



Penerapan Metode Clustering Dalam Upaya Pencegahan Penyakit Demam Berdarah Menggunakan Algoritma K-Means (Studi Kasus: Kota Tasikmalaya)

Muhamad Rivalda¹, Erti Maulina Hidayat², Muhamad Azhar Gunawan³, Diaz Defriyanto⁴

¹Universitas Bina Sarana Informatika
e-mail: rivaldamuhamad96@gmail.com

²Universitas Bina Sarana Informatika
e-mail: ertymaulina30@gmail.com

³Universitas Bina Sarana Informatika
e-mail: azharg@gmail.com

⁴Universitas Bina Sarana Informatika
e-mail: diaz.defriyanto@yahoo.com

Abstrak-- Penyakit DBD merupakan penyakit yang menular yang ditularkan oleh nyamuk *Aedes Aegypti*, dan penyakit ini terjadi terus menerus di sepanjang tahun, menimbulkan wabah dan kematian. Di Indonesia, kesadaran masyarakat dalam menjaga kebersihan dan kurangnya antisipasi masyarakat yang terjangkit DBD, sehingga penyakit DBD termasuk jenis penyakit yang mudah terjangkit pada semua kalangan usia. Metodologi penelitian yang digunakan pada penelitian ini menggunakan metode clustering. Tujuannya adalah melakukan klastering dan evaluasi model algoritma k-means untuk mengetahui akurasi algoritma dalam mengklastering penyakit DBD. K-Means Clustering, K yang berarti konstanta untuk jumlah cluster yang diinginkan, sedangkan Means yang berarti rata-rata kelompok data didalam cluster. Oleh karena itu, penelitian yang menggunakan Algoritma K-Means ini akan mengelompokkan daerah-daerah yang ada di Kota Tasikmalaya sesuai tingkat terjadinya kasus penyakit DBD agar dengan tepat dan cepat sasaran dalam upaya pencegahan penyakit DBD. Rumus yang digunakan pada algoritma K-Means ada 4 tahapan yaitu : Menentukan jumlah cluster, menghitung jarak, mengelompokkan data, serta menghitung pusat cluster. Manfaat dari penelitian ini adalah bisa mempercepat proses upaya pencegahan dalam penyakit DBD, mengetahui upaya pencegahan penyakit DBD secara cepat dan tepat, sistem klastering untuk upaya pencegahan dan dapat mengurangi angka kematian, serta menambah wawasan bagi pembaca yang ingin mempelajari tentang klasterisasi serta algoritma K-Means.

Kata Kunci: Demam Berdarah, Clustering, K-Means, Pencegahan

Abstract – Dengue fever is a contagious disease transmitted by the *Aedes Aegypti* mosquito, and this disease occurs continuously throughout the year, causing outbreaks and deaths. In Indonesia, public awareness in maintaining cleanliness and lack of anticipation of people infected with Dengue Fever, so Dengue Fever is a disease easily infected for all ages. The research methodology used in this study used the clustering method. The goal is to cluster and evaluate the k-means algorithm model to determine the algorithm's accuracy in classifying Dengue Fever disease. K-Means Clustering, K, which means a constant for the desired number of clusters, while Means which means the average of the data groups in the cluster. Therefore, research using the K-Means Algorithm will group the areas in Tasikmalaya City according to the rate of occurrence of dengue cases so that they are precisely and quickly targeted in efforts to prevent Dengue Fever. The formula used in the K-Means algorithm has four stages, namely: Determining the number of clusters, calculating distances, grouping data, and calculating the center of the cluster. The benefits of this research are that it can speed up the process of prevention efforts in Dengue Fever,

* Korenpondensi.

Alamat E-mail : jurnal.larik@bsi.ac.id.

Diterima 30 July 2022; Direvisi 15 Desember 2022; Diterima 1 Juli 2023

© 2023 Jurnal Larik.

know efforts to prevent Dengue fever quickly and precisely, clustering systems for prevention efforts and can reduce mortality, and add insight for readers who want to learn about clustering and K-Means algorithms.

Keywords: *Dengue fever, clustering, k-means, prevention*

I. PENDAHULUAN

Penyakit Demam Berdarah Dengue (DBD) merupakan salah satu penyakit menular yang banyak menjangkiti penduduk yang hidup di daerah sub tropis dan tropis. Penyakit DBD termasuk dalam salah satu penyakit menular yang ditularkan oleh nyamuk *Aedes aegypti*. Diperkirakan 390 juta kasus infeksi dengue setiap tahunnya di dunia. Lebih dari 100 negara mejadi daerah endemis DBD salah satunya Indonesia. Pada tahun 2012 Indonesia dicatat sebagai Negara dengan kasus DBD tertinggi di dunia berdasarkan catatan World Health Organization (WHO) [1].

Kasus demam berdarah dengue di negara Indonesia masuk dalam status kejadian luar biasa yang ditandai dengan meningkatnya penyebaran demam berdarah yang menimbulkan kematian. Hal ini didasari karena situasi lingkungan yang semerawut dan penataan kebersihan yang tidak mumpuni pada musim hujan mengakibatkan munculnya beribu-ribu jentik nyamuk [2].

Semakin meningkatnya angka kematian bagi para penderita penyakit demam berdarah, dikarenakan banyak masyarakat yang belum mengenal sejak dini ciri-ciri penyakit DBD, kurangnya kesadaran masyarakat dalam menjaga kebersihan dan kurang antisipasi masyarakat yang terkena DBD, sehingga penyakit demam berdarah dengue merupakan jenis penyakit yang mudah terjangkau di semua kalangan usia [3].

II. METODE PENELITIAN

Dalam Penelitian ini metode yang digunakan yaitu metode clustering. Tujuandari penelitian ini adalah melakukan kalsifikasi dan evaluasi medel algoritma k-means untuk mengetahui akurasi algoritma k-means dalam mengklasifikasi penyakit Demam Berdarah Dengue (DBD).

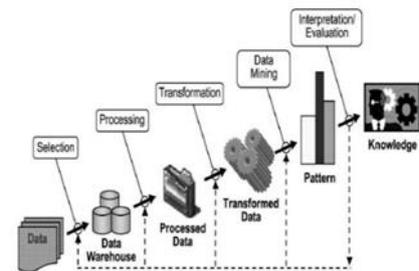
2.1 Data Mining

Data mining adalah suatu proses menemukan korelasi baru yang bermanfaat, pola dan trend dengan menambang sejumlah repository data dalam jumlah besar, menggunakan teknologi

pengenalan pola seperti statistik dan teknik matematika. Data mining merupakan suatu bidang dari seluruh bidang keilmuan yang tekniknya itu hasil penyatuan dari teknik pembelajaran mesin, identifikasi pola, statistik, basis data, dan pengenalan visualisasi untuk pengambilan permasalahan setiap informasi dari database yang besar [4].

2.1.1 Tahap Data Mining

Tahapan yang dilakukan pada proses data mining diawali dari seleksi data dari data sumber ke data target, tahap pre-processing untuk memperbaiki kualitas data, transformasi, data mining serta tahap interpretasi dan evaluasi yang menghasilkan output berupa pengetahuan baru yang diharapkan memberikan kontribusi yang lebih baik. Secara detail dijelaskan sebagai berikut [5] :



1. Data selection

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. Pre-processing /cleaning

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses Cleaning pada data yang menjadi fokus KDD. Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data.

3. Transformation

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. Data mining

Data mining adalah proses 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 sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. Interpretation/evaluation

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 ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

2.2 Clustering

Metode Clustering adalah mempartisi data ke dalam kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan kedalam satu cluster yang sama. Tujuan dari clustering ini adalah untuk meminimalisir fungsi tujuan yang diterapkan dalam proses clustering. Yang umumnya berusaha meminimalisir variasi dalam suatu cluster dan memaksimalkan variasi antar cluster [6]. Dalam penelitian ini metode utama yang digunakan dalam pengolahan data upaya pencegahan demam berdarah di Kecamatan Cipedes Kota Tasikmalaya adalah clustering.

2.3 Algoritma K-Means

Algoritma K-Means merupakan algoritma klusterisasi yang mengelompokkan data berdasarkan titik pusat kluster (centroid) terdekat dengan data. Tujuan K-Means adalah mengelompokkan data dengan memaksimalkan kemiripan data dalam satu kluster dan meminimalkan kemiripan data antara kluster. ukuran kemiripan yang digunakan dalam kluster adalah fungsi jarak. sehingga pemaksimalan kemiripan data didapatkan berdasarkan jarak terpendek antara data titik centroid. (Adiputra 2022) Tahapan algoritma k-means adalah sebagai berikut :

1. Pilih secara acak k buah data sebagai pusat cluster.
2. Jarak antara data dan pusat cluster dihitung menggunakan Euclidian Distance. 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 atribut data ke k

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

3. Data ditempatkan dalam cluster yang terdekat, dihitung dari tengah cluster.
4. Pusat cluster baru akan ditentukan bila semua data telah ditetapkan dalam cluster terdekat.
5. Proses penentuan pusat cluster dan penempatan data dalam cluster diulangi sampai nilai centroid tidak berubah lagi.

III. HASIL DAN PEMBAHASAN

3.1 Hasil dan Pembahasan

Pada bab ini akan membahas mengenai pengolahan data yang akan digunakan dalam penelitian, data tersebut akan diproses menggunakan algoritma K-Means dengan metode Clustering. Pada penelitian ini, data yang digunakan adalah Data Kasus Jumlah DBD di Kota Tasikmalaya. Data tersebut didapat dari Open Data Kota Tasikmalaya yaitu Dinkes - Jumlah Kasus Demam Berdarah Dengue (DBD) Tahun 2021 data set. Variabel yang digunakan pada penelitian ini ialah sebanyak 11 atribut dengan jumlah data sebanyak 22 record yang akan dilakukan pengujian.

JUMLAH KASUS DEMAM BERDARAH DENGUE (DBD) TAHUN 2021

© 13 Mei 2022 1 Min Read

No	Rekamans	Peskesmas	Jumlah Laki-Laki	Jumlah Perempuan	Jumlah	Meninggal (L)	Meninggal (P)	Jumlah Meninggal	CR (%)	CR
1	Cibadung	Cibadung	23	34	57	0	3	3	0.0	88.2

Gambar 3.1 Dataset Jumlah Kasus DBD Tahun 2021

Pengujian ini dilakukan menggunakan platform Laptop: Intel Core i7-8570H CPU @ 2.20GHz 2.21GHz, 8 GB RAM, OS Windows 10 Pro 64-bit,

RapidMiner Studio Versi 10.00 sebagai analisis data. Variabel dataset Dinkes - Jumlah Kasus Demam Berdarah Dengue (DBD) Tahun 2021.

Tabel 3.1 Deskripsi Variabel Dataset

No	Atribut	Tipe Data
1	Kecamatan	Polynomial
2	Puskesmas	Polynomial
3	Jumlah Laki-laki	Integer
4	Jumlah Perempuan	Integer
5	Jumlah	Integer
6	Meninggal (L)	Integer
7	Meninggal (P)	Integer
8	Jumlah Meninggal	Integer
9	CFR % (L)	Integer
10	CFR % (P)	Integer
11	Jumlah CFR %	Integer

3.2 Preprocessing

Berikut adalah tahapan data preprocessing pada data Mining:

1. Pembersihan data (Data Cleansing)

Pembersihan data ini akan memberikan hasil performa dari sistem data mining dikarenakan data yang digunakan akan berkurang jumlah dan kompleksitasnya. Dalam studi kasus Jumlah Kasus DBD di Kota Tasikmalaya tahun 2021 ini terdapat 0 missing data dari 22 data record.

2. Seleksi Data (Data Selection)

Selanjutnya data diseleksi untuk memilih variabel apa saja yang akan diambil agar tidak terjadi kesamaan dan perulangan yang tidak diperlukan dalam pengolahan teknik data mining. Dalam pengerjaan pada studi kasus dataset Jumlah kasus DBD di Kota Tasikmalaya tahun 2021 tidak terdapat noise atau kesalahan data.

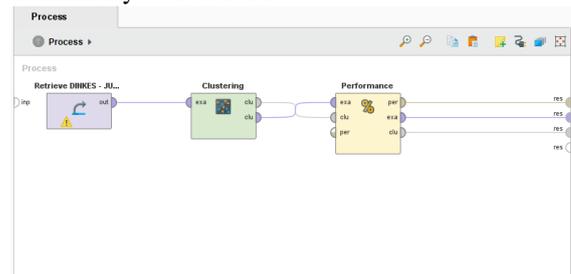
3.3 Penerapan Data Mining

3.3.1 Penerapan Algoritma K-Means untuk akurasi jumlah DBD Kota Tasikmalaya Tahun 2021

Pengujian data set dari jumlah data 22record dengan 11 atribut.

3.3.2 Proses Penerapan Algoritma K-Means

Berikut merupakan proses algoritma K-Means untuk mengetahui hasil pengujian algoritma dengan dataset public Jumlah Kasus DBD Kota Tasikmalaya tahun 2021.



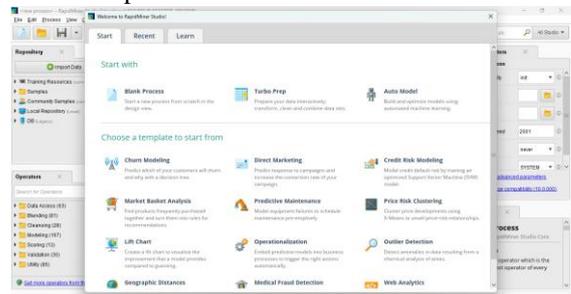
Gambar 3.2 Proses Algoritma K-Means

3.4 Langkah – Langkah proses di Rapi Miner

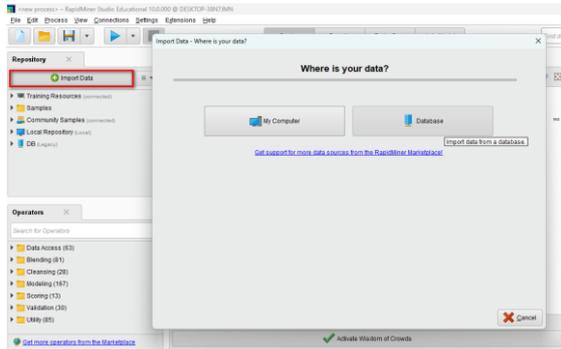
1. Buka Aplikasi Rapid Miner



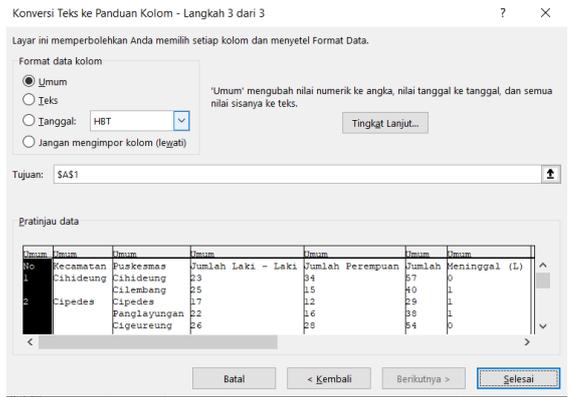
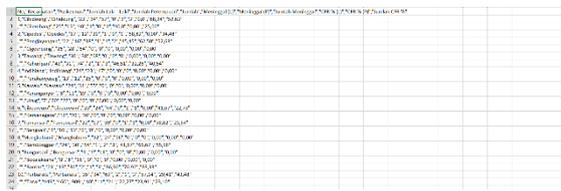
2. Lalu pilih *Blank Process*



3. Lalu Pilih *Import Data* atau *Retrieve Data* yang berbentuk excel untuk data yang ingin dimasukkan dan sebelumnya sudah diunduh pada website Open Data Kota Tasikmalaya.

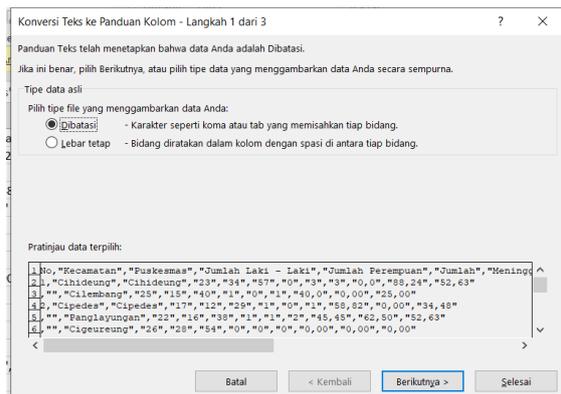


4. Jika data excel tersebut masih belum rapih seperti pada format csv, maka lakukanlah *cleansing* terlebih dahulu dengan cara mengubah teks menjadi kolom sebagai berikut:

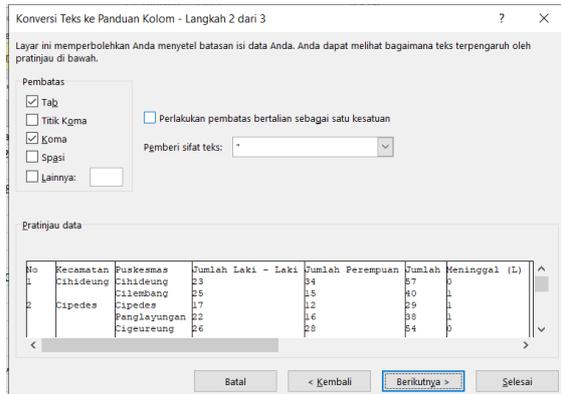


No	Kecamatan	Puskesmas	Jumlah Lal	Jumlah Pei	Jumlah Meninggal	Meninggal	Jumlah Me	CFR % (L)	CFR % (P)	Jumlah CFR %
1	Cihideung	Cihideung	23	34	57	0	3	0	88,24	52,63
2	Cihideung	Cihideung	25	15	40	1	0	1	40	0
3	Cipedeas	Cipedeas	17	12	29	1	0	1	58,82	34,48
4	Cipedeas	Cipedeas	17	12	29	1	0	1	58,82	34,48
5	Panglayuran	Panglayuran	22	16	38	1	1	2	45,45	62,5
6	Cigeureung	Cigeureung	26	28	54	0	0	0	0	0
7	Tawang	Tawang	30	38	68	0	0	0	0	0
8	Kahuripan	Kahuripan	43	31	74	2	1	3	46,51	32,26
9	Indihiang	Indihiang	24	23	47	0	0	0	0	0
10	Parakanya	Parakanya	13	12	25	0	0	0	0	0
11	Kawalu	Kawalu	24	31	55	0	0	0	0	0
12	Karangany	Karangany	8	11	19	0	0	0	0	0
13	Urug	Urug	7	20	27	0	0	0	0	0
14	Cibeureur	Cibeureur	20	24	44	0	1	1	0	41,67
15	Kersnaga	Kersnaga	16	20	36	0	0	0	0	0
16	Tamansari	Tamansari	22	17	39	0	1	1	0	58,82
17	Sangkal	Sangkal	3	10	13	0	0	0	0	0
18	Mangkubu	Mangkubu	32	24	56	0	0	0	0	0
19	Sambong	Sambong	24	30	54	1	2	3	41,67	66,67
20	Bungarsari	Bungarsari	4	9	13	0	0	0	0	0
21	Sukalaskar	Sukalaskar	8	8	16	0	0	0	0	0
22	Bantar	Bantar	23	13	36	2	1	3	86,96	76,92
23	Purbaratu	Purbaratu	35	34	69	2	1	3	57,14	29,41
24	Total		449	460	909	10	11	21	22,27	23,91

5. Lalu hilangkan kolom “no” agar tidak terjadi penumpukan *row* ketika dimasukan ke aplikasi Rapid Miner



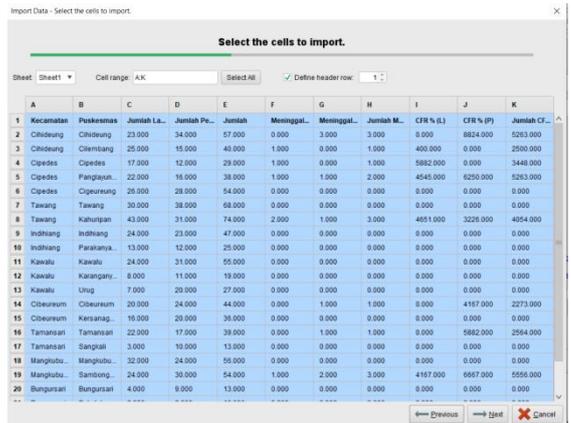
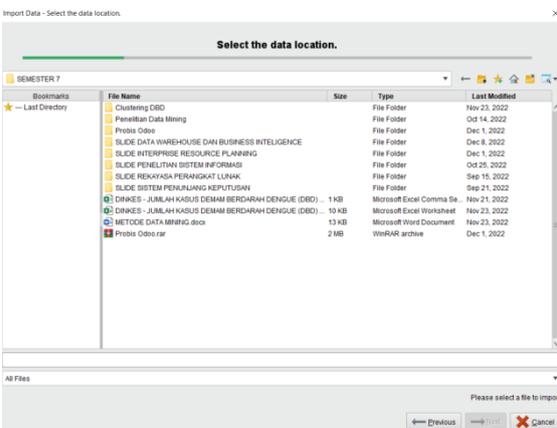
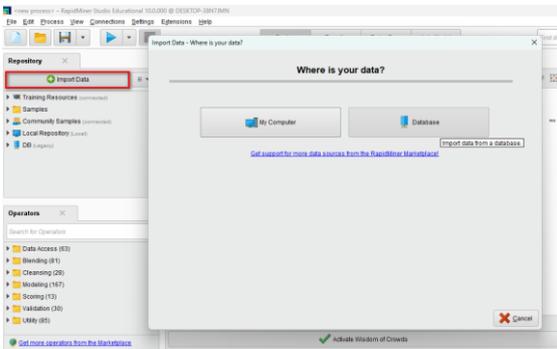
Kecamatan	Puskesmas	Jumlah Lal	Jumlah Pei	Jumlah Meninggal	Meninggal	Jumlah Me	CFR % (L)	CFR % (P)	Jumlah CFR %
Cihideung	Cihideung	23	34	57	0	3	0	88,24	52,63
Cihideung	Cihideung	25	15	40	1	0	1	40	0
Cipedeas	Cipedeas	17	12	29	1	0	1	58,82	34,48
Cipedeas	Cipedeas	17	12	29	1	0	1	58,82	34,48
Panglayuran	Panglayuran	22	16	38	1	1	2	45,45	62,5
Cigeureung	Cigeureung	26	28	54	0	0	0	0	0
Tawang	Tawang	30	38	68	0	0	0	0	0
Kahuripan	Kahuripan	43	31	74	2	1	3	46,51	32,26
Indihiang	Indihiang	24	23	47	0	0	0	0	0
Parakanya	Parakanya	13	12	25	0	0	0	0	0
Kawalu	Kawalu	24	31	55	0	0	0	0	0
Karangany	Karangany	8	11	19	0	0	0	0	0
Urug	Urug	7	20	27	0	0	0	0	0
Cibeureur	Cibeureur	20	24	44	0	1	1	0	41,67
Kersnaga	Kersnaga	16	20	36	0	0	0	0	0
Tamansari	Tamansari	22	17	39	0	1	1	0	58,82
Sangkal	Sangkal	3	10	13	0	0	0	0	0
Mangkubu	Mangkubu	32	24	56	0	0	0	0	0
Sambong	Sambong	24	30	54	1	2	3	41,67	66,67
Bungarsari	Bungarsari	4	9	13	0	0	0	0	0
Sukalaskar	Sukalaskar	8	8	16	0	0	0	0	0
Bantar	Bantar	23	13	36	2	1	3	86,96	76,92
Purbaratu	Purbaratu	35	34	69	2	1	3	57,14	29,41
Total		449	460	909	10	11	21	22,27	23,91



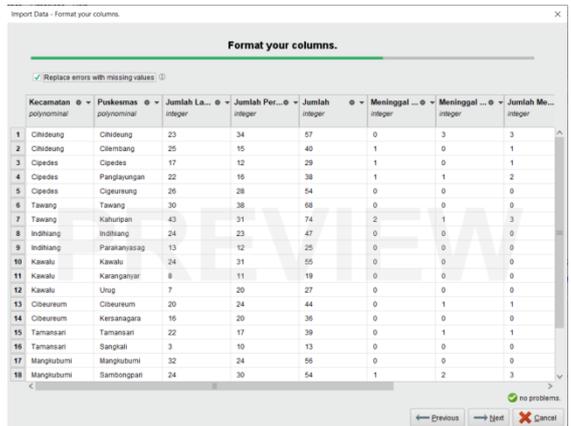
6. Selanjutnya hapus kembali kolom “TOTAL” agar ketika Rapid Miner menampilkan visualisasi data tidak terjadi ambiguitas visual data yang membingungkan

	Kecamatan	Puskesmas	Jumlah L	Jumlah P	Jumlah M	Meninggal	Jumlah Me	CFR % (L)	CFR % (P)	Jumlah CFR %	
1	Cibeureum	Cibeureum	23	34	57	0	3	3	0	88,24	52,63
2	Cibeureum	Cibeureum	25	15	40	1	0	1	40	0	25
3	Cibeureum	Cibeureum	17	12	29	1	0	1	58,82	0	34,48
4	Cibeureum	Cibeureum	22	16	38	1	1	2	45,45	62,5	52,63
5	Cibeureum	Cibeureum	26	28	54	0	0	0	0	0	0
6	Cibeureum	Cibeureum	30	38	68	0	0	0	0	0	0
7	Cibeureum	Cibeureum	43	31	74	2	1	3	46,51	32,26	40,54
8	Cibeureum	Cibeureum	24	23	47	0	0	0	0	0	0
9	Cibeureum	Cibeureum	13	12	25	0	0	0	0	0	0
10	Cibeureum	Cibeureum	24	31	55	0	0	0	0	0	0
11	Cibeureum	Cibeureum	8	11	19	0	0	0	0	0	0
12	Cibeureum	Cibeureum	7	20	27	0	0	0	0	0	0
13	Cibeureum	Cibeureum	20	24	44	0	1	1	0	41,67	22,73
14	Cibeureum	Cibeureum	16	20	36	0	0	0	0	0	0
15	Cibeureum	Cibeureum	22	17	39	0	1	1	0	58,82	25,64
16	Cibeureum	Cibeureum	3	10	13	0	0	0	0	0	0
17	Cibeureum	Cibeureum	32	24	56	0	0	0	0	0	0
18	Cibeureum	Cibeureum	24	30	54	1	2	3	41,67	66,67	55,56
19	Cibeureum	Cibeureum	4	9	13	0	0	0	0	0	0
20	Cibeureum	Cibeureum	8	8	16	0	0	0	0	0	0
21	Cibeureum	Cibeureum	23	13	36	2	1	3	86,96	76,92	83,33
22	Cibeureum	Cibeureum	35	34	69	2	1	3	57,14	29,41	43,48

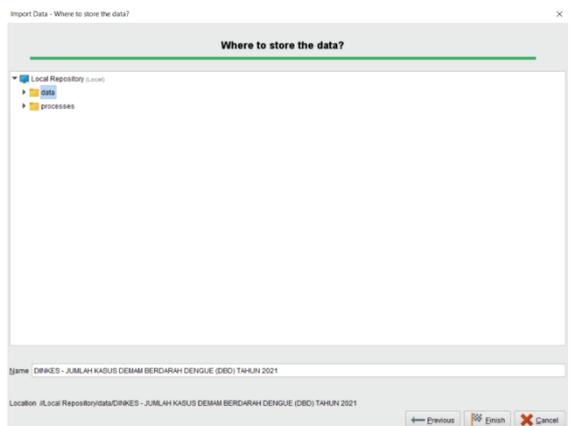
7. Setelah proses *cleansing* selesai, langkah selanjutnya yaitu retrieve data tersebut ke Rapid Miner



8. Centang pilihan *replace error with missing values*, dan ubah tipe datanya jika masih ada yang belum sesuai

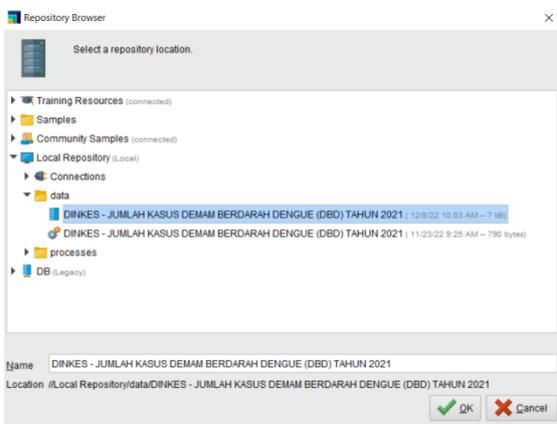
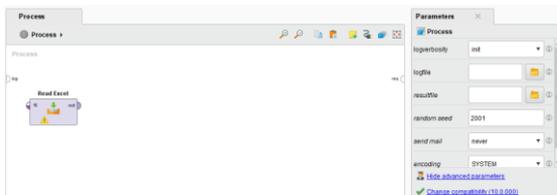
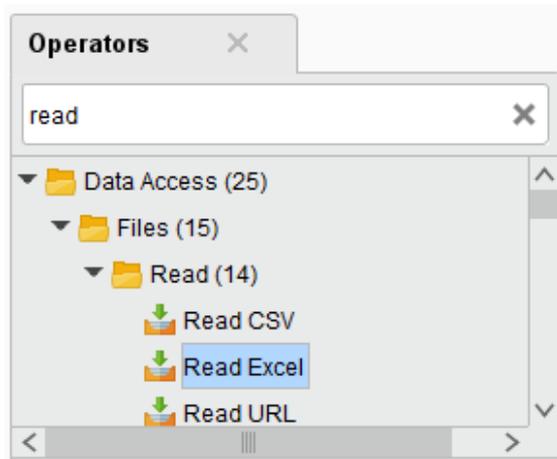


9. Lalu pilih penyimpanan pada *Local Repository* dan ubah namanya jika diperlukan



10. Setelah proses untuk *import* data selesai. Lalu kita masukan operator *read excel* dengan file

yang sudah kita masukkan sebelumnya pada local repository data



11. Sebelum ke proses selanjutnya, kita cek data excel tersebut dengan menghubungkan out pada read excel ke result yang ada di sebelah pojok kanan



12. Setelah itu kita bisa melihat *statistic* data DBD excel. Langkah selanjutnya ialah mengecek apakah ada *missing data* yang terdapat pada data tersebut, jika tidak ada maka data tersebut aman dan dapat dilakukan proses *clustering data*

Name	Type	Missing	Statistics	Filter (11 / 11 attributes)
Kecamatan	Nominal	0	Most Purwaratu (1)	Most Bungursari (3), Cipede...
Puskesmas	Nominal	0	Least Urug (1)	Most Banjar (1), Bungursari (...)
Jumlah Laki - Laki	Integer	0	Min 3	Max 43
Jumlah Perempuan	Integer	0	Min 8	Max 38
Jumlah	Integer	0	Min 13	Max 74
Meninggal (L)	Integer	0	Min 0	Max 2
Meninggal (P)	Integer	0	Min 0	Max 3
Jumlah Meninggal	Integer	0	Min 0	Max 3
CFR % (L)	Integer	0	Min 0	Max 8696

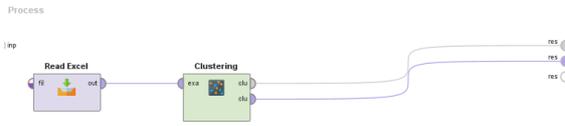
13. Selanjutnya kita tambahkan operator k-means, dan masukkan ke dalam design untuk di proses. Setelah itu hubungkan *out* dari *retrieve* data DBD dengan *exa* pada operator k-means, lalu atur jumlah clustering atau “k” pada menu *parameters* disebelah kiri dan ubah dari 5 cluster menjadi 3 cluster



14. Selanjutnya ubah measure type menjadi *mixed measures* dan otomatis *mixed measures* yang digunakan ialah *euclidean distance*



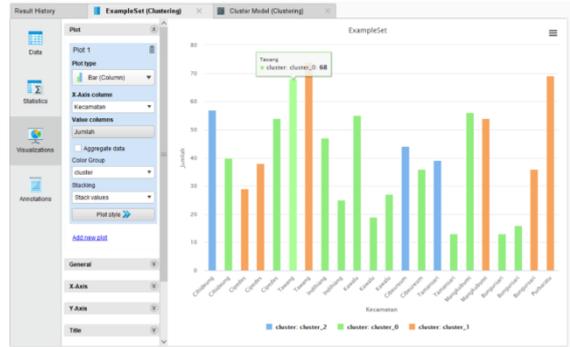
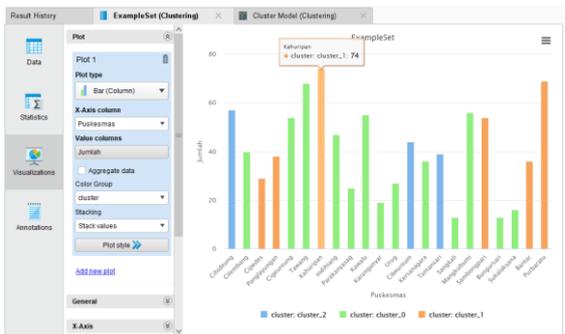
15. Setelah itu hubungkan *clu* pada *result* dan pilih *play* atau menjalankan proses



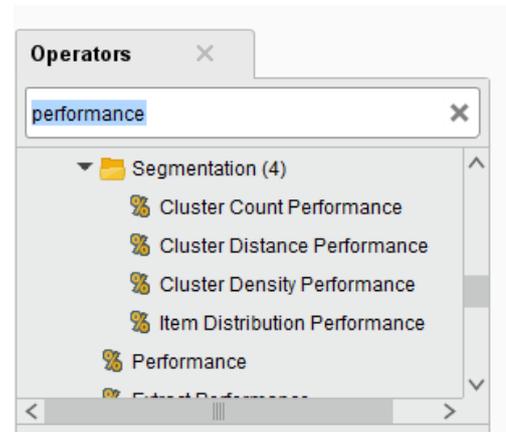
16. Dan kita dapat melihat ada 2 result yaitu *cluster model* dan *example set* clustering, dimana *cluster model* berisi data deskripsi berupa item yang didapat dari proses clustering. Pada *example set* kita dapat melihat data, statistic, dan *visualizations*, yang dapat kita atur menggunakan plot type bar (*column*) untuk melihat jumlah orang yang terjangkau DBD baik laki-laki maupun perempuan pada puskesmas yang ada di tiap Kecamatan pada Kota Tasikmalaya.

Row No.	id	cluster	Kecamatan	Puskesmas	Jumlah Laki.	Jumlah Per.	Jumlah	Meninggal (L)	Meninggal (P)
1	1	cluster_2	Cibadung	Cibadung	23	34	57	0	3
2	2	cluster_3	Cibadung	Cibadung	25	15	40	1	0
3	3	cluster_1	Ciendes	Ciendes	17	12	29	1	0
4	4	cluster_1	Ciendes	Pangajeneung	22	16	38	1	1
5	5	cluster_3	Ciendes	Ciendes	26	28	54	0	0
6	6	cluster_3	Tawang	Tawang	30	38	68	0	0
7	7	cluster_1	Tawang	Kahuripan	43	31	74	2	1
8	8	cluster_3	Imahing	Imahing	24	23	47	0	0
9	9	cluster_3	Imahing	Panaraswang	13	12	25	0	0
10	10	cluster_3	Karasik	Karasik	24	31	55	0	0
11	11	cluster_3	Karasik	Karangpanar	8	11	19	0	0
12	12	cluster_3	Karasik	Ung	7	20	27	0	0
13	13	cluster_2	Cibeureum	Cibeureum	20	24	44	0	1
14	14	cluster_3	Cibeureum	Keranasgara	16	20	36	0	0
15	15	cluster_2	Tamanasari	Tamanasari	22	17	39	0	1
16	16	cluster_3	Tamanasari	Sangkal	3	10	13	0	0
17	17	cluster_3	Mangroveum	Mangroveum	32	24	56	0	0
18	18	cluster_1	Manunggulm	Manunggulm	34	40	74	1	3

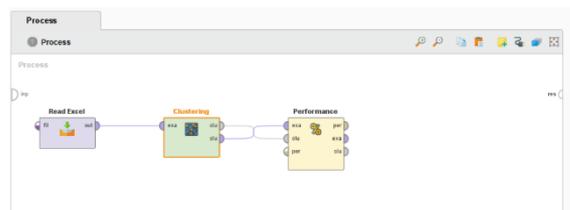
17. Kita dapat melihat bahwa puskesmas yang ada di kecamatan Kahuripan dan Tawang memiliki jumlah pasien laki-laki dan perempuan terbanyak yaitu 74 + 68.

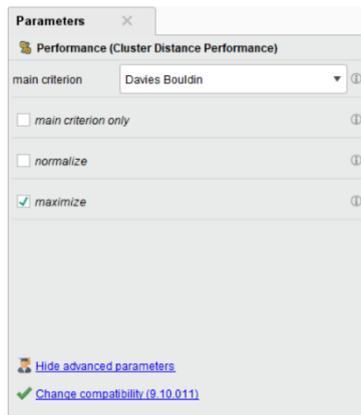


18. Selanjutnya kita akan menggunakan operator *Cluster Performance Distance*, untuk mengetahui berapakah cluster yang baik untuk kasus DBD ini



19. Dan kita atur untuk Parameter *cluster performance distance* ini dengan mencentang pilihan *maximize*





20. Hasil akhir dari Performance Vector (*Average within centroid distance*) dan *Davies Bouldin Index*



IV. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma K-Means dapat digunakan untuk menerapkan upaya pencegahan penyakit DBD dengan mempertimbangkan tingkat kejadian DBD di suatu wilayah. Berdasarkan data yang diolah didapatkan 2 kelompok data, yaitu:

1. Kelompok data kecamatan, meliputi:
 - a. Cluster yang pertama memiliki pusat cluster (5.231)
 - b. Cluster yang kedua memiliki pusat cluster (5.667)
 - c. Cluster yang ketiga memiliki pusat cluster (4.667)
2. Kelompok data puskesmas, meliputi:
 - a. Cluster yang pertama memiliki pusat cluster (11.462)
 - b. Cluster yang kedua memiliki pusat cluster (12.500)

- c. Cluster yang ketiga memiliki pusat cluster (9.667)

Dan untuk mengetahui jumlah kluster yang terbaik untuk sebuah kasus dalam penelitian ini kami menggunakan teknik Elbow Method (Metode Siku) dengan melihat persentase hasil perbandingan antara jumlah cluster yang akan membentuk siku di suatu titik. Hasil teknik elbow method pada penelitian yang kami dapat ialah jumlah kluster (k) sama dengan tiga.

4.2 Saran

Untuk mengembangkan lebih jauh dari penelitian ini, disarankan untuk mengembangkan lebih lanjut adanya pemetaan kesesuaian antara tingkat kejadian DBD di Kota Tasikmalaya dan melakukan pengujian dengan menggunakan algoritma yang berbeda sebagai pembanding, untuk mengetahui metode tersebut dengan tingkat keakuratan optimal sesuai dengan kebutuhan.

REFERENSI

- [1] N. Istiqomah and H. Santoso, "Penyuluhan Demam Berdarah Dengue (DBD) dan Penggunaan Tanaman Selasih Mekah Sebagai Penolak Nyamuk di Desa Bakalan Kecamatan Grogol Kabupaten Kediri," in *SENIAS - Seminar Nasional Pengabdian Masyarakat*, 2019.
- [2] M. F. C. Da Cunha, M. Samin and A. Rahmawati, "Hubungan Perilaku Warga di Rumah dengan Kondisi Lingkungan Tempat Tinggal Sebagai Faktor Penyebab Demam Berdarah Dengue di Kelurahan Kota Uneng Kecamatan Alok Kabupaten Sikka," *Jurnal Geografi*, vol. 18, no. 1, pp. 126-146, 21 Juni 2022.
- [3] B. Pratama and R. J. Sagala, "Sistem Pakar Mendiagnosa Penyakit Demam Berdarah Akibat Virus Nyamuk Aedes Aegepty dengan Menggunakan Metode Certainty Factor," *Excellent Midwifery Journal*, vol. 2, no. 2, pp. 68-73, Oktober 2019.
- [4] M. Hariyanto and R. T. Shita, "Clustering Pada Data Mining Untuk Mengetahui Potensi Penyebaran Penyakit DBD Menggunakan Metode Algoritma K-Means dan Metode Perhitungan Jarak

- Euclidean Distance," *SKANIKA Sistem Komputer dan Teknik Informatika*, vol. 1, no. 1, pp. 117-122, Maret 2018.
- [5] K. Fatmawati and A. P. Windarto, "Data Mining: Penerapan RapidMiner dengan K-MEANS Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (DBD) Berdasarkan Provinsi," *CESS (Journal of Computer Engineering System and Science)*, vol. 3, no. 2, pp. 173-178, Juli 2018.
- [6] I. N. M. Adiputra, "Clustering Penyakit DBD Pada Rumah Sakit Dharma Kerti Menggunakan Algoritma K-Means," *INSERT : Information System and Emerging Technology Journal*, vol. 2, no. 2, pp. 99-105, Desember 2021.