

# PENERAPAN ALGORITMA K-MEANS UNTUK CLUSTERING TEMPAT MAKAN DI BATUBARA

*Application of the K-Means algorithm for clustering places to eat in Batubara*

Riski Askia Kurniawan<sup>1</sup>, Muhammad Siddik Hasibuan<sup>2</sup>, Piramida<sup>3</sup>, Rizky Syahrul Ramadhan<sup>4</sup>

<sup>1,3,4</sup>Ilmu Komputer, Universitas Islam Negeri Sumatera Utara

<sup>2</sup>Dosen Ilmu Komputer, Universitas Islam Negeri Sumatera Utara

E-mail: <sup>1</sup>[riskiaskiak@gmail.com](mailto:riskiaskiak@gmail.com), <sup>2</sup>[muhammadsiddik@uinsu.ac.id](mailto:muhammadsiddik@uinsu.ac.id), <sup>3</sup>[piramidapira@gmail.com](mailto:piramidapira@gmail.com),

<sup>4</sup>[riskisyahrul99@gmail.com](mailto:riskisyahrul99@gmail.com)

## Abstrak

Batubara adalah kabupaten yang menyimpan potensi wisata yang banyak. Tempat wisata yang banyak memungkinkan untuk menarik wisatawan dalam berkunjung, selain mempertimbangkan tempat wisata tentu saja wisatawan melirik tempat kuliner yang disajikan ketika berkunjung. Rumah makan, resto dan cafe yang terdapat di Kabupaten Batubara merupakan diantaranya. Kami melakukan penelitian dengan mengumpulkan jumlah tempat makan (Rumah makan, Resto, dan Cafe) setiap kecamatan di Kabupaten Batubara pada tahun 2016, 2017, dan 2021 dan menerapkan algoritma K-Means Clustering yaitu mengelompokkan jumlah tempat makan berdasarkan tahun. Tujuannya agar wisatawan mudah mencari tempat kuliner dan wisatawan juga mengetahui tempat makan berdasarkan kecamatan yang sesuai dengan tempat wisata di Kabupaten Batubara. Kegiatan mengumpulkan data tempat makan (Rumah makan, Resto dan Cafe) ini didasarkan kegiatan Kerja Praktik yang diselenggarakan selama 1 bulan dengan menerapkan apa yang dipelajari selama perkuliahan. Nantinya informasi yang diperoleh dalam penelitian ini digunakan untuk melihat perkembangan ekonomi dan tempat kuliner di Kabupaten Batubara

Kata kunci : Data Mining, Algorithm K-Means Clustering, Tempat Makan

## Abstract

*Batubara is a district that has a lot of tourism potential. Many tourist attractions make it possible to attract tourists to visit, besides considering tourist attractions, of course tourists glance at the culinary places served when visiting. Food stalls, restaurants and cafes in Batubara Regency are among them. We conducted research by collecting the number of places to eat (food stalls, restaurants, and cafes) in each sub-district in Batubara Regency in 2016, 2017, and 2021 and applying the K-Means Clustering algorithm, namely grouping the number of places to eat by year. The goal is for tourists to easily find culinary places and tourists also know where to eat based on the sub-districts that are in accordance with tourist attractions in Batubara Regency. The activity of collecting data on places to eat (food stalls, restaurants and cafes) is based on practical work activities held for 1 month by applying what was learned during lectures. Later the information obtained in this research is used to see the economic development and culinary places in Batubara Regency*

Keywords : Data Mining, Algorithm K-Means Clustering, Tempat Makan

## 1. PENDAHULUAN

Kabupaten Batu Bara adalah salah satu kabupaten yang berada di provinsi Sumatra Utara. Kabupaten ini merupakan hasil pemekaran dari Kabupaten Asahan dan beribu kota di Kecamatan Limapuluh. Kabupaten Batu Bara salah kabupaten dengan penuh tempat wisatawan dikarenakan berdekatan dengan laut. Sehingga tidak heran pertumbuhan tempat makan di Kabupaten Batubara

perlu diperhatikan supaya para pengunjung dapat memilih tempat makan terdekat dan dapat memantau perkembangan ekonomi tempat makan di Kab. Batu Bara.

Salah satu tugas Dinas Pariwisata di Kabupaten ialah untuk melihat perkembangan tempat makan seperti rumah makan, cafe, restoran. Tepat tanggal 27 september – 27 october 2021, Kami mahasiswa/mahasiswi melakukan salah satu kegiatan yaitu kerja praktik di salah satu instansi yaitu Dinas Pariwisata. kami ditugaskan untuk mendata tingkat perkembangan tempat makan di Kab. Batu Bara di masing-masing Kecamatan. Kami juga melakukan pengelompokan di masing masing kecamatan untuk melihat di kecamatan mana tempat makan paling banyak dan sedikit dengan data yang sudah didapatkan sebelum nya di tahun 2016, 2017 oleh instansi tersebut. Pengelompokan yang kami lakukan akan menggunakan Algoritma K-Means Clustering.

Data mining merupakan proses dengan kecerdasan buatan, teknik statistik, matematika, dan machine learning untuk mengidentifikasi dan mengekstraksi dan suatu informasi yang bermanfaat dan pengetahuan terkait dari berbagai database yang besar. Clustering termasuk salah satu metode dari data mining dan clustering telah menjadi instrumen yang valid untuk memecahkan masalah kompleks ilmu komputer dan statistik. Clustering adalah proses pengelompokan titik-titik data kedalam dua kelompok atau lebih sehingga titik-titik data yang termasuk didalam kelompok yang sama lebih mirip satu sama lain dari pada didalam kelompok yang berbeda, hanya berdasarkan informasi yang tersedia dengan poin data[7].

K-Means merupakan salah satu metode clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster. Metode ini mempartisi data ke dalam cluster sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karateristik yang berbeda di kelompokkan ke dalam cluster yang lain. K-Means Clustering adalah suatu metode penganalisaan data atau metode Data Mining yang melakukan proses pemodelan tanpa supervisi (unsupervised) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Terdapat dua jenis data clustering yang sering dipergunakan dalam proses pengelompokan data yaitu Hierarchical dan Non-Hierarchical, dan K-Means merupakan salah satu metode data clustering non-hierarchical atau Partitional Clustering. Metode K-Means Clustering berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain[1].

Pada dasarnya clustering merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteritktik (similarity) antara satu data dengan data yang lain. Clustering merupakan salah satu metode data mining yang bersifat tanpa arahan (unsupervised), maksudnya metode ini diterapkan tanpa adanya latihan (training) dan tanpa ada guru (teacher) serta tidak memerlukan target output. Dalam data mining ada dua jenis metode clustering yang digunakan dalam pengelompokan data, yaitu hierarchical clustering dan non-hierarchical clustering[2][3].

Hierarchical clustering adalah suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan kedua. Demikian seterusnya sehingga cluster akan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah cluster. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut[2][3][4].

Berbeda dengan metode hierarchical clustering, metode non-hierarchical clustering justru dimulai dengan menentukan terlebih dahulu jumlah cluster yang diinginkan (dua cluster, tiga cluster, atau lain sebagainya). Setelah jumlah cluster diketahui, baru proses cluster dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan K-Means Clustering[2][4].

langkah-langkah melakukan clustering dengan metode K-Means adalah sebagai berikut:

- a. Pilih jumlah cluster k.
- b. Inisialisasi k pusat cluster ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat cluster diberiduberi

nilai awal dengan angka-angka random,

- c. Alokasikan semua data/ objek ke cluster terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke cluster tertentu ditentukan jarak antara data dengan pusat cluster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat cluster. Jarak paling antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster mana. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

dimana:

$D(i, j)$  = Jarak data ke i ke pusat cluster j

$X_{ki}$  = Data ke i pada atribut data ke k

$X_{kj}$  = Titik pusat ke j pada atribut ke k

- d. Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata-rata dari semua data/ objek dalam cluster tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.
- e. Tugaskan lagi setiap objek memakai pusat cluster yang baru. Jika pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat cluster tidak berubah lagi[2].

## 2. METODE PENELITIAN

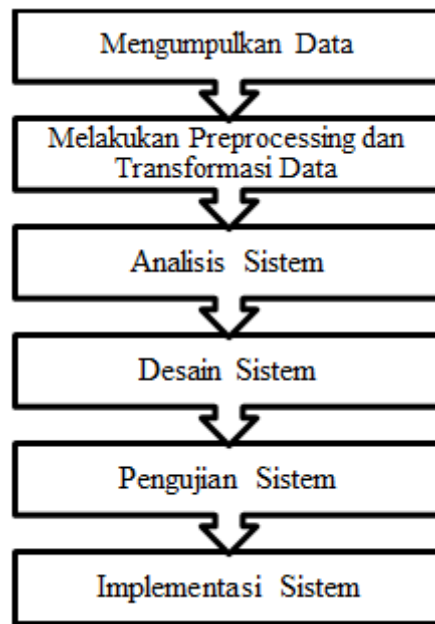
Pengumpulan data dilakukan secara kolektif berdasarkan Kecamatan se Kabupaten Batu Bara dan kelompokan pada tahun 2016, 2017, dan 2021 melihat perkembangan Tempat makan (Rumah makan, Cafe, Resto)

### 2.1 Metodologi

Pada Metodologi penelitian ini dilakukan secara sistematis menjelaskan proses bagaimana kerangka kerja penelitian yang dilaksanakan. Beberapa tahapan yang digambarkan dalam bentuk kerangka gambar kemudian dapat di pahami mulai dari proses analisa kebutuhan sampai dengan hasil dari penelitian ini. Tahapan-tahapan dalam melakukan penelitian, yakni menganalisa kebutuhan data, proses pengumpulan data, analisa data menggunakan algoritma K-Means, kemudian dilakukan pengolahan data dengan RapidMiner, dan hasil dari penelitian ini.

### 2.2 Kerangka Kerja Penelitian

Tahapan pada poin ini yakni dimana kerangka kerja berfungsi untuk menggambarkan proses yang sistematis dalam penelitian ini, dari tahap menganalisa kebutuhan data hingga menyimpulkan hasil[5].



Gambar 1. Kerangka Kerja Penelitian

### 3. HASIL DAN PEMBAHASAN

Data awal adalah data yang akan diolah berdasarkan data yang kami peroleh dari instansi terkait dan melakukan pendataan mandiri se Kabupaten Batubara. Data ini menampilkan jumlah rumah makan, restoran, dan cafe pada setiap kecamatan di Kabupaten Batubara.

| No | Kecamatan          | 2016 | 2017 | 2021 |
|----|--------------------|------|------|------|
| 1  | Sei Suka           | 17   | 17   | 11   |
| 2  | Lau Tador          | 0    | 0    | 2    |
| 3  | Air Putih          | 12   | 13   | 6    |
| 4  | Tanjung Tiram      | 15   | 15   | 7    |
| 5  | Talawi             | 7    | 7    | 6    |
| 6  | Lima Puluh Pesisir | 0    | 0    | 2    |
| 7  | Medang Deras       | 10   | 10   | 7    |
| 8  | Datuk Lima Puluh   | 0    | 0    | 2    |
| 9  | Datuk Tanah Datar  | 0    | 0    | 1    |
| 10 | Sei Balai          | 11   | 12   | 2    |
| 11 | Lima Puluh         | 13   | 14   | 6    |
| 12 | Nibung Hangus      | 0    | 0    | 0    |
|    | Jumlah             | 85   | 88   | 52   |

Tabel 1. Data Awal

Setelah semua data telah disajikan dalam bentuk tabel, maka data tersebut dapat diolah menggunakan Algoritma K-Means Clustering, algoritma ini sendiri terdiri dari beberapa tahapan.

1. Tahapan pertama dari proses ini adalah penentuan jumlah cluster. Cluster sendiri yaitu menentukan berapa banyak kelompok data yang diinginkan. Pada penelitian ini telah ditentukan sebanyak 2 cluster.
2. Penentuan titik pusat (centroid) awal dari cluster. Pada penelitian ini, kami menentukan titik pusat awal secara random. Dapat dilihat pada table di bawah ini.

| Titik Pusat awal | Kecamatan            | 2016 | 2017 | 2021 |
|------------------|----------------------|------|------|------|
| cluster 1        | Tanjung Tiram (4)    | 15   | 15   | 7    |
| cluster 2        | Datuk Lima Puluh (8) | 0    | 0    | 2    |

Tabel 2. Titik Pusat(centroid) dari cluster

3. Mengalokasikan setiap data ke dalam suatu cluster, sehingga data akan dimasukkan ke dalam suatu cluster dengan jarak paling dekat dengan titik pusat cluster. Untuk mengetahui nya, diperlukan perhitungan jarak setiap data dengan titik pusat setiap cluster. Lebih jelasnya, akan dihitung jarak dari data setiap kecamatan pada tahun 2016, 2017, dan 2021 dengan pusat cluster pertama dan kedua.

$$D(\text{data1}, \text{cluster1}) = \sqrt{(data\ 1\ dimensi\ 1 - centroid\ 1\ dimensi\ 1)^2 + (data\ 1\ dimensi\ 2 - centroid\ 1\ dimensi\ 2)^2 + (data\ 1\ dimensi\ 3 - centroid\ 1\ dimensi\ 3)^2}$$

$$D(1,1) = \sqrt{(17 - 15)^2 + (17 - 15)^2 + (11 - 7)^2} = 4.8$$

Dari hasil perhitungan di atas, bahwa jarak data kecamatan pertama dengan pusat cluster pertama adalah 4.8

$$D(\text{data1}, \text{cluster2}) = \sqrt{(data\ 1\ dimensi\ 1 - centroid\ 2\ dimensi\ 1)^2 + (data\ 1\ dimensi\ 2 - centroid\ 2\ dimensi\ 2)^2 + (data\ 1\ dimensi\ 3 - centroid\ 2\ dimensi\ 3)^2}$$

$$D(1,2) = \sqrt{(17 - 0)^2 + (17 - 0)^2 + (11 - 2)^2} = 25.67$$

$$D(\text{data1}, \text{cluster1}) = \sqrt{(data\ 1\ dimensi\ 1 - centroid\ 1\ dimensi\ 1)^2 + (data\ 1\ dimensi\ 2 - centroid\ 1\ dimensi\ 2)^2 + (data\ 1\ dimensi\ 3 - centroid\ 1\ dimensi\ 3)^2}$$

$$D(1,1) = \sqrt{(17 - 15)^2 + (17 - 15)^2 + (11 - 7)^2} = 4.8$$

Dari hasil perhitungan di atas, bahwa jarak data kecamatan pertama dengan pusat cluster pertama adalah 4.8

$$D(\text{data1}, \text{cluster2}) = \sqrt{(data\ 1\ dimensi\ 1 - centroid\ 2\ dimensi\ 1)^2 + (data\ 1\ dimensi\ 2 - centroid\ 2\ dimensi\ 2)^2 + (data\ 1\ dimensi\ 3 - centroid\ 2\ dimensi\ 3)^2}$$

$$D(1,2) = \sqrt{(17 - 0)^2 + (17 - 0)^2 + (11 - 2)^2} = 25.67$$

Dari hasil perhitungan di atas, bahwa jarak data kecamatan pertama dengan pusat cluster kedua adalah 25.67

Berdasarkan hasil perhitungan di atas, bahwa jarak data kecamatan yang pertama lebih dekat dengan cluster 1 dengan nilai 4.8 sedangkan pada cluster kedua nilainya adalah 25.67. Kemudian hasil perhitungan data selanjutnya, dapat dilihat pada table di bawah ini

| No | Kecamatan          | 2016 | 2017 | 2021 | C1    | C2    | Jarak terdekat ke cluster |
|----|--------------------|------|------|------|-------|-------|---------------------------|
| 1  | Sei Suka           | 17   | 17   | 11   | 4.9   | 25.67 | 1                         |
| 2  | Lau Tador          | 0    | 0    | 2    | 21.8  | 0     | 2                         |
| 3  | Air Putih          | 12   | 13   | 6    | 3.74  | 18.13 | 1                         |
| 4  | Tanjung Tiram      | 15   | 15   | 7    | 0     | 21.79 | 1                         |
| 5  | Talawi             | 7    | 7    | 6    | 11.36 | 10.67 | 2                         |
| 6  | Lima Puluh Pesisir | 0    | 0    | 2    | 21.8  | 0     | 2                         |
| 7  | Medang Deras       | 10   | 10   | 7    | 7.07  | 15    | 1                         |
| 8  | Datuk Lima Puluh   | 0    | 0    | 2    | 21.8  | 0     | 2                         |
| 9  | Datuk Tanah Datar  | 0    | 0    | 1    | 22.04 | 1     | 2                         |
| 10 | Sei Balai          | 11   | 12   | 2    | 7.07  | 16.28 | 1                         |
| 11 | Lima Puluh         | 13   | 14   | 6    | 2.5   | 19.51 | 1                         |
| 12 | Nibung Hangus      | 0    | 0    | 0    | 2.34  | 2     | 2                         |

Table 3. Iterasi Pertama

4. Penentuan nilai titik pusat (centroid) baru. Caranya dengan menghitung rata-rata setiap data pada dimensi yang sudah dikelompokkan pada setiap cluster terdekat. Untuk cluster pertama memiliki 6 data. Berikut ini adalah perhitungannya:

$$\text{Cluster 1, dimensi 1} = \frac{17+12+15+10+11+13}{6} = 13$$

$$\text{Cluster 1, dimensi 2} = \frac{17+13+15+10+12+14}{6} = 13.5$$

$$\text{Cluster 1, dimensi 3} = \frac{11+6+7+7+2+6}{6} = 6.5$$

Cluster 2 memiliki 6 data, berikut ini adalah perhitungannya:

$$\text{Cluster 2, dimensi 1} = \frac{0+7+0+0+0+0}{6} = 1.16$$

$$\text{Cluster 2, dimensi 2} = \frac{0+7+0+0+0+0}{6} = 1.16$$

$$\text{Cluster 2, dimensi 3} = \frac{2+6+2+2+1+0}{6} = 2.16$$

Pada tahap ini didapati nilai centroid yang baru, yaitu :

- Centroid 1 = 13, 13.5, 6.5
  - Centroid 2 = 1.16, 1.16, 2.16
5. Perhitungan iterasi kedua yaitu melakukan perhitungan jarak setiap data yang sama dengan proses pengolahan data yang ketiga, namun menggunakan nilai centroid yang baru. Berikut table hasil perhitungan Iterasi kedua

| No | Kecamatan          | 2016 | 2017 | 2021 | C1    | C2    | Jarak terdekat ke cluster |
|----|--------------------|------|------|------|-------|-------|---------------------------|
| 1  | Sei Suka           | 17   | 17   | 11   | 6.96  | 23.03 | 1                         |
| 2  | Lau Tador          | 0    | 0    | 2    | 19.27 | 1.65  | 2                         |
| 3  | Air Putih          | 12   | 13   | 6    | 1.22  | 16.50 | 1                         |
| 4  | Tanjung Tiram      | 15   | 15   | 7    | 2.55  | 20.16 | 1                         |
| 5  | Talawi             | 7    | 7    | 6    | 8.86  | 9.11  | 1                         |
| 6  | Lima Puluh Pesisir | 0    | 0    | 2    | 19.27 | 1.65  | 2                         |
| 7  | Medang Deras       | 10   | 10   | 7    | 4.63  | 13.40 | 1                         |
| 8  | Datuk Lima Puluh   | 0    | 0    | 2    | 19.27 | 1.65  | 2                         |
| 9  | Datuk Tanah Datar  | 0    | 0    | 1    | 19.53 | 2.01  | 2                         |
| 10 | Sei Balai          | 11   | 12   | 2    | 5.15  | 14.64 | 1                         |
| 11 | Lima Puluh         | 13   | 14   | 6    | 0.71  | 17.88 | 1                         |
| 12 | Nibung Hangus      | 0    | 0    | 0    | 19.83 | 2.71  | 2                         |

Table 4. Iterasi Kedua

Kemudian dilakukan perhitungan untuk menentukan centroid yang baru dan didapati nilainya :

- a. Centroid 1 = 12.14, 12.57, 6.42
  - b. Centroid 2 = 0, 0, 1.4
6. Lakukan perhitungan iterasi ketiga. Setelah diketahui nilai centroid yang baru, lakukan lagi langkah ketiga. Berikut tabel hasil perhitungan dari iterasi ketiga.

| No | Kecamatan          | 2016 | 2017 | 2021 | C1    | C2    | Jarak terdekat ke cluster |
|----|--------------------|------|------|------|-------|-------|---------------------------|
| 1  | Sei Suka           | 17   | 17   | 11   | 8.01  | 25.89 | 1                         |
| 2  | Laut Tador         | 0    | 0    | 2    | 18.02 | 0.6   | 2                         |
| 3  | Air Putih          | 12   | 13   | 6    | 0.62  | 18.28 | 1                         |
| 4  | Tanjung Tiram      | 15   | 15   | 7    | 3.80  | 21.94 | 1                         |
| 5  | Talawi             | 7    | 7    | 6    | 7.59  | 10.92 | 1                         |
| 6  | Lima Puluh Pesisir | 0    | 0    | 2    | 18.02 | 0.6   | 2                         |
| 7  | Medang Deras       | 10   | 10   | 7    | 3.39  | 15.21 | 1                         |
| 8  | Datuk Lima Puluh   | 0    | 0    | 2    | 18.02 | 0.6   | 2                         |
| 9  | Datuk Tanah Datar  | 0    | 0    | 1    | 18.30 | 0.4   | 2                         |
| 10 | Sei Balai          | 11   | 12   | 2    | 4.60  | 16.29 | 1                         |
| 11 | Lima Puluh         | 13   | 14   | 6    | 1.87  | 19.65 | 1                         |
| 12 | Nibung Hangus      | 0    | 0    | 0    | 18.62 | 1.4   | 2                         |

Table 5. Iterasi Ketiga

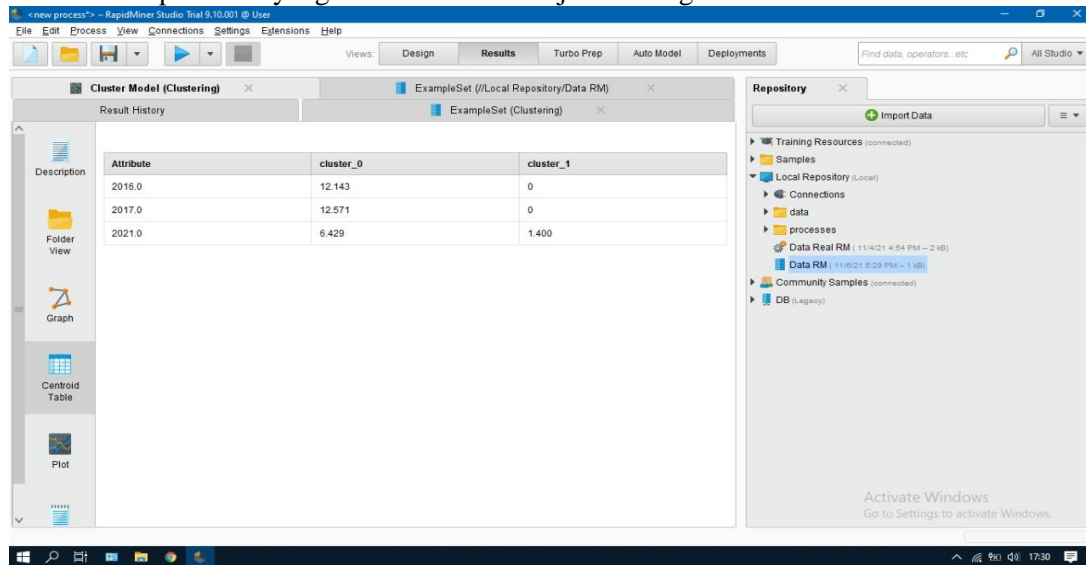
Pada tahap perhitungan nilai centroid yang baru didapati bahwa nilai nya sama dengan nilai centroid yang sebelumnya, yaitu

- a. Centroid 1 = 12.14, 12.57, 6.42



b. Centroid 2 = 0, 0, 1.4

Disini kami juga melakukan perhitungan dengan menggunakan Rapid Miner dan mendapati hasil yang sama. Untuk lebih jelas lihat gambar di bawah ini



The screenshot shows the Rapid Miner Studio interface. The main window displays a 'Cluster Model (Clustering)' result. Below the model name, there is a table with the following data:

| Attribute | cluster_0 | cluster_1 |
|-----------|-----------|-----------|
| 2016.0    | 12.143    | 0         |
| 2017.0    | 12.571    | 0         |
| 2021.0    | 6.429     | 1.400     |

The interface also shows a sidebar with various tools like 'Description', 'Folder View', 'Graph', 'Centroid Table', and 'Plot'. On the right, there is a 'Repository' pane showing connected data sources like 'Training Resources', 'Samples', 'Local Repository', 'Connections', 'data', 'processes', 'Data Real RM', 'Data RM', 'Community Samples', and 'DB (Legacy)'.

Gambar 2. Data Rapid Miner

Dalam penelitian ini, terjadi iterasi sebanyak tiga kali hingga tidak ada perubahan nilai centroid dan tidak ada perubahan data dari satu cluster ke cluster yang lain. Jika tidak ada perubahan, maka proses iterasi dianggap tidak perlu dilakukan lagi.

#### 4. KESIMPULAN DAN SARAN

Berdasar kan hasil yang diperoleh dari penelitian ini dengan metode K-Means clustering didapatkan Centroid 1(12.14, 12.57, 6.42) dan Centroid 2 (0, 0, 1.4), Kelompok Cluster 1 yaitu Sei Suka, Air Putih, Tanjung Tiram, Talawi, Medang Deras, Sei Balai, Lima Puluh merupakan kelompok tempat makan yang banyak dan Cluster 2 yaitu Laut Tador, Lima Puluh Pesisir, Datuk Lima Puluh, Datuk Tanah Datar, Nibung Hangus merupakan kelompok tempat makan yang sedikit maka penelitian sudah didapatkan kelompok yang banyak dan sedikit

Pada penelitian selanjutnya diharuskan mendapatkan data yang lebih banyak dari tahun ke tahun untuk clustering bisa memberikan pembobotan kriteria agar data yang dihasilkan lebih akurat. Selain itu, juga perlu meningkatkan akurasi clustering pada data.

#### DAFTAR PUSTAKA

- [1] H Sy. 2019. *Data Mining Menggunakan Algoritma K-Means Pengelompokan Penyebaran Diare Di Kota Makassar*. Vol. 8, No. 1. 73-83 : <https://ejurnal.dipaneegara.ac.id/index.php/sisiti/article/download/73%20-82/74>
- [2] Ong, Johan Oscar. 2013. *Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University* . ISSN 1412-6869. Vol. 12, No. 1. 13-14 : <https://journals.ums.ac.id/index.php/jiti/article/viewFile/651/392>.
- [3] Anggraini, Amelia Ayu.2017. *Pengelompokan Kecamatan Menggunakan Metode K-Means Cluster*. 3-4 : <http://repository.unmuhjember.ac.id/591/1/ARTIKEL.pdf>
- [4] Alfina, T. Santosa, B. & Barakbah, A. R. 2012. *Analisa Perbandingan Metode Hierarchical Clustering, K-Means dan Gabungan Keduanya dalam Membentuk Cluster Data (Studi Kasus: Problem Kerja Praktek Jurusan Teknik Industri)* (1) : 1-5.
- [5] Baginda Harahap. 2019. *Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)*.ISSN: 2620-6048. Vol 2, No 2 : <https://ptki.ac.id/jurnal/index.php/readystar/article/download/82/pdf>



- [6] M. N. M. Ediyanto. 2013. *Pengklasifikan Karakteristik Dengan Metode K-Means Cluster Analysis*, BIMASTER, Vol. 2, no. 02.
- [7] Gustientiedina, M.Hasmil Adiya, Yenny Deslita. 2019. *Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru*. ISSN : 2476-8812, 1 : <http://teknosi.fti.unand.ac.id/>