

# Penerapan Algoritma K-Means Clustering Pada Penyebaran Penyakit Infeksi Saluran Pernapasan Akut (ISPA) di Provinsi Riau

Ninaria Purba<sup>1</sup>, Poningsih<sup>2</sup>, Heru Satria Tambunan<sup>1</sup>

<sup>1</sup> Prodi Sistem Infomasi, STIKOM Tunas Bangsa, Pematangsiantar

<sup>2</sup> AMIK Tunas Bangsa, Pematangsiantar

Email: <sup>1</sup>Ninaria77@gmail.com, <sup>2</sup>poningsihamiktb@gmail.com, <sup>3</sup>heru@amiktunasbangsa.ac.id

**Abstrak**—Kesehatan merupakan hal yang berharga bagi manusia karena siapa saja dapat mengalami gangguan kesehatan, begitu pula pada manusia yang sangat rentan terhadap berbagai macam penyakit namun penyebabnya tidak kita sadari. Algoritma K-means tidak terpengaruh terhadap urutan objek yang digunakan, hal ini dibuktikan ketika penulis mencoba menentukan secara acak titik awal pusat cluster dari salah satu objek pada permulaan perhitungan. Jumlah keanggotaan cluster yang dihasilkan berjumlah sama ketika menggunakan objek yang lain sebagai titik awal pusat cluster tersebut. Namun, hal ini hanya berpengaruh pada jumlah iterasi yang dilakukan. Pengelompokan objek (objek clustering) adalah salah satu proses dari objek mining yang bertujuan untuk mempartisi objek yang ada kedalam satu atau lebih cluster objek berdasarkan karakteristiknya. Penelitian ini mengkaji bagaimana penggunaan Algoritma K-means Cluster Analysis dalam studi kasus penyakit menular manusia yaitu Infeksi Saluran Pernafasan Akut. Penelitian ini mengkaji metode K-means Cluster Analysis dalam penyakit Infeksi Saluran Pernafasan Akut berdasarkan set variabel yang dibentuk per kotamadya Di Provinsi Riau.

**Kata Kunci:** Algoritma; K-Means; Clustering; ISPA

**Abstract**—Health is a valuable thing for humans because anyone can experience health problems, as well as humans who are very susceptible to various kinds of diseases but we don't realize the cause. The K-means algorithm is not affected by the order of objects used, this is proven when the author tries to randomly determine the starting point of the cluster center of one of the objects at the beginning of the calculation. The resulting number of cluster membership is the same when using another object as the starting point of the center of the cluster. However, this only affects the number of iterations performed. Object clustering (object clustering) is a process of object mining which aims to partition existing objects into one or more object clusters based on their characteristics. This study examines how to use the K-means Cluster Analysis Algorithm in case studies of human infectious diseases, namely Acute Respiratory Infection. This study examines the K-means Cluster Analysis method in Acute Respiratory Infection based on a set of variables established per municipality in Riau Province.

**Keywords:** Algorithm; K-Means; Clustering; ISPA

## 1. PENDAHULUAN

Kesehatan merupakan hal yang berharga bagimanusia karena siapa saja dapat mengalami gangguankesehatan, begitu pula pada manusia yang sangat rentan terhadap berbagai macam penyakittetapi penyebabnya tidak kita sadari. Hambatan-hambatan yang menyebabkan sulitnya melakukankonsultasi penyakit oleh dokter sekarang ini dapat diatasi dengan adanya program komputer. Dalam hal ini, data informasi dapat membantu pemecahan masalah terhadap penyakit-penyakit dengan diberikannasihat kepada pembaca dan menemukan solusi terhadap berbagai permasalahan yang spesifik. Penyakit infeksi saluran pernapasan akut (ISPA) adalah penyebab utama morbilitas dan mortalitas penyakit menular didunia.

Penyakit ISPA juga penyebab kematian terbesar ketiga didunia dan pembunuh utama di negara berpenghasilan rendah dan menengah. salah satu faktor yang berpengaruh terhadap faktor-faktor resiko penyakit ISPA yaitu faktor lingkungan. Lingkungan yang dimaksud adalah pencemaran udara baik didalam ruangan maupun diluar ruangan serta sanitasi rumah. Pencemaran udara dalam rumah seperti asap hasil pembakaran bahan bakar untuk memasak dengan konsentrasi yang tinggi, asap rokok, ventilasi rumah dan kepadatan hunian, sedangkan pencemaran diluar ruangan seperti pembakaran, transportasi dan hasil pembuangan asap pabrik. ISPA menular melalui kontak langsung atau tidak langsung dan benda yang telah cemari virus dan bakteri penyebab ISPA dapat juga ditularkan melalui udara tercemar. Permasalahan yang dikaji dalam penelitianini adalah bagaimana penggunaan metode *K-means Cluster* terhadap penyebaran penyakit ISPA. Tujuan yang ingin penulis capai adalah mengkaji metode *K-means Cluster* dalam penyebaran penyakit ISPA berdasarkan set variabel yang dibentuk oleh Dinas Kesehatan Provinsi Riau.

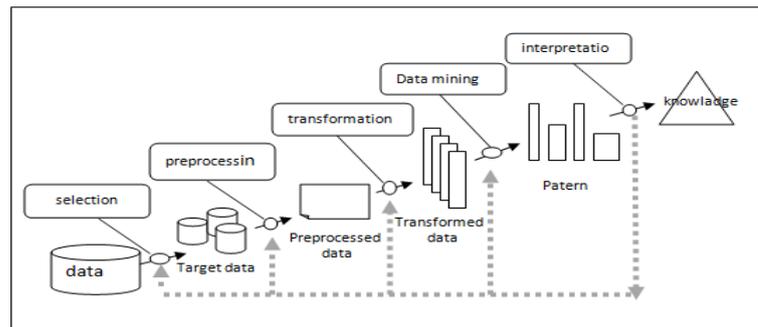
Dari data tersebut, Dinas Kesehatan Provinsi Riau membutuhkan pengolahan data penyakit lebih lanjut untuk menemukan informasi mengenai penyakit ISPA yang diderita oleh masyarakat di wilayah Provinsi Riau menggunakan data mining algoritma *K-means Cluster*. Dengan mengetahui kondisi ini, pihak Dinas Kesehatan Provinsi Riau dapat mengambil tindakan kebijakan dalamantisipasi pengobatan dan pencegahan penyakit salah satunya yaitu dengan melakukan penyuluhan atau sosialisasi, sehingga kegiatan tersebut akan tepat sasaran terhadap Penyebaran penyakit ISPA yang ada di masyarakat.

## 2. METODOLOGI PENELITIAN

### 2.1 Data Mining

Data Mining adalah suatu metode pengelolaan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengelolaan data dengan metode data mining ini dapat digunakan untuk mengambil keputusan di masa depan.

Data mining ini juga dikenal dengan istilah *pattern recognition* (Sulastrri and Gufroni, 2017). Berikut ini tahapan proses data mining dalam penemuan pengetahuan berulang dalam *database* dapat kita lihat pada gambar 1



**Gambar 1.** Tahapan Proses data Mining

## 2.2 Clustering

Metode *Clustering* adalah mempartisi data ke dalam kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan kedalam satu *cluster* yang sama (Muhammad Jazuli). 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* (Sudiman, Windarto, 2018).

## 2.3 Algoritma K-Means

Algoritma *K-Means* merupakan algoritma klasterisasi yang mengelompokkan data berdasarkan titik pusat klaster (*centroid*) terdekat dengan data. Tujuan *K-Means* adalah pengelompokkan data dengan memaksimalkan kemiripan data dalam satu klaster dan meminimalkan kemiripan data antara klaster. ukuran kemiripan yang digunakan dalam klaster adalah fungsi jarak. sehingga pemaksimalan kemiripan data didapatkan berdasarkan jarak terpendek antara data titik *centroid*.

## 2.3 Infeksi Saluran Pernafasan (ISPA)

ISPA adalah penyakit infeksi yang mengenai saluran pernafasan bagian atas dan bawah yang disebabkan oleh masuknya kuman berupa virus, bakteri, atipikal (atipikal plasma) atau aspirasi substansi asing yang menyerang organ pernafasan. Kasus penyakit ispa ini sangat berkaitan dengan perubahan kondisi lingkungan akibat erupsi merapi, perilaku manusia dan faktor lingkungan meliputi sanitasi fisik rumah, kebakaran hutan, sarana air bersih, sarana pembuangan limbah, dan kesehatan lingkungan pada musim kemarau. Penularan penyakit ISPA dapat terjadi melalui udara yang tercemar, bibit penyakit masuk kedalam tubuh melalui pernapasan, melalui udara dimaksudkan adalah cara penularan yang terjadi tanpa kontak dengan penderita maupun dengan benda terkontaminasi.

## 2.4 Teknik Pengumpulan Data

Data penelitian ini diperoleh dari Badan Pusat Statistik (BPS) Nasional dengan situs <https://www.bps.go.id> mengenai data penyebaran ISPA di Provinsi Riau tahun 2015-2019. Untuk menyelesaikan masalah pada penelitian ini adalah penulis melakukan observasi ke badan pusat statistik dengan tujuan untuk mendapatkan validasi data yang jelas.

## 2.5 Pengolahan Data

Pada tahapan ini, data diolah untuk mendapatkan hasil yang kemudian dapat dikelola ke tahap berikutnya sehingga menghasilkan informasi yang tepat. proses inialisasi data terlebih dahulu ke dalam bentuk angka/numerikal. Kemudian mengelompokkan data yang sudah ada dalam dua kelompok yaitu tinggi, rendah, dengan metode *K-Means Clustering*. Dalam pengujian ini penguji menggunakan sebuah *software Rapidminer*, dengan pengujian data menggunakan *software* kita akan membandingkan bagaimana hasil pengolahan data secara manual dengan hasil pengolahan data menggunakan *software*.

**Tabel 1.** Data Penelitian

No	Nama Kabupaten	Data Penyakit ISPA (Jumlah/orang)				
		2015	2016	2017	2018	2019
1	Bengkalis	2918	1162	402	705	420
2	Inhil	1245	1576	615	212	1804
3	Inhul	2264	975	565	721	404
4	Kampar	2137	1040	810	795	4152
5	Meranti	471	312	412	830	610
6	Kuansing	4571	517	517	517	4571
7	Pelalawan	1950	892	962	313	2244

No	Nama Kabupaten	Data Penyakit ISPA (Jumlah/orang)				
		2015	2016	2017	2018	2019
8	Rohil	1092	1051	451	866	670
9	Rohul	3515	541	841	418	2462
10	Siak	4539	910	610	720	4616
11	Dumai	3764	870	670	6103	3932
12	Pekanbaru	8861	7215	1715	1779	7377

### 3. HASIL DAN PEMBAHASAN

Dalam melakukan Clustering, data yang diperoleh akan dihitung terlebih dahulu berdasarkan jumlah Kabupaten yang ada di Provinsi Riau. Hasil penjumlahan berdasarkan set variabel yang digunakan adalah 2015 sampai 2019.

Berikut langkah-langkah perhitungan *K-Means*.

- a) Menentukan jumlah *cluster* dilakukan dengan menggunakan data penyebaran penyakit ISPA dalam sebanyak 12 Kabupaten. Berikut ini adalah cara mencari nilai rata-rata dengan menjumlahkan data setiap tahun dibagi dengan jumlah banyak tahun yang diperoleh.

$$R_1 = \frac{2918 + 1162 + 402 + 705 + 420}{5} = 1121,4$$

$$R_2 = \frac{1245 + 1576 + 615 + 212 + 1804}{5} = 1090,4$$

$$R_3 = \frac{2264 + 975 + 565 + 721 + 404}{5} = 985,8$$

$$R_4 = \frac{2137 + 1040 + 810 + 795 + 4152}{5} = 1786,8$$

$$R_5 = \frac{471 + 312 + 412 + 830 + 610}{5} = 527$$

Dan seterusnya sampai dengan perhitungan  $R_{12}$ . Sehingga diperoleh tabel hasil rata-rata sebagai berikut :

**Tabel 2.** Nilai rata-rata

No	Kabupaten	Total
1	Bengkalis	1121,4
2	Inhil	1090,4
3	Inhul	985,8
4	Kampar	1786,8
5	Meranti	527
6	Kuansing	2138,6
7	Pelalawan	1272,2
8	Rohil	826
9	Rohul	1555,4
10	Siak	2279
11	Dumai	3067,8
12	Pekan Baru	5389,4

- b) Menentukan *centroid* awal *cluster* ditentukan secara random yang diambil dari data yang ada dalam range. Nilai *cluster* 1 diambil dari data paling tinggi dan nilai *cluster* 2 diambil dari nilai data yang paling rendah. Berikut ini cara mencari nilai *centroid* awal :

$$C = 1121,4/1090,4/985,8/1786,8/527/2138,6/1272,2/826/1555,4/2279/3067,8/5389,4 = 5389,4$$

$$C = 1121,4/1090,4/985,8/1786,8/527/2138,6/1272,2/826/1555,4/2279/3067,8/5389,4 = 527$$

**Tabel 3.** Centroid Awal

Cluster	Nilai
C1= Max	5389,4
C2= Min	527

- c) Langkah selanjutnya yaitu mengalokasikan masing-masing data ke *centroid* atau rata-rata terdekat dengan menghitung jarak masing-masing data terhadap pusat *cluster* dengan menggunakan rumus perhitungan sebagai berikut :

Perhitungan untuk jarak dari titik pusat (*centroid*) pertama adalah sebagai berikut.

$$D_{x1,c1} = \sqrt{(5389,4 - 1121,4)^2} = 4268$$

$$D_{x2,c1} = \sqrt{(5389,4 - 1090,4)^2} = 4299$$

$$D_{x3,c1} = \sqrt{(5389,4 - 985,8)^2} = 4403,6$$

$$D_{x4,c1} = \sqrt{(5389,4 - 1786,8)^2} = 3602,6$$

$$D_{x5,c1} = \sqrt{(5389,4 - 527)^2} = 4862$$

Dan seterusnya sampai dengan  $D_{x12,c1}$  . selanjutnya perhitungan jarak dengan titik pusat(*centroid*) ke-2 adalah sebagai berikut :

$$D_{x1,c2} = \sqrt{(1121,4 - 527)^2} = 594,4$$

$$D_{x2,c2} = \sqrt{(1090,4 - 527)^2} = 563,4$$

$$D_{x3,c2} = \sqrt{(985,8 - 527)^2} = 458,8$$

$$D_{x4,c2} = \sqrt{(1786,8 - 527)^2} = 1259,8$$

$$D_{x5,c2} = \sqrt{(527 - 527)^2} = 0$$

Dan selanjutnya sampai dengan  $D_{x12,c2}$  . sehingga diperoleh tabel jarak dari *centroid* dan mencari nilai minimal dari kedua *centroid*. Tabel jarak *centroid* adalah sebagai berikut :

**Tabel 4.** Hasil Perhitungan Jarak Antar Pusat Cluster

No	Kabupaten	Rata-Rata	Clustering		Jarak Terpendek
			C1	C2	
1	Bengkalis	1121,4	4268	594,4	594,4
2	Inhil	1090,4	4299	563,4	563,4
3	Inhul	985,8	4403,6	458,8	458,8
4	Kampar	1786,8	3602,6	1259,8	1259,8
5	Meranti	527	4862,4	0	0
6	Kuansing	2138,6	3250,8	1611,6	1611,6
7	Pelalawan	1272,2	4117,2	745,2	745,2
8	Rohil	826	4563,4	299	299
9	Rohul	1555,4	3834	1028,4	1028,4
10	Siak	2279	3110,4	1752	1752
11	Dumai	3067,8	2321,6	2540,8	2321,6
12	Pekan Baru	5389,4	0	4862,4	0

- d) Menentukan *Cluster* atau pengelompokkan dengan mencari nilai cluster berdasarkan nilai minimal dari *cluster* C1 dan C2 kemudian bandingkan nilai tersebut, lihat nilai terendah dan letakkan pada *cluster* yang sesuai dengan nilai minimal pada iterasi 1, dengan bantuan tanda (1) yang digunakan untuk mengatakan data menjadi anggota pada *cluster*. Berikut tabel *cluster* pada iterasi 1 sebagai berikut :

**Tabel 5.** Cluster iterasi ke-1

No	Kabupaten	Clustering		Hasil
		C1	C2	
1	Bengkalis		1	C2
2	Inhil		1	C2
3	Inhul		1	C2
4	Kampar		1	C2
5	Meranti		1	C2
6	Kuansing		1	C2
7	Pelalawan		1	C2
8	Rohil		1	C2
9	Rohul		1	C2
10	Siak		1	C2
11	Dumai	1		C1
12	Pekan Baru	1		C1

**Tabel 6.** Centroid baru

Cluster	Kabupaten
C1	(11,12)
C2	(1,2,3,4,5,6,7,8,9,10)

- e) Selanjutnya ulangi langkah 3 dan 4. Jika nilai *centroid* pada iterasi dengan nilai *centroid* sebelumnya bernilai sama atau apabila nilai *centroid* sudah optimal serta posisi *cluster* data penyebaran ISPA yang tidak mengalami perubahan lagi maka proses iterasi sudah selesai atau berhenti.

### 3.1 Menghitung Titik Pusat Baru

Mengitung titik pusat baru menggunakan hasil dari setiap anggota masing-masing cluster. Berikut mencari titik pusat baru pada cluster x adalah :

$$C1x = \frac{3067,8 + 5389,4}{2} = 4228,6$$

$$C2x = \frac{1121,4 + 1090,4 + 985,8 + 1786,8 + 527 + 2138,6 + 1272,2 + 826 + 1555,4 + 2279}{10} = 1358,26$$

Hasil perhitungan cluster baru telah didapatkan dapat dilihat pada tabel 7 berikut:

**Tabel 7.** Pusat Cluster Iterasi 2

Pusat Cluster	Nilai
C1 = Max	4228,6
C2 = Min	1358,26

Kemudian setelah mendapatkan nilai pusat *cluster* iterasi 2 maka langkah selanjutnya yaitu melakukan ulang untuk dengan titik pusat *cluster(centroid)* pada *cluster* pertama, sebagai berikut :

$$D_{x1,c1} = \sqrt{(4228,6 - 1121,4)^2} = 3107,2$$

$$D_{x2,c1} = \sqrt{(4228,6 - 1090,4)^2} = 3138,2$$

$$D_{x3,c1} = \sqrt{(4228,6 - 985,8)^2} = 3242,8$$

$$D_{x4,c1} = \sqrt{(4228,6 - 1786,8)^2} = 2441,8$$

$$D_{x5,c1} = \sqrt{(4228,6 - 527)^2} = 3701,6$$

Dan seterusnya sampai dengan  $D_{x12,c1}$  . selanjutnya perhitungan jarak dengan titik pusat(*centroid*) ke-2 adalah sebagai berikut :

$$D_{x1,c2} = \sqrt{(1121,4 - 1358,26)^2} = 236,86$$

$$D_{x2,c2} = \sqrt{(1090,4 - 1358,26)^2} = 267,86$$

$$D_{x3,c2} = \sqrt{(985,8 - 1358,26)^2} = 372,46$$

$$D_{x4,c2} = \sqrt{(1786,8 - 1358,26)^2} = 428,54$$

$$D_{x5,c2} = \sqrt{(527 - 1358,26)^2} = 831,26$$

Dan selanjutnya sampai dengan  $D_{x12,c2}$  . sehingga diperoleh tabel jarak dari *centroid* dan mencari nilai minimal dari kedua *centroid*. Tabel jarak *centroid* adalah sebagai berikut

**Tabel 8.** Perhitungan titik pusat iterasi 2

No	Kabupaten	C1	C2	Rata-Rata	C1	C2
1	Bengkalis		1	1121,4	3107,2	236,86
2	Inhil		1	1090,4	3138,2	267,86
3	Inhul		1	985,8	3242,8	372,46
4	Kampar		1	1786,8	2441,8	428,54
5	Meranti		1	527	3701,6	831,26
6	Kuansing		1	2138,6	2090	789,34
7	Pelalawan		1	1272,2	2956,4	86,06
8	Rohil		1	826	3402,6	532,26
9	Rohul		1	1555,4	2673,2	197,14
10	Siak		1	2279	1949,6	920,74
11	Dumai	1		3067,8	1160,8	1709,54
12	Pekan Baru	1		5389,4	1160,8	4031,14

Dari tabel diatas, data yang memiliki jarak terpendek dengan *centroid* akan menjadi anggota pada kelompok tersebut. Posisi tiap *cluster* pada iterasi 2 dengan bantuan tanda(1) yang digunakan untuk mengatakan data menjadi anggota pada *cluster*. Perhatikan tabel 9 dibawah ini:

**Tabel 9.** Data berdasarkan jarak terdekat iterasi 2

No	Kabupaten	Clustering		Hasil
		C1	C2	
1	Bengkalis	1		C2
2	Inhil	1		C2
3	Inhul	1		C2

No	Kabupaten	Clustering		Hasil
		C1	C2	
4	Kampar		1	C2
5	Meranti		1	C2
6	Kuansing		1	C2
7	Pelalawan		1	C2
8	Rohil		1	C2
9	Rohul		1	C2
10	Siak		1	C2
11	Dumai	1		C1
12	Pekan Baru	1		C1

Tabel 10. Pusat cluster baru

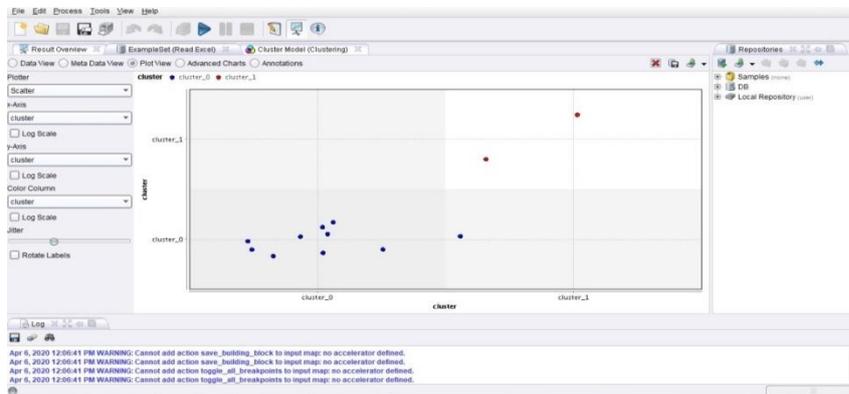
Cluster	Kabupaten
C1	(11,12)
C2	(1,2,3,4,5,6,7,8,9,10)

Dari hasil perhitungan yang telah dicoba maka dapat diperoleh hasil akhir dimana hasil perhitungan pada cluster iterasi 1 dan iterasi 2 memiliki nilai cluster yang sama atau tidak berubah. maka proses perhitungan dihentikan iterasi dihentikan dan hasil yang diperoleh yaitu :

- a) Cluster 1 (C1) memiliki 2 data yang diartikan bahwa 2 kelompok pertama adalah kategori kabupaten penyebaran penyakit ISPA yang tinggi yaitu: kabupaten Dumai dan Pekanbaru.
- b) Cluster 2 (C2) memiliki 10 data yang diartikan bahwa 10 kelompok kedua adalah kategori kabupaten penyebaran penyakit ISPA yang rendah yaitu : kabupaten Bengkalis, Inhil, Inhil, Kampar, Meranti, Kuansiang, Pelalawan, Rohil, Rohul, Siak.
- c) Total seluruh kabupatensetiap itemyang telah dikelompokkan adalah 12 data.

### 3.2 Pengujian Rapidminer

Hasil pengelompokkan pada *rapidminer* dapat dilihat pada gambar diagram *scater* berikut ini :

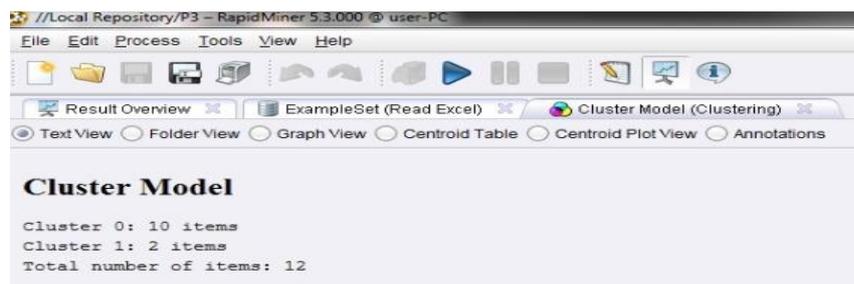


Gambar 2. Hasil Pengelompokkan

Berdasarkan gambar 2 diatas, dapat diketahui bahwa :

- a) Warna Biru : menunjukkan kelompok rendah sebanyak 10 node
- b) Warna Merah : menunjukkan kelompok tinggi memiliki 2 node.

Hasil pada *Cluster Model* sebagai berikut:



Gambar 3. Hasil Pada Cluster Model

Keterangan :

- a) Jumlah Cluster 0 (Rendah) berjumlah 10 items

- b) Jumlah *Cluster* 1 (Tinggi) berjumlah 2 *items*
- c) Jumlah keseluruhan *items* adalah 12.

#### 4. KESIMPULAN

Berdasarkan pembahasan sebelumnya dapat disimpulkan penelitian yang telah dilakukan dengan menggunakan algoritma *K-Means* mampu mengelompokkan penyebab penyebaran penyakit ISPA kedalam 2 *cluster*, yang mana *cluster* 1 memberikan rekomendasi tinggi berjumlah 10 Kabupaten, *cluster* 2 memberikan rekomendasi rendah berjumlah 2 Kabupaten. Dalam pengujian ini dilakukan dengan menggunakan dua aplikasi, yang pertama dengan menggunakan aplikasi pengolah angka *Microsoft Excel* dan dengan menggunakan aplikasi *Rapidminer 5.3* menggunakan *K-Means* dapat menampilkan 2 *cluster* dari hasil klasifikasi.

#### REFERENCES

- [1] Bastian, A., Sujadi, H. and Febrianto, G. (2018) 'Penerapan Algoritma K-Means Clustering Analysis Pada Penyakit Menular Manusia (Studi Kasus Kabupaten Majalengka)', *Jurnal Sistem Informasi (Journal of Information System)*, 14(1), pp. 26–32.
- [2] Dhuhita, W. M. P. (2015) 'Clustering Menggunakan Metode K-Means Untuk Menentukan Status Gizi Balita', *Jurnal Informatika*, 15(2), pp. 160–174.
- [3] Hakim, L. (2019) 'Universitas Xyz Menggunakan Algoritma K-Means Clustering', *JURNAL SISTEM INFORMASI DAN TEKNOLOGI*, 2(1), pp. 80–86.
- [4] Jaroji, Danuri and Putra, F. P. (2016) 'K-Means Untukmenentukan Calon Penerima Beasiswa Bidikmisi Di Polbeng', *JURNAL INOVTEKPOLBENG - SERI INFORMATIKA*, 1(1), pp. 87–94.
- [5] Sari, R. W., Wanto, A. and Windarto, A. P. (2018) 'Implementasi Rapidminer Dengan Metode K-Means (Study Kasus: Imunisasi Campak Pada Balita Berdasarkan Provinsi)', *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 2(1), pp. 224–230.
- [6] Setiawan, R. (2016) 'Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru ( Studi Kasus : Politeknik LP3I Jakarta )', *JURNAL LENTERA ICT*, 3(1), pp. 76–92.
- [7] Windarto, A. P. (2017) 'Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering', *Techno.COM*, 16(4), pp. 348–357.