

## Implementasi K-Means Clustering dengan Menggunakan Data Transaksi Penjualan untuk Penentuan Reward pada Agen Aqua dan Gas LPG FF Tirta

Refika Ayuna Sari<sup>1</sup>, Henki Bayu Seta<sup>2</sup>, I Gede Susrama Mas Diyasa<sup>3</sup>

<sup>1,2</sup>Program Studi Informatika

<sup>3</sup>Program Studi Data Science

<sup>1,2</sup>Universitas Pembangunan Nasional Veteran Jakarta

<sup>3</sup>Universitas Pembangunan Nasional “Veteran” Jawa Timur

<sup>1,2</sup>Jalan RS. Fatmawati Raya, Pd. Labu, Kec. Cilandak, Kota Depok, Jawa Barat 12450

<sup>3</sup>Jl.Raya Rungkut Madya, Gunung Anyar, Surabaya

refikaayuna15@gmail.com<sup>1</sup>, henkiseta@upnvj.ac.id<sup>2</sup>, igsusrama.if@upnjatim.ac.id<sup>3</sup>

**Abstrak.** Target pelanggan bagi Agen Aqua dan Gas LPG sangatlah penting karena persaingan antara perusahaan sejenis mungkin terjadi dan mengakibatkan hilangnya kepuasan pelanggan. Penelitian ini memiliki tujuan yaitu membangun sistem yang dapat mengelompokkan dan mengetahui tingkat target pelanggan berdasarkan transaksi pembelian agar pihak agen dapat mempertahankan pelanggan yang potensial tersebut dengan cara memberikan hadiah berupa promo spesial (reward). Hasil penelitian ini ada pengelompokan pelanggan yang mendapatkan reward dan tidak dapat dilihat perbulannya, yaitu pada bulan Januari ada 75 pelanggan mendapatkan reward dan 100 tidak mendapatkan reward, Februari ada 70 pelanggan mendapatkan reward dan 106 tidak mendapatkan reward, Maret ada 80 pelanggan mendapatkan reward dan 96 tidak mendapatkan reward, April ada 35 pelanggan mendapatkan reward dan 141 tidak mendapatkan reward, Mei ada 65 pelanggan yang mendapatkan reward dan 111 tidak yang mendapatkan reward, Juni ada 43 pelanggan yang mendapatkan reward dan 133 tidak mendapatkan reward, Juli ada 77 pelanggan yang mendapatkan reward dan 99 tidak mendapatkan reward, Agustus ada 48 pelanggan yang mendapatkan reward dan 128 tidak mendapatkan reward, September ada 94 pelanggan mendapatkan reward dan 82 tidak mendapatkan reward, dan Oktober ada 94 pelanggan mendapatkan reward dan 82 tidak mendapatkan reward.

**Kata Kunci:** *Clustering, K-Means Clustering, Reward, Davies-Bouldin Index*

### 1 Pendahuluan

Ketika dihadapkan dengan banyak pesaing dalam bisnis yang sama, pemilik agen harus melakukan upaya untuk mencegah penggantian pelanggan, karena pelanggan yang puas dapat membawa banyak keuntungan bagi agen, seperti memungkinkan tercapainya reward pelanggan pada agen Aqua dan Gas LPG FF Tirta yang merupakan salah satu agen yang berada di Jalan Radar Baru Rt 04 Rw 04, Kelurahan Kalisari, Kecamatan Pasar Rebo, Jakarta timur, kompleks pertokoan DITHUBAD. Oleh karena itu pemilik memberikan penghargaan kepada pelanggan yang bertransaksi dengan jumlah minimal 100 galon atau lebih gas dalam bentuk pelayanan yang unggul dan promosi khusus, namun permasalahan yang muncul saat ini adalah pemilik agen belum mengetahui mana pelanggan dengan data transaksi yang sesuai kriteria tersebut, sehingga pemilik agen kesulitan dalam memberikan reward kepada para pelanggannya dengan teknik data mining menggunakan *clustering* dalam penelitian ini menggunakan algoritma *K-Means Clustering* [1]. Tujuan data mining untuk memberikan solusi nyata bagi para pengambil keputusan di dunia bisnis, untuk mengembangkan bisnisnya, salah satu metode yang terdapat dalam data mining adalah *clustering* dimana metode ini dapat mengidentifikasi objek yang memiliki kesamaan karakteristik tertentu, dan kemudian menggunakan karakteristik tersebut sebagai vector karakteristik atau *centroid* [2].

Selanjutnya, uji clustering Davies-Bouldin Index (DBI) merupakan teknik clustering yang mengukur efektivitas cluster. Ukuran ini menggunakan indeks Davies-Bouldin memaksimalkan jarak antar cluster dengan meminimalkan jarak antara titik dalam cluster. Berdasarkan penjelasan di atas, maka dilakukan penelitian yang berjudul “Implementasi K-Means Clustering dengan menggunakan data transaksi penjualan untuk penentuan reward pada agen Aqua dan Gas LPG FF Tirta”. Tujuan penelitian ini untuk mengembangkan sistem yang dapat mengklasifikasikan dan menentukan jumlah pembelian pelanggan berdasarkan transaksi pembelian sehingga agen dapat mempertahankan prospek tersebut melalui layanan dan penghargaan yang unggul dalam bentuk promosi khusus.

## 2 Tinjauan Pustaka

*K-Means Clustering*, teknik clustering data non-hierarchical yang mengelompokkan data menjadi satu atau lebih cluster/grup. Data dengan karakteristik yang sama dikelompokkan ke dalam klaster/kelompok, dan data dengan karakteristik yang berbeda dikelompokkan dengan klaster/kelompok lain, sehingga data di dalam klaster/kelompok memiliki variabilitas yang lebih kecil. Ada dua metode *clustering* yang kita kenal, yaitu hierarchical clustering dan partitioning[3].

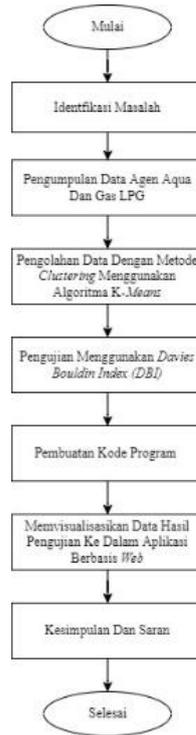
*Davies-Bouldin Index (DBI)*, metode pengelompokan yang mengukur efektivitas sebuah cluster, dan kohesi didefinisikan sebagai jumlah kedekatan data dengan pusat cluster dari cluster yang dilacak. Pemisahan disisi lain, didasarkan pada jarak dari pusat cluster ke cluster. Pengukuran menggunakan indeks Davies-Bouldin ini memaksimalkan jarak antara cluster  $C_i$  dan  $C_j$  sambil mencoba meminimalkan jarak antar titik dalam cluster. Jarak maksimum antar cluster.

Pada penelitian ini, terdapat beberapa penelitian terdahulu dengan metode yang sama salah satunya yaitu dilakukan oleh [4] terhadap tingkat kepuasan pelanggan pada perusahaan PT. TIKI Jalur Nugraha Ekakurir (JNE) Agen Mastrip Jember menggunakan algoritma *K-Means Clustering*. Hasil dari penelitian ini menghasilkan 4 Cluster. Penelitian lain juga dilakukan oleh [3] yang membahas tentang segmentasi pelanggan yang berfungsi untuk mengelompokkan pelanggan ke dalam beberapa kelompok guna mengetahui karakteristik pelanggan. Algoritma K-Means digunakan untuk pembentukan cluster berdasarkan model RFM (*Recency, Frequency, Monetary*) dengan menggunakan tools Weka 3.9. Metode Elbow berfungsi untuk menentukan jumlah cluster (k) terbaik. Penelitian selanjutnya yaitu pada PT. PLN (Persero) ULP Maros yang dilakukan oleh Ardimansyah, Firdaus, Iqram, dan Annah [1] yang memiliki lebih dari 50.000 pelanggan pascabayar dan lebih dari 40.000 pelanggan Prabayar.

Dari ketiga penelitian di atas, masing-masing menerapkan *K-Means* secara berbeda-beda, yaitu penelitian yang dilakukan oleh [4] pada perusahaan PT. TIKI Jalur Nugraha Ekakurir (JNE) Agen Mastrip Jember menggunakan *K-Means* clustering sebagai algoritma pengelompokan responden dengan uji validitas *Product Moment Pearson Correlation*. Penelitian kedua yang dilakukan oleh [3] yaitu menggunakan *K-Means* clustering sebagai segmentasi pelanggan untuk mengelompokkan pelanggan yang mempunyai karakteristik dengan uji model RFM. Penelitian ketiga yang dilakukan oleh [1] yaitu menggunakan *K-Means* clustering sebagai pengelompokan pelanggan untuk membedakan antara pelanggan setia dan pelanggan tidak jujur dalam hal tagihan listrik bulanan dengan pengujian *Blackbox*. Akan tetapi, ketiga penelitian tersebut tidak melakukan pengujian menggunakan *Davies-Bouldin Index (DBI)*. Sehingga keterbaruan penelitian ini adalah dilakukannya pengujian menggunakan *Davies-Bouldin Index (DBI)* berdasarkan 2 (dua) faktor yang digunakan yaitu faktor jumlah kedatangan dan jumlah pembelian. Pengujian ini bertujuan untuk membuktikan bahwa terdapat perbedaan metode pembayaran dalam menentukan pusat *cluster* berdasarkan jumlah *squared error*. *Davies-Bouldin Index (DBI)* digunakan menggunakan untuk mengukur validitas *cluster*, dan dapat memaksimalkan jarak inter-cluster. Untuk itu, diajukanlah penelitian ini dengan metode *K-Means Clustering* dengan pengujian menggunakan *Davies-Bouldin Index (DBI)*.

## 3 Metode Penelitian

Data yang digunakan pada penelitian ini adalah data pelanggan agen Aqua dan Gas LPG FF Tirta sebanyak 176 pelanggan. Metode pengembangan pada penelitian ini menggunakan metode *K-Means* clustering dan diuji menggunakan *DBI*. Proses pada penelitian ini dilakukan yaitu dengan cara identifikasi masalah, pengumpulan data Agen Aqua Dan Gas LPG, pengolahan data dengan metode clustering menggunakan algoritma *K-Means*, pengujian menggunakan *Davies Bouldin Index (DBI)*, pembuatan kode program, memvisualisasikan data hasil pengujian ke dalam aplikasi berbasis web, dan menarik kesimpulan dan saran.



**Gambar. 1.** Tahapan Penelitian

Yang pertama yaitu tentukan jumlah cluster (k) pada data set. Jadi, nilai k dapat ditentukan dengan syarat dimana nilai k harus lebih kecil dari jumlah data. Lalu tentukan nilai pusat (centroid). Jadi, penentuan nilai centroid pada tahap awal dilakukan secara acak dari masing-masing kelompok (*cluster*). Yaitu sebagai berikut ini:

$$\frac{(\text{Jumlah Data})}{(\text{Jumlah Class}+1)} \quad (2.1)$$

Keterangan :

Jumlah data : Jumlah data yang akan digunakan.

Jumlah class : Jumlah kelompok yang telah ditentukan sebelumnya

Jadi, pada tahap iterasi digunakan rumus rata-rata, sebagai berikut ini:

$$V_{ij} = \frac{1}{N_i} = \sum_{k=0}^{N_i} x_{kj} \quad (2.2)$$

Keterangan:

$V_{ij}$ : centroid rata-rata cluster ke-i untuk variabel ke-j  $N_i$  : jumlah anggota cluster ke-I

$i, k$ : indeks dari cluster  $j$  : indeks dari variabel

$x_{kj}$ : nilai data ke-k variabel ke-j dalam cluster tersebut.

Pada masing-masing record, kita hitung jarak terdekat dengan centroid. Jarak centroid yang digunakan adalah dengan menggunakan rumus *Euclidean Distance*, dengan rumus seperti pada persamaan 2.3 :

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (2.3)$$

Keterangan:

De: *Euclidean Distance*

I: banyaknya objek

(x, y): koordinat objek (nilai datanya)

(s, t): koordinat centroid (nilai centroid)

Tentukan rata-rata dari data yang ada di masing-masing cluster. Lalu kita lihat, apakah posisi centroid berpindah atau tidak.

## 4 Pembahasan

### 1. Data

Tabel 1. Data Jumlah Pembelian dan Jumlah Kedatangan

| No     | Nama Pelanggan | Oktober 2020 - Oktober 2021 |     |                  |     |
|--------|----------------|-----------------------------|-----|------------------|-----|
|        |                | Jumlah Kedatangan           |     | Jumlah Pembelian |     |
|        |                | Galon                       | Gas | Galon            | Gas |
| 1      | Gumilereng     | 30                          | 4   | 96               | 4   |
| 2      | Aipangso       | 30                          | 4   | 91               | 4   |
| 3      | Tumangger      | 30                          | 4   | 94               | 4   |
| 4      | Waworantu      | 30                          | 4   | 89               | 4   |
| 5      | Mutangon       | 30                          | 4   | 91               | 4   |
| ...176 | ...            | ...                         | ... | ...              | ... |

Pada tabel di atas ini terdapat data pelanggan pada bulan Oktober 2020 sampai Oktober 2021, yaitu data sebanyak 176 pelanggan dengan jumlah kedatangan dan jumlah pembelian pelanggan per bulan.

### 2. Proses Clustering

Menentukan nilai parameter pada algoritme K-Means yaitu jumlah cluster yang akan dibentuk, maksIterasi untuk menentukan batas iterasi. Data yang digunakan pada perhitungan algoritme K-Means adalah data *customer* yang berjumlah 176 data dengan 2 fitur. Fitur ke-1 merepresentasikan jumlah pembelian pelanggan. Kemudian fitur ke-2 akan merepresentasikan jumlah pembelian kedatangan pelanggan selama sebulan. Berikut ini adalah data yang digunakan pada proses perhitungan manualisasi K-Means yaitu sample data pada bulan September 2021 yang ditunjukkan pada Table 4.2 Yang mana Fitur 1 dan Fitur 2 merupakan inisialisasi dari jumlah pembelian dan jumlah kedatangan.

Tabel 2. Data Yang Akan Diproses

| No     | Nama Pelanggan | September 2021             |     |                             |     |
|--------|----------------|----------------------------|-----|-----------------------------|-----|
|        |                | Jumlah Pembelian (Fitur 1) |     | Jumlah Kedatangan (Fitur 2) |     |
|        |                | Galon                      | Gas | Galon                       | Gas |
| 1      | Gumilereng     | 97                         | 4   | 30                          | 4   |
| 2      | Aipangso       | 92                         | 4   | 30                          | 4   |
| 3      | Tumangger      | 90                         | 4   | 30                          | 4   |
| 4      | Waworantu      | 98                         | 4   | 30                          | 4   |
| 5      | Mutangon       | 99                         | 4   | 30                          | 4   |
| ...176 | ...            | ...                        | ... | ...                         | ... |

### 3. Proses Clustering

*Clustering* berdasarkan Jumlah Pembelian Pelanggan dapat dijelaskan sebagai berikut:

Menentukan nilai parameter pada algoritme K-Means yaitu jumlah cluster yang akan dibentuk, maksIterasi untuk menentukan batas iterasi. Menentukan jumlah cluster sebanyak K setiap data. Berikut ini adalah tabel hasil penentuan klaster secara acak ditunjukkan pada Tabel 3.

Tabel 3. Penentuan Cluster

| No     | Nama Pelanggan | September 2021 |     |         |     |
|--------|----------------|----------------|-----|---------|-----|
|        |                | Fitur 1        |     | Fitur 2 |     |
|        |                | Galon          | Gas | Galon   | Gas |
| 1      | Gumilereng     | 97             | 4   | 30      | 4   |
| 2      | Aipangso       | 92             | 4   | 30      | 4   |
| 3      | Tumangger      | 90             | 4   | 30      | 4   |
| 4      | Waworantu      | 98             | 4   | 30      | 4   |
| 5      | Mutangon       | 99             | 4   | 30      | 4   |
| ...176 | Juprianto      | 86             | 28  | 30      | 28  |

Untuk setiap iterasi lakukan perhitungan nilai centroid pada setiap cluster.

**Tabel 4.** Hasil Penjumlahan Fitur Untuk Setiap Cluster

| Fitur 1                 | Fitur 2 | K1 | K2 | K1x   | K1y  | K2x  | K2y  |
|-------------------------|---------|----|----|-------|------|------|------|
| 101                     | 34      | 1  |    | 101   | 34   |      |      |
| 96                      | 34      |    | 1  |       |      | 96   | 34   |
| 94                      | 34      |    | 1  |       |      | 94   | 34   |
| 102                     | 34      | 1  |    | 102   | 34   |      |      |
| 103                     | 34      | 1  |    | 103   | 34   |      |      |
| 114                     | 58      | 1  |    | 114   | 58   |      |      |
| Total s/d 176 pelanggan |         | 91 | 75 | 17511 | 4382 | 7617 | 3097 |

Lalu kita cari nilai centroidnya. Berikut ini adalah perhitungan nilai centroid untuk clusternya:

- Cluster 1 Fitur 1

$$C_{(i,j)} = \frac{\text{Jumlah Cluster } i}{\text{Jumlah Fitur } j \text{ Pada Cluster } i}$$

$$C_{(2,2)} = \frac{91}{17511}$$

$$C_{(2,2)} = 0,0051$$

- Cluster 1 Fitur 2

$$C_{(i,j)} = \frac{\text{Jumlah Cluster } i}{\text{Jumlah Fitur } j \text{ Pada Cluster } i}$$

$$C_{(2,2)} = \frac{91}{7617}$$

$$C_{(2,2)} = 0,0119$$

- Cluster 2 Fitur 1

$$C_{(i,j)} = \frac{\text{Jumlah Cluster } i}{\text{Jumlah Fitur } j \text{ Pada Cluster } i}$$

$$C_{(2,2)} = \frac{75}{4382}$$

$$C_{(2,2)} = 0,0171$$

- Cluster 2 Fitur 2

$$C_{(i,j)} = \frac{\text{Jumlah Cluster } i}{\text{Jumlah Fitur } j \text{ Pada Cluster } i}$$

$$C_{(2,2)} = \frac{75}{3097}$$

$$C_{(2,2)} = 0,02$$

**Tabel 5.** Nilai Centroid

| Kelompok | Centroid Fitur ke-1 | Centroid Fitur ke-2 |
|----------|---------------------|---------------------|
| 1        | 0,0051              | 0,0119              |
| 2        | 0,0171              | 0,0242              |

Melakukan perhitungan jarak data dengan nilai centroid menggunakan perhitungan euclidean distance.

❖ Bulan September 2021

- Fitur 1 Cluster 1 pada pelanggan

Gumilereng

$$d(x1, C1) = \sqrt{(x(1,1) - C(1,1))^2 + (y(1,2) - C(1,2))^2}$$

$$d(x1, C1) = \sqrt{(101 - 0,0051)^2 + (34 - 0,0119)^2}$$

$$d(x1, C1) = \sqrt{(100,99)^2 + (33,98)^2}$$

$$d(x1, C1) = \sqrt{(10.198,9801) + (1.154,6404)}$$

$$d(x1, C1) = \sqrt{11.353,6205}$$

$$d(x1, C1) = 106,55336$$

- Fitur 2 Cluster 2 pada pelanggan Gumilereng

$$d(y2, C2) = \sqrt{(x(1,2) - C(1,2))^2 + (y(2,2) - C(2,2))^2}$$

$$d(y2, C2) = \sqrt{(101 - 0,0171)^2 + (34 - 0,0242)^2}$$

$$d(y2, C2) = \sqrt{(100,9829)^2 + (33,9758)^2}$$

$$d(y2, C2) = \sqrt{(10.197,5461) + (1.154,354)}$$

$$d(y2, C2) = \sqrt{11.351,900}$$

$$d(y2, C2) = 106,54529$$

Nantinya semua perhitungan berdasarkan contoh hasil di atas dimasukkan ke dalam table berikut ini:

**Tabel 6.** Hasil Perhitungan Jarak *Euclidean Distance*

| Jarak C1          | Jarak C2          | Jarak Minimal     |
|-------------------|-------------------|-------------------|
| 1256,173842       | 1255,349886       | 1255,349886       |
| 1251,173842       | 1250,349886       | 1250,349886       |
| 1249,173842       | 1248,349886       | 1248,349886       |
| 1257,173842       | 1256,349886       | 1256,349886       |
| 1258,173842       | 1257,349886       | 1257,349886       |
| s/d 176 pelanggan | s/d 176 pelanggan | s/d 176 pelanggan |

Menentukan cluster baru berdasarkan data yang dekat dengan nilai centroid.

Berikut ini adalah tabel untuk menentukan cluster baru berdasarkan jarak terdekat dengan nilai centroid ditunjukkan pada Tabel 7.

**Tabel 7.** Penentuan Cluster Baru

| Jarak C1          | Jarak C2          | Jarak Minimal     | Cluster Baru      |
|-------------------|-------------------|-------------------|-------------------|
| 1256,173842       | 1255,349886       | 1255,349886       | C2                |
| 1251,173842       | 1250,349886       | 1250,349886       | C2                |
| 1249,173842       | 1248,349886       | 1248,349886       | C2                |
| 1257,173842       | 1256,349886       | 1256,349886       | C2                |
| 1258,173842       | 1257,349886       | 1257,349886       | C2                |
| s/d 176 pelanggan | s/d 176 pelanggan | s/d 176 pelanggan | s/d 176 pelanggan |

Berdasarkan tabel 7 dapat disimpulkan bahwa algoritma K-Means dapat dengan baik menentukan cluster pada data jumlah pembelian customer dengan hasil semua cluster tidak berpindah yaitu semua masuk ke dalam cluster 2. Oleh karena itu proses iterasi pun berhenti.

### Pengujian DBI

Data Centroid, menggunakan Data Centroid yang ada diproses K-Means Clustering di atas.

**Tabel 8.** Data Centroid

| Kelompok | Centroid Fitur ke-1 | Centroid Fitur ke-2 |
|----------|---------------------|---------------------|
| 1        | 0,0171              | 0,0119              |
| 2        | 0,0051              | 0,0242              |

Perhitungan SSW (*Sum of Square Within-cluster*)

**Tabel 9.** Perhitungan SSW

| No.    | Jarak C1    | Jarak C2    | Jarak Minimal | Cluster Baru |
|--------|-------------|-------------|---------------|--------------|
| 1      | 1256,173842 | 1255,349886 | 1255,349886   | C2           |
| 2      | 1251,173842 | 1250,349886 | 1250,349886   | C2           |
| 3      | 1249,173842 | 1248,349886 | 1248,349886   | C2           |
| 4      | 1257,173842 | 1256,349886 | 1256,349886   | C2           |
| 5      | 1258,173842 | 1257,349886 | 1257,349886   | C2           |
| ...176 | 3476,602642 | 3475,188286 | 3475,188286   | C2           |
|        |             | Total       | 353369,0759   |              |

SSW Cluster 1:

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

$$SSW_i = \frac{\text{Semua yg termasuk cluster 1}}{\text{Jumlah cluster 1}}$$

$$SSW_i = \frac{0}{0}$$

$$SSW_i = 0$$

SW Cluster 2:

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

$$SSW_i = \frac{\text{Semua yg termasuk cluster 2}}{\text{Jumlah cluster 2}}$$

$$SSW_i = \frac{353369,0759}{176}$$

$$SSW_i = 2007,70$$

Perhitungan *Sum of Square Between-cluster* (SSB)

**Tabel 10.** Perhitungan SSB

| Kelompok | Centroid Fitur ke-1 | Centroid Fitur ke-2 |
|----------|---------------------|---------------------|
| 1        | 0,0171              | 0,0119              |
| 2        | 0,0051              | 0,0242              |

- Jarak cluster 1 ke cluster 1 (0,0100 dan 0,0102):  

$$\sqrt{(0,0100 - 0,0100)^2 + (0,0102 - 0,0102)^2} = 0$$
- Jarak cluster 1 ke cluster 2 (0,0100, 0,0102) dan (0,0284, 0,0289):  

$$= \sqrt{(0,0100 - 0,0284)^2 + (0,0102 - 0,0289)^2}$$

$$= \sqrt{(-0,0184)^2 + (-0,0187)^2}$$

$$= \sqrt{(0,00033) + 0,00034}$$

$$= 0,00067$$
- Jarak cluster 2 ke cluster 1 (0,0284, 0,0289) dan (0,0100, 0,0102):  

$$= \sqrt{(0,0284 - 0,0100)^2 + (0,0289 - 0,0102)^2}$$

$$= \sqrt{(0,0184)^2 + (0,0187)^2}$$

$$= \sqrt{(0,00033856) + 0,00034969}$$

$$= 0,000678825$$
- Jarak cluster 2 ke cluster 2 (0,0284, 0,0289):  

$$\sqrt{(0,0284 - 0,0284)^2 + (0,0289 - 0,0289)^2} = 0$$

**Tabel 11.** Hasil Perhitungan SSB (*Sum of Square Between-cluster*)

| SSB | Centroid   |            |
|-----|------------|------------|
|     | 1          | 2          |
| 1   | 0          | 0,01215129 |
| 2   | 0,01215129 | 0          |

1) Perhitungan Ratio

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}}$$

$$R_{i,j} = \frac{0 + 2007,7}{0,01215129}$$

$$R_{i,j} = 0,679132$$

$$DBI = \frac{1}{K} \sum_{j=1}^k \max_{1 \neq j} R_{i,j}$$

$$DBI = \frac{1}{\text{Banyak Cluster}} * \text{Ratio}$$

$$DBI = \frac{1}{2} * 0,679132$$

$$DBI = 0,3420$$

Perhitungan *Davies Bouldin Index* (DBI)

Hasil dari perhitungan DBI di atas yaitu sejumlah 0.3420. Hasil DBI tersebut didapatkan hasil yang kecil yaitu mendekati 0, hal itu membuktikan bahwa data yang digunakan merupakan data yang bagus. Dikarenakan semakin kecil nilai yang didapatkan, maka semakin bagus datanya. Presentase Keberhasilan Menggunakan K Means Clustering Dan Diuji Dengan DBI.

**Tabel 12.** Presentase Keberhasilan

| No.   | Jumlah Data Pelanggan | Hasil Pengurangan Centroid | %   |
|-------|-----------------------|----------------------------|-----|
| 1.    | 176                   | 0,029                      | 60% |
| 2.    | 176                   | 0,0191                     | 38% |
| Total |                       |                            | 98% |

Hasil presentase keberhasilan menggunakan *K Means clustering* dan diuji dengan DBI di atas yaitu sebesar 98% yang artinya semua data yang diproses termasuk data yang sangat baik yaitu mendekati 100%.

### Hasil Implementasi

Pada gambar di bawah ini yaitu semua tampilan visualisasi berbasis web yang akan menghubungkan semua sistem berdasarkan tampilan yang diinginkan, yaitu terdapat tampilan halaman login, halaman dashboard, halaman user, halaman barang, halaman pelanggan, halaman transaksi dan halaman K-Means. Berikut adalah tampilan halaman Login seperti pada gambar 4.5



Gambar. 2. Tampilan Visualisasi Berbasis Web

## 5 Kesimpulan

Implementasi metode k-means clustering dapat dilakukan untuk penentuan dengan berdasarkan jumlah pembelian dan kedatangan pelanggan terbanyak di agen aqua dan gas LPG FF Tirta. Hasil implementasi dapat memvisualisasikan dalam penentuan pemberian reward kepada pelanggan menggunakan data transaksi Agen Aqua dan Gas LPG yang diterapkan pada aplikasi berbasis web. Yang mana terdapat hasil pelanggan yang mendapatkan reward dan pelanggan yang tidak mendapatkan reward serta hasil DBI. Hasil presentase keberhasilan menggunakan *K Means clustering* dan diuji dengan DBI di atas yaitu sebesar 98% yang artinya semua data yang diproses termasuk data yang sangat baik yaitu mendekati 100%. Hasil olahan data pelanggan menggunakan metode k-means yang menghasilkan pembagian pelanggan yang mendapatkan reward dan pelanggan yang tidak mendapatkan reward per bulan-nya, sehingga memudahkan pemilik agen dalam memberikan promo spesial kepada pelanggan dengan menggunakan proses perhitungan *K Means clustering* dan diuji oleh DBI. Hasil perhitungan *K Means clustering* berjumlah 353369,0759 dan hasil DBInya berjumlah 0,3420 yang artinya nilai DBInya bagus yaitu mendekati 0, serta hasil persentase keberhasilannya sebesar 98% yang artinya sangat baik yaitu mendekati 100%.

## Referensi

- [1] A. Ardiansyah, Firdaus, A. T. Iqam, dan Annah. "Implementasi Metode K-Means Clustering Pada Aplikasi Analisis Loyalitas Pelanggan PT. PLN (Persero) ULP Maros Berbasis Progressive Web Apps." dalam *Seminar Ilmiah Sistem Informasi Dan Teknologi Informasi X(1)*, 2021, hal. 98-107.
- [2] R. Hamdani, "Pengelompokan Loyalitas Pelanggan Dengan Menggunakan Kombinasi Rfm Dan Algoritma K-Means," *Journal Of Informatic Pelita Nusantara*, vol. 5, no. 1, hal. 7-13, 2020.
- [3] I. Maskanah, A. Primajaya, dan A. Rizal, "Segmentasi Pelanggan Toko Purnama dengan Algoritma K-Means dan Model RFM untuk Perancangan Strategi Pemasaran," *Inovtek Polbeng - Seri Informatika*, vol. 5, no. 2, hal. 218-228, 2020.
- [4] M. B. Maulana, Slamain, dan O. Juwita, "250XRancang Bangun Aplikasi Customer Relationship Management (CRM) Untuk Identifikasi Tingkat Kepuasan Pelanggan Pada Perusahaan PT. TIKI Jalur Nugraha Ekakurir (JNE) Agen Mastrip Jember Menggunakan Metode K-Means Clustering," *Informatics Journal*, vol. 2, no. 2, hal. 92-100, 2017.