak yaitu Kambing dan jumlah ternak yang dimiliki oleh peternak ialah 51-100 Ekor, untuk *Group 2 Centroid* (2.42;3.64;3.57) dengan data populasi hewan ternak padagroup wilayah adalah Secanggang dengan jenis ternak yaitu Sapi dan jumlah ternak yang dimiliki oleh peternak ialah 101-150 Ekor, dan untuk *Group 3 Centroid* (8.78;3.03;3.06) dengan data populasi hewan ternak padagroup wilayah adalah Kuala dengan jenis ternak yaitu Sapi dan jumlah ternak yang dimiliki oleh peternak ialah 101-150 Ekor.

**Kata Kunci :** Data Mining, Clustering, Populasi, Hewan Ternak.

### Abstract

*Data mining contains the desired pattern search in a large database to help with future decision-making. Simply put data mining is the activity of extracting or mining knowledge from large amounts of data and generating information. Populations are a group of living beings that have the same characteristics and live in the same geographical region at any given time and can produce among other living beings. Due to the increasing development of livestock populations every year and some that decline in a given year, there is a lot of data needed that is ril and actual. And the result obtained based on this analysis is the number of groups formed as many as 3 groups. Group 1 Centroid (5.45;2.35;2.60) with data on the population of farm animals in the group of regions is Salapian with the breed of cattle namely Goats and the number of livestock owned by farmers is 51-100 Tails, for Group 2 Centroid (2.42;3.64;3.57) with data on the population of farm animals in the group of regions is Secanggang with the breed of cattle namely Cows and the number of cattle owned by farmers is 101-150 Tails , and for Group 3 Centroid (8.78;3.03;3.06) with data on the population of farm animals in the group of regions is Kuala with the breed of cattle namely Cows and the number of cattle owned by farmers is 101-150 Tails.*

**Keywords :** Data Mining, Clustering, Population, Farm Animals.

## 1. Pendahuluan

Di Indonesia peternakan merupakan sektor usaha yang dapat menyerap tenaga kerja, selain dapat menyerap tenaga kerja, juga dapat dijadikan sebagai sumber penghasilan utama. Sehingga dapat menjanjikan untuk kelangsungan hidup peternaknya. Salah satunya yaitu dengan melakukan usaha pembibitan dan penggemukan hewan ternak. Perkembangan dunia peternakan beberapa tahun ini sangat cepat, terutama dalam produksi hewan ternak. Peningkatan produksi ini diharapkan akan membawa dampak terhadap peningkatan pendapatan peternak, memperbaiki keadaan lingkungan, meningkatkan kesempatan berusaha, membuka lapangan kerja baru dan memperluas kesempatan kerja yang telah ada.

Populasi merupakan sekumpulan makhluk hidup yang memiliki karakteristik yang sama dan hidup di wilayah geografis yang sama pada waktu tertentu dan bisa berproduksi antara sesama makhluk hidup lainnya. dan terdiri dari manusia, benda, tumbuhan dan hewan ataupun nilai tetes sebagai suatu sumber data yang memiliki karakteristik tertentu dalam suatu penelitian.

Dari data Dinas Pertanian dan Ketahanan Pangan Kabupaten Langkat tahun 2019, populasi hewan ternak mengalami penurunan. Contohnya pada tahun 2018 populasi hewan ternak seperti kambing jumlahnya mencapai 234.589 ekor dibandingkan dari tahun 2017 menurun sebesar 3 %. Dan meningkat kembali pada tahun berikutnya. Namun dari data yang ada setiap tahunnya di Dinas Pertanian dan Ketahanan Pangan Kabupaten Langkat, terdapat kesulitan atau masalah yang terjadi mengenai populasi hewan ternak yaitu sulitnya pihak Dinas Pertanian dan Ketahanan Pangan Kabupaten Langkat dalam mendapatkan suatu data yang ril/nyata dan aktual pada hewan ternak.

Karena perkembangan populasi hewan ternak yang semakin meningkat setiap tahunnya dan ada pula yang menurun pada tahun tertentu, maka sangat dibutuhkan data yang ril dan aktual. Agar mempermudah pihak Dinas Pertanian dan Ketahanan Pangan Kabupaten Langkat dalam mengatasi populasi hewan ternak setiap tahunnya.

## 2. Metode Penelitian

Dalam melakukan penelitian pada skripsi ini, penulis mengikuti tahapan metodologi dalam penelitian ini yaitu sebagai berikut :

### 1. Identifikasi Masalah

Tahap ini merupakan tahap awal yang digunakan untuk mengidentifikasi masalah dengan tujuan untuk mengamati dan mencari permasalahan yang sedang dihadapi pada objek penelitian yaitu Dinas Pertanian Kabupaten Langkat.

### 2. Mengumpulkan Teori Pendukung

Pengumpulan teori-teori yang berhubungan dengan pokok permasalahan seperti teori tentang *data mining*, teori populasi, metode yang digunakan dan aplikasi perancangan dari sistem yang diperlukan. Dalam tahap ini, teori dikumpulkan dari beberapa sumber seperti buku-buku, jurnal, artikel dan referensi lainnya.

### 3. Pengujian Metode

Pada tahap ini peneliti akan menguji metode yang digunakan dalam proses pengelompokan yang tepat, dengan panduan yang sudah ada pada teori-teori pendukung dari buku-buku maupun jurnal terkait dengan pokok permasalahan.

### 4. Perancang Sistem

Pada tahap ini dilakukan perancangan sistem terhadap masalah yang sedang diteliti, bisa berupa tahap untuk merancang alur kerja dari sistem dan juga merancang desain dari tampilan tatap muka (*interface*) dari sistem yang akan dibuat.

### 5. Implementasi Metode

Mengimplementasikan metode yang sudah diuji sebelumnya dengan rancangan sistem yang telah dibuat serta melakukan pengkodean (*coding*) sesuai dengan bahasa pemrograman yang digunakan untuk membuat sistem tersebut.

### 6. Pengujian Sistem

Pada tahap akhir, dilakukan serangkaian pengujian terhadap sistem yang telah dibuat, pengujian-pengujian dilakukan agar dapat menemukan kesalahan-kesalahan (*error*) pada sistem dan melakukan perbaikan-perbaikan yang diperlukan.

## 2.1. Pengertian *Data Mining*

*Data mining* berisi pencarian pola yang diinginkan dalam database besar untuk membantu pengambilan keputusan diwaktu akan datang. Pola-pola ini dikenal oleh perangkat tertentu yang dapat memberikan suatu analisa data yang berguna dan kemudian dapat dipelajari dengan lebih teliti, yang mungkin saja menggunakan perangkat pendukung keputusan yang lain.

Menurut Hermawati (2013,h.3) bahwa: “*Data mining* adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*mechine learning* ) untuk menganalisa dan mengekstrasi pengetahuan (*knowladge*) secara otomatis”.

Menurut Prasetyo (2012, h.2) bahwa: “*Data mining* merupakan proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar”. Data mining juga dapat diartikan sebagai pengekstrakan informasi yang baru diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan.

Istilah *Data Mining* disebut juga *Knowledge discovery* ataupun *pattern recognition*. Istilah tersebut memiliki ketepatan masing-masing..

Istilah *Knowledge discovery*atau penemuan pengetahuna tepat digunakan karena tujuan utama dari data mining memang untuk mendapatkan pengetahuan yang tersembunyi dari bongkahan data. Istilah *pattern recognition*atau pengenalan pola yang tepat untuk digunakan, karena pengetahuan yang hendak digali memang berbentuk pola-pola yang mungkin juga masih perlu digali dari pola bongkahan data yang tengah dihadapi.

Alasan utama mengapa data mining sangat menarik perhatian industri informasi dalam beberapa tahun belakangan ini adalah karena tersedianya data dalam jumlah yang besar dan semakin besarnya kebutuhan untuk mengolah data tersebut menjadi informasi dan pengetahuan yang berguna. *Data Mining* juga merupakan gabungan dari berbagai ilmu, antara lain basis data, kecerdasan buatan, statistik, *information retrieval*, statistika, *machine learning* dan sebagainya.

Secara sederhana *Data Mining* adalah kegiatan mengekstraksi atau menambah pengetahuan dari data yang berjumlah besar dan menghasilkan informasi, informasi inilah yang nantinya sangat berguna untuk pengembangan.

### 2.1.1. Tahapan *Data Mining*

Adapun tahapan *data mining* adalah sebagai berikut :

#### 1. Tahap Pembersihan Data / *Selection*

Yaitu dimana dilakukan proses pemilihan data yang akan digali (*field* yang dibutuhkan dalam proses *data mining*). Pemilihan (seleksi) data dari sekumpulan dataoperasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining disimpan dalam suatu berkas, terpisah dari basisdata operasional.

#### 2. Tahap Preprocessing / *Data Warehouse*

Megeliminasi data yang tidak konsisten. Contohnya menghapus data yang kosong. Sebelum proses data mining dapat dilaksanakan,perlu dilakukan proses *cleaning* pada data yang menjadi fokus KDD.

#### 3. Tahap Trasformasi / *Task-relevant Data*

Proses pengubahan data menjadi bentuk lain, seperti jenis kelamin yang diganti menjadi 1 dan 0. Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses coding dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

#### 4. Tahap *Data Mining*

Data yang telah diolah diawal siap untuk digali, sehingga dapat menghasilkan informasi yang baru. Data mining adalah proses untuk mencari pola atau informasi menarik dalam data terpilih

dengan menggunakan teknik atau metode tertentu. Teknik, metode atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat tergantung pada tujuan dan proses KDD secara keseluruhan.

#### 5. Tahap Evaluasi

Setelah didapatkan hasil dari penambangan dan penggalian data maka hasil dari pengolahan data tersebut harus dievaluasi. Al Fattah (2007, h.13). Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut interpretation. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditentukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

### 2.1.2. Pengelompokan *Data Mining*

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu:

#### 1. Deskripsi

Terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data.

#### 2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun dengan menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi.

#### 3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang.

#### 4. Klasifikasi

Dalam klasifikasi, terdapat terget variabel kategori, sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang dan pendapatan rendah.

#### 5. Pengklusteran

Pengklusteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. *Cluster* adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidak miripan dengan *record-record* dalam cluster lain.

## 2.2. Pengertian Populasi

Populasi merupakan sekumpulan makhluk hidup yang memiliki karakteristik yang sama dan hidup di wilayah geografis yang sama pada waktu tertentu dan bisa berproduksi antara sesama makhluk hidup lainnya. dan terdiri dari manusia, benda, tumbuhan dan hewan ataupun nilai tes sebagai suatu sumber data yang memiliki karakteristik tertentu dalam suatu penelitian.

Menurut Hartono (2011, h. 46), populasi dengan karakteristik tertentu ada yang jumlahnya terhingga dan ada yang tidak terhingga, penelitian hanya dapat dilakukan pada populasi yang jumlahnya terhingga saja.

### 2.2.1. Ciri-Ciri Populasi

Populasi mempunyai ciri-ciri secara biologi dan statistik, yaitu:

#### 1. Lingkup Biologi

Merupakan ciri-ciri yang terdapat pada sekumpulan individu yang membentuk suatu populasi yang diantaranya adalah:

- a. Terdapat suatu organisasi dan struktur organisasi yang bersifat konstan maupun yang fluktuasi sesuai waktu.
- b. Memiliki suatu sejarah kehidupan (ontogenetik) mulai dari lahir, tumbuh, berdiferensiasi, menjadi tua dan mati.

- c. Terpengaruh oleh dampak lingkungan dan dapat merespon perubahan lingkungan.
2. Lingkup Statistik  
Merupakan suatu ciri-ciri kelompok dan merupakan hasil penggabungan berbagai karakteristik dari individu di dalam populasi yang diantaranya:
  - a. Kepadatan atau ukuran besarnya suatu populasi termasuk berbagai parameter utama ini dipengaruhi oleh kelahiran dan kematian.
  - b. Sebaran suatu populasi dipengaruhi oleh persebaran suatu objek tertentu, kondisi iklim dan cuaca, struktur, dan umum objek tersebut.
  - c. Populasi ini juga dipengaruhi oleh komposisi genetik

### 2.2.2. Jenis-Jenis Populasi

Secara umum populasi di klasifikasikan dalam tiga jenis, yaitu:

1. Berdasarkan Jumlah Populasi
  - a. Populasi terbatas, yaitu suatu sumber data yang relatif dapat dihitung jumlahnya karena batasanya jelas secara kuantitatif.
  - b. Populasi tak terbatas, yaitu suatu sumber daya yang tidak dapat dinyatakan dalam bentuk jumlah karena tidak terdapat jumlah batasan.
2. Berdasarkan Sifat Populasi
  - a. Populasi Homogen, yaitu suatu populasi dimana unsurnya mempunyai sifat yang sama sehingga jumlahnya secara kuantitatif tidak perlu dipermasalahkan.
  - b. Populasi Heterogen, yaitu suatu populasi dimana unsurnya terdapat sifat yang bervariasi sehingga harus ditetapkan batasan secara kuantitatif dan kualitatif.
3. Berdasarkan Perbedaan Lain
  - a. Populasi target, yaitu suatu jenis populasi yang telah ditentukan sesuai dengan masalah penelitian.
  - b. Populasi survey, yaitu suatu jenis populasi yang terliput dalam penelitian yang dilaksanakan.

### 2.2.3. Faktor Yang Mempengaruhi Populasi

Perubahan pada jumlah populasi yang terjadi biasanya dipengaruhi oleh beberapa faktor, yaitu sebagai berikut:

1. Natalitas  
Adalah suatu kemampuan suatu populasi untuk meningkatkan jumlahnya dengan berproduksi. Pada umumnya natalitas ini dinyatakan dalam bentuk angka jumlah kelahiran individu baru dibagi dengan waktu. Dan terdapat 2 jenis natalitas, yaitu:
  - a. Natalitas maksimum, yaitu suatu angka kelahiran atau reproduksi maksimum dalam keadaan yang ideal dimana faktor fisiologis diabaikan.
  - b. Natalitas ekologi, yaitu suatu angka kelahiran atau pertambahan populasi karena kondisi lingkungan yang spesifik.
2. Mortalitas  
Adalah suatu tingkat angka kematian individu di dalam suatu populasi dalam kurun waktu tertentu. Mortalitas dibagi dua, yaitu sebagai berikut:
  - a. Mortalitas ekologi, yaitu suatu angka kematian individu dalam populasi karena kondisi lingkungan tertentu.
  - b. Mortalitas minimum, yaitu suatu angka kematian individu dalam lingkungan yang ideal, atau kematian karena usia tua.
3. Densitas  
Adalah suatu tingkat kepadatan populasi yang berhubungan dengan satuan ruang atau area. Kepadatan populasi dibedakan menjadi dua, yaitu:
  - a. Kerapatan kasar, yaitu jumlah kepadatan populasi persatuan ruang total.
  - b. Kerapatan ekologi, yaitu jumlah kepadatan persatuan ruang habitat

### 2.3 Algoritma K-Means

Menurut Prasetyo (2012, h. 178) algoritma K-Means merupakan metode analisis kelompok yang mengarah kepada partisipasi objek pengamatan kedalam kelompok (*cluster*) dimana setiap objek pengamatan dimiliki oleh sebuah kelompok dengan rata-rata (*mean*) terdekat, mirip dimana keduanya mencoba untuk menemukan pusat dari kelompok dalam data sebanyak iterasi perbaikan yang dilakukan oleh kedua algoritma.

K-Means merupakan salah satu metode pengelompokan data (*nonhierarchical*) yang berusaha mempartisi data yang ada kedalam bentuk dua atau lebih kelompok. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diset dalam proses pengelompokan. Pengelompokan data dengan menggunakan metode K-Means ini secara umum dilakukan dengan algoritma dasar sebagai berikut :

1. Tentukan jumlah kelompok (*cluster*).
  2. Alokasikan data ke dalam cluster secara acak (*random*).
  3. Hitung *centroid*/ rata-rata dari data yang ada di masing-masing *cluster*.
- Alokasikan masing-masing data ke *centroid*/ rata-rata terdekat.

#### 2.3.1. Karakteristik K-Means

K-Means dapat diringkas menjadi seperti berikut:

1. K-Means merupakan metode pengelompokan yang sederhana dan dapat digunakan dengan mudah.
2. Pada jenis set data tertentu, K-Means tidak dapat melakukan segmentasi data dengan baik. dimana hasil segmentasinya tidak dapat memberikan pola kelompok yang mewakili karakteristik bentuk alami data.

## 3. Hasil dan Pembahasan

Analisis sistem sangat penting dilakukan dalam penelitian. Hal ini, dilakukan bertujuan untuk menggali data dan mengolahnya sebagai informasi untuk mengambil keputusan. Analisis sistem ini merupakan tahap yang sangat kritis dan sangat penting, karena kesalahan dalam tahap ini akan menyebabkan kesalahan pada tahap berikutnya.

Tugas utama analisis sistem dalam tahap ini adalah menemukan kelemahan-kelemahan dari sistem yang berjalan sehingga dapat diusulkan perbaikannya.

### 3.1. Perhitungan Clustering

Dalam permasalahan metode yang digunakan adalah *clustering*, proses awal yang dilakukan dalam pembentukan *cluster* adalah transformasi data kedalam bentuk numeric dengan kode-kode yang telah ditentukan, tentukan jumlah group (K), hitung *centroid*, hitung jarak objek ke *centroid* dan kemudian groupkan berdasarkan jarak terdekat, jika tidak ada objek yang pindah group maka iterasi selesai.

Tabel 1.Populasi Hewan Ternak Yang Akan Diolah

No	Nama	Wilayah	Jenis Ternak	Jumlah Ternak
1	A	Stabat	Kambing	120 Ekor
2	B	Batang Serangan	Itik	211 Ekor
3	C	Selesai	Domba	52 Ekor
4	D	Wampu	Ayam	50 Ekor
5	E	Kuala	Sapi	125 Ekor
6	F	Salapian	Sapi	100 Ekor
7	G	Bahorok	Domba	45 Ekor
8	H	Sei Bingai	Ayam	138 Ekor
9	I	Sawit Seberang	Kambing	115 Ekor
10	J	Secanggang	Itik	97 Ekor

11	K	Wampu	Itik	154 Ekor
12	L	Selesai	Kambing	36 Ekor
13	M	Salapian	Sapi	109 Ekor
14	N	Sei Bingai	Ayam	161 Ekor
15	O	Stabat	Domba	235 Ekor
16	P	Batang Serangan	Itik	48 Ekor
17	Q	Bahorok	Domba	84 Ekor
18	R	Sawit Seberang	Kambing	146 Ekor
19	S	Secanggang	Ayam	265 Ekor
20	T	Kuala	Sapi	94 Ekor

### 1. Inisialisasi Data

Dari data yang ada maka dapat dilakukan inisialisasi data sesuai dengan kebutuhan variabel sebagai berikut:

#### a. Inisialisasi Kriteria Wilayah

Berikut dibawah ini adalah tabel inisialisasi kriteria wilayah yang dapat dilihat pada tabel dibawah ini:

Tabel 2. Inisialisasi Kriteria Wilayah

Kode	Wilayah
1	Sei Bingai
2	Secanggang
3	Batang Serangan
4	Wampu
5	Salapian
6	Stabat
7	Sawit Seberang
8	Kuala
9	Selesai
10	Bahorok

#### b. Inisialisasi Kriteria Jenis Ternak

Berikut dibawah ini adalah tabel inisialisasi kriteria jenis ternak yang dapat dilihat pada tabel dibawah ini:

Tabel 3. Inisialisasi Kriteria Jenis Ternak

Kode	Jenis Ternak
1	Domba
2	Kambing
3	Sapi
4	Itik
5	Ayam

#### c. Inisialisasi Kriteria Jumlah Ternak

Berikut dibawah ini adalah tabel inisialisasi kriteria jumlah ternak yang dapat dilihat pada tabel dibawah ini:

Tabel 4. Inisialisasi Kriteria Jumlah Ternak

Kode	Jumlah Ternak
1	<=50 Ekor
2	51 – 100 Ekor
3	101 – 150 Ekor
4	151 – 200 Ekor
5	201- 250 Ekor
6	> 250 Ekor

Untuk menentukan group dari satu objek, pertama yang harus dilakukan adalah mengukur jarak *Deeuclidean* antara dua titik atau objek atau X, Y dan Z yang didefinisikan sebagai berikut:

$$\text{Deeuclidean} (X, Y, Z) = \sqrt{\sum(X_1 - X_2)^2 + (Y_1 - Y_2)^2 + (Z_1 - Z_2)^2}$$

Lakukan *cluter* menjadi 3 kelompok ( $K=3$ ) dan tentukan titik pusat *centroid*.

### 3.2. Hasil Pembahasan Perhitungan 2 Group

Setelah data di-*import* kedalam *Matlab* setelah data di-*import* kedalam *Matlab* dengan syntax yang sudah ditentukan berdasarkan jarak *centroid* terkecil benda. Hasil dari perhitungan group di bagi menjadi 3 *group* dengan data wilayah (X), jenis ternak (Y), dan jumlah ternak (Z), yaitu dapat dilihat pada masing-masing *group* yaitu sebagai berikut:

1. *Group 1* sebanyak 305 data
2. *Group 2* sebanyak 295 data

Jadi total data = 600

Dari hasil analisa diatas proses *Replicate* ditentukan sebanyak 5 kali perulangan dimana *cluster* ditentukan sebanyak 2 (X, Y, dan Z) maka total iterasi sebanyak 10 kali, hal ini menunjukkan bahwa proses iterasi berhenti jika total dan jarak dengan iterasi sebelumnya sampai pada *group* yang tidak berubah lagi. Hasil iterasi yang diperoleh dari perhitungan jarak objek ke *centroid* menggunakan program *Matlab* adalah sebagai berikut:

4 iterations, total sum of distances = 3556.63  
 7 iterations, total sum of distances = 3556.63  
 5 iterations, total sum of distances = 3556.63  
 6 iterations, total sum of distances = 3556.63  
 4 iterations, total sum of distances = 3556.63

Adapun penentuan hasil jumlah *centroid* untuk setiap *group* adalah sebagai berikut:

*Centroid 1* = total *group 1*/banyak *group 1*  
 C1 = 917 / 305 = 3.00  
 C2 = 989 / 305 = 3.24  
 C3 = 1013 / 305 = 3.32

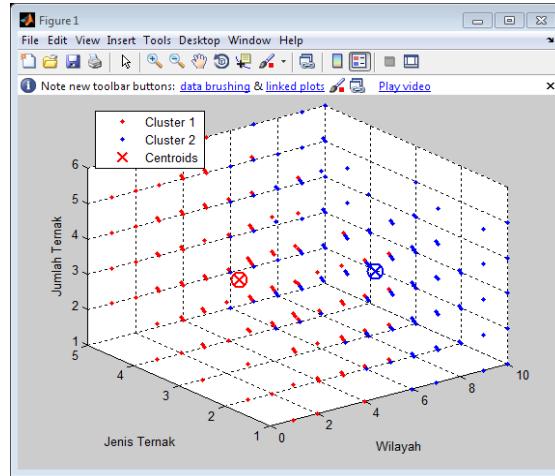
*Centroid 2* = total *group 2*/banyak *group 2*  
 C1 = 2367 / 295 = 8.02  
 C2 = 847 / 295 = 2.87  
 C3 = 856 / 295 = 2.90

Dari hasil perhitungan *centroid* diatas dapat dilihat pada tabel berikut:

Tabel 5. Perhitungan *Centroid*

Wilayah (X)	Jenis Ternak (Y)	Jumlah Ternak (Z)	Keterangan
3.00	3.24	3.32	<i>Centroid 1</i>
8.02	2.87	2.90	<i>Centroid 2</i>

Hasil akhir penentuan kelompok populasi hewan ternak mana yang termasuk dalam *group 1*, dan *group 2*, dapat dilihat pada grafik dibawah ini:



Gambar 1. Grafik Pengelompokan

Pusatnya =

$$\begin{array}{lll} 3(3.00) & 3(3.24) & 3(3.32) \\ 8(8.02) & 3(2.87) & 3(2.90) \end{array}$$

Keterangan :

$$1. 3(3.00) \quad 3(3.24) \quad 3(3.32)$$

Dapat diketahui bahwasannya pada *cluster* 1. Kelompok data populasi hewan ternak padagroup wilayah adalah Batang Serangan dengan jenis ternak yaitu Sapi dan jumlah ternak yang dimiliki oleh peternak ialah 101-150 Ekor.

$$2. 8(8.02) \quad 3(2.87) \quad 3(2.90)$$

Dapat diketahui bahwasannya pada *cluster* 1. Kelompok data populasi hewan ternak padagroup wilayah adalah Kuala dengan jenis ternak yaitu Sapi dan jumlah ternak yang dimiliki oleh peternak ialah 101-150 Ekor.

### 3.3. Hasil Pembahasan Perhitungan 3 Group

Setelah data di-import kedalam *Matlab* dengan *syntax* yang sudah ditentukan berdasarkan jarak *centroid* terkecil benda. Hasil dari perhitungan group di bagi menjadi 3 *group* dengan data wilayah (X), jenis ternak(Y), dan jumlah ternak (Z), yaitu dapat dilihat pada tabel berikut:

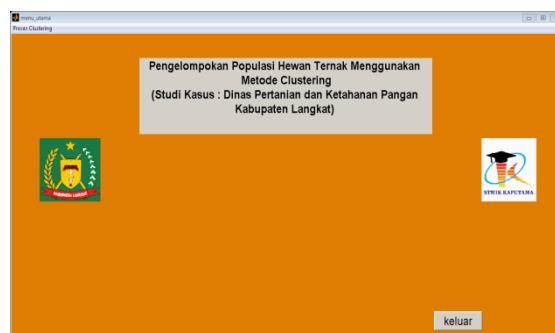
Dari hasil tabel diatas dapat dijumlahkan masing-masing *group* yaitu sebagai berikut :

1. *Group* 1 sebanyak 178 data
2. *Group* 2 sebanyak 219 data
3. *Group* 3 sebanyak 203 data

Jadi total data = 600

### 3.4. Interface Sistem

Berikut gambar *interface Home / Menu Utama*:



Gambar 2. Halaman Utama

2. Menu Proses *Clustering*, disini akan terlihat keseluruhan proses *data mining* sampai pada pemunculan grafik dan keterangan *centroid* sebagai hasil dari perhitungan dengan menggunakan metode *clustering* menggunakan algoritma *k-means*. Untuk lebih jelasnya dapat dilihat pada gambar berikut:



Gambar 3. Tampilan Proses Data

#### 4. Kesimpulan

Dari hasil analisa berdasarkan *clustering* kelompok populasi hewan ternak yang didapat, maka akan diambil suatu kasimpulan yaitu sebagai berikut:

1. Dari pengujian yang dilakukan menggunakan metode *clustering* dengan algoritma K-Means ini, dapat diketahui bahwa kelompok wilayah, jenis ternak dan jumlah ternak mana saja yang memiliki kelompok paling tinggi dan paling sering meningkat tiap tahunnya sehingga dapat diketahui hubungan terdekat antara kelompok populasi hewan ternak tersebut.
2. Diketahui hasil 600 data yang telah dijelaskan sebelumnya pada bab IV bahwa untuk *Centroid 1* (5.45 2.35 2.60) Kelompok data populasi hewan ternak *padagroup* wilayah adalah Salapian dengan jenis ternak yaitu Kambing dan jumlah ternak yang dimiliki oleh peternak ialah 51-100 Ekor, serta *Centroid 2* (2.42 3.64 3.57) Kelompok data populasi hewan ternak *padagroup* wilayah adalah Secanggang dengan jenis ternak yaitu Sapi dan jumlah ternak yang dimiliki oleh peternak ialah 101-150 Ekor, dan *Centroid 3* (8.78 3.03 3.06) Kelompok data populasi hewan ternak *padagroup* wilayah adalah Kuala dengan jenis ternak yaitu Sapi dan jumlah ternak yang dimiliki oleh peternak ialah 101-150 Ekor.

#### Daftar Pustaka

- [1] Alfatta, H , 2007, Analisis dan Perancang Sistem Informasi, Penerbit CV.Andi, Yogyakarta.
- [2] Darmi, Y. 2016. Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk.
- [3] Fina, N. 2013. Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru.
- [4] Hermawati, A, F, 2013. Data Mining, Penerbit CV. Andi, Yogyakarta.
- [5] Kusrini dan Luthfi, 2009, Algoritma Data Mining, Penerbit CV. Andi, Yogyakarta.
- [6] Nugroho, A, 2013, T. Konsep dan Pemrograman GUI, Penerbit CV. Andi, Yogyakarta.
- [7] Prasetyo, E, 2012, Data Mining – Konsep dan Aplikasi Menggunakan Matlab
- [8] Santosa, B. 2007. Data Mining Terapan Dengan Matlab, Yogyakarta.
- [9] Sitorus, L. 2015, Algoritma dan Pemrograman, Yogyakarta.
- [10] Sri, S. 2016. Analisa Dan Penerapan Data Mining Untuk Menentukan Kubikasi Air Terjual Berdasarkan Pengelompokan Pelanggan Menggunakan Algoritma K-MeansClustering.