

## PENERAPAN METODE *K-MEANS CLUSTERING* DALAM MENGELOMPOKKAN JUMLAH PESERTA BPJS KESEHATAN JKN/KIS DI KABUPATEN CIREBON

Tri Ananda Widyaningsih<sup>1)</sup>, Denni Pratama<sup>2)</sup>

<sup>1,2)</sup>Jurusan Manajemen Informatika, STMIK IKMI Cirebon

JL. Perjuangan No.10,B, Karyamulya, Kec.Kesambi, Kota Cirebon,Jawa Barat 45131

Email : <sup>1)</sup> [trianandawidyaningsih28@gmail.com](mailto:trianandawidyaningsih28@gmail.com) <sup>2)</sup> [Pratamadenni@gmail.com](mailto:Pratamadenni@gmail.com)

### ABSTRAK

BPJS Kesehatan adalah Badan Penyelenggara Jaminan Sosial Kesehatan yang dilaksanakan oleh pemerintah sesuai dengan UU No.40 Tahun 2004 Peserta yang menjadi anggota akan mendapatkan Kartu Indonesia Sehat ,karena Jumlah Peserta BPJS Kesehatan JKN/KIS yang belum merata di setiap wilayah yang ada di kabupaten Cirebon ada yang sedikit dan banyak. Penelitian ini bertujuan untuk mengelompokkan peserta BPJS Kesehatan Jkn/Kis di Kabupaten Cirebon ke dalam beberapa kelompok *sample* penelitian ini di peroleh dari Dataset Open Data Jabar yaitu Jumlah Peserta BPJS Kesehatan dengan Jumlah 412 *Dataset* Peserta BPJS Kesehatan JKN/KIS di Kabupaten Cirebon. Metode *K-Means* adalah metode yang tepat untuk di gunakan untuk mengelompokkan jumlah Peserta BPJS Kesehatan Jkn/Kis Di Kabupaten Cirebon yang cukup banyak dengan waktu yang relatif cepat dan efisien dengan menggunakan *machine learning* dengan *tools Rapidminer*. Hasil pengelompokan dinilai dengan *Davies Bouildin Index* untuk mengetahui hasil optimasi terhadap algoritma *K-Means*. Hasil Clustering didapatkan kelompok terbanyak peserta BPJS kesehatan maka Cluster\_0: Kelompok Rendah sejumlah 86 Peserta, Cluster\_1: Kelompok sedang Peserta BPJS Kesehatan sejumlah 156 Peserta, dan Cluster\_2: Kelompok Banyak sejumlah 170 Peserta. Dengan nilai K=3 sebagai nilai Optimum dengan nilai DBI = 0.164.

**Kata kunci** : bpjs, jkn, *data mining*, *clustering*

### ABSTRACT

*This research aims to find solutions related to the existence of queues at the BNI Head Office, and this is because, at the Head Office, problems related to customers or consumers often cannot be resolved at Branch Offices, especially associated with the approval of the Bank's Regional Leaders, transactions over one billion, product sales banking and so on. Through the analysis of the initial situation, consumers often complain because there is a queue that is quite long, namely the average customer waits between 40 minutes and 1 hour at customer service. This, besides giving the impression of discomfort and disappointment. As a service provider, banking is more complex in providing service products followed by product, price, place, promotion, people, process and physical evidence. This study uses a queue simulation approach to determine the effectiveness of the service with the arrival distribution pattern is Poison. This is because the arrival of customers with irregular and random arrival intervals and the number of customers arriving with available servers seem unbalanced. The results of this study are that it is necessary to change the pattern of service from M/M/1 to M/M/S, a minimum addition of one server is required, a gap of about half an hour is required before the end of operating hours for service closure and the need for increased skills of waiters.*

**Keywords** : bpjs, jkn, *data mining*, *clustering*

## 1. PENDAHULUAN

Kondisi kesehatan di Indonesia masih cukup buruk karena belum meratanya pelayanan fasilitas kesehatan serta kesenjangan sosial ekonomi yang berbeda di setiap wilayah sehingga untuk masyarakat yang memiliki tingkat ekonomi cukup dapat memenuhi fasilitas kesehatan yang layak [1]. Kesehatan sangat penting untuk semua orang yang hidup di dunia, khususnya di Indonesia yang memiliki tingkat populasi penduduk terbanyak selain Cina sehingga tingkat fasilitas dan pelayanan kesehatan di suatu negara karena sebagai aspek terpenting dalam proses pembangunan serta mempengaruhi *produktivitas* sumber daya manusia yang ada [2], serta tubuh yang sehat akan membuat manusia dapat melakukan berbagai aktivitas sehari-hari untuk memenuhi kebutuhan hidup sandang, pangan, papan maupun fasilitas kesehatan saat sedang sakit, oleh karena itu pemerintah membuat suatu sistem dan pelayanan untuk memenuhi fasilitas kesehatan yang layak, bermutu dan berkualitas [3]. Karena negara Indonesia adalah salah satu negara berkembang penting untuk meningkatkan dan melaksanakan Program Jaminan Sosial berdasarkan *Funded Social Security* yang artinya jaminan sosial yang masih dibiayai oleh peserta yang memiliki pekerjaan formal [4].

Sebelum Badan Penyelenggara Jaminan Sosial Kesehatan terbentuk pemerintah awalnya membentuk Jaminan Sosial Nasional (JKN) dengan program yang ada di dalamnya yaitu Jaminan Kesehatan Masyarakat (Jamkesmas) kemudian berubah berganti nama menjadi Badan Penyelenggara Jaminan Sosial Kesehatan (BPJS Kesehatan) [5].

BPJS Kesehatan memiliki tujuan untuk memberikan fasilitas dan pelayanan kesehatan yang layak dan bermutu untuk semua orang sesuai dengan Undang-Undang No 36 tahun 2009 yang menyatakan bahwa "Setiap orang mempunyai hak yang sama untuk mendapatkan fasilitas kesehatan yang terjangkau, bermutu dan terjangkau, serta

setiap orang berkewajiban ikut serta dalam program jaminan Kesehatan sosial". Badan Penyelenggara Jaminan Sosial terbagi menjadi dua yaitu BPJS Kesehatan dan BPJS Ketenagakerjaan [6]. BPJS Ketenagakerjaan memiliki tujuan untuk menyelenggarakan program jaminan saat kecelakaan kerja terjadi Jaminan Hari Tua, Jaminan Pensiun dan Jaminan akibat kematian yang disebabkan faktor apapun [7].

Penelitian terdahulu yang dilakukan oleh Salvan, Handikristanto, dan Edora (2022) berjudul Klasterisasi Data Penggunaan Layanan BPJS Kesehatan Menggunakan Algoritma *K-Means* dengan permasalahan yang akan diteliti yaitu penyetaraan pengguna Layanan BPJS Kesehatan dan menggunakan 1000 *dataset* kategori nama faskes, Menggunakan metode klasterisasi algoritma *K-Means* untuk mengelompokkan penggunaan layanan BPJS kesehatan ke dalam beberapa *cluster*, dan pada penelitian ini akan terbentuk kelompok *Cluster Medium/Cluster 0* yang memiliki 474 anggota kelompok, *Cluster Low /Cluster 1* memiliki 353 anggota kelompok dan *Cluster High/Cluster 2* memiliki 173 anggota kelompok [8].

## 2. DASAR TEORI

Dasar-Dasar teori yang digunakan dalam penelitian ini yaitu BPJS Kesehatan JKN/KIS, *Data Mining*, *Clustering*, *K-Means*, *Rapidminer*, *Davies-Bouldin Index*.

### 2.1. BPJS Kesehatan JKN/KIS

BPJS Kesehatan adalah jaminan sosial yang di bentuk sesuai UU.No 36 Tahun 2009 dan ditunjuk untuk melaksanakan jaminan sosial sesuai UU.No 24 Tahun 2011[9]. Tujuan BPJS agar terwujudnya hak dan fasilitas kesehatan yang layak, bermutu, terjamin dan terjangkau [10]. Memiliki prinsip *Managed Care* yaitu sistem yang menggabungkan pembiayaan serta pelayanan untuk memenuhi pelayanan kesehatan yang layak dan bermutu bagi setiap orang [11].

## 2.2. Data Mining

*Data Mining* yaitu penyaringan data secara tersirat [12]. untuk mengetahui nilai dari data dengan menggali pola baru dari suatu *dataset* yang tersedia [13]. *Data mining* yaitu proses analisa untuk menyimpulkan hipotesis dengan memahami data sehingga bermanfaat untuk peliti saat hasil sudah di temukan dan didapatkan kesimpulan serta untuk pengguna sebagai bahan referensi [14].

## 2.3. K-Means

Algoritma *K-Means* merupakan suatu pengelompokan untuk mengetahui jumlah kelompok melalui nilai *centroid* awal dengan mengulang secara berkala [15].

Langkah langkah untuk menentukan Algoritma *K-Means* adalah sebagai berikut:

1. Tentukan nilai k atau jumlah *cluster* pada *dataset*
2. Penentuan nilai *centroid* pada tahap awal di lakukan secara random sedangkan pada tahap iterasi di harus digunakan dengan rumus seperti di bawah ini:

$$v_{IJ} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \quad (1)$$

Keterangan

$v_{ij}$  = *Centroid* rata -rata *cluster* ke-I untuk variable 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 untuk *cluster* tersebut

3. Menghitung jarak antara titik *centroid* dengan titik objek menggunakan *Euclidean Distance* dengan rumus seperti di bawah ini

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

Keterangan:

$De$  = *Euclidean Distance*

$i$  = Banyaknya Objek

$(x, y)$  = Koordinat Objek

$(s, t)$  = Koordinat *Centroid*

- 4). Kelompokkan objek berdasarkan jarak ke *Centroid* terdekat dengan cara melakukan iterasi hingga *centroid* bernilai optimal dengan menggunakan rumus pada point ke 2 dan 3.

## 2.4. Clustering

*Clustering* yaitu suatu proses pengelompokan pada data menjadi beberapa kelompok yang terdapat pada data yang memiliki kemiripan yang telah dilakukan pengujian [16]. *Clustering* yaitu proses pembentukan *cluster*/kelompok yang sebelumnya belum diketahui oleh siapaun berdasarkan *dataset* yang digunakan [17].

## 2.5. Knowledge

*Knowledge Discovery in Database* (KDD) adalah proses untuk menghasilkan suatu pola serta model dengan menggunakan algoritma untuk mengembangkan pola dan model terbaru, Algoritma yang digunakan untuk menghasilkan pola serta model nya juga berbeda-beda baik itu *Random Forest*, *Naïve Naves*, *Decision tree*, KNN, *K-Medoid*, *K-Means* atau pun Klasifikasi [18].

## 2.6. RapidMiner

*RapidMiner* yaitu suatu perangkat/software yang berguna untuk melakukan proses pengolahan *data mining* dengan cara memeriksa dan menganalisa terlebih dahulu pola dan atribut- atribut yang akan di gunakan nantinya [19]. *RapidMiner* juga merupakan Perangkat lunak yang dapat digunakan oleh siapa saja serta bersifat terbuka (*Open Source*) yang digunakan untuk berbagi pengolahan *data mining* [20].

## 2.7. Davies Bouildin Index

*Davies-Bouldin Index* adalah salah satu metode evaluasi internal yang dapat mengukur hasil evaluasi *cluster* suatu pengelompokan yang akan dilakukakan, Semakin kecil nilai DBI yang diperoleh dari hasil percobaan dengan nilai yang bukan *negative*, yaitu nilai yang mendekati 0 maka

hasilnya akan semakin baik untuk jumlah *cluster* yang ditentukan.

### 3. METODE PENELITIAN

#### 1. Jenis Penelitian

Pada penelitian ini peneliti menggunakan pendekatan deskriptif kualitatif yaitu penelitian yang memiliki tujuan untuk memberikan penjelasan terhadap fenomena yang terjadi saat ini untuk menggambarkan secara sistematis, faktual dan akurat mengenai sifat dan data yang diteliti.

#### 2. Sumber Data

Data yang akan di teliti bersumber dari data open data jabar mengenai data jumlah peserta BPJS Kesehatan JKN/KIS di Jawa Barat dengan dan data yang akan di teliti khusus pada jumlah peserta BPJS Kesehatan JKN/KIS di Kabupaten Cirebon <https://opendata.jabarprov.go.id/id>.

#### 3. Metode Penelitian

Metode Penelitian yang digunakan pada peneliti ini adalah metode *K-Means clustering* dengan mengklaster jumlah Peserta BPJS Kesehatan JKN/KIS.

### 4. HASIL DAN DISKUSI

Data Penelitian diperoleh dari data Jumlah peserta BPJS Kesehatan JKN/KIS di Kabupaten Cirebon tahun 2021 dengan jumlah *record data* yang diambil untuk penelitian 412 *dataset*.

#### 1. Data Selection

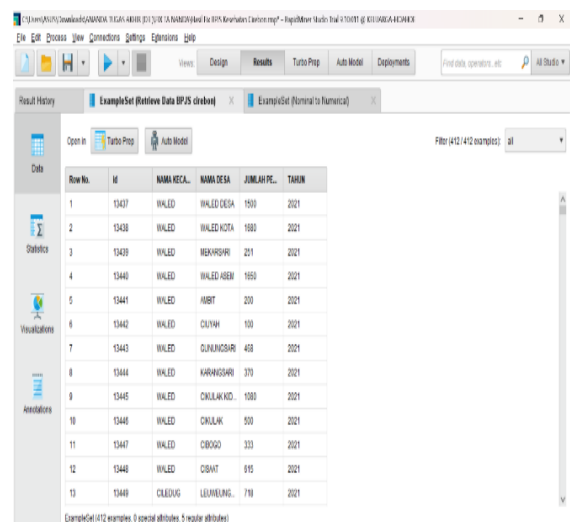
Proses pemilihan dari *dataset* yang relevan untuk mempermudah proses analisa data menggunakan Algoritma *K-Means* dengan cara membersihkan data agar tidak memiliki nilai (*null*). Data yang digunakan yaitu Data Jumlah Peserta BPJS Kesehatan JKN/KIS di Kabupaten Cirebon dengan *Record* 412 *dataset* yang di peroleh dari *Open Jabar* yang akan diolah menggunakan *Rapidminer*.

**Tabel 1.** *Data Sample* Peserta BPJS Kesehatan

No	Id	Nama Kecamatan	Nama Desa	Jumlah Peserta	tahun
1	13437	Waled	Waled Desa	1500	2021
2	13438	Waled	Waled Kota	1680	2021
3	13439	Waled	Mekarsari	251	2021
4	13440	Waled	Waled Asem	1650	2021
5	13441	waled	Ambit	200	2021
6	13442	Waled	Ciuyah	100	2021
7	13443	Waled	Gunungsari	82	2021
8	13444	Waled	Karangsari	370	2021
9	13445	Waled	Cikulak Kidul	1080	2021
10	13446	Waled	Cikulak	500	2021
...	...	...	...	...	...
411	13848	Jamblang	Bakung Kidul	1870	2021
412	13847	Jamblang	Bakung Lor	4500	2021

#### 2. Data Cleaning

Proses *cleaning* dilakukan sebelum tahap data mining di lakukan untuk membuang duplikat data ,memeriksa data yang missing value dan data yang kosong agar tidak mengurangi hasil dari nilai akurasi yang akan di hasilkan nanti.



**Gambar 1.** Hasil *Data Cleaning*

Gambar 1 merupakan hasil dari proses pemersihan pada data yang kosong dengan cara mengaktifkan *replace*

*errors with missing values* pada saat proses memasukan data ke dalam *rapidminer* dan menghasilkan tidak ada data yang kosong dan *missing*, sehingga jumlah data peserta BPJS Kesehatan JKN/KIS di Kabupaten Cirebon tetap sebanyak 412 data.

### 3. Data Transformation

Pada tahap ini di lakukan proses perubahan data dengan cara memasukkan dua operator yaitu *Operator Nominal to Numeric* untuk merubah bentuk nominal menjadi numerik atau angka agar dapat dibaca saat proses pengolahan data melalui *Rapidminer*.

Row No.	Id	NAMA KECAMATAN	NAMA DESA	JUMLAH PESERTA	TAHUN
1	13437	0	0	1000	2021
2	13438	0	1	1980	2021
3	13439	0	2	201	2021
4	13440	0	3	1650	2021
5	13441	0	4	200	2021
6	13442	0	5	100	2021
7	13443	0	6	458	2021
8	13444	0	7	370	2021
9	13445	0	8	1000	2021
10	13446	0	9	500	2021
11	13447	0	10	333	2021
12	13448	0	11	615	2021
13	13449	1	12	718	2021

Gambar 2. Hasil Transformasi Data

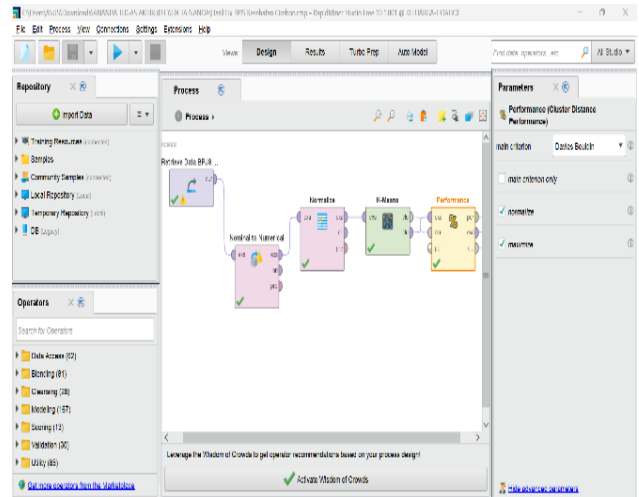
Dan tahap kedua dalam Data Transformasi yaitu dengan melakukan normalisasi menggunakan *Operator Normalize* yang berfungsi untuk memperkecil jarak antar nominal pada ketetapan Jumlah Peserta BPJS Kesehatan JKN/KIS di Kabupaten Cirebon karena jika jarak yang ada terlalu besar maka akan mempengaruhi hasil dari pemodelan dan kurang optimal menggunakan *rapidminer* pada metode transformasi yang digunakan adalah metode *range transformation* yaitu metode dimana nilai jarak yang diinginkan bisa ditentukan. Pada penelitian ini parameter yang digunakan adalah nilai min = 0.0 dan nilai max = 1.0.

Row No.	NAMA KECA.	NAMA DESA	Id	JUMLAH PE.	TAHUN
1	0	0	0	0.189	1
2	0	0.002	0.002	0.212	1
3	0	0.005	0.005	0.032	1
4	0	0.007	0.007	0.206	1
5	0	0.010	0.010	0.025	1
6	0	0.012	0.012	0.313	1
7	0	0.015	0.015	0.059	1
8	0	0.017	0.017	0.047	1
9	0	0.020	0.019	0.136	1
10	0	0.022	0.022	0.053	1
11	0	0.025	0.024	0.042	1
12	0	0.027	0.027	0.070	1
13	0.029	0.030	0.029	0.091	1

Gambar 3. Hasil Normalize

### 4. Data Mining

Pada tahap ini penelitian di lakukan dengan menggolongkan Jumlah Peserta BPJS Kesehatan JKN/KIS di Kabupaten Cirebon Pada Tahun 2021 dengan menggunakan Metode *K-Means Clustering* dengan menggunakan *Davies Bouildin Index (DBI)*



Gambar 4. Pengujian Pada Rapidminer

Dengan menerapkan metode *K-Means Clustering* dimana diketahui nilai  $k = 3$ , *max run* sebanyak 100 kali putaran dan dengan jenis *measure typenya* adalah *Euclidean Distance* dan *max optimization steps* berjumlah 100.

Row No.	ID	cluster	NAMA KECAHATAN	NAMA DESA	JUMLAH PESERTA	TUMAH
1	1	cluster_0	0	0	0.169	1
2	2	cluster_2	0	0.002	0.212	1
3	3	cluster_2	0	0.005	0.002	1
4	4	cluster_2	0	0.007	0.008	1
5	5	cluster_2	0	0.010	0.003	1
6	6	cluster_2	0	0.012	0.010	1
7	7	cluster_2	0	0.015	0.009	1
8	8	cluster_2	0	0.017	0.007	1
9	9	cluster_2	0	0.020	0.105	1
10	10	cluster_2	0	0.022	0.063	1
11	11	cluster_2	0	0.029	0.042	1
12	12	cluster_2	0	0.027	0.019	1
13	13	cluster_2	0.026	0.030	0.091	1

Gambar 5. Hasil Clustering

**Cluster Model**

Description

- Cluster 0: 86 Items
- Cluster 1: 96 Items
- Cluster 2: 170 Items
- Total number of items: 412

Gambar 6. Cluster Model

Terlihat bahwa model *cluster* yang dihasilkan meliputi *cluster* 0 dengan 86 Items Jumlah Peserta BPJS Kesehatan JKN/KIS di Kabupaten Cirebon, *Cluster* 1 dengan 156 Items Jumlah Peserta BPJS Kesehatan JKN/KIS di Kabupaten Cirebon, *Cluster* 2 dengan 170 items Jumlah Peserta BPJS Kesehatan JKN/KIS di Kabupaten Cirebon.

Attribute	cluster_0	cluster_1	cluster_2
NAMA KECAHATAN	0.883	0.878	0.183
NAMA DESA	0.887	0.600	0.205
JUMLAH PESERTA	0.268	0.005	0.181
TUMAH	1	1	1

Gambar 7. Centroid Table

Gambar 7 menampilkan nilai *centroid* masing-masing atribut pada tiap *cluster* yang ada. Nilai tersebut yang menjadi acuan pada perhitungan di setiap *dataset* dengan cara mengukur kedekatan nilai masing-masing *cluster*.

**Davies Bouldin**

Davies Bouldin: 0.164

Gambar 8. Nilai Davies Bouildin Index

Gambar 8 diatas merupakan hasil nilai Optimum pada *Davies Bouildin Index* yaitu 0.164.

Tabel 2. Hasil Percobaan dengan Nilai Parameter *Max Run/Iterations* =10

No	Nilai K	Avg. within centroid distance	Davies Bouldin Index (DBI)	Cluster
1	2	0.018	0.172	Cluster 0: 204 items Cluster 1: 208 items
2	3	0.011	0.164	Cluster 0 : 86 items Cluster 1: 156 items Cluster 2: 170 items
3	4	0.008	0.200	Cluster 0: 89 items Cluster 1: 86 items Cluster 2: 84 items Cluster 3: 153 items
4	5	0.006	0.181	Cluster 0: 79 items Cluster 1: 91 items Cluster 2: 57 items

				Cluster 3: 90 items Cluster 4: 95 items
5	6	0.020	0.740	Cluster 0: 88 items Cluster 1: 98 items Cluster 2: 22 items Cluster 3: 57 items Cluster 4: 62 items Cluster 5: 85 items
6	7	0.005	0.200	Cluster 0: 56 items Cluster 1: 62 items Cluster 2: 76 items Cluster 3: 68 items Cluster 4: 41 items Cluster 5: 27 items Cluster 6: 82 items
7	8	0.004	0.207	Cluster 0: 27 items Cluster 1: 50 items Cluster 2: 58 items Cluster 3: 11 items Cluster 4: 80 items Cluster 5: 46 items Cluster 6: 79 items Cluster 7: 61 items
8	9	0.003	0.217	Cluster 0: 50 items Cluster 1: 16 items Cluster 2: 83 items Cluster 3: 38 items Cluster 4: 28 items Cluster 5: 79 items Cluster 6: 13 items Cluster 7: 61 items Cluster 8: 44 items

9	10	0.003	0.199	Cluster 0: 34 items Cluster 1: 28 items Cluster 2: 54 items Cluster 3: 80 items Cluster 4: 55 items Cluster 5: 43 items Cluster 6: 24 items Cluster 7: 75 items Cluster 8: 17 items Cluster 9: 2 items
---	----	-------	-------	---

Dari hasil penelitian yang telah di lakukan dengan menggunakan nilai k=2 sampai k=10 untuk mengetahui nilai *Avg.within centroid distance*, *Davies Bouildin Index (DBI)*, dan *Cluster* pada *Rapidminer* untuk memperoleh k optimum maka di hasil dari penelitian di peroleh sebagai berikut :

K=2 dengan nilai *Avg.within centroid distance* yaitu 0.018 serta nilai *Davies Bouildin Index (DBI)* yang di hasilkan yaitu 0.172 dan menghasilkan 2 *Cluster* yaitu = *Cluster 0* : 204 items dan *Cluster 1*= 208 items.

K=3 dengan nilai *Avg.within centroid distance* yaitu 0.011 serta nilai *Davies Bouildin Index (DBI)* yang di hasilkan yaitu 0.164 dan menghasilkan 3 *Cluster* yaitu = *Cluster 0* = 86 items, *Cluster 1*= 156 items dan *Cluster 2*=170 items.

K=4 dengan nilai *Avg.within centroid distance* yaitu 0.008 serta nilai *Davies Bouildin Index (DBI)* yang di hasilkan yaitu 0.200 dan menghasilkan 4 *Cluster* yaitu = *Cluster 0*: 89 items, *Cluster 1*: 86 items *Cluster 2*: 84 items, dan *Cluster 3*: 153 items.

K=5 dengan nilai *Avg.within centroid distance* yaitu 0.008 serta nilai *Davies Bouildin Index (DBI)* yang di hasilkan yaitu

0.181 dan menghasilkan 5 Cluster yaitu = Cluster 0: 79 items, Cluster 1: 86 items, Cluster 2: 84 items dan Cluster 4: 153 items.

k=6 dengan nilai *Avg.within centroid distance* yaitu 0.020 serta nilai *Davies Bouildin Index (DBI)* yang di hasilkan yaitu 0.740 dan menghasilkan 6 Cluster yaitu = Cluster 0: 88 items, Cluster 1: 98 items, Cluster 2: 22 items, Cluster 3: 57 items, Cluster 4: 62 items, dan Cluster 5: 85 items.

k=7 dengan nilai *Avg.within centroid distance* yaitu 0.005 serta nilai *Davies Bouildin Index (DBI)* yang di hasilkan yaitu 0.200 dan menghasilkan 7 Cluster yaitu = Cluster 0: 56 items, Cluster 1: 62 items, Cluster 2: 76 items, Cluster 3: 68 items, Cluster 4: 41 items, Cluster 5: 27 items dan Cluster 6: 82 items.

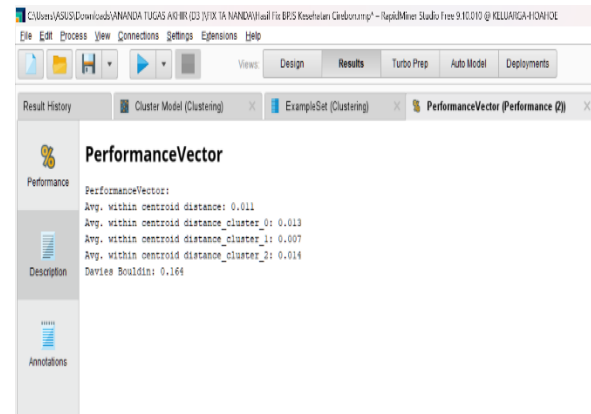
k=8 dengan nilai *Avg.within centroid distance* yaitu 0.004 serta nilai *Davies Bouildin Index (DBI)* yang di hasilkan yaitu 0.207 dan menghasilkan 8 Cluster yaitu = Cluster 0: 27 items, Cluster 1: 50 items, Cluster 2: 58 items, Cluster 3: 11 items, Cluster 4: 80 items, Cluster 5: 46 items, Cluster 6: 79 items, dan Cluster 7: 61 items.

k=9 dengan nilai *Avg.within centroid distance* yaitu 0.003 serta nilai *Davies Bouildin Index (DBI)* yang di hasilkan yaitu 0.217 dan menghasilkan 9 Cluster yaitu = Cluster 0: 50 items, Cluster 1: 16 items, Cluster 2: 83 items, Cluster 3: 38 items, Cluster 4: 28 items, Cluster 5: 79 items, Cluster 6: 13 items, Cluster 7: 61 items dan Cluster 8: 44 items.

k=10 dengan nilai *Avg.within centroid distance* yaitu 0.003 serta nilai *Davies Bouildin Index (DBI)* yang di hasilkan yaitu 0.199 dan menghasilkan 10 Cluster yaitu = Cluster 0: 34 items, Cluster 1: 28 items, Cluster 2: 54 items, Cluster 3: 80 items, Cluster 4: 55 items, Cluster 5: 43 items, Cluster 6: 24 items, Cluster 7: 75 items, Cluster 8: 17 items, dan Cluster 9: 2 items.

Dari hasil 10 kali percobaan dengan nilai parameter run max mulai dari k=2 sampai k=10 maka nilai optimum yang di peroleh terdapat pada k=3 dengan nilai *Avg.within centroid distance* yaitu 0.011 serta nilai *Davies Bouildin Index (DBI)* yang di hasilkan yaitu 0.164 dan menghasilkan 3 Cluster yaitu = Cluster 0 = 86 items, Cluster 1= 156 items dan Cluster 2=170 items.

### 5. Interpretasi/Evaluation



**Gambar 8.** Deskripsi Performance Vector

Dari beberapa percobaan terhadap nilai k= 2 sampai k=10 di peroleh nilai *Davies Bouildin* yang optimum sebesar 0.164 pada nilai k = 3 pada algoritma *K-Means*, semakin kecil nilai indeksnya maka akan semakin baik hasil pengelompokannya. Untuk mengetahui anggota kelompok dari k, maka dapat dilakukan perhitungan selisih dari rata-rata *centroid distance* seperti tampak pada tabel berikut ini:

**Tabel 3.** Perhitungan Selisih Centroid Distance pada *K-Means Clustering*

<i>K-Means</i>	0	1	2
<i>Avg.within centroid distance</i>	0.011	0.011	0.011
<i>Avg.within centroid distance cluster</i>	0.013	0.007	0.014
Selisih	-0.002	0.004	-0.003

Dari hasil perhitungan selisih *centroid distante* yang dihitung dengan cara hasil dari



Avg.within *centroid distance* dikurangi dengan hasil Avg.within *centroid distance cluster* sehingga menghasilkan selisih. Pada *cluster 0* didapatkan selisih -0,002, pada *cluster 1* didapatkan selisih 0.004 dan pada *cluster 2* didapatkan selisih -0.003.

**Tabel 4.** Anggota *Cluster* Berdasarkan Jumlah Peserta BPJS Kesehatan Jkn/Kis Di Kabupaten Cirebon

Cluster	Jumlah	Tingkatan
0	86 Items	Rendah
1	156 Items	Sedang
2	170 Items	Banyak
Total	412 Items	

Dari hasil penelitian yang telah dilakukan maka di peroleh *cluster 0* dengan jumlah 86 items termasuk dalam tingkatan rendah, *cluster 1* dengan jumlah 156 items termasuk dalam tingkatan sedang, dan *cluster 2* dengan jumlah 170 items termasuk dalam tingkatan banyak.

## 5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan oleh penulis maka dapat disimpulkan bahwa:

1. Penelitian yang di lakukan terhadap data Jumlah Peserta BPJS Kesehatan JKN/KIS di Kabupaten Cirebon menggunakan metode *K-Means Clustering* dengan *Tools RapidMiner* maka hasil penggelompokkan menghasilkan *cluster* yang optimal dengan nilai DBI = 0.0655 dengan nilai K=3
2. Penelitian yang dilakukan terhadap data Jumlah Peserta BPJS Kesehatan JKN/KIS di Kabupaten Cirebon menggunakan metode *K-Means Clustering* dengan *Tools RapidMiner* maka hasil penggelompokkan menghasilkan 3 *Cluster* yaitu *Cluster 0* =86 *Items* termasuk jumlah peserta dengan tingkatan rendah, *Cluster 1*= 156 *Items* termasuk jumlah peserta BPJS Kesehatan dengan tingkatan sedang dan *Cluster 2*=170 *Items* termasuk jumlah peserta BPJS

Kesehatan dengan tingkatan terbanyak dengan Total keseluruhan 412 *Dataset*

## 6. DAFTAR PUSTAKA

- [1] K. I. L. Dewi, N. N. Yulianthini, and N. L. W. S. Telagawathi, “Pengaruh Dimensi Kualitas Pelayanan Terhadap Kepuasan Pelanggan Pengguna Bpjs Kesehatan Di Kota Singaraja,” *Bisma: Jurnal Manajemen*, vol. 5, no. 2, pp. 82–92, 2019, Accessed: May 08, 2023. [Online]. Available: <https://ejournal.undiksha.ac.id/index.php/BISMA-JM/article/view/22011>
- [2] I. P. Sari, H. Sawiji, and T. Murwaningsih, “Analisis Kepuasan Pelanggan di Kantor BPJS Kesehatan Karanganyar,” *JIKAP (Jurnal Informasi Dan Komunikasi Administrasi Perkantoran)*, vol. 3, no. 4, pp. 10–18, 2019, Accessed: May 08, 2023. [Online]. Available: <https://jurnal.uns.ac.id/JIKAP/article/view/35467>
- [3] N. Tou and P. M. Endraswari, “IMPLEMENTASI DATA MINING DALAM KLASIFIKASI HASIL DIAGNOSA PASIEN BPJS MENGGUNAKAN ALGORITMA CART,” *JIKA (Jurnal Informatika)*, vol. 6, no. 2, p. 170, Jun. 2022, doi: 10.31000/jika.v6i2.6164.
- [4] Z. Zema, Y. Maulita, and L. A. N. Kadim, “PENERAPAN DATA MINING PENGELOMPOKAN PESERTA BPJS KETENAGAKERJAAN BERDASARKAN PROGRAM YANG DIAMBIL MENGGUNAKAN METODE CLUSTERING,” *Jurnal Sistem Informasi Kaputama (JSIK)*, vol. 6, no. 2, pp. 152–164, 2022, Accessed: May 08, 2023. [Online]. Available: <http://jurnal.kaputama.ac.id/index.php/JSIK/article/view/839>

- [5] R. S. Dewi, “Perlindungan Hukum terhadap Peserta BPJS Kesehatan dalam Pelayanan kesehatan di rumah sakit,” *Focus hukum upmi*, vol. 1, no. 2, pp. 74–79, 2020, Accessed: May 08, 2023. [Online]. Available: <https://www.ejournal.warmadewa.ac.id/index.php/public-inspiration/article/view/801>
- [6] I. Widiastuti, “Pelayanan Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan di Jawa Barat,” *Public Inspiration: Jurnal Administrasi Publik*, vol. 2, no. 2, pp. 91–101, 2017.
- [7] E. K. Astuti, “Peran BPJS Kesehatan Dalam Mewujudkan Hak Atas Pelayanan Kesehatan Bagi Warga Negara Indonesia,” *JPeHI: Jurnal Penelitian Hukum Indonesia*, vol. 1, no. 1, 2020.
- [8] S. Salvan N, W. Hadikristanto, and E. Edora, “Klasterisasi Data Penggunaan Layanan BPJS Kesehatan Menggunakan Algoritma K-Means,” *Prosiding Sains dan Teknologi*, vol. 1, no. 1, pp. 121–127, 2022.
- [9] F. Ariyanti and M. T. Gifari, “Analisis Persetujuan Klaim BPJS Kesehatan pada Pasien Rawat Inap,” *Jurnal Ilmu Kesehatan Masyarakat*, vol. 8, no. 04, pp. 156–166, Dec. 2019, doi: 10.33221/jikm.v8i04.415.
- [10] R. Silfiani, “Peran Demokrasi Dalam Pelayanan Kesehatan Bagi Masyarakat Pengguna BPJS Kesehatan,” 2021.
- [11] S. Suprpto and A. A. Malik, “Implementasi Kebijakan Diskresi Pada Pelayanan Kesehatan Badan Penyelenggara Jaminan Kesehatan (BPJS),” *Jurnal Ilmiah Kesehatan Sandi Husada*, vol. 8, no. 1, pp. 1–8, Jun. 2019, doi: 10.35816/jiskh.v8i1.62.
- [12] R. W. Nasution, S. Suhada, I. O. Kirana, I. Gunawan, and I. P. Sari, “Penerapan Data Mining Untuk Pengelompokan Minat Konsumen Terhadap Pengguna Jasa Pengiriman Pada PT. Jalur Nugraha Ekakurir (JNE) Pematangsiantar,” *Resolusi: Rekayasa Teknik Informatika dan Informasi*, vol. 1, no. 4, pp. 274–281, 2021.
- [13] S. M. Hutabarat and A. Sindar, “Data Mining Penjualan Suku Cadang Sepeda Motor Menggunakan Algoritma K-Means,” *Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)*, vol. 2, no. 2, p. 126, Oct. 2019, doi: 10.32672/jnkti.v2i2.1555.
- [14] F. M. Basysyar, A. Y. Wijayya, I. Ali, and S. Anwar, “Clustering Data Disabilitas menggunakan Algoritma K-Means di Kabupaten Cirebon,” *JURSIMA (Jurnal Sistem Informasi dan Manajemen)*, vol. 9, no. 3, pp. 247–255, 2021, Accessed: May 07, 2023. [Online]. Available: <https://ejournal.stmikgici.ac.id/index.php/jursima/article/view/305>
- [15] R. A. Indraputra and R. Fitriana, “K-Means clustering data COVID-19,” *Jurnal Teknik Industri*, vol. 10, no. 3, pp. 275–282, 2020.
- [16] V. Herlinda, D. Darwis, and D. Dartono, “Analisis Clustering Untuk Recredesialing Fasilitas Kesehatan Menggunakan Metode Fuzzy C-Means,” *Jurnal Teknologi Dan Sistem Informasi*, vol. 2, no. 2, pp. 94–99, 2021.
- [17] H. Priyatman, F. Sajid, and D. Haldivany, “Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa,” *Jurnal*

*Edukasi Dan Penelitian Informatika (JEPIN)*, vol. 5, no. 1, p. 62, 2019.

- [18] M. Iqbal, “Klasterisasi data jamaah umroh pada auliya tour & travel menggunakan metode k-means clustering,” *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 5, no. 2, pp. 97–104, 2019.
- [19] M. Faid, M. Jasri, and T. Rahmawati, “Perbandingan Kinerja Tool Data Mining Weka dan Rapidminer Dalam Algoritma Klasifikasi,” *Teknika*, vol. 8, no. 1, pp. 11–16, 2019.
- [20] Y. R. Sari, A. Sudewa, D. A. Lestari, and T. I. Jaya, “Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer,” *CESS (Journal of Computer Engineering, System and Science)*, vol. 5, no. 2, p. 192, Jul. 2020, doi: 10.24114/cess.v5i2.18519.